Photometric stereo with central panoramic cameras

Jordan Caracotte, Fabio Morbidi, El Mustapha Mouaddib

HAL Id: hal-03086537
https://hal.science/hal-03086537

Photometric stereo with central panoramic cameras

Jordan Caracotte[a,**], Fabio Morbidi[a], El Mustapha Mouaddib[a]

[a]*MIS laboratory, Université de Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens, France*

ABSTRACT

In this paper we introduce and rigorously solve for the first time the photometric stereo problem for *central panoramic systems*. In particular, by leveraging the unified camera model, we present a *new* spherical image irradiance equation and analyze its mathematical properties in detail. The discrete gradient field issued from this equation is fed into a drift-free normal integration algorithm tailored to the spherical image geometry, and the light direction is estimated using the specular highlights observed on mirror balls inside the scene. Extensive experiments conducted with analytical surfaces, and synthetic and real-world images captured by central panoramic cameras, show that the proposed 3D reconstruction pipeline is effective and tolerant to noise. The image database and the code developed are publicly available at the address: mis.u-picardie.fr/~fabio/ustereo.html

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Motivation and related work

In the last decades, photometric stereo and multi-view stereo have known a growing success in computer vision (Ackermann and Goesele, 2015; Furukawa and Hernández, 2015). Photometric stereo approaches leverage the observed intensity variations caused by illumination changes to reconstruct the shape of a 3D object. In its simplest incarnation, a diffuse surface is illuminated from at least three known directions and images are captured by a static camera: in his pioneer work, Woodham showed that under suitable conditions, the per-pixel surface normals can be thus recovered (Woodham, 1980). Besides the Lambertian model (i.e. surface brightness looks the same from any viewing direction) and the neglect of shadows and inter-reflections, other simplifying assumptions

include light sources at infinity and orthographic projection of the scene onto the imaging sensor. The data is typically collected at small scale, in controlled indoor environments, using multiple static lamps or robotic gantries.

Numerous extensions to the basic problem formulation have been proposed in the literature. For example, in (Abrams et al., 2012; Ackermann et al., 2012; Abrams et al., 2013; Jung et al., 2015), the lighting constraints have been relaxed and outdoor images (from webcams) under natural sunlight illumination have been used for 3D reconstruction. When no prior knowledge about the illumination, geometry and reflectance of the surface is available, the so-called "uncalibrated" photometric stereo problem arises, which is well-known to be ill-posed, see (Hayakawa, 1994; Basri et al., 2007; Quéau et al., 2015, 2017; Papadhimitri and Favaro, 2013). Other extensions include more sophisticated surface reflectance models (Tagare and Defigueiredo, 1991; Higo et al., 2009; Weinmann et al., 2013; Lu et al., 2015), (i.e. "Bidirectional

---

[**]Corresponding author: Tel.: +33 03 22 82 59 02;

e-mail: jordan.caracotte@etud.u-picardie.fr (Jordan Caracotte)
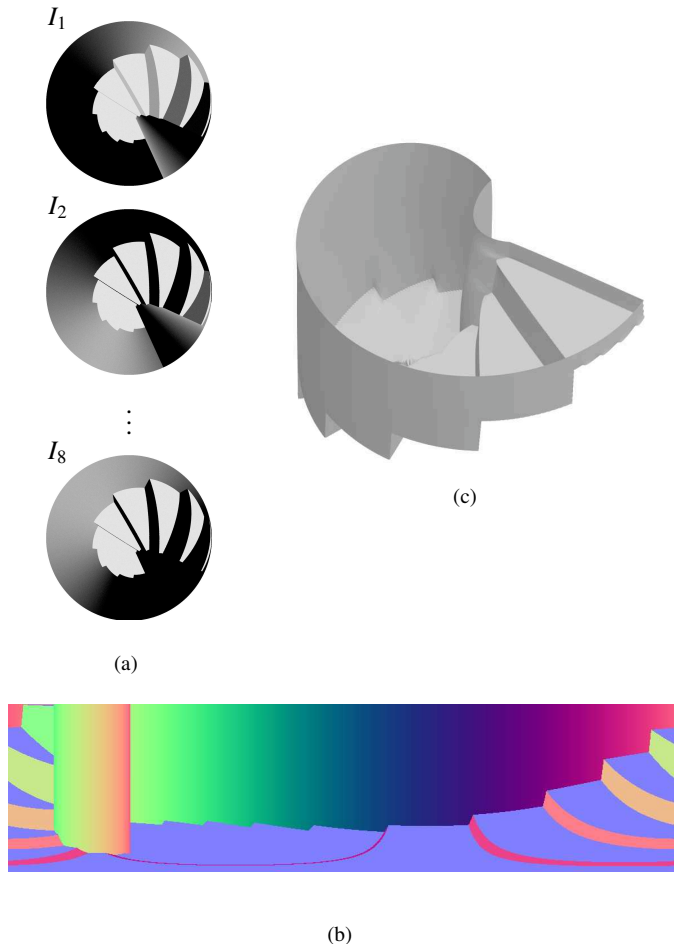
I_1

I_2

$\vdots$

I_8

(a)

(b)

(c)

Fig. 1. Example output from our photometric stereo algorithm: (a) Input 360° images of a spiral staircase under eight different illumination conditions; (b) Estimated normal map in false colors (turned upside down, for ease of visualization), obtained with the panoramic images; (c) 3D reconstruction of the scene.

era coupled with a hyperbolic, parabolic or elliptical mirror). Thanks to their large field of view, panoramic systems offer distinctive advantages over standard pinhole cameras, for photometric stereo. First of all, a single 360° image is more informative about the content of an observed scene, thus speeding up data collection and 3D reconstruction. Therefore, it is no longer necessary to create large datasets of images taken from multiple vantage viewpoints for recording the appearance of a surface from all around. Second, severe visibility issues, e.g. due to self-occlusion, are alleviated. These benefits are particularly evident when the camera is placed at the center of a complex 3D environment, as that reported in Fig. 1, which would require multiple shots from a perspective system to be fully covered. Central panoramic cameras also suffer from some fundamental limitations: in fact, they provide images with non-uniform spatial resolution and large distortions. These issues can be partially alleviated by non-central systems, such as panoramic line-scan cameras (Benosman et al., 1996; Ait-Aider and Berry, 2019).

As shown in (Geyer and Daniilidis, 2000), projection systems with a single effective viewpoint can be described by a unifying image spherical model. However, whereas this model has been very successful in robot vision (e.g. for visual odometry (Zhang et al., 2016; Matsuki et al., 2018) or visual servoing (Mariottini and Prattichizzo, 2008)), it has not yet been exploited in the photometric stereo literature.

1.2. Original contributions, organization and notation

In this paper, we make standard assumptions on the illumination and reflectance models (i.e. Lambertian surfaces), but we push the boundaries of research by considering *more general camera models*. More specifically, we leverage the unified camera model of Geyer & Daniilidis, and present a *novel* spherical image irradiance equation whose mathematical properties are studied in detail. The measured gradient field is fed into a drift-free normal integration algorithm which explicitly accounts for the geometry of spherical images. Adapting the approach in (Schnieders and Wong, 2013) to our generic setting, the light direction in real images is estimated using the

Reflectance Distribution Functions" or BRDF, other than the classical Lambertian reflectance), near-field (i.e. local) lights (Mecca et al., 2014; Logothetis et al., 2017), the relaxation of constraints on camera placement (multi-view stereo) (Hernández et al., 2008; Vlasic et al., 2009; Logothetis et al., 2019) and combinations thereof (Shi et al., 2019).

However, while significant efforts have been made in the literature to robustify Woodham's approach and to extend it from orthographic to perspective projection (Tankus and Kiryati, 2005), to the best of our knowledge (with the exception of an early attempt in (Kondo et al., 2008)), no results exist for *central panoramic cameras*, i.e. for vision systems which preserve the uniqueness of projection viewpoint (as a pinhole cam-

specular highlights observed on mirror balls inside the scene. A large-scale validation campaign performed with analytical surfaces, synthetic and real-world images captured by central catadioptric cameras illustrates the theory and shows the effectiveness of our 3D reconstruction pipeline in terms of accuracy and robustness.

In summary, the *two main contributions* of this paper are:

1. A unifying theory of photometric stereo, encompassing widely-used camera models in computer vision,

2. A complete 3D reconstruction pipeline, including a drift-free normal integration algorithm tailored to the spherical image geometry.

The rest of this paper is organized as follows. In Sect. 2, we briefly review the unified camera model and the perspective image irradiance equation. In Sect. 3, we present the new spherical image irradiance equation, we study its geometric properties, and describe our geometry-aware normal integration method. In Sect. 4, we discuss the results of extensive numerical and real-world experiments. Finally, in Sect. 5, the main contributions of the paper are summarized and some possible avenues for future research are outlined.

Assumptions: For the reader's convenience, we recapitulate here the assumptions made through this work. We consider a 3D environment which includes a central panoramic camera observing a scene with Lambertian surface. The scene and the calibrated camera are static. A moving homogeneous light source without fall-off is set at infinity, and the direction of parallel light rays is assumed to be known (or it is estimated, e.g. using mirror balls inside the scene). **Notation**: Throughout this article, we use the symbol $\mathbb{R}^n$ to denote the n-dimensional Euclidean space, and $\mathbb{R}^{m \times n}$ the space of $m \times n$ matrices. SO(3) indicates the special orthogonal group in dimension three, $\mathbb{S}^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : \|\mathbf{x}\| = \rho\}$ the n-sphere of radius $\rho > 0$ where $\|\mathbf{x}\|$ denotes the Euclidean norm of vector $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{R}_y(\theta)$ and $\mathbf{R}_z(\varphi)$ the 3×3 elementary rotations of an angle θ and φ about the y- and z-axis, respectively. The tangent space of the differentiable manifold $\mathcal{M}$ at point $\mathbf{x}$ is denoted by $\mathrm{T}_{\mathbf{x}} \mathcal{M}$. Moreover, $\widehat{\mathbf{x}}$ denotes a normalized vector i.e. $\widehat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$, and

$\overline{\mathbf{x}}$ an estimate of vector $\mathbf{x}$. Finally, $\langle \mathbf{x}, \mathbf{y} \rangle$ indicates the scalar product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \times \mathbf{y}$ the vector product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, and $\triangleq$ the equality by definition.

2. Preliminaries

2.1. Unified camera model

For the reader's convenience, in this section we briefly review the unified theory of central catadioptric projection (Geyer and Daniilidis, 2000, 2001). Theorem 1 in (Geyer and Daniilidis, 2000) establishes that catadioptric projection with a single effective viewpoint is equivalent to projection to a sphere, followed by projection to a plane from a point. Consider a unit sphere centered at the origin of the world frame $\{O; x, y, z\}$, and let the plane $z = -m$ be the image plane (see Fig. 2). A 3D point $X = [x, y, z]^T$ is first projected to two antipodal points Q and Q' of coordinates $\pm X/\|X\|$ $\triangleq \pm [x_s, y_s, z_s]^T$ on the unit sphere. For the second part of the map, we need to determine the perspective projection to the plane $z = -m$ from the projection point $[0, 0, \ell]^T$ on the z-axis of the sphere (point O' in Fig. 2). If we confine ourselves to point Q, the projection of X onto the image plane is then:

$$\mathbf{x} = \left[\frac{(\ell + m)x}{\ell \|X\| - z}, \quad \frac{(\ell + m)y}{\ell \|X\| - z}, \quad -m \right]^T. \qquad (1)$$

Fig. 2. A 3D point $X = [x, y, z]^T$ is projected to two antipodal points $\pm X/\|X\|$ on the unit sphere (Geyer and Daniilidis, 2000, Sect. 2.1). The two antipodal points Q and Q' are projected to the image plane $z = -m$ via projection from the point O' (in the interest of clarity, only the projection of Q to x on the image plane is shown in the figure). The normal vector to the 3D surface $\mathcal{S}$ at point X is denoted by N, and L is a generic illumination vector.

Let us now describe the projections using *conical section mirrors* (i.e. parabolic, hyperbolic and elliptical mirrors). In this case, the non-negative parameters ℓ and m have a clear physical meaning: ℓ is a function of the eccentricity of the conic, and m is a function of both its scale and eccentricity. When $\ell = 1$ and $m = 0$, i.e. the point of projection is the North pole, we have a case of *stereographic projection* (when X is restricted to the sphere) or, equivalently, parabolic projection (Coxeter, 1969, Sect. 6.9). On the other hand, when $\ell = 0$ and $m = 1$, we have a *perspective projection*.

Note that the model above can be applied to cameras with fisheye lenses as well (Ying and Hu, 2004). However, it does not fit most fisheye lenses perfectly, and it is often supplemented with a distortion model, see (Barreto, 2006; Kannala and Brandt, 2006; Usenko et al., 2018).

2.2. Perspective image irradiance equation

To make a parallel with the results that we will present in Sect. 3, we report below the perspective image irradiance equation derived in (Tankus et al., 2005, Th. 1). The following customary assumptions hold throughout the paper. A 3D surface $\mathcal{S}$ (see Fig. 2) can be represented by a function of world coordinates or image coordinates (Tankus and Kiryati, 2005). We denote by $\check{z}(x, y)$ the depth function written in $\{O;\, x,\, y,\, z\}$. If $(x, y, \check{z}(x, y))$ is projected onto the image point (u, v) (normalized coordinates), then its depth is denoted by $z(u, v)$ (by definition, $z(u, v) = \check{z}(x, y)$). $I(u, v)$ indicates the intensity at image point (u, v). The scene is Lambertian[1] and is illuminated from direction $\mathbf{L} \triangleq [p_d, q_d, -1]^T$ by a light source at infinity (cf. Fig. 2). $\mathbf{N}(x, y)$ denotes the surface normal at point X. In our notation, we use the convention that the light sources and normals point towards the camera (hence, the negative third coordinate). Finally, $\check{\alpha}(x, y) \in [0, 1]$ denotes the diffuse albedo (intrinsic reflectivity of the surface (Smith and Fang, 2016)) at

point $(x, y, \check{z}(x, y))$, and $\alpha(u, v)$ the albedo projected onto image point (u, v).

Theorem 1 (Perspective image irradiance equation (Tankus and Kiryati, 2005))

Under the previous notation and assumptions, the perspective image irradiance equation is,

$$I(u, v) = \alpha(u, v) \langle \widehat{\mathbf{N}}(u, v), \widehat{\mathbf{L}} \rangle =$$

$$\frac{-\alpha(u, v)\,\mathbf{L}^T}{\|\mathbf{L}\|\,\sqrt{(up + vq + 1)^2 + f^2(p^2 + q^2)}} \begin{bmatrix} fp \\ fq \\ up + vq + 1 \end{bmatrix}, \quad (2)$$

where

$$p \triangleq \frac{1}{z}\frac{\partial z}{\partial u} = \frac{\partial \ln z}{\partial u}, \quad q \triangleq \frac{1}{z}\frac{\partial z}{\partial v} = \frac{\partial \ln z}{\partial v}. \quad \blacksquare$$

Note that by definition, $I(u, v) \in [0, 1]$.

3. Photometric stereo for central panoramic systems

In this section, we will derive an image irradiance equation which complies with the unified camera model presented in Sect. 2.1. To this end, we will assume that the perspective projection step in equation (1) has been already performed, and we will restrict our attention to the unit sphere. Our ultimate goal will be to estimate the *radial distance* (or range) ρ from multiple images of a 3D surface $\mathcal{S}$ observed from the same viewpoint but under different illumination conditions. We will conclude the section with a description of the method developed to integrate the measured gradient field.

3.1. Spherical image irradiance equation

To represent an image on the sphere, it is expedient to introduce the spherical coordinates (ρ, θ, φ) (Weber and Arfken, 2003, Sect. 2.5):

$$\begin{cases} x = \rho \sin\theta \cos\varphi, \\ y = \rho \sin\theta \sin\varphi, \\ z = \rho \cos\theta, \end{cases}$$

where $\rho \in [0, \infty)$ is the radial distance, $\theta \in (0, \pi)$ the polar angle and $\varphi \in [0, 2\pi)$ the azimuthal angle (see Fig. 3). Note that

---

[1] Lambertian surfaces do *not* change appearance depending on the viewing direction. For instance, matte surfaces are to a large extent well approximated by the Lambertian model, since they diffuse light almost uniformly in all directions. On the other hand, metal, mirrors, and other shiny surfaces, do not (Ma et al., 2004, Sect. 3.A).

Fig. 3. Spherical coordinates (ρ, θ, φ) and spherical coordinate system surfaces.

at $\rho = 0$, the angles θ and φ are not defined. The coordinate surfaces θ = constant, are right-circular cones with apex at the origin having the z-axis as their axis of revolution, and the coordinate surfaces φ = constant, are vertical half-planes passing through the origin (cf. Fig. 3). Let now $\mathbf{r} = [x, y, z]^T$. Then, the basis set of unit vectors $\{\mathbf{e}_\rho, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$ for the spherical coordinates is given by (Weber and Arfken, 2003, Sect. 2.3):

$$\mathbf{e}_\rho = \frac{1}{h_\rho}\frac{\partial \mathbf{r}}{\partial \rho} = \begin{bmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{bmatrix},$$

$$\mathbf{e}_\theta = \frac{1}{h_\theta}\frac{\partial \mathbf{r}}{\partial \theta} = \begin{bmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{bmatrix}, \quad \mathbf{e}_\varphi = \frac{1}{h_\varphi}\frac{\partial \mathbf{r}}{\partial \varphi} = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix},$$

where the *scale factors* (a.k.a. metrical or Lamé coefficients):

$$(h_\rho, h_\theta, h_\varphi) = \left(\left\| \frac{\partial \mathbf{r}}{\partial \rho} \right\|, \left\| \frac{\partial \mathbf{r}}{\partial \theta} \right\|, \left\| \frac{\partial \mathbf{r}}{\partial \varphi} \right\| \right) = (1, \rho, \rho\sin\theta).$$

Note that the unit vectors $\mathbf{e}_\rho$, $\mathbf{e}_\theta$ and $\mathbf{e}_\varphi$ *vary* in direction as the angles θ and φ vary. We are now ready to state the main result of this section (cf. Theorem 1).

Theorem 2 (Spherical image irradiance equation)

Let $\mathbf{v} = [v_1, v_2]^T$ and $\mathbf{w} = [w_1, w_2]^T$ be two vectors lying on the tangent plane of $\mathbb{S}^2$ at $\mathbf{r}$, and let $\det(\mathbf{A}) > 0$ where

$\mathbf{A} = [\mathbf{v}, \mathbf{w}]$. Then, under the previous assumptions, the spherical image irradiance equation for the directions $\mathbf{v}$ and $\mathbf{w}$ is,

$$I(\theta, \varphi) = \frac{\alpha(\theta, \varphi)\,\mathbf{L}^T}{\|\mathbf{L}\|\ \sqrt{(v_1 p - w_1 q)^2 + \left(\frac{w_2 q - v_2 p}{\sin\theta}\right)^2 + \det(\mathbf{A})^2}}$$

$$\cdot \mathbf{R}_z(\varphi)\,\mathbf{R}_y(\theta) \begin{bmatrix} v_1\,p - w_1\,q \\ \dfrac{w_2\,q - v_2\,p}{\sin\theta} \\ -\det(\mathbf{A}) \end{bmatrix},$$

(3)

where $\alpha(\theta, \varphi) \in [0, 1]$ denotes the albedo at point (θ, φ),

$$p \triangleq \frac{1}{\rho}\nabla_{\mathbf{w}}\rho = \nabla_{\mathbf{w}}\ln\rho, \quad q \triangleq \frac{1}{\rho}\nabla_{\mathbf{v}}\rho = \nabla_{\mathbf{v}}\ln\rho, \quad (4)$$

and $\nabla_{\mathbf{w}}\rho = \langle \nabla\rho, \mathbf{w} \rangle$ is the directional derivative[2] of ρ along vector $\mathbf{w}$.

Proof: The proof follows the same outline as that of Theorem 1 in (Tankus et al., 2005). By considering the directions $\mathbf{v}, \mathbf{w} \in T_{\mathbf{r}}\mathbb{S}^2$, the surface normal $\mathbf{N}$ can be computed as (see Fig. 2):

$$\begin{aligned}
\mathbf{N}(\theta, \varphi) &= \left(v_1\frac{\partial \mathbf{r}}{\partial \varphi} + v_2\frac{\partial \mathbf{r}}{\partial \theta} \right) \times \left(w_1\frac{\partial \mathbf{r}}{\partial \varphi} + w_2\frac{\partial \mathbf{r}}{\partial \theta} \right) \\
&= \left[\left(v_1\frac{\partial \rho}{\partial \varphi} + v_2\frac{\partial \rho}{\partial \theta} \right)\mathbf{e}_\rho + v_1 h_\varphi \mathbf{e}_\varphi + v_2 h_\theta \mathbf{e}_\theta \right] \times \\
&\quad \left[\left(w_1\frac{\partial \rho}{\partial \varphi} + w_2\frac{\partial \rho}{\partial \theta} \right)\mathbf{e}_\rho + w_1 h_\varphi \mathbf{e}_\varphi + w_2 h_\theta \mathbf{e}_\theta \right] \\
&= \rho[\sin\theta\,(v_1\nabla_{\mathbf{w}}\rho - w_1\nabla_{\mathbf{v}}\rho)\,\mathbf{e}_\theta + (w_2\nabla_{\mathbf{v}}\rho - v_2\nabla_{\mathbf{w}}\rho)\,\mathbf{e}_\varphi \\
&\quad - \rho\sin\theta\,\det(\mathbf{A})\,\mathbf{e}_\rho],
\end{aligned}$$

(5)

where the last equality follows from the definition of directional derivative and the identities:

$$\mathbf{e}_\rho \times \mathbf{e}_\rho = \mathbf{0}, \quad \mathbf{e}_\varphi \times \mathbf{e}_\varphi = \mathbf{0}, \quad \mathbf{e}_\theta \times \mathbf{e}_\theta = \mathbf{0},$$

$$\mathbf{e}_\rho \times \mathbf{e}_\theta = \mathbf{e}_\varphi, \quad \mathbf{e}_\varphi \times \mathbf{e}_\rho = \mathbf{e}_\theta, \quad \mathbf{e}_\theta \times \mathbf{e}_\varphi = \mathbf{e}_\rho.$$

---

[2]The *directional derivative* of a scalar function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ along a vector $\mathbf{w} \in \mathbb{R}^n$ is the function $\nabla_{\mathbf{w}} f(\mathbf{x})$ defined by the limit $\nabla_{\mathbf{w}} f(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x}+h\mathbf{w}) - f(\mathbf{x})}{h}$. If f is differentiable at $\mathbf{x}$, then the directional derivative exists along any vector $\mathbf{w}$, and one has $\nabla_{\mathbf{w}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{w} \rangle$.

Equation (5) can be equivalently rewritten as:

$$\mathbf{N}(\theta, \varphi) = \rho \sin\theta\, \mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta) \begin{bmatrix} v_1 \nabla_{\mathbf{w}} \rho - w_1 \nabla_{\mathbf{v}} \rho \\ \frac{1}{\sin\theta}(w_2 \nabla_{\mathbf{v}} \rho - v_2 \nabla_{\mathbf{w}} \rho) \\ -\det(\mathbf{A})\rho \end{bmatrix},$$

where $\mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta) = [\mathbf{e}_\theta, \mathbf{e}_\varphi, \mathbf{e}_\rho] \in \mathrm{SO}(3)$. Note that given a vector written in $\{\mathbf{e}_\rho, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$, the rotation matrix $\mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta)$ allows to express it in $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, the basis set of unit vectors for the Cartesian coordinates (cf. Fig. 3). Since for $\mathbf{R} \in \mathrm{SO}(3)$, $\|\mathbf{R}\,\mathbf{x}\| = \|\mathbf{x}\|$, $\forall\, \mathbf{x} \in \mathbb{R}^3$, we have that:

$$\widehat{\mathbf{N}}(\theta, \varphi) = \frac{\mathbf{N}(\theta, \varphi)}{\|\mathbf{N}(\theta, \varphi)\|} =$$

$$\frac{1}{\sqrt{(v_1 \nabla_{\mathbf{w}} \rho - w_1 \nabla_{\mathbf{v}} \rho)^2 + \left(\frac{w_2 \nabla_{\mathbf{v}} \rho - v_2 \nabla_{\mathbf{w}} \rho}{\sin\theta}\right)^2 + \det(\mathbf{A})^2 \rho^2}}$$

$$\cdot \mathbf{R}_z(\varphi)\mathbf{R}_y(\theta) \begin{bmatrix} v_1 \nabla_{\mathbf{w}} \rho - w_1 \nabla_{\mathbf{v}} \rho \\ \frac{1}{\sin\theta}(w_2 \nabla_{\mathbf{v}} \rho - v_2 \nabla_{\mathbf{w}} \rho) \\ -\det(\mathbf{A})\rho \end{bmatrix}.$$

If the scene is Lambertian and it is illuminated from direction $\mathbf{L} = [p_d, q_d, -1]^T$ by a point source at infinity, the irradiance equation for the directions $\mathbf{v}$ and $\mathbf{w}$ is given by:

$$I(\theta, \varphi) = \alpha(\theta, \varphi) \langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}} \rangle =$$

$$\frac{\alpha(\theta, \varphi)\,\mathbf{L}^T}{\|\mathbf{L}\|\sqrt{(v_1 \nabla_{\mathbf{w}} \rho - w_1 \nabla_{\mathbf{v}} \rho)^2 + \left(\frac{w_2 \nabla_{\mathbf{v}} \rho - v_2 \nabla_{\mathbf{w}} \rho}{\sin\theta}\right)^2 + \det(\mathbf{A})^2 \rho^2}}$$

$$\cdot \mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta) \begin{bmatrix} v_1 \nabla_{\mathbf{w}} \rho - w_1 \nabla_{\mathbf{v}} \rho \\ \frac{1}{\sin\theta}(w_2 \nabla_{\mathbf{v}} \rho - v_2 \nabla_{\mathbf{w}} \rho) \\ -\det(\mathbf{A})\rho \end{bmatrix}, \tag{6}$$

where $\alpha(\theta, \varphi)$ is the albedo at point (θ, φ). From equation (4), it follows that $\nabla_{\mathbf{v}} \rho = q\rho$ and that $\nabla_{\mathbf{w}} \rho = p\rho$. By substituting these expressions in (6) and rearranging, the radial distance ρ cancels out, and equation (3) in the statement is obtained. ∎

The set of points on the surface of the sphere where $I(\theta, \varphi) = \kappa$ for constant $\kappa \in [0, 1]$, is called an *isophote* (Pottmann and Wallner, 2010, Sect. 6.3). In other words an isophote is a curve on the surface of the sphere that connects points of equal intensity.

If $\{\mathbf{v}, \mathbf{w}\}$ is the standard basis set of unit vectors, i.e. $\mathbf{v} =$ $[1, 0]^T$ and $\mathbf{w} = [0, 1]^T$, then equation (3) simply reduces to,

$$I(\theta, \varphi) = \frac{\alpha(\theta, \varphi)\,\mathbf{L}^T}{\|\mathbf{L}\|\sqrt{p^2 + \left(\frac{q}{\sin\theta}\right)^2 + 1}}\, \mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta) \begin{bmatrix} p \\ \frac{q}{\sin\theta} \\ -1 \end{bmatrix},$$

where

$$p \triangleq \frac{1}{\rho}\frac{\partial \rho}{\partial \theta} = \frac{\partial \ln \rho}{\partial \theta}, \quad q \triangleq \frac{1}{\rho}\frac{\partial \rho}{\partial \varphi} = \frac{\partial \ln \rho}{\partial \varphi}. \tag{7}$$

Remark 1 (Properties of the spherical image irradiance equation)
As in the perspective case (cf. equation (2)), the spherical image irradiance equation (3) only depends on the directional derivatives of $\ln \rho(\theta, \varphi)$ along $\mathbf{v}$ and $\mathbf{w}$, but not on $\ln \rho(\theta, \varphi)$ itself. As a consequence, the problem of recovering the radial distance $\rho(\theta, \varphi)$ from the image irradiance equation reduces to the problem of recovering $\ln \rho(\theta, \varphi)$ from equation (3). Since the natural logarithm is a bijective function and $\rho(\theta, \varphi) > 0$, recovering $\ln \rho(\theta, \varphi)$ is equivalent to recovering $\rho(\theta, \varphi) = \exp(\ln \rho(\theta, \varphi))$. We also observe that as (2), equation (3) is *invariant to scale changes*, meaning that the intensity functions of $c\,\rho(\theta, \varphi)$ (for constant c) and $\rho(\theta, \varphi)$ are identical. In other words, the radial map can only be reconstructed up to a scale factor. ◇

Photometric stereo relies on several images of the same object observed from an identical viewpoint, under n different illumination conditions. Let us denote the images and the corresponding illumination vectors, $I_i(\theta, \varphi)$ and

$$\mathbf{L}_i \triangleq [p_{d_i}, q_{d_i}, -1]^T, \quad i \in \{0, 1, \dots, n-1\},$$

respectively. Using (3), the ith image irradiance equation is then:

$$I_i(\theta, \varphi) = \frac{\alpha(\theta, \varphi)\,\mathbf{L}_i^T}{\|\mathbf{L}_i\|\sqrt{(v_1 p - w_1 q)^2 + \left(\frac{w_2 q - v_2 p}{\sin\theta}\right)^2 + \det(\mathbf{A})^2}}$$

$$\cdot \mathbf{R}_z(\varphi)\, \mathbf{R}_y(\theta) \begin{bmatrix} v_1 p - w_1 q \\ \frac{w_2 q - v_2 p}{\sin\theta} \\ -\det(\mathbf{A}) \end{bmatrix}.$$

By dividing the ith image by the kth (assuming that the latter is

non-zero everywhere), we obtain:

$$\frac{I_i(\theta, \varphi)}{I_k(\theta, \varphi)} = \frac{\|\mathbf{L}_k\| \mathbf{L}_i^T \mathbf{R}_z(\varphi) \mathbf{R}_y(\theta) [v_1 p - w_1 q, \frac{w_2 q - v_2 p}{\sin \theta}, -\det(\mathbf{A})]^T}{\|\mathbf{L}_i\| \mathbf{L}_k^T \mathbf{R}_z(\varphi) \mathbf{R}_y(\theta) [v_1 p - w_1 q, \frac{w_2 q - v_2 p}{\sin \theta}, -\det(\mathbf{A})]^T}.$$

(8)

By expanding the numerator and the denominator, and by collecting the terms containing p and q, the photometric ratio (8) can be rewritten as,

$$A_{i,k}\, p + B_{i,k}\, q + C_{i,k} = 0, \quad i, k \in \{0, 1, \dots, n-1\}, \quad (9)$$

where

$$A_{i,k} = \Big[I_i \|\mathbf{L}_i\| \langle \mathbf{L}_k, \mathbf{e}_\theta \rangle - I_k \|\mathbf{L}_k\| \langle \mathbf{L}_i, \mathbf{e}_\theta \rangle,$$
$$- \frac{1}{\sin \theta} \big(I_i \|\mathbf{L}_i\| \langle \mathbf{L}_k, \mathbf{e}_\varphi \rangle - I_k \|\mathbf{L}_k\| \langle \mathbf{L}_i, \mathbf{e}_\varphi \rangle \big) \Big] \mathbf{v},$$

$$B_{i,k} = -\Big[I_i \|\mathbf{L}_i\| \langle \mathbf{L}_k, \mathbf{e}_\theta \rangle - I_k \|\mathbf{L}_k\| \langle \mathbf{L}_i, \mathbf{e}_\theta \rangle,$$
$$- \frac{1}{\sin \theta} \big(I_i \|\mathbf{L}_i\| \langle \mathbf{L}_k, \mathbf{e}_\varphi \rangle - I_k \|\mathbf{L}_k\| \langle \mathbf{L}_i, \mathbf{e}_\varphi \rangle \big) \Big] \mathbf{w},$$

$$C_{i,k} = -\det(\mathbf{A})(I_i \|\mathbf{L}_i\| \langle \mathbf{L}_k, \mathbf{e}_\rho \rangle - I_k \|\mathbf{L}_k\| \langle \mathbf{L}_i, \mathbf{e}_\rho \rangle),$$

and I_i is a shorthand for $I_i(\theta, \varphi)$. One may observe that system (9) is *linear* in p and q. Since two images are necessary to construct each one of its equations, then three images are sufficient to recover the two unknowns p and q under *ideal conditions*. However, in the presence of image noise, the least-squares solution to system (9) is clearly preferable. Note that the number of equations in system (9) does not exceed $N_{\text{eq}} = \binom{n}{2} = n(n-1)/2$.

Some remarks are in order at this point.

Remark 2 (Albedo estimation)

Note that equation (8) is *independent* of albedo (in fact, by computing the photometric ratio, the coefficients $\alpha(\theta, \varphi)$ cancel out). Hence, the approach described above recovers the gradient field (p, q) without providing any information about the albedo. Computing the albedo amounts to a simple inverse rendering problem: to this end, one can use, for example, the closed-form formula in (Smith and Fang, 2016, Sect. 4.1). ⬦

Remark 3 (Illumination vectors: special configurations)

With reference to system (9) and Fig. 3, we observe the following special configurations for the illumination vectors $\mathbf{L}_i$, $\mathbf{L}_k$:

- If $\mathbf{L}_i$, $\mathbf{L}_k$ are orthogonal to $\mathbf{e}_\rho$ for all i, k, then system (9) only admits the trivial solution $p = q = 0$. In fact, in this case $\langle \mathbf{L}_k, \mathbf{e}_\rho \rangle = \langle \mathbf{L}_i, \mathbf{e}_\rho \rangle = 0$ and $C_{i,k} = 0$. Geometrically speaking, this means that the illumination vectors lie in a tangent plane to the sphere (raking light), and that the isophotes satisfying $I(\theta, \varphi) = 0$, are lines of latitude (i.e. circles on the sphere parallel to the equator).

- If $\mathbf{L}_i$, $\mathbf{L}_k$ are orthogonal to $\mathbf{e}_\theta$ for all i, k and $v_2 = 0$, then $A_{i,k} = 0$ and only q can be computed (p is undetermined). This means that the illumination vectors belong to a plane tangent to the right-circular cone with apex at origin, shown in Fig. 3.

- If $\mathbf{L}_i$, $\mathbf{L}_k$ are orthogonal to $\mathbf{e}_\varphi$ for all i, k and $w_1 = 0$, then $B_{i,k} = 0$ and only p can be computed (q is undetermined). This means that the illumination vectors are radial to the sphere in a meridian half-plane.

- If either $\mathbf{L}_i$ or $\mathbf{L}_k$ is orthogonal to both $\mathbf{e}_\varphi$ and $\mathbf{e}_\theta$ for all i, k, system (9) simplifies, but p and q can be both computed. In this case, the illumination vectors are generically radial to the sphere. ⬦

Remark 4 (Ambient illumination)

Note that by introducing a third image $I_j(\theta, \varphi)$, equation (8) can be made *invariant* to ambient illumination. In fact, if the irradiance equation relative to direction $\mathbf{L}_i$ is $I_i(\theta, \varphi) = \alpha(\theta, \varphi)(\langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}}_i \rangle + a)$ where $a > 0$ is the constant ambient light, then by taking the ratio of differences between images, i.e. by computing, for example,

$$\frac{I_i(\theta, \varphi) - I_j(\theta, \varphi)}{I_k(\theta, \varphi) - I_j(\theta, \varphi)} = \frac{\langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}}_i \rangle - \langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}}_j \rangle}{\langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}}_k \rangle - \langle \widehat{\mathbf{N}}(\theta, \varphi), \widehat{\mathbf{L}}_j \rangle},$$

the ambient term cancels out. However, a known pitfall of this approach (Smith and Fang, 2016), is that the estimation of the gradient field becomes more sensitive to noise. ⬦

Remark 5 (Light attenuation)

Let f(d) be a generic *light attenuation function*, where d is the distance between the light source and a point on the surface $\mathcal{S}$. This is a decreasing function of the distance: for example

in (Prados and Faugeras, 2005; Liao et al., 2007), the authors considered a simple inverse square law, i.e. $\mathrm{f}(d) = 1/d^2$. Then, by defining $I_i^{\mathrm{f}}(\theta, \varphi) = \mathrm{f}(d_i)\, I_i(\theta, \varphi)$ where d_i is the distance of the ith light source, the photometric ratio (8) becomes,

$$\frac{I_i^{\mathrm{f}}(\theta, \varphi)}{I_k^{\mathrm{f}}(\theta, \varphi)} = \frac{\mathrm{f}(d_i)}{\mathrm{f}(d_k)} \frac{I_i(\theta, \varphi)}{I_k(\theta, \varphi)}, \tag{10}$$

from which a set of linear equations in p and q similar to (9) can be derived. However, to solve this system of equations, the distances d_i, d_k need to be known, and to accurately estimated them is not a trivial task in practice. If all distances are identical (i.e. $d_i = d_k, \forall i, k$), then the attenuation factor in (10) cancels out, and we obtain equation (8). $\qquad \diamond$

To construct system (9), we simply discarded the zero intensities. However, in the last decade, a popular class of approaches has emerged in which only a subset of observed intensities for each pixel is chosen (Barsky and Petrou, 2003; Hernández et al., 2010). The idea behind these *selection-based approaches* is to exclude observations that are believed to deviate from the modeling assumptions made by the photometric stereo algorithm (e.g. by discarding shadowed pixels). While these methods could be potentially used to appropriately choose the "best" combinations, we decided not to apply them here in order to keep our 3D reconstruction pipeline simple and the message focused on our generic camera model.

3.2. Integration of gradient field

System (9) allows to recover the gradient field (p, q) of the 3D surface $\mathcal{S}$. However, to reconstruct the unknown *radial map* $\rho(\theta, \varphi)$ (in contrast to the unknown *depth map* $z(\mathrm{u}, \mathrm{v})$ in the perspective case), the gradient field needs to be integrated. A plethora of integration methods exists in the literature (Quéau et al., 2018a). The variational approach is among the earliest and most popular (Quéau et al., 2018b): it considers a least-squares cost function in its continuous form, and it solves the corresponding 2D Poisson equation. In (Horovitz and Kiryati, 2004), iterative methods based on the Gauss-Seidel scheme or its variants (Successive Over-Relaxation and Full Multigrid) are used for numerical integration. Path integrals and grid-based basis functions have

been also proposed by some authors for the integration step (Ackermann and Goesele, 2015).

The direct (i.e. non-iterative) algebraic approach proposed in (Harker and O'Leary, 2008) has several attractive features: uniqueness of solution (up to a constant of integration) and numerical stability, which allows for reliable surface reconstruction even for high-resolution images, among others. Assuming an orthographic image projection model and a gradient field corrupted by Gaussian noise, the authors in (Harker and O'Leary, 2008) formulated the problem as that of reconstructing a discrete surface whose derivatives are equal to p and q in the least-squares sense. With a matrix definition of the cost function, standard linear algebra is used to find the unique least-squares minimum, which satisfies a Sylvester matrix equation (see (Harker and O'Leary, 2015) for further details). Finally, numerical differentiation, which takes the form of a matrix multiplication, yields the depth of the reconstructed surface.

We will use the method of Harker & O'Leary in Sect. 4, for comparison purposes. In the next section, we will describe a drift-free approach for the integration of the gradient field, tailored to the spherical image geometry.

Geometry-aware normal integration method

To account for the constraints on the (θ, φ) grid, we designed a normal integration method which explicitly handles them. We have seen in Sect. 3.1 that if $\{\mathbf{v}, \mathbf{w}\}$ is the standard basis set of unit vectors, then p and q are defined as in equation (7). Using a simple finite-difference method to approximate the two partial derivatives in (7), we obtain:

$$
\begin{aligned}
p &= \frac{\partial \ln \rho}{\partial \theta} \simeq \frac{\ln \rho(\theta + \Delta\theta, \varphi) - \ln \rho(\theta, \varphi)}{\Delta\theta}, \\
q &= \frac{\partial \ln \rho}{\partial \varphi} \simeq \frac{\ln \rho(\theta, \varphi + \Delta\varphi) - \ln \rho(\theta, \varphi)}{\Delta\varphi},
\end{aligned}
\tag{11}
$$

where $\Delta\theta$ and $\Delta\varphi$ denote the step-sizes along the θ- and φ-direction, respectively. From (11), a linear system of equations (one for each node of the grid) can be constructed and solved using the least-squares method. The sparsity of this linear system can be exploited for fast numerical solution (a speed
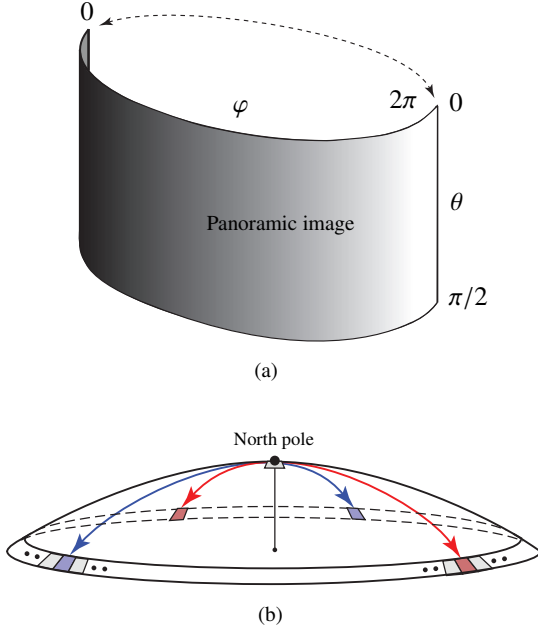
Fig. 4. Spherical constraints imposed in our drift-free normal integration method: (a) *Circularity constraint* (note that only the northern hemisphere was considered in our implementation, see Sect. 4), and (b) *North-pole constraint*. The distance between the North pole and the surrounding nodes of the discrete grid has been exaggerated for ease of illustration.

up of a few orders of magnitude was observed in our tests). Note that the two equations in (11) are well-defined everywhere except for the nodes on the boundary of the grid (a von Neumann neighborhood is considered). To address this issue, we augmented the system of equations with two additional constraints. The first one, cf. (Frankot and Chellappa, 1988), is a periodic boundary condition referred to as the *Circularity constraint*. In fact, since the radial distance must take on the same values along opposite azimuth boundaries, we enforced $\rho(\theta, \varphi + \Delta\varphi) = \rho(\theta, 0)$ for $\varphi = 2\pi$ with $\theta \in (0, \pi/2]$ (i.e., the rectangular grid is wrapped around, see Fig. 4(a)). The second constraint is imposed on the contiguous nodes lying on the annulus around the North pole (i.e. the node with $\theta = 0$, a singularity of the spherical coordinate representation): all these nodes are considered to be neighbors. This constraint is referred to as *North-pole constraint* (see Fig. 4(b)). In Sect. 4, we will show that the introduction of the Circularity and North pole constraints is beneficial in drastically reducing the drift experienced with the method of Harker & O'Leary.

Fig. 5 reports the complete flowchart of our photometric stereo algorithm. For visualization purposes, note that the reconstructed radial map is displayed in Cartesian coordinates, and that $\overline{X}$ in Fig. 5 denotes the estimated Cartesian coordinates of a generic point on it (up to a scale factor).

4. Experimental validation

In this section, we present the results of an extensive experimental campaign. We evaluated our photometric stereo algorithm in a number of different scenarios, going from analytical surfaces (Sect. 4.1) to synthetic (Sect. 4.2) and real-world images (Sect. 4.3) taken by perspective and catadioptric cameras. The use of synthetic images allowed us to arbitrarily control the illumination conditions, and easily test the robustness of our approach. For the sake of simplicity, in Sect. 4.1 and Sect. 4.2 we ignored the second projection step (from the unit sphere to the image plane, cf. Sect. 2.1). In fact, as long as the intrinsic parameters ℓ and m are known, the process is fully reversible. Moreover, the light directions were assumed to be perfectly known. Therefore, only the last three blocks of the flowchart in Fig. 5 were considered. Finally, in all cases, we assumed that $\{\mathbf{v}, \mathbf{w}\}$ is the standard basis set of unit vectors.

4.1. Analytical surfaces

The proposed method has been validated using two test surfaces whose mathematical expression is known. By illuminating the surfaces from different directions, a set of synthetic images was generated, which we injected into our 3D reconstruction pipeline. The first surface will be referred to as *Volcano*, and its analytical expression, given here expediently in spherical coordinates, is (see Fig. 6(a)):

$$\rho(\theta, \varphi) = \frac{9}{\cos\theta} + 100\sin\theta. \quad (12)$$

Note that since (12) is axially symmetric, it does not depend on parameter φ. The second surface (see Fig. 6(d)), called *Starfish*, has the following expression:

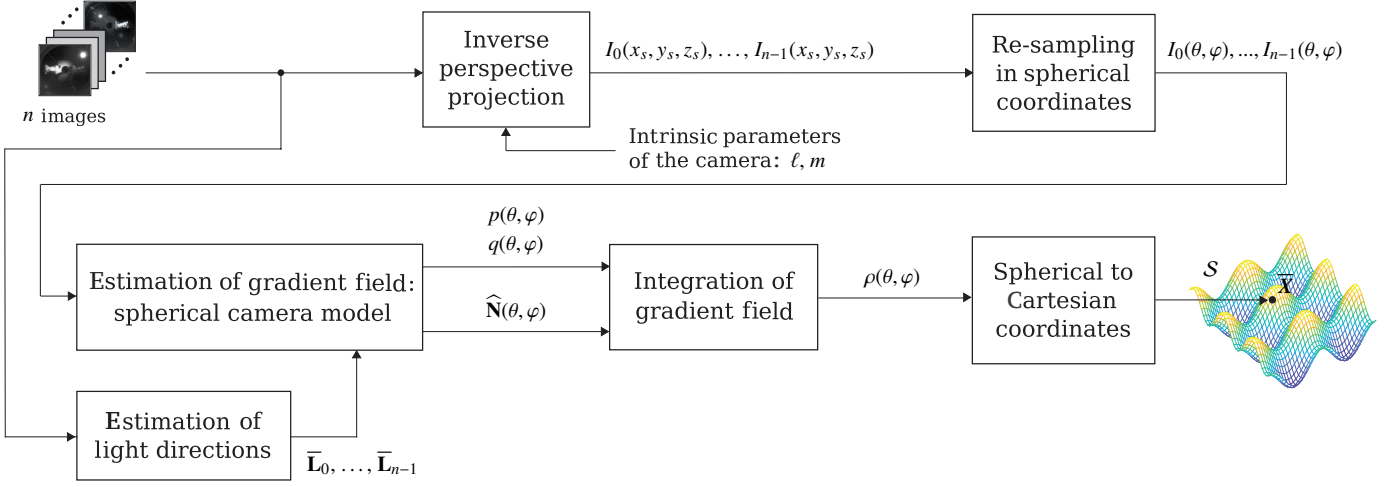$$\rho(\theta, \varphi) = 6 + \sin\left(\frac{\theta}{10} + 5\varphi\right).$$

Fig. 5. Flowchart of the proposed photometric stereo algorithm. The algorithm takes n images from a central panoramic camera, as input.

To precisely assess the 3D reconstruction quality, we considered the angular deviation between the actual and estimated normal vectors:

$$\varepsilon(\theta, \varphi) = \arccos(\langle \mathbf{N}(\theta, \varphi), \overline{\mathbf{N}}(\theta, \varphi)\rangle), \qquad (13)$$

with $\theta \in \left\{\frac{\pi}{\lambda}, \frac{2\pi}{\lambda}, \ldots, \frac{\pi}{2} - \frac{\pi}{\lambda}\right\}$, $\varphi \in \left\{0, \frac{\pi}{\lambda}, \ldots, 2\pi - \frac{\pi}{\lambda}\right\}$, where $\overline{\mathbf{N}}(\theta, \varphi)$ is the normal at (θ, φ) estimated by our method, and $1/\lambda$ is the sampling interval ($\lambda = 100$ in our tests). Note that only the northern hemisphere was considered in our tests, i.e. θ is between 0 and $\pi/2$. For the *Volcano* and *Starfish*, the image intensities $I_i(\theta, \varphi)$, $i \in \{0, 1, 2, 3\}$, corresponding to the light directions,

$$\mathbf{L}_0 = [1/4, \ 1/4, \ -1]^T, \qquad \mathbf{L}_1 = [-1/4, \ 1/4, \ -1]^T,$$

$$\mathbf{L}_2 = [-1/4, \ -1/4, \ -1]^T, \quad \mathbf{L}_3 = [1/4, \ -1/4, \ -1]^T,$$

have been corrupted with additive zero-mean white Gaussian noise with variance σ^2 (the four illumination vectors were placed in symmetric positions to irradiate the two surfaces as uniformly as possible). A saturation function was applied to the noisy image intensities to keep them within the [0, 1] interval. For the sake of simplicity and without loss of generality, we assumed that all the 3D points of the test surfaces were visible from the camera, i.e. no self-occlusions. Figs. 6(b),(f) report the boxplots (Matlab) of the error $\varepsilon(\theta, \varphi)$ (in degrees) over the two test surfaces, for an increasing magnitude of the image noise ($\sigma \in \{0, 0.1, \ldots, 0.5\}$). Note that $\varepsilon(\theta, \varphi) \neq 0$ for $\sigma = 0$.

In fact, the information about the 3D scene provided by the four illumination vectors is not rich enough to guarantee a perfect reconstruction. However, with 16 light directions, we observed a 25% error reduction. Finally, Figs. 6(c) and 6(g) show the distribution of the error $\varepsilon(\theta, \varphi)$ (bin size 0.5°) for $\sigma = 0.3$, and Figs. 6(d) and 6(h) the distribution of the *scaled* radial distance error $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ (bin size 0.01) for $\sigma = 0.3$, where,

$$\rho_s(\theta, \varphi) \triangleq \frac{\rho(\theta, \varphi)}{\max\limits_{\theta, \varphi} \rho(\theta, \varphi)},$$

by using the our normal integration method with step-sizes $\Delta\theta = \Delta\varphi = \pi/\lambda$. By definition, $\rho_s(\theta, \varphi) \in [0, 1]$. An inspection of Fig. 6(h) reveals that the radial distance error is higher in correspondence with the central narrow ridges of the *Starfish*, as expected.

4.2. Synthetic images

To test the effectiveness of the proposed reconstruction method under more realistic conditions, we used Blender (Blender Foundation, 2018) to create a 3D scene of known geometry observed from a given viewpoint (our ground truth). We then generated panoramic images under 8 different illumination conditions from this unique viewpoint. As in Sect. 4.1, we assumed that the illumination vectors $\mathbf{L}_0$, $\ldots$, $\mathbf{L}_7$ are known (and arranged symmetrically about the scene), and that all the points on the 3D surfaces are visible from the camera. To limit the number of input images to
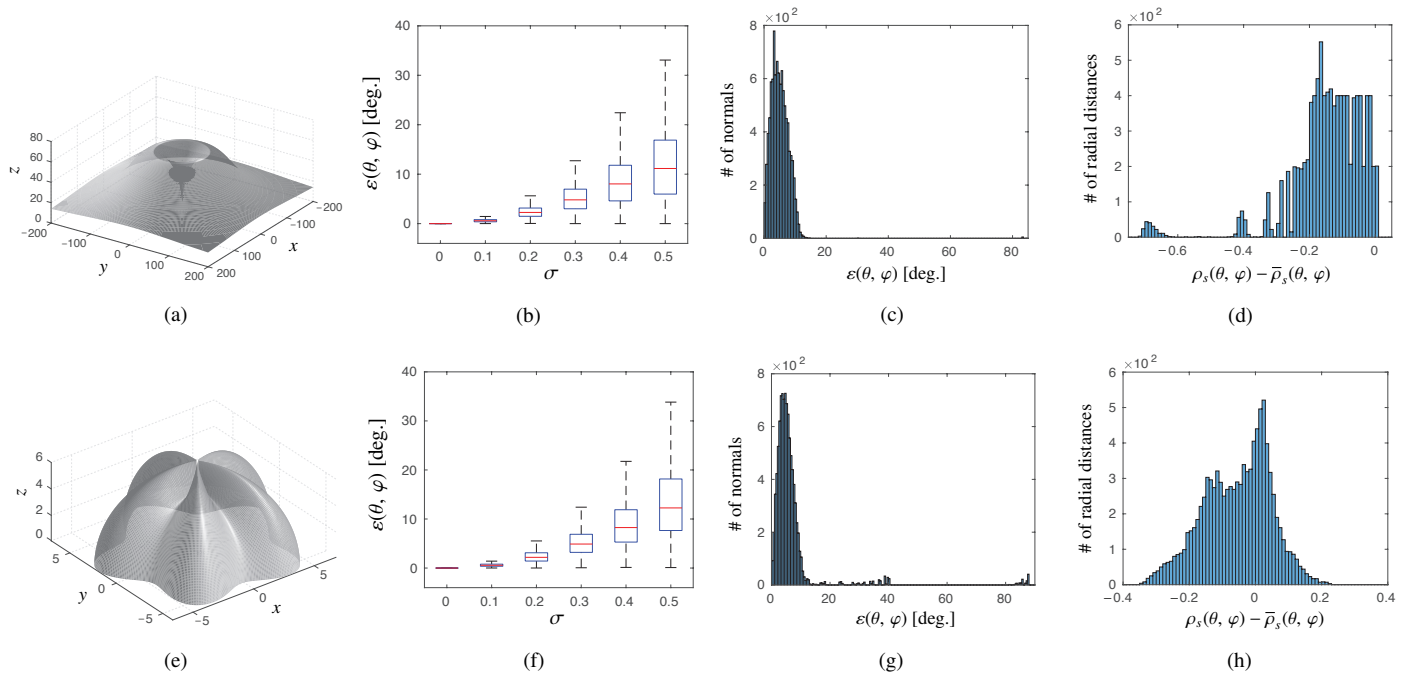
Fig. 6. Analytical surfaces: (a) *Volcano* **and (d)** *Starfish*; **(b),(f) Boxplots of the error** $\varepsilon(\theta, \varphi)$ **for an increasing magnitude of the additive Gaussian noise on the image intensities; (c),(g) Distribution of the error** $\varepsilon(\theta, \varphi)$ **(bin size** $0.5°$**) for** $\sigma = 0.3$**; (d),(h) Distribution of the scaled radial distance error** $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ **(bin size** 0.01**) for** $\sigma = 0.3$**.**

the algorithm, the self-occlusions were not taken into account. Finally, we set $\lambda = 1024$, and for more realism, we simulated the quantization effect on 8-bit intensity images by mapping $I_i(\theta, \varphi) \in [0, 1] \subset \mathbb{R}$, $i \in \{0, 1, \dots, 7\}$, onto the discrete set $\{0, 1, \dots, 255\}$.

In our first test, we considered the *Stanford Bunny*[3]. Fig. 7(a) shows the reconstructed 3D surface for $\sigma = 0.3$, Fig. 7(b) the boxplot of the error $\varepsilon(\theta, \varphi)$ for an increasing magnitude of the image noise, and Fig. 7(c) the distribution of the error $\varepsilon(\theta, \varphi)$ (bin size $0.5°$) for $\sigma = 0.3$. Finally, Fig. 7(d) reports the distribution of the error $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ (bin size 0.01) for $\sigma = 0.3$, relative to the area of *Stanford Bunny* inside the red circle shown in Fig. 7(a) (the vertical axes in Figs. 7(c) and 7(d) refer to the number of normals and radial distances in the reconstructed surface, respectively). We focused on this portion of the scene to eliminate the artifacts due to depth discontinuities and kinks (i.e. where the surface is continuous, but non-differentiable), and perform a pertinent statistical eval-

uation of our normal integration algorithm (the step-sizes are $\Delta\theta = \Delta\varphi = \pi/\lambda$).

For the sake of comparison, we modified the method based on the orthographic image projection model in (Harker and O'Leary, 2008), to fit our spherical formulation. In particular, the (x, y) rectangular grid was replaced by a (θ, φ) grid, and the solution of the Sylvester matrix equation becomes, in our case, the radial distance $\rho(\theta, \varphi)$ between the camera center O and the points on the 3D surface, instead of the depth $z(u, v)$. Compared to the same 3D reconstruction of the *Stanford Bunny* obtained with the method of Harker & O'Leary, the Circularity and North pole constraints are instrumental to reduce the drift and the Gibbs phenomena around the kinks (see Fig. 8).

In our second test, we emulated the image acquisition process of a panoramic camera placed at the center of a 3D synthetic scene, a *Spiral staircase*. As already mentioned in the Introduction (see Fig. 1), the reconstruction of this scene in its entirety pose a significant challenge for standard perspective cameras, because of their limited field of view. Similarly to
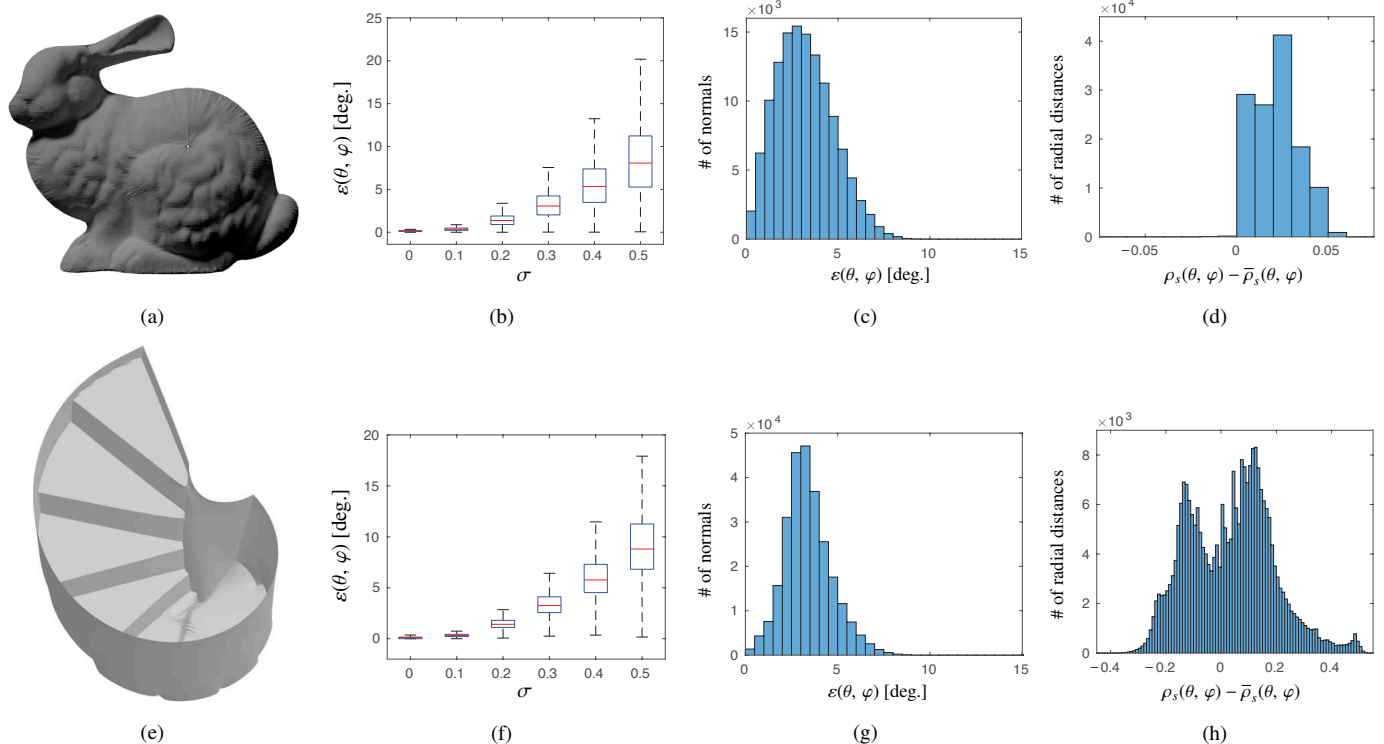
---

[3]The Stanford 3D Scanning Repository.

Fig. 7. Synthetic images: *Stanford Bunny*, **(a) Reconstruction of the 3D surface for** $\sigma = 0.3$, **(b) Boxplot of the error** $\varepsilon(\theta, \varphi)$ **for an increasing magnitude of the additive Gaussian noise on the image intensities, (c) Distribution of the error** $\varepsilon(\theta, \varphi)$ **(bin size: 0.5°) for** $\sigma = 0.3$, **(d) Distribution of the error** $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ **(bin size: 0.01) relative to the red circled area shown in (a), for** $\sigma = 0.3$; *Spiral staircase*, **cf. Fig. 1, (e) Reconstruction of the 3D surface for** $\sigma = 0.3$, **(f) Boxplot of** $\varepsilon(\theta, \varphi)$ **for an increasing magnitude of the additive Gaussian noise on the image intensities, (g) Distribution of the error** $\varepsilon(\theta, \varphi)$ **(bin size: 0.5°) for** $\sigma = 0.3$, **(h) Distribution of the error** $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ **(bin size: 0.01) for** $\sigma = 0.3$.

Stanford Bunny, Fig. 7(e) reports the reconstructed 3D surface for $\sigma = 0.3$, Fig. 7(f) the boxplot of $\varepsilon(\theta, \varphi)$ for an increasing magnitude of the image noise, and Figs. 7(g) and 7(h), the distribution of the errors $\varepsilon(\theta, \varphi)$ and $\rho_s(\theta, \varphi) - \overline{\rho}_s(\theta, \varphi)$ for $\sigma = 0.3$, respectively. Again, the step-sizes are $\Delta\theta = \Delta\varphi = \pi/\lambda$. In both tests, the proposed reconstruction method provided accurate results in the presence of noisy measurements.

4.3. Real-world images

In the next subsection, we will describe the hardware used in our experiments and the method to estimate the light directions from the specular highlight on a mirror ball. We will then report the 3D reconstruction results obtained with real images from a perspective and a catadioptric camera. We will conclude the section with a discussion about some possible improvements of our 3D reconstruction pipeline.

4.3.1. Material

To evaluate the performance of our 3D reconstruction pipeline in a real setting, we built a 2 m × 1.25 m × 1.12 m booth consisting of an aluminum frame covered by a black curtain, to shield any parasitic room light (see Fig. 9). Two expanded polystyrene foam objects, whose white matte surfaces agree reasonably well with the Lambertian model, *Cat* (23 cm tall and 14 cm at the widest point) and *Teddy Bear* (19 cm tall and 12 cm at the widest point), were placed inside the booth and they were illuminated with a 220 V, 5.3 W OSRAM 5-LED lamp with the following optical specifications:

- Luminous flux: 350 lm,

- CIE R_a: 80,

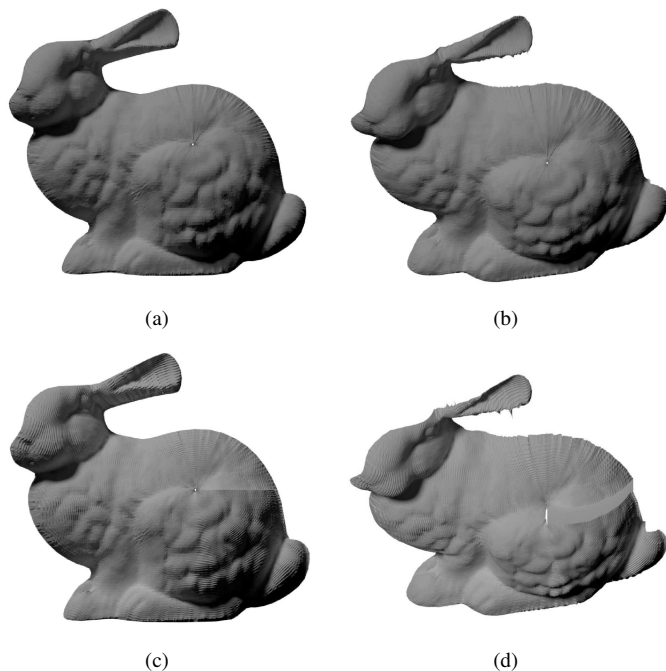- Temperature: 4000 K (cool white),

- Beam angle: 36°.

(a)   (b)

(c)   (d)

Fig. 8. Synthetic images, Stanford Bunny: Reconstruction of the 3D surface obtained with our method (top row), and the method of Harker & O'Leary (bottom row), for $\sigma = 0.3$; (a),(c) frontal viewing angle, and (b),(d) high viewing angle. As evident from (d), the drift deforms the back of the Bunny.

The images were captured by an IDS UI-1245LE-C-HQ RGB camera: we used a Tamron objective (8 mm, $\varnothing$ 25.5 mm) for the acquisition of the perspective images (size: 1280 pixels $\times$
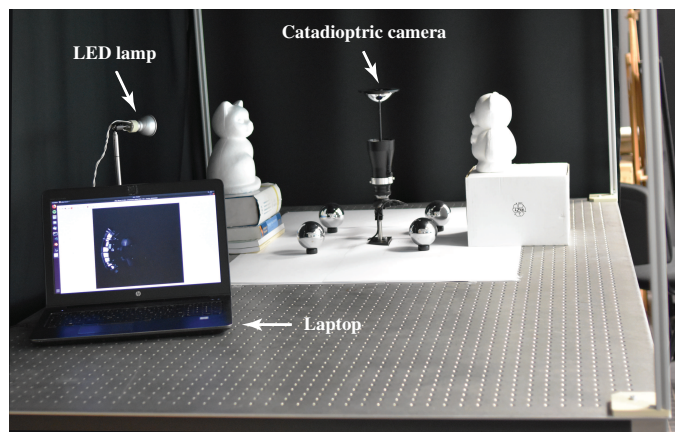

Fig. 9. Experimental setup: 5-LED lamp, test objects (*Cat*, left, and *Teddy Bear*, right) and catadioptric camera. The IDS camera faces upwards and it is connected to a laptop via a USB port for the acquisition of the panoramic images. During the experiments, we drew the black curtains which fully cover the booth: they are shown open for illustration purposes only.

1024 pixels) and a RemoteReality parabolic mirror with a telecentric lens screwed on the IDS camera for the catadioptric ones (size: 1024 pixels $\times$ 1024 pixels). The 24-bit RGB images were converted to 8-bit grayscale before being injected into the 3D reconstruction pipeline. In order to estimate the Camera Response Function (CRF), we took a sequence of images of the same scene with 16 different exposure times (we fixed gamma to 1). We then applied the method by (Grossberg and Nayar, 2004), but this led to unsatisfactory results. In fact, we observed a large discrepancy between the estimated CRF and the curve obtained by plotting the pixel intensities as a function exposure time, which is indeed linear. Hence, following (Diaz and Sturm, 2013), we simply fitted an affine function to the data points to determine the CRF.

In order to estimate the illumination vectors $\mathbf{L}_0, \ldots, \mathbf{L}_{n-1}$ using real catadioptric images (see the bottom left block in Fig. 5), we adapted the light calibration method proposed in (Schnieders and Wong, 2013) for perspective cameras, to the generic setting considered in this paper. This method consists in observing the specular highlight produced by the LED lamp on a mirror ball of *known* radius (for more details, see (Caracotte et al., 2021)). It turned out to be quite reliable in our experiments, but other (possibly less invasive) methods can used for the estimation of light directions as well.
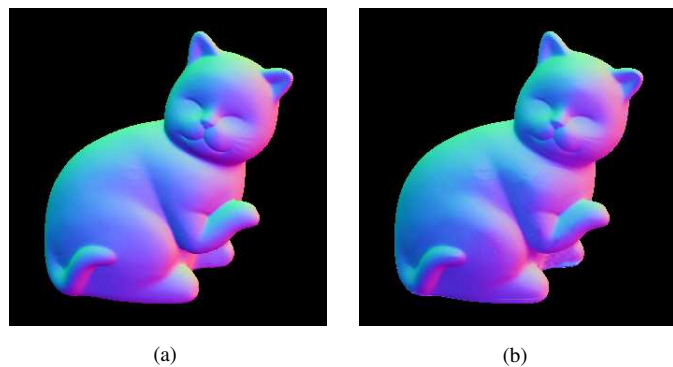

(a)   (b)

Fig. 10. Perspective images of a cat from the DiLiGenT dataset (Shi et al., 2019): (a) Ground truth of the normal map in false colors; (b) Normal map estimated with our generic camera model.

Table 1. Perspective images of a cat from the DiLiGenT dataset (Shi et al., 2019): Mean angular error (MAE) in degrees of some state-of-the-art algorithms and ours.

Algorithm	Baseline	WG10	IW12	GC10	AZ08
MAE	8.41	6.72	7.21	8.21	6.53
Algorithm	HM10	ST12	ST14	IA14	**Ours**
MAE	8.40	12.34	6.12	6.74	8.36

4.3.2. Perspective images

In our first test, we quantified the accuracy of the estimated gradient field on perspective images (obviously, our generic camera model includes pinhole cameras). To this end, we considered the 96 images of a cat under different illumination conditions, available in the 'DiLiGenT' dataset (Shi et al., 2019), hence, $N_{eq} = 4560$, cf. Sect. 3.1. The image intensities were scaled in the [0, 1] interval, and the values below the 0.05 threshold were set to zero to filter out image noise. Fig. 10(a) reports the ground truth of the normal map in false colors, and Fig. 10(b) the normal map estimated with our generic model. With the exception of few minor differences on the left hind leg and right shoulder, the two normal maps are identical. Table 1 reports a more quantitative evaluation of the results. In fact, the *mean angular error* (MAE, recall (13)), obtained with our method is compared with that of some state-of-the-art photometric stereo algorithms: Baseline (Woodham, 1980), WG10 (Wu et al., 2010), IW12 (Ikehata et al., 2012), GC10 (Goldman et al., 2010), AZ08 (Alldrin et al., 2008), HM10 (Higo et al., 2010), ST12 (Shi et al., 2012), ST14 (Shi et al., 2014), IA14 (Ikehata and Aizawa, 2014), see (Shi et al., 2019, Sect. 2) for more details. The accuracy of the normal vectors estimated with our generic camera model is comparable to that of the Baseline, as expected. In our second test, we performed a qualitative evaluation of our 3D reconstruction algorithm using 20 perspective images of the polystyrene foam *Cat*, taken inside the booth shown in Fig. 9 with our IDS camera. For the estimation of the light directions, we placed two steel pin-

ball balls with a radius $r_m = 12.5$ mm in front of the object (see Fig. 11(a)). HySCaS (Caron and Eynard, 2011) was used for the calibration of the camera, yielding $\xi = 0.82655$ where ξ is the parameter of Barreto's unifying model for central projection systems (Barreto, 2006). Fig. 11(b) shows the estimated normal map in false colors after the application of a mask, and Fig. 11(c) the albedo map estimated with the method in (Smith and Fang, 2016, Sect. 4.1). Finally, Fig. 11(d) reports the 3D reconstruction of the *Cat* obtained with our geometry-aware normal integration method with step-sizes $\Delta\theta = \Delta\varphi = \pi/\lambda$, $\lambda = 1024$. Since the node of the integration grid with $\theta = 0$ lies in the middle of the scene, we could apply the Circularity and North-pole constraints. In spite of a slight leveling and distortion along the radial directions towards the North pole, the 3D shape of the *Cat* has been correctly retrieved by the normal integration method.
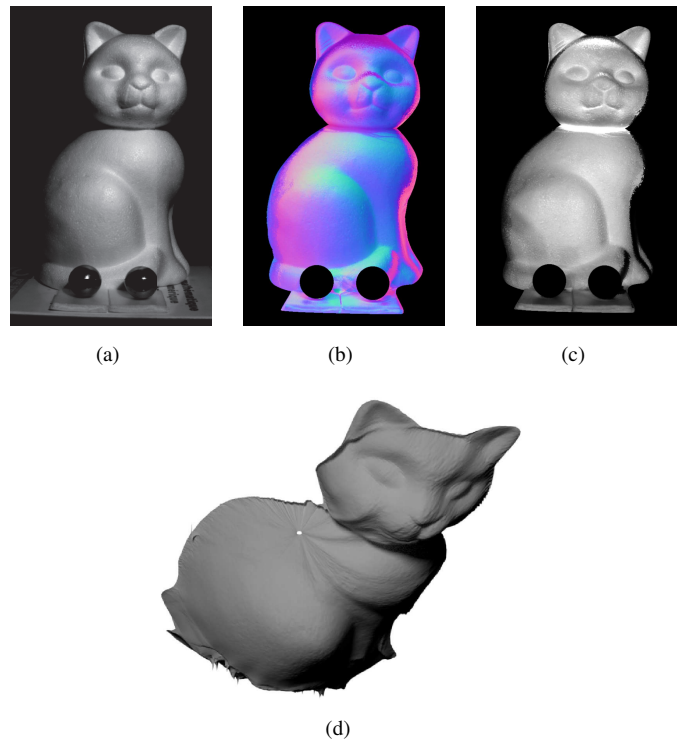

(a) (b) (c)

(d)

Fig. 11. Perspective images: (a) One out of the 20 input images of the *Cat*; (b) Normal map in false colors, estimated with our generic camera model; (c) Albedo map, and (d) 3D reconstruction of the object.

4.3.3. Catadioptric images

A battery of tests was performed with catadioptric images of the *Cat* (26 images) and *Teddy Bear* (16 images). Our experimental setup is shown in Fig. 12. The camera points towards the ceiling of the booth, to provide a better coverage of the two objects. Four steel balls with a radius $r_m = 35.5$ mm were placed inside the scene for the determination of the illumination vectors. At least two balls were visible in each image, providing redundant information for robust estimation. The direction of the illumination vectors is shown in Fig. 13. Again, the camera was calibrated with HySCaS. Fig. 14 illustrates the whole reconstruction procedure. Figs. 14(a), (b) report two catadioptric images of the *Cat* and the *Teddy Bear*. The LED lamp is visible as a bright spot on the center left and top right of Fig. 14(a) and Fig. 14(b), respectively, and that the maximum distance from the camera, corresponding to one of the top corners of the booth, is 170 cm. The estimated normal map in false colors and the albedo map estimated with the method in (Smith and Fang, 2016) are shown in Figs. 14(c), (d), respectively, while the 3D reconstruction of the two objects obtained with our normal integration method for $\Delta\theta = \Delta\varphi = \pi/\lambda$ and $\lambda = 1024$, is reported in Fig. 14(e). To handle the discontinuities due to the background and apply the Circularity and North-pole constraints, in the integration phase we introduced five mutually-orthogonal synthetic planes of normals around each object (four walls and the floor), and then combined the 3D reconstructions of the *Cat*

Fig. 12. Real-world experiments: catadioptric camera, *Cat* and *Teddy Bear*, and the four mirror balls used for the estimation of the illumination vectors.
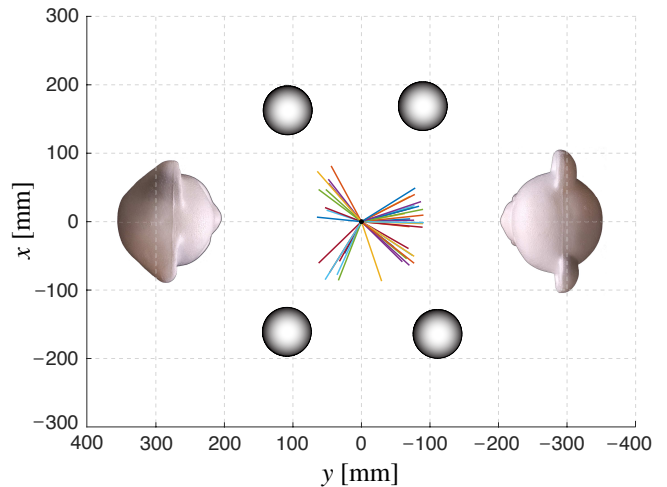

Fig. 13. Top view of the illumination vectors (colored) in the camera frame.

and *Teddy Bear*. From Fig. 14(e) we observe that, while slightly flattened, the 3D shape of the two objects has been correctly estimated. This is all the more remarkable, considering a number of practical limitations of our proof-of-concept implementation: the image distortion has not been corrected, the illumination is not spatially homogeneous (the light cone of the LED lamp has an aperture of $36°$ and produced an undesirable bright disk on the objects), the light attenuation has not been taken into account (cf. Remark 5), and the projection model of the camera is quasi-central (as a consequence, the ears of the *Cat* and *Teddy Bear* are slightly trimmed).

4.3.4. Discussion

We conclude this section with a discussion about some possible improvements of our 3D reconstruction pipeline. Our focus will be on the impact of non-directional lighting, image re-sampling, and two-step recovery of radial distance, on the quality of the recovered surface.

On directional lighting

In our experiments, the assumption of directional lighting is not strictly satisfied. It is then worth quantifying which is the effect of the distance of the point light source on the 3D reconstruction accuracy. To evaluate to which extent the theoretical assumption of a point light source at infinity is critical for photometric stereo with central catadioptric cameras, some
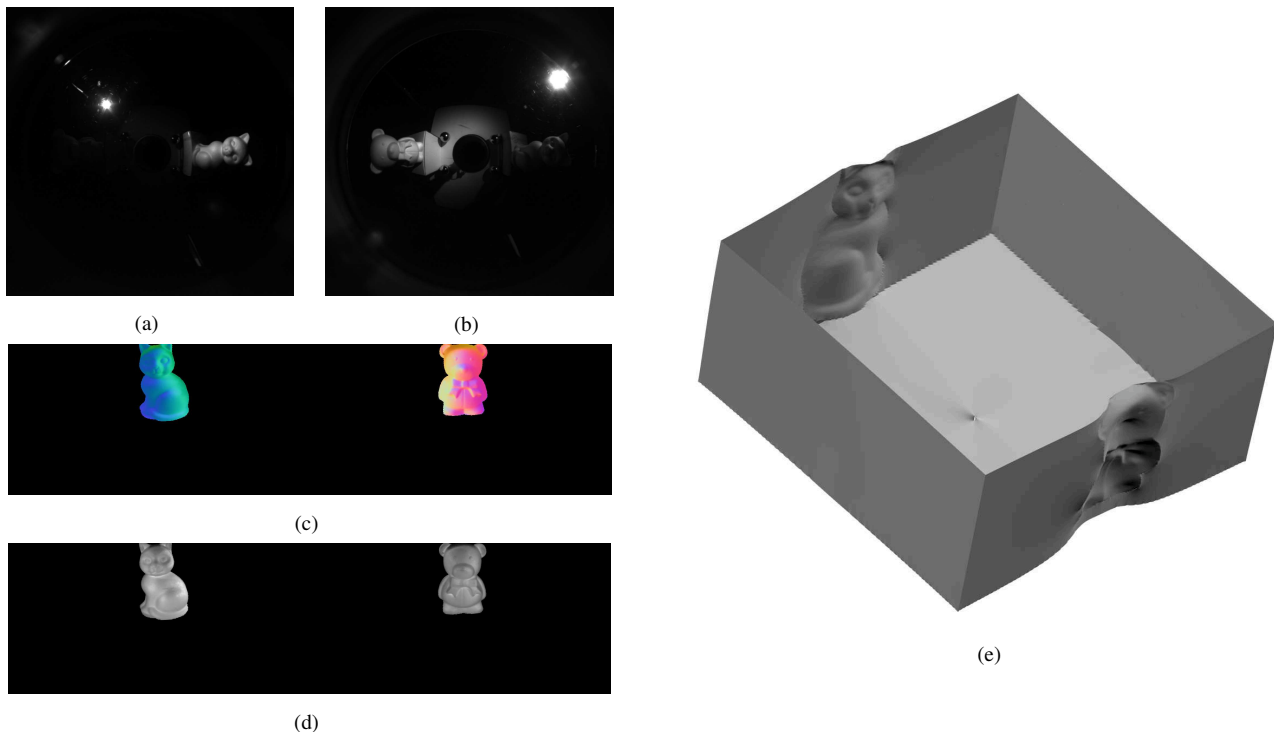
(a)

(b)

(c)

(d)

(e)

Fig. 14. Catadioptric images: (a),(b) Sample input images of the *Cat* and *Teddy Bear*, respectively; (c) Normal map in false colors, estimated with our generic camera model, and (d) albedo map; (e) 3D reconstruction of the objects. The images in (c) and (d) were turned upside down, for ease of visualization.
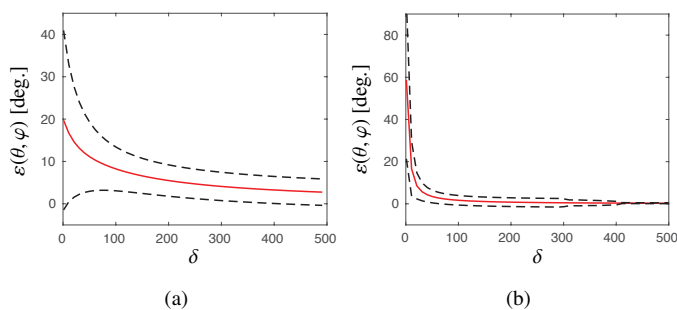

(a)

(b)

Fig. 15. Impact of the distance δ of the point light source on the normal estimation error: mean (solid red) and standard deviation (± 1 sigma, dashed black) of the error $\varepsilon(\theta, \varphi)$, for (a) the *Stanford Bunny* (size: 100 units) in the perspective case, considering the whole object, and (b) the *Spiral staircase* (size: 30 units) in the catadioptric case.

tests have been performed. Figs. 15(a),(b) show the mean (solid red) and standard deviation (± 1 sigma, dashed black) of the error $\varepsilon(\theta, \varphi)$ for the *Stanford Bunny* and the *Spiral staircase*, respectively, as a function of the distance δ of the light source (no image noise). The size of the first object (length of the diagonal of the bounding box) is 100 units, and the size of the second, 30 units. From Fig. 15(b), we observe that for $\delta > 450$

(i.e. 15 times the size of the *Spiral staircase*), the light source can be deemed to be at "infinity", and it has virtually no effect on the error $\varepsilon(\theta, \varphi)$.

On the partition of the sphere

The choice of *constant* polar and azimuthal angular steps $\Delta\theta$ and $\Delta\varphi$, yields a rectangular grid which can be easily processed by our normal integration method. However, it requires an interpolation (the re-sampling step) that might alter the input data. Moreover, the resulting tessellation of the sphere is *not spatially uniform* (in a geodesic sense), and this might affect the quality of the reconstructed 3D surface. By choosing "small" step-sizes in our experiments, we did not observe any appreciable quantitative degradation in performance. However, in future works, we plan to rigorously address this issue by taking advantage of the free tangent directions **v** and **w** introduced in Theorem 2, to define grids of arbitrary shape. An alternative option is to uniformly partition the sphere with one of the numerous methods available in the literature, for example considering a Quaternary Triangular Mesh (see (Guan and Smith, 2017) and

(Horn, 1984, Sect. V)), or the *geodesic grids* recently proposed in (Zhao et al., 2015; Adarve and Mahony, 2017).

On "global" methods

Note that (9) is a pixel-wise linear system in $(p, q) = (\nabla_{\mathbf{w}} \ln \rho, \nabla_{\mathbf{v}} \ln \rho)$, or equivalently, a linear system of partial differential equations in $\ln \rho$. Instead of adopting a two-step approach which consists in first solving in terms of (p, q) and then integrating the normals, one could discretize the gradient operator directly with finite differences, and turn (9) into a single large (but sparse) linear least-squares system in $\ln \rho$, thus circumventing the problem of integration. This "global" approach has been considered in a series of recent papers (see e.g. (Logothetis et al., 2016; Smith and Fang, 2016)), and it makes a legitimate direction for future research. In fact, directly solving for the radial distance is expected to limit the propagation of artifacts throughout the 3D reconstruction pipeline.

5. Conclusions and future work

The reduced field of view of perspective cameras calls for vision sensors with enhanced imaging capabilities for the 3D reconstruction of complex scenes. In this paper, we have presented a *unifying theory* of photometric stereo, which encompasses the most popular camera models in computer vision (notably perspective and central catadioptric systems), and complements the existing literature. In particular, closed-form solutions based on a new generic spherical image irradiance equation (Theorem 2), have been proposed and experimentally tested over a large database of synthetic and real-world panoramic images.

This work opens up several interesting avenues for future research. Since the focus of this article was on camera modeling, a classic calibrated photometric stereo setting in a controlled environment was considered. For further studies, we plan to relax our assumptions on the nature of the light source, and to address the case of general illumination conditions in an outdoor scenario. More sophisticated normal integration methods which can handle non-rectangular reconstruction domains, will be

evaluated as well (Quéau et al., 2018a). The concurrent estimation of the gradient field, albedo and illumination vectors without mirror balls inside the scene, and the extension of our results to multi-color surfaces (Quéau et al., 2016), and to glossy and hybrid surfaces (i.e. surfaces which are neither purely Lambertian nor purely specular (Weinmann et al., 2013)) considering more general BRDF, are other subjects of future research. Finally, the case of twin-fisheye cameras has been recently addressed in (Caracotte et al., 2021).

Supplementary Material

To foster reproducible research, the image datasets and the code developed to obtain the results reported in Sect. 4, are publicly available on the Internet at the address: mis.u-picardie.fr/~fabio/ustereo.html

Acknowledgments

References

Abrams, A., Hawley, C., Pless, R., 2012. Heliometric Stereo: Shape from Sun Position, in: Proc. Eur. Conf. Comp. Vis., pp. 357–370.

Abrams, A., Miskell, K., Pless, R., 2013. The Episolar Constraint: Monocular Shape from Shadow Correspondence, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 1407–1414.

Ackermann, J., Goesele, M., 2015. A Survey of Photometric Stereo Techniques. Found. Trends Comput. Graphics and Vision 9, 149–254.

Ackermann, J., Langguth, F., Fuhrmann, S., Goesele, M., 2012. Photometric Stereo for Outdoor Webcams, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 262–269.

Adarve, J., Mahony, R., 2017. Spherepix: a Data Structure for Spherical Image Processing. IEEE Robot. Autonom. Lett. 2, 483–490.

Ait-Aider, O., Berry, F., 2019. A flexible calibration method for the intrinsic and mechanical parameters of panoramic line-scan cameras. Comput. Vis. Image Und. 180, 47–58.

Alldrin, N., Zickler, T., Kriegman, D., 2008. Photometric stereo with non-parametric and spatially-varying reflectance, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn.

Barreto, J., 2006. A unifying geometric representation for central projection systems. Comput. Vis. Image Und. 103, 208–217.

Barsky, S., Petrou, M., 2003. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. IEEE Trans. Pattern Anal. 25, 1239–1252.

Basri, R., Jacobs, D., Kemelmacher, I., 2007. Photometric Stereo with General, Unknown Lighting. Int. J. Comput. Vision 72, 239–257.

Benosman, R., Manière, T., Devars, J., 1996. Multidirectional Stereovision Sensor, Calibration and Scenes Reconstruction, in: Proc. IEEE Int. Conf. Pattern Recogn., pp. 161–165.

Blender Foundation, 2018. Blender 2.79b: 3D Computer Graphics Software. www.blender.org.

Caracotte, J., Morbidi, F., Mouaddib, E., 2021. Photometric Stereo with Twin-Fisheye Cameras, in: Proc. IEEE Int. Conf. Pattern Recogn. To appear.

Caron, G., Eynard, D., 2011. Multiple Camera Types Simultaneous Stereo Calibration, in: Proc. IEEE Int. Conf. Robot. Automat., pp. 2933–2938.

Coxeter, H., 1969. Introduction to Geometry. 2nd ed., John Wiley & Sons.

Diaz, M., Sturm, P., 2013. Estimating Photometric Properties from Image Collections. J. Math. Imaging Vis. 47, 93–107.

Frankot, R., Chellappa, R., 1988. A Method for Enforcing Integrability in Shape from Shading Algorithms. IEEE Trans. Pattern Anal. 10, 439–451.

Furukawa, Y., Hernández, C., 2015. Multi-View Stereo: A Tutorial. Found. Trends Comput. Graphics and Vision 9, 1–148.

Geyer, C., Daniilidis, K., 2000. A Unifying Theory for Central Panoramic Systems and Practical Implications, in: Proc. Eur. Conf. Comp. Vis., pp. 445–461.

Geyer, C., Daniilidis, K., 2001. Catadioptric Projective Geometry. Int. J. Comput. Vision 45, 223–243.

Goldman, D., Curless, B., Hertzmann, A., Seitz, S., 2010. Shape and Spatially-Varying BRDFs from Photometric Stereo. IEEE Trans. Pattern Anal. 32, 1060–1071.

Grossberg, M., Nayar, S., 2004. Modeling the space of camera response functions. IEEE Trans. Pattern Anal. 26, 1272–1282.

Guan, H., Smith, W., 2017. BRISKS: Binary Features for Spherical Images on a Geodesic Grid, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 4516–4524.

Harker, M., O'Leary, P., 2008. Least Squares Surface Reconstruction from Measured Gradient Fields, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 1–7.

Harker, M., O'Leary, P., 2015. Regularized Reconstruction of a Surface from its Measured Gradient Field: Algorithms for Spectral, Tikhonov, Constrained, and Weighted Regularization. J. Math. Imaging Vis. 51, 46–70.

Hayakawa, H., 1994. Photometric stereo under a light source with arbitrary motion. J. Opt. Soc. Am. A 11, 3079–3089.

Hernández, C., Vogiatzis, G., Cipolla, R., 2008. Multiview Photometric Stereo. IEEE Trans. Pattern Anal. 30, 548–554.

Hernández, C., Vogiatzis, G., Cipolla, R., 2010. Overcoming Shadows in 3-Source Photometric Stereo. IEEE Trans. Pattern Anal. 33, 419–426.

Higo, T., Matsushita, Y., Ikeuchi, K., 2010. Consensus photometric stereo, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 1157–1164.

Higo, T., Matsushita, Y., Joshi, N., Ikeuchi, K., 2009. A Hand-held Photometric Stereo Camera for 3-D Modeling, in: Proc. IEEE Int. Conf. Comp. Vis., pp. 1234–1241.

Horn, B., 1984. Extended Gaussian Images. Proc. of the IEEE 72, 1671–1686.

Horovitz, I., Kiryati, N., 2004. Depth from gradient fields and control points: bias correction in photometric stereo. Image Vision Comput. 22, 681–694.

Ikehata, S., Aizawa, K., 2014. Photometric Stereo using Constrained Bivariate Regression for General Isotropic Surfaces, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 2187–2194.

Ikehata, S., Wipf, D., Matsushita, Y., Aizawa, K., 2012. Robust photometric stereo using sparse regression, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 318–325.

Jung, J., Lee, J.Y., Kweon, I., 2015. One-day outdoor photometric stereo via skylight estimation, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 4521–4529.

Kannala, J., Brandt, S., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE Trans. Pattern Anal. 28, 1335–1340.

Kondo, K., Mukaigawa, Y., Yagi, Y., 2008. Free-Form Mirror Design Inspired by Photometric Stereo, in: Proc. 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras.

Liao, M., Wang, L., Gong, M., 2007. Light Fall-off Stereo, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn.

Logothetis, F., Mecca, R., Cipolla, R., 2017. Semi-calibrated near field photometric stereo, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 941–950.

Logothetis, F., Mecca, R., Cipolla, R., 2019. A Differential Volumetric Approach to Multi-View Photometric Stereo, in: Proc. IEEE Int. Conf. Comp. Vis., pp. 1052–1061.

Logothetis, F., Mecca, R., Quéau, Y., Cipolla, R., 2016. Near-Field Photometric Stereo in Ambient Light, in: Proc. 27th British Machine Vision Conf., pp. 1–12.

Lu, F., Matsushita, Y., Sato, I., Okabe, T., Sato, Y., 2015. From Intensity Profile to Surface Normal: Photometric Stereo for Unknown Light Sources and Isotropic Reflectances. IEEE Trans. Pattern Anal. 37, 1999–2012.

Ma, Y., Soatto, S., Košecká, J., Sastry, S.S., 2004. An Invitation to 3D Computer Vision: From Images to Geometric Models. Springer.

Mariottini, G., Prattichizzo, D., 2008. Image-based Visual Servoing with Central Catadioptric Cameras. Int. J. Robot. Res. 27, 41–56.

Matsuki, H., von Stumberg, L., Usenko, V., Stückler, J., Cremers, D., 2018. Omnidirectional DSO: Direct Sparse Odometry With Fisheye Cameras. IEEE Robot. Autonom. Lett. 3, 3693–3700.

Mecca, R., Wetzler, A., Brustein, A., Kimmel, R., 2014. Near Field Photomet-

ric Stereo with Point Light Sources. SIAM J. Imag. Sci. 7, 2732–2770.

Papadhimitri, T., Favaro, P., 2013. A New Perspective on Uncalibrated Photometric Stereo, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 1474–1481.

Pottmann, H., Wallner, J., 2010. Computational Line Geometry. Springer.

Prados, E., Faugeras, O., 2005. Shape From Shading: a well-posed problem?, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 870–877.

Quéau, Y., Durou, J.D., Aujol, J.F., 2018a. Normal Integration: A Survey. J. Math. Imaging Vis. 60, 576–593.

Quéau, Y., Durou, J.D., Aujol, J.F., 2018b. Variational Methods for Normal Integration. J. Math. Imaging Vis. 60, 609–632.

Quéau, Y., Lauze, F., Durou, J.D., 2015. Solving Uncalibrated Photometric Stereo Using Total Variation. J. Math. Imaging Vis. 52, 87–107.

Quéau, Y., Mecca, R., Durou, J.D., 2016. Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 4359–4368.

Quéau, Y., Wu, T., Lauze, F., Durou, J.D., Cremers, D., 2017. A Non-Convex Variational Approach to Photometric Stereo under Inaccurate Lighting, in: Proc. IEEE Conf. Comp. Vis. Pattern Recogn., pp. 350–359.

Schnieders, D., Wong, K.Y., 2013. Camera and light calibration from reflections on a sphere. Comput. Vis. Image Und. 117, 1536–1547.

Shi, B., Mo, Z., Wu, Z., Duan, D., Yeung, S.K., Tan, P., 2019. A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo. IEEE Trans. Pattern Anal. 41, 271–284.

Shi, B., Tan, P., Matsushita, Y., Ikeuchi, K., 2012. Elevation Angle from Reflectance Monotonicity: Photometric Stereo for General Isotropic Reflectances, in: Proc. Eur. Conf. Comp. Vis., pp. 455–468.

Shi, B., Tan, P., Matsushita, Y., Ikeuchi, K., 2014. Bi-Polynomial Modeling of Low-Frequency Reflectances. IEEE Trans. Pattern Anal. 36, 1078–1091.

Smith, W., Fang, F., 2016. Height from photometric ratio with model-based light source selection. Comput. Vis. Image Und. 145, 128–138.

Tagare, H., Defigueiredo, R., 1991. A Theory of Photometric Stereo for a Class of Diffuse Non-Lambertian Surfaces. IEEE Trans. Pattern Anal. 13, 133–152.

Tankus, A., Kiryati, N., 2005. Photometric Stereo under Perspective Projection, in: Proc. IEEE Int. Conf. Comp. Vis., pp. 611–616.

Tankus, A., Sochen, N., Yeshurun, Y., 2005. Shape-from-Shading Under Perspective Projection. Int. J. Comput. Vision 63, 21–43.

Usenko, V., Demmel, N., Cremers, D., 2018. The Double Sphere Camera Model, in: Proc. Int. Conf. 3D Vision, pp. 552–560.

Vlasic, D., Peers, P., Baran, I., Debevec, P., Popović, J., Rusinkiewicz, S., Matusik, W., 2009. Dynamic Shape Capture using Multi-View Photometric Stereo. ACM Trans. Graphic 28. Article n. 174.

Weber, H., Arfken, G., 2003. Essential Mathematical Methods for Physicists. 6th ed., Academic Press.

Weinmann, M., Osep, A., Ruiters, R., Klein, R., 2013. Multi-View Normal Field Integration for 3D Reconstruction of Mirroring Objects, in: Proc. IEEE Int. Conf. Comp. Vis., pp. 2504–2511.

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

Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y., Ma, Y., 2010. Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery, in: Proc. Asian Conf. Comp. Vis., pp. 703–717.

Ying, X., Hu, Z., 2004. Can We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model?, in: Proc. Eur. Conf. Comp. Vis., pp. 442–455.

Zhang, Z., Rebecq, H., Forster, C., Scaramuzza, D., 2016. Benefit of Large Field-of-View Cameras for Visual Odometry, in: Proc. IEEE Int. Conf. Robot. Automat., pp. 801–808.

Zhao, Q., Feng, W., Wan, L., Zhang, J., 2015. SPHORB: A Fast and Robust Binary Feature on the Sphere. Int. J. Comput. Vis. 113, 143–159.