Towards Uncalibrated Photometric Stereo for Non-Lambertian Surfaces

A Dissertation Presented to the Faculty of Electrical Engineering of the Czech Technical University in Prague in Partial Fulfillment of the Requirements for the Ph.D. Degree by

Ondřej Drbohlav

September 8, 2003

Thesis Advisor

Prof. Ing. Václav Hlaváč, CSc.

Thesis Co-Advisor

Dr. Ing. Radim Šára

Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13, 121 35 Prague 2, Czech Republic
fax: +420-224-357-385, phone: +420-224-357-465
http://cmp.felk.cvut.cz

To Mary

Abstract

Photometric stereo is a method that captures local shape and reflectance of a 3D object from several intensity images of this object taken under different illuminations and a fixed viewpoint. The ability to estimate the surface reflectance parameters and local shape makes photometric stereo an invaluable tool for capturing intrinsic properties of real surfaces. The greatest benefit from this ability receive mainly augmented reality applications where realistic texture mapping and relighting of virtual objects is widely needed. The recovered reflectance parameters can also be used for visual quality control applications.

We are interested in photometric stereo setups in which the illumination conditions used to acquire individual input images are completely unknown except that they are provided by a single point illuminant at infinity. So far, such *uncalibrated* photometric stereo setup has been understood for Lambertian surfaces only. It is known that the surface parameters can be recovered up to nine degrees of freedom (9dof). In this thesis, we make a step towards answering the following question: If the object reflectance is not Lambertian, what is the ambiguity up to which the local surface parameters can be recovered by uncalibrated photometric stereo?

We choose the following approach to answer this question. We identify fundamental properties of reflectance which are shared over large classes of BRDFs. These properties include, for example, symmetries (e.g., the BRDF reciprocity or isotropy). We analyse whether a given reflectance property is able to deliver some information about the surface or light parameters. We approximate an arbitrary surface reflectance by Lambertian reflectance and perform the surface parameters recovery using uncalibrated Lambertian photometric stereo. Subsequently, we use additional information provided by a given BRDF property to reduce the 9dof ambiguity.

The main result of this thesis is that we found two fundamental reflectance properties which are useful for this goal. Firstly, we showed that if the surface reflectance is a composition of Lambertian and ideal specular reflectance then the albedo map is recovered up to a global scale factor, and the normal map is recovered up to a rotation around the viewing direction (1dof) and the change in coordinate frame handedness (a binary ambiguity). Subsequent application of the known integrability constraint which requires the normals to correspond to a continuous surface results in further reducing the ambiguity of the normal map determination to only the convex-concave (binary ambiguity). Secondly, we showed that if the BRDF is separable and isotropic then it is possible to construct a measurement which reveals the projections of the normals onto the image plane and onto the plane perpendicular to the light direction. In this case the albedo map is again recovered up to a global scale, and the normals are determined up to a convex-concave (binary) ambiguity without the need of the integrability constraint.

These results identify two important cases in which the necessity of knowing the illumination parameters can be relaxed. They are likely to become the basis of even wider use of photometric stereo in practice.

Acknowledgements

First of all, I would like to thank Radim Šára for truly being my advisor and a colleague of mine during the entire period of my PhD studies.

Václav Hlaváč deserves a special acknowledgement for wise management of our research group and for pushing me kindly towards finishing my PhD thesis in a reasonable time. A lot of thanks is addressed to him and to other colleagues from the CMP for their support. I confess I can hardly imagine a workplace with a more suitable atmosphere.

I am grateful to my thesis advisors as well as to Tomáš Pajdla and Tomáš Werner for careful reading drafts of this thesis and providing me with comments and suggestions which considerably increased the readability of the manuscript.

I wish to thank Aleš Leonardis of the Faculty of Informatics of Ljubljana who adopted me into the Slovenian computer vision group during my 'early days' in the computer vision field.

I thank my family for encouragement and motivation. I am especially grateful to my wife Mária for her continuous support. She deserves a warm thank for her patience during all those days I was writing this thesis.

I gratefully acknowledge grants GACR 102/01/1371, MSMT J04/98:212300013, CVUT 30/99068/333 and VS96049 which supported my research.

Contents

Chapter 1

Introduction

1.1 What is photometric stereo?

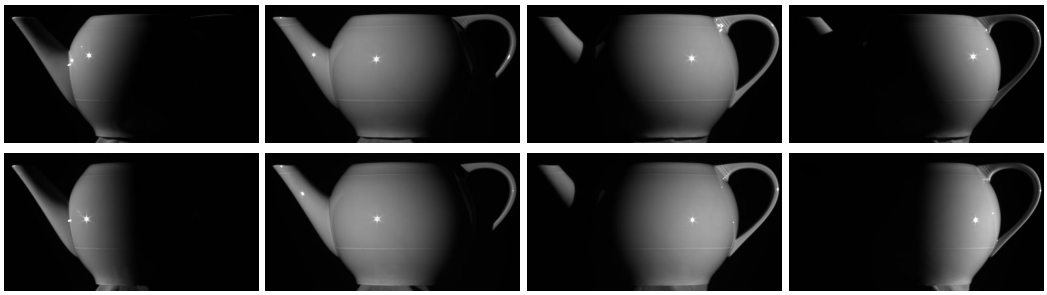
Photometric stereo is a method that captures local shape and reflectance of a 3D object from several intensity images of this object taken under different illumination and a fixed viewpoint. Figure 1.1 shows an example: The top of the figure (a) displays the input image data set. The recovered surface orientations (b) and the reflectance parameters (c) are shown in the bottom. The normals and reflectance parameters are evaluated for each image point independently.

Photometric stereo captures the surface normals with an accuracy which is generally not reachable by triangulation-based methods (e.g., binocular stereo or rangefinder methods) because such approaches estimate the depth of the surface and can deliver the surface normals only by the depth map differentiation. Photometric stereo, conversely, estimates the normal directions directly, but it is generally not able to provide an accurate depth map. If needed, the depth map can only be estimated by an integration of the normal map provided that the actual surface depth is continuous.

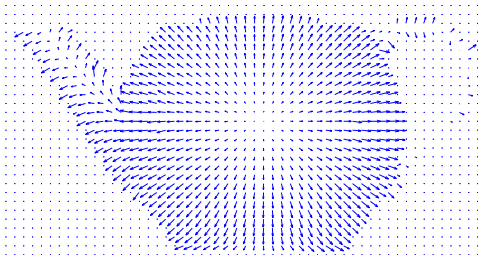
The ability to estimate the surface reflectance parameters and local shape makes photometric stereo an invaluable tool for capturing intrinsic textures of real surfaces. The applications which benefit from this ability include film post production, interactive video and other augmented reality applications where texture mapping and relighting of virtual objects is widely needed. The recovered reflectance parameters can also be used for quality control applications in, for example, food manufacturing or surface finishing industry.

1.2 The principle

This section reviews the relationship between different illumination conditions and the image content, and shows how this relationship is employed to compute orientations and reflectance properties of an object surface.



(a) Input images



(b) Normal map

(c) Reflectance map

Figure 1.1: The input and output of photometric stereo: images of a 3D object observed under different illuminations from the same viewpoint (a) are used to compute the normals (b) (shown as orthographic projections onto the retina) and reflectance parameters (c). Both entities are recovered at each image point. For the sake of readability, the resolution of the normal map was reduced by the factor of 10 in both coordinate directions.

1.2.1 Bidirectional reflectance distribution function

The principal role in photometric stereo is played by the bidirectional reflectance distribution function (BRDF) [28] which describes the surface reflectance. Consider a surface patch which is depicted in Fig. 1.2 together with an outward-pointing surface normal $\mathbf{n}$ and a *reference plane* whose basis are the vectors $\mathbf{n}_\perp^1$, $\mathbf{n}_\perp^2$ such that $[\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]$ forms an orthonormal basis of $\mathbb{R}^3$. The domain of the BRDF is the Cartesian product of all incident and reflected directions which are traditionally represented in spherical coordinates. For any incident direction (θ_i, ϕ_i) and any reflected direction (θ_r, ϕ_r), the BRDF tells what is the ratio of radiance* $\mathrm{d}L_r(\theta_r, \phi_r)$ of a surface in a reflected direction (θ_r, ϕ_r) to its irradiance $\mathrm{d}E_i(\theta_i, \phi_i)$ incident from a solid angle $\mathrm{d}\Omega_i$ centered around the incident direction (θ_i, ϕ_i):

$$f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r) = \frac{\mathrm{d}L_r(\theta_r, \phi_r)}{\mathrm{d}E_i(\theta_i, \phi_i)} = \frac{\mathrm{d}L_r(\theta_r, \phi_r)}{L_i(\theta_i, \phi_i)\cos\theta_i \mathrm{d}\Omega_i}\,, \tag{1.1}$$

where $L_i(\theta_i, \phi_i)$ is the incident radiance.

*See Section 2.4 for the definition of radiometric terms.

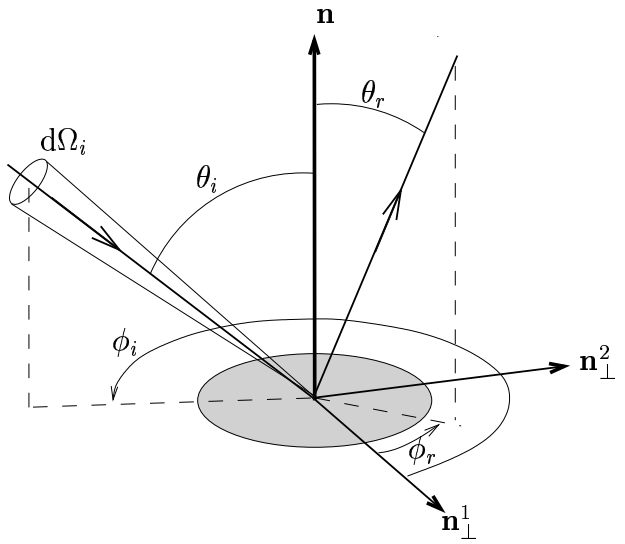


Figure 1.2: A surface patch and the parametrisation of incidence and reflected directions in spherical coordinates.

The BRDF is usually represented by a closed-form reflectance model[†]. In addition to the dependence on the incident and reflected directions, the reflectance model often accommodates a small number of parameters $\boldsymbol{\alpha} = [\alpha_1, .., \alpha_m]^T$. For example, reflectance on very rough surfaces is different from reflectance on almost smooth surfaces; if there exists a reflectance model capable of describing reflectance for both cases, it would comprise the roughness parameter. Whenever this fact is to be emphasised, we write the BRDF as $f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r; \boldsymbol{\alpha})$.

Provided that the incident radiance $L_i(\theta_i, \phi_i)$ is known, the BRDF enables to compute the surface radiance for an arbitrary reflected direction as

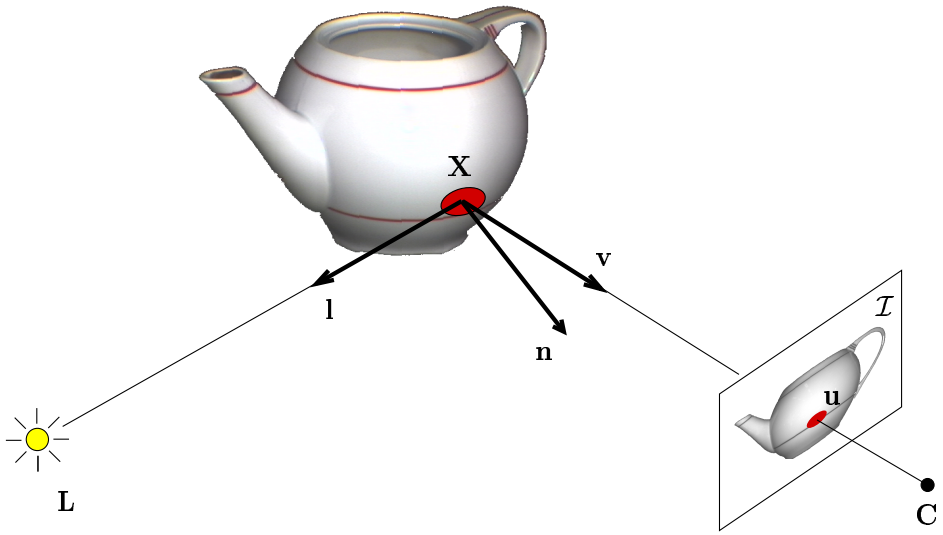
$$\begin{aligned} L_r(\theta_r, \phi_r) &= \int f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r; \boldsymbol{\alpha}) \mathrm{d}E_i(\theta_i, \phi_i) \\ &= \int f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r; \boldsymbol{\alpha}) L_i(\theta_i, \phi_i) \cos\theta_i \mathrm{d}\Omega_i \,, \end{aligned} \tag{1.2}$$

where the integration is performed over the entire upper hemisphere.

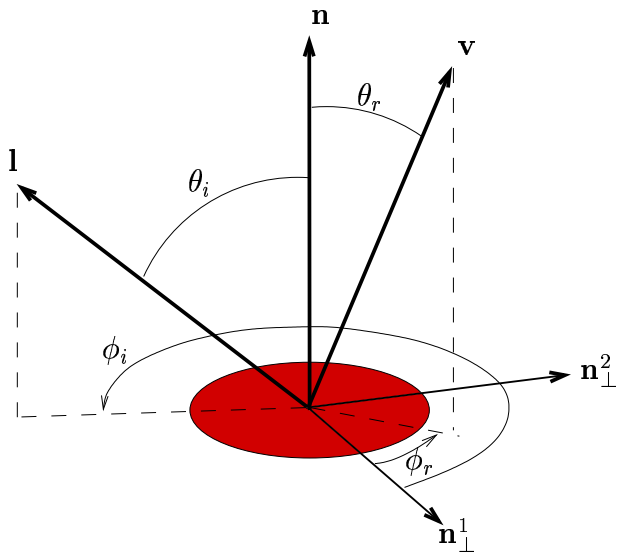
1.2.2 Observed image intensity

Let us have a look at a simple photometric stereo setup in which different illumination conditions are provided by different positions of a single point light source (see Fig. 1.3(a)). The figure depicts a 3D object observed by a camera positioned at $\mathbf{C}$ and illuminated by a point light source positioned at $\mathbf{L}$. A particular point $\mathbf{u} \in \mathcal{I}$ in the image plane $\mathcal{I}$ defines the viewing ray whose intersection with the

[†]In this thesis we use the terms 'BRDF', 'reflectance model' and even 'reflectance' interchangeably whenever there is no risk of confusion.



(a)



(b)

Figure 1.3: (a) A 3D surface is illuminated by a single point light source positioned at **L** and viewed by a camera positioned at **C**. A pixel $\mathbf{u} \in \mathcal{I}$ observes a surface patch (shown as a red ellipse) with the normal **n** positioned at **X**. Vectors **l** and **v** are unit vectors pointing from the surface patch towards the light source and the camera, respectively. (b) Representation of vectors **v** and **l** in a local coordinate frame of the surface patch $[\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]$ by spherical coordinates (θ_i, ϕ_i) and (θ_r, ϕ_r), respectively.

surface gives the position $\mathbf{X}$ of a surface patch (shown as a red ellipse) which this pixel observes. Unit vectors $\mathbf{v}$ and $\mathbf{l}$ pointing from the surface patch towards the camera and towards the light source are termed the *viewing vector* and the *light vector*, respectively. These two vectors define the reflected and incident directions (see Fig. 1.3(b) where the surface patch is shown enlarged together with its coordinate system).

Equation (1.2) enables to enumerate the patch radiance in a viewing direction resulting from its irradiance by the point light source at $\mathbf{L}$. Without loss of generality, we assume that the radiant intensity of the light source is isotropic and equal to I_e. As there is only one incident direction $\mathbf{l}$ of nonzero incident radiance and the irradiance resulting from it is $E_i = I_e \cos\theta_i / \|\mathbf{L} - \mathbf{X}\|^2$, Equation (1.2) becomes

$$L_r(\mathbf{u}) = f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r) \frac{I_e \cos\theta_i}{\|\mathbf{L} - \mathbf{X}\|^2}\,, \tag{1.3}$$

where $L_r(\mathbf{u})$ stands for the radiance of a patch being projected to $\mathbf{u}$. The incident and reflected directions (θ_i, ϕ_i) and (θ_r, ϕ_r) are given by

$$\begin{aligned} \theta_r &= \arccos(\mathbf{n}\cdot\mathbf{v})\,, & \phi_r &= \arctan\frac{\mathbf{n}_\perp^2\cdot\mathbf{v}}{\mathbf{n}_\perp^1\cdot\mathbf{v}}\,, \\ \theta_i &= \arccos(\mathbf{n}\cdot\mathbf{l})\,, & \phi_i &= \arctan\frac{\mathbf{n}_\perp^2\cdot\mathbf{l}}{\mathbf{n}_\perp^1\cdot\mathbf{l}}\,. \end{aligned} \tag{1.4}$$

What is the relationship between the radiance $L_r(\mathbf{u})$ given by (1.3) and the raw value $I_{raw}(\mathbf{u})$ output from the camera? Firstly, $I_{raw}(\mathbf{u})$ is the result of monotonic mapping of irradiance of the sensor element placed at $\mathbf{u}$. This mapping is performed by the electronic circuits of a camera. Secondly, this sensor irradiance is directly proportional to the observed surface radiance, but the factor of proportionality is dependent on a pixel position. It is relatively easy to compensate for both effects (i.e., gamma correction and vignetting) such that the corrected value which we call the *image intensity* is, up to a positive scale η, equal to the surface radiance. Denoting image intensity $I(\mathbf{u})$, that means:

$$\forall \mathbf{u} \in \mathcal{I}\colon \quad I(\mathbf{u}) = \eta\, L_r(\mathbf{u})\,, \tag{1.5}$$

We suppose that the camera is fully radiometrically calibrated such that it is able to measure absolute values of surface radiance, and we set $\eta = \pi$ for reasons which will become clear in Section 1.5.1. However we do not restrict ourselves only to this case of a full radiometric calibration of a camera. At all appropriate places in the thesis, we discuss how the results change if η is unknown.

Finally, let us denote the transformation (1.4) which performs the mapping of vectors into local spherical coordinates T, such that for any unit direction $\mathbf{x}$ within the upper hemisphere there holds

$$(\theta, \phi) = T(\mathbf{x}, [\mathbf{n}_\perp^1 | \mathbf{n}_\perp^2 | \mathbf{n}])\,. \tag{1.6}$$

According to (1.5) and (1.3), the image intensity at $\mathbf{u}$ is

$$\boxed{I(\mathbf{u}) = \eta L_r(\mathbf{u}) = \eta\, f_{i,r}(T(\mathbf{l}, [\mathbf{n}_\perp^1 | \mathbf{n}_\perp^2 | \mathbf{n}]); T(\mathbf{v}, [\mathbf{n}_\perp^1 | \mathbf{n}_\perp^2 | \mathbf{n}]); \boldsymbol{\alpha}) \frac{(\mathbf{n} \cdot \mathbf{l})\, I_e}{\|\mathbf{L} - \mathbf{X}\|^2}\,.} \tag{1.7}$$

This equation answers the question what image intensity is observed at image point $\mathbf{u}$ when the patch is illuminated by the light source at position $\mathbf{L}$.

Let us assume that all variables that appear on the right-hand side of this equation are known except for the surface patch basis $[\mathbf{n}_\perp^1 | \mathbf{n}_\perp^2 | \mathbf{n}]$ and the BRDF parameters $\boldsymbol{\alpha}$. The strategy of the photometric stereo method to find the unknowns is to acquire the image intensities for different light source positions, and then find the surface patch basis and the BRDF parameters such that the measured data and the model represented by the set of equations (1.7) are put into agreement. Finding the patch basis comprises finding the patch normal and the reference plane basis vectors $[\mathbf{n}_\perp^1, \mathbf{n}_\perp^2]$. Note that for isotropic BRDFs[‡] a particular choice of the reference plane basis does not play a role since it does not affect the validity of Equation (1.7).

1.2.3 Lights at infinity

In the previous section, we assumed that all entities present in (1.7) are known except for the surface patch basis $[\mathbf{n}_\perp^1 | \mathbf{n}_\perp^2 | \mathbf{n}]$ and the BRDF parameters $\boldsymbol{\alpha}$. The knowledge of the patch position $\mathbf{X}$ which induces the light vector direction $\mathbf{l}$ is, however, problematic. In real situations, the geometry of the surface is rarely at hand in advance, and therefore the distance of a surface point from the camera center for a given image point $\mathbf{u}$ is unknown.

It would be possible to include the unknown patch position $\mathbf{X}$ on a viewing ray into the set of parameters that have to be found by the photometric stereo. There exists a practical approach which avoids that. Consider the case in which a light source is positioned in a distance much larger than the object size. Then for all points $\mathbf{X}$ on an object surface $\mathcal{Q}$ there holds

$$\begin{aligned} \forall \mathbf{X} \in \mathcal{Q}: \quad \mathbf{l}(\mathbf{X}) = \frac{\mathbf{L} - \mathbf{X}}{\|\mathbf{L} - \mathbf{X}\|} &\approx \mathbf{l}\,(\mathbf{X}_g)\ , \\ \sigma(\mathbf{X}) = \frac{I_e}{\|\mathbf{L} - \mathbf{X}\|^2} &\approx \sigma\,(\mathbf{X}_g)\ , \end{aligned} \tag{1.8}$$

where $\mathbf{X}_g$ is the center of gravity of the surface points and $\sigma(\mathbf{X})$ denotes *light strength* [§] at $\mathbf{X}$. By moving the light source to infinity (and increasing I_e appropriately such that $\sigma\,(\mathbf{X}_g)$ does not vanish) the illumination conditions for all surface points will be the same, characterised by the light vector $\mathbf{l}$ and the light strength σ. Equation (1.7)

[‡]Isotropic BRDFs are those for which there holds that $f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_i) = f_{i,r}(\theta_i, \phi_i + \gamma; \theta_r, \phi_i + \gamma) \quad \forall \gamma \in [0, 2\pi]$, i.e., they stay constant when the incident and reflected directions are jointly rotated around the normal.

[§]This is the irradiance of an imaginary surface positioned at the same place as the actual surface but oriented directly towards the incident illumination.

is then no more dependent on the depth of the surface patch that projects to pixel $\mathbf{u}$ and there holds that

$$\boxed{I(\mathbf{u}) = \eta\, f_{i,r}(T(\mathbf{l}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]); T(\mathbf{v}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]); \boldsymbol{\alpha})\ (\mathbf{n}\cdot\mathbf{l})\,\sigma\,.} \tag{1.9}$$

1.2.4 Lambertian photometric stereo

In this section, we review an important special example of recovery of normal and reflectance maps provided that the BRDF is represented by the Lambertian reflectance model and the illumination is provided by point light sources at infinity. The Lambertian BRDF

$$f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r;\ \rho) = \frac{\rho}{\pi} \tag{1.10}$$

comprises one parameter ρ called *albedo* which states what portion of the incident radiant flux is returned back to space in a form of diffuse reflectance. Ideal ('perfectly white') diffuse reflectors have $\rho = 1$ while an ideal absorber has $\rho = 0$. The Lambertian BRDF is constant for all incident and reflected directions.

According to (1.9) the image intensity observed at image point $\mathbf{u}$ is

$$I(\mathbf{u}) = \eta\,\frac{\rho}{\pi}\sigma\mathbf{n}\cdot\mathbf{l} = \frac{\eta}{\pi}\,(\rho\mathbf{n})\cdot(\sigma\mathbf{l}) = \frac{\eta}{\pi}\,\mathbf{b}\cdot\mathbf{s}\,, \tag{1.11}$$

where $\mathbf{b} = \rho\mathbf{n}$ is the unit normal vector scaled by albedo and $\mathbf{s}$ is the unit light vector scaled by the light strength of the light. We call $\mathbf{b}$ and $\mathbf{s}$ the *scaled normal vector* and the *scaled light vector*, respectively.

Suppose images have Q pixels and there are R point light sources at infinity that are used, one after another, for illuminating the object. Image intensity observed in q-th pixel under the illumination by r-th light is

$$I_{qr} = \begin{cases} (\eta/\pi)\,\mathbf{b}_q\cdot\mathbf{s}_r: & q\text{-th patch illuminated under } r\text{-th light,} \\ 0: & q\text{-th patch shadowed under } r\text{-th light.} \end{cases} \tag{1.12}$$

These intensities form the *image intensity matrix* $\mathbf{I} \in \mathbb{R}^{Q\times R}$. Suppose for a while that th q-th surface patch is illuminated under all point light source positions. Then the q-th row of the image intensity matrix can be written compactly as

$$\mathbf{I}_q = \frac{\eta}{\pi}\,\mathbf{b}_q^T\mathbf{S}\,, \tag{1.13}$$

where the matrix $\mathbf{S} \in \mathbb{R}^{3\times R}$ stacks all scaled light vectors $\mathbf{S} = [\mathbf{s}_1|\mathbf{s}_2|\ldots|\mathbf{s}_R]$. We call $\mathbf{S}$ the *scaled lights matrix*. If $\mathrm{rank}(\mathbf{S}) = 3$ then the scaled normal $\mathbf{b}_q$ can be determined from the image data as

$$\mathbf{b}_q = \frac{\pi}{\eta}\,(\mathbf{I}_q\mathbf{S}^{-1})^T\,, \tag{1.14}$$

where $\mathbf{S}^{-1}$ is the inverse (or pseudoinverse if $R > 3$) of the scaled lights matrix $\mathbf{S}$. To compute the normal $\mathbf{n}_q$ and the albedo ρ_q, it then suffices to decompose the scaled

normal vector $\mathbf{b}_q$:

$$\begin{aligned} \mathbf{n}_q &= \frac{\mathbf{b}_q}{\|\mathbf{b}_q\|}\,, \\ \rho_q &= \|\mathbf{b}_q\|\,. \end{aligned} \tag{1.15}$$

If the q-th patch is illuminated only under illuminations $\{r_1, r_2, \ldots, r_m\}$ then the scaled normal $\mathbf{b}_q$ can be computed by a formula analogous to (1.14) where $\mathbf{I}_q$ is replaced by $(I_{qr_1}, I_{qr_2}, \ldots, I_{qr_m})$ and $\mathbf{S}$ is replaced by $[\mathbf{s}_{r_1}|\mathbf{s}_{r_2}|\ldots|\mathbf{s}_{r_m}]$ provided that the latter matrix still has rank 3.

In order to write the relation between image intensities, scaled normal vectors and scaled light vectors in a compact form which will be found useful in the subsequent text, we define the *stacked image matrix* $\mathbf{J} \in \mathbb{R}^{Q\times R}$ as follows:

$$J_{qr} = \begin{cases} I_{qr}: & q\text{-th patch illuminated under } r\text{-th light,} \\ \{\emptyset\}: & q\text{-th patch shadowed under } r\text{-th light,} \end{cases} \tag{1.16}$$

where $\{\emptyset\}$ denotes a *missing value*. The stacked image matrix $\mathbf{J}$ is equal to the image intensity data matrix $\mathbf{I}$ except for the matrix elements where observed intensity I_{qr} can not be written as $\mathbf{b}_q \cdot \mathbf{s}_r$. The relationship between the stacked image matrix $\mathbf{J}$, the scaled normal vectors and the scaled light vectors is

$$\boxed{\mathbf{J} \overset{\bullet}{=} \frac{\eta}{\pi}\left[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_Q\right]^T \left[\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_R\right] = \frac{\eta}{\pi}\,\mathbf{B}^T\mathbf{S}\,,} \tag{1.17}$$

where $\mathbf{B}$ is the *scaled normals matrix*, $\mathbf{S}$ is the scaled lights matrix as above and the relation $\overset{\bullet}{=}$ is read 'equal where matrix elements on both sides are present'. Note that when $\mathbf{J}$ and $\mathbf{S}$ are given, the equation may determine the scaled normals matrix $\mathbf{B}$ only up to a scale if the factor η is unknown (cf. (1.5)). This ambiguity is usually not considered serious because it does not affect the determination of normals but only transforms the recovered albedos by a global scale (cf. (1.15)).

1.3 Standard assumptions

Thus we define the standard assumptions for photometric stereo as follows:

i) Single point light source at infinity is used to illuminate the object surface. Individual illuminations $r \in \{1, 2, \ldots R\}$ are represented by their light vectors $\mathbf{l}_r$ and direct incident irradiances σ_r, or alternatively by the scaled light vectors $\mathbf{s}_r = \sigma_r \mathbf{l}_r$.

ii) The surfaces are opaque and their reflectance is described by a parametric reflectance model $f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r; \boldsymbol{\alpha})$. The model parameters $\boldsymbol{\alpha}$ are allowed to vary across the surface.

iii) The camera is orthographic and all its parameters (see Section 2.1) are known. As the center of the orthographic camera is at infinity, the viewing vector $\mathbf{v}$ (cf. Fig. 1.3(a)) is the same for all surface points.

1.4 Calibrated vs. uncalibrated photometric stereo

For photometric stereo, the calibratedness is a term which is related to knowing or not knowing illumination conditions used for acquiring the input images. In this thesis we restrict our attention only to illuminations which are provided by single light sources at infinity. We define the calibrated and uncalibrated stereo scenarios as follows:

- **Calibrated photometric stereo** refers to the case when the light vectors $\mathbf{l}_r$ are all known and all the light strengths σ_r are known up to a global scale[¶].
- **Uncalibrated photometric stereo** refers to the case when all the light vectors $\mathbf{l}_r$ and all the light strengths σ_r are unknown.

Photometric stereo has been usually applied in a calibrated setup. There are two practical scenarios how this is achieved: i) placing several light sources rigidly into space and switching between them while acquiring the input data, or ii) using one light source which is being moved along a controlled path with stops for acquiring individual images.

This thesis studies the uncalibrated photometric stereo. The uncalibrated stereo is of advantage because the illumination conditions can be adjusted for each 3D object as needed and no rigidly mounted light-sources or moving light sources with a well-defined path are required. The task of uncalibrated stereo is to recover both the illumination conditions and the surface parameters from the input data. We believe that identifying conditions under which the uncalibrated photometric stereo is feasible is the key to even larger use of photometric stereo in practice.

1.5 State of the art in uncalibrated photometric stereo

So far, the uncalibrated photometric stereo has been understood only for surfaces with Lambertian reflectance. This is the reason for reviewing in this section the state of the art in uncalibrated *Lambertian* photometric stereo. Subsection 1.5.1 reviews how the scaled normals matrix $\mathbf{B}$ can be found from the stacked image matrix $\mathbf{J}$ if the scaled light matrix $\mathbf{S}$ is unknown. Subsection 1.5.2 describes the ambiguity in surface parameters determination. Subsections 1.5.4—1.5.7 review the individual possibilities how the ambiguity can be reduced.

1.5.1 Solution by factorisation

It was observed by Woodham [43] and Hayakawa [16] who worked under the assumption of point light sources at infinity that the surface parameters and the light source parameters can be determined from the input data, but not uniquely. This is shown as follows.

[¶]Observe on the example of Lambertian photometric stereo (1.17) that this global scale does not affect the determination of $\mathbf{B}$ in any way if the factor η is unknown.

Let us assume for a while that the stacked image matrix $\mathbf{J}$ in (1.17) is free of missing elements. Provided that the number of pixels is higher than the number of illuminations ($Q > R$), the SVD decomposition [14] of $\mathbf{J}$ is

$$\mathbf{J} = \mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{U} \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & d_R \\ 0 & 0 & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & 0 \end{bmatrix} \mathbf{V}^T, \qquad \begin{array}{l} \mathbf{U} \in O(Q)\,, \\ \mathbf{V} \in O(R)\,, \\ d_1 > d_2 > \ldots > d_R > 0\,, \end{array} \tag{1.18}$$

where $O(N)$ is the group of real orthogonal matrices of size $N \times N$ and $d_1, d_2 \ldots d_R$ are real singular values. In a noise-free measurement, the rank of $\mathbf{J}$, and consequently also of $\mathbf{D}$, is three. This rank is higher in the presence of noise. Setting

$$\mathbf{B} = \mathbf{U}_{Q\times 3}^T\,, \qquad \mathbf{S} = \mathrm{diag}\,[d_1, d_2, d_3]\,\mathbf{V}_{R\times 3}^T\,, \tag{1.19}$$

where $\mathbf{U}_{Q\times 3}$ is a submatrix formed from first three columns of $\mathbf{U}$ (and analogously for $\mathbf{V}_{R\times 3}$), one gives the best solution to (1.17) in a least-square sense. Note that by choosing $\mathbf{B}$ and $\mathbf{S}$ as in (1.19) we set $\eta = \pi$ in (1.17). We could fix the unknown factor η in other way but the particular choice $\eta = \pi$ leads to most visually compact formula: $\mathbf{J} \overset{\bullet}{=} \mathbf{B}^T\mathbf{S}$.

If the stacked image matrix $\mathbf{J}$ comprises missing elements then the factorisation of $\mathbf{J}$ into $\mathbf{B}$ and $\mathbf{S}$ can be performed using the method of Jacobs [18].

1.5.2 Uncalibrated Lambertian photometric stereo ambiguity

It turns out that the factorisation of $\mathbf{J}$ into $\mathbf{B}$ and $\mathbf{S}$ is ambiguous. This is so because

$$\forall \mathbf{A} \in GL(3): \quad \mathbf{B}^T\mathbf{S} = \mathbf{B}^T \underbrace{\mathbf{A}^T\mathbf{A}^{-T}}_{\text{identity}}\mathbf{S} = (\mathbf{A}\mathbf{B})^T(\mathbf{A}^{-T}\mathbf{S}) = \overline{\mathbf{B}}^T\overline{\mathbf{S}}\,, \tag{1.20}$$

where $GL(3)$ is the group of all invertible matrices of size 3×3. This fact defines the relation of equivalence in the solution space $\mathcal{T} = \mathbb{R}^{3\times Q} \times \mathbb{R}^{3\times R}$ where $\mathbb{R}^{3\times Q}$ and $\mathbb{R}^{3\times R}$ represent the space of all possible scaled normals matrices $\mathbf{B}$ and the space of all possible scaled lights matrices $\mathbf{S}$, respectively. We say that two solutions $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}$ and $(\overline{\mathbf{B}}, \overline{\mathbf{S}}) \in \mathcal{T}$ are equivalent if there exists $\mathbf{A} \in GL(3)$ such that

$$\begin{aligned} \overline{\mathbf{B}} &= \mathbf{A}\mathbf{B}\,, \\ \overline{\mathbf{S}} &= \mathbf{A}^{-T}\mathbf{S}\,. \end{aligned} \tag{1.21}$$

Suppose the photometric stereo input images are represented by the stacked image matrix $\mathbf{J}$. Matrix $\mathbf{J}$ selects the appropriate equivalence class $\mathcal{T}(\mathbf{J})$ by (1.19). This class contains the true scaled normals matrix $\mathbf{B}^*$ and the true scaled lights matrix $\mathbf{S}^*$, but it is not possible to distinguish $(\mathbf{B}^*, \mathbf{S}^*)$ from other members of $\mathcal{T}(\mathbf{J})$ just based on the image content represented by the stacked image matrix $\mathbf{J}$.

1.5.3 Ways to reduce the ambiguity

The uncalibrated photometric stereo ambiguity can be reduced or even removed only if additional information about the light parameters or surface parameters is available [42, 20].

The action of additional information is that it further subdivides the equivalence classes. To give an example, suppose that one scaled normal is known, the first one, say. Let us denote it as ${\mathbf{b}_1}^*$. Let $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be an arbitrary member of $\mathcal{T}(\mathbf{J})$. First we need to find a matrix $\mathbf{Q} \in GL(3)$ such that

$$\mathbf{Q}\mathbf{b}_1 = {\mathbf{b}_1}^* , \tag{1.22}$$

where $\mathbf{b}_1$ is the first column of $\mathbf{B}$. This equation expresses the constraint in an algebraic form. It enables us to find such $\mathbf{Q}$, and it also enables us to observe new equivalence classes. We observe that $\mathbf{Q}$ is not determined uniquely; any transformation $\mathbf{A}$ which leaves ${\mathbf{b}_1}^*$ invariant can be applied to both sides of (1.22) without altering the constraint. Hence, we observe that the original ambiguity $\mathbf{A} \in GL(3)$ *reduces* to an ambiguity represented by the class of matrices which leave the vector ${\mathbf{b}_1}^*$ invariant. This class is a group again.

An extreme case of ambiguity reduction is, for example, the case when three linearly independent true scaled normals are known. Then the matrix $\mathbf{Q}$ is determined uniquely and the ambiguity reduces to a trivial group comprising the identity matrix; informally speaking, the ambiguity vanishes.

1.5.4 Uniform albedo, equal light strengths

The first possibility is to assume that some light sources are of equal light strengths, or that albedo is uniform over some set of normals. Such possibility was employed and/or discussed in [16, 44, 2]. Note that due to the duality of scaled normals and lights, these two constraints are also dual. Let us demonstrate the analysis of the constraint for the case when there is a set of scaled normals which have the same length (i.e., albedo). Let this set be denoted $\mathcal{U}$ and the albedo value be denoted ρ.

Let us take an arbitrary solution $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ given by the factorization of the stacked image matrix $\mathbf{J}$. Such representative may differ from the true scaled normals and lights by an arbitrary transformation $\mathbf{A} \in GL(3)$. A matrix $\mathbf{Q}$ is required which makes the data consistent with the prescribed albedo constraint. We require

$$\forall \mathbf{b} \in \mathcal{U}, \quad \|\mathbf{Q}\mathbf{b}\| = \rho \, . \tag{1.23}$$

Rewriting this, we have

$$\forall \mathbf{b} \in \mathcal{U}, \quad \mathbf{b}^T \mathbf{Q}^T \mathbf{Q} \mathbf{b} = \rho^2 \, . \tag{1.24}$$

We observe that the transformation $\mathbf{Q}$ is present in this equation in the form of $\mathbf{P} = \mathbf{Q}^T\mathbf{Q}$. The matrix $\mathbf{P}$ is symmetric (and positive-definite) and the equation is linear in the elements of $\mathbf{P}$. Let us write it as

$$\mathbf{P} \stackrel{\text{def}}{=} \begin{bmatrix} p_1 & p_4 & p_6 \\ p_4 & p_2 & p_5 \\ p_6 & p_5 & p_3 \end{bmatrix} . \tag{1.25}$$

Equations (1.24) can be written as

$$(\mathbf{F}[\mathbf{b}])^T\mathbf{p} = \rho^2\,, \tag{1.26}$$

where $\mathbf{p} = [p_1, p_2, p_3, p_4, p_5, p_6]^T$ and $\mathbf{F}[\mathbf{b}]$ is the vector formed from $\mathbf{b}$ with the elements appropriately computed such that this equation is equivalent to (1.24).

We observe, firstly, that to achieve a full determination of $\mathbf{P}$ (and consequently the maximum possible reduction of the ambiguity), six normals in a general configuration are sufficient. Secondly, as $\mathbf{P}$ can be determined uniquely and there holds that $\mathbf{P} = \mathbf{Q}^T\mathbf{Q}$, it follows that the matrix $\mathbf{Q}$ which converts the original solution to the one consistent with the constraint can be determined from $\mathbf{P}$ only up to an ambiguity represented by a group $O(3)$ (the group of all orthogonal transformations).

For the just discussed case, it was assumed that the absolute value of the albedo ρ was known. Usually this is not the case; the albedo is known up to a scale instead. This, however, does not change the result significantly. The only difference is that in that case the matrix $\mathbf{P}$ is determined only up to a scale. Consequently, the resulting ambiguity is represented by scaled orthogonal transformations $\mathbf{A} = \lambda\mathbf{O}$ ($\mathbf{O} \in O(3), \lambda \neq 0$).

1.5.5 Integrability

Another important possibility is represented by the *integrability constraint* that requires the normals recovered by photometric stereo to correspond to a smooth surface [2, 10]. As shown by Belhumeur et al. [2], in this case the original ambiguity is reduced into ambiguity represented by the group of four-parametric generalized bas-relief transformations. The demonstration of the ambiguity reduction presented here is much like in [44]. We only reveal all the details of the derivation.

The integrability condition is written as

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}\,, \tag{1.27}$$

where p and q are the surface depth gradients (see Section 2.2.2). This is equivalent to

$$\frac{\partial}{\partial y}\left(\frac{b_1}{b_3}\right) = \frac{\partial}{\partial x}\left(\frac{b_2}{b_3}\right)\,, \tag{1.28}$$

where b_1 is the first component of vector $\mathbf{b}$ (and analogically for the other components). Expanding and multiplying by $(b_3)^2$ we get

$$b_3\frac{\partial b_1}{\partial y} - b_1\frac{\partial b_3}{\partial y} = b_3\frac{\partial b_2}{\partial x} - b_2\frac{\partial b_3}{\partial x}\,. \tag{1.29}$$

Not all solutions contained in $\mathcal{T}(\mathbf{J})$ are consistent with this constraint, however. The condition on matrix $\mathbf{Q}$ which converts an arbitrary $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ into an integrable solution is written as

$$\frac{\partial \mathbf{b}^T}{\partial y}\underbrace{(\mathbf{Q}_1^T\mathbf{Q}_3 - \mathbf{Q}_3^T\mathbf{Q}_1)}_{\mathbf{G}}\mathbf{b} = \frac{\partial \mathbf{b}^T}{\partial x}\underbrace{(\mathbf{Q}_2^T\mathbf{Q}_3 - \mathbf{Q}_3^T\mathbf{Q}_2)}_{\mathbf{H}}\mathbf{b}\,, \tag{1.30}$$

where $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$ are the first, second and third rows of $\mathbf{Q}$, respectively. The matrices $\mathbf{G}$ and $\mathbf{H}$ are both antisymmetric. Writing $\mathbf{G}$ explicitly gives

$$\mathbf{G} = \begin{bmatrix} 0 & q_{11}q_{32} - q_{31}q_{12} & q_{11}q_{33} - q_{31}q_{13} \\ q_{12}q_{31} - q_{32}q_{11} & 0 & q_{12}q_{33} - q_{32}q_{13} \\ q_{13}q_{31} - q_{33}q_{11} & q_{13}q_{32} - q_{33}q_{12} & 0 \end{bmatrix} \tag{1.31}$$

$$\stackrel{\text{def}}{=} \begin{bmatrix} 0 & -g_1 & g_3 \\ g_1 & 0 & -g_2 \\ -g_3 & g_2 & 0 \end{bmatrix}; \tag{1.32}$$

the matrix $\mathbf{H}$ is written analogously. In terms of the unknown matrices $\mathbf{G}$ and $\mathbf{H}$, the system of equations (1.30) is linear and homogeneous. The number of unknowns is six, and hence a minimum number of equations to establish the constraint is five.

Let us observe how the integrability constraint reduces the photometric stereo ambiguity. There holds

$$\mathbf{Q}^{-T} = \frac{1}{\det \mathbf{Q}} \begin{bmatrix} \mathbf{Q}_2 \times \mathbf{Q}_3 \\ \mathbf{Q}_3 \times \mathbf{Q}_1 \\ \mathbf{Q}_1 \times \mathbf{Q}_2 \end{bmatrix} \tag{1.33}$$

$$= \frac{1}{\det \mathbf{Q}} \begin{bmatrix} q_{22}q_{33} - q_{23}q_{32} & q_{23}q_{31} - q_{21}q_{33} & q_{21}q_{32} - q_{22}q_{31} \\ q_{32}q_{13} - q_{33}q_{12} & q_{33}q_{11} - q_{31}q_{13} & q_{31}q_{12} - q_{32}q_{11} \\ q_{12}q_{23} - q_{13}q_{22} & q_{13}q_{21} - q_{11}q_{23} & q_{11}q_{22} - q_{12}q_{21} \end{bmatrix}. \tag{1.34}$$

We observe that

$$\mathbf{Q}^{-T} = \frac{1}{\det \mathbf{Q}} \begin{bmatrix} -h_2 & -h_3 & -h_1 \\ g_2 & g_3 & g_1 \\ ? & ? & ? \end{bmatrix}. \tag{1.35}$$

To find a $\mathbf{Q}$ which ensures that the transformed normals are integrable, the three unspecified elements in (1.35) are filled by arbitrary values (with the constraint, however, that the resulting matrix has to be invertible), and then $\mathbf{Q}$ is computed as the transposed inverse of the matrix in (1.35).

To answer the question of what ambiguity reduction is achieved by the application of the integrability constraint is equivalent to asking what are the only transformations $\mathbf{A}$ that can be applied to the normals transformed by $\mathbf{Q}$ (hence integrable ones) such that the integrability constraint continues to hold, i.e., the normals $\mathbf{AQB}$ are integrable. While $\mathbf{Q}$ is transformed to $\mathbf{AQ}$, $\mathbf{Q}^{-T}$ is transformed to $\mathbf{A}^{-T}\mathbf{Q}^{-T}$. The only transformations $\mathbf{A}^{-T}$ which leave the two rows of the matrix $\mathbf{Q}^{-T}$ in (1.35) unchanged are

$$\mathbf{A}^{-T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & \beta & \gamma \end{bmatrix}, \qquad \begin{array}{l} \gamma \neq 0, \\ \alpha, \beta \in \mathbb{R}, \end{array} \tag{1.36}$$

where γ has to be nonzero for the matrix to be invertible. If $\mathbf{A}^{-T}$ is as above then $\mathbf{A}$ has to be

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -\frac{\alpha}{\gamma} \\ 0 & 1 & -\frac{\beta}{\gamma} \\ 0 & 0 & \frac{1}{\gamma} \end{bmatrix}, \qquad \begin{array}{l} \gamma \neq 0, \\ \alpha, \beta \in \mathbb{R}. \end{array} \tag{1.37}$$

The final fact that has to be taken into account is that the matrices $\mathbf{G}$ and $\mathbf{H}$ are determined from (1.30) only up to a scale. The final result, then, is that the original ambiguity reduces to a group of generalized bas-relief transformations (GBR) [2] which take the form:

$$\mathbf{A} = \begin{bmatrix} \lambda & 0 & \mu \\ 0 & \lambda & \nu \\ 0 & 0 & \tau \end{bmatrix}, \qquad \begin{array}{l} \lambda, \tau \neq 0\,, \\ \mu, \nu \in \mathbb{R}\,. \end{array} \tag{1.38}$$

1.5.6 Occluding boundary

The occluding boundary is the set of points on a surface the normals of which are perpendicular to the viewing direction. Let us assume that the normals are parametrized in a 'natural' coordinate system (first two basis vectors span the image plane and the third one is the direction of viewing). Then the normals at the occluding boundary have the third coordinate zero. In addition, since the occluding boundary is visible in an image, the orthographic projections of normals at the occluding boundary are determined as well. This is because the projections of normals onto the image plane are perpendicular to the projection of the occluding boundary itself. This is equivalent to knowing the true first two coordinates of $\mathbf{b} \in \mathcal{U}$ up to a scale.

To express this constraint, let us first denote by $\mathcal{U}$ the set of normals on an occluding boundary. Algebraically, we require that

$$\forall \mathbf{b} \in \mathcal{U} \;\; \exists \lambda_{\mathbf{b}}\text{:} \quad \mathbf{Q}_i \mathbf{b} = \lambda_{\mathbf{b}} {b_i}^*\,, \qquad i = \{1, 2\}\,, \tag{1.39}$$

$$\mathbf{Q}_3 \mathbf{b} = 0\,, \tag{1.40}$$

where b_i^* are the (known) projections of the normals. The system of equations (1.39) is linear in the unknowns $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\lambda_{\mathbf{b}}$'s and hence the first two rows of $\mathbf{Q}$ can be determined up to a scale. The system of equations (1.40) fixes the direction of $\mathbf{Q}_3$ (but not its magnitude). Taken together, the only transformations that leave the occlusion boundary condition invariant are:

$$\mathbf{A} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \tau \end{bmatrix}, \qquad \lambda, \tau \neq 0\,, \tag{1.41}$$

which is the *bas-relief* ambiguity.

1.5.7 Shadows

Shadows in images are of two types: self-shadows which appear at places where the surface normals make an obtuse angle with the light direction, and cast shadows which appear when an occluder exists on a path from the surface point to the light source. In contrast to the assumptions and phenomena described above, the shadows do not represent any additional information which could be used to reduce the photometric stereo ambiguity [2].

1.6 Goals of the thesis

We aim to make a step towards answering the following question: If the object reflectance is *not* Lambertian, what is the ambiguity up to which the surface parameters can be recovered by uncalibrated photometric stereo?

This question is especially challenging due to the fact that many real surfaces strongly deviate from being Lambertian. The variability in surface reflectance is enormous, but our intention is not to analyse each reflectance model individually. Instead, our goal is to identify fundamental properties of reflectance which are shared over large classes of reflectance functions. These properties include symmetries (e.g., the BRDF reciprocity or isotropy) and the presence of both body and interface reflectance components which behave differently in a number of ways. We analyse whether a given property of reflectance can be used to construct new constraints on surface or illumination parameters which could be used in uncalibrated photometric stereo to reduce the ambiguity of surface parameters determination. Such approach has the advantage that when some result follows from a given reflectance property, it is valid for a large class of reflectance models. If, for example, we were able to get the answer for isotropic reflectance functions (i.e., those which produce the same image intensity under both the light sources and the detector rotate around the surface normal direction), the result would hold generally irrespective of whether the real reflectance would be Lambertian, mirror-like or other.

The goals of this thesis are the following. Under the standard assumptions defined in Section 1.3,

1. To review the properties of reflectance functions in a systematic way. Subsequently, to identify which of these properties enable to gain additional information on surface or illumination parameters that could be used in uncalibrated photometric stereo.

2. For the possibilities identified above, to analyse theoretically the ambiguity up to which the surface parameters can be determined from the photometric stereo method if the illumination directions are not known. To conduct experiments verifying the theoretical results.

We adopt the following strategy for analysing uncalibrated photometric stereo for non-Lambertian surfaces. We assume that the surface reflectance can be sufficiently well approximated by the Lambertian one for the majority of illumination and viewing geometries. Examples of such surfaces are shown in Fig. 1.1(a): Their surface reflectance is close to Lambertian except for bright spots corresponding to specular reflection [39]. We represent the surface properties by the normal and albedo maps, as if the object was Lambertian. Subsequently, any constraint which follows from newly identified constraints implied by a given BRDF property gives us a valuable possibility to approach the analysis of uncalibrated photometric stereo for non-Lambertian surfaces as an uncalibrated Lambertian photometric stereo with additional information.

1.7 Main results

The main result of this thesis is that we found two fundamental reflectance properties which can be exploited within the uncalibrated photometric stereo in order to deliver additional information about surface or illumination parameters. Specifically, these are the following:

1. **Presence of specular reflectance component**. If object reflectance is a sum of Lambertian reflectance and an ideal specular (mirror-like reflectance), then the original ambiguity represented by a group $GL(3)$ reduces into a two-parametric group of transformations. These transformations are compositions of isotropic scaling (1dof), rotation around the viewing vector (1dof), and change in the global coordinate frame handedness (binary ambiguity). This ambiguity reduction is implied by a condition that all lights reflected around corresponding specular normals must give the same vector (the viewing direction). We call this condition the *consistent viewpoint constraint*. We showed that specularities in as few as *four* images corresponding to four different distant point lights in general configuration are sufficient to utilize the consistent viewpoint constraint [9].

2. **Separable isotropic BRDFs.** Such BRDFs can be expressed as a product of two terms, one of which is dependent on the incidence angle and the other is dependent on the reflected angle. We show that separability enables to construct two geometrical constraints. The constraints are represented by projections of normals onto planes perpendicular to the viewing and light vectors, respectively. We constructed the constraints using polarisation measurement under the assumption of separable isotropic reflectance model for smooth dielectrics. We showed that the two constraints alone combined reduce the ambiguity to convex/concave ambiguity composed with isotropic scaling [8].

1.8 Thesis outline

Each of the chapters is concluded by a section entitled 'Summary of the chapter' which is the excerpt of that chapter.

- Chapter 2 (**Basic concepts**) reviews the geometry and radiometry of a camera, radiometric terms and reflectance models. It also reviews the relationship between the depth and normal maps and defines what we understand by tilts of normals.

- Chapter 3 (**BRDF properties**) reviews the properties of reflectance functions in a systematic way, and observes which properties could be exploited in uncalibrated photometric stereo to deliver algebraic constraints on surface or illumination parameters.

- Chapter 4 (**Separable isotropic BRDFs**) shows the results described in Section 1.7 for Main Result 2.

- Chapter 5 (**Presence of specular reflectance component**) shows the results described in Section 1.7 for Main Result 1.
- Chapter 6 (**Summary & Conclusions**) summarises the thesis and lists open problems for future research.

Chapter 2

Basic concepts

This chapter reviews the concepts which are essential for understanding the other parts of the thesis. A reader with a fair background in computer vision needs to read only the following parts of this Chapter: Section 2.3 where the *tilt of a normal* is defined, the end of Section 2.5 where we describe what is meant by the term *image intensity*, and the definition of *light strength* in Section 2.4. All the other concepts are standard. Section 2.1 reviews how points in space are mapped onto the image plane. Section 2.2 reviews the concepts of depth and normal maps, and shows the relationship between them. Section 2.4 reviews the basic radiometric terms. Section 2.5 reviews the relationship between the values obtained from an electronic camera and observed surface radiances. Section 2.6 reviews surface reflectance and Section 2.7 briefly reviews the reflectance models.

2.1 Geometry of imaging

An ordinary camera consists of a lens and an array of light-sensitive sensors (see Fig. 2.6). By letting the lens aperture radius to go to zero we get a pinhole camera model which performs the central projection of points in space into the image plane [15, 11]. Such camera model is called *perspective*. Let $\tilde{\mathbf{X}} \in \mathbb{P}^3$ be the homogeneous vector representing a point $\mathbf{X} \in \mathbb{R}^3$ in the Euclidean space and let $\tilde{\mathbf{u}} \in \mathbb{P}^2$ denote the homogeneous vector representing the projection $\mathbf{u} \in \mathcal{I}$ of $\mathbf{X}$ onto an image plane $\mathcal{I} \subset \mathbb{R}^2$. For $\tilde{\mathbf{u}}$ there holds

$$\tilde{\mathbf{u}} = \mathbf{P}\tilde{\mathbf{X}}\,, \qquad \mathbf{P} = [\mathbf{M}|\mathbf{a}],\ \ \mathbf{M} \in GL(3),\ \ \mathbf{a} \in \mathbb{R}^{3\times 1} \tag{2.1}$$

where $\mathbf{P}$ is the *camera matrix*. The 'anatomy' of $\mathbf{P}$ (its relationship with the center of projection and other projective camera parameters) is well understood. There is a limiting case of a perspective camera which is obtained when the center of projection is moved to infinity. For *cameras at infinity* (the affine cameras, see [15], Section 5.3.1) the last row of $\mathbf{M}$ is a zero vector and $\mathbf{a} = (t_1, t_2, 1)^T$:

$$\mathbf{P}_{\mathrm{A}} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & t_1 \\ m_{21} & m_{22} & m_{23} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}\,, \tag{2.2}$$

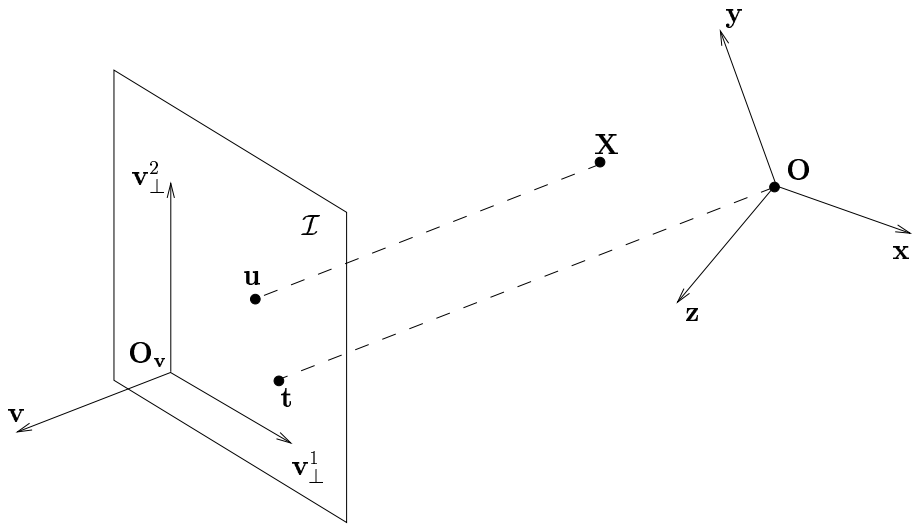


Figure 2.1: Orthographic camera. A point $\mathbf{X}$ in space is mapped to the point $\mathbf{u}$ in an image plane by a ray parallel with the viewing direction $\mathbf{v} = \mathbf{v}_\perp^1 \times \mathbf{v}_\perp^2$. Point $\mathbf{t}$ is the projection of a global coordinate frame origin $\mathbf{O}$. The image basis vectors $\mathbf{v}_\perp^1$ and $\mathbf{v}_\perp^2$ are orthonormal.

where $\mathbf{t} = (t_1, t_2)^T$ are the coordinates of the projection of the global coordinate frame origin. There is a restriction that the first two rows of $\mathbf{M}$ above have to have rank 2. The mapping (2.1) can be written in a compact form for inhomogeneous vectors $\mathbf{X} \in \mathbb{R}^3$ and $\mathbf{u} \in \mathcal{I}$:

$$\mathbf{u} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{bmatrix} \mathbf{X} + \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \tag{2.3}$$

which can be rewritten as

$$\mathbf{u} = \mathbf{M}_{2\times 3}\mathbf{X} + \mathbf{t}\,. \tag{2.4}$$

In photometric stereo, a specialisation of a general affine camera is of interest: an orthographic camera (see Fig. 2.1). This is an affine camera with $\mathbf{M}_{2\times 3}$ formed by an orthonormal pair $(\mathbf{v}_\perp^1, \mathbf{v}_\perp^2)$. For orthographic camera (2.4) becomes

$$\mathbf{u} = \begin{bmatrix} \mathbf{v}_\perp^{1\,T} \\ \mathbf{v}_\perp^{2\,T} \end{bmatrix} \mathbf{X} + \mathbf{t}\,. \tag{2.5}$$

2.2 Depth and normal maps

Let us consider a 3D object surface viewed by an orthographic camera and let the global coordinate frame coincide with the camera coordinate frame (see Fig. 2.2). The *depth map* is a mapping $Z{:}\mathcal{I} \mapsto \mathbb{R}^1 \cup \{\emptyset\}$ which for each image point $\mathbf{u} = (x, y)^T$

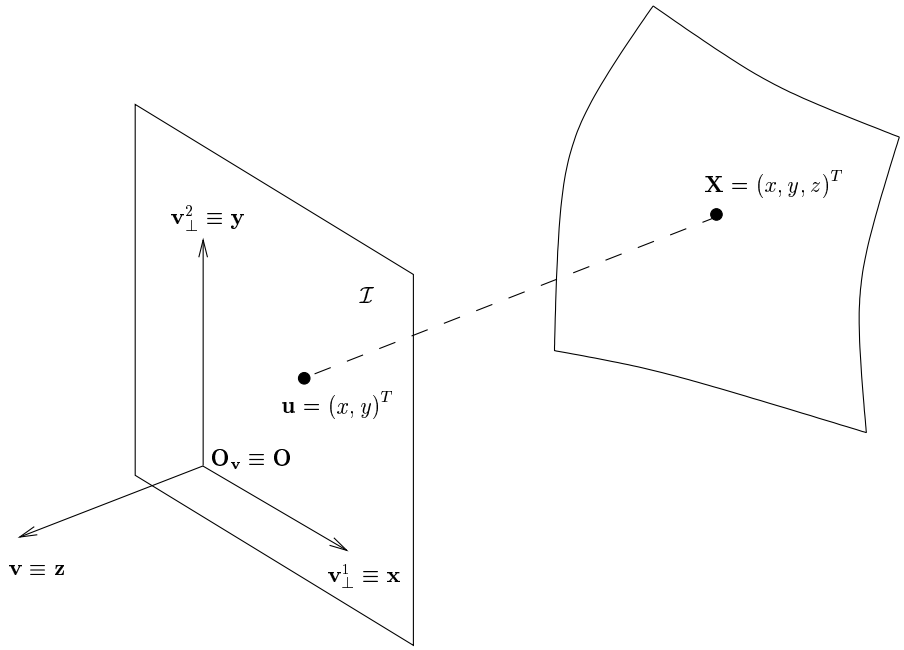


Figure 2.2: A surface viewed by an orthographic camera.

stores the third coordinate of the surface point $(x, y, z)^T$ which is projected to $\mathbf{u}$ if such point exists, and if a point projecting to $\mathbf{u}$ does not exist then $Z(\mathbf{u}) = \{\emptyset\}$.

If the depth map has a gradient at $(x, y)^T$ then the surface normal $\mathbf{n}$ at a point $(x, y, Z(x, y))^T$ can be computed as follows. The surface is defined by the zero level set of a function g

$$g(x, y, z) = z - Z(x, y) = 0\,. \tag{2.6}$$

The normal direction to this surface is given by the gradient of g at its zero level set:

$$\mathbf{n}(x, y) \simeq \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right)\Bigg|_{g=0} = \left(-\frac{\partial Z}{\partial x}, -\frac{\partial Z}{\partial y}, 1\right)\Bigg|_{(x,y)} = (-p, -q, 1)|_{(x,y)}\,, \tag{2.7}$$

where $p = \frac{\partial Z}{\partial x}$ and $q = \frac{\partial Z}{\partial y}$ are the components of the surface height gradient. After normalisation,

$$\mathbf{n} = \frac{(-p, -q, 1)}{\sqrt{p^2 + q^2 + 1}}\,. \tag{2.8}$$

The *normal map* is the mapping $\mathbf{N}{:}\,\mathcal{I} \mapsto \mathcal{S}_\mathbf{v}^{1/2} \cup \{\emptyset\}$ where $\mathcal{S}_\mathbf{v}^{1/2}$ is the set of all visible normals $\mathcal{S}_\mathbf{v}^{1/2} = \{\mathbf{n}{:}\,\|\mathbf{n}\| = 1,\ \ \mathbf{n} \cdot \mathbf{v} \geq 0\}$. For each image point $(x, y)^T$, the mapping stores the normal $\mathbf{N}(x, y)$ at the surface point $(x, y, Z(x, y))^T$ if it exists, and if it does not exist (i.e., either the depth map is not smooth or there is no visible surface for that image point) then $\mathbf{N}(x, y) = \{\emptyset\}$.

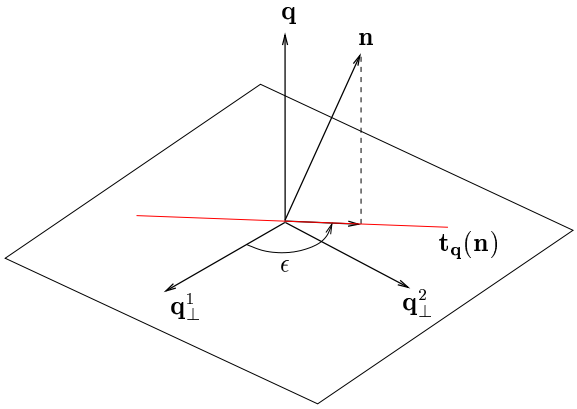


Figure 2.3: Definition of the **q**-tilt of the normal **n**. The tilt $\mathbf{t}^{\mathbf{q}}(\mathbf{n}) \in \mathbb{P}^1$ corresponds to the direction of the red line drawn in the figure and can be represented by an angle $\epsilon \in [0, \pi)$.

2.2.1 Relationship between the depth and normal maps

It has been demonstrated above how the normal map can be obtained from the depth map. It turns out that under specific conditions the depth map can be reconstructed from the normal map as well.

Let $\Gamma \subset \mathcal{I}$ be a region within the image plane $\mathcal{I}$ and let the normal map $\mathbf{n}(x, y)$ is defined everywhere in this region. If the true depth map $Z(x, y)$ is free of discontinuities in this region and if it is known at some point $(x_0, y_0)^T \in \Gamma$ then there holds

$$\forall (x, y)^T \in \Gamma: \quad Z(x, y) = Z(x_0, y_0) + \int\limits_{C(x_0, y_0; x, y)} p\mathrm{d}x + q\mathrm{d}y\,, \tag{2.9}$$

where $C(x_0, y_0; x, y) \subset \Gamma$ is a curve connecting $(x_0, y_0)^T$ and $(x, y)^T$. The surface gradients p, q are computed from normals simply as

$$p = -\frac{n_1}{n_3}\,, \qquad q = -\frac{n_2}{n_3}\,. \tag{2.10}$$

2.2.2 Integrability

Let $\Gamma \subset \mathcal{I}$ be a region as above. There is no condition on the normal map which would be sufficient for the continuity of the true depth map in Γ. But there is a necessary condition on the normal map if the depth map Z has continuous second derivatives in Γ. For such depth maps there holds

$$\forall (x, y)^T \in \Gamma: \qquad \frac{\partial^2 Z}{\partial x \partial y} = \frac{\partial^2 Z}{\partial y \partial x}\,. \tag{2.11}$$

This defines the constraint on surface gradients:

$$\forall (x, y)^T \in \Gamma: \qquad \frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}\,, \tag{2.12}$$

or rewritten using the normal components as in (2.10)

$$\forall (x,y)^T \in \Gamma: \qquad \frac{\partial}{\partial y}\left(\frac{n_1}{n_3}\right) = \frac{\partial}{\partial x}\left(\frac{n_2}{n_3}\right). \tag{2.13}$$

2.3 Tilts of normals

Let $[\mathbf{q}_\perp^1|\mathbf{q}_\perp^2|\mathbf{q}]$ be an orthonormal basis. Let $\mathbf{n}$ be a normal which is not collinear with $\mathbf{q}$. We define the $\mathbf{q}$-tilt of normal $\mathbf{n}$ to be a homogeneous vector $\mathbf{t}^{\mathbf{q}}(\mathbf{n}) \in \mathbb{P}^1$:

$$\mathbf{t}^{\mathbf{q}}(\mathbf{n}) = [\mathbf{q}_\perp^1|\mathbf{q}_\perp^2]^T\mathbf{n}\,. \tag{2.14}$$

The $\mathbf{q}$-tilt thus represents the direction given by the projection of $\mathbf{n}$ onto $[\mathbf{q}_\perp^1, \mathbf{q}_\perp^2]$ (see Fig. 2.3 where the direction is shown in red). The information carried by $\mathbf{t}^{\mathbf{q}}(\mathbf{n})$ is equivalent to knowing the angle $\epsilon \in [0, \pi)$ which the line makes with $\mathbf{q}_\perp^1$.

2.4 Radiometric terms

This section reviews the basic terms of radiometry. Table 2.1 gives an overview of the terms together with their symbols, relations used to define them, and units. The individual terms are described in a more detail afterwards.

Entity	Symbol	Unit
Radiant flux	Φ	W
Radiant intensity	$I_e = \mathrm{d}\Phi/\mathrm{d}\Omega$	$\mathrm{W\,sr^{-1}}$
Light strength	$\sigma = I_e/r^2$	$\mathrm{W\,sr^{-1}\,m^{-2}}$
Irradiance	$E = \mathrm{d}\Phi/\mathrm{d}S$	$\mathrm{W\,m^{-2}}$
Radiance	$L_r = \mathrm{d}^2\Phi/(\mathrm{d}S\cos\theta_r\mathrm{d}\Omega_r)$	$\mathrm{W\,sr^{-1}\,m^{-2}}$
Incident radiance	$L_i = \mathrm{d}^2\Phi/(\mathrm{d}S\cos\theta_i\mathrm{d}\Omega_i)$	$\mathrm{W\,sr^{-1}\,m^{-2}}$

Table 2.1: The radiometric terms.

Radiant flux Φ is a power propagated as optical electromagnetic radiation.

Radiant intensity $I_e(\mathbf{e}) = \mathrm{d}\Phi/\mathrm{d}\Omega$ of a point light source in a direction $\mathbf{e}$ is the radiant flux $\mathrm{d}\Phi$ emerging into a solid angle $\mathrm{d}\Omega$ centered around $\mathbf{e}$ (see Fig. 2.4), normalised per unit solid angle.

Irradiance $E = \mathrm{d}\Phi/\mathrm{d}S$ is the radiant flux $\mathrm{d}\Phi$ incident on a surface $\mathrm{d}S$ (see Fig. 2.4), normalised per unit surface area.

Light strength at position $\mathbf{X}$ resulting from a point light source of intensity I_e in position $\mathbf{L}$ is $\sigma = I_e(\mathbf{e})/r^2$ where $I_e(\mathbf{e})$ is the radiant intensity of the source in a direction of $\mathbf{X}$ and $r = \|\mathbf{X} - \mathbf{L}\|$ is the distance between the point of light strength evaluation and the light source (see Fig. 2.4). This entity tells us what would be the irradiance of an imaginary surface patch oriented perpendicularly to $\mathbf{e}$ and placed at $\mathbf{X}$.

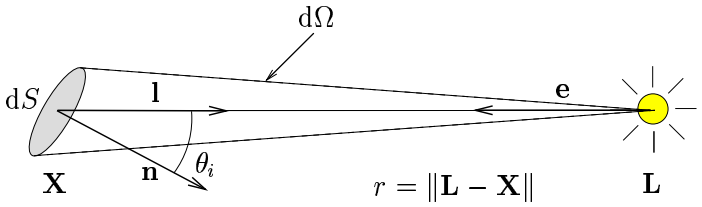


Figure 2.4: An elementary surface patch $\mathrm{d}S$ irradiated by a point light source. This illustration is used for explaining the majority of radiometric terms.

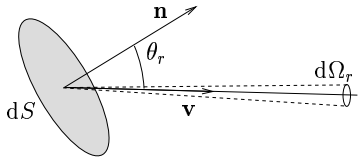


Figure 2.5: Illustrating the definition of radiance.

Radiance $L_r(\mathbf{v}) = \mathrm{d}^2\Phi/(\mathrm{d}S\cos\theta_r\mathrm{d}\Omega_r)$ in a direction $\mathbf{v}$ is the radiant flux $\mathrm{d}^2\Phi$ emerging from an area $\mathrm{d}S$ into solid angle $\mathrm{d}\Omega_r$ centered around $\mathbf{v}$, normalised to unit solid angle and unit foreshortened area.

Incident radiance $L_i(\mathbf{l}) = \mathrm{d}^2\Phi/(\mathrm{d}S\cos\theta_i\mathrm{d}\Omega_i)$ is defined analogously to the above described radiance, with the only difference that it involves the incident (rather than emerging) flux.

2.5 Radiometry of imaging

Consider the situation in Fig. 2.6. A pixel projects onto an object surface, defining a surface patch whose radiance affects the radiant flux incident on this pixel.

Suppose that such surface is considered planar and that the radiance is constant across the patch, that the camera is distant so that the solid angle Ω_r given by the lens aperture viewed from any point within the patch is very small, and that the solid angle Ω_P given by the pixel extent viewed from the lens center is small as well (the situation depicted in the figure does not obey these assumptions for the sake of readability). By definition of radiance, the radiant flux Φ incident on the lens aperture is well approximated by

$$\Phi = L_r\Omega_r S\cos\theta_r\,, \tag{2.15}$$

where L_r is the radiance in a direction of the camera and $S\cos\theta_r$ is the foreshortened

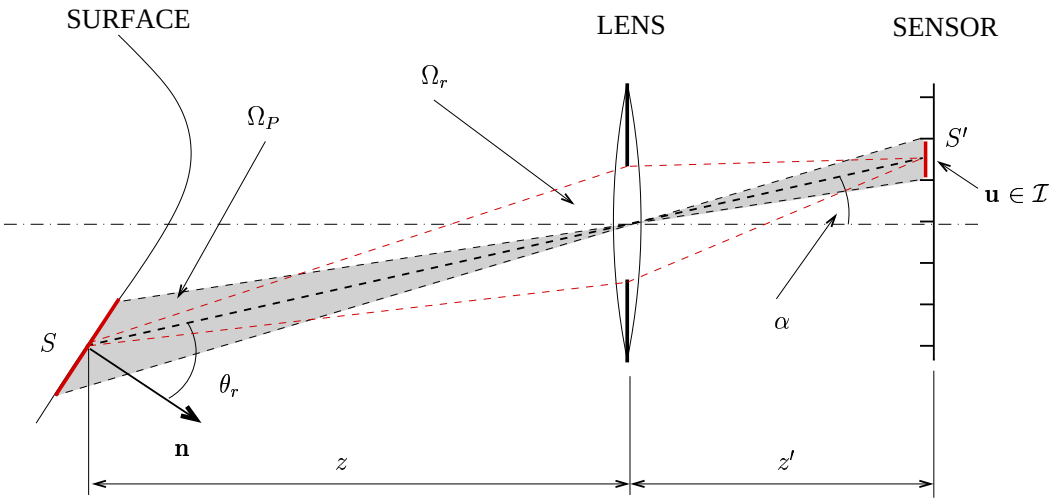


Figure 2.6: A camera.

area of the patch. Let us expand (2.15). For Ω_r there holds

$$\Omega_r = \frac{\pi \frac{D^2}{4} \cos\alpha}{\left(\frac{z}{\cos\alpha}\right)^2}\,, \tag{2.16}$$

where D is the lens aperture diameter and α is the angle which the imaging beam makes with the optical axis. Next, it holds

$$\frac{S\cos\theta_r}{S'\cos\alpha} = \left(\frac{z}{z'}\right)^2\,, \tag{2.17}$$

where z' is the distance of the pixel array from the lens centre, and $S'\cos\alpha$ is the foreshortened area of a pixel. Putting these two results to (2.15) gives

$$\Phi = L_r \frac{\pi}{4}\left(\frac{D}{z'}\right)^2 S' \cos^4\alpha\,. \tag{2.18}$$

Finally, the irradiation of the pixel $\mathbf{u}$ is

$$\boxed{E(\mathbf{u}) = L_r(\mathbf{u})\frac{\pi}{4}\left(\frac{D}{z'}\right)^2 \cos^4\alpha(\mathbf{u})\,,} \tag{2.19}$$

which is a well-known result [34, 12]. In this equation, we explicitly indicated the dependency of L_r and α on the pixel position $\mathbf{u}$. The irradiance of a pixel is proportional to the observed surface patch radiance, but the factor of proportionality is dependent on the pixel position $\mathbf{u}$ (as α is dependent on a pixel position). This dependence is known as *vignetting*.

The irradiance $E(\mathbf{u})$ of a pixel $\mathbf{u}$ is converted using the electronic circuits of a camera to a *raw image intensity* $I_{raw}(\mathbf{u})$. The way this is done corresponds to a

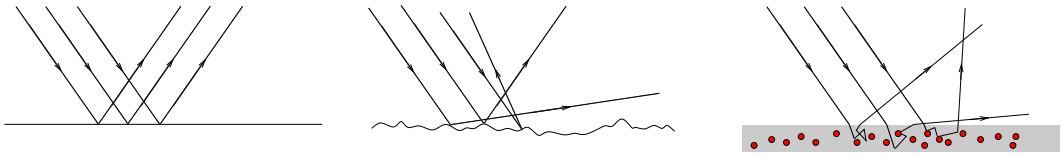


Figure 2.7: Interaction of light with surfaces: (a) a pure reflection on a smooth interface, (b) reflection/diffraction on a rough surface, (c) subsurface scattering. Cases (a) and (b) are the examples of the *interface reflection*. The subsurface scattering in (c) is the essence of the *body reflection*.

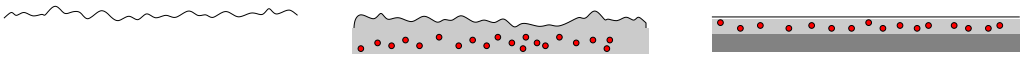


Figure 2.8: Three examples of object surfaces: (a) a rough metal surface accommodating the interface reflection only, (b) a rough dielectric with inhomogeneities inside which cause the light scattering (e.g. colored plastic, rubber), (c) glazed ceramics with a smooth interface and two subsurface layers.

mapping $I_{raw}(\mathbf{u}) = f(E(\mathbf{u}))$. The mapping f is usually monotonic and close to linear.

Both the vignetting effect and the monotonic mapping of pixel irradiance to raw image intensity can be corrected by radiometric calibration of the camera, such that the result of this correction which we call the *image intensity* $I(\mathbf{u})$ is directly proportional to the radiance of a patch which the pixel observes:

$$\boxed{I(\mathbf{u}) = V^{-1}(f^{-1}(I_{raw}(\mathbf{u}))) = \eta L_r(\mathbf{u})\,,} \tag{2.20}$$

where V^{-1} corrects vignetting and f^{-1} transform the raw image intensities back to pixel irradiance values. It is relatively easy to compensate for both effects up to a scale, such that (2.20) holds but η is unknown. Only when the camera is fully radiometrically calibrated such that it is able to give absolute measurements of radiance, the factor η is known.

2.6 Reflectance characteristics of surfaces

In Section 2.4 we observed that image irradiance is proportional to the observed surface radiance. Let us now focus on how surfaces interact with light, and consequently, what the surface radiance is.

The reflectance on a real surface is often advantageously regarded as composed of two components: the *interface* and *body* ones. The interface component follows from interaction of light with the air-object boundary. It is responsible for the visual

appearance of metals, whenever polished ones (see Fig. 2.7(a)) exhibiting specular reflectance, or roughened ones (see Fig. 2.7(b)) for which more diffuse reflectance is observed. The body component (see Fig. 2.7(c)) is caused by light which returns from the inside of the object, after being scattered by local inhomogeneities in the object body material.

This partitioning was proposed in the computer vision community by Shafer [32]. Such decomposition is natural since the reflection processes governing the two components 'happen' at two different loci: the interface component is governed by the air-surface boundary properties, while the body component results from the light penetrating to the subsurface, scattering there, and then coming back to air. It follows that the two components also behave differently; while the spectral profile of the interface component is often effectively unchanged by the reflection process, the body component inherits the color of particles in a subsurface. In addition, the body component tends to be diffuse, while the interface component tends to be directional, centered around the specular direction (the direction of light as reflected by a mirror, cf. Fig. 2.7(a)).

Note that the images shown in Fig. 2.7 are just schematic illustrations. This is most obvious in the figure (c) where the process of scattering is depicted. It is illustrated by a line that the light is multiply scattered before going back into the air. However, by no way this means that the effect of scattering is just to force the ray to go in a different direction. On the contrary, the scattering distributes the radiant flux incident from one direction to multiple directions. Similar comments apply to figures (a) and (b). These two images differ just by the shape of the optical interface. In case (a), it is smooth; more precisely, that means that the incident light wavelength is much larger than the irregularities of the object surface. For the case (b), if the size of the height function modulation is less then or comparable to the wavelength then diffraction takes place. It is only the case when the undulation size is much larger then the wavelength when the situation precisely corresponds to the image depicted, since then the process of interaction can be regarded as a reflection (with the local angle of incidence equal to the local angle of reflection).

To illustrate the concept of reflectance decomposition into the interface and body components, Fig. 2.8 shows the examples of three different surfaces. The first example (a) displays a rough surface of opaque body; that could be a rough metal. The second example (b) shows the structure of a dielectric with inhomogeneities; all sorts of colored plastics or rubber have this structure. The color of such materials can be set by spectral behavior of the inhomogeneities. The third example (c) corresponds to a glazed ceramics; the interface reflectance is specular, and the body component is again caused by the light refracting into the body, scattering there, and then refracting back. The interval between the two refractions may include the interaction with the second layer. This example is included just to demonstrate the fact that there are no limitations to the subsurface; it can be composed of several layers (here the layer of colored glazing, and the clay layer), or highly structured.

2.7 Reflectance models

In this section we review the models which are appropriate in scope of this thesis only. For an excellent review the reader is referred to a brief document by Mungan [26].

This brief review is made in the perspective of reflectance decomposition into the *body* and *interface* components. For the body component of BRDF, the oldest model is the Lambert's one [23] which is constant and accommodates a single reflectance parameter called *albedo*. Note that according to (1.2) the constancy of $f_{i,r}$ implies that the observed radiance is independent on the viewing direction. The albedo parameter denoted ρ characterizes how much light is reflected back from the surface. Precisely speaking, if Φ_i and Φ_r are the total radiant flux incident on and reflected from a surface patch of area S, respectively, then

$$\rho = \frac{\Phi_r}{\Phi_i} = \frac{\int L_r \cos\theta_r \mathrm{d}\Omega_r}{\int L_i \cos\theta_i \mathrm{d}\Omega_i}\,, \tag{2.21}$$

where L_r is radiance, L_i is incident radiance and integration is performed over the entire upper hemisphere. If $f_{i,r}$ is constant then according to (1.2) $L_r = f_{i,r} \int L_i(\theta_i, \phi_i) \cos\theta_i \mathrm{d}\Omega_i$; hence the above equation becomes

$$\rho = f_{i,r} \int \cos\theta_r \mathrm{d}\Omega_r = \pi f_{i,r}\,, \tag{2.22}$$

or, equivalently, the Lambertian BRDF is

$$f_r = \frac{\rho}{\pi}\,. \tag{2.23}$$

According to (1.11) the observed image intensity for a point characterised by the scaled normal $\mathbf{b}$ and illuminated by a light source represented by the scaled light vector $\mathbf{s}$ is linear in both $\mathbf{b}$ and $\mathbf{s}$. This is one of the reasons for the Lambertian model popularity. Besides that, many real matte surfaces can be well approximated by the Lambertian model. In addition, the Lambertian model can be also regarded as a first term in a series used for representation of the BRDF [21, 29] and can then serve as a first-order approximation even to non-Lambertian models of reflectance. For smooth dielectrics, Wolff [39] constructed a physically correct reflectance model by simply adding two Fresnel transmission coefficients to the Lambertian term (one for rays coming from the light source to the object body, and the other for rays refracting from the object body towards the viewer) and proved applicability of the model to smooth ceramics, plastics and rubber.

For the interface reflection, the simplest model is the ideal specular reflectance given by

$$f_{i,r} = \frac{\delta(\theta_i - \theta_r)\delta(\phi_r - \phi_i - \pi)}{\sin\theta_i \cos\theta_i}\,; \tag{2.24}$$

that states that the reflected direction equals the incident direction rotated around the normal by 180° (this corresponds to the mirror-like reflection). The factor $\sin\theta_i \cos\theta_i$ in the denominator is implied by (1.2), requiring that $L_r(\theta_r, \phi_r) = L_i(\theta_i, \phi_i + \pi)$ (see for example [17]).

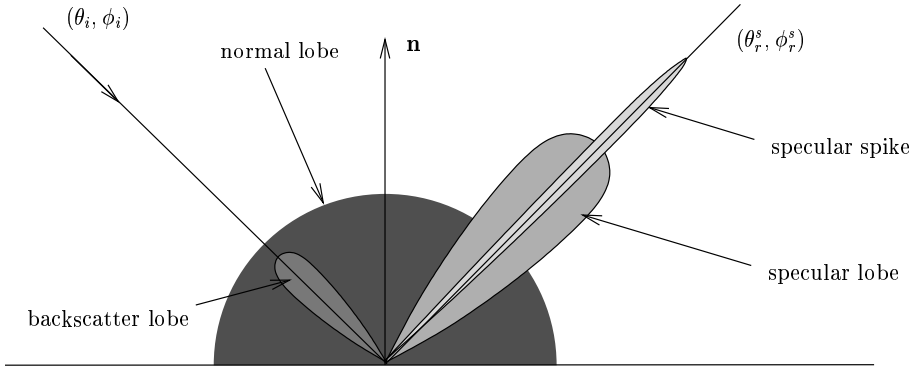


Figure 2.9: The illustration of the reflectance decomposition. The figure shows the graph of dependence of radiance on the reflected direction for a given incident direction. The direction (θ_r^s, ϕ_r^s) represents the specular direction given by mirror-like reflection.

The surface reflectance is often modeled as a superposition of individual lobes (see for example [35]). Figure 2.9 depicts an example. The normal lobe corresponds to the body component modeled by the Lambertian model. Specular spike corresponds to mirror-like reflection in a specular direction (θ_r^s, ϕ_r^s). Specular lobe is spread around the specular direction and corresponds to interface reflectance of a rough surface. The backscatter lobe is spread around the incident direction. It is usually negligible except for, e.g., metals or paints.

2.8 Summary of the chapter

This chapter reviewed the following facts:

- The image formation process consists of two aspects: the geometric one (2.5) which describes how a point is space is mapped onto the image plane, and the radiometric one (2.19) which explains what is the irradiation at the projected point.

- A value observed at a pixel is proportional to the radiance of a surface patch the given pixel observes. If the camera is radiometrically corrected then the factor of proportionality η is independent on the pixel position in an image (2.20). By *image intensity* we understand the corrected values.

- For $[\mathbf{q}_\perp^1|\mathbf{q}_\perp^2|\mathbf{q}]$ being an orthonormal basis, the $\mathbf{q}$-tilt of a normal $\mathbf{n}$ was defined as a homogeneous vector $\mathbf{t}^{\mathbf{q}}(\mathbf{n})$ given by the projection of $\mathbf{n}$ onto the plane given by vectors $\mathbf{q}_\perp^1$ and $\mathbf{q}_\perp^2$ (2.14).

- The light strength at position $\mathbf{X}$ resulting from a point light source of intensity I_e in position $\mathbf{L}$ was defined to be $\sigma = I_e(\mathbf{e})/r^2$ where $I_e(\mathbf{e})$ is the radiant intensity of the source in a direction of $\mathbf{X}$ and $r = \|\mathbf{X} - \mathbf{L}\|$ is the distance between the point of light strength evaluation and the light source (Section 2.4).

- Under appropriate conditions, the normal map can be computed from the depth map (2.7), as well as the depth map can be computed from the normal map (2.9). If the depth map has continuous second derivatives then the normal map necessarily obeys the integrability condition (2.13).

- Surface reflectance can be regarded as a superposition of body and interface components (Section 2.6). Lambertian BRDF (2.23) is an example of the body reflectance, while the ideal specular BRDF (2.24) is a representative of the interface component.

Chapter 3

BRDF properties

This chapter reviews the properties of reflectance which are shared over large classes of the BRDFs. The classes discussed are: the BRDFs possessing some symmetry (i.e., Helmholtz reciprocity, isotropy, etc.) and the BRDFs whose body and interface components are both nonzero. In addition to these, we discuss the fact which is true for all BRDFs: the observed image intensity vanishes where the surface is not illuminated.

We briefly review how the properties of reflectance have been exploited in computer vision methods in order to grasp useful information about the observed objects. If such possibilities exist they generate algebraic constraints incorporating surface shape or lighting parameters. We then turn to the principal question: Do there exist any previously known constraints on normals and light directions which could help in uncalibrated photometric stereo? Or, do there even exist novel constraints that were previously unknown?

3.1 The review of BRDF properties

A BRDF may exhibit one or more of the following symmetries (also see Fig. 3.1):

(a) *Identity.* This trivial symmetry states that the BRDF stays constant when the incident and reflected directions are not changed. We review that this symmetry, although a trivial one, has an impact in computer vision.

(b) *Helmholtz reciprocity.* This symmetry states the invariance of BRDF with respect to swapping the incidence and reflected directions.

(c) *Isotropy.* If the properties of a surface patch (interface geometry and subsurface composition, etc.) are rotation-invariant within the reference plane, the BRDF inherits this symmetry. The BRDF is then invariant with respect to rotation of the surface patch around its normal.

(d) *Bilateral symmetry.* Similarly, if the properties of a surface patch are invariant when reflected around a plane containing the normal, then the BRDF inherits this symmetry.

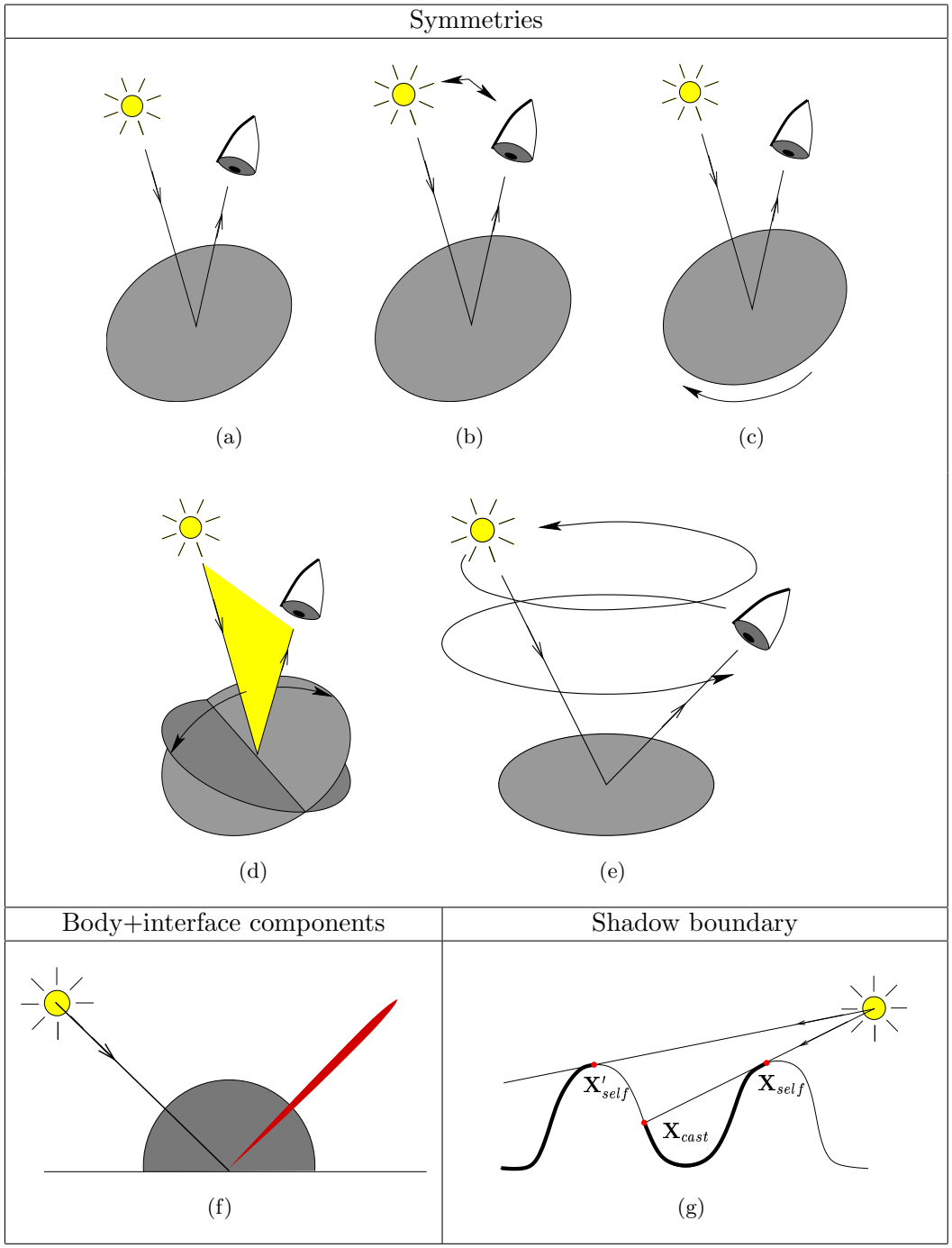


Figure 3.1: The list of discussed reflectance properties. For symmetries, the arrows indicate the change under which the BRDF stays invariant: (a) Identity; (b) Reciprocity (swapping the incident and viewing directions); (c) Isotropy (rotation of surface around the normal); (d) Bilateral symmetry (mirroring the surface by the principal plane); (e) Separability (rotation of the incident and/or viewing directions around the normal). Two more important facts: (f) Decomposition of reflectance to body and interface components which possess different geometric properties; (g) Shadow boundaries: $\mathbf{X}_{cast}$ is a cast shadow boundary, and $\mathbf{X}_{self}$, $\mathbf{X}'_{self}$ are self-shadow boundaries.

(e) *Separability.* This property means that the BRDF can be decomposed to a product of two functions, one of which is dependent on the incidence direction (θ_i, ϕ_i) and the other on the reflected direction (θ_r, ϕ_r). The physics interpretation behind is that the light is multiply scattered in the body of the subsurface before being transmitted back to air, such that the shape of distribution of reflected light does not depend on the incident direction. Hence the scattering is the essence of 'forgetting' the direction of incident illumination.

In addition to the symmetries, we discuss two more notable facts which carry geometric information:

(f) *Presence of specular reflectance component.* Interface component of the BRDF tends to be centered around the specular direction (cf. Fig. 2.9), and hence carries information about the direction of illumination.

(g) *Shadow boundary.* The observed image intensity vanishes at the shadow boundary. The shadow boundary occurs where the incident direction becomes tangential to the surface (a *self*-shadow boundary), or where other object part prevents the surface point from being visible from the light source (so-called *cast*-shadows).

We review the properties one after another.

3.1.1 Identity

Identity is a trivial symmetry which can be exploited if both the incident and reflected directions are kept fixed. If image intensity is recorded for two positions of a light source which lie on the same incident ray but in two different distances from the surface patch, then according to (1.7), the two values must follow the squared distance fall-off. This fact has been exploited by Magda et al. [25]. The camera was kept fixed, and an isotropic light source was moved across two non-intersecting imaginary surfaces: the outer one $\mathcal{Q}_o$ and the inner one $\mathcal{Q}_i$ (see Fig. 3.2(a)). The shape of $\mathcal{Q}_o$ and $\mathcal{Q}_i$ was known. Consider a point $\mathbf{X}$ positioned anywhere in object space. An arbitrary line containing $\mathbf{X}$ defines an incident direction, and the intersection of this line with the two surfaces $\mathcal{Q}_o$ and $\mathcal{Q}_i$ gives two light source positions $\mathbf{L}_o \in \mathcal{Q}_o$ and $\mathbf{L}_i \in \mathcal{Q}_i$ for which the image intensities were recorded. If the hypothesised point $\mathbf{X}$ lies on a surface then the recorded intensities must be consistent with the squared distance fall-off for any ray incident on $\mathbf{X}$. This criterion enables us to discriminate between the surface and non-surface points, and therefore to recover the depth map.

3.1.2 Reciprocity

Reciprocity, originally due to Helmholtz [37], is implied by time reversibility of reflection mechanisms on a microscopic level [5]. It can be exploited when the light source and the camera are swapped. Reciprocity was used only recently [25, 47] for shape reconstruction. Consider the camera is positioned at $\mathbf{A}$ and a light source

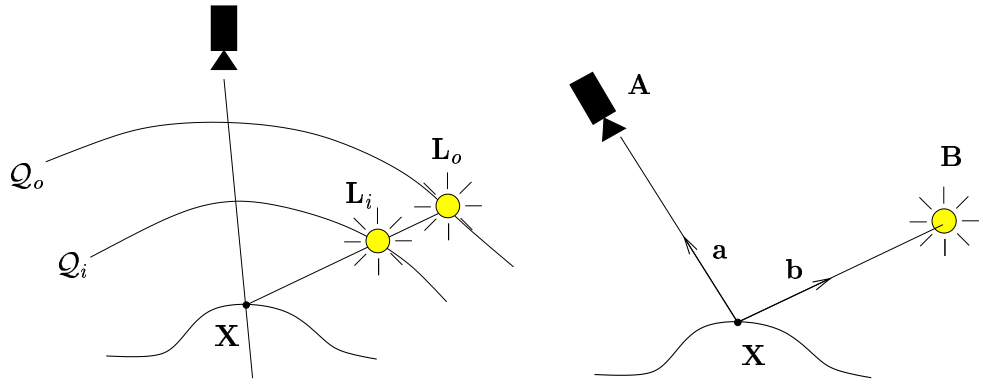


Figure 3.2: The experimental setups for exploiting identity (a) and reciprocity (b).

is positioned at $\mathbf{B}$ (see Fig. 3.2(b)). Assume point $\mathbf{X}$ lies on a surface. In this configuration, $\mathbf{a}$ is the viewing vector and $\mathbf{b}$ is the light vector (see the figure). According to (1.7), the observed intensity $I_{\mathbf{A}}$ of the point $\mathbf{X}$ is

$$I_{\mathbf{A}} = \eta\, f_{i,r}(T(\mathbf{b}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]); T(\mathbf{a}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]))\,(\mathbf{n}\cdot\mathbf{b})\,\sigma_{\mathbf{B}}\,, \tag{3.1}$$

where the irradiance term is written using the light strength $\sigma_{\mathbf{B}}$ at $\mathbf{X}$ resulting from light source positioned at $\mathbf{B}$ (cf. (1.8)). If the camera and the light source are now swapped such that the camera is at $\mathbf{B}$ and the light source is at $\mathbf{A}$ then the image intensity $I_{\mathbf{B}}$ observed for the same point is

$$I_{\mathbf{B}} = \eta\, f_{i,r}(T(\mathbf{a}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]); T(\mathbf{b}, [\mathbf{n}_\perp^1|\mathbf{n}_\perp^2|\mathbf{n}]))\,(\mathbf{n}\cdot\mathbf{a})\,\sigma_{\mathbf{A}}\,, \tag{3.2}$$

where $\mathbf{b}$ is now the viewing vector and $\mathbf{a}$ is now the light vector. As the BRDF is to be invariant with respect to swapping the incident and reflected directions, the two equations can be combined to eliminate the BRDF term:

$$I_{\mathbf{A}}\sigma_{\mathbf{A}}(\mathbf{n}\cdot\mathbf{a}) - I_{\mathbf{B}}\sigma_{\mathbf{B}}(\mathbf{n}\cdot\mathbf{b}) = 0\,, \tag{3.3}$$

or, with substitution $\mathbf{s}_{\mathbf{A}} = \sigma_{\mathbf{A}}\mathbf{a}$ and $\mathbf{s}_{\mathbf{B}} = \sigma_{\mathbf{B}}\mathbf{b}$

$$(I_{\mathbf{A}}\mathbf{s}_{\mathbf{A}} - I_{\mathbf{B}}\mathbf{s}_{\mathbf{B}})\cdot\mathbf{n} = 0\,. \tag{3.4}$$

This equation represents a constraint on surface orientation at $\mathbf{X}$, while the BRDF at $\mathbf{X}$ may be arbitrary. If the position of $\mathbf{X}$ in space is known, then two non-degenerate reciprocal configurations give the normal direction $\mathbf{n}$ at that point. More importantly, if $\mathbf{X}$ is not known then it can be hypothesised and at least three reciprocal pairs can be used to validate it by checking whether the matrix

$$\mathbf{W} = \begin{bmatrix} I_{\mathbf{A}_1}\mathbf{s}_{\mathbf{A}_1} - I_{\mathbf{B}_1}\mathbf{s}_{\mathbf{B}_1} \\ I_{\mathbf{A}_2}\mathbf{s}_{\mathbf{A}_2} - I_{\mathbf{B}_2}\mathbf{s}_{\mathbf{B}_2} \\ \cdots \\ I_{\mathbf{A}_n}\mathbf{s}_{\mathbf{A}_n} - I_{\mathbf{B}_n}\mathbf{s}_{\mathbf{B}_n} \end{bmatrix} \tag{3.5}$$

is of rank 2. In [25, 47], a simple condition number tests were applied to $\mathbf{W}$. If the matrix is close to singular then the hypothesized point is likely to lie on the surface. The normal orientation $\mathbf{n}$ can be then determined by requiring

$$\mathbf{W} \cdot \mathbf{n} = \mathbf{0}\,. \tag{3.6}$$

The method developed is quite elegant and produces supreme reconstruction results. This is due to the fact that the method enables to compute both the depth *and* the surface normal at the same time. This new method is also quite intriguing due to the fact that the surface reflectance reciprocity is valid for almost all BRDFs.

3.1.3 Isotropy and bilateral symmetry

Isotropic parametric models of reflectance are often employed to represent reflectance. However, isotropy is a BRDF symmetry which is difficult to be exploited directly. This is because to exploit isotropy, the incident and reflected directions have to be rotated jointly around a surface normal. But a curved surface accommodates many different surface orientations, and it is impossible to rotate the viewing and light directions jointly around all surface orientations at the same time.

If the BRDF is isotropic then the underlying reasons for isotropy (the surface geometry, subsurface composition etc.) most often also imply bilateral symmetry of the BRDF with respect to *any* plane containing the normal. To exploit this property, the viewing and light directions have to be reflected around this plane. The bilateral symmetry then could be exploited for all normals which lie in that plane. It is, however, an open question whether this fact could be exploited efficiently in computer vision in some applications.

3.1.4 Separability

Separable BRDFs are those which can be written as the product of two functions. One function is dependent on the incident direction and the other one on the reflected direction:

$$f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r) = f_i(\theta_i, \phi_i) f_r(\theta_r, \phi_r)\,. \tag{3.7}$$

The underlying assumption leading to separability is that there exists multiple scattering within the (sub)surface that provides the 'forgetting' of the incident light direction [39]. The separability of BRDF was exploited by Wolff [40] in shape reconstruction of (possibly) texture-less surfaces by binocular stereo. The idea of the method is the following. Consider a point $\mathbf{X}$ on the surface. The point is observed by two cameras and illuminated by a single point light source positioned at two locations in turn (see Fig. 3.3). Consider the ratio of intensities recorded by the first camera for the two light sources. Obviously, due to (1.7) and (3.7) this ratio will be

$$\frac{f_i(T(\mathbf{l}_1, [\mathbf{n}^1_\perp | \mathbf{n}^2_\perp | \mathbf{n}]))\,(\mathbf{n} \cdot \mathbf{l}_1)\sigma_1}{f_i(T(\mathbf{l}_2, [\mathbf{n}^1_\perp | \mathbf{n}^2_\perp | \mathbf{n}]))\,(\mathbf{n} \cdot \mathbf{l}_2)\sigma_2}\,, \tag{3.8}$$

where $\mathbf{n}$ is the normal orientation at $\mathbf{X}$ and σ_1, σ_2 are the light strengths of the light source at the two positions. It turns out that this ratio is independent on the

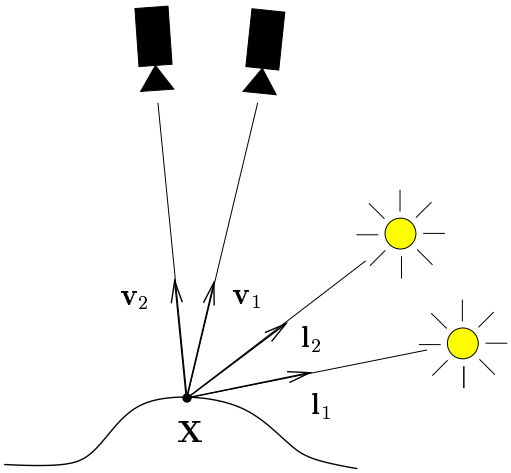


Figure 3.3: Exploiting separability.

viewing direction and hence for a given point it should be the same in both cameras. The ratio is then used as an entity to be matched by geometric stereo algorithm.

3.1.5 Shadow boundary

The shadow boundary occurs in two different situations. The first one is when the incident direction becomes tangent to the surface. That case corresponds to self-shadow (also termed attached shadow) boundary. In Fig. 3.1(g), the self-shadow boundaries are encountered at $\mathbf{X}_{self}$ and $\mathbf{X}'_{self}$. The second situation is when the path from the light source to the point is occluded by another part of the object. That case corresponds to a cast shadow boundary which occurs at $\mathbf{X}_{cast}$ in Fig. 3.1(g). The employment of information carried by shadows has a long history and it is not our aim to bring a detailed overview here. A traditional problem of computer vision is the *shape from darkness* [22, 19, 7] the goal of which is to reconstruct the 3D surface shape from shadow boundaries in images captured under different illuminations and a fixed viewpoint. Another example of shadow boundary employment is the work of Bouguet and Perona [4]. Their work is similar to range-finder methods: they use a single point light source to illuminate the object, and move a thin stick in front of the light source. The surface shape is then recovered from a temporal sequence of shadows which are cast by the stick in the images.

3.1.6 Presence of specular reflectance component

The distribution of the specular component of reflectance is usually centered around the direction corresponding to mirror-like reflection (cf. Fig. 3.1(f)). It carries a valuable geometric information. Consider the case of ideal specular (mirror-like) reflectance given by (2.24). If a specularity occurs in an image then it means that the normal at the point accommodating the specularity is the bisector of the viewing

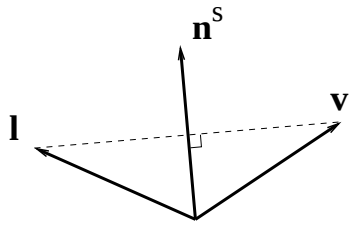


Figure 3.4: Specular geometry configuration. Surface normal $\mathbf{n}^S$ is a bisector between the viewing direction $\mathbf{v}$ and the illumination direction $\mathbf{l}$.

and light directions (Figure 3.4 shows the situation). We call this normal a *specular normal*.

The information carried by the specular configuration of viewing vector, light vector, and the normal was exploited in a number of works. A recent one is by Zheng and Murata [46] who employ it for reconstructing specular (e.g., metal) curved surfaces. They use a fixed camera, a fixed circular-shaped light source and rotate the object on a turntable to capture a sequence. The specularities detected in this sequence are then used to recover the surface shape.

3.2 Implications for uncalibrated photometric stereo

Which of the discussed properties can be exploited in uncalibrated photometric stereo? The photometric stereo can use only such properties that can be exploited without alteration of the object position and the viewpoint. We discuss them in turn:

- Identity has already been used in a setup which corresponds to photometric stereo with careful control of the light source position [25].
- Reciprocity could be exploited only if the reciprocal pairs are captured with the camera center and the light source position co-located in which a case it does not deliver any constraint.
- Isotropy and bilateral symmetry have been discussed to be hard to be exploited on a curved surface.
- Shadows has been shown to represent no further information which could be used to reduce the uncalibrated photometric stereo ambiguity [2].

The properties described above seem to deliver no novel result. Following two properties do.

3.2.1 Separable BRDFs

There is no obvious approach which would enable to exploit separability discussed in Section 3.1.4 in photometric stereo. We noticed, however, that if a BRDF is not only

separable but also isotropic, then there exists a possibility to determine additional information about normals on a general (possibly curved) surface. For a separable isotropic BRDF, the functions f_i and f_r in (3.7) must be independent on ϕ_i and ϕ_r, respectively, such that the BRDF can be written as

$$f_{i,r}(\theta_i, \phi_i; \theta_r, \phi_r) = f_i(\theta_i) f_r(\theta_r)\,. \tag{3.9}$$

In this section we demonstrate that if the BRDF is separable and isotropic then it is possible to design experimental setups which enable to measure the projections of normals onto the image plane and the light plane. These projections can be determined only up to a scale and correspond to the **v**-tilts and **l**-tilts, respectively, which were defined in Section 2.3. We start by illustrating how the **l**-tilt of normals can be recovered.

Observation (a thought experiment)

Let there be a normal $\mathbf{n}$ and vector $\mathbf{l}$ which is not collinear with $\mathbf{n}$ but otherwise it is arbitrary. Let $\mathbf{l}^1_\perp$, $\mathbf{l}^2_\perp$ be two vectors which form an orthonormal basis together with $\mathbf{l}$.

Consider an experiment depicted in Fig. 3.5. A light source $\mathbf{l}_{var}$ characterised by its light strength σ rotates within a cone whose axis is $\mathbf{l}$. The figure shows the situation from a perspective (bottom) and viewed orthographically from the direction of $\mathbf{l}$ (top). The angle of rotation is measured within the reference basis $[\mathbf{l}^1_\perp, \mathbf{l}^2_\perp]$. The incident angle θ_i is dependent on the angle of rotation ($\theta_i = \theta_i(\epsilon)$). Obviously there holds that

$$\begin{aligned} \theta_i(\epsilon^{\mathbf{n}+} + \gamma) &= \theta_i(\epsilon^{\mathbf{n}+} - \gamma)\,, \qquad \forall \gamma \in \mathbb{R}\,, \\ \theta_i(\epsilon^{\mathbf{n}-} + \gamma) &= \theta_i(\epsilon^{\mathbf{n}-} - \gamma)\,, \end{aligned} \tag{3.10}$$

where $\epsilon^{\mathbf{n}+}$ is the angle of the orthographic projection of the normal $\mathbf{n}$ onto the reference basis, and $\epsilon^{\mathbf{n}-} = \epsilon^{\mathbf{n}+} + \pi$. The observed image intensity $I(\epsilon)$ as a function of the rotation angle ϵ is

$$I(\epsilon) = \eta f_i(\theta_i(\epsilon)) f_r(\theta_r)\, \sigma \cos\theta_i(\epsilon)\,, \tag{3.11}$$

and inherits this symmetry, such that there holds

$$\begin{aligned} I(\epsilon^{\mathbf{n}+} + \gamma) &= I(\epsilon^{\mathbf{n}+} - \gamma)\,, \qquad \forall \gamma \in \mathbb{R}\,, \\ I(\epsilon^{\mathbf{n}-} + \gamma) &= I(\epsilon^{\mathbf{n}-} - \gamma)\,. \end{aligned} \tag{3.12}$$

If we measure $I(\epsilon)$ for a whole circle then it may be possible to identify two points of symmetry $\epsilon^{\mathbf{n}+}$ and $\epsilon^{\mathbf{n}-}$ in $I(\epsilon)$ such that (3.12) is valid. The two points correspond to the orientation of the orthonormal projection of the normal onto the reference basis:

$$\begin{aligned} (\cos\epsilon^{\mathbf{n}+}, \sin\epsilon^{\mathbf{n}+})^T &= \alpha [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{n}\,, \\ (\cos\epsilon^{\mathbf{n}-}, \sin\epsilon^{\mathbf{n}-})^T &= -\alpha [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{n}\,, \end{aligned} \tag{3.13}$$

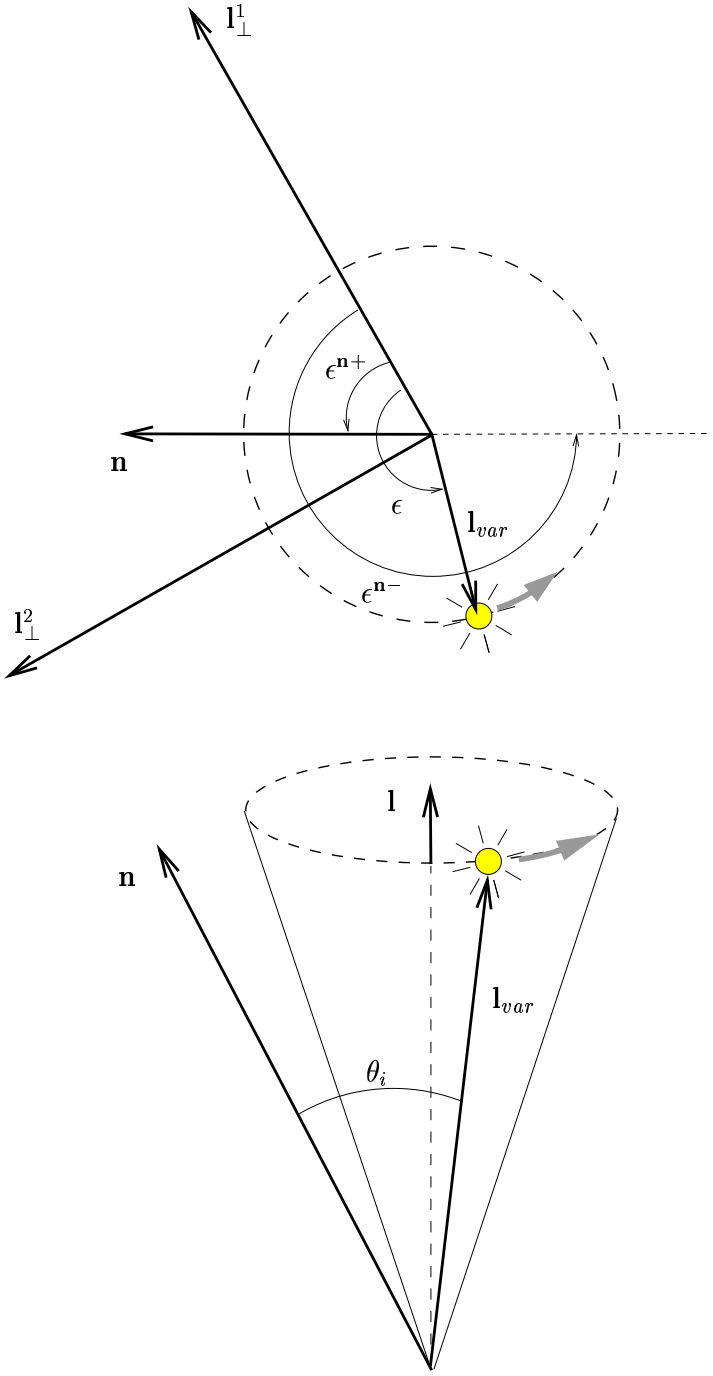


Figure 3.5: Illustration of **l**-tilt determination of normal **n**. A light source $\mathbf{l}_{var}$ of light strength σ is rotated around a cone the axis of which is **l**. The situation is shown viewed from a perspective (bottom) and from above (top). If the radiance at the normal location is measured as a function of ϵ then this function is symmetric around $\epsilon^{\mathbf{n}+}$, as well as about $\epsilon^{\mathbf{n}-} = \epsilon^{\mathbf{n}+}+\pi$. Identifying these two points of symmetry in the measured radiance gives the **l**-tilt.

where $\alpha > 0$ is $\alpha = 1/\|[\mathbf{l}_\perp^1|\mathbf{l}_\perp^2]^T\mathbf{n}\|$. Without further knowledge about $f_i(\theta)$ it is impossible to determine which of the two points of symmetry $\epsilon^{\mathbf{n}+}$ and $\epsilon^{\mathbf{n}-}$ corresponds to the proximal position of $\mathbf{l}_{var}$ and $\mathbf{n}$. We conclude that the it is possible to determine the normal projection onto the reference plane basis $[\mathbf{l}_\perp^1|\mathbf{l}_\perp^2]$ only up to a nonzero scale. Such entity is the previously defined **l**-tilt.

The construction of the **v**-tilt can be demonstrated analogously, for rotation of the viewing direction within a cone the axis of which is **v**.

Practical implementation of the tilts measurement

The thought experiment showed how to construct the **l**-tilts and **v**-tilts from by a measurement employing the rotation of light around a principal direction **l**, and rotating the viewpoint around a principal direction **v**, respectively.

It would be quite cumbersome to realise such experiment in practice, however. This is especially true for the **v**-tilt determination because the experimental procedure would require the camera to move around a cone. The principal problem is that the points in individual images would first have to be put into correspondence to get the dependence $I(\epsilon)$ for each point on a surface. This, however, is a non-trivial task if surface geometry is not known in advance.

If there exists an experimental approach which would be capable of determining the tilts without actually moving the camera or the light source then it would be much more useful. It turns out that there is a possibility to determine the tilts without actually moving the camera or the light source. This possibility exists for a special class of dielectric smooth surfaces if polarisation of light is employed. The polarisation reflectance model for body reflectance on such surfaces possesses the essential property of being separable with respect to the incident and reflected light parameters. The separability follows from the assumption that the light which penetrates under the smooth optical boundary becomes completely depolarised by internal scattering [41]. The model is constructed by enriching the Lambert's model by two transmissivity terms $\mathcal{T}(\theta_i, n)$, $\mathcal{T}(\theta_r, n)$ such that for the observed intensity there holds

$$I = \frac{\eta}{\pi}\,\mathcal{T}(\theta_i, n)\mathcal{T}(\theta_r, n)\sigma\rho\,(\mathbf{n}\cdot\mathbf{l})\,, \tag{3.14}$$

where n is the index of refraction and the other variables are as in (1.11). The two transmissivity terms correspond to the processes of (1) incident light refraction into the sub-surface (term $\mathcal{T}(\theta_i, n)$) and, after internal scattering, part of it (2) refracting back out into the air (term $\mathcal{T}(\theta_r, n)$). It turns out from elementary optics [3] that if a linear polarising filter is mounted in front of the light source and the pixel intensity I is measured as a function of the polariser orientation ϵ, the intensity will be maximum for orientation ϵ_l^{max} when the polariser orientation corresponds to the projection of normal onto the plane perpendicular to the light vector (see Fig. 3.6). It holds [3] that the dependence of intensity on the polariser orientation is described by the following substitution in (3.14):

$$\mathcal{T}(\theta_i, n) \to \frac{1}{2}\left\{\mathcal{T}_\parallel(\theta_i, n)\cos^2\Delta\epsilon_l + \mathcal{T}_\perp(\theta_i, n)\sin^2\Delta\epsilon_l\right\}, \tag{3.15}$$

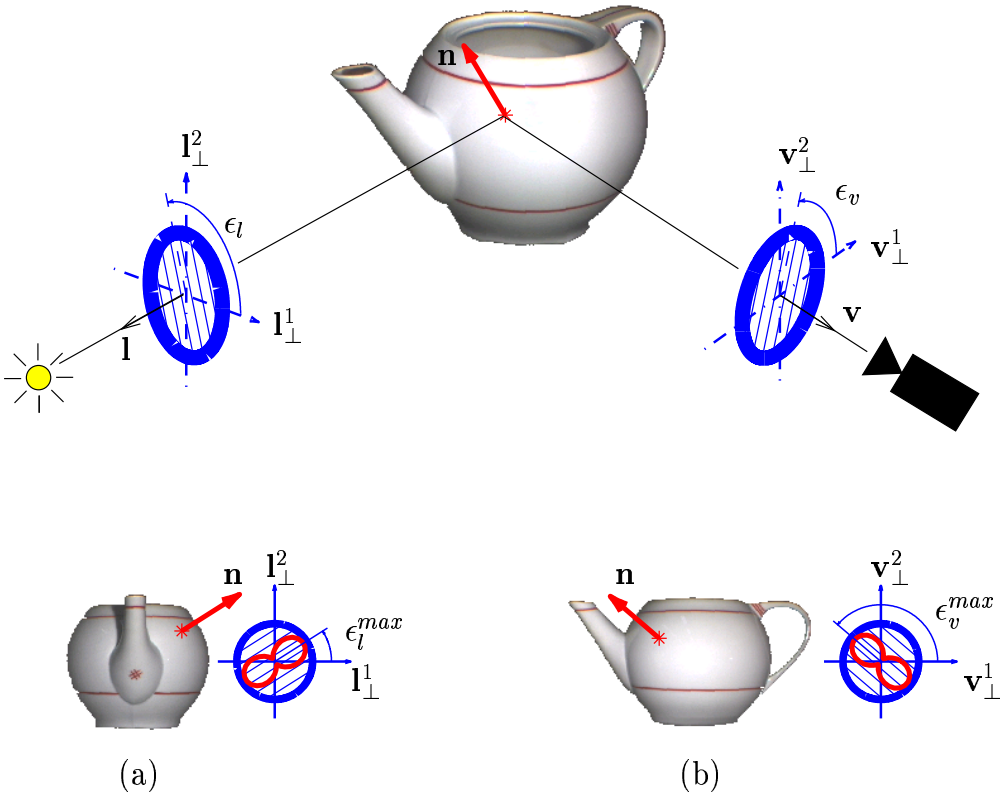
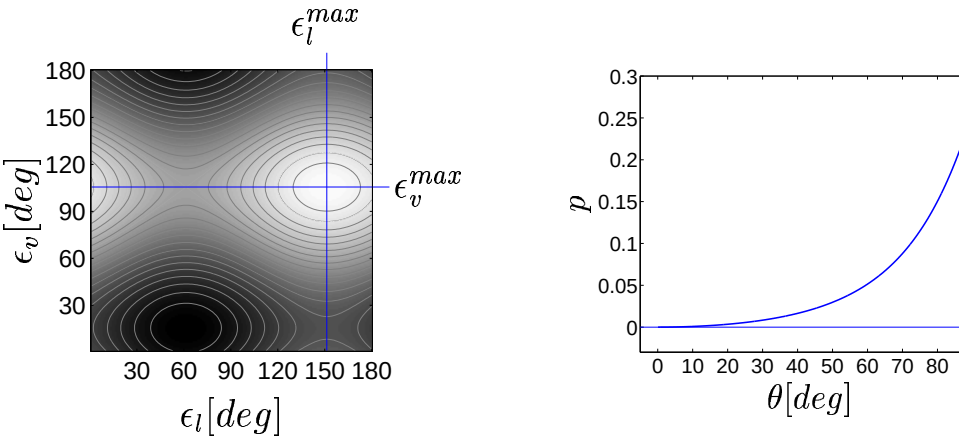
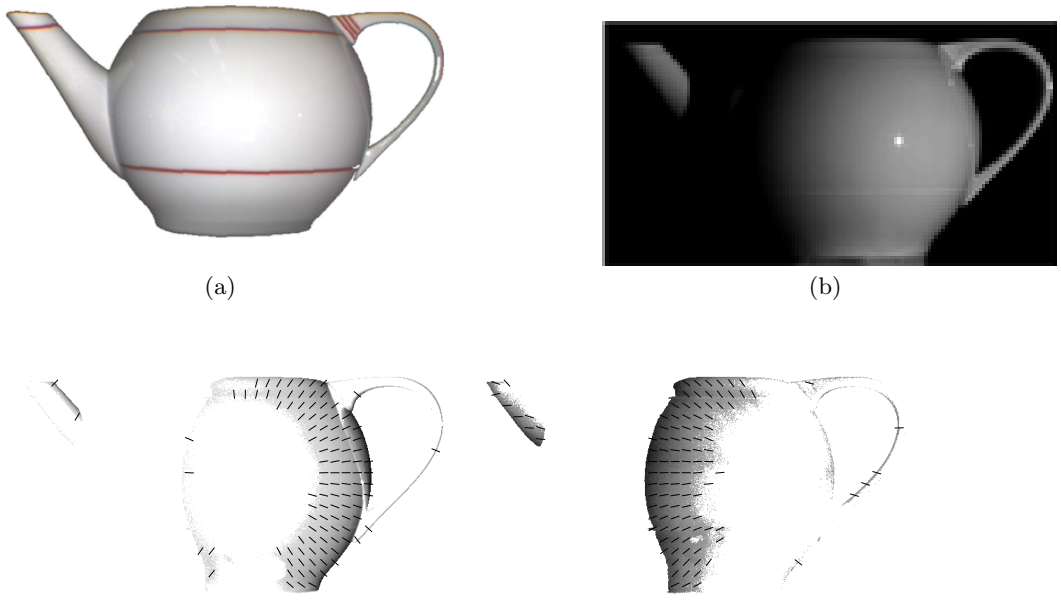


Figure 3.6: Evaluation of the **l**-tilt and **v**-tilt for a normal **n** on an object surface. If the linear polariser placed in front of the light source is rotated then the dependence of observed image intensity is as shown by the 8-shaped red plot in subfigure (a). The maximum gives the **l**-tilt. The **v**-tilt is found analogously from dependence of image intensity on the angle of rotation of a linear polariser placed in front of the camera (b). The subfigures (a) and (b) also show the teapot and the normal **n** as the light and camera see it, respectively.



(a) (b)

Figure 3.7: (a) Image intensity as a function of orientations ϵ_l and ϵ_v of the two polarisers. (b) Dependence of the partial polarisation on the incident/reflected angle, displayed for refraction index $n = 1.3$.



(a) (b)

(c) (d)

Figure 3.8: A real 3D object (a) illuminated from the right (b), the **v**-tilt (c) and the **l**-tilt (d) showed as needle maps superimposed on the **v**-partial polarisation and the **l**-partial polarisation values, respectively (dark points correspond to high partial polarisation values). The tilts are displayed only where the partial polarisation exceeds 3%, and the resolution of the tilt map is reduced by the factor of 20 for the sake of readability.

where the terms $\mathcal{T}_\perp$ and $\mathcal{T}_\parallel$ are the transmissivities for light polarised perpendicularly and parallelly to illumination plane [3], and $\Delta\epsilon_l = \epsilon_l - \epsilon_l^{max}$. The image intensity hence varies harmonically with the polariser orientation ϵ_l which is measured as ordinary polar angle within the polariser coordinate system $[\mathbf{l}_\perp^1 | \mathbf{l}_\perp^2]$. It is therefore sufficient to capture images for several (at least three) orientations of the polariser and then apply a harmonic fit in each pixel. The phase of this fit then gives the ϵ_l^{max} in each pixel, and consequently the **l**-tilt. The reliability of the **l**-tilt determination depends on the relative magnitude of the harmonic variation:

$$p_{\mathbf{l}} = \frac{\mathcal{T}_\parallel(\theta_i, n) - \mathcal{T}_\perp(\theta_i, n)}{\mathcal{T}_\parallel(\theta_i, n) + \mathcal{T}_\perp(\theta_i, n)} \,. \tag{3.16}$$

We call $p_{\mathbf{l}}$ the **l**-partial polarisation. Fig. 3.7(b) shows the dependence of the partial polarisation on the angle of incidence θ_i.

The **v**-tilt is then constructed analogously using a polarising filter mounted in front of the camera. We assume that the reference basis of the polariser equals the image plane basis $[\mathbf{v}_\perp^1 | \mathbf{v}_\perp^2]$. The dependence of image intensity on the camera polariser orientation ϵ_v can be described by the substitution [41]

$$\mathcal{T}(\theta_r, n) \rightarrow \frac{1}{2} \left\{ \mathcal{T}_\parallel(\theta_r, n) \cos^2 \Delta\epsilon_v + \mathcal{T}_\perp(\theta_r, n) \sin^2 \Delta\epsilon_v \right\} . \tag{3.17}$$

where $\Delta\epsilon_v = \epsilon_v^{max} - \epsilon_v$ is the angular difference between the polariser orientation corresponding to the projection of the normal onto the image plane ϵ_v^{max} and current polariser orientation ϵ_v (see Fig. 3.5). The analogous form of (3.15) and (3.17) enables us to quickly come to the result that the **v**-tilt can be determined as the orientation of the camera polariser under which the maximum image intensity is observed. This orientation can again be determined as a phase of the harmonic fit applied to image intensities measured for several polariser orientations. The reliability of the **v**-tilt evaluation raises with the reflectance angle θ_r. The reliability is represented by the **v**-partial polarisation $p_{\mathbf{v}}$ that is defined in analogy with (3.16).

To give this section a summary, the observed image intensity is separable with respect to orientation of the light polariser and the camera polariser (see Fig. 3.7(a)). It is possible to determine the **l**-tilt constraint and then the **v**-tilt constraint for each normal by first rotating the polariser placed in front of the light source and then in front of the camera. The **l**-partial polarisation and **v**-partial polarisation serve as factors that express the reliability of the tilt determination. The **v**-tilts and **l**-tilts of normals are shown in Fig. 3.8. The tilts are displayed superimposed on the partial polarisation map where the partial polarisation exceeds 3%.

The question of how the information carried by the tilts of normals can be used for reducing the uncalibrated photometric stereo ambiguity is the topic of Chapter 4.

3.2.2 Presence of specular reflectance component

Consider a specularity which is observed in an image captured under illumination represented by a scaled light vector **s**. The specularity is observed at a point accommodating the scaled normal **b**. We know that both **s** and **b** are determined

from uncalibrated photometric stereo ambiguously, and hence reflection of $\mathbf{s}$ around $\mathbf{b}$ does not necessarily lead to correct viewpoint. Imposing the constraint that the true viewpoint has to be obtained leads to reduction of uncalibrated photometric stereo ambiguity. This will be shown in Chapter 5.

3.3 Summary of the chapter

This chapter reviewed the properties a BRDF may have. Two of the properties were identified which could be useful in photometric stereo.

Firstly, we observed that if the BRDF is separable and isotropic then it is possible to construct an experiment which measures the orthographic projections of normals onto the light plane. Projection of each normal can be evaluated only up to a nonzero scale and is termed the $\mathbf{l}$-tilt. A practical realisation of $\mathbf{l}$-tilt evaluation employs polarisation measurement with a linear polarising filter mounted such that its axis coincides with the light direction. Analogously, it was shown that it is possible to evaluate orthographic projections of normals onto the image plane. These projections are again determined up to a scale and can be evaluated by a measurement employing a linear polarising filter installed in front of a camera. Whether knowing the $\mathbf{v}$-tilts and $\mathbf{l}$-tilts represents a helpful knowledge in uncalibrated photometric stereo is the topic of the next chapter.

Secondly, it was reviewed that specularities occur in images when the specularly reflecting normal is a bisector of the viewing and illumination directions. The consequence of this fact for uncalibrated photometric stereo will be observed in Chapter 5.

Chapter 4

Separable isotropic BRDFs

In this chapter we show how the uncalibrated Lambertian photometric stereo ambiguity can be reduced when additional information represented by the **v**-tilts and **l**-tilts of normals is exploited*. Section 4.1 presents the details regarding the **v**-tilts and **l**-tilts determination and defines the terms of the **v**-*tilt constraint* an the **l**-*tilt constraint*. Section 4.2 then shows how the photometric stereo ambiguity is reduced when these constraints are applied individually, and in combination with each other or with the integrability constraint. Section 4.3 then presents the experiments we made in order to test the applicability of theoretical results in practice. Section 4.4 brings the summary of this chapter.

4.1 Tilt constraints

In this section, we briefly review the details of the tilts evaluation and define the **v**-tilt constraint and the **l**-tilt constraint.

- The **v**-tilts of normals are evaluated by rotating a polarising filter positioned in front of the camera as described in Section 3.2.1. The tilt is measured in a local coordinate frame of the camera polariser. We consider only the case in which the camera polariser basis coincides with the image plane basis $(\mathbf{v}_\perp^1, \mathbf{v}_\perp^2)$. This situation is depicted in Fig. 3.6†.

  We define the **v-tilt constraint** as follows. Let $\mathcal{V}$ be the set of indices of normals $\mathbf{n}$ for which the **v**-tilt is measured: $\mathcal{V} = \{i\text{: } \mathbf{t}^{\mathbf{v}}(\mathbf{n}_i) \text{ is measured}\}$. For the sake of brevity, we denote $\mathbf{t}^{\mathbf{v}}(\mathbf{n}_i)$ by $\mathbf{t}_i^{\mathbf{v}}$. Let $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be the solution to (1.17). We say that $(\mathbf{B}, \mathbf{S})$ obeys the **v**-*tilt constraint* if

  $$\forall i \in \mathcal{V} \;\; \exists \lambda_i \neq 0\text{:} \qquad [\mathbf{v}_\perp^1 | \mathbf{v}_\perp^2]^T \mathbf{b}_i = \lambda_i \mathbf{t}_i^{\mathbf{v}}\,, \tag{4.1}$$

  where $\mathbf{v}_\perp^1$ and $\mathbf{v}_\perp^2$ are the image plane basis vectors. Equation (4.1) requires the measured $\mathbf{t}_i^{\mathbf{v}}$ and the tilt evaluated by the projection of the scaled normal $\mathbf{b}_i$ onto the image plane basis be collinear.

*The definition of a tilt can be found in Section 2.3. A principle of the **v**-tilt and **l**-tilt measurement is described in Section 3.2.1.

†Aligning the two coordinate system is achieved by calibration.

- The **l**-tilts of normals are evaluated by rotating the polarising filter positioned in front of the light source the light vector of which is **l**. The **l**-tilt is measured in a local coordinate frame of the light polariser. We assume the basis of the light polariser $(\mathbf{l}_\perp^1, \mathbf{l}_\perp^2)$ to be unknown.

  We define the **l-tilt constraint associated with the r-th illuminant** as follows. Let $\mathbf{l}^*$ be the true light vector of the r-th illuminant. Let $\mathcal{L}$ be the set of indices of normals for which the $\mathbf{l}^*$-tilt is measured: $\mathcal{V} = \{i\colon \mathbf{t}^{\mathbf{l}^*}(\mathbf{n}_i) \text{ is measured}\}$. We denote $\mathbf{t}^{\mathbf{l}^*}(\mathbf{n}_i)$ by $\mathbf{t}_i^{\mathbf{l}^*}$. Let $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be the solution to (1.17). Let $\mathbf{l}$ be the light vector given by the r-th scaled light vector normalised to unity, $\mathbf{l} = \mathbf{s}_r / \|\mathbf{s}_r\|$. We say that the solution $(\mathbf{B}, \mathbf{S})$ obeys the **l**-tilt constraint associated with the r-th illuminant if

$$\exists \mathbf{l}_\perp^1, \mathbf{l}_\perp^2\colon \quad [\mathbf{l}_\perp^1, \mathbf{l}_\perp^2, \mathbf{l}] \in O(3) \wedge \left(\forall i \in \mathcal{L} \;\; \exists \lambda_i \neq 0\colon \quad [\mathbf{l}_\perp^1 | \mathbf{l}_\perp^2]^T \mathbf{b}_i = \lambda_i \mathbf{t}_i^{\mathbf{l}^*} \right), \tag{4.2}$$

  which means that there exists an orthonormal basis $(\mathbf{l}_\perp^1, \mathbf{l}_\perp^2)$ of the plane perpendicular to the light vector $\mathbf{l}$ such that the measured $\mathbf{t}_i^{\mathbf{l}^*}$ and the tilt evaluated by projection of $\mathbf{b}_i$ onto this plane basis are collinear.

4.2 Transformations that preserve tilt constraints

Consider scaled normals and lights which obey the **v**-tilt constraint. In this section, we identify the classes of linear invertible transformations which can be applied to them according to (1.21) such that the **v**-tilt constraint is preserved. Subsequently, we do the same for the **l**-tilt constraint.

We will make three key observations afterwards. First, constraining the normals to be consistent with both the **v**-tilts and **l**-tilts reduces the photometric stereo ambiguity into the composition of isotropic scaling and a convex/concave ambiguity. Second, the **l**-tilt constraint combined with the integrability constraint results in the same reduction of ambiguity as above. Third, the **v**-tilt constraint combined with the integrability constraint results in a bas-relief ambiguity.

Throughout the analysis we work in the natural coordinate frame represented by the vectors $[\mathbf{v}_\perp^1 | \mathbf{v}_\perp^2 | \mathbf{v}]$. We thus set

$$[\mathbf{v}_\perp^1 | \mathbf{v}_\perp^2 | \mathbf{v}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.3}$$

Theorem 4.1 *Let $\mathcal{V}$ be the set of indices of normals for which the* **v***-tilt is measured. Let $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be the solution to (1.17), and let $(\mathbf{B}, \mathbf{S})$ obey the* **v***-tilt constraint represented by (4.1).*

The only linear invertible transformations which can be applied to $\mathbf{B}$ *and* $\mathbf{S}$ *according to (1.21) such that the* **v***-tilt constraint is preserved are:*

$$\mathbf{A} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ \mu & \nu & \tau \end{bmatrix} \qquad \begin{array}{l} \lambda, \tau \neq 0\,, \\ \mu, \nu \in \mathbb{R}\,. \end{array} \tag{4.4}$$

Proof. The proof is led as follows. Instead of $(\mathbf{B},\mathbf{S}) \in \mathcal{T}(\mathbf{J})$ which obey the **v**-tilt constraint, we start with arbitrary $(\hat{\mathbf{B}},\hat{\mathbf{S}}) \in \mathcal{T}(\mathbf{J})$ possibly not consistent with the measured **v**-tilts. We show how to find a matrix $\mathbf{Q}$ which transforms $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ according to (1.21) such that the transformed scaled normals $\mathbf{B} = \mathbf{Q}\hat{\mathbf{B}}$ obey the **v**-tilt constraint. Then, we show that $\mathbf{Q}$ is determined up to an ambiguity represented by matrices in (4.4).

Requiring that the transformed scaled normals $\mathbf{B}$ obey the **v**-tilt constraint means (cf. Equations (4.1) and (4.3)):

$$\forall i \in \mathcal{V} \;\; \exists \lambda_i \neq 0\text{:} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{Q}\hat{\mathbf{b}}_i = \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix} \hat{\mathbf{b}}_i = \lambda_i \;\; \mathbf{t}_i^{\mathbf{v}} \,, \tag{4.5}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the first two rows of $\mathbf{Q}$. The relation (4.5) represents a set of homogeneous linear equations for the first two rows of $\mathbf{Q}$ and for the unknown scalars λ_i. Knowing the **v**-tilt for N normals represents $2N$ homogeneous equations in $N+6$ unknown parameters; hence the minimum number of tilts needed to compute the unknowns is five. The rows $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are determined only up to a scale $\lambda \neq 0$, and the third row $\mathbf{Q}_3$ of $\mathbf{Q}$ is undetermined. It turns out that the only linear invertible transformations $\mathbf{A}$ which can be applied from the left to $\mathbf{Q}$ such that the first two rows are at most scaled are those written in (4.4). The parameter τ has to be nonzero for the matrix to be invertible. □

Theorem 4.2 *Let $\mathbf{l}^*$ be the true light vector of the r-th illuminant and let $\mathcal{L}$ be the set of indices of normals for which the $\mathbf{l}^*$-tilt is measured. Let $(\mathbf{B},\mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be the solution to (1.17). Let $\mathbf{l}$ be the r-th scaled light vector normalised to unity, i.e., $\mathbf{l} = \mathbf{s}_r/\|\mathbf{s}_r\|$. Let (4.2) be valid for $(\mathbf{B},\mathbf{S})$, and let $(\mathbf{l}^1_\perp, \mathbf{l}^2_\perp)$ be the orthonormal basis of a plane perpendicular to $\mathbf{l}$ for which (4.2) holds.*

The only linear invertible transformations which can be applied to $\mathbf{B}$ and $\mathbf{S}$ according to (1.21) such that the $\mathbf{l}$-tilt constraint associated with r-th illuminant is preserved are:

$$\mathbf{A} = \left[\bar{\mathbf{l}}^1_\perp | \bar{\mathbf{l}}^2_\perp | \bar{\mathbf{l}}\right] \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} \left[\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}\right]^T , \qquad \begin{array}{l} \alpha, \beta \neq 0 \,, \\ [\bar{\mathbf{l}}^1_\perp | \bar{\mathbf{l}}^2_\perp | \bar{\mathbf{l}}] \in O(3) \,. \end{array} \tag{4.6}$$

The free parameters of this class of transformations are $\alpha, \beta \neq 0$ and an orthonormal matrix $\left[\bar{\mathbf{l}}^1_\perp | \bar{\mathbf{l}}^2_\perp | \bar{\mathbf{l}}\right] \in O(3)$. The interpretation of this orthonormal matrix is the following. The last column $\bar{\mathbf{l}}$ is the light vector of the transformed scaled light vector $\bar{\mathbf{s}}_1 = \mathbf{A}^{-T}\mathbf{s}_1$. The first two columns of this matrix $(\bar{\mathbf{l}}^1_\perp, \bar{\mathbf{l}}^2_\perp)$ define the transformed local coordinate system of the light polariser.

Proof. Like in the proof of Theorem 4.1, we start with $(\hat{\mathbf{B}},\hat{\mathbf{S}}) \in \mathcal{T}(\mathbf{J})$ which is possibly not consistent with the measured $\mathbf{l}$-tilts of normals. We show how to find a matrix $\mathbf{Q}$ which transforms $\hat{\mathbf{B}}$ and $\hat{\mathbf{S}}$ according to (1.21) such that the transformed scaled normals $\mathbf{B} = \mathbf{Q}\hat{\mathbf{B}}$ obey the $\mathbf{l}$-tilt constraint evaluated for the transformed scaled light $\mathbf{s}_r = \mathbf{Q}^{-T}\hat{\mathbf{s}}_r$. Then, we show that $\mathbf{Q}$ is determined up to an ambiguity represented by matrices in (4.6).

Requiring that the transformed normals are consistent with the known tilts means finding a transformation matrix $\mathbf{Q}$ and an orthonormal double $(\mathbf{l}^1_\perp, \mathbf{l}^2_\perp)$ such that (cf. (4.2))

$$\forall i \in \mathcal{L} \;\; \exists \lambda_i \neq 0\text{:} \quad [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{Q} \hat{\mathbf{b}}_i = \lambda_i \mathbf{t}^{\mathbf{l}^*}_i \,. \tag{4.7}$$

Let $\hat{\mathbf{l}} = \hat{\mathbf{s}_r}/\|\hat{\mathbf{s}_r}\|$ denote the original light vector. The orthonormal double $(\mathbf{l}^1_\perp, \mathbf{l}^2_\perp)$ is constrained to be orthogonal to the transformed light vector

$$\mathbf{l} = \beta \mathbf{Q}^{-T} \hat{\mathbf{l}}, \qquad \beta > 0 \,, \tag{4.8}$$

where β is the ratio of the light strengths before and after the transformation. Writing the orthogonality condition explicitly gives

$$[\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{l} = \mathbf{0} \,. \tag{4.9}$$

Let us analyse (4.7) first. The entity $[\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{Q}$ on the left-hand side is an unknown 2×3 matrix which we denote $\mathbf{C}^T$. Then, (4.7) is written as $\mathbf{C}^T \hat{\mathbf{b}}_i = \lambda_i \mathbf{t}^{\mathbf{l}^*}_i$ which is formally equivalent to (4.5). Thus, the matrix $\mathbf{C}$ is determined up to an unknown scaling factor if at least five tilts in a general configuration are known. This means

$$\alpha \mathbf{C}^T = \; [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T \mathbf{Q} \,, \qquad \alpha \neq 0 \,, \tag{4.10}$$

where $\alpha \neq 0$ is an unknown scaling factor. Wanting to obtain possible solutions for $\mathbf{Q}$ and $[\mathbf{l}^1_\perp | \mathbf{l}^2_\perp]^T$, matrix $\mathbf{C}$ has to be factorised according to this equation.

If we choose α and $(\mathbf{l}^1_\perp, \mathbf{l}^2_\perp)$ freely then (4.10) partly determines the $\mathbf{Q}$ matrix. At the same time, (4.8) is equivalent to

$$\beta \hat{\mathbf{l}}^T = \mathbf{l}^T \mathbf{Q} \,, \qquad \beta > 0 \,. \tag{4.11}$$

Hence, to compute $\mathbf{Q}$ it is possible to proceed in the following way: We freely choose the orthonormal matrix $[\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}]$ which fixes the polariser coordinate system and the transformed light direction. Then we choose the parameters $\alpha \neq 0$, $\beta > 0$ arbitrarily. Concatenating (4.10) and (4.11) we get

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} [\mathbf{C} | \hat{\mathbf{l}}]^T = [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}]^T \mathbf{Q} \,, \qquad \begin{array}{l} \alpha \neq 0, \beta > 0 \\ [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}] \in O(3) \,, \end{array} \tag{4.12}$$

which enables to compute $\mathbf{Q}$ since all the other variables are either known or fixed:

$$\mathbf{Q} = [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}] \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} [\mathbf{C} | \hat{\mathbf{l}}]^T \,, \qquad \begin{array}{l} \alpha \neq 0, \beta > 0 \\ [\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}] \in O(3) \,. \end{array} \tag{4.13}$$

If we now ask what are the only transformations $\mathbf{A}$ that preserve the $\mathbf{l}$-tilt constraint associated with the r-th illuminant, this is equivalent to asking what matrices $\mathbf{A}$ may be applied from the left to $\mathbf{Q}$ such that the new solution corresponds to different

choices of the free variables α, β and $[\mathbf{l}_\perp^1 | \mathbf{l}_\perp^2 | \mathbf{l}] \in O(3)$. It turns out immediately that these will have the form

$$\mathbf{A} = [\bar{\mathbf{l}}_\perp^1 | \bar{\mathbf{l}}_\perp^2 | \bar{\mathbf{l}}] \begin{bmatrix} \alpha' & 0 & 0 \\ 0 & \alpha' & 0 \\ 0 & 0 & \beta' \end{bmatrix} [\mathbf{l}_\perp^1 | \mathbf{l}_\perp^2 | \mathbf{l}]^T , \qquad \begin{array}{l} \alpha', \beta' \neq 0 , \\ [\bar{\mathbf{l}}_\perp^1 | \bar{\mathbf{l}}_\perp^2 | \bar{\mathbf{l}}] \in O(3) , \end{array} \tag{4.14}$$

since this will correspond to a substitution $\alpha \mapsto \alpha\alpha'$, $\beta \mapsto \beta\beta'$ and $[\mathbf{l}_\perp^1 | \mathbf{l}_\perp^2 | \mathbf{l}] \mapsto [\bar{\mathbf{l}}_\perp^1 | \bar{\mathbf{l}}_\perp^2 | \bar{\mathbf{l}}]$ in (4.13). As the change $\beta' \mapsto -\beta'$ corresponds to the allowable change $\bar{\mathbf{l}} \mapsto -\bar{\mathbf{l}}$, we removed the condition $\beta' > 0$ and require only $\beta' \neq 0$. The result (4.14) is equivalent to (4.6) and the proof is complete. □

None of the **v**-tilt and **l**-tilt constraints alone removes the photometric stereo ambiguity completely. Combining one with the other, or one of them with an independent constraint may result in a various degree of ambiguity reduction. Our intention is to work primarily with constraints that can be widely applied and we therefore restrict our attention to three of them: the **v**-tilt, the **l**-tilt, and the integrability constraints.

What follows is the identification of classes of transformations that represent the ambiguity that remains after pairwise applications of the three constraints.

Theorem 4.3 *Let $\mathbf{l}^*$ be the true light vector of the r-th illuminant and let this vector is not collinear with the viewing direction, i.e., $\mathbf{v} \not\simeq \mathbf{l}^*$. Joint application of the* **v***-tilt constraint and the* **l***-tilt constraint associated with the r-th illuminant reduces the uncalibrated photometric stereo ambiguity into the composition of isotropic scaling and a convex/concave ambiguity:*

$$\mathbf{A} = \lambda \mathrm{diag}\,[s, s, 1] , \qquad \lambda \neq 0, s = \pm 1 . \tag{4.15}$$

Proof. To prove this statement, it is essential to show that the intersection of the two transformation classes (4.4) and (4.6) is given by (4.15). The proof is based on the fact that if two matrices are equal, then they have equal singular values[‡]. The singular values of the **l**-tilt preserving matrix (4.6) are obviously $\{|\alpha|, |\alpha|, |\beta|\}$, i.e. two of the singular values are equal. Let us check when the **v**-tilt preserving matrix (4.4) has a two-fold singular value. First, $\mathbf{A}^T\mathbf{A}$ for (4.4) is

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} \lambda^2 + \mu^2 & \mu\nu & \mu\tau \\ \mu\nu & \lambda^2 + \nu^2 & \nu\tau \\ \mu\tau & \nu\tau & \tau^2 \end{bmatrix} \qquad \begin{array}{l} \lambda, \tau \neq 0 , \\ \mu, \nu \in \mathbb{R} . \end{array} \tag{4.16}$$

The determinant of this matrix is $\det(\mathbf{A}^T\mathbf{A}) = \det(\mathbf{A})^2$ which, according to (4.4), is

$$\det(\mathbf{A}^T\mathbf{A}) = (\lambda^2\tau)^2 = \lambda^4\tau^2 . \tag{4.17}$$

[‡]Let $\mathbf{M}$ be a 3×3 matrix and let $\{d_1, d_2, d_3\}$ be the eigenvalues of $\mathbf{M}^T\mathbf{M}$. The singular values of $\mathbf{M}$ are $\{\sqrt{d_1}, \sqrt{d_2}, \sqrt{d_3}\}$.

We show that λ^2 is always an eigenvalue of (4.16). Indeed,

$$\begin{bmatrix} \mu^2 & \mu\nu & \mu\tau \\ \mu\nu & \nu^2 & \nu\tau \\ \mu\tau & \nu\tau & \tau^2 - \lambda^2 \end{bmatrix}, \qquad \begin{array}{l} \lambda, \tau \neq 0\,, \\ \mu, \nu \in \mathbb{R}\,, \end{array} \tag{4.18}$$

is always singular because the first two rows of this matrix are linearly dependent for any $\mu, \nu \in \mathbb{R}$ and $\tau \neq 0$. The matrix is required to have a two-fold eigenvalue. There are two possibilities for that:

- λ^2 is a one-fold eigenvalue and $\sqrt{\frac{\det(\mathbf{A}^T\mathbf{A})}{\lambda^2}} = |\lambda\tau|$ is at least two-fold eigenvalue[§]. This requires

$$\operatorname{rank}\left(\begin{bmatrix} \lambda^2 + \mu^2 - |\lambda\tau| & \mu\nu & \mu\tau \\ \mu\nu & \lambda^2 + \nu^2 - |\lambda\tau| & \nu\tau \\ \mu\tau & \nu\tau & \tau^2 - |\lambda\tau| \end{bmatrix}\right) \leq 1 \qquad \begin{array}{l} \mu, \nu \in \mathbb{R}, \\ \lambda, \tau \neq 0\,. \end{array} \tag{4.19}$$

  But this can be true only if $|\lambda| = |\tau|$ and $\mu = \nu = 0$ which is argued as follows.

  - $\mu, \nu \neq 0$. Then (4.19) holds if and only if

$$\begin{bmatrix} \frac{\lambda^2+\mu^2-|\lambda\tau|}{\mu} & \nu & \tau \\ \mu & \frac{\lambda^2+\nu^2-|\lambda\tau|}{\nu} & \tau \\ \mu & \nu & \frac{\tau^2-|\lambda\tau|}{\tau} \end{bmatrix} \tag{4.20}$$

    is of rank at most 1. By observation, this does not hold for any $\mu, \nu \neq 0$ and $\tau \neq 0$.

  - $\mu\nu = 0$ (at least one of μ, ν is zero). For $\mu = 0$, for example, (4.19) becomes

$$\begin{bmatrix} \lambda^2 - |\lambda\tau| & 0 & 0 \\ 0 & \lambda^2 + \nu^2 - |\lambda\tau| & \nu\tau \\ 0 & \nu\tau & \tau^2 - |\lambda\tau| \end{bmatrix}. \tag{4.21}$$

    By observation, this matrix is of rank at most 1 only if $|\lambda| = |\tau|$ and $\nu = 0$ (in that case, the rank is zero and $|\lambda\tau| = \lambda^2$ is a three-fold eigenvalue). The case $\nu = 0$ is shown analogously. The case $\mu, \nu = 0$ implies $|\lambda| = |\tau|$ again.

- λ^2 is at least a two-fold eigenvalue. This requires (4.18) to be of rank at most 1. By observation, this condition implies $\mu, \nu = 0$. Parameters $\lambda, \tau \neq 0$, on the other hand, may be arbitrary.

So far, we have the result that requiring the two eigenvalues of (4.16) to be the same implies

$$\mu = 0, \qquad \nu = 0\,. \tag{4.22}$$

[§]By saying 'at least two-fold' we take into account the fact that it could happen $|\lambda| = |\tau|$. In that case $|\lambda\tau|$ is a three-fold eigenvalue.

To complete the proof we must show $|\lambda| = |\tau|$ has to hold. Consider the true scaled normals $\mathbf{B}^*$ and the true scaled lights $\mathbf{S}^*$. The vector $\mathbf{l}^*$ from the theorem statement is given by $\mathbf{l}^* = {\mathbf{s}_r}^*/\|{\mathbf{s}_r}^*\|$. Let $({\mathbf{l}^*}^1_\perp, {\mathbf{l}^*}^2_\perp)$ denote the actual polariser coordinate system basis. The matrix (4.4) with $\mu = 0$ and $\nu = 0$ represents an overall scaling plus extra scaling in the direction of $\mathbf{v}$. In contrast, (4.6) represents an overall scaling plus extra scaling in the direction of $\mathbf{l}^*$, followed by an orthogonal transformation. As $\mathbf{v} \not\simeq \mathbf{l}^*$, the extra scalings act along different directions and the necessary condition for the two mappings to be identical is that $|\alpha| = |\beta| = |\lambda| = |\tau|$. We observe that:

- For $\lambda = \tau$, and taking (4.22) into account, (4.4) becomes

$$\mathbf{A} = \lambda \mathrm{diag}\,[1, 1, 1]\,, \qquad \lambda \neq 0. \tag{4.23}$$

  This is a scaled identity which is a member of transformations (4.6) as well. This is because in the case considered $[\,\mathbf{l}^1_\perp | \mathbf{l}^2_\perp | \mathbf{l}\,] = [\,{\mathbf{l}^*}^1_\perp | {\mathbf{l}^*}^2_\perp | \mathbf{l}^*\,]$ in (4.6), and it suffices to set $\alpha = \lambda$, $\beta = \lambda$ and $[\,\bar{\mathbf{l}}^1_\perp | \bar{\mathbf{l}}^2_\perp | \bar{\mathbf{l}}\,] = [\,{\mathbf{l}^*}^1_\perp | {\mathbf{l}^*}^2_\perp | \mathbf{l}^*\,]$ to make (4.6) equal to (4.23).

- For $\lambda = -\tau$, and taking (4.22) into account, (4.4) becomes

$$\mathbf{A} = \lambda \mathrm{diag}\,[-1, -1, 1]\,, \qquad \lambda \neq 0. \tag{4.24}$$

  This is a transformation is a composition of isotropic scaling by λ and reflection of the first two coordinates. It is a member of (4.6) as well: it suffices to set all parameters as above except $[\,\bar{\mathbf{l}}^1_\perp | \bar{\mathbf{l}}^2_\perp | \bar{\mathbf{l}}\,] = \mathrm{diag}\,[-1, -1, 1]\,[\,{\mathbf{l}^*}^1_\perp | {\mathbf{l}^*}^2_\perp | \mathbf{l}^*\,]$. This concludes the proof. □

Theorem 4.4 *Let $\mathbf{l}^*$ be the true light vector of the r-th illuminant and let this vector is not collinear with the viewing direction, i.e., $\mathbf{v} \not\simeq \mathbf{l}^*$. Joint application of the integrability constraint and the $\mathbf{l}$-tilt constraint associated with the r-th illuminant reduces the uncalibrated photometric stereo ambiguity into the composition of isotropic scaling and a convex/concave ambiguity:*

$$\mathbf{A} = \lambda \mathrm{diag}\,[s, s, 1]\,, \qquad \lambda \neq 0, s = \pm 1\,. \tag{4.25}$$

Proof. To prove this statement, it is essential to show that the intersection of the class (4.6) and the class of generalised bas-relief transformations (1.38) which preserve the integrability constraint is given by (4.25). Note that for given $\lambda, \tau \neq 0$ and $\mu, \nu \in \mathbb{R}$ the $\mathbf{v}$-tilt preserving transformation (4.4) and the GBR transformation (1.38) are just a transponse of each other. Employing the first part of the proof of Theorem 4.3 on transponses of (1.38) and (4.6) gives the result that it must hold $\mu, \nu = 0$ in (1.38). The rest of the proof of Theorem 4.3 then makes the proof of this theorem. □

Theorem 4.5 *Joint application of the* **v***-tilt and integrability constraints results in reduction of the uncalibrated photometric stereo ambiguity to a bas-relief ambiguity represented by transformations*

$$\mathbf{A} = \mathrm{diag}\left[\lambda, \lambda, \tau\right] , \qquad \lambda, \tau \neq 0 \,. \tag{4.26}$$

Proof. Obviously, the intersection of classes (4.4) and (1.38) is represented by matrices in (4.26). □

This theorem concludes the theoretical analysis of constraints studied in this chapter. The next section presents the results of application of these constraints in a real experiment.

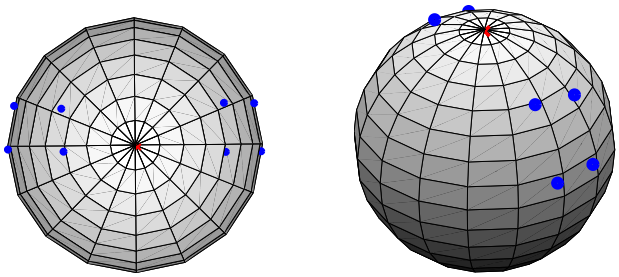
4.3 Experiment

The object we used to test the applicability of the theoretical results was a glazed china teapot. The teapot was illuminated from eight directions in turn (see Fig. 4.1; the ground truth positions of light sources were measured by a geometric stereo-based method). For each light source position, we grabbed two sets of images by a 14-bit camera. The first and second sets were formed by images taken for six orientations of the camera polariser and for six orientations of the light polariser, respectively. Both the camera polariser and the light polariser were rotated in 30° steps. For each illumination we computed the **v**-tilt and **l**-tilt of every illuminated surface patch by applying the harmonic fit to the first and second set of images, respectively, in a pixel-wise manner. An example of the **v**-tilt and the **l**-tilt determined for the teapot illuminated from the right is shown in Fig. 3.8. Only tilts whose partial polarisation exceeded 3% were trusted. If a given surface patch was illuminated by more than one light during the measurement, there were multiple evaluations of **v**-tilt available from which the one with highest magnitude of harmonic brightness variation was kept. The tilts of specularly reflecting surface patches were marked as invalid. Specularities were identified by simple thresholding.

Figure 4.2 shows the photometric stereo input images. Image of the teapot illuminated by a given illumination was computed by averaging all images grabbed under that illumination. This gave us the image intensity matrix (1.12). To compute the scaled normals matrix **B** and the scaled lights matrix **S**, the data was processed in the following consecutive steps:

1. The elements of the image intensity matrix were marked as missing if their intensity fell below or over respective thresholds. This simple procedure identified points accommodating a shadow or a specularity. By that, the stacked image matrix **J** (1.16) was acquired. The Lambertian-consistent masks resulting from this step are shown in Fig. 4.3.

2. The stacked image matrix **J** was used to compute the scaled lights matrix **S** by Jacobs algorithm [18] for factorisation with missing data.



(a) (b)

Figure 4.1: (a) The directions of eight light sources, displayed on a sphere from the camera point of view. The viewing direction shown as the central dot. (b) The same, rotated for better illustration.

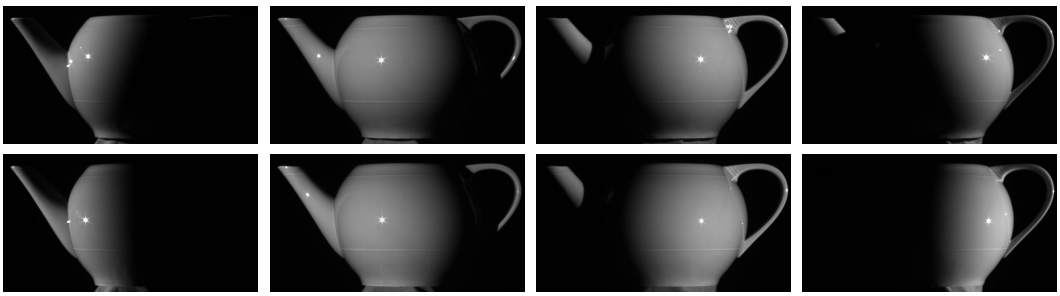


Figure 4.2: The images of a teapot under a point light source illuminations from eight different directions.

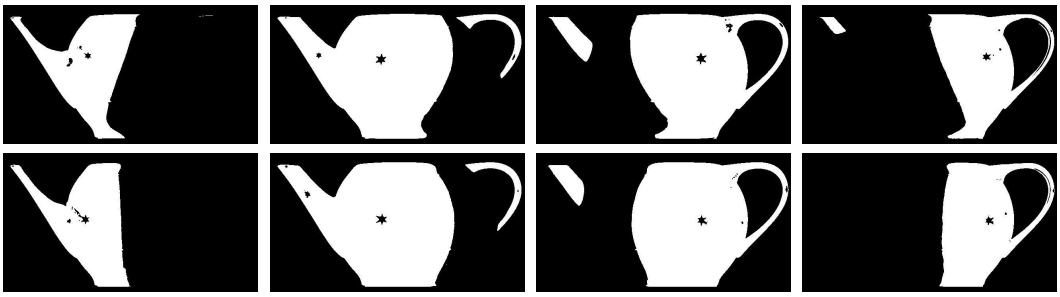


Figure 4.3: Lambertian-consistent masks for the input dataset. The non-Lambertian pixels (shadowed or specular) were determined by simple thresholding.

(a)

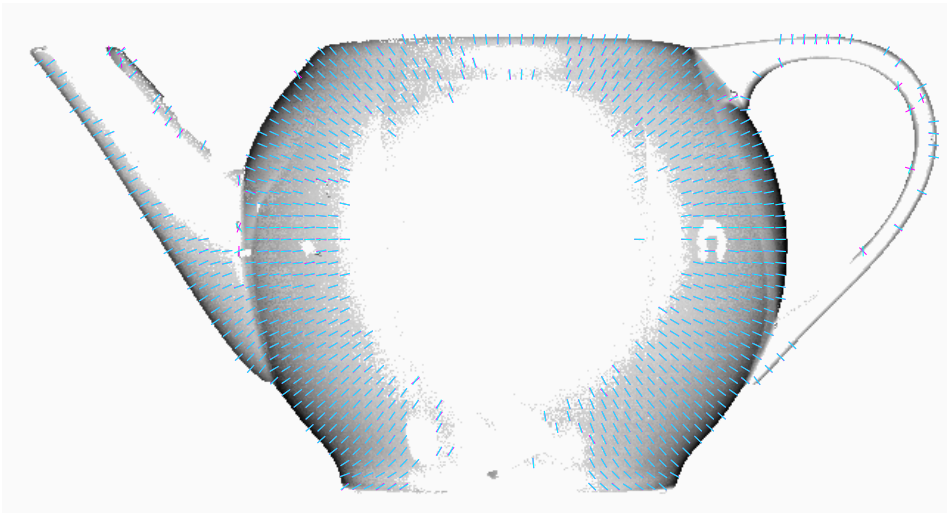
(b)

Figure 4.4: The normals obtained by factorisation of the Lambertian portion of the input data. The first image row (a) shows the normals relit by lights of equal light strengths from the left, top, and viewpoint directions (first three) and the recovered albedo (the last one). The result is obtained by the Jacobs method [18]. The second image row (b) shows the same for the data refined by a nonlinear optimisation.

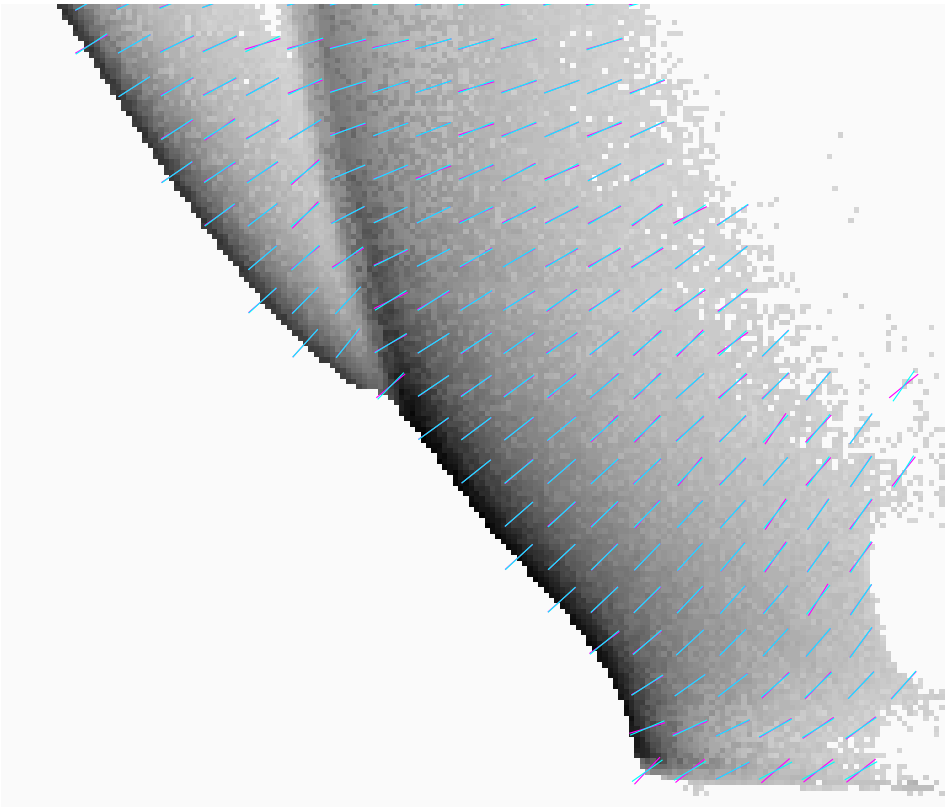
3. The scaled normals matrix $\mathbf{B}$ was computed using the stacked image matrix $\mathbf{J}$ and the scaled lights matrix $\mathbf{S}$ obtained in the previous step. Each normal was fit individually by least-squares. The result of this step is shown in Fig. 4.4(a). After that, scaled lights were re-evaluated by an analogous procedure, and this iterative process (alternating between re-computation of scaled normals and scaled lights as in [13]) was repeated 10 times. The residual (the sum of squared differences between the predicted and observed intensities over the valid image regions) converged to about 1/2 of its initial value. The factorisation solution refined by the non-linear optimisation procedure is shown in Fig. 4.4(b).

Note that the relit normals shown in Fig. 4.4(b) do not correspond to how the normals illuminated from the directions specified look in reality. Similar observation is true about the albedo: In reality, the teapot albedo is uniform. These facts are in agreement with the uncalibrated photometric stereo ambiguity up to which the surface parameters are recovered. In order to reduce this ambiguity, different sets of constraints were applied subsequently. The results are shown in Fig. 4.6.

- **Integrability constraint**. The integrability constraint was employed first, using the algorithm published in [44] (see the first image row in Fig. 4.6). We observe that the normals relit by the light of constant light strength from three directions (the retinal plane axes directions and the camera direction) do not correspond to the appearance of a teapot illuminated from the right, top and the viewpoint direction. The albedo is not correct either. This is consistent with the remaining ambiguity (1.38).



(a)



(b)

Figure 4.5: The agreement of the measured **v**-tilts (shown in green) and the projections of the scaled normals made consistent with the **v**-tilt constraint onto the retina (magenta). The entire object (a), and its bottom-left part (b).

- **v-tilt constraint**. The **v**-tilt constraint was applied by solving (4.5) for **Q**. The results are shown in the second row of Fig. 4.6. The normals relit from the retinal axes directions look correctly, but this is not true about the normals relit from the viewpoint direction nor about the albedo. This is consistent with the remaining ambiguity represented by (4.4).

  Figure 4.5 shows the measured **v**-tilts and the normals resulting from this step projected onto the retinal plane. The two entities match very well in the majority of image points.

- **v-tilt + integrability constraints**. We subsequently let $\lambda = 1$ in (4.4) and found μ,ν and τ such that the normals consistent with the **v**-tilts obeyed the integrability constraint. This was done by finding $\mathbf{Q}_3 = (\mu, \nu, \tau)$ in (1.30) by linear least-squares for known $\mathbf{Q}_1 = (1, 0, 0)$ and $\mathbf{Q}_2 = (0, 1, 0)$.

  The result is shown in the third row. The first two images exhibit correct shading. The albedo is not uniform, but appears 'centered'. This is consistent with the remaining ambiguity (4.26) which scales the scaled normal vectors anisotropically in a direction of the viewpoint. The normals relit from the viewing direction look flat.

- **l-tilt + v-tilt constraints**. The **l**-tilt associated with a given illuminant was applied first. This was done by solving (4.7) by linear least-squares. Parameters α, β and $[\bar{\mathbf{l}}^1_\perp|\bar{\mathbf{l}}^2_\perp|\bar{\mathbf{l}}]$ for (4.6) were then found such that (4.6) transformed the normals consistent with the **l**-tilts to those obeying the **v**-tilt constraint as well. This was done using a simple non-linear optimisation.

  We tried **l**-tilt constraints associated with different illuminants to be combined with the **v**-tilt constraint. This necessarily leads to different results. Two of them are shown in the fourth and fifth rows of Fig. 4.6¶. In order to eliminate the dependence on selected illuminant, we applied **l**-tilt constraints associated with all illuminants (hence eight constraints) together with the **v**-tilt constraint. This was done by non-linear optimisation, with the optimization function defined as squared sum of angular differences between measured tilts and tilts evaluated by projection of normals onto respective planes. The result is shown in sixth row of Fig. 4.6. The appearance of the relit normals corresponds is correct. The same is true about the albedo.

- **l-tilt + integrability constraints**. Similarly, a nonlinear optimisation was employed to obtain the normals consistent with the **l**-tilt constraints associated with all illuminants and the integrability constraint. The result is shown in the last row of Fig. 4.6. Both the normals and albedo are recovered correctly as in the previous case.

To further illustrate the accurateness of determination of normals under different sets of constraints, the shadow boundaries are shown in Fig. 4.7. Note that in all

¶The difference between the results may be due to imperfect experimental settings or slight discrepancy between assumed separability of the model of reflectance and the real object reflectance.

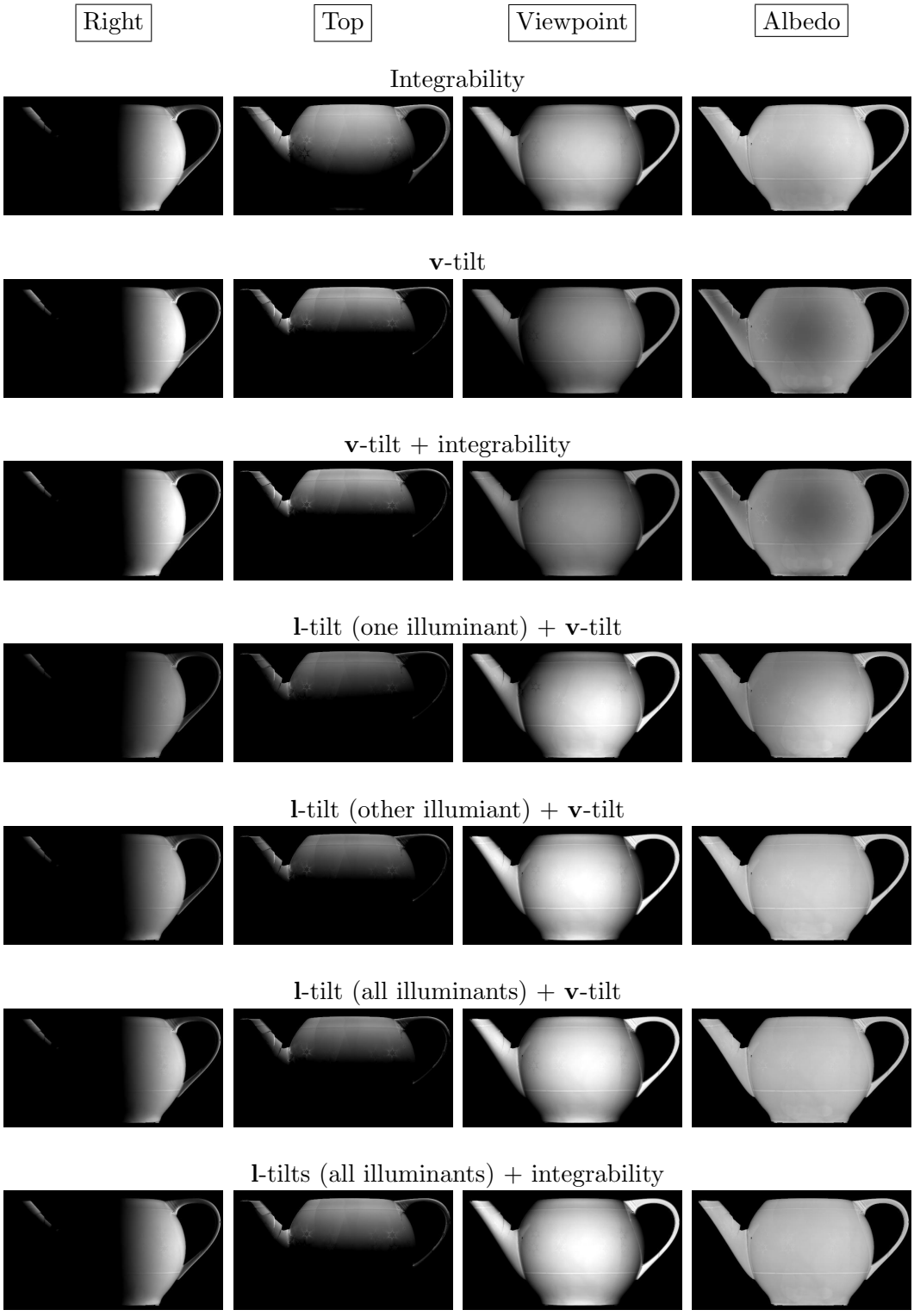


Figure 4.6: The results of normal transformations conforming to the constraints that are written above each image row. The images show the computed normals relit by illuminant of constant light strength from directions of the first and second retinal plane axes and the viewing direction, respectively. Finite ambiguities were resolved by hand.

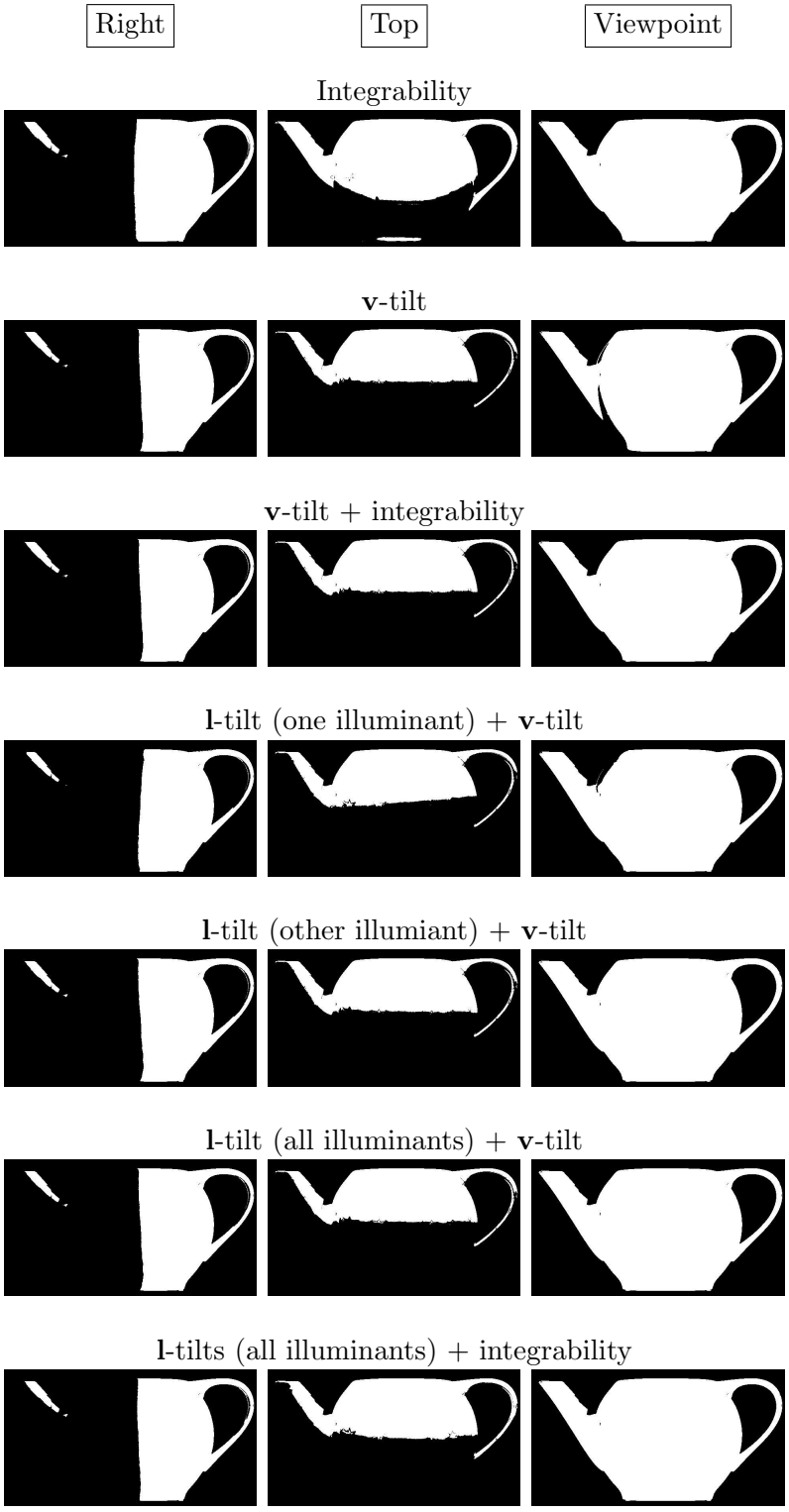


Figure 4.7: Shadow boundaries of normals illuminated as in the previous figure.

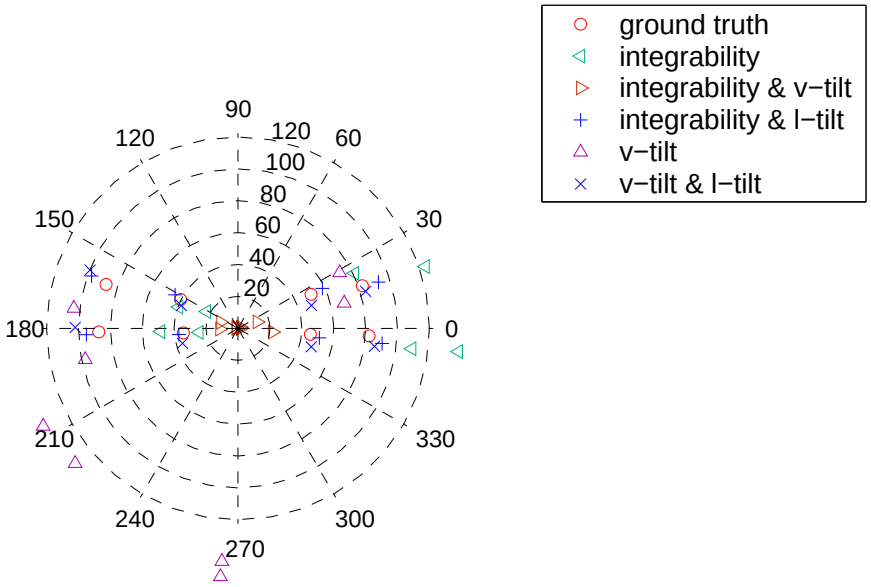


Figure 4.8: Illumination directions estimated under different sets of constraints imposed on normals. Directions are displayed in spherical coordinates. Only the combination of the **v**-tilt and **l**-tilt constraints, as well as of the integrability and **l**-tilt constraints determines the illumination directions correctly.

cases, finite ambiguities (e.g., the convex-concave ambiguity) were resolved by hand in order the results could be shown in a consistent way.

To summarize, only the **v**-tilt and **l**-tilt constraints combined, or the integrability and **l**-tilt constraints combined can recognize the normals *and* the lights correctly, as expected. In Fig. 4.8 we show directions of the eight light sources estimated by the five methods together with the ground truth.

4.4 Summary of the chapter

This chapter showed that making the solution of uncalibrated photometric stereo consistent with the **v**-tilt and **l**-tilt constraints results in reducing the uncalibrated photometric stereo ambiguity to isotropic scaling composed with a convex-concave ambiguity. The same result is obtained if the **l**-tilt constraint is imposed together with the integrability constraint but such approach requires an explicit assumption about the surface continuity.

To recover tilts we employed partial polarisation of light reflected on smooth dielectric surface. Previous work exploiting polarisation vision to acquire 3D shape used only the fact that the *specular* reflection component is polarised [31, 38]. The polarisation of diffusely reflected light is too weak to recover useful dense information about the object shape but, as we showed, may be used to provide disambiguation constraints for uncalibrated photometric stereo.

Chapter 5

Presence of specular reflectance component

In this chapter we show that if the object reflectance is a sum of Lambertian reflectance and mirror-like reflectance, then the original ambiguity represented by a group $GL(3)$ reduces into a two-parametric group of transformations. These transformations are compositions of isotropic scaling (1dof), rotation around the viewing vector (1dof), and change in the global coordinate frame handedness (two-fold ambiguity). This ambiguity reduction is implied by a condition that all lights reflected around corresponding specular normals must give the viewing direction. We call this condition the *consistent viewpoint constraint*. We show that specularities in as few as *four* images corresponding to four different point lights in general configuration are sufficient to utilize the consistent viewpoint constraint.

By this result, we make a step towards uncalibrated photometric stereo for objects whose reflectance includes not only body (diffuse) component, but also interface (specular) component. Such composite reflectance models are certainly not new to photometric stereo applications, see e.g. [6, 33, 27, 24], but in those methods the illumination conditions are supposed to be known.

The specific representative of composite reflectance model (the superposition of Lambertian and mirror-like reflectance) is selected because as specularities are sparse in the images, they can be treated as *outliers* to the Lambertian reflectance model. This gives us a possibility to study the problem as Lambertian photometric stereo with additional information represented by the consistent viewpoint constraint.

5.1 Consistent viewpoint constraint

It is well understood that specularities occur at loci where light reflects on a smooth surface in a mirror-like manner towards the observing sensor. Hence, mirror-like reflection carries valuable information about the geometrical configuration of the viewing vector, the light vector, and the surface normal: if a specularity is observed in an image, then at the corresponding surface point, surface normal is a bisector of the viewing and illumination vectors (see Fig. 3.4). For the viewing vector there

holds

$$\mathbf{v} = 2(\mathbf{l}\cdot\mathbf{n}^S)\mathbf{n}^S - \mathbf{l} = \frac{2\left(\mathbf{s}\cdot\mathbf{b}^S\right)\mathbf{b}^S}{\|\mathbf{s}\|\|\mathbf{b}^S\|^2} - \frac{\mathbf{s}}{\|\mathbf{s}\|}\,, \tag{5.1}$$

where $\mathbf{n}^S$ is a normal which induces the occurrence of specularity in an image while illuminated from the direction $\mathbf{l}$. The right-most part of the equation essentially states the same fact in "natural" photometric stereo variables. Without the risk of confusion we call both $\mathbf{n}^S$ and $\mathbf{b}^S$ the *specular normals*, and the pair of vectors $(\mathbf{l}, \mathbf{n}^S)$ as well as of vectors $(\mathbf{s}, \mathbf{b}^S)$ we call a *specular pair*.

The equation may be viewed as a formula for computing viewpoint direction from known light $\mathbf{s}$ and specular normal $\mathbf{b}^S$. The key fact this relation states is the following: No matter which specular pair is used for viewing direction evaluation, all give the same result.

We define the **consistent viewpoint constraint** as follows. Let $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ be the solution to (1.17). Let the viewing vector be denoted by $\mathbf{v}$. For each scaled light vector $\mathbf{s}_r$, $r \in \{1, 2, .., R\}$, let $\mathbf{b}_r^S$ be the scaled normal vector at an image point in which the specularity is observed. Let $\mathcal{P}$ be a set of all such specular pairs. We say that the solution $(\mathbf{B}, \mathbf{S})$ obeys the consistent viewpoint constraint if (5.1) holds for all specular pairs $(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}$.

5.1.1 Transformations which preserve the consistent viewpoint constraint

Consider the scaled normals matrix and the scaled lights matrix $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ representing a solution to (1.17). Suppose that this particular solution obeys the consistent viewpoint constraint. We will analyse what are the only transformations $\mathbf{A} \in GL(3)$ which transform $\mathbf{B}$ and $\mathbf{S}$ according to (1.21) such that the consistent viewpoint constraint is preserved. The analysis is led as follows: We first derive an equivalent of (5.1) for transformed scaled normals and scaled lights (5.7). Then we show in Lemma 5.1 that the consistent viewpoint constraint implies that the matrix $\mathbf{P}$ in (5.7) may be only a positively scaled identity. Utilising this observation enables us to prove the principal statement formulated in Theorem 5.1 (page 65).

First, let us write the equivalent of (5.1) for the transformed lights and normals:

$$\mathbf{v} = \frac{2\left[(\mathbf{A}^{-T}\mathbf{s})\cdot(\mathbf{A}\mathbf{b}^S)\right]\mathbf{A}\mathbf{b}^S}{\|\mathbf{A}^{-T}\mathbf{s}\|\|\mathbf{A}\mathbf{b}^S\|^2} - \frac{\mathbf{A}^{-T}\mathbf{s}}{\|\mathbf{A}^{-T}\mathbf{s}\|}\,. \tag{5.2}$$

Left-multiplying both sides of the equation by $\|\mathbf{A}^{-T}\mathbf{s}\|\|\mathbf{A}\mathbf{b}^S\|^2\mathbf{A}^T$ we obtain

$$\|\mathbf{A}^{-T}\mathbf{s}\|\|\mathbf{A}\mathbf{b}^S\|^2\mathbf{A}^T\mathbf{v} = 2[(\mathbf{A}^{-T}\mathbf{s})\cdot(\mathbf{A}\mathbf{b}^S)]\mathbf{A}^T\mathbf{A}\mathbf{b}^S - \|\mathbf{A}\mathbf{b}^S\|^2\mathbf{s}\,. \tag{5.3}$$

With substitutions

$$\alpha(\mathbf{s}, \mathbf{b}^S) = \|\mathbf{A}^{-T}\mathbf{s}\|\|\mathbf{A}\mathbf{b}^S\|^2\|\mathbf{A}^T\mathbf{v}\|\,, \tag{5.4}$$

$$\mathbf{w} = \mathbf{A}^T\mathbf{v}/\|\mathbf{A}^T\mathbf{v}\|\,, \tag{5.5}$$

$$\mathbf{P} = \mathbf{A}^T\mathbf{A} \tag{5.6}$$

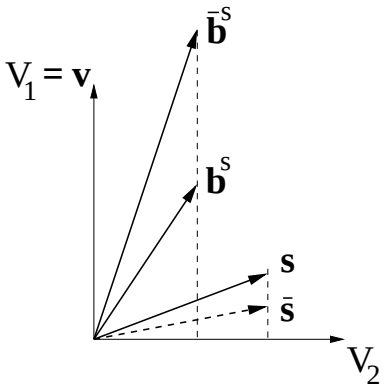


Figure 5.1: Scaling the $\mathbf{V}_1$ component of normals by λ_1 while scaling the same component of lights by $1/\lambda_1$ results in that the new specular pair violates consistent viewpoint constraint if $|\lambda_2| \neq |\lambda_1|$.

this equation is rewritten as

$$\boxed{\alpha(\mathbf{s}, \mathbf{b}^S)\mathbf{w} = 2(\mathbf{s} \cdot \mathbf{b}^S)\mathbf{P}\mathbf{b}^S - (\mathbf{b}^S \cdot \mathbf{P}\mathbf{b}^S)\mathbf{s}\,,} \tag{5.7}$$

where we applied the fact that $(\mathbf{A}^{-T}\mathbf{s}) \cdot (\mathbf{A}\mathbf{b}^S) = \mathbf{s} \cdot \mathbf{b}^S$, and rewrote the factor $\|\mathbf{A}\mathbf{b}^S\|^2$ as $\|\mathbf{A}\mathbf{b}^S\|^2 = (\mathbf{b}^S \cdot \mathbf{A}^T\mathbf{A}\mathbf{b}^S) = (\mathbf{b}^S \cdot \mathbf{P}\mathbf{b}^S)$. Having sufficient number of specular pairs, (5.7) represents a set of nonlinear equations. We make the following observation: Matrix $\mathbf{P}$ in (5.7) can only be a posively scaled identity, $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$. This fact is formulated in the following lemma.

Lemma 5.1 *Let $\mathcal{P}$ be a set of specular pairs which obey the consistent viewpoint constraint, i.e., for any specular pair $(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}$ there holds (5.1). Let this set be* sufficient *in a sense that adding more specular pairs to $\mathcal{P}$ by no way affects the solution of the system of equations (5.7). Then, the only matrices $\mathbf{P}$ which the solution to (5.7) may comprise are $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$, where $\mathbf{I}$ is the identity.*

Proof of the Lemma. The current version of the proof requires the set of specular pairs $\mathcal{P}$ to be *maximal*. This means that: i) the object is illuminated from all directions $\mathbf{l} \in \mathcal{S}$, where $\mathcal{S}$ is a unit sphere in $\mathbb{R}^3$, and ii) the object is required to accommodate all surface orientations $\mathbf{n} \in \mathcal{S}^{1/2}_{\mathbf{v}}$, where $\mathcal{S}^{1/2}_{\mathbf{v}} = \{\mathbf{x} \in \mathcal{S}\colon \mathbf{v} \cdot \mathbf{x} \geq 0\}$ is a unit half-sphere of all orientations visible from the viewing direction $\mathbf{v}$. In such case, for each normal $\mathbf{n} \in \mathcal{S}^{1/2}_{\mathbf{v}}$ there exists the light vector $\mathbf{l} \in \mathcal{S}$ under which this normal is specular, such that $(\mathbf{l}, \mathbf{n})$ form a specular pair. For sake of notational consistency with (5.7), we set $\mathbf{b}^S = \mathbf{n}$ and $\mathbf{s} = \mathbf{l}$ and denote this pair $(\mathbf{s}, \mathbf{b}^S)$. Let $\mathcal{P} = \{(\mathbf{s}, \mathbf{b}^S)\colon \mathbf{b}^S \in \mathcal{S}^{1/2}_{\mathbf{v}},\ \mathbf{s} = 2(\mathbf{v} \cdot \mathbf{b}^S)\mathbf{b}^S - \mathbf{v}\}$ be the set of all specular pairs. To deliver the proof of the Lemma, let us first observe the coefficients of vectors in (5.7):

$$\alpha(\mathbf{s}, \mathbf{b}^S) \neq 0\colon \quad \forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}\,, \tag{5.8}$$

$$(\mathbf{b}^S \cdot \mathbf{P}\mathbf{b}^S) \neq 0\colon \quad \forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}\,, \tag{5.9}$$

$$(\mathbf{s} \cdot \mathbf{b}^S) \neq 0\colon \quad \forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}'\,, \quad \mathcal{P}' = \mathcal{P}\backslash\{(\mathbf{s}', \mathbf{b}'^S) \in \mathcal{P}\colon \mathbf{v} \cdot \mathbf{b}'^S = 0\}\,. \tag{5.10}$$

The result (5.8) is obvious from (5.4) and relation (5.9) follows from that $\mathbf{P}$ is positively definite. The last result (5.10) holds because $\mathbf{s} \cdot \mathbf{b}^S = 0$ if and only if the specular normal $\mathbf{b}^S$ is located at the occluding boundary and illuminated by a light source of direction $\mathbf{s} = -\mathbf{v}$. From this observation we have that the necessary conditions for the validity of (5.7) are

$$\forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}: \qquad \mathbf{w} \in span(\mathbf{P}\mathbf{b}^S, \mathbf{s})\,, \tag{5.11}$$

$$\forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}': \qquad \mathbf{P}\mathbf{b}^S \in span(\mathbf{w}, \mathbf{s})\,. \tag{5.12}$$

The matrix $\mathbf{P}$ is symmetric and positive definite and thus its effect on $\mathbf{b}^S$ in (5.7) represents anisotropic scaling in certain three orthogonal directions. To see that, let us write the SVD decompositions of $\mathbf{A}$ and $\mathbf{P}$:

$$\mathbf{A} = \mathbf{U}\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)\mathbf{V}^T \quad \lambda_1, \lambda_2, \lambda_3 > 0; \quad \mathbf{U}, \mathbf{V} \in O(3)\,, \tag{5.13}$$

$$\mathbf{P} = \mathbf{V}\mathrm{diag}(\lambda_1^2, \lambda_2^2, \lambda_3^2)\mathbf{V}^T\,. \tag{5.14}$$

Thus $\mathbf{P}$ scales along the direction of eigenvector $\mathbf{V}_i$ (i-th column of $\mathbf{V}$) by the respective λ_i^2:

$$\mathbf{P}\mathbf{V}_i = \lambda_i^2 \mathbf{V}_i\,, \tag{5.15}$$

which means that the normals $\mathbf{b}_i^S = \pm\mathbf{V}_i$ (where $\pm$ is properly selected such that the selected choice is in $\mathcal{S}_{\mathbf{v}}^{1/2}$) are mapped, up to a scale, onto themselves. Let the light vectors under which these normals are specular be denoted $\mathbf{s}_i$. Then (5.11) states that for the normals $\mathbf{b}_i^S = \pm\mathbf{V}_i$ it holds $\mathbf{w} \in span(\mathbf{b}_i^S, \mathbf{s}_i)$, $i = \{1, 2, 3\}$ which implies

$$\mathbf{w} \in span(\mathbf{b}_1^S, \mathbf{s}_1) \cap span(\mathbf{b}_2^S, \mathbf{s}_2) \cap span(\mathbf{b}_3^S, \mathbf{s}_3) \simeq \mathbf{v}\,. \tag{5.16}$$

In the same time, (5.12) must hold and we therefore require

$$\forall(\mathbf{s}, \mathbf{b}^S) \in \mathcal{P}': \quad \mathbf{P}\mathbf{b}^S \in span(\mathbf{v}, \mathbf{s}) = span(\mathbf{v}, \mathbf{b}^S)\,, \tag{5.17}$$

where the latter equality $span(\mathbf{v}, \mathbf{s}) = span(\mathbf{v}, \mathbf{b}^S)$ follows from coplanarity of the viewing, normal and light directions in specular configuration. The equation (5.17) says that the transformed normal $\mathbf{P}\mathbf{b}^S$ must be mapped into the plane given by the viewing direction $\mathbf{v}$ and the original normal $\mathbf{b}^S$. The only way to achieve this is to align one of the scaling directions (say, $\mathbf{V}_1$) with the viewing direction $\mathbf{v}$, and to set the scalings along the other two directions equal ($\lambda_2^2 = \lambda_3^2$). To conclude the proof, we need to show that it must hold that $\lambda_1^2 = \lambda_2^2 = \lambda_3^2$. We make that by showing the effect of transformation $\mathbf{A}$ on the normals and of transformation $\mathbf{A}^{-T}$ on the light vectors.

First, let us complete (5.13) and (5.14) by writing the decomposition of $\mathbf{A}^{-T}$:

$$\mathbf{A}^{-T} = \mathbf{U}\mathrm{diag}(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3})\mathbf{V}^T\,. \tag{5.18}$$

A particular choice of orthonormal matrix $\mathbf{U}$ has no fundamental effect on the validity of the consistent viewpoint constraint (it corresponds to rotating all specular pairs, and hence also the viewing direction $\mathbf{v}$, by a global rotation). Let us show the

effect of transforming the normals and light vectors for a case when $\mathbf{U} = \mathbf{I}$. Figure 5.1 shows a specular pair $(\mathbf{s}, \mathbf{b}^S)$ lying in a plane given by first two eigenvectors, and the result of its transformation into $(\bar{\mathbf{s}}, \bar{\mathbf{b}}^S)$ by (1.21). The situation is shown for a particular choice $\lambda_1 = 2$, $\lambda_2 = 1$ and $\mathbf{V}_1 = \mathbf{v}$. The transformed normal $\bar{\mathbf{b}}^S$ makes a smaller angle with $\mathbf{v}$ as compared to the original normal, while the transformed light $\bar{\mathbf{s}}$ makes a greater angle with $\mathbf{v}$ as compared with the original light. From that it follows that for transformed normals and lights $\mathbf{v}$ is not a consistent viewpoint. The necessary condition for $\mathbf{v}$ to be the consistent viewpoint, is $|\lambda_1| = |\lambda_2|$. As it was discussed above that $\lambda_2^2 = \lambda_3^2$, we have the result that $\mathbf{P}$ may only be the scaled identity $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$. □

Theorem 5.1 *Let* $(\mathbf{B}, \mathbf{S}) \in \mathcal{T}(\mathbf{J})$ *be the solution to (1.17). Let the viewing vector be denoted by* $\mathbf{v}$. *Let* $(\mathbf{B}, \mathbf{S})$ *obey the consistent viewpoint constraint and let* $\mathcal{P}$ *be the set of specular pairs extracted from the image data. Let this set be* sufficient *in the same sense as in Lemma 5.1. The only linear invertible transformations* $\mathbf{A}$ *which transform* $(\mathbf{B}, \mathbf{S})$ *according to (1.21) such that the consistent viewpoint constraint is preserved are*

$$\mathbf{A} = \lambda \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{R}_z(\xi)\,, \qquad \xi \in \langle 0, 2\pi \rangle\,, \lambda > 0\,. \tag{5.19}$$

In this equation, $\mathbf{R}_z(\xi)$ *stands for rotation around the z-axis (the viewing direction) by angle* ξ. *The ambiguity in sign of the first coordinate corresponds to reflecting the transformed vectors by the* $y - z$ *plane (and hence to the change in the coordinate frame handedness).*

Proof. It has been demonstrated that (5.1) is equivalent to (5.7). We will show that in (5.7), the vector $\mathbf{w}$ is determined uniquely, and it holds that $\mathbf{w} = \mathbf{v}$. The statement about uniqueness holds because by Lemma 5.1 we have $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$. Hence the right-hand side of (5.7) gives a vector the magnitude of which is unknown (due to scaling ambiguity of $\mathbf{P}$) but the orientation of which is unique (as $\mathbf{P}$ is determined up to a positive scale). The factors $\alpha(\mathbf{s}, \mathbf{b}^S)$ on a left-hand side are always positive. This implies that both the direction and orientation of $\mathbf{w}$ are determined uniquely. For any particular choice of $\lambda \neq 0$, Equation (5.7) is written as

$$\alpha(\mathbf{s}, \mathbf{b}^S)\mathbf{w} = 2\lambda(\mathbf{s} \cdot \mathbf{b}^S)\mathbf{b}^S - \lambda(\mathbf{b}^S \cdot \mathbf{b}^S)\mathbf{s}\,.$$

Dividing it by $\lambda\|\mathbf{s}\|\|\mathbf{b}^S\|^2$ gives the right-most part of Equation (5.1) which is equal to $\mathbf{v}$. Hence $\mathbf{w} = \mathbf{v}$.

Next, according to (5.6), $\mathbf{A}$ is given by $\mathbf{P} = \mathbf{A}^T\mathbf{A}$. Due to Lemma 5.1, $\mathbf{P} = \lambda^2\mathbf{I}$, $\lambda \neq 0$, and this implies

$$\mathbf{A} = \lambda\mathbf{O}\,, \qquad \mathbf{O} \in O(3),\, \lambda \neq 0\,. \tag{5.20}$$

In addition we showed that $\mathbf{w} = \mathbf{v}$. Using this fact together with (5.20), Equation (5.5) becomes

$$\mathbf{v} = \frac{\lambda\mathbf{O}^T\mathbf{v}}{\lambda\|\mathbf{O}^T\mathbf{v}\|} = \mathbf{O}^T\mathbf{v}\,. \tag{5.21}$$

This equation represents further constraint on $\mathbf{A} = \lambda\mathbf{O}$: The only orthogonal transformations allowable within this equality are only those which fix the viewpoint direction $\mathbf{v}$. These can be only rotations around $\mathbf{v}$ or reflections around planes containing $\mathbf{v}$. Taken together with the scaling, this gives (5.19). □

5.1.2 Four specular pairs establish the constraint

Let us observe how many specular pairs are needed to establish the consistent viewpoint constraint. The original ambiguity $GL(3)$ comprises 9dof. The ambiguity represented by matrices (5.19) comprises 2dof. Hence the number of degrees of freedom to fix is seven. For each specular pair, Equation (5.7) represents three scalar equations. After eliminating the scale-factor $\alpha(\mathbf{s}, \mathbf{b}^S)$, there remain two independent equations per specular pair. Hence, our expectation is that:

- At least *four* specular pairs in general configuration are needed to establish the consistent viewpoint constraint,
- As four specular pairs give 8 constraints on 7dof, the system of equations to be solved will be over-constrained.

We performed numerical experiments which support this observation. So far, we have not analyzed which configurations of four specular pairs are singular, nor the problem of (possible) finite solution multiplicity for non-singular configurations. However, in experiments we did observe unique solution in all cases. Analysis of which sets of four specular pairs give well-conditioned solution is the matter of future research.

5.2 Combinations with other constraints

The consistent viewpoint constraint fixes the photometric stereo ambiguity up to a composition of isotropic scaling (1dof), the rotation around the viewing direction (1dof) and the change in coordinate frame handedness (binary ambiguity). The consistent viewpoint constraint can be combined with other constraints in order to further reduce the ambiguity.

We decided to use the integrability constraint for resolving the remaining ambiguity. Integrability constraint fixes both the coordinate frame handedness *and* the rotation angle ξ modulo π. This implies that the resulting ambiguity is convex/concave ambiguity composed with isotropic scaling. This result follows from the fact that the intersection of the scaled orthogonal group $O(3)$ (of which transformations (5.19) are a sub-group) with the generalized bas-relief transformations (1.38) is isotropic scaling composed with a two-element set of the identity transformation and the transformation that reflects the first two coordinates [2]. Algebraically,

$$\mathbf{A} = \lambda \mathrm{diag}\left[s, s, 1\right] , \qquad \lambda > 0,\ s = \pm 1. \tag{5.22}$$

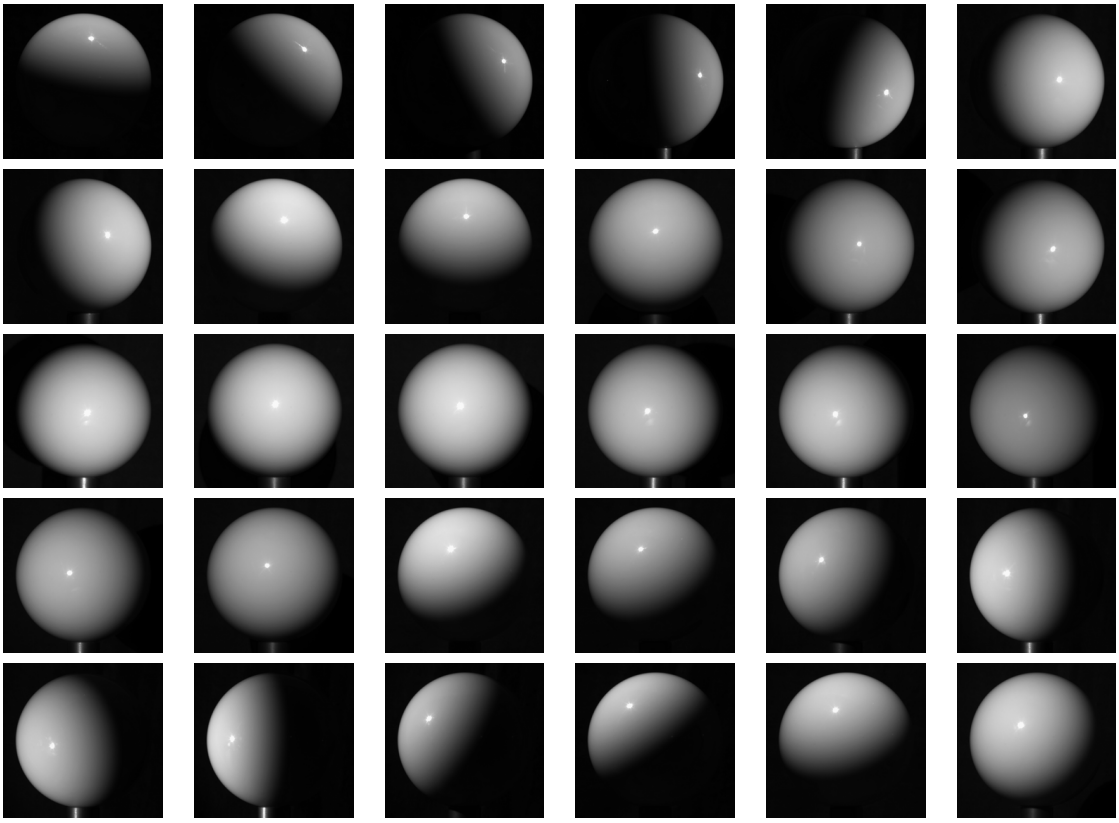


Figure 5.2: Input data for the WhiteBall object.

5.3 Experiment

This section demonstrates the applicability of the theoretical results obtained in practice. We performed the experiments with two objects:

1. *WhiteBall* which is a highly polished billiard ball of uniform albedo,

2. *ChinaPot* which is a glazed china tea pot with painted motif.

Images were acquired by 12 bit cooled camera (COOL-1300 by Vosskühler, standard Computar 75mm lens) under tungsten illumination (150W, stabilized direct current). The light was moved by hand around the object. The distance between object and light was not kept constant. No information about lights has been measured nor recorded. Input images for the WhiteBall object are shown in Fig. 5.2.

Obtaining the scaled normals matrix $\mathbf{B}$ and the scaled lights matrix $\mathbf{S}$ was performed as described in detail in Section 4.3 except we applied a more sophisticated procedure for identifying the Lambertian portion of the data (Step 2 below):

1. Image points whose intensity fell below or over respective thresholds were labeled as non-Lambertian, other as *candidates* for Lambertian points.

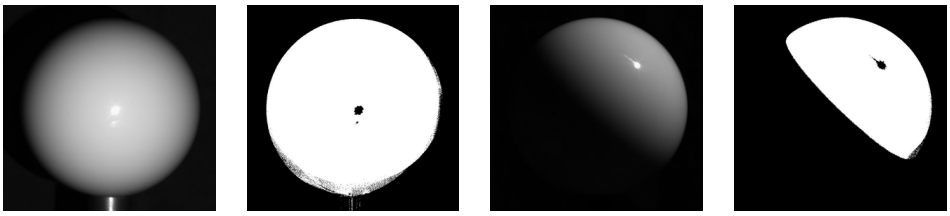


Figure 5.3: Selected Lambertian behaviour masks for the WhiteBall object.

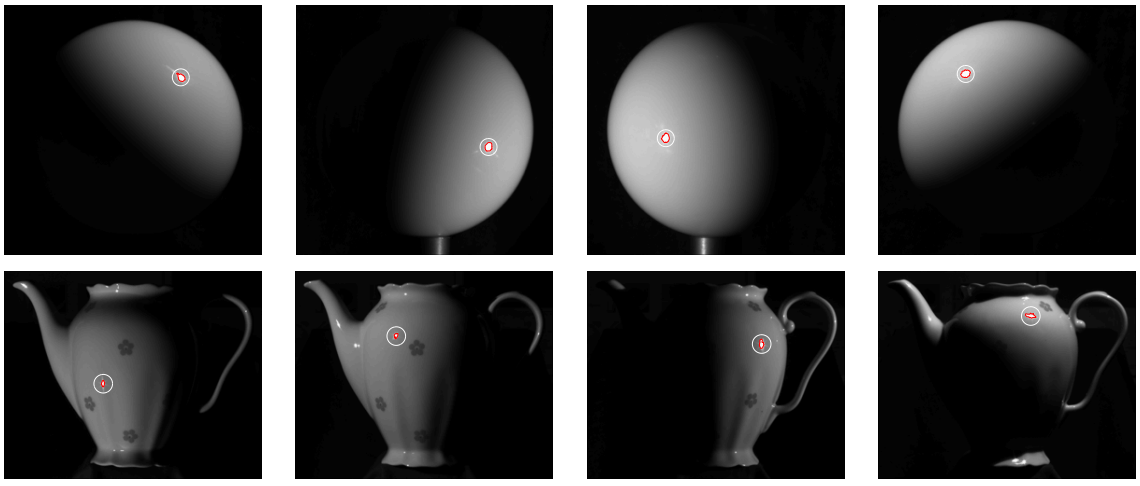
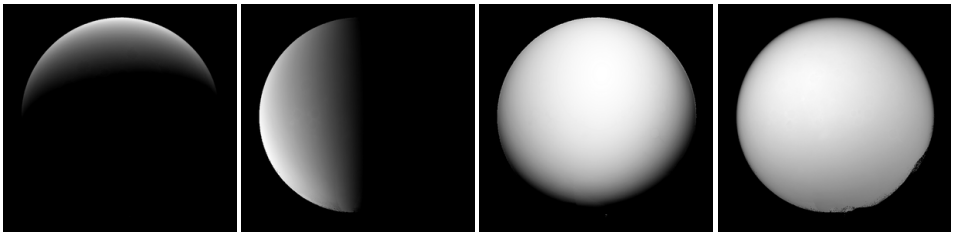
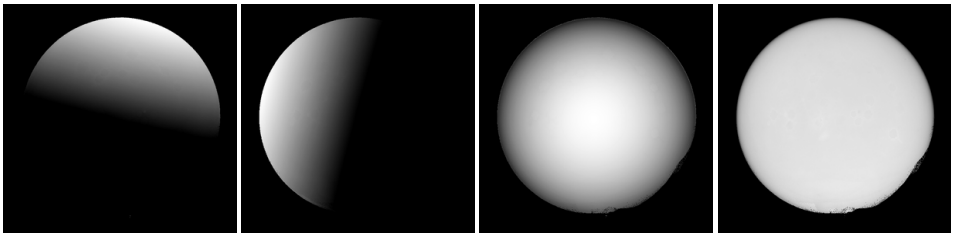


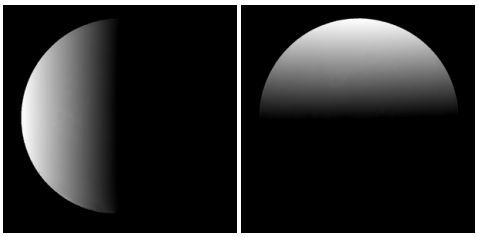
Figure 5.4: Selected specularities (marked red) used for photometric stereo ambiguity reduction for both objects.



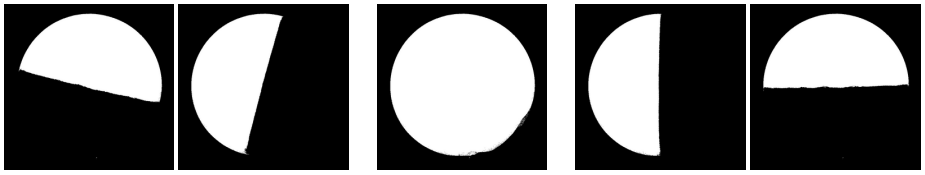
(a) Normals illuminated from the coordinate axis directions (left, top, viewpoint) and albedo (last image) reconstructed by Jacobs method from segmented input data assuming Lambertian reflectance.



(b) Normals illuminated as above and albedo conforming to the consistent viewpoint constraint.



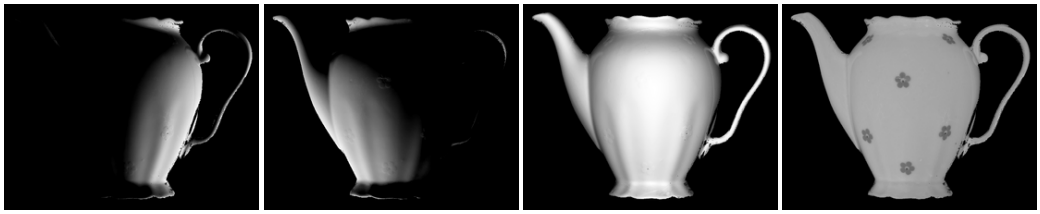
(c) Normals illuminated from the left and top directions conforming to both the consistent viewpoint and integrability constraints. The viewpoint-illuminated image and albedo are identical to those in Fig. 5.5(b).



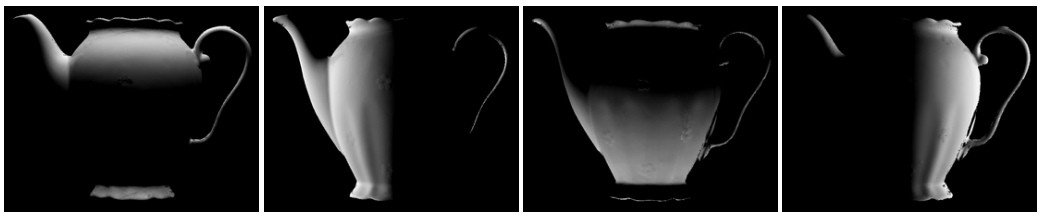
(d) Shadow boundaries for illuminations as above after applying the consistent viewpoint constraint (first three) and both consistent viewpoint and integrability constraints (last three).

Figure 5.5: Results on the WhiteBall object.

(a) Normals and albedo as in Fig. 5.5(a).



(b) Normals and albedo as in Fig. 5.5(b).



(c) Normals illuminated from four directions (top, left, below, right). The viewpoint illuminated image and albedo are identical to those in Fig. 5.6(b).

(d) Shadow boundaries for illuminations as above after applying the consistent viewpoint constraint (first three) and both consistent viewpoint and integrability constraints (last five).

Figure 5.6: Results on the ChinaPot object.

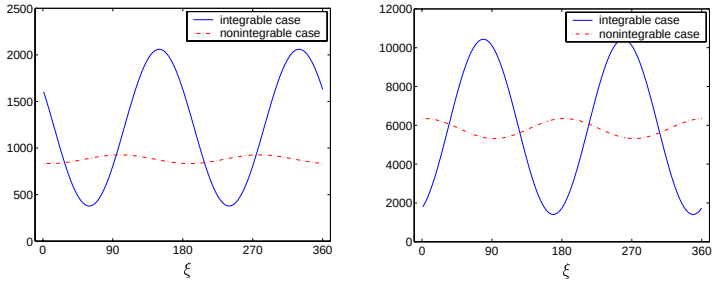


Figure 5.7: Integrability violation measure for two possible coordinate frame handednesses as a plot of rotation angle around the viewing axis. The WhiteBall object (left) and the ChinaPot object (right).

2. Candidates for Lambertian points were confirmed to be Lambertian in the case that in four randomly selected images, they belonged to sufficiently large image pixel sets with Lambertian behavior (i.e., if intensities of the pixel set in any of four images could be sufficiently well expressed as a linear combination of intensities of the pixel set in the other three images). Only such quadruples of images were involved whose any three corresponding light directions were sufficiently far from being coplanar (automatic check of this condition was done using simple conditioning number tests). Two selected Lambertian-consistent masks resulting from this step are shown in Fig. 5.3.

3. Lambertian portion of data was factorized by Jacobs algorithm [18]. From the factorization the lights **S** were obtained.

4. Normals **B** were computed using Lambertian image regions and the lights obtained in the previous step. Each normal was fit individually by using least-square fit. After that, lights were re-evaluated by an analogous procedure, and this iterative process (alternating between re-computation of normals and lights as in [13]) was repeated 10 times. The residual (the sum of squared differences between the predicted and observed intensities over the valid image regions) converged to about 1/3 of its initial value. The result of this step is shown in Fig. 5.5(a) for the WhiteBall object and in Fig. 5.6(a) for the ChinaPot object. Note neither the illuminated normals nor albedo do correspond to our expectations.

5. Four specularly reflecting normals were determined. Specular regions were selected by hand in four images. Normals from the previous step were averaged over the whole extension of the respective specularity. The selections are shown in Fig. 5.4.

6. A transformation **A** was found that maps specular normals and corresponding lights onto those which fulfill the consistent viewpoint constraint. The idea of the algorithm was to design 'ideal' specular pairs (which follow the consistent viewpoint constraint *exactly*), and look for transformation **A** that maps

experimentally obtained specular normals and lights *closest* (in a least-square sense) to these ideal specular pairs. The algorithm was essentially of the same type as the well-known bundle-adjustment one and we acknowledge the article of Triggs et al. [36] that helped us to design it. The transformation itself was parametrized as $\mathbf{A} = \mathrm{diag}\,[1, \lambda_2, \lambda_3]\,\mathbf{V}^{T*}$. A unique solution existed in both objects.

7. This transformation was applied to normals and lights output from Step 4. The consistent viewpoint direction resulting from the previous step was rotated to $[0,\ 0,\ 1]^T$.

   The results of this step are shown in Figs. 5.5(b) and 5.6(b). Note that the resulting albedo is uniform, as expected. It is already the disambiguated albedo of the object. Note also the viewpoint-illuminated normal fields are already symmetric in both objects, as expected.

8. Integrability constraint was applied to find the rotation angle around the viewpoint $[0,\ 0,\ 1]^T$. The integrability constraint violation measure was constructed as a sum of squared integrals over elementary loops in the normal field. As normals are rotated around the viewing direction, the integrability violation measure varies harmonically (see Fig. 5.7). Thus a simple harmonic fit locates the minima (corresponding to the convex/concave ambiguity). The same fit was done for normals that are reflected (the y-component stays the same, the x-component is inverted). This was done to resolve ambiguity in coordinate frame handedness. The harmonic fits for normal fields with/without the reflection are shown in Fig. 5.7. Solution that is able to fulfill the integrability condition closer was selected automatically. The remaining convex/concave ambiguity was resolved by hand.

   The result of this step is shown in Figs. 5.5(c) (illuminated from left and top directions) and 5.6(c) (illuminated from four directions as indicated). Note the normals look as expected. This is confirmed by self-shadow boundaries as shown in Fig. 5.5(d) and 5.6(d).

5.4 Implications

In this section we show that there exist problems in computer vision that are formally equivalent, or very similar to the discussed problem of uncalibrated photometric stereo with consistent viewpoint constraint.

*This corresponds to the SVD decomposition of $\mathbf{A}$ (5.13) with $\mathbf{U}$ set to the identity and λ_1 set to 1. Specifically, the optimized parameters were: viewpoint direction $\mathbf{v}$, four ideal specular normals $\mathbf{n}^S$, and the parameters of transformation $\mathbf{A}$. Ideal specular lights were computed by reflecting the (optimized) viewing direction $\mathbf{v}$ around (optimized) ideal specular normals $\mathbf{n}^S$. Initial parameter values were: $\lambda_2 = \lambda_3 = 1$, $\mathbf{V}$ initiated randomly (random rotation, solution was proven not to be dependent on this initial choice), ideal specular normals were set to normalized specular normals, and $\mathbf{v}$ to normalized average of initial specular normals. The iterative algorithm converged quickly (in about 15 iterations) into a unique solution.

5.4.1 Uncalibrated affine stereopsis

In Section 5.1, we analyzed the problem of uncalibrated photometric stereo, and the role of consistent viewpoint direction in constraining essentially affine ambiguity of uncalibrated photometric stereo. Here we show the validity of the following statement:

> Let affine reconstruction of an object be evaluated by geometric stereo with affine uncalibrated cameras. If four cameras in a general configuration observe specularities being reflected by the object surface from one distant point light source, then the original affine ambiguity reduces into similarity (composition of rotation, overall scaling, and change in coordinate frame handedness).

This statement follows from that if we make substitution *light direction* $\leftrightarrow$ *viewing direction* in the analysis given in Section 5.1, all the results apply.

Let us check the validity of this observation in detail. Let there be a surface $\mathbf{X}$ that is parametrized by u and v, $\mathbf{X} = \mathbf{X}(u, v)$. Geometric stereo with uncalibrated affine projection matrices evaluates the shape up to an affine transformation because the projections $\mathbf{x}^j(u, v)$ of point $\mathbf{X}(u, v)$ in the j-th affine camera $\mathbf{C}^j$ are[†]

$$\mathbf{x}^j(u,v) = \begin{bmatrix} C^j_{1,1} & C^j_{1,2} & C^j_{1,3} \\ C^j_{2,1} & C^j_{2,2} & C^j_{2,3} \end{bmatrix} \begin{pmatrix} X_1(u,v) \\ X_2(u,v) \\ X_3(u,v) \end{pmatrix} = \left[\mathbf{C}^j_1, \mathbf{C}^j_2\right]^T \mathbf{X}(u,v) = \mathbf{C}^{jT}\mathbf{X}(u,v) \tag{5.23}$$

and thus $\mathbf{x}^j(u, v)$'s are invariant under transformation $\mathbf{C}^j \mapsto \mathbf{A}\mathbf{C}^j$, $\mathbf{X} \mapsto \mathbf{A}^{-T}\mathbf{X}$, where $\mathbf{A} \in GL(3)$. It is known (see Yuille et al. [45]) that under these affine transformations the camera viewing vectors $\mathbf{v}_j \sim \mathbf{C}^j_1 \times \mathbf{C}^j_2$ are transformed covariantly and the surface normals $\mathbf{n} \sim \frac{\partial \mathbf{X}}{\partial u} \times \frac{\partial \mathbf{X}}{\partial v}$ are transformed contravariantly: $\mathbf{v} \sim \mathbf{A}^{-T}\mathbf{v}$, and $\mathbf{n} \sim \mathbf{A}\mathbf{n}$ ($\sim$ means "up to a scaling factor"). Thus the normals in affine geometrical stereo transform like in photometric stereo, and the camera viewing vectors behave just like the illumination directions. But, in addition, the specular geometry condition (see Fig. 3.4) is also symmetric with respect to the change $\mathbf{l} \leftrightarrow \mathbf{v}$. After we formulate the equivalent of the consistent viewpoint constraint (in this case, it could be called *consistent specular illumination constraint*), the mathematics of the problem is the same. The affine ambiguity is therefore reduced into composition of scaling, rotation around the (unknown) illumination direction, and change in coordinate frame handedness. But this corresponds to the similarity ambiguity.

5.4.2 Uncalibrated affine Helmholtz stereopsis

The principle of the Helmholtz stereopsis has been reviewed in Section 3.1.2. The special case is considered here in which the cameras are affine and uncalibrated (as in the previous section).

[†]Origins of image frames in all cameras are aligned.

The Helmholtz stereopsis setup requires that the light sources are co-located with camera centers which implies that they must also be at infinity. They are indeed represented by the scaled light vectors $\mathbf{s}_i$ like in photometric stereo. The co-locatedness requires that

$$\forall i \;\; \exists \lambda_i > 0: \quad \mathbf{s}_i = \lambda_i \mathbf{v}_i \,, \tag{5.24}$$

where $\mathbf{v}_i$ is the viewing vector of the i-th affine camera. It is known that affine uncalibrated Helmholtz stereopsis deteremines only the affine structure of the object reconstructed [48]. This is because for affine cameras and directional light sources the reciprocity constraint (3.6) written as

$$\begin{bmatrix} I_{\mathbf{A}_1}\mathbf{s}_{\mathbf{A}_1} - I_{\mathbf{B}_1}\mathbf{s}_{\mathbf{B}_1} \\ I_{\mathbf{A}_2}\mathbf{s}_{\mathbf{A}_2} - I_{\mathbf{B}_2}\mathbf{s}_{\mathbf{B}_2} \\ \cdots \\ I_{\mathbf{A}_n}\mathbf{s}_{\mathbf{A}_n} - I_{\mathbf{B}_n}\mathbf{s}_{\mathbf{B}_n} \end{bmatrix} \cdot \mathbf{n} = 0 \tag{5.25}$$

is indeed invariant when the scaled light vectors (colinear with viewing directions) are transformed covariantly and the normals are transformed contravariantly by any transformation from $GL(3)$.

Consider a Helmholtz stereopsis setup: a camera with a viewing direction $\mathbf{v_A}$ and another camera with viewing direction $\mathbf{v_B}$. Let $\mathbf{n}^S$ denote the normal which is specular when seen from $\mathbf{v_B}$ under illumination from $\mathbf{s_A} = \lambda_\mathbf{A}\mathbf{v_A}$. Then (5.7) is indeed equivalent to

$$\left\{2(\mathbf{s_A} \cdot \mathbf{n}^S)\mathbf{P}\mathbf{n}^S - (\mathbf{n}^S \cdot \mathbf{P}\mathbf{n}^S)\mathbf{s_A}\right\} \times \mathbf{v_B} = \mathbf{0}\,. \tag{5.26}$$

The vectors $\mathbf{v_B}$ and $\mathbf{s_A}$ are both determined from affine Helmholtz stereopsis. For different Helmholtz pairs this relation represents a set of equations. Compared with (5.7), the equations are linear in the unknown $\mathbf{P}$, and the number of degrees of freedom to fix is five. This observation suggests that three specularities observed in three Helmholtz pairs in general configurations are sufficient to solve (5.26). The affine ambiguity is then again reduced to similarity as in the previous case.

5.5 Summary of the chapter

As a basic result of this chapter we have shown that if object reflectance is a sum of Lambertian and specular terms, the uncalibrated photometric stereo ambiguity is reduced into a 2dof group of transformations (compositions of rotation around the viewing vector, isotropic scaling and change in coordinate frame handedness). For that, identification of specularities in images corresponding to four different distant point lights in general configuration is sufficient. We expect a similar result will hold if the specular spike is blurred by isotropic surface roughness. The good applicability of the approach was verified experimentally on two real objects made of different material.

Note that albedo is obtained without imposing the integrability constraint. The integrability is used to fix only 1dof of the normal field. Since integrability must be

computed on normal derivatives, any reduction of the number of parameters to be found significantly improves the accuracy of the resulting normals.

As we noted, lights and cameras play a symmetric role in the consistent viewpoint constraint. By interchanging lights and cameras, the constraint may also be applied to the case of uncalibrated geometric stereo with four affine cameras in a general configuration observing specularities from a single distant point light source. The affine ambiguity is then reduced to similarity ambiguity. The analysis also showed that interpretation of specularities observed in three Helmholtz image pairs taken by affine uncalibrated cameras results in the same ambiguity reduction as in the previous case.

Chapter 6

Summary & Conclusions

6.1 Main results and additional contributions

This thesis opened a novel question in uncalibrated photometric stereo: Provided that the surface reflectance is *non-Lambertian*, what is the ambiguity up to which the surface parameters can be determined? We made following two steps towards the general answer:

A If the surface reflectance is a superposition of Lambertian and ideal specular reflectance and specularities are visible in (at least) four images then the albedo map can be determined up to a scale and the surface orientations are subject to composition of rotation around the viewing vector (1dof) and the change in coordinate frame handedness (binary ambiguity). To obtain this result, we constructed a new constraint termed *the consistent viewpoint constraint* which requires that the light directions reflected by the normals for which specularities are observed must give the viewing direction. If previously known integrability constraint on normals is applied subsequently, the albedo stays determined up to a scale and the normals get determined up to a convex/concave (binary) ambiguity.

B If the surface reflectance is separable and isotropic then it is possible to determine the orthographic projections of normals onto the image and light planes. These projections are given up to a scale and we termed them the **v**-tilt and the **l**-tilt, respectively. It was shown that making the scaled normals and scaled lights consistent with the measured **v**-tilts and **l**-tilts determine the albedo map up to a scale and the normals up to a convex/concave ambiguity. The same ambiguity reduction is obtained when the **l**-tilt and integrability constraints are combined. When **v**-tilt and integrability constraints are combined then the resulting ambiguity is the bas-relief ambiguity which contains 2dof.

The process leading to these results delivered interesting novel 'by-products':

- Theoretical analysis made within Result **A** showed that:

  - The uncalibrated affine stereopsis can be disambiguated provided that specular reflections from a single point light source in infinity are observed in (at least) four images. The resulting ambiguity is the similarity.
  - The same ambiguity reduction is achieved for a case of uncalibrated affine Helmholtz stereopsis with at least three reciprocal pairs.

- Procedure for obtaining the **v**-tilts and **l**-tilts within Result **B** based on polarisation measurement employed the dependence of observed image intensity on the incident light polarisation. The possibility to get the **l**-tilt using the polarisation measurement was not perceived in computer vision field before, in contrast to the **v**-tilt evaluation [41].

6.2 Impact

It is especially the Result **A** which has a potential to stimulate further research in the area of uncalibrated photometric stereo for non-Lambertian surfaces because it uses just the information available in the images (as opposed to Result **B** which requires to measure the **v**-tilts and **l**-tilts). Although Result **A** has been acquired under the assumption of reflectance given by a sum of Lambertian and ideal specular reflectance, an intuition suggests that the same result will hold for objects the interface reflectance of which is not such sharp but still centered around the specular direction. As many real surfaces possess reflectance of nonzero interface component, this result pushes the limits of applicability of photometric stereo. One may envisage a scenario in which a cheap camera is used and a hand-held light source is waved around the object and the shape and reflectance of the object is revealed. Relaxing the need of knowing the positions of light sources used for acquiring individual images represents a significant step towards even wider use of photometric stereo in practice.

6.3 Future work

We see two main directions of future research. First, we aim to analyse the ambiguities in surface parameters determination for general reflectance models. Second, it is an interesting question how the novel results obtained in this work could be incorporated within the framework which represents illumination in a spherical harmonics basis [30, 1]. Such approach enables to deal with complex illumination conditions within uncalibrated photometric stereo. We believe that it is the combination of these two streams—the signal processing framework for lighting and the non-Lambertian models of reflectance—that holds the promise of opening up many new and exciting research avenues and numerous applications.

Bibliography

[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, February 2003.

[2] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, November 1999.

[3] M. Born and E. Wolf. *Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light.* Cambridge University Press, 1980.

[4] J.-Y. Bouguet and P. Perona. 3D photography using shadows in dual-space geometry. *International Journal of Computer Vision*, 35(2):129–149, 1999.

[5] F. J. J. Clarke and D. J. Parry. Helmholtz reciprocity: Its validity and application to reflectometry. *Lighting Research & Technology*, 17(1):1–11, 1985.

[6] E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, 1982.

[7] M. Daum and G. Dudek. On 3-D surface reconstruction using shape from shadows. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 461–468, 1998.

[8] O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc. International Conference on Computer Vision*, volume 1, pages 581–586, Los Alamitos, USA, July 2001.

[9] O. Drbohlav and R. Šára. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. European Conference on Computer Vision*, volume 2, pages 46–60, 2002.

[10] J. Fan and L. B. Wolff. Surface curvature and shape reconstruction from unknown multiple illumination and integrability. *Computer Vision and Image Understanding*, 65(2):347–359, February 1997.

[11] O. Faugeras, Quang Tuan Luong, and T. Papadopoulo. *The Geometry of Multiple Images : The Laws That Govern the Formation of Multiple Images of a*

Scene and Some of Their Applications. MIT Press, Cambridge, Massachusetts, 2001.

[12] D. A. Forsyth and J. Ponce. *Computer Vision : A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, USA, 2003.

[13] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 52–58, 1998.

[14] G. H. Golub and C. F. Van Loan. *Matrix Computation*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, USA, 3rd edition, 1996.

[15] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000. Chapter 2 : Projective Geometry and Transformation of 3D.

[16] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of Optical Society of America A*, 11(11):3079–3089, November 1994.

[17] B. K. P. Horn and M. J. Brooks. *Shape from shading*, chapter 8. The MIT Press, 1989.

[18] D. Jacobs. Linear fitting with missing data: Applications to structure from motion and to characterizing intensity images. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 206–212, 1997.

[19] J. R. Kender and E. M. Smith. Shape from darkness: deriving surface information from dynamic shadows. In *Proc. IEEE International Conference on Computer Vision*, pages 539–546, 1987.

[20] J. J. Koenderink and A. J. van Doorn. The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, 23(3):217–234, 1997.

[21] J.J. Koenderink, A.J. van Doorn, and M. Stavridi. Bidirectional reflection distribution function expressed in terms of surface scattering modes. In *Proc. European Conference on Computer Vision*, volume 2, pages 28–39, 1996.

[22] D. J. Kriegman and P. N. Belhumeur. What shadows reveal about object structure. In *Proc. European Conference on Computer Vision*, volume 1407 of *Lecture Notes in Computer Science*, pages 399–414, 1998.

[23] J. H. Lambert. *Photometria sive de mensura et gradibus luminis, colorum et umbra*. Augustae Vindelicorum, Basel, 1760.

[24] S. Lin and S. W. Lee. Estimation of diffuse and specular appearance. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 855–860, 1999.

[25] S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur. Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 391–398, 2001.

[26] C. E. Mungan. Bidirectional reflectance distribution functions describing first-surface scattering. Technical report, AFOSR Final Report for the Summer Faculty Research Program, 1998. Available at `http://www.altavista.com/r?ck_sm=a51e9b19&rpos=1&rpge=1&rsrc=R&ref=1000%20080&uid=2d796d1750fe3bb9&r=http%3A%2F%2Fusna.edu%2FUsers%2Fphysics%2Fmungan%%2FPublications%2FBRDFreview.pdf`.

[27] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Journal of Robotics and Automation*, 6(4):418–431, August 1990.

[28] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. Geometrical considerations and nomeclature for reflectance. U.S. Department of Commerce, National Bureau of Standards, Washington D.C., USA, 1977.

[29] R. Ramamoorthi and P. Hanrahan. Analytic PCA construction for theoretical analysis of lighting variability, including attached shadows, in a single image of a convex Lambertian object. In *Proc. CVPR Workshop on identifying objects across variations in lighting: Psychophysics and computation*, 2001. To appear. Available on-line at `http://graphics.stanford.edu/papers/pca/pca.pdf`.

[30] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *Proc. International SIGGRAPH Conference*, pages xx–yy, 2001. Available on-line at `http://graphics.stanford.edu/papers/invrend/invrend.pdf`.

[31] M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi. Measurement of surface orientations of transparent objects using polarization in highlight. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 381–386, 1999.

[32] S. A. Shafer. Using color to separate reflection components. *Color Res. App.*, 10(4):210–218, 1985.

[33] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):449–454, April 1996.

[34] M. Šonka, V. Hlaváč, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Chapman and Hall Computing, 1993.

[35] H. D. Tagare and R. J. P. deFigueiredo. A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):133–152, February 1991.

[36] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. F. Fitzgibbon. Bundle adjustment-a modern synthesis. In *Proc. Vision Algorithms: Theory and Practice. International Workshop on Vision Algorithms.*, volume 1883 of *Lecture Notes in Computer Science*, pages 298–372, 1999.

[37] H. von Helmholtz. *Optique physiologique*, volume I, pages 231–232. Éditions Jaques Gabay, Sceaux, France, 1989. Available on-line from the digital library of *Bibliothèque nationale de France* at `http://gallica.bnf.fr/`.

[38] L. B. Wolff. Surface orientation from two camera stereo with polarizers. In *Proc. Conference of Optics, Illumination and Image Sensing for Machine Vision IV*, volume 1194 of *Proc. of SPIE - The International Society for Optical Engineering*, pages 287–297, November 1990.

[39] L. B. Wolff. Diffuse-reflectance model for smooth dielectric surfaces. *Journal of Optical Society of America A*, 11(11):2956–2968, November 1994.

[40] L. B. Wolff and E. Angelopoulou. Three-dimensional stereo by photometric ratios. *Journal of Optical Society of America A*, 11(11):3069–3078, November 1994.

[41] L. B. Wolff and T. E. Boult. Constraining object features using a polarization reflectance model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):635–657, July 1991.

[42] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[43] R. J. Woodham, Y Iwahori, and R. A. Barman. Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength. Technical Report 91-18, Laboratory for Computational Intelligence, Univ. of British Columbia, Vancouver, BC, Canada, August 1991.

[44] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 158–164, 1997.

[45] A. L. Yuille, J. M. Coughlan, and S. Konishi. The KGBR viewpoint-lighting ambiguity and its resolution by generic constraints. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 376–382, 2001.

[46] Jiang Yu Zheng and A. Murata. Acquiring a complete 3D model from specular motion under the illumination of circular-shaped light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):913–920, 2000.

[47] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. In *Proc. European Conference on Computer Vision*, volume 3, pages 869–884, May 2002.

[48] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Toward a stratification of Helmholtz stereopsis. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 548–555, June 2003.