CHAPTER 4

Photometric Stereo

The appearance of a surface in an image results from the effects of illumination, shape and reflectance. Reflectance models have been developed to characterise image radiance with respect to the illumination environment, viewing angles and material properties described in *Chapter 3*. These models provide a local description of reflection mechanisms that can serve as a foundation for appearance representations. *Photometric stereo* approaches utilise reflection models for estimating surface properties from transformations of image intensities that arise from illumination changes [Woodham80]. Furthermore, photometric stereo methods are simple and elegant for Lambertian diffuse models.

4.1. Candidate Surface Recovering Methods

4.1.1. Motivation

The effect of variation in illumination direction on the appearance of textures has already been discussed in previous chapters. As most texture classification schemes depend on the texture's appearance instead of topology, they are more likely to suffer from tilt induced classification error [Chantlter94a]. In the case of rough surface classification, it is therefore better to use surface properties rather than image properties as the basis for our rotation invariant texture classification. In order to do so, an intrinsic characteristic of a surface has to be recovered prior to the classification process.

Given that we are assuming a Lambertian reflectance model, the image intensity of a surface facet at a point *(x, y)* can be determined from the orientation [*p(x,y), q(x,y)*]. On the other hand, a unique surface orientation can not be determined from a single image intensity or radiance value, because there is an infinite number of surface orientations that can give rise to the same value of image intensity. Furthermore, the image intensity has only one degree of freedom and the surface orientation *(p, q)* has two. Therefore, to determine local surface orientation we need additional information. One technique that uses additional information from multiple images is called *photometric stereo*.

As we stated before, classification of surfaces should ideally be carried out on the basis of the texture's surface *s(x, y)* rather than its image *i(x,y)*. Several candidate approaches exist for the recovery of surface topography, including binocular stereo, shape from shading, and photometric stereo.

4.1.2. Binocular Stereo

Binocular stereo is a means of recovering depth by identifying corresponding points in two images taken from different viewpoints. Although binocular stereo has been used successfully in cartography to generate topographic maps of the surface of the Earth, several drawbacks make it unsuitable for this thesis.

1) Additional hardware is necessary, as this method required two cameras.
2) The difficulty in applying binocular stereo arises from reliably determining the corresponding features between two separate images. It is essential that the view position remains fixed during the image acquisition phase, in order to prevent the so called correspondence problem. Implementing the matching algorithm also results in additional computation.
3) The depth of surface is recovered rather than the surface orientation, as illustrated in *Figure 4. 1*. This will introduce noise and artefacts.

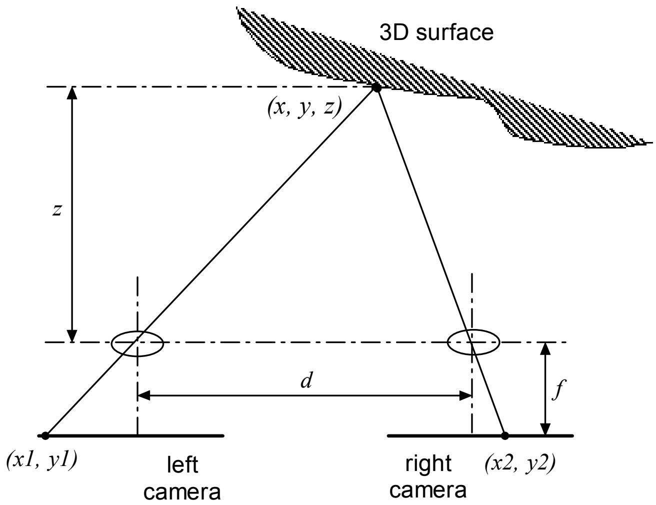

Figure 4. 1 The depth of surface is recovered by binocular stereo.

Given that camera focal length *f* and camera spacing *d*. Then the surface depth *z* can be obtained as follow:

$$z = \frac{d \cdot f}{x1 - x2} \tag{4. 1}$$

4) It is conversely best performed using objects containing discontinuities and shape features, easily corresponded between views. On the other hand, photometric stereo is suited to the objects with uniform surfaces and smooth varying topography [Horn86].

4.1.3. Shape from shading from a single image

The topic of shape from shading (SFS) is concerned with determining the shape of an object solely from the intensity variation in the image plane. Unfortunately, measurements of brightness at a single point in the image only provide one constraint whereas describing surface orientation requires two variables. The problem is ill-posed unless further assumptions are made. It was one of the first areas of study in computer vision and the initial work was carried out by Horn [Horn75] [Horn86] [Horn89].

Assuming a surface with no discontinuities [Horn75], we need knowledge concerning the reflectance properties of the surface that we are trying to describe. In addition the method must have knowledge of the reflectance properties of the surface. In other words, the shading of a surface depends on both how it is illuminated and its reflectance properties. Horn's method relies on a 2-D representation called gradient space [Horn89].

Here is the problem. We have an intensity function *I(x, y)*, the image, and an assumed reflectance function *R(p, q)*. We have:

$$I(x,y) = R(p,q) \tag{4. 2}$$

This is an equation with two unknowns. All we know is that the surface orientation that has produced *I(x, y)* must lie somewhere on a contour of *R(p, q)*. As an infinite number of surface orientations can lie on a single contour, we need further constraints. 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 method then grows a solution by moving a small amount in the image plane along the contour.

The single-image shape from shading algorithm is still limited even if the exact lighting condition and surface reflectivity are known. One extreme case is that the 3D surface information may be totally lost under certain lighting conditions, and so there is no way to recovery the surface orientation. The problems with extracting shape from shading with a single image are that:

1) It is an approach that relies on having a known reflectance function for a surface.
2) One constraint exists in terms of the mathematical solution for this method. The method relies on *I(x, y)* being continuous. This means there are no discontinuities on the surface, and is therefore unsuitable for 3D texture surface estimation.
3) Another problem is that we really need a starting point to grow a solution. Because the equations that are solved are not over-constrained the method is extremely susceptible to noise in the image.

4.1.4. Photometric stereo

- ***What is photometric stereeo?***

Photometric stereo gives us ability to estimate local surface orientation by using several images of the same surface taken from the same viewpoint but under illumination from different directions (*Figure 4. 2* and *Figure 4. 3*). It was first introduced by Woodham [Woodham80]. 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 from which to measure surface orientation. Therefore, the change of the intensities in the images depends on both local surface orientation and illumination direction.

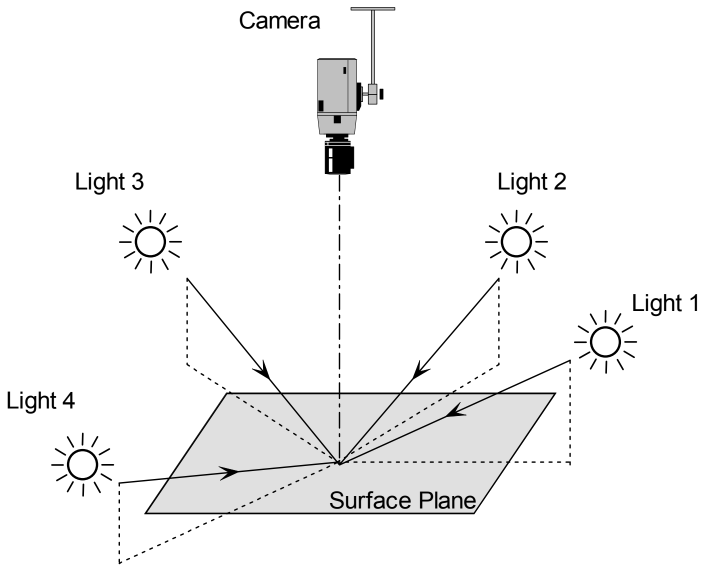

Figure 4. 2 Illustration of photometrc stereo geometry.

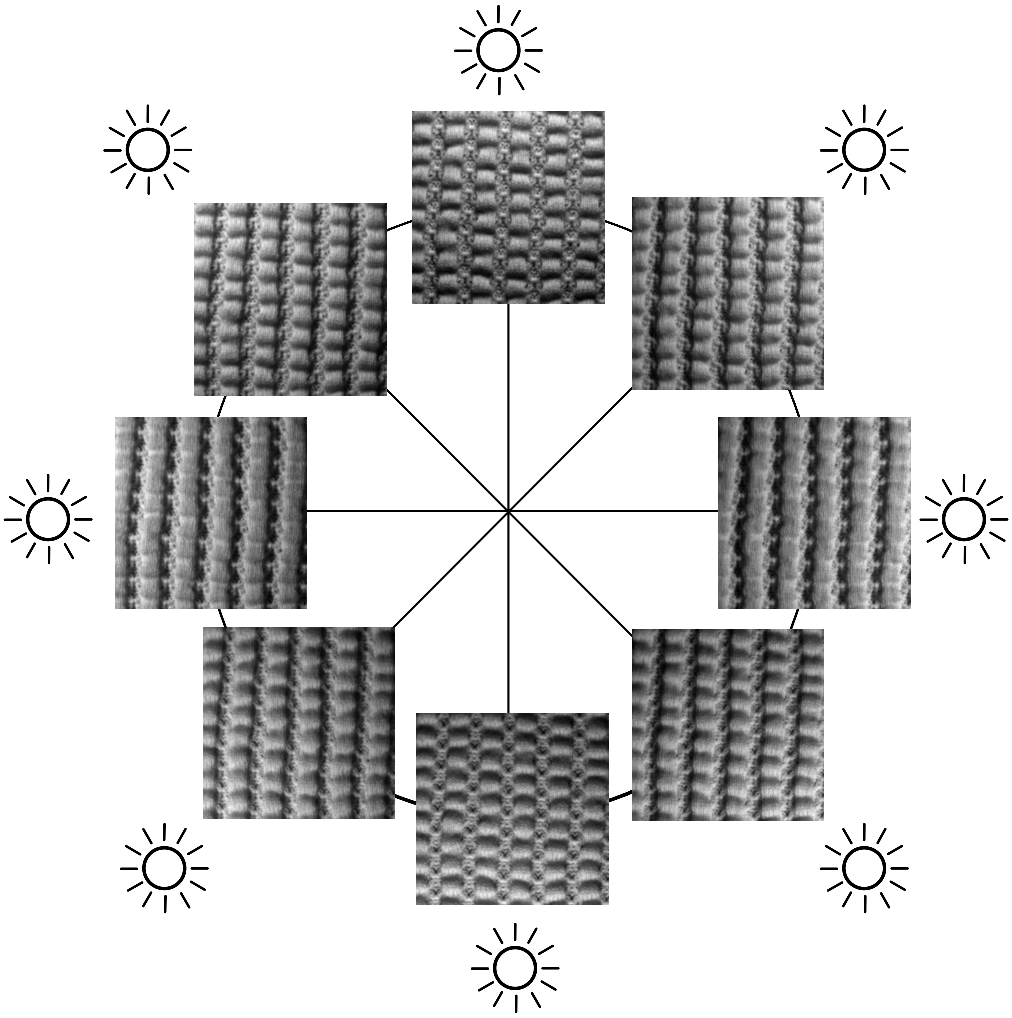

Figure 4. 3 Example of photometric images illuminated by different lighting sources.

- ***Why choose photometric stereo?***

Photometric stereo is a way in which the ill-posed problems in shading from shading can be resolved. It uses several images of the same surface under different illumination directions. The advantages of photometric stereo are:

1) Unlike single image shape from shading algorithms, photometric stereo makes no assumption of the smoothness of the surface (p and q vary continuously over the surface).
2) Furthermore, it requires only additional lighting and can be easily implemented at a reasonable computational cost.
3) Each image brings along its own unique reflectance map, therefore each image will define a unique set of possible orientations for each point.

4) Photometric stereo can recover not only surface orientation *(p, q)*, but also surface albedo ρ [Woodham80].

All of above points make photometric stereo the most suitable candidate relevant to this thesis.

4.2. A General Review of the Development of Photometric Stereo

Woodham [Woodham80] was the first to introduce photometric stereo. He proposed a method which was simple and efficient, but only dealt with Lambertian surfaces and was sensitive to noise. In his method, surface gradient can be solved by using two photometric images, assuming that the surface albedo is already known for each point on surface.

To determine local surface orientation, we need additional information. The simplest approach is to take two images which are of the same surface scene but with different light sources. Therefore we obtain two values of image intensity, $I_1(x, y)$ and $I_2(x, y)$ at each point *(x, y)*. In general, the image intensity values of each light source correspond to two points on the reflectance map, as follow

$$I_1(x,y) = R_1(p,q) \quad \text{and} \quad I_2(x,y) = R_2(p,q) \tag{4. 3}$$

Thus we can determine the surface normal parameters from two images. Defining the two light source vectors as $[p_1, q_1, -1]$ and $[p_2, q_2, -1]$, and equations *(4. 3)* as linear and independent, there will be a unique solution for *p* and *q* [Horn86] shown as follow:

$$p = \frac{(I_1^2 r_1 - 1)q_2 - (I_2^2 r_2 - 1)q_1}{p_1 q_2 - q_1 p_2} \tag{4. 4}$$

$$q = \frac{(I_2^2 r_2 - 1)p_1 - (I_1^2 r_1 - 1)p_2}{p_1 q_2 - q_1 p_2} \tag{4. 5}$$

where provided $p_1 / q_1 \neq p_2 / q_2$; $r_1 = \sqrt{1 + p_1^2 + q_1^2}$ and $r_2 = \sqrt{1 + p_2^2 + q_2^2}$. This gives a unique solution for surface orientation at all points in the image.

If the equations *(4. 3)* are non-linear, there are either no solutions or several solutions. In the case of a Lambertian reflectance function, we have to introduce another image to remove such ambiguities. This image enables us to estimate another surface parameter, *albedo*. It is especially useful in some cases where a surface is not uniform in its reflectance properties.

Lee and Kou [Lee93] were the first ones to introduce parallel and cascade photometric stereo for more accurate surface reconstruction. Parallel photometric stereo combined all of the photometric images together in order to produce the best estimation of the surface. Cascade would take the images, one after another, in a cascading manner. Compared with the conventional photometric stereo method, their iterative method has two major advantages. Firstly, this method determines surface heights directly but surface orientation as the conventional photometric stereo, therefore the integrability problem does not arise in this method. Second, this method is a global method that minimises the intensity errors over all points so that it is insensitive to noise. However, our task is to estimate the surface orientation rather than surface heights.

Cho and Minamitani [Cho93] have applied photometric stereo with three point light sources to recover textured and/or specular surfaces in closed environments like the gastric tract. Their concern was to reduce three-dimensional reconstruction errors due to specularities. Specular reflection produces incorrect surface normal by elevating the image intensity. Facets with estimated reflectivities greater than two standard deviations above the distribution mean are classified as being specular. Therefore they readjusted the pixel with greatest intensity by re-scaling with a modified reflectivity. In that way, the 3-D reconstruction errors may be reduced.

Iwahori and Woodham [Iwahori95] used principal components analysis (PCA) to extract a reduced dimensionality subspace from many more than the theoretical minimum number of images required. It applied two neural networks that were trained on a calibration sphere. The first one maps image irradiance to surface normal and estimates the surface derivatives, whereas the second one takes these estimates and forms an estimate of the intensity for the facet. The comparison yields a confidence estimate.

Kay and Caelli [Kay95] not only use photometric stereo to estimate the surface normal but also the roughness parameters associated with the Torrance-Sparrow (TS) reflectance model. They assume the reflectance map is a simplified Torrance-Sparrow map with additional Lambertian and mirror-like specular terms. No smoothness or regularization assumptions need be made. The basic approach they use is to apply non-linear regression techniques to photometric stereo. It is noted that in this thesis we only take account into diffuse Lambertian reflectance model, and both glossy specular and mirror specular factors are ignored.

McGunnigle [McGunnigle98] introduces a simple photometric stereo scheme which only considers a Lambertian reflectance model, where the self and cast shadow as well as inter-reflections are ignored. Three images at tilt angle of $90°$ increments are captured and the linear functions mapping surface partial derivatives to image intensity are required. Furthermore his approach does not require albedo information. He suggests using his scheme to provide a first estimate for an iterative procedure. This is in fact a simplified version of Woodham's scheme in which the illumination directions are chosen in order to simplify the mathematics.

Another more difficult problem is that of estimating a surface with an unknown reflectance map. Nayar [Nayar90] used a linear combination of Lambertian and an impulse specular component. He used distributed light sources for photometric stereo of surfaces whose reflection is a sum of specular and Lambertian components.

A technique that separates the effects of geometry and surface coloration/texture in "*tri-luminal*" environment is developed by Angelopoulou and Williams [Angelopoulou99]. Precision calibration of the illumination is not required. They demonstrate that the tri-luminal environment supports a broad range of analysis techniques: Differential geometric properties of smooth surfaces can be calculated directly from photometric data, albedo can be isolated and geometry-based recognition can be performed without recovering surfaces. Avoiding the recovery step increases both speed and robustness.

Belhumeur *et al* show that the set of images produced by arbitrary illumination of an object is the same as the set of images produced by what they call a "generalized bas-relief transformation" of the object [Belhumeur99]. A generalized bas-relief transformation is a transformation of both the surface shape and the surface albedo for an arbitrary Lambertian surface. While it has been thought that photometric stereo with unknown light source direction could be solved by first estimating the light source directions and then estimating the surface structure, they have shown that these estimates are coupled through an unsolvable generalized bas-relief transformation.

Other robust methods using photometric stereo with more images are reviewed in the latter *section 4.4*.

4.3. Three Image Based Photometric Stereo

In this section, we propose a simple three-image-based photometric stereo solution.

4.3.1. Three Photometric Images

We use three photometric images to recover the local surface orientation and albedo information for a Lambertian surface where the shadow effect is absent. The geometry of the camera / lighting setup can be seen in *Figure 4. 4*. Note that, during image capture, we keep the slant angle constant ($\sigma=50°$). Furthermore, we apply the same assumptions described in *section 3.3.1*, that is that the surface is parallel to the image plane of the camera and approximately flat; the light source and camera are located far away from the test sample; and the illumination direction and viewing direction are uniform at each point on the surface.

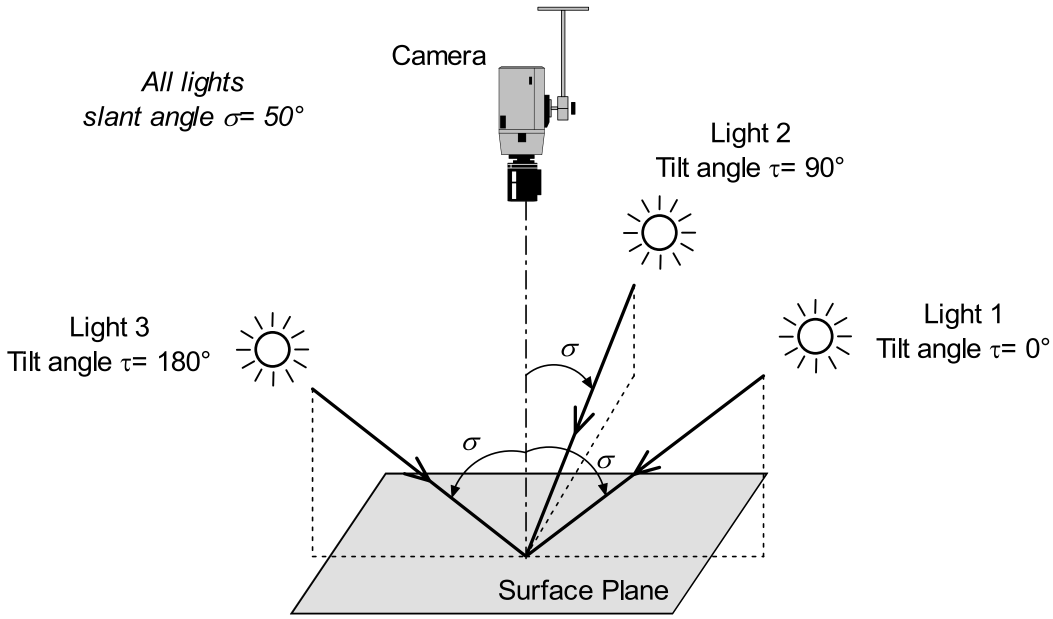

Figure 4. 4 Geometry of three-image based photometric stereo.

4.3.2. Equations of Photometric Stereo

We transfer the Lambertian surface model denoted in *section 3.2.2* into matrix format. Therefore the intensity of pixel in an image can be expressed as follow

$$i_1(x,y) = \rho(\mathbf{L_1} \cdot \mathbf{N}) \tag{4. 6}$$

where

- $i_1(x, y)$ is image intensity at the point (x, y);
- $\boldsymbol{N} = \frac{(p,q,-1)^T}{\sqrt{p^2+q^2+1}} = \left(\frac{-p}{\sqrt{p^2+q^2+1}}, \frac{-q}{\sqrt{p^2+q^2+1}}, \frac{1}{\sqrt{p^2+q^2+1}} \right)^T$ is a surface normal unit vector to the surface $s(x, y)$ at the point (x, y), and $p = \frac{\partial s(x,y)}{\partial x}$ and

$q = \frac{\partial s(x,y)}{\partial y}$ are surface partial derivatives measured in the x and y directions, respectively;

- $\mathbf{L_1}$ = [l_{x1}, l_{y1}, l_{z1}]^T is a unit illumination vector, which is pointing from the surface towards to the light source;
- ρ is surface albedo at the given point *(x, y)*.

Now we consider three light sources with illumination vectors $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$. The equation *(4. 6)* can be rewritten in matrix form

$$\mathbf{I} = \rho \cdot \mathbf{L} \cdot \mathbf{N} \tag{4. 7}$$

where

- $\mathbf{I}$ = [i_1, i_2, i_3]^T is image intensity vector;
- $\mathbf{L} = [\, \mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3} \,]^T$ is photometric illumination matrix which incorporates the light intensity for each light source.

Provided that all of three illumination vectors $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$ are not lying in the same plane (non-coplanar), then the photometric illumination matrix $\mathbf{L}$ is non-singular and its inverse matrix, $\mathbf{L^{-1}}$ exists and

$$\mathbf{M} = \mathbf{L}^{-1} \cdot \mathbf{I} = \rho \cdot \mathbf{N} \tag{4. 8}$$

where $\mathbf{M}$ = [m_1, m_2, m_3]^T

Therefore the three image based photometric stereo method can be summarised as follow:

1) For each given point *(x, y)* on the surface, the image intensity vector $\mathbf{I}$ is firstly formed by capturing three images under different illumination directions $\mathbf{L_1}$, $\mathbf{L_2}$ and $\mathbf{L_3}$.
2) The vector $\mathbf{M}$ = [m_1, m_2, m_3]^T is obtained by the production of $\mathbf{I}$ and $\mathbf{L^{-1}}$.
3) The *surface gradient components* can be calculated via

$$p = -\frac{m_1}{m_3} \quad \text{and} \quad q = -\frac{m_2}{m_3} \tag{4. 9}$$

4) Finally, the *surface albedo* is recovered by finding the length of vector **M**.

$$\rho = \sqrt{m_1^2 + m_2^2 + m_3^2} \tag{4. 10}$$

Those computation is straightforward in this case, and a unique result is assured.

4.3.3. Separating Gradient and Albedo using Photometric Stereo

An example of the results that can be obtained using this approach is shown below in *Figure 4. 5*.

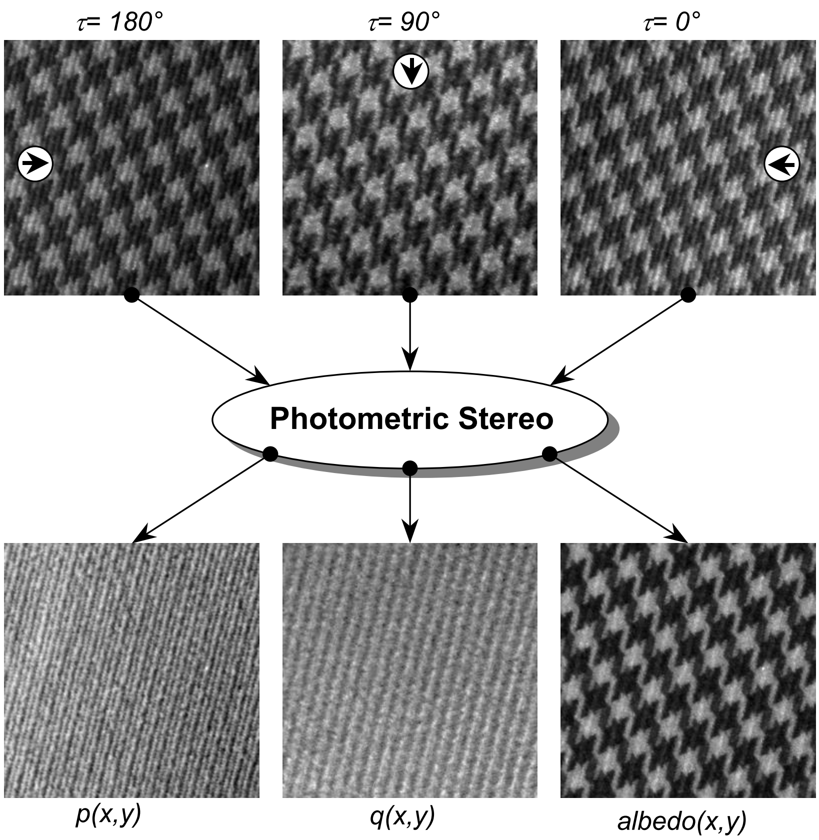

Figure 4. 5 Separating gradient and albedo information using photometric stereo.

4.4. Improvement on Three Image Based Photometric Stereo

There are two well-known problems with the traditional photometric stereo methods:

1) Surfaces are not ideally Lambertian and can contain a *specular* component; and
2) There will be some points that will be in *shadow* for one or more of the images. Therefore, there will not always be the three non-zero values of **I** to use in solving equation *(4. 8)*.

There are many approaches to deal with these problems documented in the computer vision literature.

As stated before, while three photometric images are used to estimate surface orientation and surface albedo, if four or more images are captured using further illumination sources, more information can be obtained. While the concept of photometric stereo is fairly simple, great difficulties arise when considering a realistic illumination model, in particular, a model that can represent diffuse and specular reflection. Only by considering such a model can an accurate representation of a wide range of material properties be achieved.

Ikeuchi [Ikeuchi81] first applied the photometric stereo method to specular surfaces, by using three extended light sources and the reflectance maps for each source in the form of lookup tables. In his research, he used a distributed light source obtained by uneven illumination of a diffusely reflecting planar surface and three input images. This method assumed a known object position and required accurate measurement of reflected brightness.

Coleman and Jain [Coleman82] extended the method and used four-light-source photometric stereo to determine the shape of surfaces that are non-Lambertian. It was based on the assumption that specular highlight areas between images do not overlap; therefore, they used relative deviation to determine the specular source. They proposed to calculate four albedo values based on the four possible combinations of three light sources. For a perfectly Lambertian surface, the four albedo would be identical. However, for surfaces that exhibit some specularities, this is not the case. Their method reduces the problem to Woodham's photometric stereo solution for three sources. However, they do not recover any specular parameters, only geometric and Lambertian ones.

Compared to Coleman and Jain's method, Tagare and deFigueriredo [Tagare90] do consider Lambertian and glossy specular objects. They show the uniqueness and the completeness of photometric stereo using their m-lobed reflectance maps. They give some results for a synthetic sphere using the reflectance maps of the Torrance-Sparrow model. They used eight light sources to recover the shape of real objects. Their work was continued by Kay and Caelly [Kay95] who investigated the problem of simultaneous estimation of surface normals and surface reflectance parameters from a practical point of view.

Solomon and Ikeuchi [Solomon92] extended Coleman and Jain's method by using four light sources. The shape information is produced directly by three light and four light photometric stereo methods. After they have shape information, statistical segmentation techniques can be applied to determine which pixels are specular and which are non-specular. Then, they can use the specular pixels and shape information, in conjugation with a simplified Torrance-Sparrow reflectance model to determine the surface roughness.

Nayar, Ikeuchi and Kanade [Nayar90] also developed a theory which accounted for specularities. They developed another reflectance model, called a *hybrid reflectance model*, which is the weighted sum of a diffuse lobe represented by a cosine function, and a specular spike, modelled as a delta-function. The relative strength of the specular and diffuse models was not known, and the proposed method could recover not only the local gradient but also the parameters of the reflectance model.

Rushmeier and Taubin [Rushmeier97] designed a system for obtaining bump maps from small sets of images captured under controlled lighting conditions. The bump map capture system complements inexpensive techniques for obtaining complex input data for rendering. The resulting maps can be used to re-render objects without reconstructing the original geometry. The map can also represent fine scale and self-shadowing geometry that would be difficult to recover from traditional photometric stereo.

Drbohlav and Leonardis [Drbohlav98] presented a global approach to the problem of removing shadows and highlights which is based on photometric image sets. The method exploits the basic properties of Lambertian surfaces and treats shadows and specularities as outliers. They showed that when three images exist in which the brightness value of a given pixel behaves Lambertian, it is enough to predict the brightness of this pixel in any image taken under arbitrary illumination. They also demonstrated that the performance of the principal component analysis can not improve even while increasing the amount of the input image data.

We note that there is no analysis of these above techniques with regard to the techniques of removing the artefacts of shadow and specularity. In *Chapter 8*, we propose a simple but efficient method to reduce effect of artefacts, which is similar to Rushmeier's method [Rushmeier97].

4.5. Summary

In this chapter, we introduce the technique of photometric stereo which gives us the ability to estimate surface properties using several images of a surface taken from the same viewpoint but under illuminations from different directions.

Firstly, some of candidate techniques are compared with photometric stereo. Binocular stereo is not suitable for us mainly due to correspondence problem. Shape from shading is based on smoothness and continuity, although it only needs one image. Photometric stereo technique does not have this problem. It allows us to obtain a surface description from several images of the same texture surface under various illumination directions. It is also a useful means that is suitable to recovery of the surface characteristics and matches our needs due to the problem caused by variation in illuminant direction.

Furthermore, a general review of development on photometric stereo is presented. It is followed by a solution to photometric stereo, where three photometric images are used to recover both the local surface orientation and albedo information given the assumption that Lambertian surface and both shadow and specular factors are absent. The equations of photometric stereo are expressed in the format of matrix equations.

Finally, some of techniques used to improve photometric stereo by considering shadow and specular effects are reviewed.