CHAPTER 1

Photometric Stereo: An Overview

Vasileios Argyriou and **Maria Petrou***

Contents

* Communications and Signal Processing Group, Electrical and Electronic Engineering Department, Imperial College London SW7 2AZ, United Kingdom

Advances in Imaging and Electron Physics, Volume 156, ISSN 1076-5670, DOI: 10.1016/S1076-5670(08)01401-8.

1. INTRODUCTION

Although it has long been known that shading in images provides an important depth cue, only relatively recently have surface orientation and depth recovery problems been properly formulated and solved sufficiently reliably for commercial applications. The earliest work on the quantitative use of shading information was apparently in the mid-1960s (Hapke, 1963; Orlova, 1956; Rindfleisch, 1966) on recovering the shape of parts of the lunar surface in preparation for human exploration of the moon.

This chapter presents an overview of the state-of-the-art methods for three-dimensional (3D) shape reconstruction, starting with a tutorial on the basic understanding of the problem with emphasis on photometric stereo technology. *Photometric stereo* is a method for recovering local surface shape and albedo from a number of images captured under different illumination directions. The photometric stereo method is simple to implement. However, a brief introduction on radiometry, image formation, and surface description models is essential to understanding its underlying principles.

Section 2 defines the imaging geometry and provides some basic definitions of the necessary geometric equations. Section 3 defines some basic concepts of radiometry and derives the fundamental equation that connects the brightness of an imaged surface patch with the image brightness of the corresponding pixel. The interaction of light with a surface is modeled by the so-called bidirectional reflectance distribution function (BRDF), which is discussed in Section 4. Some well-known in computer vision reflectance models are introduced in Section 5. Surface recovering methods—either from binocular stereo or a single image (shape from shading)—are presented in Section 6. In Section 6 we also refine the term reflectance map, which relates image irradiance to surface orientation, for a given light source direction and surface reflectance, since it is fundamental in photometric stereo and image rendering for the description of the shape of an object. Section 7 discusses extensions of photometric stereo for color images. Sections 8 and 9 deal with the relaxation of some of the most restrictive assumptions of photometric stereo: reconstruction methods when highlights and shadows are present (Section 8) or the lights are at a finite distance from the surface (Section 9). Photometric stereo requires the reconstructed object to remain still during the acquisition stage, and this limitation is discussed in Section 10. Section 11 reviews the solutions proposed for the problems introduced due to the invalid assumption of orthographic projection. In Section 12 the error analysis literature on photometric stereo is reviewed in relation to errors due to acquisition or illumination calibration, and some ways of reducing them are reported. Finally, solutions to the problem of optimal illumination configuration are

reviewed in Section 13, based on image rendering and the properties of the reflectance maps. Conclusions are drawn in Section 14.

2. DEFINITIONS

Figure 1 shows a simple imaging system. The viewing direction is aligned with the z-axis of the coordinate system, pointing toward the camera, since a right-handed coordinate system is assumed. The x and y axes are perpendicular to the viewing direction with the origin of the system at the intersection of the z-axis and the observed plane. The shape of a surface may be modeled either by using the height of every point above a reference plane or the gradient vector of the surface at each position.

Various active vision methods (e.g., laser scanners or radar-based systems) usually recover a height map for the imaged surface. Conventional stereo also recovers a height (or depth) map using disparity measures. Conversely, photometric stereo recovers the normal vector field of the surface, modeling the surface as a collection of flat patches, each corresponding to a pixel of the captured images.

The general equation of a 3D plane is given by

$$Ax + By + Cz + D = 0 \Rightarrow \frac{A}{C}x + \frac{B}{C}y + z + \frac{D}{C} = 0, \tag{1}$$

where the first partial derivatives of z with respect to x and y correspond to the components of the slope of the surface, p_n and q_n, respectively, as

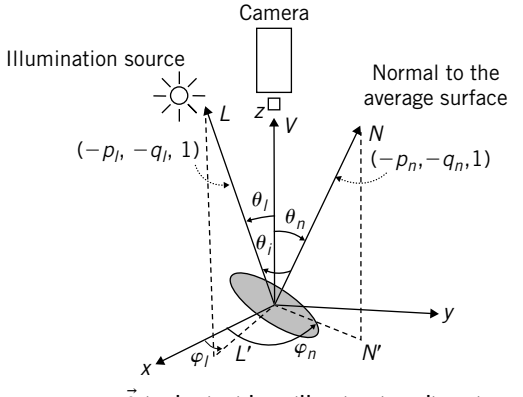


FIGURE 1 Imaging geometry. $\vec{L}$ is the incident illumination direction vector, and $\vec{N}$ is the normal to the surface vector. $\vec{V}$ is the viewing direction.

follows:

$$\frac{\partial z}{\partial x} = -\frac{A}{C} = -p_n$$

$$\frac{\partial z}{\partial y} = -\frac{B}{C} = -q_n. \tag{2}$$

Vector $\left(-\frac{A}{C}, -\frac{B}{C}, 1\right)^T = (-p_n, -q_n, 1)^T$ is normal to the plane (Figure 2a).

An alternative definition, which is also useful in this chapter, is based on the representation of a plane by equation $\vec{r} = \vec{a} + \alpha\vec{b} + \beta\vec{c}$, where $\vec{r}$ is the position vector of any point on the plane, $\vec{a}$ is the position vector of a specific known point on the plane, and $\vec{b}$ and $\vec{c}$ are two nonparallel vectors lying on the plane (Figure 2b). Then the normal to the plane is given by the cross product $\vec{b} \times \vec{c}$. For example, if the nonparallel vectors lying on the plane are chosen as

$$\vec{b} = (1, 0, p_n)^T$$
$$\vec{c} = (0, 1, q_n)^T, \tag{3}$$

a normal vector to the surface is

$$\vec{b} \times \vec{c} = (-p_n, -q_n, 1)^T. \tag{4}$$

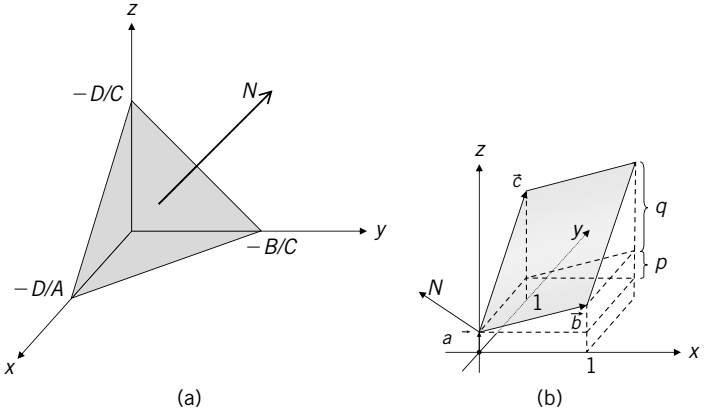


FIGURE 2 (a) A plane and its normal vector. (b) The slope of a surface patch may be expressed in terms of p and q. See text for details.

We define the surface normal to be of unit length, so

$$\vec{N} = \frac{(-p_n, -q_n, 1)^T}{\sqrt{p_n^2 + q_n^2 + 1}}. \tag{5}$$

In terms of azimuth angle φ_n (measured in the xy-plane counterclockwise from the x-axis) and the zenith angle θ_n (measured between the surface normal and the z-axis), the surface slope components may be written as (Figure 1)

$$\begin{aligned} p_n &= -\cos\varphi_n \tan\theta_n \\ q_n &= -\sin\varphi_n \tan\theta_n. \end{aligned} \tag{6}$$

Furthermore, orthographic projection is assumed; that is, the camera is considered to be far from the surface relative to the size of the surface. The light source is regarded as a point source at infinite distance and therefore constant illumination is assumed over the entire scene. These assumptions are relaxed in Sections 9 through 11, when we review some methods of photometric stereo designed to cope with non-orthographic projection and an illumination source of finite extent at a finite distance from the surface. The normal vector of a surface perpendicular to the light rays is used to specify the direction of the illumination source $(-p_l, -p_l, 1)$. The unit light source direction vector is then given by

$$\vec{L} = \frac{(-p_l, -q_l, 1)^T}{\sqrt{p_l^2 + q_l^2 + 1}}. \tag{7}$$

The components of this vector in terms of angles φ_l and θ_l (defined in Figure 1) are

$$\begin{aligned} p_l &= -\cos\varphi_l \tan\theta_l \\ q_l &= -\sin\varphi_l \tan\theta_l \end{aligned} \tag{8}$$

In terms of the zenith and azimuth angles, the unit illumination vector may be written as

$$\vec{L} = (\sin\theta_l \cos\varphi_l, \sin\theta_l \sin\varphi_l, \cos\theta_l), \tag{9}$$

where we used Eqs. (7) and (8) and the fact that $\cos\theta_l = 1/(\sqrt{1 + p_l^2 + q_l^2})$.

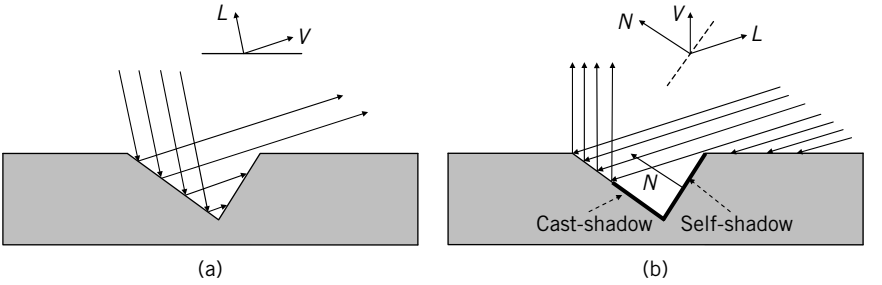


FIGURE 3 Masking (a) and self- and cast-shadows (b). $\vec{V}$ is the viewing direction, $\vec{L}$ is the illumination direction, and $\vec{N}$ is the surface normal.

The effect of obstructions in the path of the reflected light is called *masking* or *occlusion*, whereas the term *shadowing* is used to describe obstructions in the path of the incident light, (Figure 3). The masking effects may be neglected only when the viewer is in the zenith of a flat surface. Shadows may be distinguished as *self-* and *cast-shadows*. *Self-shadows* occur when the angle between the normal vector $\vec{N}$ to the surface facet and the illumination vector $\vec{L}$ is larger then 90°. The term *cast-shadow* is used when the shadow is projected from another facet (Figure 3).

3. RADIOMETRY AND IMAGE FORMATION OVERVIEW

The image intensity of a pixel is a function of the orientation of the corresponding surface patch (surface normal), the reflectance of the material from which the surface is made or with which the surface is coated, and the spectrum and direction of the illumination used to capture the image. In order to recover the shape and the surface reflectance properties of an object, we need to understand how the intensity of each image pixel depends on all these factors.

Some concepts relevant to this chapter are introduced here. First, we introduce the concept of the solid angle subtended by a surface patch. A solid angle, subtended by an area dA' on the surface of a sphere of radius R, has its vertex at the center of the sphere and is equal to $d\omega$ steradian (sr) measured as $d\omega = dA'/R^2$. Thus, the solid angle subtended by a surface patch dA with normal vector forming angle θ_k with the radius of the sphere is

$$d\omega = \frac{dA \cos\theta_k}{R^2}, \tag{10}$$

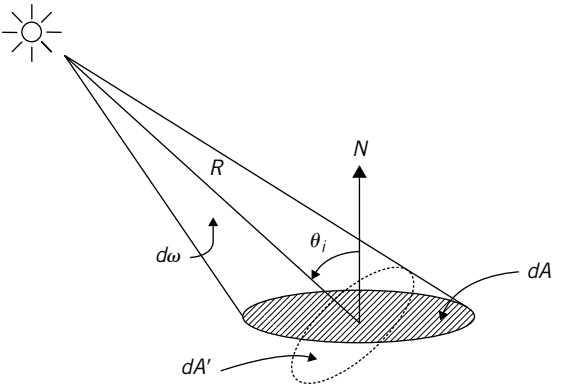


FIGURE 4 Solid angle subtended by a surface patch of area *dA*.

since $dA\cos\theta_k$ is the projected area of patch dA on the plane orthogonal to the radius.

According to the National Bureau of Standards (Nicodemus *et al.*, 1977), the radiant flux $d\Phi$ is defined as the power propagated as optical electromagnetic radiation and is measured in watts (W). The incident light power per unit surface area is called irradiance E and is measured in watts per square meter of surface (Wm^{-2})

$$E = \frac{d\Phi}{dA}, \tag{11}$$

where $d\Phi$ is the light power received by surface dA.

Radiance is the amount of light emitted by a particular area within a given solid angle in a specified direction. Surface radiance (brightness) I, in a direction forming angle θ_n with respect to the surface normal, is the energy emitted per unit solid angle and per unit area orthogonal to direction θ_n:

$$I = \frac{d\Phi}{dA\cos\theta_n d\omega} \Rightarrow d\Phi = IdA\cos\theta_n d\omega. \tag{12}$$

Surface radiance is measured in watts per square meter per steradian ($Wm^{-2}sr^{-1}$), (Figure 5).

First, we show that the brightness recorded by the sensor used is directly proportional to the brightness of the imaged surface. This then allows use of the captured digital image to draw conclusions about the physical imaged surface. For the purpose of this argument, the E in Eq. (11) will be taken to represent the light that falls on the imaging sensor, while the I in Eq. (12)

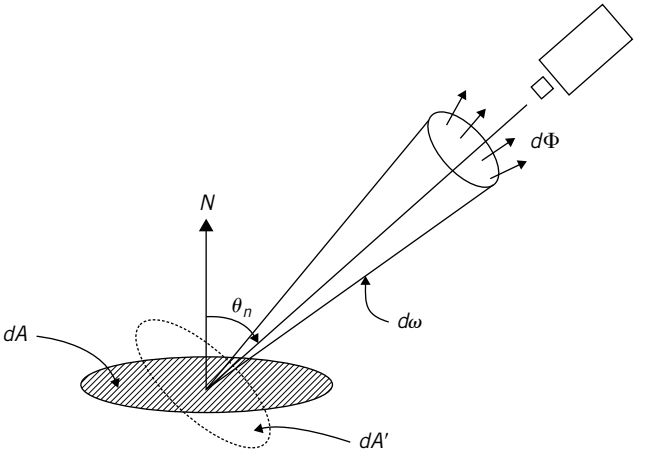


FIGURE 5 Surface radiance (brightness).

will be taken to represent the light that leaves the imaged surface. We shall make this explicit by using the right suffixes in E and I. Assuming that the lens does not absorb any light, the $d\Phi$ in Eqs. (11) and (12) will be the same.

In order to find the relationship between the surface radiance I and image irradiance E, the area of the region in the image dA_i and the area of the corresponding patch on the surface dA_s must be determined. Based on Eqs. (11) and (12) the irradiance of the image is given by

$$E_{on\_image} = \frac{d\Phi}{dA_i} = \frac{I_{from\_surface}\ dA_s \cos\theta_n\ d\omega}{dA_i}, \tag{13}$$

where $d\omega$ is the solid angle subtended by the lens, as seen from the scene patch. For a surface patch in direction α with respect to the lens axis, the surface of a lens of diameter d orthogonal to the direction of the patch is $\pi(\frac{d}{2})^2 \cos\alpha$, whereas the distance of the surface patch from the center of the lens is $z_0/\cos\alpha$. Thus, $d\omega$ is given by

$$d\omega = \frac{\pi\left(\frac{d}{2}\right)^2 \cos\alpha}{\left(\frac{z_0}{\cos\alpha}\right)^2} \Rightarrow d\omega = \frac{\pi}{4}\left(\frac{d}{z_0}\right)^2 \cos^3\alpha, \tag{14}$$

where z_0 is the distance of the scene patch from the lens along the z-axis, (Figure 6).

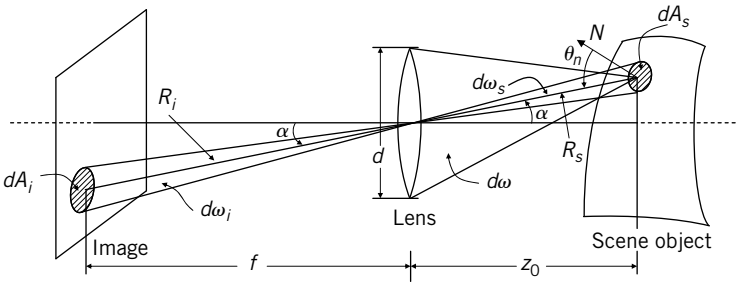


FIGURE 6 The image irradiance is proportional to the scene radiance.

Because the solid angles subtended by the image patch and the scene object patch are equal (Figure 6), we obtain

$$d\omega_i = d\omega_s \Rightarrow \frac{dA_i \cos\alpha}{(f/\cos\alpha)^2} = \frac{dA_s \cos\theta_n}{(z_0/\cos\alpha)^2} \Rightarrow \frac{dA_s}{dA_i} = \frac{\cos\alpha}{\cos\theta_n}\left(\frac{z_0}{f}\right)^2. \quad (15)$$

Substituting from Eqs. (14) and (15) into Eq. (13), we obtain

$$E_{on\_image} = I_{from\_surface}\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\alpha. \quad (16)$$

Thus, the image irradiance E, which corresponds to our notion of "brightness," is proportional to the scene radiance I. Therefore, E depends on the amount of light that falls on the imaged surface and the reflected fraction of it. It also depends through I on the illumination and viewing directions.

Here E is what falls on the film and I is what leaves the surface. Therefore, all we do is modeling what happens from the moment the light leaves the surface until the moment it reaches the sensor. The next section discusses what happens to the light from the moment it leaves the light source until the moment it leaves the surface. It is that part of the light's journey from the light source to the sensor that depends on the physical properties of the imaged surface and the imaging geometry.

4. BIDIRECTIONAL REFLECTANCE DISTRIBUTION FUNCTION

The reflectance models developed in the field of computer vision are classified into two main categories: *physical models* and *geometrical models*. The latter have simple forms and are derived by analyzing the surface and illumination geometry. In contrast, the physical models are based on the

electromagnetic wave theory providing accurate and detailed descriptions, but requiring complex manipulations and significant prior information concerning the imaged scene. This makes them inappropriate for computer vision applications.

Reflectance models are generally presented in terms of the BRDF, relating the incident light energy from the direction of illumination $E_{on\_surface}(\lambda, \theta_l, \varphi_l)$, to the reflected brightness $I_{from\_surface}(\lambda, \theta_n, \varphi_n)$ in the direction of the viewer:

$$f_{\text{BRDF}}(\lambda, \theta_l, \varphi_l, \theta_n, \varphi_n) \equiv \frac{I_{from\_surface}(\lambda, \theta_n, \varphi_n)}{E_{on\_surface}(\lambda, \theta_l, \varphi_l)}. \tag{17}$$

Note that f_{BRDF} has units of sr^{-1} because I is energy per unit area per unit solid angle, measured in $W\ m^{-2}sr^{-1}$, whereas E is energy per unit area, measured in $W\ m^{-2}$. In this expression, I is the same as the I in Eqs. (12) and (16), but we have made explicit here its dependence on the wavelength λ and the viewing direction with respect to the normal to the surface. $E(\lambda, \theta_l, \varphi_l)$ is different from the E in Eq. (16), and again we have made explicit here its dependence on the wavelength λ and the direction of illumination (θ_l, φ_l) (Figure 7). The E in Eq. (16) is the light that falls on the sensor cell, whereas the E in Eq. (17) is the light that falls on the imaged surface.

The dependence of f_{BRDF} on λ, θ_l, φ_l, θ_n, and φ_n is complicated. As it is difficult to define models based on the exact physical processes that take place, empirical models of f_{BRDF} are used. Accurate models of the BRDF are

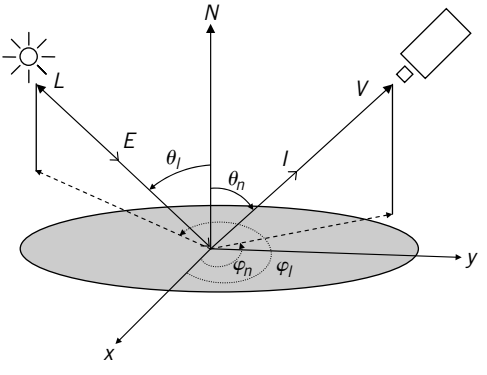


FIGURE 7 Geometry of light reflection. $E(\theta_l, \varphi_l)$, irradiance due to source in direction (θ_l, φ_l). $I(\theta_n, \varphi_n)$, radiance of the surface in direction (θ_n, φ_n). $\vec{V}$ is the vector of the viewing direction, and $\vec{L}$ is the vector of the direction of the illuminating source.

generally complex with many parameters. However, various simplifying assumptions may be made.

One of the simplifications is separating the radiometric from the geometric dependence on the right-hand side of Eq. (17), which therefore may be written as

$$f_{\text{BRDF}}(\lambda, \theta_l, \varphi_l, \theta_n, \varphi_n) = r(\lambda) \frac{\tilde{I}_{from\_surface}(\theta_n, \varphi_n)}{\tilde{E}_{on\_surface}(\theta_l, \varphi_l)}. \tag{18}$$

Here $r(\lambda)$, also known as *surface albedo*, is the reflectance coefficient that expresses the fraction of light energy reflected by the surface as a function of the wavelength of the incident light. We have used a tilde to distinguish between radiometric quantities I and E, that depend on wavelength and geometry, from radiometric quantities $\tilde{I}$ and $\tilde{E}$, that depend only on the geometry. Albedo $r(\lambda)$ is often treated as a constant, independent of λ. In the section on photometric stereo, we shall see how it may be recovered up to a multiplicative factor, for the wavelengths at which the sensors we use have nonzero response.

Another simplifying assumption in modeling f_{BRDF} is that each surface facet is isotropic, since for many surfaces the radiance is not altered if the surface is rotated about its normal (He *et al.*, 1991). In that case, the BRDF depends on the difference of the two azimuth angles $\varphi_n - \varphi_l$, and not explicitly on each angle.

Based on these simplifications, many BRDF models have been presented over the past two decades, but only the Lambertian and the Phong models have enjoyed widespread use for modeling smooth surfaces and surfaces with specular and diffuse reflections, respectively.

5. REFLECTANCE MODELS

The Lambertian and Phong models have been widely used to model surface reflectance properties in computer vision. The Lambertian model, often used to model reflectances of matte appearance, is simpler and predicts that light incident on a surface is reflected equally in all directions. Phong's model (Phong 1975) has found great acceptance within the computer graphics community and has become the industry standard. In this section, we discuss both models in detail. Finally, we present the dichromatic reflectance model, which is a physical model, and we show that both the Lambertian and Phong models are special cases of it (Klinker, Shafer, and Midha, 1988; Shafer, 1985).

5.1. Lambertian Reflectance Model

A Lambertian surface appears equally bright from any direction, for any illumination direction. Therefore, its reflected intensity is independent of the viewing direction. The BRDF function that can be deduced from this definition must be equal to a constant: $f_{\text{BRDF}}(\lambda, \theta_l, \varphi_l, \theta_n, \varphi_n) = r(\lambda) \times constant$. In order to determine the value of this constant, we proceed as follows. An elementary patch dB on the surface of a sphere of radius R subtends a solid angle $d\omega = \frac{dB}{R^2}$ with respect to the center A of the sphere (Figure 8). However, an elementary patch of the same area ($dA = dB$), placed at the center of the sphere, subtends an angle $d\Omega = \frac{dA \cos\theta_n}{R^2}$ as seen from point B. Thus, we may say $d\Omega = d\omega \cos\theta_n$. In terms of spherical coordinates, $d\omega = \sin\theta_n d\theta_n d\varphi_n$, so $d\Omega = \sin\theta_n \cos\theta_n d\theta_n d\varphi_n$. Note that $d\Omega$ is the solid angle of the *viewer*. To estimate how much light point B receives, we must integrate f_{BRDF} over all values of $d\Omega$, since f_{BRDF} expresses the fraction of energy received by the viewer per unit sterad. Under the Lambertian assumption, $\tilde{I}_{from\_surface}(\theta_n, \varphi_n) = \tilde{E}_{on\_surface}(\theta_l, \varphi_l)$ (i.e., the light the viewer sees is independent of the imaging geometry), and so, from Eq. (18) we have

$$\begin{aligned}
\int f_{\text{BRDF}}(\lambda) d\Omega &= r(\lambda) \\
\Rightarrow f_{\text{BRDF}}(\lambda) \int_{-\pi}^{\pi} \int_{0}^{\pi/2} \sin\theta_n \cos\theta_n d\theta_n d\varphi_n &= r(\lambda) \\
\Rightarrow 2\pi f_{\text{BRDF}}(\lambda) \int_{0}^{\pi/2} \sin\theta_n \cos\theta_n d\theta_n &= r(\lambda) \\
\Rightarrow \pi f_{\text{BRDF}}(\lambda) \int_{0}^{\pi/2} \sin 2\theta_n d\theta_n &= r(\lambda) \\
\Rightarrow f_{\text{BRDF}}(\lambda) = \frac{r(\lambda)}{\pi}. \qquad (19)
\end{aligned}$$

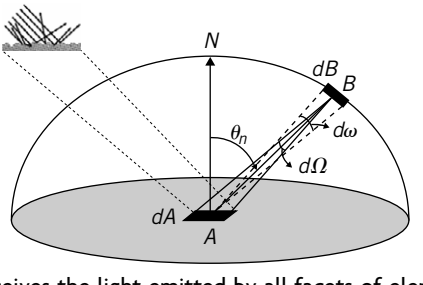


FIGURE 8 Point *B* receives the light emitted by all facets of elementary patch *dA*.

Thus, using the BRDF function for a Lambertian surface, we obtain from Eq. (17) that the surface radiance is proportional to the surface irradiance

$$I_{from\_surface}(\lambda) = \frac{r(\lambda)}{\pi} E_{on\_surface}(\lambda) = \frac{r(\lambda)}{\pi} I_{from\_source}(\lambda) \cos\theta_i \quad \text{for } \cos\theta_i \geq 0, \tag{20}$$

where I_{from_source} indicates the source radiance. The dependence on the incident angle θ_i comes from the geometry between the illuminating source and the surface. If the surface is face-on opposite the illuminating source (i.e., $\vec{N} = \vec{L}$ in Figure 7), whatever light energy leaves the source reaches the surface. If, however, $\vec{N} \neq \vec{L}$, the surface will receive only as much light as its effective area is in the path of the light. If $\vec{N} \perp \vec{L}$, the light from the source will just graze the surface ($\cos\theta_i = 0$) and the surface will receive no light energy. In that case, the surface is expected to appear totally black to the viewer. Equation (20) may be combined with Eq. (13) to allow us to say that the brightness observed on the image is directly proportional to the cosine of the angle between the surface normal and the direction of illumination. We wrap up all constants of proportionality that appear in these equations into a single function $\rho(\lambda)$ and refer to this as the albedo of the imaged surface, all other factors ignored.[1] Thus, we may write

$$\begin{aligned} I(\lambda, x, y) &= I_{from\_source}(\lambda)\rho(\lambda)\cos\theta_i = \rho(\lambda)(\vec{N} \cdot \vec{L}) \\ &= \rho(\lambda)\frac{1 + p_l p_n + q_l q_n}{\sqrt{1 + p_n^2 + q_n^2}\sqrt{1 + p_l^2 + q_l^2}} \\ &= \rho(\lambda)\frac{-p_n \cos\varphi_l \sin\theta_l - q_n \sin\varphi_l \sin\theta_l + \cos\theta_l}{\sqrt{p_n^2 + q_n^2 + 1}}, \end{aligned} \tag{21}$$

where the following apply:

- $I(x, y)$ is the image intensity;
- $\vec{N}$ is the unit vector normal to the surface $z(x, y)$ at point (x, y);
- p_n and q_n are partial derivatives of the surface with respect to the x and y coordinates, respectively [see Eq. (2)];
- $\vec{L}$ is the unit vector toward the light source (Figure 1);
- p_l and q_l are the partial derivatives of a virtual planar surface orthogonal to the direction of the illumination, with respect to the x and the y coordinates, respectively [see Eq. (7)];

[1] We may ignore factors of proportionality in brightness because the gray values of an image have significance only in relation to each other (i.e., in relative terms and not in absolute terms). Thus, $r(\lambda)$ of Eq. (18) and $\rho(\lambda)$ of Eq. (21) are both the surface albedo, up to a constant of proportionality.

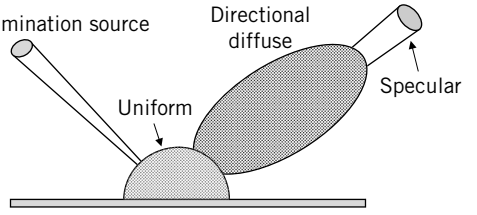


FIGURE 9 Components of the Phong reflectance model.

- θ_l and φ_l are the illumination vector's zenith and azimuth angles defined in Figure 1;
- $\rho(\lambda)$ is the surface albedo, and
- λ is the wavelength of the light.

5.2. Phong Reflectance Model

Phong's model is more complete and realistic compared with the diffuse reflectance model. The linear combination of three components—namely, diffuse, specular, and ambient components—constitute the Phong model. The *diffuse component* is nothing but a Lambertian component. The *specular component* forms a lobe of reflected light that spreads out around the specular direction and is modeled by a cosine function raised to a power. The *ambient component* accounts for the ambient light caused by interreflections. It is this light that causes the shadows not to be black.[2] This model is shown schematically in Figure 9 and mathematically expressed as

$$\begin{aligned} I(\lambda, n, \beta) &= I_\alpha(\lambda)k_\alpha(\lambda) + I_{from\_source}(\lambda)\rho(\lambda)\cos\theta_i \\ &\quad + E_{on\_surface}(\lambda)k_s(\lambda)\cos^n\beta \\ &= I_\alpha(\lambda)k_\alpha(\lambda) + I_{from\_source}(\lambda)\rho(\lambda)(\vec{L}\cdot\vec{N}) \\ &\quad + E_{on\_surface}(\lambda)k_s(\lambda)(\vec{R}\cdot\vec{V})^n, \end{aligned} \tag{22}$$

where $I_\alpha(\lambda)$ is the ambient constant light intensity, $k_\alpha(\lambda)$ is the reflectance coefficient of the surface, and $k_s(\lambda)$ is the specular reflectance coefficient. β denotes the angle between the viewing direction $\vec{V}$ and the perfect reflector vector $\vec{R}$, and n controls the width of the specular lobe.

It can be determined that direction $\vec{R}$ is given by $\vec{R} = 2\vec{N}(\vec{N}\cdot\vec{L}) - \vec{L}$ (Figure 10). This computation is time consuming, so angle β in the model often is replaced by angle α (Figure 10). Vector $\vec{H}$ is halfway between

[2] In the Lambertian model, the shadows are black.

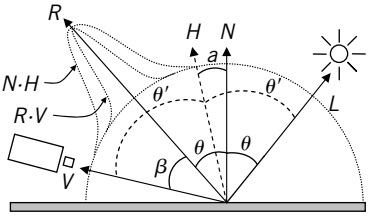


FIGURE 10 Angle β between the specular $\vec{R}$ and the viewing direction, and its approximation by the angle α between the halfway vector $\vec{H}$ and normal $\vec{N}$.

vectors $\vec{L}$ and $\vec{V}$

$$\vec{H} = \frac{\vec{L} + \vec{V}}{2}. \tag{23}$$

This is the required orientation for a surface to reflect light along direction $\vec{V}$.

If $\vec{V}$, $\vec{L}$, and $\vec{N}$ lie in the same plane (Blinn, 1977), α is equal to half of angle β: $\alpha = \theta' - \theta = \frac{\theta+\theta+\beta}{2} - \theta = \frac{\beta}{2}$. Thus, using angle α instead of angle β in Eq. (22) changes the model. However, because this is an empirical model, exponent n can be adjusted to achieve the desirable effect. As α is smaller than β, $\cos\alpha$ is larger than $\cos\beta$, and to achieve the same effect as when using β, we must increase exponent n to $\tilde{n} > n$:

$$\begin{aligned} I(\lambda, n, \alpha) = {} & I_\alpha(\lambda) k_\alpha(\lambda) + I_{from\_source}(\lambda)\rho(\lambda)(\vec{L} \cdot \vec{N}) \\ & + E_{on\_surface}(\lambda) k_s(\lambda)(\vec{N} \cdot \vec{H})^{\tilde{n}}. \end{aligned} \tag{24}$$

This approximation means that intensity I is solely a function of the surface normal $\vec{N}$. If we assume the viewer and the light are at infinity, then the half-angle vector $\vec{H}$ is independent of position and surface curvature. Therefore, it can be calculated once for each light source and used for the entire frame, or while the viewpoint and the light remain in the same relative position. In Phong's original model [Eq. (22)], the reflected light vector $\vec{R}$ depends on the surface curvature and must be calculated for each pixel of the rendered image.

5.3. Dichromatic Reflectance Model

The Dichromatic reflectance model is a physical model that attempts to model the physical process that occurs when light reaches the surface of an opaque object. According to this theory, the light I, which is reflected from a point of the surface of an object, is a mixture of the light I_{spec} reflected by the surface itself (surface, or specular, reflection) and the light I_{body} reflected

by the body of the object (body, or Lambertian, or matte reflection):

$$\begin{aligned} I(\lambda, \theta_i, \theta_l, \theta_n) &= I_{spec}(\lambda, \theta_i, \theta_l, \theta_n) + I_{body}(\lambda, \theta_i, \theta_l, \theta_n) \\ &= m_s(\theta_i, \theta_l, \theta_n) w(\lambda) + m_b(\theta_i, \theta_l, \theta_n) c(\lambda). \end{aligned} \tag{25}$$

Function $I_{spec}(\lambda, \theta_i, \theta_l, \theta_n)$ is assumed separable (Nicodemus *et al.*, 1977) and it may be written as the product of a spectral power distribution, $w(\lambda)$, and a geometric scale factor, $m_s(\theta_i, \theta_l, \theta_n)$. Similarly, the body reflection component $I_{body}(\lambda, \theta_i, \theta_l, \theta_n)$ is also assumed to be a separable function, and it may be written as the product of a spectral power distribution, $c(\lambda)$, and a geometric scale factor, $m_b(\theta_i, \theta_l, \theta_n)$.

The light reflected by the surface has approximately the same spectral power distribution as the light source. The light that is not reflected by the surface penetrates into the material body where it is scattered and selectively absorbed. Some fraction of the light arrives again at the surface and is re-emitted. The light traveling through the body is increasingly absorbed at wavelengths that are characteristic of the material of the viewed object. The body reflectance provides the characteristic object color (Klinker, 1993).

By comparing Eq. (25) with Eq. (21), we may easily infer that the Lambertian model assumes that

$$\begin{aligned} w(\lambda) &= 0 \\ m_b(\theta_i, \theta_l, \theta_n) &= \cos\theta_i \\ c(\lambda) &= I_{from\_source}(\lambda)\rho(\lambda). \end{aligned} \tag{26}$$

Note that under the assumption of a white light illumination source, $I_{from\_source}(\lambda) = constant$, and the color of the object $c(\lambda)$ is truly the observed color $\rho(\lambda)$ by the imaging sensor. Under any other light source, we need to know exactly the spectrum of the light $I_{from\_source}(\lambda)$ to work out $c(\lambda)$ from $\rho(\lambda)$.

Comparing Eq. (25) with Eq. (22), we see that according to the Phong model

$$\begin{aligned} w(\lambda) &= E_{on\_surface}(\lambda) k_s(\lambda) \\ m_s(\theta_i, \theta_l, \theta_n) &= (2\cos\theta_i \cos\theta_n - \cos\theta_l)^n \\ c(\lambda) &= I_{from\_source}(\lambda)\rho(\lambda) \\ m_b(\theta_i, \theta_l, \theta_n) &= \cos\theta_i. \end{aligned} \tag{27}$$

The Phong model adds also a term of uniformly distributed ambient light, which could be due to the object itself or to other objects. One might consider $I_\alpha(\lambda)$ as an extra illuminating source, which, however, is diffuse and

nondirectional. The effect of such a light on the surface cannot possibly depend on angles θ_i and θ_l, because these angles cannot be defined. Thus, this term may be considered another body (Lambertian) component, with geometric factor $m_{b\alpha} = 1$ and spectral factor $c_\alpha(\lambda) = I_\alpha(\lambda)c(\lambda)$.

6. SURFACE-RECOVERING METHODS

We may state that we know the shape of a surface if we know the normal vector $\vec{N}$ or vector $(-p_n(x,y), -q_n(x,y), 1)^T$ of every facet (x,y) that makes up the surface [see Eqs. (4) and (5)]. Alternatively, we may say that we know the shape of a surface if we know the distance of every point (x,y) of the surface from some reference plane. However, a unique facet orientation or point distance cannot be determined from a single image intensity or radiance value observed, since an infinite number of surface orientations can give rise to the same value of image intensity. This is obvious since the recorded brightness of a facet is only one piece of information, whereas the facet orientation $(-p_n(x,y), -q_n(x,y), 1)$ has two degrees of freedom. Therefore, we need additional information to determine local surface orientation. Several approaches exist for the recovery of surface relief, including binocular stereo, shape from shading, and photometric stereo.

6.1. Binocular Stereo

Binocular stereo is used to recover surface distance by identifying corresponding points in two images taken from different viewpoints. *Stereoscopic vision* has been used successfully in cartography and robot navigation but it has several drawbacks. The difficulty of applying binocular stereo arises from reliably determining the corresponding features between two separate images and, if surface detail is the main concern, the problem becomes more evident. In addition, surface depth is recovered rather than surface orientation, as illustrated in Figure 11. This introduces noise and artifacts. The surface depth z is obtained as

$$z = \frac{f\ d}{x_1 - x_2}, \tag{28}$$

where f is the camera focal length, d the distance between the two cameras, and x_1 and x_2 are the coordinate positions of the images of the same physical point on the two focal planes of the cameras, with respect to coordinate systems that are fully aligned, but each is centered on the axis of the corresponding camera lens.

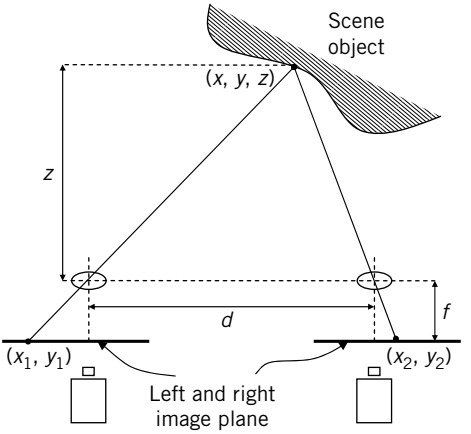


FIGURE 11 Geometry of binocular stereo method.

6.2. Shape From Shading from a Single Image

The topic of shape from shading (SFS) deals with the determination of an object's shape solely from the intensity variation in the image plane. For a Lambertian surface and for a given illumination direction (φ_l, θ_l), Eq. (21) may be viewed as a function that relates the brightness of a pixel (x, y) with the orientation $(-p_n, -q_n, 1)$ of the facet of the surface depicted by the pixel.

If we multiply both sides of Eq. (21) with the sensitivity function $s(\lambda)$ of the imaging sensor we use and integrate over all λ, the left-hand side becomes the brightness of pixel (or sensor element) (x, y) as recorded by the imaging sensor we use. On the right-hand side integration over λ results in yet another version of the albedo of the surface, this time closer to what can actually be measured. Thus, Eq. (21) may eventually be simplified to

$$I(x,y) = \rho_S(x,y)\frac{-p_n(x,y)\cos\varphi_l \sin\theta_l - q_n(x,y)\sin\varphi_l \sin\theta_l + \cos\theta_l}{\sqrt{p_n^2(x,y) + q_n^2(x,y) + 1}}, \tag{29}$$

where we have made explicit the dependence of p_n, q_n, and ρ_S on the coordintates of the facet they refer to, and $\rho_S \equiv \int_\lambda \rho(\lambda)s(\lambda)d\lambda$. Note that ρ_S depends on the sensor characteristics and I now may be treated as the gray value of the image produced by this sensor. If the sensor sensitivity function $s(\lambda)$ were a delta function, ρ_S would have been the value of the true albedo $r(\lambda)$ of Eq. (18) of the imaged surface at point (x, y), at the wavelength to which the sensor we use is sensitive, times some constant of proportionality. The right-hand side of Eq. (29) is a function of $p_n(x, y)$ and

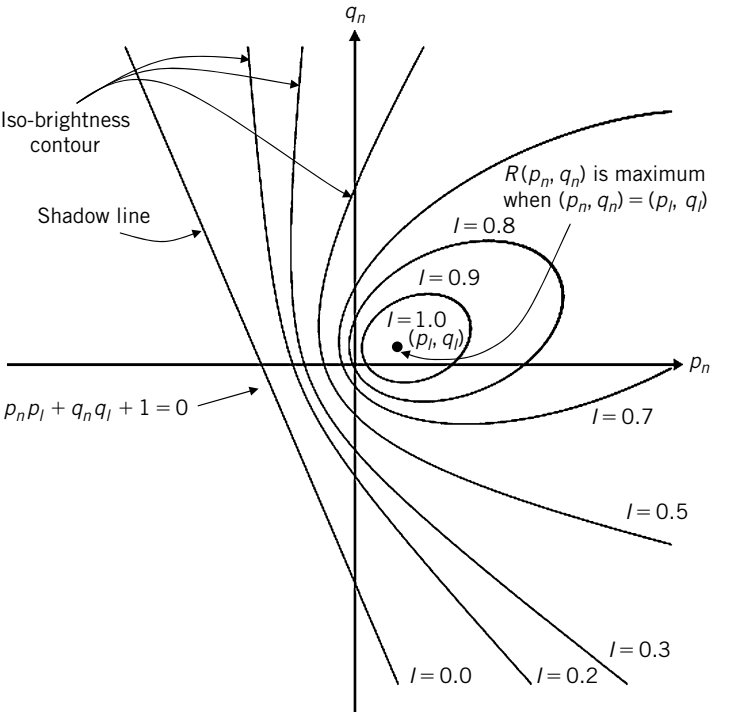


FIGURE 12 Reflectance map $R(p_n, q_n)$ for a Lambertian surface illuminated from a direction perpendicular to the point with $p_l = 0.7$ and $q_l = 0.3$ and for $\rho_S = 1$. All points on each iso-brightness contour are solutions of Eq. (29) that relates image intensity to surface orientation. The straight line corresponds to the so-called shadow line. Points along this line have gradient vector orthogonal to the illumination direction, so they receive only grazing light. All points with orientations on the left of that line receive no light and so they have $I = 0$.

$q_n(x, y)$ and may be called $R(p_n(x, y), q_n(x, y))$. If we plot $q_n(x, y)$ versus $p_n(x, y)$ for certain values of $I(x, y)$, we produce the so-called reflectance map (Figure 12). An analogous analysis may be made for the Phong model, expressed either by Eq. (22) or Eq. (24). In all cases, Eqs. (21), (22), or (24) may be expressed as

$$I(x, y) = R(p_n(x, y), q_n(x, y)). \tag{30}$$

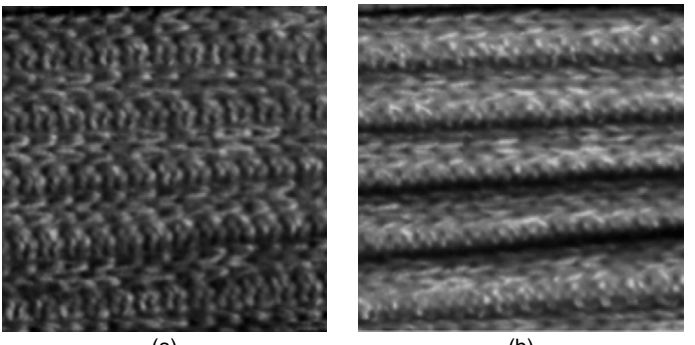
The above equation has two unknowns, namely, $p_n(x, y)$ and $q_n(x, y)$, and, therefore, determining object shape from image intensity alone is impossible. For a given reflectance model and given imaging geometry, function $R(p_n(x, y), q_n(x, y))$ is a function of the two unknowns and it can be plotted in a two-dimensional (2D) $(p_n(x, y), q_n(x, y))$ coordinate system. Figure 12

shows the contours of constant value of R as functions of (p_n, q_n) for the Lambertian model. Therefore, if we know the brightness of a pixel, its $(p_n(x,y), q_n(x,y))$ values could be any point on the contour of that brightness. Additional constraints are required to calculate object shape (i.e., to identify the $(p_n(x,y), q_n(x,y))$ values of every pixel in the image of the object). This was one of the first areas of study in computer vision. The initial work was done by Horn (Horn, 1975; Horn, 1986; Horn and Brooks, 1989; Horn and Sjoberg, 1979). Horn's method of solving this problem relies on growing a solution by starting at a single point in the image plane $I(x_0, y_0)$ where the surface orientation is known. The single-image shape from shading algorithm is still limited even if the exact lighting conditions and surface reflectivity are known. The original shape from shading method has been advanced by many researchers in recent years. Most of these methods rely on the use of a regularization term that controls the smoothness of the created surface (e.g., Bors, Hancock, and Wilson, 2003, 2000a,b,c; Worthington and Hancock, 1997; Worthington and Hancock, 1999; Zhang *et al.*, 1999) or the roughness of the created surface when the surface is known to exhibit fractal characteristics (Liao, Petrou, and Zhao, 2008).

6.3. Photometric Stereo

Photometric stereo is a reflectance map–based technique that uses two or more images to solve the underdetermined problem of recovering surface shape from a single image. Photometric stereo allows estimation of local surface orientation by using several images of the same surface taken from the same viewpoint but under illuminations from different directions (Woodham, 1980). The light sources are ideally point sources, some distance away in different directions, so that in each case there is a well-defined light source direction. The variation of the intensities observed in an image depends on variation in both surface reflectance and surface relief. Although the reflectance properties are intrinsic to a surface, the surface relief produces a pattern of shadings that depends strongly on the direction of the illumination. The appearance of a 3D surface changes drastically with illumination. An example is shown in Figure 13. The idea is to use this information to recover the intrinsic surface parameters—that is, local surface orientation and reflectance—independent of the illumination direction.

Woodham (1980) was the first to introduce photometric stereo. He proposed a method that was simple and efficient, but it dealt only with Lambertian surfaces and was sensitive to noise. In Woodham's method, the surface gradient may be recovered by using two photometric images, assuming that the surface albedo is already known for each point on the surface. Coleman and Jain (1982) extended photometric stereo to four light sources, where specular reflections were discarded and estimation



(a) (b)

FIGURE 13 An example that shows how the appearance of the same 3D surface changes drastically with illumination direction.

of surface shape could be performed by means of diffuse reflections and the Lambertian model. Nayar, Ikeuchi, and Kanade (1990) developed the photometric approach, which uses a linear combination of Lambertian and an impulse specular component to obtain the shape and reflectance information for a surface.

The simplest approach is to take two images of the same surface under different illumination directions. Therefore, two values of image intensity, $I^1(x,y)$ and $I^2(x,y)$, at each point (x,y) can be obtained. In general, the image intensity values of each light source correspond to two contours on the reflectance map, namely,

$$\begin{aligned} I^1(x,y) &= R^1(p_n(x,y), q_n(x,y)) \\ I^2(x,y) &= R^2(p_n(x,y), q_n(x,y)). \end{aligned} \tag{31}$$

This approach gives a unique solution for surface orientation at almost all points in the image, since the $(p_n(x,y), q_n(x,y))$ values of a point are determined by the intersection of the iso-brightness contours (reflectance functions) on which it lies (Figure 14).

In general, Eqs. (31) are nonlinear, so more than one solution is possible (contours may intersect at two points; twofold ambiguity). One solution would be to construct the reflectance function for a third illumination vector (Figure 15)

$$I^3(x,y) = R^3(p_n(x,y), q_n(x,y)) \tag{32}$$

to remove such ambiguities (Figure 16). This also allows estimation of the third surface parameter, $\rho_S(x,y)$, which is especially useful in cases where a surface is not uniform in its reflectance properties.

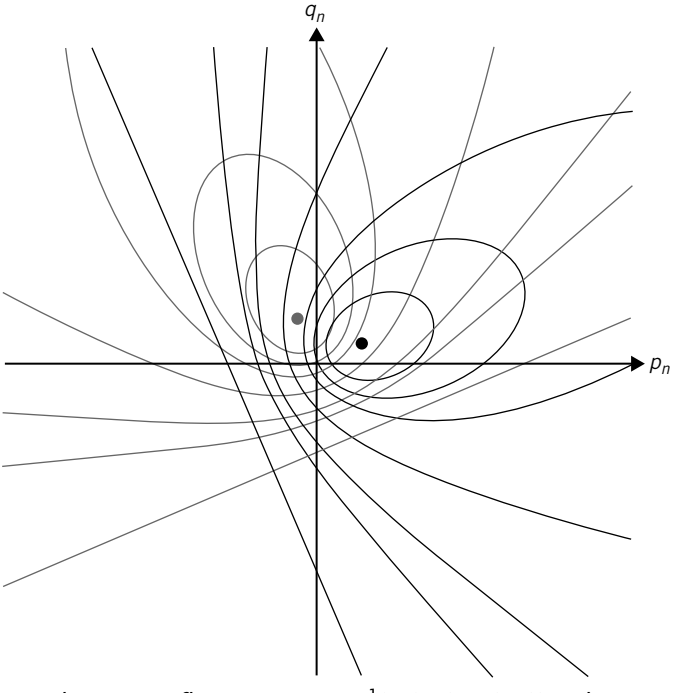


FIGURE 14 At least two reflectance maps—$R^1(p_n(x,y), q_n(x,y))$ and $R^2(p_n(x,y), q_n(x,y))$—are required to determine the surface orientation if the reflectance factor $\rho_S(x,y)$ is known.

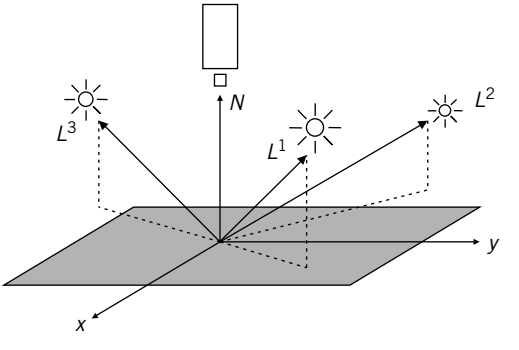


FIGURE 15 Illustration of photometric stereo geometry with three lighting directions.

Let us consider a Lambertian surface illuminated in turn by three illumination sources with directions $\vec{L}^1$, $\vec{L}^2$, and $\vec{L}^3$. In this case, the intensities of the obtained pixels can be expressed as

$$I^k(x,y) = \rho_S(x,y)\vec{L}^k \cdot \vec{N}(x,y), \text{ where } k = 1,2,3. \tag{33}$$

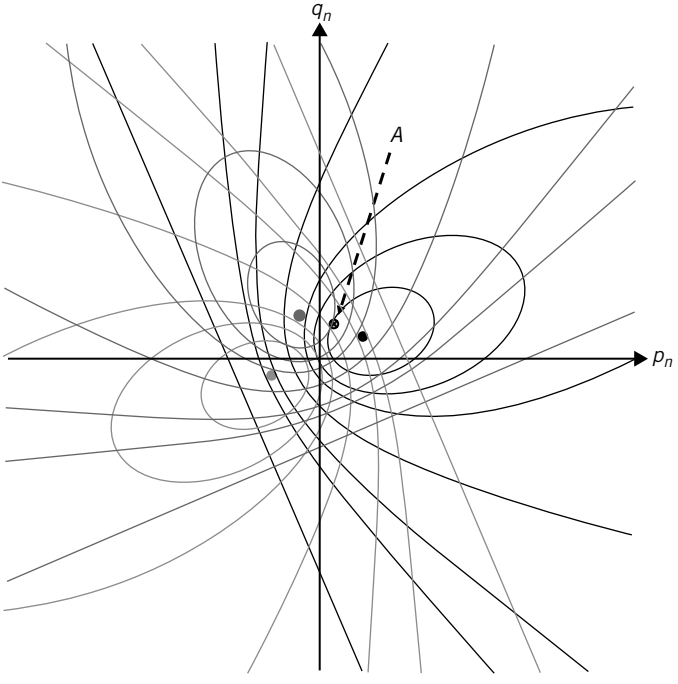


FIGURE 16 Three views are sufficient to determine both surface orientation and surface albedo uniquely.

We stack the intensities and the illumination vectors to form the pixel intensity vector $\vec{I}(x,y) = (I^1(x,y), I^2(x,y), I^3(x,y))$ and the illumination matrix $[L] = (\vec{L}^1, \vec{L}^2, \vec{L}^3)^T$. Then Eq. (33) may be rewritten in matrix form as

$$\vec{I}(x,y) = \rho_S(x,y)[L]\vec{N}(x,y). \tag{34}$$

If the three illumination vectors $\vec{L}^k$ do not lie in the same plane (noncoplanar), then the photometric illumination matrix $[L]$ is nonsingular and can be inverted, giving

$$\vec{M}(x,y) = [L]^{-1}\vec{I}(x,y) = \rho_S(x,y)\vec{N}(x,y), \tag{35}$$

where $\vec{M}(x,y) = (m_1(x,y), m_2(x,y), m_3(x,y))^T$. The surface gradient components could be obtained from $p_n(x,y) = -m_1(x,y)/m_3(x,y)$ and $q_n(x,y) = -m_2(x,y)/m_3(x,y)$, and the surface albedo is recovered by calculating the length of vector $\vec{M}$, $\rho_S(x,y) = \sqrt{m_1^2(x,y) + m_2^2(x,y) + m_3^2(x,y)}$.

Equation system (34) may also be written using the gradient representation of the surface orientation and direction of light source as

$$I^1(x,y) = \rho_S(x,y)\frac{1+p_n(x,y)p_{l_1}+q_n(x,y)q_{l_1}}{\sqrt{1+p_n^2(x,y)+q_n^2(x,y)}\sqrt{1+p_{l_1}^2+q_{l_1}^2}}$$
$$I^2(x,y) = \rho_S(x,y)\frac{1+p_n(x,y)p_{l_2}+q_n(x,y)q_{l_2}}{\sqrt{1+p_n^2(x,y)+q_n^2(x,y)}\sqrt{1+p_{l_2}^2+q_{l_2}^2}}$$
$$I^3(x,y) = \rho_S(x,y)\frac{1+p_n(x,y)p_{l_3}+q_n(x,y)q_{l_3}}{\sqrt{1+p_n^2(x,y)+q_n^2(x,y)}\sqrt{1+p_{l_3}^2+q_{l_3}^2}}. \quad (36)$$

Hence, the values of $p_n(x,y)$ and $q_n(x,y)$ may be determined uniquely by solving the above set of three nonlinear equations, given the image intensity values $I^1(x,y)$, $I^2(x,y)$, and $I^3(x,y)$ and the gradients of the three light sources $(-p_{l_1}, -q_{l_1}, 1)$, $(-p_{l_2}, -q_{l_2}, 1)$, and $(-p_{l_3}, -q_{l_3}, 1)$, respectively.

We define $A_1 \equiv \sqrt{1+p_{l_1}^2+q_{l_1}^2}$, $A_2 \equiv \sqrt{1+p_{l_2}^2+q_{l_2}^2}$, and $A_3 \equiv \sqrt{1+p_{l_3}^2+q_{l_3}^2}$. By pairwise division by parts, we may eliminate $\rho_S(x,y)$ from Eq. (36). Then

$$\frac{I^1(x,y)A_1}{I^2(x,y)A_2} = \frac{1+p_n(x,y)p_{l_1}+q_n(x,y)q_{l_1}}{1+p_n(x,y)p_{l_2}+q_n(x,y)q_{l_2}}$$
$$\frac{I^2(x,y)A_2}{I^3(x,y)A_3} = \frac{1+p_n(x,y)p_{l_2}+q_n(x,y)q_{l_2}}{1+p_n(x,y)p_{l_3}+q_n(x,y)q_{l_3}}$$
$$\frac{I^3(x,y)A_3}{I^1(x,y)A_1} = \frac{1+p_n(x,y)p_{l_3}+q_n(x,y)q_{l_3}}{1+p_n(x,y)p_{l_1}+q_n(x,y)q_{l_1}}. \quad (37)$$

Cross-multiplication and transposition allows us to write the above equations as

$$p_n(I^2A_2p_{l_1} - I^1A_1p_{l_2}) = I^1A_1 - I^2A_2 + q_n(q_{l_2}I^1A_1 - q_{l_1}I^2A_2)$$
$$p_n(I^3A_3p_{l_2} - I^2A_2p_{l_3}) = I^2A_2 - I^3A_3 + q_n(q_{l_3}I^2A_2 - q_{l_2}I^3A_3)$$
$$p_n(I^1A_1p_{l_3} - I^3A_3p_{l_1}) = I^3A_3 - I^1A_1 + q_n(q_{l_1}I^3A_3 - q_{l_3}I^1A_1), \quad (38)$$

where for simplicity, we dropped the explicit dependence on (x,y). Any pair of the above set of equations may be used to yield a value of q_n by division by parts and rearrangement. We adopt as value of q_n the average

of these three solutions:

$$q_n = \frac{1}{3}\left(\frac{(I^1A_1 - I^2A_2)(I^3A_3p_{l_2} - I^2A_2p_{l_3}) - (I^2A_2 - I^3A_3)(I^2A_2p_{l_1} - I^1A_1p_{l_2})}{(I^2A_2p_{l_1} - I^1A_1p_{l_2})(I^2A_2q_{l_3} - I^3A_3q_{l_2}) - (I^3A_3p_{l_2} - I^2A_2p_{l_3})(I^1A_1q_{l_2} - I^2A_2q_{l_1})}\right.$$
$$+ \frac{(I^2A_2 - I^3A_3)(I^1A_1p_{l_3} - I^3A_3p_{l_1}) - (I^3A_3 - I^1A_1)(I^3A_3p_{l_2} - I^2A_2p_{l_3})}{(I^3A_3p_{l_2} - I^2A_2p_{l_3})(I^3A_3q_{l_1} - I^1A_1q_{l_3}) - (I^1A_1p_{l_3} - I^3A_3p_{l_1})(I^2A_2q_{l_3} - I^3A_3q_{l_2})}$$
$$\left. + \frac{(I^3A_3 - I^1A_1)(I^2A_2p_{l_1} - I^1A_1p_{l_2}) - (I^1A_1 - I^2A_2)(I^1A_1p_{l_3} - I^3A_3p_{l_1})}{(I^1A_1p_{l_3} - I^3A_3p_{l_1})(I^1A_1q_{l_2} - I^2A_2q_{l_1}) - (I^2A_2p_{l_1} - I^1A_1p_{l_2})(I^3A_3q_{l_1} - I^1A_1q_{l_3})}\right). \tag{39}$$

Each of Eqs. (38) may be solved for p_n in terms of q_n. The value of p_n we adopt is the average of these three solutions:

$$p_n = \frac{1}{3}\left(\frac{I^3A_3 - I^1A_1 + q_n(I^3A_3q_{l_1} - I^1A_1q_{l_3})}{I^1A_1p_{l_3} - I^3A_3p_{l_1}}\right.$$
$$+ \frac{I^1A_1 - I^2A_2 + q_n(I^1A_1q_{l_2} - I^2A_2q_{l_1})}{I^2A_2p_{l_1} - I^1A_1p_{l_2}}$$
$$\left. + \frac{I^2A_2 - I^3A_3 + q_n(I^2A_2q_{l_3} - I^3A_3q_{l_2})}{I^3A_3p_{l_2} - I^2A_2p_{l_3}}\right). \tag{40}$$

From the knowledge of $p_n(x,y)$ and $q_n(x,y)$ then one may derive from Eq. (36) three values of $\rho_S(x,y)$, which may be averaged to yield an estimate of the albedo of facet (x,y). Clearly, a more elegant solution would be to solve system (38) in the least-square error sense. Section 11 shows how the least-square error solution may be obtained in the more general case of photometric stereo with perspective projection.

The above equations were derived under the following simplifying assumptions:

1. The images used are gray.
2. All pixels have the same illumination vector in each image.
3. The object size is much smaller than the viewing distance.
4. The observed object is directly illuminated without the presence of cast- or self-shadows.
5. The surface of the object is Lambertian.
6. The imaged surface is static during image capture.

In the sections that follow, we examine how each of these assumptions may be relaxed.

7. COLOR PHOTOMETRIC STEREO [ASSUMPTION (1)]

Color images typically consist of three bands identified with the red, green, and blue parts of the electromagnetic spectrum. In general, if more than one type of sensor is used to image the same scene, there will be several gray bands for each illumination direction, one photometric set for each sensor type. Therefore, we shall, have as many sets of Eqs. (36) as different sensors, identified by index S. We may solve each such set of equations following the method of the previous section, each time recovering the orientation of each facet and its albedo in the band of the corresponding sensor. The orientations derived in this manner may be averaged to yield a single orientation for each facet (Schluns and Wittig, 1993), while the albedos will correspond to the color of the surface in the bands of the sensors we used. The main problem of this approach is that in most cases the color bands have different intensities and the low-intensity bands are much more prone to noise corruption. Therefore, Barsky and Petrou (2001a,b, 2006) suggested an algorithm that simultaneously recovers the optimal gradient and color estimates taking into consideration the relationship between the color bands. This algorithm is based on the fact that for a Lambertian surface patch, the position of a pixel in the 3D color space is shifted along a line as the level of brightness changes. Introduced errors may disturb the collinearity of the positions of a pixel in the different photometric sets, and, therefore, the principal direction corresponding to the chromaticity of the body color is obtained using principal component analysis, (Figure 17). The problem is then reduced to the monochromatic case by projecting all input pixels on the principal color line. The use of the first principal component instead of the individual color bands reduces the problem to that of gray image photometric stereo.

8. PHOTOMETRIC STEREO WITH HIGHLIGHTS AND SHADOWS [ASSUMPTIONS (4) AND (5)]

Photometric stereo, in order to recover color and gradient information of a surface patch, assumes that neither highlights nor shadows are present in the images used. In case this assumption is not valid, the surface recovery will be affected, with the body color appearing different and the normal leaning more toward the light source that produced the highlight (Figure 18a and b) or away from the source that produced the shadow (Figure 18c). Coleman and Jain (1982), and Solomon and Ikeuchi (1996) proposed a method of determining highlights in the absence of shadows utilizing four images of the same surface. Combining all the recovered albedos from all four possible triplets of pixels, the surface patch was regarded as a highlight if the albedos differed significantly from each other. Barsky

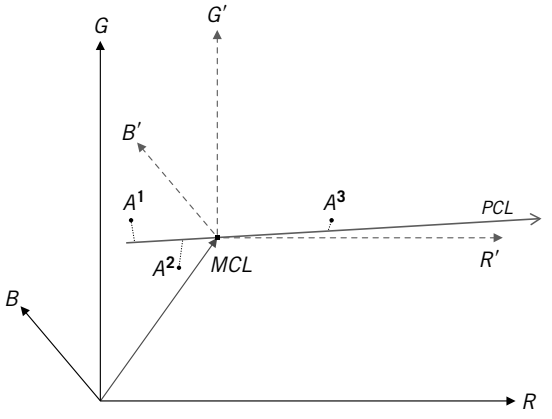


FIGURE 17 An example of reducing color photometric stereo to the monochromatic case based on the mean and the principal chromaticity line. Points A^1, A^2, and A^3 identify the positions of the same pixel in the red-green-blue (RGB) color space under the three different illumination directions. Vectors $\overrightarrow{MCL}$ and $\overrightarrow{PCL}$ are the mean and the principal color lines, respectively. The dotted lines indicate the projections of points A^1, A^2, and A^3 on the principal line. The R, G, and B axes correspond to the red, green, and blue components of the colored pixels. The R', G', and B' axes correspond to the color coordinate system shifted to the center of points A^1, A^2, and A^3.

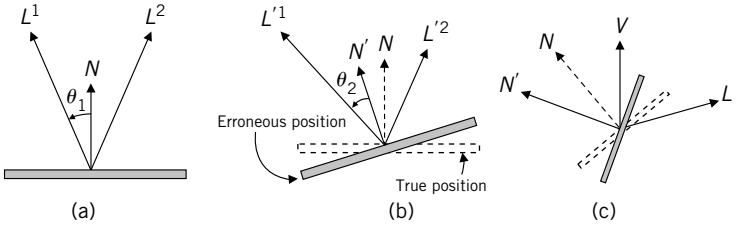


FIGURE 18 The effect of a highlight on the estimated normal of a facet. (a) With no highlights; (b) highlighted facet, where the erroneous normal $\vec{N}'$ leans toward the light source $\vec{L}'^1$ that produced the highlight; (c) the effect of a shadow on the estimated normal of a facet, where the erroneous normal $\vec{N}'$ leans away from the source $\vec{L}$, which produced the shadow.

and Petrou (2001b, 2003) presented an algorithm for separating the local gradient and color information by using four-source color photometric stereo in the presence of highlights and shadows. Based on the fact that any four vectors in a 3D space are linearly dependent, there must exist constants a_1, a_2, a_3, and a_4 such that the following equation holds for the four illumination direction vectors:

$$a_1\vec{L}^1 + a_2\vec{L}^2 + a_3\vec{L}^3 + a_4\vec{L}^4 = 0. \tag{41}$$

A similar equation for the four pixel intensities I^k may be obtained by multiplying the above equation with the albedo and the normal of the surface patch using vector dot product, as follows:

$$a_1\rho_S(\vec{L}^1 \cdot \vec{N}) + a_2\rho_S(\vec{L}^2 \cdot \vec{N}) + a_3\rho_S(\vec{L}^3 \cdot \vec{N}) + a_4\rho_S(\vec{L}^4 \cdot \vec{N}) = 0 \Rightarrow$$
$$a_1 I^1 + a_2 I^2 + a_3 I^3 + a_4 I^4 = 0. \quad (42)$$

The problematic pixel quadruples, containing either shadows or highlights, do not satisfy Eq. (42) and therefore this method allows ruling out the majority of quadruples that are not purely Lambertian. The use of color information was suggested to separate shadows from highlights. The color of a highlighted pixel is the combination of "matte" (body) I_{body} color, and the color of the illuminant (specular component) I_{spec} as expressed by Eq. (25).

Thus, the chromaticity of the brightest pixel was compared with the chromaticity of the darker pixels, and if the difference exceeded a certain threshold, the pixel was labeled as a highlight. This method cannot provide reliable classification if the chromaticity of the surface color is close to the chromaticity of the incident light. In this case, Barsky and Petrou (2003) suggested discarding the brightest pixel and reconstructing the normal using only the darkest three pixels; the pixel was regarded as a highlight if the recovered normal was close to the specular direction of the corresponding imaging configuration (vector $\vec{N}$ near vector $\vec{R}$ in Figure 10). In the rest of the problematic quadruples, the darkest pixel was discarded as a shadow, and the color and the normal were recovered using the brightest three pixels.

9. PHOTOMETRIC STEREO WITH AN EXTENDED LIGHT SOURCE [ASSUMPTION (2)]

In conventional photometric stereo, the viewing and lighting positions are assumed to be distant, with the lighting source modeled as a collimated point source. This assumption appears limiting for practical applications. Various authors (Farooq *et al.*, 2005; Lee *et al.*, 2005; Smith, 1999; Smith and Smith, 2005) have made a first attempt to solve the problem of the extended light source. This work is in relation to an industrial inspection problem, where strip lights are used to illuminate objects on a conveyor belt. Following their approach, we show here that under certain assumptions, an extended strip light may be replaced by a virtual point source at infinity.

The imaged object is on a conveyor belt, illuminated by a strip of light, and viewed by a line scan camera. Thus, the problem may be treated as a one-dimensional (1D) one. The following discussion is in relation to

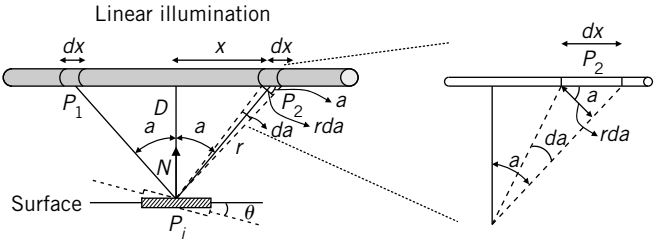


FIGURE 19 Distributed linear illumination.

Figure 19. Let us consider a strip light source with length much longer than the size of the imaged object. Let us consider a point P_i on the imaged object, at distance D from the line of the light. Let us consider two points P_1 and P_2 on the light source, symmetrically placed on either side of the point of interest P_i, their positions forming an angle α with the normal from point P_i to the linear light source.

Let us consider two infinitesimal elements at these points, each of length dx. If the brightness of the light source is I_0 per unit length, the light one of these elements produces is $I_0 dx$ and at point P_i creates an intensity $\frac{I_0 dx}{r^2}\cos\alpha$, where r is the distance $P_iP_1 = P_iP_2$ and factor $\cos\alpha$ appears because the normal of the surface patch at P_i is at angle α with respect to either direction P_iP_1 or direction P_iP_2. Figure 19 shows that $dx\cos\alpha = rd\alpha \Rightarrow dx = \frac{rd\alpha}{\cos\alpha}$ and $r\cos\alpha = D$. So, the light point P_i receives from segments P_1 and P_2 is $\frac{2I_0}{r^2}\frac{rd\alpha}{\cos\alpha}\cos\alpha = \frac{2I_0}{r}d\alpha = \frac{2I_0}{D}\cos\alpha d\alpha$. The total light, therefore, the surface element at point P_i receives is the integral of this expression over angle α. If the largest value angle α obtains is α_0, the total light at point P_i is $\frac{2I_0}{D}\sin\alpha_0$.

Next, let us assume that the imaged surface element is rotated by an angle θ, about the normal of the page, so the facet at point P_i now receives different amounts of light from segments P_1 and P_2. To avoid shadowing, $\theta < (\pi/2) - \alpha$. The intensity at point P_i of the surface is then given by

$$\begin{aligned}
&\frac{I_0 dx}{r^2}\cos(\alpha+\theta) + \frac{I_0 dx}{r^2}\cos(\alpha-\theta)\\
&\quad = \frac{I_0 dx}{r^2}[\cos(\alpha+\theta)+\cos(\alpha-\theta)]\\
&\quad = \frac{I_0 dx}{r^2}[2\cos\alpha\cos\theta]\\
&\quad = 2\frac{I_0 d\alpha}{D}\cos\alpha\cos\theta. \qquad (43)
\end{aligned}$$

Integrating then over α, we obtain the total illumination point P_i receives: $\frac{2I_0}{D}\sin\alpha_0\cos\theta$. This result indicates that the strip of light behaves as if

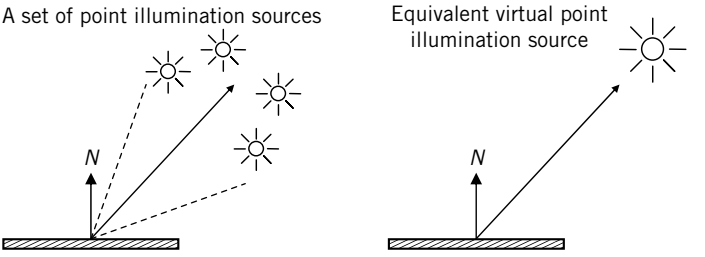


FIGURE 20 Approximating an extended light source.

a single point source is placed at the zenith of the imaged surface and at distance D from it, with intensity $I' \equiv 2I_0 \sin\alpha_0 D$, so that an elementary part of the surface tilted away from the average surface normal by an angle θ, receives light $\frac{I'}{D^2}\cos\theta = \frac{2I_0}{D}\sin\alpha_0\cos\theta$. Note that as long as the extent of the imaged object is much smaller than the length of the strip of light, this expression is valid for all points on the same line of the imaged surface parallel to the strip of light. Thus, all imaged points will receive equal intensity of light from a virtual point source that is above each of them and at distance D. This is equivalent to saying that the point source is at infinity with its rays arriving parallel at all imaged points. Note also that different lines on the surface, parallel to the strip of light, are at different distances D from that strip of light and thus receive different amounts of light.

The approximation of a linear distributed illumination source (long strip of light) by an equivalent virtual point source has practical interest for spatial multiplexing approaches (Figure 20). In this case, each light source directs the light onto a different area of the acquired surface at the same time, avoiding any photometric overlap. Therefore, the only practical way to use spatially multiplexed photometric stereo is in line scan fashion (Figure 21). The main drawbacks of this method are the need for registration among the interlaced images and the restriction to rigid objects avoiding self-motion between successive line scans.

10. DYNAMIC PHOTOMETRIC STEREO AND MOTION [ASSUMPTION (6)]

Many algorithms for shape reconstruction from multiple images of a moving Lambertian object have recently appeared in the literature (Lim, Yang, and Kriegman, 2005; Maki, Watanabe, and Wiles, 2002; Simakov and Basri, 2003; Smith and Smith, 2005; Zhang *et al.*, 2003). Smith and Smith (2005) considered the sixth main constraint of photometric stereo, namely, the limiting assumption that the relative position of the illumination configurations and the imaged surface must be fixed during image capture. This assumption is violated for rapid or uncontrolled relative movement between the

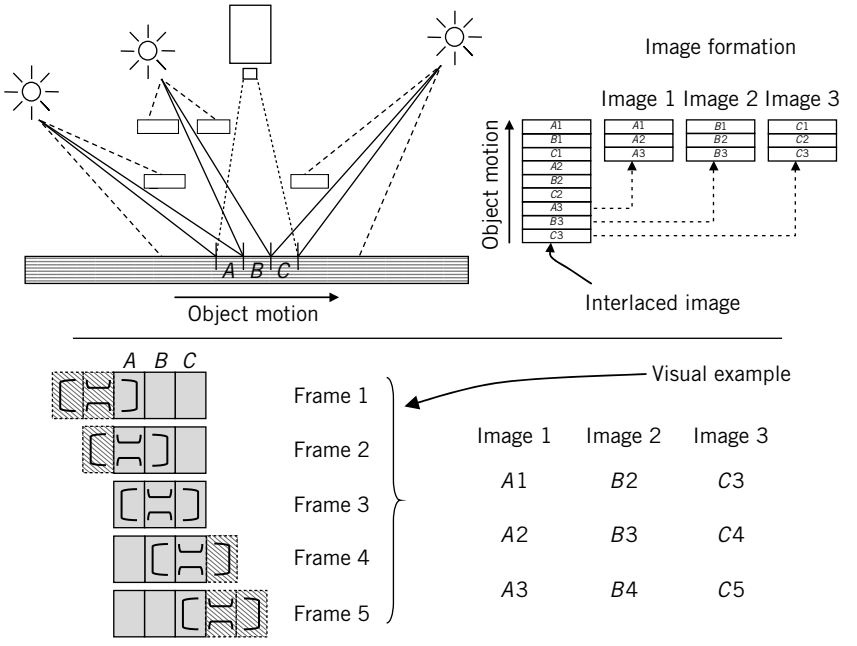


FIGURE 21 A schematic representation for spectral multiplexing. The shaded blocks in the top-left depiction of the imaging rig indicate barriers that prevent the light of one source illuminating the surface outside its designating area on the conveyor belt. Camera recordings *A*1, *B*1, and *C*1 are images of the same strip of the imaged surface under three different illumination directions. As this strip of surface moves out of the field of view of the camera, the next strip enters it and creates image segments *A*2, *B*2, and *C*2, which when added to the corresponding previously imaged strips help build the three images for the three different illumination directions.

camera and the surface, since the simultaneous acquisition of multiple images under different lighting conditions is required.

Several approaches for achieving dynamic photometric stereo have been proposed. Temporal multiplexing represents an adaptation of conventional photometric stereo to dynamic applications, in which separate lighting configurations are deployed and images are rapidly acquired at closely spaced intervals in time. This approach is inherently complex, and there are significant implications in terms of achievable camera frame rates and the high intensity of lighting needed to realize very short image acquisition periods, particularly if more than three images are required.

Spatial multiplexing involves separate images of the same surface location acquired from different points in space. Image capturing occurs simultaneously; therefore, the images must be registered between viewing positions and synchronized with the speed of the moving surface. This case is shown in Figure 21.

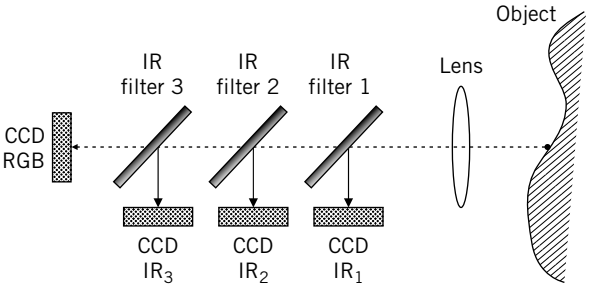


FIGURE 22 Narrow infrared (IR) photometric stereo. CCD, charge-coupled device.

Spectral multiplexing allows multiple images to be captured from a single point in space and time, since the cross-talk of image-illumination configuration can be minimized in terms of light frequency separation.

Smith and Smith (2005) suggest a narrow-band infrared photometric stereo technique utilizing narrow frequency channels that are closely spaced. This approach diminishes a number of key limitations that the use of broad-band color photometric stereo introduces. For example, similar components of colored light may be reflected in similar proportions, both for a surface of a particular color or inclination (e.g., a red surface with a certain inclination may appear similar to a pink one with a different inclination). Therefore, to reduce the sensitivity to surface color changes and decouple the gradient data from color, medium to long-wave infrared (IR) light was suggested. In addition, if required, a white light source may be simultaneously included to provide fully registered color data (Figure 22).

This approach, however, replaces the assumption that the optical albedo of the imaged surface is constant with the assumption that its IR albedo is constant. Any variation in the surface albedo will interfere with the spectral multiplexing of the lights used. In addition, the recovery of color by a synchronized color camera is imprecise, as the color inferred from a single image only is illumination dependent.

Another approach followed by many (Lim, Yang, and Kriegman, 2005; Maki, Watanabe, and Wiles, 2002; Simakov and Basri, 2003; Zhang *et al.*, 2003) is to assume an initial surface and then use an algorithm that iteratively estimates a new surface based on the previously reconstructed estimates. The algorithm assumes that the object can be segmented from the background and the scene illumination is modeled by a constant ambient illumination plus a directional source. One camera is used to capture the moving object. A sequence I_t of F frames, indexed by $t = 1, \ldots, F$ and m manually selected scene points, indexed by $p = 1, \ldots, m$, are considered. Let $\{P_1, \ldots, P_F\}$ denote the F orthogonal projections of points in $\Re^3$ to images $\{I_1, \ldots, I_F\}$, respectively. Each pair of (I_t, P_t) defines a mapping

from 3D to a gray value on the image plane. The object is assumed to be both rigid and Lambertian, and the observed intensity of a point p on the surface at frame t is given by Eq. (22) without the highlight term. Making explicit the application of the equation to a specific point p, and integrating out the dependence on the wavelength, we may write:

$$I_t(p) = I_\alpha + \rho(p)\vec{L}_t \cdot \vec{N}(p). \tag{44}$$

The m manually selected feature points are used to estimate camera projections, light source directions, and an initial piecewise surface. The camera projection matrix P_t and the 3D positions of the feature points with respect to the image plane of the initial frame I_1, can be recovered using the Tomasi–Kanade factorization algorithm for all frames (Tomasi and Kanade, 1992). An initial depth map is calculated by linearly interpolating the depth values of the selected m feature points. Given the initial surface S_t and its associated depth map $z_t(x, y)$, the intensity matrix $I_t(p) = I_t(P_t(x, y, z_t(x, y)))$ may be computed for all the pixels inside the segmented object region, which initially was not possible because the pixel correspondences across different frames were not known *a priori*.

Since the correspondence problem and the intensity matrix have been calculated, an integrable normal vector field $\vec{N}$ is then produced based on photometric stereo by minimizing the following function:

$$\sum_{t=1}^{F} \sum_{(x,y)\in\Omega} (I_t(P_t(x, y, z_t(x, y))) - \rho\vec{L}_t \cdot \vec{N}_p - I_\alpha)^2, \tag{45}$$

where Ω is a region in I_1 containing the segmented object. The normals may be recovered in the least-square error sense (see Barsky and Petrou, 2001b, 2003) and they may be integrated to obtain a new surface S_{t+1} and its associated depth map $z_{t+1}(x, y)$, while the known 3D positions of the m scene points may be used to refine the estimated surface. This process is repeated indefinitely and a sequence of the reconstructed surfaces can be estimated. Furthermore, other methods utilize a coarse-to-fine refinement step using image pyramids to improve the accuracy of the optical flow and the normal vector field (Zhang *et al.*, 2003).

11. PERSPECTIVE PHOTOMETRIC STEREO [ASSUMPTION (3)]

A common assumption in the field of computer vision and photometric stereo is that image projection is orthographic. This assumption is valid only for distant viewing positions. A perspective model should be introduced for positions closer to the observed object. Tankus *et al.* (Tankus,

Sochen, and Yeshurun, 2003, 2004, 2005; Tankus and Kiryati, 2005) and Prados and Faugeras (2003) reexamined the basic set of equations of photometric stereo under the assumption of perspective projection. In these papers, the scene object is assumed Lambertian and is illuminated from directions $\vec{L}^k = (-p_{l_k}, -q_{l_k}, 1)$, where $k = 1, 2, 3$. The acquired surfaces are assumed representable by functions of real-world coordinates $(x, y, z(x, y))$, as well as of image coordinates $(u, v, h(u, v))$. Function $z(x, y)$ denotes the depth in a Cartesian coordinate system, the origin of which is on the camera plane. Projecting the real coordinates $(x, y, z(x, y))$ onto the image point (u, v), with the depth denoted by $h(u, v)$, would result in $h(u, v) = z(x, y)$ by definition. Let f denote the known focal length of the camera, $\vec{N}(x, y)$ the surface normal, and $\rho_W(x, y)$ the albedo at point $(x, y, z(x, y))$, while $\rho_I(u, v)$ is the albedo projected onto image point (u, v), with $\rho_W(x, y) = \rho_I(u, v)$.

The image irradiance equation for a Lambertian surface in its general form is given by Eq. (21). The real-world coordinates $(x, y, z(x, y))$ under a perspective projection model are related to image coordinates (u, v) by

$$x = -\frac{u\, z(x, y)}{f}, \quad y = -\frac{v\, z(x, y)}{f}. \tag{46}$$

Equation (21) may be expressed with the current notation as

$$I(u, v) = \rho_W(x, y) \frac{1 + p_l z_x + q_l z_y}{\sqrt{1 + p_l^2 + q_l^2}\sqrt{1 + z_x^2 + z_y^2}}. \tag{47}$$

In order to derive the perspective image irradiance equation, the authors define a surface $S = \{(x, y, z(x, y))\}$ and its projection as $\tilde{S} = \left\{\left(\frac{-uh}{f}, \frac{-vh}{f}, h(u, v)\right)\right\}$. In addition, the surface is assumed differentiable with respect to (u, v) and (x, y). The setup is shown schematically in Figure 23.

A curve on the image plane with parameter s is defined as

$$c(s) = (u(s), v(s), -f\ ) \tag{48}$$

and due to the perspective projection the real-world curve may be written as

$$C(s) = \left(-\frac{u(s)h(s)}{f}, -\frac{v(s)h(s)}{f}, h(s)\right) = \frac{h(s)}{f}(-u, -v, f), \tag{49}$$

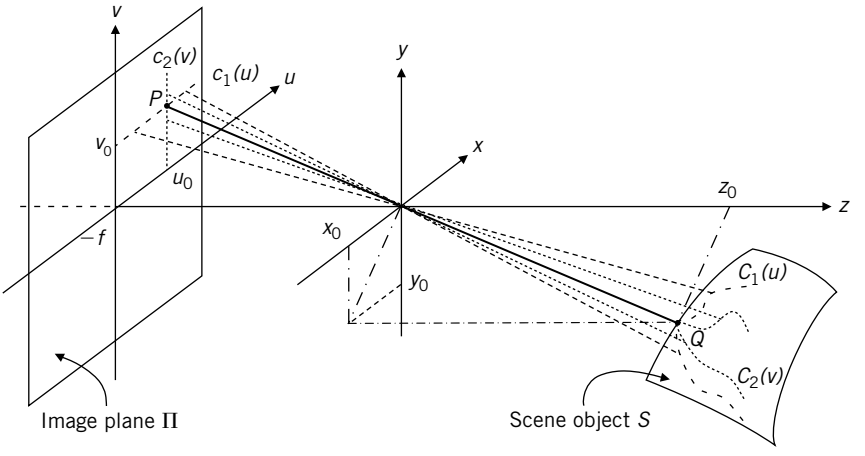


FIGURE 23 The image plane is $\Pi = \{(u, v, -f)\}$, where f is the focal length. Point $\vec{Q} = (x_0, y_0, z(x, y))$ on the surface S is projected onto point $\vec{P} = (u_0, v_0, -f)$ on image plane Π. The curves $c_1(u)$ and $c_2(v)$ are parallel to axes x, y (or u, v). The curves $C_1(u), C_2(v) \in S$ are the curves on the object, the perspective projections of which on Π are curves $c_1(u)$ and $c_2(v)$, respectively. The tangents to $C_1(u)$ and $C_2(v)$ at point $\vec{Q}$ are computed from $\vec{P}$ and the perspective projection equations. The normal to the two tangents is the normal to S at point $\vec{Q}$.

The gradient of the real-world curve $C(s)$ is

$$\frac{dC(s)}{ds} = \frac{1}{f}[-u_s(s)h(s) - u(s)h_s(s), -v_s(s)h(s) - v(s)h_s(s), f\, h_s(s)]. \quad (50)$$

where subscript s indicates differentiation with respect to s.

The authors consider two different curves parallel to the x and y axes, through the image point $\vec{P} = (u_0, v_0, -f)$. Thus,

$$c_1(u) = (u, v_0, -f)$$
$$c_2(v) = (u_0, v, -f). \quad (51)$$

We may calculate the two tangents at point $\vec{P}$ as

$$\left.\frac{dC_1(u)}{du}\right|_{u=u_0} = \left.\frac{1}{f}(-h - uh_u, -vh_u, fh_u)\right|_{u=u_0}$$

$$\left.\frac{dC_2(v)}{dv}\right|_{v=v_0} = \left.\frac{1}{f}(-uh_v, -h - vh_v, fh_v)\right|_{v=v_0}. \quad (52)$$

The normal to the surface is parallel to the cross-product

$$\left.\frac{dC_1(u)}{du}\right|_{u=u_0} \times \left.\frac{dC_2(v)}{dv}\right|_{v=v_0} = \left.\frac{h}{f^2}\big(f\,h_u, f\,h_v, uh_u + vh_v + h\big)\right|_{u=u_0,\,v=v_0}, \tag{53}$$

and dropping the subscript 0 to make point P generic, we deduce that the unit normal is given by

$$\vec{N} = \frac{(f\,h_u, f\,h_v, uh_u + vh_v + h)}{\sqrt{(uh_u + vh_v + h)^2 + f^2(h_u^2 + h_v^2)}}. \tag{54}$$

For unit illumination direction vector $\vec{L}$, the image irradiance equation thus becomes

$$\begin{aligned} I(u,v) &= \rho_I(u,v)\vec{N}\cdot\vec{L} \\ &= \rho_I(u,v)\frac{(-p_l, -q_l, 1)\cdot\big(f\,h_u, f\,h_v, uh_u + vh_v + h\big)}{\sqrt{1+p_l^2+q_l^2}\sqrt{(uh_u + vh_v + h)^2 + f^2(h_u^2 + h_v^2)}} \\ &= \rho_I(u,v)\frac{(u - fp_l)h_u + (v - fq_l)\,h_v + h}{\sqrt{1+p_l^2+q_l^2}\sqrt{(uh_u + vh_v + h)^2 + f^2(h_u^2 + h_v^2)}}. \end{aligned} \tag{55}$$

The above equation shows direct dependence of the image irradiance on both $h(u,v)$ and its first derivatives. Note that in Eq. (55) $h(u,v)$ and its two derivatives explicitly appear in the equation. This makes the problem rather difficult to solve. If, however, $\ln(h(u,v))$ is used instead of $h(u,v)$, the problem becomes simpler, as the irradiance equation is expressed as

$$I(u,v) = \rho_I(u,v)\frac{(u - fp_l)\tilde{p}_n(u,v) + (v - fq_l)\tilde{q}_n(u,v) + 1}{\sqrt{1+p_l^2+q_l^2}\sqrt{(u\tilde{p}_n(u,v) + v\tilde{q}_n(u,v) + 1)^2 + f^2\big(\tilde{p}_n^2(u,v) + \tilde{q}_n^2(u,v)\big)}}, \tag{56}$$

where $\tilde{p}_n(u,v) \equiv \frac{h_u}{h} = \frac{\partial \ln h}{\partial u}$ and $\tilde{q}_n(u,v) \equiv \frac{h_v}{h} = \frac{\partial \ln h}{\partial v}$, and it depends on the derivatives of $\ln(h(u,v))$, but not on $\ln(h(u,v))$ itself. Therefore, the problem of recovering $h(u,v)$ from the image irradiance equation is reduced to the problem of recovering the surface $\ln(h(u,v))$, since the natural logarithm is a bijective mapping and $h(u,v) > 0$.

In the case of photometric stereo, several images $I^i(u,v)$ of the same object under different illumination directions $\vec{L}^i$ are acquired and the

perspective image irradiance equations are as follows:

$$I_I^i(u,v) = \rho(u,v)\frac{(u-fp_{l_i})\tilde{p}_n(u,v) + (v-fq_{l_i})\tilde{q}_n(u,v)+1}{|\vec{L}^i|\sqrt{(u\tilde{p}_n(u,v)+v\tilde{q}_n(u,v)+1)^2+f^2(\tilde{p}_n^2(u,v)+\tilde{q}_n^2(u,v))}}. \tag{57}$$

Dividing the ith image by the kth, assuming the latter is nonzero everywhere, yields

$$\frac{I^i(u,v)}{I^k(u,v)} = \frac{|\vec{L}^k|[(u-fp_{l_i})\tilde{p}_n(u,v) + (v-fq_{l_i})\tilde{q}_n(u,v)+1]}{|\vec{L}^i|[(u-fp_{l_k})\tilde{p}_n(u,v) + (v-fq_{l_k})\tilde{q}_n(u,v)+1]}, \tag{58}$$

which may be written as

$$A_{ik}\tilde{p}_n + B_{ik}\tilde{q}_n + C_{ik} = 0, \tag{59}$$

where

$$\begin{aligned} A_{ik} &= I^i(u,v)|\vec{L}^i|(u-fp_{l_k}) - I^k(u,v)|\vec{L}^k|(u-fp_{l_i}) \\ B_{ik} &= I^i(u,v)|\vec{L}^i|(v-fq_{l_k}) - I^k(u,v)|\vec{L}^k|(v-fq_{l_i}) \\ C_{ik} &= I^i(u,v)|\vec{L}^i| - I^k(u,v)|\vec{L}^k|. \end{aligned} \tag{60}$$

If we have three images, we will have three equations like Eq. (58), which correspond to Eq. (37) for the orthographic projection:

$$\begin{aligned} A_{01}\tilde{p}_n + B_{01}\tilde{q}_n + C_{01} &= 0 \\ A_{12}\tilde{p}_n + B_{12}\tilde{q}_n + C_{12} &= 0 \\ A_{02}\tilde{p}_n + B_{02}\tilde{q}_n + C_{02} &= 0. \end{aligned} \tag{61}$$

This system of equations may be solved like Eq. (37) (i.e., considering all three pairs of them and solving for $\tilde{q}_n$, averaging the three answers and doing the same for $\tilde{p}_n$). Alternatively, we may solve the system in the least-square error sense. In a matrix form linear system (61) may be written as

$$\begin{pmatrix} A_{01} & B_{01} \\ A_{12} & B_{12} \\ A_{02} & B_{02} \end{pmatrix} \begin{pmatrix} \tilde{p}_n \\ \tilde{q}_n \end{pmatrix} = \begin{pmatrix} -C_{01} \\ -C_{12} \\ -C_{02} \end{pmatrix} \Rightarrow \mathbf{Wx} = \mathbf{Y}. \tag{62}$$

Multiplying both sides with the transpose of $\mathbf{W}$, we obtain

$$\mathbf{W}^T\mathbf{W}\mathbf{x} = \mathbf{W}^T\mathbf{Y}, \tag{63}$$

and solving for $\tilde{p}_n$ and $\tilde{q}_n$ we have

$$\mathbf{Q}\mathbf{x} = \mathbf{R} \Rightarrow \begin{pmatrix} M_1 & N_1 \\ M_2 & N_2 \end{pmatrix} \begin{pmatrix} \tilde{p}_n \\ \tilde{q}_n \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} \tilde{p}_n = \frac{N_2R_1 - N_1R_2}{N_2M_1 - N_1M_2} \\ \tilde{q}_n = -\frac{M_2R_1 - M_1R_2}{N_2M_1 - N_1M_2}, \end{cases} \tag{64}$$

where $\mathbf{Q} \equiv \mathbf{W}^T\mathbf{W}, \mathbf{R} \equiv \mathbf{W}^T\mathbf{Y}, M_1 = A_{01}^2 + A_{12}^2 + A_{02}^2, N_2 = B_{01}^2 + B_{12}^2 + B_{02}^2$, and $M_2 = N_1 = A_{01}B_{01} + A_{12}B_{12} + A_{02}B_{02}$. Substituting from (60) in the previous equations, we obtain

$$p_n(u,v) = \frac{I^0|\vec{L}^0|\Delta q_l^{21} + I^1|\vec{L}^1|\Delta q_l^{02} + I^2|\vec{L}^2|\Delta q_l^{10}}{I^0|\vec{L}^0|D_{12} + I^1|\vec{L}^1|D_{20} + I^2|\vec{L}^2|D_{01}}$$

$$q_n(u,v) = \frac{I^0|\vec{L}^0|\Delta p_l^{21} + I^1|\vec{L}^1|\Delta p_l^{02} + I^2|\vec{L}^2|\Delta p_l^{10}}{I^0|\vec{L}^0|D_{12} + I^1|\vec{L}^1|D_{20} + I^2|\vec{L}^2|D_{01}}. \tag{65}$$

Here $\Delta p_l^{ik} = p_{l_i} - p_{l_k}$, $\Delta q_l^{ik} = q_{l_i} - q_{l_k}$, and $D_{ik} = u(q_{l_i} - q_{l_k}) + v(p_{l_k} - p_{l_i}) + f(p_{l_i}q_{l_k} - p_{l_k}q_{l_i})$ with $i,k = 0,1,2$. Albedo ρ_I is obtained by substitution of $\tilde{p}_n$ and $\tilde{q}_n$ into one of the image irradiance equations. The final step of the perspective photometric stereo method is taking the exponent of the result, since the derivatives of the natural logarithm of the depth have been calculated rather than the depth itself.

12. PHOTOMETRIC STEREO ERROR ANALYSIS

The photometric stereo method is susceptible to systematic errors derived from a number of factors, such as errors in the image acquisition stage, inaccurate measurement of light source orientations, effects of specular reflections and shadows, the spatial and spectral distribution of incident light, imaging and illumination geometry, and surface size and material. Jiang and Bunke (1991) and Ray, Birk, and Kelley (1983) analyzed some of the above error sources and performed a sensitivity analysis of surface

normals computed by the photometric stereo method with respect to the majority of the errors. Furthermore, Barsky and Petrou (2006) presented a detailed performance and error analysis for four-light color photometric stereo. Their error analysis was based on the assumptions that errors may arise due to Gaussian image noise and inaccuracies with which the geometry of the illuminating setup was known.

12.1. Errors in the Image Acquisition Stage

Errors that occur during the acquisition stage are due mostly to theoretical assumptions, limitations of the available equipment, and the general system architecture. For example, the assumption that the light rays reaching the surface of the object are parallel requires the distance between the illumination sources and the object to be significantly large. Furthermore, the illumination direction vectors should be non-coplanar, and a perfect match for all the light sources in their spatial and spectral properties is required. The reflected light of a surface depends on its albedo and the wavelength of the incident light; therefore, the photometric stereo cannot be applied with light sources of different wavelengths, since different albedo factors then are involved in each image.

The limitations of the available equipment introduce additional errors. The presence of noise and faults in the CCD arrays alter the intensity values. The focal lens may introduce distortions (e.g., pin cushion and barrel) mainly at the image borders. In order to achieve orthographic projection, which is assumed in photometric stereo, the distance between the acquired object and the viewing point should be much larger than the size of the object itself, with a good approximation being achieved for a ratio of 1/30. In contradiction, by placing the camera close to the object, higher resolution and precision may be achieved. Since the resolution of the reconstructed surface using photometric stereo depends on the camera resolution and its relative position with respect to the acquired object, the optimal distance regarding the desired resolution of the surface may be accurately defined using camera calibration, which also can eliminate the focal lens distortions.

Camera calibration consists of the estimation of a model for an uncalibrated camera. The objective is to find the external parameters (position and orientation relative to a world coordinate system), and the internal parameters of the camera (principal point or image center, focal length, and distortion coefficients). One of the most commonly used camera calibration techniques is the one proposed by Tsai (1986). During calibration, a target of known geometry is imaged and correspondences between target points and their images are obtained. These form the basic data on which the calibration is based.

12.2. Sensitivity Analysis of Surface Normals

Errors in measurement of light source orientations and image intensity result in errors in the estimated surface orientation. The sensitivity of photometric stereo in various input parameters may be analyzed using the gradient representation $(p_n(x,y), q_n(x,y))$ of the reflectance maps. The accuracy of the computed values of p_n and q_n depends on the accuracy in measuring the zenith θ_{l_i} and azimuth angles φ_{l_i}.

Light calibration methods are required to measure the angles of the illumination sources. Therefore, we consider lights placed symmetrically about the optical axis z. Based on Eqs. (7) and (8), in that case it can be observed that all the zenith angles are equal $\theta_l = \theta_{l_1} = \theta_{l_2} = \theta_{l_3}$ with $\cos\theta_l = 1/\left(\sqrt{1 + p_l^2 + q_l^2}\right)$. The logical consequence of the above symmetric arrangement is that the acquired image intensities from the three light sources are the same if the observed surface is directly facing the camera $(p_n(x,y) = q_n(x,y) = 0)$.

As shown in Figure 24, the ratio of the distance of the camera from the three light sources d_l over the distance between the object and the camera d_c corresponds to the tangent value of the zenith angle θ_l. Hence,

$$\theta_l = \tan^{-1}\frac{d_l}{d_c}, \quad 0 \le \theta_l \le 90^\circ. \tag{66}$$

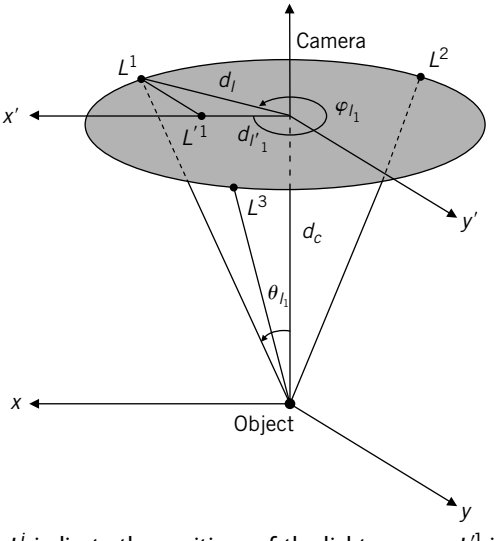


FIGURE 24 Points L^i indicate the positions of the light sources. L'^1 is the projection of L^1 on the zx-plane, θ_{l_1} is its zenith angle, and φ_{l_1} is the azimuth angle of L^1.

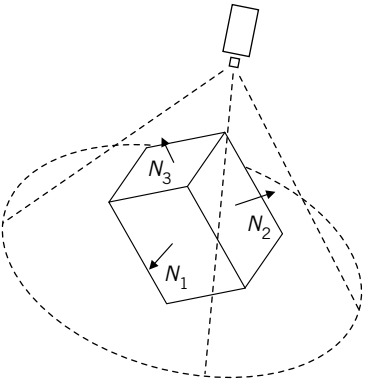


FIGURE 25 A rectangular parallelepiped placed so that three distinct surfaces are clearly visible from the viewing position simultaneously.

In a similar manner, the azimuth angle may be obtained by measuring the distance between the projection of the light sources on the xz-plane and the center of the axes $d_{l'_i}$, assuming that the camera is on the z-axis. The values of φ_{l_i} are given by

$$\cos\varphi_{l_i} = \frac{d_{l'_i}}{d_l} \quad \text{and} \quad \sin\varphi_{l_i} = \frac{L^i L'^i}{d_l} \tag{67}$$

so that $0 \leq \varphi_{l_i} < 360°$.

A second method that is often selected for illumination directions calibration requires positioning a rectangular parallelepiped such that three distinct surfaces are clearly visible from the viewing position simultaneously (Figure 25). The normals of the three visible surfaces are known *a priori* and, using the values of the acquired image intensities for a single illumination direction, we obtain

$$\begin{aligned}
I^1(x,y) &= \rho_S(x,y)\frac{1 + p_{n_1}(x,y)p_{l_i} + q_{n_1}(x,y)q_{l_i}}{\sqrt{1 + p^2_{n_1}(x,y) + q^2_{n_1}(x,y)}\sqrt{1 + p^2_{l_i} + q^2_{l_i}}} \\
I^2(x,y) &= \rho_S(x,y)\frac{1 + p_{n_2}(x,y)p_{l_i} + q_{n_2}(x,y)q_{l_i}}{\sqrt{1 + p^2_{n_2}(x,y) + q^2_{n_2}(x,y)}\sqrt{1 + p^2_{l_i} + q^2_{l_i}}} \\
I^3(x,y) &= \rho_S(x,y)\frac{1 + p_{n_3}(x,y)p_{l_i} + q_{n_3}(x,y)q_{l_i}}{\sqrt{1 + p^2_{n_3}(x,y) + q^2_{n_3}(x,y)}\sqrt{1 + p^2_{l_i} + q^2_{l_i}}},
\end{aligned} \tag{68}$$

where the surface gradients $(p_{n_1}(x,y), q_{n_1}(x,y))$, $(p_{n_2}(x,y), q_{n_2}(x,y))$, and $(p_{n_3}(x,y), q_{n_3}(x,y))$ correspond to the three visible surfaces. Assuming

diffuse reflection and invariant albedo for equivalent point locations, the orientation (p_{l_i}, q_{l_i}) of the ith light source is determined by solving the system of equations in the same manner as system (36). The above procedure is repeated for the remaining light sources to calculate the remaining zenith and azimuth angles, $\theta_l, \varphi_{l_1}, \varphi_{l_2}$, and φ_{l_3}.

The sensitivity of the photometric stereo algorithm with respect to the changes in various input parameters can be analyzed based on the gradient representation of the normal vectors and Eqs. (6) and (36). Solving these equations algebraically (Jiang and Bunke, 1991; Ray, Birk, and Kelley, 1983), the surface gradients are given by

$$p_n(x,y) = g(I^1, I^2, I^3, \theta_l, \varphi_{l_1}, \varphi_{l_2}, \varphi_{l_3}) \tag{69}$$

and

$$q_n(x,y) = f(I^1, I^2, I^3, \theta_l, \varphi_{l_1}, \varphi_{l_2}, \varphi_{l_3}). \tag{70}$$

In order to estimate the deviation of the normal vector components p_n and q_n, with respect to small changes in the input parameters, we obtain

$$p_n(x,y) = \sum_{i=1}^{3} \frac{\partial g}{\partial I^i} dI^i + \frac{\partial g}{\partial \theta_l} d\theta_l + \sum_{i=1}^{3} \frac{\partial g}{\partial \varphi_{l_i}} d\varphi_{l_i} \tag{71}$$

and

$$q_n(x,y) = \sum_{i=1}^{3} \frac{\partial f}{\partial I^i} dI^i + \frac{\partial f}{\partial \theta_l} d\theta_l + \sum_{i=1}^{3} \frac{\partial f}{\partial \varphi_{l_i}} d\varphi_{l_i}. \tag{72}$$

The expression of $p_n(x,y)$ and $q_n(x,y)$ in gradient space leads to complicated formulae. For this reason, in Jiang and Bunke (1991) the surface orientation was represented in terms of unit normal vectors, reducing the complexity of the sensitivity analysis process.

Based on Eqs. (34) and (35)

$$\vec{N} = \frac{1}{\rho}[L]^{-1}\vec{I}, \tag{73}$$

where $[L]$ is a 3×3 matrix, and its inverse is given by

$$[L]^{-1} = \frac{1}{[\vec{L}^1 \ \vec{L}^2 \ \vec{L}^3]}(\vec{L}^2 \times \vec{L}^3 \ \ \vec{L}^3 \times \vec{L}^1 \ \ \vec{L}^1 \times \vec{L}^2), \tag{74}$$

with $[\vec{L}^1\ \vec{L}^2\ \vec{L}^3]$ representing the triple product $\vec{L}^1 \cdot (\vec{L}^2 \times \vec{L}^3)$ (i.e., the volume of a parallelepiped formed by $\vec{L}^1$, $\vec{L}^2$, and $\vec{L}^3$), Eq. (73) may be written as

$$\vec{N} = \frac{1}{\rho[\vec{L}^1\ \vec{L}^2\ \vec{L}^3]}(I^1(\vec{L}^2 \times \vec{L}^3) + I^2(\vec{L}^3 \times \vec{L}^1) + I^3(\vec{L}^1 \times \vec{L}^2)) = \frac{1}{\rho[\vec{L}^1\ \vec{L}^2\ \vec{L}^3]}\vec{W}. \tag{75}$$

Considering, without loss of generality, that $[\vec{L}^1\ \vec{L}^2\ \vec{L}^3] > 0$ and observing that both $\vec{N}$ and $\vec{W}$ point in the same direction, the azimuth φ_n and the zenith θ_n angles are identical for the two vectors $\vec{N}$ and $\vec{W}$. Instead of using $\vec{N}$ to compute the deviation of the normal vector components, $\vec{W}$ may be used, resulting in much simpler sensitivity formulas.

The surface normal vector $\vec{W}$ is given by

$$\vec{W} = (x_n, y_n, z_n) \equiv I^1(\vec{L}^2 \times \vec{L}^3) + I^2(\vec{L}^3 \times \vec{L}^1) + I^3(\vec{L}^1 \times \vec{L}^2), \tag{76}$$

where

$$\begin{aligned} x_n &= f_x(I^1, I^2, I^3, \theta_{l_1}, \theta_{l_2}, \theta_{l_3}, \varphi_{l_1}, \varphi_{l_2}, \varphi_{l_3}) \\ y_n &= f_y(I^1, I^2, I^3, \theta_{l_1}, \theta_{l_2}, \theta_{l_3}, \varphi_{l_1}, \varphi_{l_2}, \varphi_{l_3}) \\ z_n &= f_z(I^1, I^2, I^3, \theta_{l_1}, \theta_{l_2}, \theta_{l_3}, \varphi_{l_1}, \varphi_{l_2}, \varphi_{l_3}). \end{aligned} \tag{77}$$

To the first order, the errors in the component of $\vec{W}$ with respect to errors in the input parameters may be written as

$$\begin{aligned} d\vec{W} = (dx_n, dy_n, dz_n) &= \sum_{i=1}^{3} \frac{\partial \vec{W}}{\partial I^i} dI^i + \sum_{i=1}^{3} \frac{\partial \vec{W}}{\partial \theta_i} d\theta_i + \sum_{i=1}^{3} \frac{\partial \vec{W}}{\partial \varphi_i} d\varphi_i = \\ &= \sum_{i=1}^{3} \frac{\partial \vec{W}}{\partial I^i} dI^i + \sum_{i=1}^{3} \frac{\partial \vec{W}}{\partial \vec{L}^i} \left(\frac{\partial \vec{L}^i}{\partial \theta_i} d\theta_i + \frac{\partial \vec{L}^i}{\partial \varphi_i} d\varphi_i \right), \end{aligned} \tag{78}$$

where the ratio $\partial \vec{W} / \partial \vec{L}^i$ is a matrix consisting of partial derivatives of the components of vector $\vec{W}$ with respect to the components of vector $\vec{L}^i = (x_{l_i}, y_{l_i}, z_{l_i})$:

$$\frac{\partial \vec{W}}{\partial \vec{L}^i} = \begin{bmatrix} \frac{\partial x_n}{\partial x_{l_i}} & \frac{\partial x_n}{\partial y_{l_i}} & \frac{\partial x_n}{\partial z_{l_i}} \\ \frac{\partial y_n}{\partial x_{l_i}} & \frac{\partial y_n}{\partial y_{l_i}} & \frac{\partial y_n}{\partial z_{l_i}} \\ \frac{\partial z_n}{\partial x_{l_i}} & \frac{\partial z_n}{\partial y_{l_i}} & \frac{\partial z_n}{\partial z_{l_i}} \end{bmatrix} \quad \text{for } i = 1, 2, 3. \tag{79}$$

In order to estimate the errors in $\vec{W}$, all the partial derivatives in Eq. (78) must be computed. For the first part of Eq. (78), we obtain

$$\begin{aligned}\frac{\partial \vec{W}}{\partial I^1} &= \vec{L}^2 \times \vec{L}^3 \\ \frac{\partial \vec{W}}{\partial I^2} &= \vec{L}^3 \times \vec{L}^1 \\ \frac{\partial \vec{W}}{\partial I^3} &= \vec{L}^1 \times \vec{L}^2. \end{aligned} \tag{80}$$

Using the matrix notation of cross-product

$$\vec{L}^i \times \vec{L}^j = \begin{bmatrix} 0 & -z_{l_i} & y_{l_i} \\ z_{l_i} & 0 & -x_{l_i} \\ -y_{l_i} & x_{l_i} & 0 \end{bmatrix} \begin{bmatrix} x_{l_j} \\ y_{l_j} \\ z_{l_j} \end{bmatrix}, \tag{81}$$

it can be proven (Jiang and Bunke, 1991) that

$$\frac{\partial \vec{L}^i \times \vec{L}^j}{\partial \vec{L}^j} = \begin{bmatrix} 0 & -z_{l_i} & y_{l_i} \\ z_{l_i} & 0 & -x_{l_i} \\ -y_{l_i} & x_{l_i} & 0 \end{bmatrix} \equiv H_i, \text{ for } i \neq j, \ i = 1, 2, 3. \tag{82}$$

Based on the above equation, we obtain

$$\begin{aligned}\frac{\partial \vec{W}}{\partial \vec{L}^1} &= I^2 H_3 - I^3 H_2 \\ \frac{\partial \vec{W}}{\partial \vec{L}^2} &= I^3 H_1 - I^1 H_3 \\ \frac{\partial \vec{W}}{\partial \vec{L}^3} &= I^1 H_2 - I^2 H_1. \end{aligned} \tag{83}$$

The partial derivatives of the illumination vectors using the azimuth and zenith angles for representation (see Eq. (9)) are given by

$$\begin{aligned}\frac{\partial \vec{L}^i}{\partial \theta_{l_i}} &= (\cos\theta_{l_i} \cos\varphi_{l_i}, \cos\theta_{l_i} \sin\varphi_{l_i}, -\sin\theta_{l_i}) \\ \frac{\partial \vec{L}^i}{\partial \varphi_{l_i}} &= (-\sin\theta_{l_i} \sin\varphi_{l_i}, \sin\theta_{l_i} \cos\varphi_{l_i}, 0). \end{aligned} \tag{84}$$

Now, substituting all the partial derivatives in Eq. (78), an estimation of errors in surface normals $\vec{W}$ can be found. Furthermore, representing the surface normal vectors $\vec{W} = (x_n, y_n, z_n)$ using the zenith θ_n and azimuth φ_n angles from Eq. (8), we obtain

$$\theta_n = \tan^{-1} \sqrt{p_n^2(x,y) + q_n^2(x,y)} = \tan^{-1} \frac{\sqrt{x_n^2 + y_n^2}}{z_n} \tag{85}$$

$$\varphi_n = \tan^{-1} \frac{p_n(x,y)}{q_n(x,y)} = \tan^{-1} \frac{y_n}{x_n}. \tag{86}$$

and the errors in surface normal vectors are estimated by calculating their derivatives $d\theta_n$ and $d\varphi_n$ with respect to x_n, y_n, and z_n, as

$$\begin{aligned} d\theta_n &= \frac{1}{1 + (x_n^2 + y_n^2)/z_n^2} \left(\frac{x_n}{z_n \sqrt{x_n^2 + y_n^2}} dx_n + \frac{y_n}{z_n \sqrt{x_n^2 + y_n^2}} dy_n - \frac{\sqrt{x_n^2 + y_n^2}}{z_n^2} dz_n \right) \\ &= \frac{1}{x_n^2 + y_n^2 + z_n^2} \left(\frac{x_n z_n}{\sqrt{x_n^2 + y_n^2}} dx_n + \frac{y_n z_n}{\sqrt{x_n^2 + y_n^2}} dy_n - \sqrt{x_n^2 + y_n^2} dz_n \right) \\ &= \frac{1}{x_n^2 + y_n^2 + z_n^2} \left(\frac{x_n z_n}{\sqrt{x_n^2 + y_n^2}}, \frac{y_n z_n}{\sqrt{x_n^2 + y_n^2}}, -\sqrt{x_n^2 + y_n^2} \right) \cdot d\vec{W} \end{aligned} \tag{87}$$

$$\begin{aligned} d\varphi_n &= \frac{1}{1 + (y_n/x_n)^2} \left(-\frac{y_n}{x_n^2} dx_n + \frac{1}{x_n^2} dy_n \right) \\ &= \frac{1}{x_n^2 + y_n^2} \left(-y_n dx_n + x_n dy_n \right) \\ &= \frac{1}{x_n^2 + y_n^2} \left(-y_n, x_n, 0 \right) \cdot d\vec{W}. \end{aligned} \tag{88}$$

Substituting Eq. (78) in the above equations yields the magnitude of the errors in azimuth and zenith angles of the surface normals as functions of the input parameters of the photometric stereo.

The most significant outcome of the above sensitivity analysis is that, in general, a 1-degree error in the calculation of light source orientation or 1% error in image intensity, results in a 1-degree error in the orientation of the surface normal vectors. However, errors due to highlights or shadows

are not covered by this analysis. A discussion on such errors may be found in Barsky and Petrou (2006).

13. OPTIMAL ILLUMINATION CONFIGURATION FOR PHOTOMETRIC STEREO

The influence of a lighting arrangement on the accuracy of surface reconstruction based on photometric stereo has been considered by Lee and Kuo (1992), Spence and Chantler (2003) and Woodham (1980) and suggestions for optimal illumination configuration in terms of azimuth φ_l and zenith θ_l angles were made. Woodham (1980), using reflectance maps, recommended dense iso-intensity contours to obtain maximum accuracy, since in this case a small change in the surface gradients p_n and q_n results in a large intensity change. In order to achieve dense iso-intensity contours the zenith angle θ_l has to be increased, but this results in increasing the presence of shadows, which is undesirable. Regarding the azimuth angles of the light sources, Woodham pointed out that the illumination vectors must not be coplanar, otherwise the illumination matrix in Eq. (34) is not invertible.

Lee and Kuo (1992), using two reflectance maps for their two-image photometric stereo algorithm, deduced that it is desirable to incorporate reflectance maps that compensate for each other's weaknesses, in order to determine the optimal illumination configuration. Observing that the azimuth angle φ_l determines the orientation of the reflectance map around the origin, whereas the zenith angle θ_l determines the distance between the origin and (p_l, q_l) in the gradient space (Figures 26 and 1), the angular difference between two reflectance maps would be given by $|\varphi_{l_1} - \varphi_{l_2}|$. The above observations can be easily proven using Eq. (8) and calculating the angle between the line passing through the $(0, 0)$ and (p_l, q_l) points in the gradient space and the p-axis. Obviously,

$$\phi = \arctan\left(\frac{q_l}{p_l}\right) = \arctan(\tan\varphi_l) = \varphi_l, \tag{89}$$

and the distance d between points $(0, 0)$ and (p_l, q_l) is

$$d = \sqrt{q_l^2 + p_l^2} = \tan\theta_l. \tag{90}$$

Angle ϕ is the same as the azimuth angle φ_l, and the distance from the origin of the gradient space depends on the zenith angle θ_l.

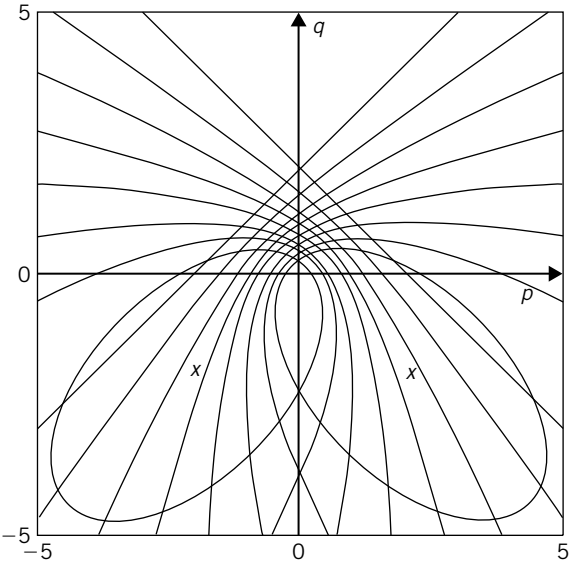


FIGURE 26 Contour plots of Lambertian reflectance maps with $(\varphi_{l_1}, \theta_{l_1}) = (45°, 45°)$ and $(\varphi_{l_2}, \theta_{l_2}) = (135°, 45°)$. The *x*'s mark points (p_{l_1}, q_{l_1}) and (p_{l_2}, q_{l_2}).

Assuming that the zenith angle is in the range of 30° to 60°, the reflectance map would cover the central region of the gradient space. Therefore, the azimuth angles of the illumination sources define the optimal lighting condition. Taking into consideration that the reflectance map provides good sensitivity along the gradient direction but poor sensitivity along the tangential direction, the optimal illumination directions of two sources are obtained when the gradient directions of one reflectance map correspond to the tangential directions of the other reflectance map, $|\varphi_{l_1} - \varphi_{l_2}| = 90°$. To understand this better, the value of $q_n(x, y)$ of the reflectance map is regarded fixed and equal to $q_{n_0}(x, y)$. Now, viewing the reflectance map as a function of $p_n(x, y)$, the sensitivity of $p_n(x, y)$ with respect to changes in $I(x, y)$ is given by

$$\left| \frac{\partial p_n}{\partial I} \right| = \left| \left[\frac{\partial R(p_n(x, y), q_{n_0}(x, y))}{\partial p_n} \right]^{-1} \right| \tag{91}$$

and is inversely proportional to the slope of the reflectance map at $p_n(x, y)$. Thus, for a fixed value of ∂I, the value of ∂p_n is smaller where $R(p_n(x, y), q_{n_0}(x, y))$ is steepest and the function changes most rapidly.

Similarly, if $p_n(x,y) = p_{n_0}(x,y)$, we obtain

$$\left|\frac{\partial q_n}{\partial I}\right| = \left|\left[\frac{\partial R(p_{n_0}(x,y), q_n(x,y))}{\partial q_n}\right]^{-1}\right|. \tag{92}$$

For a given point $(p_{n_0}(x,y), q_{n_0}(x,y))$ and according to the above, the sensitivity is greatest along the direction perpendicular to the contour; that is, $\nabla R(p_{n_0}(x,y), q_{n_0}(x,y)) = [R_p(p_{n_0}(x,y), q_{n_0}(x,y)), R_q(p_{n_0}(x,y), q_{n_0}(x,y))]$ and lowest along the tangential direction; that is $\nabla R(p_{n_0}(x,y), q_{n_0}(x,y)) = [-R_p(p_{n_0}(x,y), q_{n_0}(x,y)), R_q(p_{n_0}(x,y), q_{n_0}(x,y))]$.

Based on the work by Gullon (2002), Lee and Kuo (1992) further confirmed that the two-light photometric stereo is more sensitive to the azimuth rather than the zenith angle difference and that the optimal value is 90°. Gullon moves to the three-image photometric stereo and suggests that distributing the illumination azimuth angles equally through 360° is optimal. A theoretical analysis of Gullon's arrangement was presented by Spence and Chantler (2003, 2006) based on the sensitivity analysis of photometric stereo deriving expressions of each surface normal vector with respect to image intensities as shown in the previous section.

The illumination matrix $[L]$, represented using the zenith and azimuth angles (see Eq. (9)), is defined as

$$[L] = \begin{bmatrix} \sin\theta_{l_1}\cos\varphi_{l_1} & \sin\theta_{l_1}\sin\varphi_{l_1} & \cos\theta_{l_1} \\ \sin\theta_{l_2}\cos\varphi_{l_2} & \sin\theta_{l_2}\sin\varphi_{l_2} & \cos\theta_{l_2} \\ \sin\theta_{l_3}\cos\varphi_{l_3} & \sin\theta_{l_3}\sin\varphi_{l_3} & \cos\theta_{l_3} \end{bmatrix} \tag{93}$$

and depends on six parameters. The scaled surface normals $\rho\vec{N} = [x_n\ y_n\ z_n]^T$ are given by Eq. (35). Calculating the inverse of matrix $[L]$ and substituting into Eq. (35), the surface normals are obtained as

$$\begin{aligned} x_n = -\frac{1}{h}((&\sin\varphi_{l_3}\cos\theta_{l_2}\sin\theta_{l_3} - \sin\varphi_{l_2}\sin\theta_{l_2}\cos\theta_{l_3})I^1 \\ &+ (\sin\varphi_{l_1}\cos\theta_{l_3}\sin\theta_{l_1} - \sin\varphi_{l_3}\sin\theta_{l_3}\cos\theta_{l_1})I^2 \\ &+ (\sin\varphi_{l_2}\cos\theta_{l_1}\sin\theta_{l_2} - \sin\varphi_{l_1}\sin\theta_{l_1}\cos\theta_{l_2})I^3) \end{aligned} \tag{94}$$

$$\begin{aligned} y_n = -\frac{1}{h}((&\cos\varphi_{l_2}\cos\theta_{l_3}\sin\theta_{l_2} - \cos\varphi_{l_3}\sin\theta_{l_3}\cos\theta_{l_2})I^1 \\ &+ (\cos\varphi_{l_3}\cos\theta_{l_1}\sin\theta_{l_3} - \cos\varphi_{l_1}\sin\theta_{l_1}\cos\theta_{l_3})I^2 \\ &+ (\cos\varphi_{l_1}\cos\theta_{l_2}\sin\theta_{l_1} - \cos\varphi_{l_2}\sin\theta_{l_2}\cos\theta_{l_1})I^3) \end{aligned} \tag{95}$$

$$z_n = \frac{1}{h}((\sin(\varphi_{l_3} - \varphi_{l_2}) \sin\theta_{l_2} \sin\theta_{l_3})I^1 + (\sin(\varphi_{l_1} - \varphi_{l_3}) \sin\theta_{l_3} \sin\theta_{l_1})I^2 + (\sin(\varphi_{l_2} - \varphi_{l_1}) \sin\theta_{l_1} \sin\theta_{l_2})I^3), \tag{96}$$

where $h \equiv \sin(\varphi_{l_3} - \varphi_{l_2}) \cos\theta_{l_1} \sin\theta_{l_2} \sin\theta_{l_3} + \sin(\varphi_{l_1} - \varphi_{l_3}) \cos\theta_{l_2} \sin\theta_{l_3} \sin\theta_{l_1} + \sin(\varphi_{l_2} - \varphi_{l_1}) \cos\theta_{l_3} \sin\theta_{l_1} \sin\theta_{l_2}$.

Assuming independent Gaussian noise of variance σ_i^2 in each image, the variance of the noise in the surface normal components is given by

$$\sigma_{x_n}^2 = \bar{\sigma}_i^2 \sqrt{\left(\frac{\partial x_n}{\partial I^1}\right)^2 + \left(\frac{\partial x_n}{\partial I^2}\right)^2 + \left(\frac{\partial x_n}{\partial I^3}\right)^2} \Rightarrow$$

$$\frac{\sigma_{x_n}^2}{\bar{\sigma}_i^2} = \sqrt{\left(\frac{\partial x_n}{\partial I^1}\right)^2 + \left(\frac{\partial x_n}{\partial I^2}\right)^2 + \left(\frac{\partial x_n}{\partial I^3}\right)^2} \tag{97}$$

$$\sigma_{y_n}^2 = \bar{\sigma}_i^2 \sqrt{\left(\frac{\partial y_n}{\partial I^1}\right)^2 + \left(\frac{\partial y_n}{\partial I^2}\right)^2 + \left(\frac{\partial y_n}{\partial I^3}\right)^2} \Rightarrow$$

$$\frac{\sigma_{y_n}^2}{\bar{\sigma}_i^2} = \sqrt{\left(\frac{\partial y_n}{\partial I^1}\right)^2 + \left(\frac{\partial y_n}{\partial I^2}\right)^2 + \left(\frac{\partial y_n}{\partial I^3}\right)^2} \tag{98}$$

$$\sigma_{z_n}^2 = \bar{\sigma}_i^2 \sqrt{\left(\frac{\partial z_n}{\partial I^1}\right)^2 + \left(\frac{\partial z_n}{\partial I^2}\right)^2 + \left(\frac{\partial z_n}{\partial I^3}\right)^2} \Rightarrow$$

$$\frac{\sigma_{z_n}^2}{\bar{\sigma}_i^2} = \sqrt{\left(\frac{\partial z_n}{\partial I^1}\right)^2 + \left(\frac{\partial z_n}{\partial I^2}\right)^2 + \left(\frac{\partial z_n}{\partial I^3}\right)^2}. \tag{99}$$

Calculating the derivatives of Eqs. (94)–(96) with respect to the three image intensities, expressions describing the sensitivity of surface normals to errors in image intensity may be obtained:

$$\frac{\partial x_n}{\partial I^1} = -\frac{\sin\varphi_{l_3} \cos\theta_{l_2} \sin\theta_{l_3} - \sin\varphi_{l_2} \sin\theta_{l_2} \cos\theta_{l_3}}{h} \tag{100}$$

$$\frac{\partial x_n}{\partial I^2} = -\frac{\sin\varphi_{l_1}\cos\theta_{l_3}\sin\theta_{l_1} - \sin\varphi_{l_3}\sin\theta_{l_3}\cos\theta_{l_1}}{h} \tag{101}$$

$$\frac{\partial x_n}{\partial I^3} = -\frac{\sin\varphi_{l_2}\cos\theta_{l_1}\sin\theta_{l_2} - \sin\varphi_{l_1}\sin\theta_{l_1}\cos\theta_{l_2}}{h} \tag{102}$$

$$\frac{\partial y_n}{\partial I^1} = -\frac{\cos\varphi_{l_2}\cos\theta_{l_3}\sin\theta_{l_2} - \cos\varphi_{l_3}\sin\theta_{l_3}\cos\theta_{l_2}}{h} \tag{103}$$

$$\frac{\partial y_n}{\partial I^2} = -\frac{\cos\varphi_{l_3}\cos\theta_{l_1}\sin\theta_{l_3} - \cos\varphi_{l_1}\sin\theta_{l_1}\cos\theta_{l_3}}{h} \tag{104}$$

$$\frac{\partial y_n}{\partial I^3} = -\frac{\cos\varphi_{l_1}\cos\theta_{l_2}\sin\theta_{l_1} - \cos\varphi_{l_2}\sin\theta_{l_2}\cos\theta_{l_1}}{h} \tag{105}$$

$$\frac{\partial z_n}{\partial I^1} = \frac{\sin(\varphi_{l_3} - \varphi_{l_2})\sin\theta_{l_2}\sin\theta_{l_3}}{h} \tag{106}$$

$$\frac{\partial z_n}{\partial I^2} = \frac{\sin(\varphi_{l_1} - \varphi_{l_3})\sin\theta_{l_3}\sin\theta_{l_1}}{h} \tag{107}$$

$$\frac{\partial z_n}{\partial I^3} = \frac{\sin(\varphi_{l_2} - \varphi_{l_1})\sin\theta_{l_1}\sin\theta_{l_2}}{h}. \tag{108}$$

Substituting the derivatives in Eqs. (97)–(99) the noise ratio for all the surface normal components is obtained, which depends on the illumination configuration. The optimal illumination arrangement is estimated by minimizing each of the three noise ratios in Eqs. (97)–(99). In order to provide a single objective function, Spence and Chantler (2003, 2006), assumed a Lambertian surface and generated the rendered intensities I^r under given illumination directions (θ_l, φ_l), as

$$I^r(x,y) = x_n\cos\varphi_l\sin\theta_l + y_n\sin\varphi_l\sin\theta_l + z_n\cos\theta_l. \tag{109}$$

This indicates that the rendered intensities are weighted sums of the surface normal components x_n, y_n, and z_n, since the trigonometric terms are scalars. Therefore, the sum of the variances of the normal components is selected

as the objective function to be minimized, given by

$$F \equiv \sqrt{\left(\frac{\partial x_n}{\partial I^1}\right)^2 + \left(\frac{\partial x_n}{\partial I^2}\right)^2 + \left(\frac{\partial x_n}{\partial I^3}\right)^2} + \sqrt{\left(\frac{\partial y_n}{\partial I^1}\right)^2 + \left(\frac{\partial y_n}{\partial I^2}\right)^2 + \left(\frac{\partial y_n}{\partial I^3}\right)^2} + \sqrt{\left(\frac{\partial z_n}{\partial I^1}\right)^2 + \left(\frac{\partial z_n}{\partial I^2}\right)^2 + \left(\frac{\partial z_n}{\partial I^3}\right)^2}. \quad (110)$$

Substituting for the partial derivatives, the above equation becomes a function of the azimuth and zenith angles of the light sources and the minimum value of F determines the optimal illumination configuration. It should be noted that the effect of inaccuracies in the measurement of the light source angles was not considered in this theoretical analysis.

Spence and Chantler (2003) found that the optimal azimuth and zenith angles cannot be specified and that the configuration that results in the minimum noise ratio is not unique. Instead, it was determined that an orthogonal arrangement of the illumination vectors (with an angle of 90° to each other) is the only restriction to obtaining the optimal configuration. If the common zenith angle was constrained, the optimal values for azimuth angles were estimated and the use of a 120-degree difference in a three-light Lambertian photometric stereo was suggested. This result is in agreement with the work of Gullon (2002), where a uniform distribution of illumination azimuth angles was recommended.

The optimal zenith angle in case of equally distributed light sources was found to be around 55°, but the angle should be reduced if shadows are present. In contradiction, if the surface is smooth and shadows are not an issue, the zenith angle may be increased. Furthermore, Drbohlav and Chantler (2005) extended the above for n light sources and proved the same optimal zenith angle when the sources are equidistantly spaced in azimuth angle, $360/n$ degrees apart.

14. CONCLUSIONS

This chapter analyzed the basic concepts of radiometry and image formation and discussed the most well-known surface models. A short review of photometric stereo and a literature review of algorithms improving the standard technique were presented. Methods identifying problematic pixels, containing shadows and highlights, or taking into consideration object

motion and perspective projection were analyzed. A sensitivity analysis of photometric stereo with respect to the image acquisition process and the accuracy of the illumination angles was presented. Finally, a theoretical analysis of the optimal illumination configuration was discussed and recommendations were reported.

The field of photometric stereo has not yet reached its maturity, and its full potential in relation to applications has not been realized. Many advances are expected in the upcoming years with a resultant increased use in applications as diverse as medicine, robotics, and security.

ACKNOWLEDGMENTS

This work was supported by EPSRC grant EP/E028659/1 "Face Recognition using photometric Stereo (PhotoFace)."

REFERENCES

Barsky, S., and Petrou, M. (2001a). Color photometric stereo: simultaneous reconstruction of local gradient and color of rough textured surfaces. *Proc. Int. Conf. Comput. Vis. ICCV* **2,** 600–605.

Barsky, S., and Petrou, M. (2001b). Shadows and highlights detection in 4-source color photometric stereo. *Int. Conf. Image Process.* **3,** 967–970.

Barsky, S., and Petrou, M. (2003). The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. Patt. Anal. Machine Intell.* **25**(10), 1239–1252.

Barsky, S., and Petrou, M. (2006). Design issues for a color photometric stereo system. *J. Math. Imaging Vis.* **24**(1), 143–162.

Blinn, J. (1977). Models of light reflection for computer synthesized picture. *Comput. Graphics* **11**(2), 192–198.

Bors, A., Hancock, E., and Wilson, R. (2000a). 3-D terrain from synthetic aperture radar images. *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications,* pp. 63–72.

Bors, A., Hancock, E., and Wilson, R. (2000b). A Bayesian framework for radar shape-from-shading. *Proc. IEEE Int. Conf. Comput. Vis. Patt. Recogn. (CVPR '00)* **1,** 262–268.

Bors, A., Hancock, E., and Wilson, R. (2000c). Terrain modeling in synthetic aperture radar images using shape-from-shading. *Proc. Int. Conf. Patt. Recogn. (ICPR '00)* **1,** 798–801.

Bors, A., Hancock, E., and Wilson, R. (2003). Terrain analysis using radar shape-from shading, *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(8), 974–992.

Coleman, E., and Jain, R. (1982). Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Comput. Vis. Graph. Image Process.* **18,** 309–328.

Drbohlav, O., and Chantler, M. (2005). On optimal light configurations in photometric stereo. *IEEE Int. Conf. Comput. Vis. ICCV* **2,** 1707–1712.

Farooq, A., Smith, M., Smith, L., and Midha, P. (2005). Dynamic photometric stereo for on line quality control of ceramic tiles. *Computers in Industry* (Special edition, Machine Vision) **56**(8), 918–934.

Gullon, C. (2002). "Height Recovery of Rough Surfaces from Intensity Images." Doctoral thesis, School of Engineering and Physical Sciences, Heriot-Watt University.

Hapke, B. (1963). A theoretical photometric function for the lunar surface. *J. Geophys. Res.* **68**(15), 4571–4586.

He, X., Torrance, K., Sillon, F., and Greenberg, D. (1991). A comprehensive physical model for light reflection. *Comput. Graph.* **25**(4), 175–186.

Horn, B. (1975). Obtaining shape from shading information. *In* "The Psychology of Computer Vision" (P. H. Winston, ed.). McGraw-Hill, New York.

Horn, B. (1986). "Robot Vision." The MIT Press, Cambridge, MA.

Horn, B., and Brooks, M. (1989). "Shape from Shading." Cambridge, MA, The MIT Press.

Horn, B., and Sjoberg, R. (1979). Calculating the reflectance map. *Appl. Optics* **18**(11), 1770–1779.

Jiang, X., and Bunke, H. (1991). On error analysis for surface normals determined by photometric stereo. *Signal Proc.* **23,** 221–226.

Klinker, G. (1993). "A Physical Approach to Color Image Understanding." A K Peters, Ltd, Wellesley, MA.

Klinker, G., Shafer, S., and Kanade, T. (1988). The measurement of highlights in color images. *Int. J. Comput. Vis.* **2,** 7–32.

Lee, J., Smith, M., Smith, L., and Midha, P. (2005). A mathematical morphology approach to image based 3D particle shape analysis. *Machine Vision Appl.* **16**(5), 282–288.

Lee, K., and Kuo, C. (1992). Shape reconstruction from photometric stereo. *Comput. Vision Patt. Recogn.*, pp. 479–484.

Liao, Y., Petrou, M., and Zhao, R. (2008). A fractal-based relaxation algorithm for shape from terrain image. *Comput. Vis. Image Und.* **109**(3), 227–243.

Lim, J., Yang, J., and Kriegman, D. (2005). Passive photometric stereo from motion. *IEEE Int. Conf. Comput. Vis. ICCV* **2,** 1635–1642.

Maki, A., Watanabe, M., and Wiles, C. (2002). Geotensity: combining motion and lighting for 3D surface reconstruction. *Int. J. Comput. Vis.* **48**(2), 75–90.

Nayar, S., Ikeuchi, K., and Kanade, T. (1990). Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. Robot. Automat.* **6**(4), 418–431.

Nicodemus, F., Richmond, J., Hsia, J., Ginsberg, I., and Limperis, T. (1977). "Geometrical Considerations and Nomenclature for Reflectance" (Tech. Report NBS Monograph 160). National Bureau of Standards, Gaithersburg, MD.

Orlova, N. (1956). Photometric relief of the lunar surface. *Astron. Z* **33**(1), 93–100.

Phong, B. T. (1975). Illumination for computer generated images. *Comm. ACM* **18**(6), 311–317.

Prados, E., and Faugeras, O. (2003). Perspective shape from shading and viscosity solutions. *IEEE int. Conf. Comput. Vis. ICCV* **2,** 826–831.

Ray, R., Birk, J., and Kelley, R. (1983). Error analysis of surface normals determined by radiometry. *IEEE Trans. Patt. Anal. Machine Intell.* **5**(6), 631–641.

Rindfleisch, T. (1966). Photometric method for lunar topography. *Photogramm. Eng.* **32**(2), 262–277.

Schluns, K., and Wittig, O. (1993). Photometric stereo for non-Lambertian surfaces using color information. *Proc. 5th Int. Conf. Comput. Anal. Images Patt.,* pp. 444–451.

Shafer, S. (1985). Using color to separate reflection components. *Color Res. Appl.* **10,** 210–218.

Simakov, D., Frolova, D., and Basri, R. (2003). Dense shape reconstruction of a moving object under arbitrary, unknown lighting. *Proc. Ninth ICCV Conf. Comput. Vis.* **2,** 1202–1209.

Smith, M. (1999). The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping. *Image Vision Comput.* **17**(14), 1009–1019.

Smith, M., and Smith, L. (2005). Dynamic photometric stereo—a new technique for moving surface analysis. *Image Vision Comput.* **23,** 841–852.

Solomon, F., and Ikeuchi, K. (1996). Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Trans. Patt. Anal. Machine Intell.* **18**(4), 449–454.

Spence, A., and Chantler, M. (2003). Optimal illumination for three-image photometric stereo acquisition of texture. *Proceedings of the 3rd International Workshop on Texture Analysis and Synthesis,* pp. 89–94.

Spence, A., and Chantler, M. (2006). Optimal illumination for three-image photometric stereo using sensitivity analysis. *IEEE Proc. Vis. Image Sign. Proc.* **153**(2), 149–159.

Tankus, A., and Kiryati, N. (2005). Photometric stereo under perspective projection. *IEEE Int. Conf. Comput. Vis. ICCV* **1,** 611–616.

Tankus, A. Sochen, N., and Yeshurun, Y. (2005). Shape-from-shading under perspective projection. *Int. J. Comput. Vis.* 63(1), 21–43.

Tankus, A., Sochen N, and Yeshurun Y. (2004). Perspective shape-from-shading by fast marching. *IEEE Comput. Soc. Conf. Comput. Vis. Patt. Recogn. CVPR* **1,** 43–49.

Tankus, A., Sochen, N., and Yeshurun, Y. (2003). A new perspective [on] shape-from-shading. *Proc. Ninth IEEE Int. Conf. Comput. Vis.* **2,** 862–869.

Tomasi, C., and Kanade, T. (1992). Shape and motion from image streams under orthography: a factorization method. *Int. J. Comput. Vis.* **9**(2), 137–154.

Tsai, R. (1986). An efficient and accurate camera calibration technique for 3D machine vision. *Proc. IEEE Conf. Comput. Vis. Patt. Recogn.,* pp. 364–374.

Woodham, R. (1980). Photometric method for determining surface orientation from multiple images. *Opt. Eng.* **19**(1), 139–144.

Worthington, P., and Hancock, E. (1997). Shape from shading using robust statistics. *13th International Conference on Digital Signal Processing Proceedings* **2,** 1145–1148.

Worthington, P., and Hancock, E. (1999). Data-driven shape-from-shading using curvature consistency. *IEEE Comput. Soc. Conf. Comput. Vis. Patt. Recogn.* **1,** 288–293.

Zhang, L., Curless, B., Hertzmann, A., and Seitz, S. (2003). Shape and motion under varying illumination: unifying structure from motion, photometric stereo, and multi-view stereo. *Proc. ICCV*., pp. 618–625.

Zhang, R., Tsai, P., Cryer, J., and Shah, M. (1999). Shape-from-shading: a survey. *IEEE Trans. Patt. Anal. Machine Intell.* **21**(8), 690–706.