



3D-spline reconstruction using shape from shading: Spline from shading

F. Courteille *, A. Cruzil, J.-D. Durou, P. Gurdjos

IRIT – Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 9, France

Received 9 March 2006; received in revised form 14 December 2006; accepted 13 February 2007

Abstract

In this work, we describe an original method of solving the shape from shading (SFS) problem, which relies on the marriage of two simple ideas. On the one hand, we propose to modelize the scene by a parametric surface, namely a 3D-spline. The key advantage is that boundary conditions are no longer required to render the problem well-posed. On the other hand, we introduce the concept of “useful domain”: this is a sufficient set of pixels, whose greylevels are in accordance with the SFS hypotheses, which allows to solve the SFS problem on the whole reconstruction domain. The proposed method is described for both formulations of SFS, corresponding to orthographic and perspective projections. We can thus validate our method on synthetic as well as on real images.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Shape from shading; 3D-spline

1. Introduction

The shape from shading (SFS) problem consists in recovering the shape of a scene from a single greylevel image. Probably because of rather disappointing results on real images [1], this technique has been considered for a long time as a “textbook case”. This is owing to five recurrent problems.

1.1. *Pb1 – lack of realism of the modeling*

During a long time, SFS has suffered from the lack of realism of its modeling. However, several recent works [2–4] permit the emergence of more realistic modelings, in particular regarding the imaging geometry: encouraging results have been obtained on real images, for face reconstruction [5], organ reconstruction [6] or curved document flattening [4].

1.2. *Pb2 – necessity of boundary conditions*

All SFS modelings are devised in terms of partial differential equations of the first order, the unknown of the

problem being the depth function, which describes the height of each point in the scene. Classical methods of resolution of such equations have been used over the years: characteristic strips method [7]; level sets method [8]; search for viscosity solutions, through finite differences schemes [9,2] or semi-Lagrangian schemes [10], or through the use of the fast marching algorithm [6]; resolution of an equivalent optimal control problem [11,12]; optimization methods [13]. The important point to remind is that, even if the unknown is supposed to be fully regular, the problem is not well-posed in the absence of boundary conditions [14]. In order to render the problem well-posed, it is usual to add some *a priori* knowledge on the depth function (Dirichlet’s condition) or on its gradient (Neumann’s condition). Unfortunately, such a knowledge is generally not available on real images, causing problems of technical practicability, even if, in a recent work [5], a less arbitrary boundary condition (“state constraint”) has been successfully applied to real images.

1.3. *Pb3 – very large number of unknowns*

Each method of resolution from the previous (non-exhaustive) list can be transformed into a discrete problem that includes one or several variables at each pixel, depend-

* Corresponding author. Tel.: +33 61556882.

E-mail address: courteille@irit.fr (F. Courteille).

ing on whether the unknown is the depth function, its gradient, or even both simultaneously [15]. Hence, even for images of small sizes, one has to cope with a very large number of discrete unknowns. This entails that we have to expect the algorithms to run slowly, although some attempts have been done to accelerate the algorithms [16] or to reduce the number of unknowns, by replacing the pixel grid with a network [17] or by developing multi-resolution strategies [18,19].

1.4. Pb4 – presence of unreliable pixels

In a real image, some pixels have greylevels that contradict the SFS hypotheses, which can be caused by different phenomena [20]: if the scene is not purely Lambertian, some pixels are brighter or darker than they should; if the surface includes textured or painted patterns, the albedo is not uniform; the scene can include shadows; if the scene is not convex, some pixels are brighter than they should, due to secondary reflections; in the case of occlusions, pixels can be wrongly considered as pixels of the occluded object. Moreover, the inevitable problem of noise in real images has long been taken into account, particularly as regards the SFS methods based upon optimization [13].

1.5. Pb5 – ambiguities

The last, but not least, recurrent problem of SFS, concerns possible ambiguities on the solution. A first category of such ambiguities gathers the classical “concave/convex ambiguity”, the “non-visible deformations” [14] and the “bas-relief ambiguity” [21], which all make the assumption of orthographic projection. It has been shown recently in [5] that dealing with another modeling, which was originally introduced in [22], strongly reduces the number of ambiguous situations. Apart from these *exact* ambiguities, “quasi-ambiguities” often occur, when several surfaces, whose shapes may be completely different, produce very similar images, under the same illumination. Regarding optimization, this second kind of ambiguities is concerned with local minima. Recently, a stochastic optimization method has dispelled these quasi-ambiguities, but exclusively on synthetic images [19].

1.6. Motivation and contribution

Our motivation is to find a new method of resolution of SFS that works on real images and simultaneously overcomes most of the different problems previously listed. We observed the fact that, quite often, existing methods deal with a general surface model, while, regarding the potential application on real images to which they may apply, the best results are obtained with respect to geometrically constrained 3D scenes. One interesting example is the recovery of the shape of a curved document (see [4] and the references herein). To be flattened, the obtained surface has to be smoothed by fitting the reconstructed

point cloud to some 3D developable surface. Our first idea is thus to modelize the scene surface by a 3D parametric model, namely a 3D-spline. One of the important advantages of such a modeling is that we can expect to have more equations than unknowns *i.e.*, to have a well-posed problem. This paves the way of our second idea, which is to be selective about the considered pixels, by only dealing with a sufficient set of these *i.e.*, with those whose greylevels are in accordance with the SFS hypotheses. We introduce the notion of “useful domain”, which allows to solve the SFS problem on the whole reconstruction domain without using the greylevels of all the pixels it contains. To sum up, the proposed method is the marriage of two simple ideas, that bypasses all of the different problems previously listed except Pb5.

1.7. Organization

The remainder of the paper is organized as follows: in Section 2, we present the two ideas which underlie our new method; in Section 3, we show how to put these ideas into practice; Section 4 validates the method on synthetic images, then on real images; finally, in Section 5, we conclude on some perspectives.

2. Two ideas for the resolution of SFS

2.1. Introducing a 3D parametric model

2.1.1. State of the art

The idea of using a 3D parametric model in order to approximate the scene surface is not new, even if few papers report it. The first paper related to this idea [23] uses a quadratic model with five parameters. Although their authors claim that several patches can be separately considered, the experiments only consider a single patch over the whole image, yielding results of average quality. In [24], the same model is estimated by solving a set of algebraic equations. This new formulation is elegant but sounds difficult to be generalized to higher orders. Recently, this model has been used on multiple patches [3], but much effort is then required to smooth the connections between patches. In order to avoid possible problems when connecting patches, finite elements methods are of the greatest interest. In [25], a regular triangular grid is used, as well as a linear model over each triangular element (C^0 continuity constraint). As a result, a fine grid is thus required: in the tests, all the pixels are considered as nodes. A major drawback is slowness, as a consequence of problem Pb3, which is partially avoided through linearization of the reflectance map. Beside this problem, the results are of rather good quality, and an extension to perspective images has even been done in [26]. The problem of slowness has been partially solved in [27], where the same model is implemented, but using a non-regular grid, which makes it possible to locally refine the grid, if necessary, through a multi-resolution process. This work reformulates SFS within the frame-

work of deformable models: the SFS equation is expressed as a constraint, all along the evolution. This last method, which provides rather good results, has been extended to perspective images in [28,29]. Albeit we consider this latter work clearly as that yielding the best results, the problem of slowness (e.g., 10,000 iterations of 2 s each, at the most precise level with 10,000 triangles) makes this method technically impracticable to some extent. Another kind of deformable models, namely, 2D-snakes, has been used in [30], but once again, a special effort has to be done afterwards, in order to connect the different snakes, which correspond to parallel sections of the shape. The results are not so good, to say at least.

A few papers have dealt with more complex models than linear or quadratic ones. In [31], a superquadrics is used, whose 10 parameters are estimated via genetic algorithms, but no convincing result is shown. On the other hand, the results reported in [32] are of good quality, even if the method is intrinsically dedicated to face reconstruction. Dealing with other types of scenes would require to reformulate the main part of the problem.

Finally, let us mention some papers [33,34] in which the use of a 3D (quadratic) model is not aimed at modeling the surface, but at enforcing some “coherence criteria” on the normals at each step of an iterative process.

All of existing works, as discussed below, aim at avoiding the necessity of boundary conditions, thus solving Pb2. The methods described in [3,26,27,29], respectively, in [3,23,24,31,32], more or less successfully grapple with Pb1, respectively, with Pb3. On the other hand, Pb4 was not addressed in these works and only [31] dealt with Pb5, to some extent.

2.1.2. Our contribution

Using a 3D parametric model for the scene representation in SFS clearly prevents from the necessity of boundary conditions, which are implicitly replaced by constraints on the scene surface: Pb2 is solved at once. Nevertheless, as this comes out again from the previous state of the art, the choice of the model has great consequences upon the other aforementioned problems. In this paper, we aim at simultaneously circumscribe problems Pb1, Pb2, Pb3 but also Pb4 (as explained in Section 2.2.2). We will consider a 3D-spline to model the surface shape which surprisingly enough, has never been done before. It will be shown that this model allows us to reach this aim, contrary to the previously cited works.

2.2. Selecting the reliable pixels

2.2.1. State of the art

Aiming at applying SFS methods not only to synthetic, but also to real images, one must take into account the fact that a number of pixels could be “unreliable” i.e., not in accordance with the SFS hypotheses (Pb4). The classical methods of resolution which consider one unknown per pixel (the height of the conjugate point in the scene) cannot

straightforwardly discard the unreliable pixels, except either by imposing smoothness on the height, or by making holes in the reconstruction domain and by using a *posteriori* interpolation. To cope with problem Pb4, interesting alternatives have been proposed in some papers, where the unreliable pixels, whose greylevels are yet biased, are kept in the reconstruction domain, at the expense of great efforts to make the methods more robust. We list now what we think to be the most significant contributions dealing with Pb4 in classical SFS. In [35], the Lambertian and specular reflection components are separated, using probabilistic tools, in order to consider more realistic photometric modelings than the purely Lambertian one. In [36], the assumption of uniform albedo is questioned: the scene is segmented into regions having different albedos. In [37], shadow areas, which could be confused with areas of lower albedos, are identified through the so-called “shading flow field” procedure. The simpler situation of “black shadows” is astutely taken into account in [10], where it is considered that the limit between light and shadow belongs to the shape to be reconstructed. An iterative algorithm is reported in [38], which takes secondary reflections into account. Finally, in [39,40], the standard regularization theory is extended, in order to take visual discontinuities such as occlusions into account.

2.2.2. Our contribution

It is then judicious to consider the parametric approach to SFS versus the classical approach, since the former *intrinsically* solves the problem of interpolation, and it also prevents from the use of a smoothness term, which usually moves the solution away from a minimizer of the intensity error term. In parallel with the concept of reconstruction domain Ω_r , we introduce that of “useful domain” Ω_u . We define the “maximal useful domain” $\overline{\Omega}_u$ as the set containing all the reliable pixels. To illustrate this notion, consider the following example. Let us imagine that the rectangles in Fig. 1-a represent unreliable areas in the image. We aim at reconstructing the surface on the reconstruction domain Ω_r represented in black in Fig. 1-b, which does not contain holes. The maximal useful domain $\overline{\Omega}_u$, which contains holes, is represented in black in Fig. 1-c. In order to reconstruct the scene on the whole, while decreasing the computing time, we choose as useful domain Ω_u only a subset of $\overline{\Omega}_u$, made up for example by the pixels represented in black in Fig. 1-d. Only the pixels of Ω_u will be taken into account to check the concordance of the 3D parametric model with the image. This simple idea makes it possible to give a satisfactory response to problem Pb4, while reducing the computing time, as this will be seen in the experiments.

3. Spline from shading

3.1. Two SFS modelings

We attach a 3D frame ($Cxyz$) to the camera, whose origin C is the optical center and such that axis C_z coincides

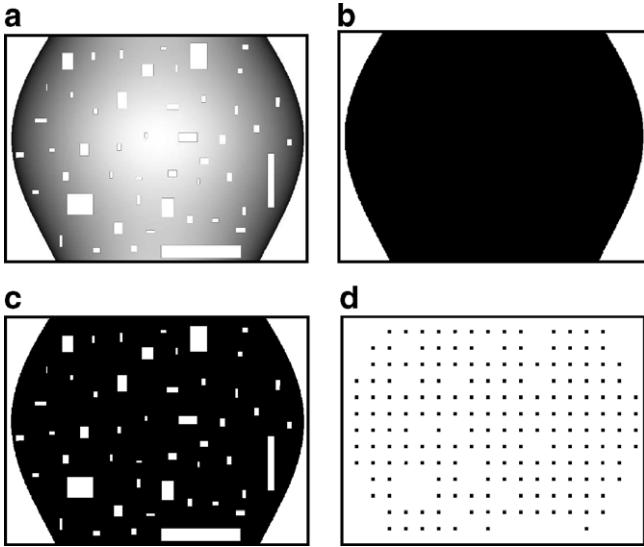


Fig. 1. (a) Greylevel image including unreliable pixels (white areas); (b) reconstruction domain Ω ; (c) maximal useful domain $\bar{\Omega}_u$; (d) example of non-maximal useful domain Ω_u .

with the optical axis (cf. Fig. 2). If f is the focal length, the sensor coincides with the plane Π' of equation $z = f$. This plane is conjugate with a plane Π of equation $z = -d$, called the “focusing plane”.

3.1.1. Weak perspective projection

Under the assumption of orthographic projection, also called “weak perspective projection”, a pixel $Q = (x, y, f)$ is conjugate with object point P^o . What we only have at our disposal is the function I such that $I(x, y)$ represents the greylevel at pixel Q . The coordinates of point P^o are $(-xd/f, -yd/f, u^o(x, y))$, where function u^o is the unknown of the problem. For the sake of simplicity, we will assume in the sequel that the light source is located at infinity and its direction is that of the optical axis. Under usual hypoth-

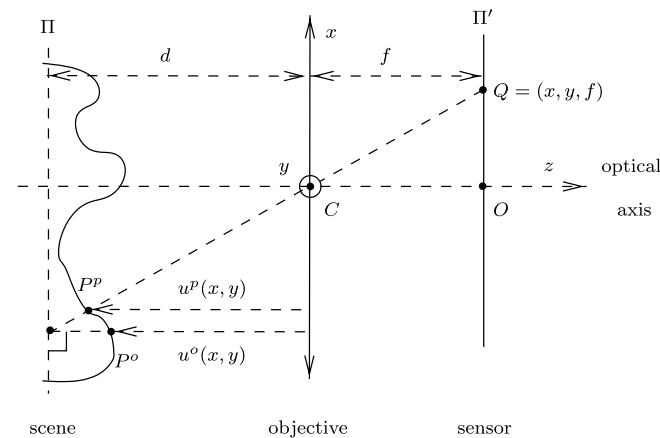


Fig. 2. Weak perspective projection (Q conjugate with P^o) and perspective projection (Q conjugate with P^p).

eses [4] the most popular formulation of SFS is the so-called *eikonal equation*:

$$g^2 \|\nabla u^o(x, y)\|^2 = \left[\frac{I_{\max}}{I(x, y)} \right]^2 - 1, \quad (1)$$

where $g = -f/d$ is the transverse magnification and I_{\max} the maximal value of I , reached at the “singular points”. It is a non-linear partial derivative equation of the first order. The case of arbitrary source location gives a more general equation than (1) which is of the same type [2].

3.1.2. Perspective projection

Shape from shading under perspective projection had already been outlined, first in the photogrammetric literature [41], and then in the computer vision literature [42,26,43], but none of these pioneering works had expressed the new modeling under its most general differential form. Recently, three groups of authors have independently established, at the same time, the SFS partial differential equation that takes perspective projection into account [2–4]. The object point conjugate with pixel Q is now P^p , whose coordinates are $(xu^p(x, y)/f, yu^p(x, y)/f, u^p(x, y))$, where function u^p is the new unknown of the problem. The perspective eikonal equation is [2–4]:

$$\hat{g}(x, y)^2 \|\nabla u^p(x, y)\|^2 = \left[\frac{I_{\max}}{I(x, y)} \right]^2 - 1, \quad (2)$$

where $\hat{g}(x, y) = f/(u^p(x, y) + x\partial_x u^p(x, y) + y\partial_y u^p(x, y))$. One can notice the strong similarity between (1) and (2).

3.2. 3D-splines

A 3D-spline of degrees (m, n) [44] is a polynomial function $S_{m,n}$ of two real variables $(x, y) \in [0, 1]^2$, of degrees m in x and n in y , defined in relation to two bases $\{B_{i,m}\}_{i \in [0, m]}$ and $\{B_{j,n}\}_{j \in [0, n]}$ of the vector spaces made of the polynomial functions of one real variable, of degrees m and n :

$$S_{m,n}(x, y) = \sum_{i=0}^m \sum_{j=0}^n P_{i,j} B_{i,m}(x) B_{j,n}(y), \quad (3)$$

where the real values $P_{i,j}$ are the $(m + 1) \times (n + 1)$ parameters of the model, called the “control parameters” of the 3D-spline. Consequently, it is easy to calculate analytically the partial derivatives of $S_{m,n}$:

$$\begin{cases} \frac{\partial S_{m,n}}{\partial x}(x, y) = \sum_{i=0}^m \sum_{j=0}^n P_{i,j} \frac{dB_{i,m}}{dx}(x) B_{j,n}(y), \\ \frac{\partial S_{m,n}}{\partial y}(x, y) = \sum_{i=0}^m \sum_{j=0}^n P_{i,j} B_{i,m}(x) \frac{dB_{j,n}}{dy}(y). \end{cases} \quad (4)$$

The knowledge of the control parameters makes it possible to define the 3D-spline without ambiguity, as well as its partial derivatives. This property is particularly interesting with respect to SFS, since Eqs. (1) and (2) are partial derivative equations. The control parameters become the new

unknowns of the problem, which we propose to reappoint “spline from shading”.

In the sequel, without loss of generality, we will assume that the pixel coordinates (x, y) are normalized so that they vary within $[0, 1]^2$, with respect to some square image.

3.3. Energy functions

We can thus formulate the spline from shading problem as an optimization problem with respect to the control parameters as unknowns.

3.3.1. Weak perspective projection

Under the assumption of weak perspective projection, we have seen that the relation between image and shape is given by the eikonal equation (1). Therefore, if we call $S_{m,n}^o$ the 3D-spline of degrees (m, n) which characterizes the shape of the scene, the spline from shading problem is that of minimizing the following energy function:

$$E^o(P_{i,j}^o) = \sum_{(x,y) \in \Omega_u} \left[I(x, y) - \frac{I_{\max}}{\sqrt{g^2 \|\nabla S_{m,n}^o(x, y)\|^2 + 1}} \right]^2, \quad (5)$$

which depends only on parameters $P_{i,j}^o$. Note that only the pixels in Ω_u are taken into account in the definition of E^o .

3.3.2. Perspective projection

We now focus on the perspective formulation of SFS. Let us note $S_{m,n}^p$ the 3D-spline of degrees (m, n) which characterizes the shape of the scene. According to Eq. (2), the problem is that of minimizing the following energy function:

$$E^p(P_{i,j}^p) = \sum_{(x,y) \in \Omega_u} \left[I(x, y) - \frac{I_{\max}}{\sqrt{\hat{g}(x, y)^2 \|\nabla S_{m,n}^p(x, y)\|^2 + 1}} \right]^2, \quad (6)$$

where $\hat{g}(x, y)$ can be expressed in terms of intrinsic parameters of the camera, $S_{m,n}^p$ and its partial derivatives:

$$\hat{g}(x, y) = \frac{f}{S_{m,n}^p(x, y) + x \frac{\partial S_{m,n}^p}{\partial x}(x, y) + y \frac{\partial S_{m,n}^p}{\partial y}(x, y)}. \quad (7)$$

Therefore, the energy function E^p depends on the focal length f , on the position of the principal point O (through the variables x and y , since these variables are related to a frame in the image having O as origin) and on the control parameters $P_{i,j}^p$ of the spline.

3.3.3. Discussion

Let us point out that the energy functions E^o and E^p have the same complexity with respect to their unknowns, which shows that our SFS method gives a satisfactory answer to problem Pb1, since its implementation is not more complicated for a more realistic modeling of SFS. The minimization of E^o or of E^p can be carried out with

usual methods of non-linear optimization. With regard to the chosen minimization algorithm (cf. Section 4), we remind the reader the following facts. On the one hand, a Levenberg–Marquardt iteration basically has complexity $\mathcal{O}(p^3)$ in the number p of the unknown parameters $P_{i,j}$ (which designate the parameters $P_{i,j}^o$ or $P_{i,j}^p$). On the other hand, by increasing by one each degree of the 3D-spline, we only change the number of parameters to estimate from $(m+1) \times (n+1)$ to $(m+2) \times (n+2)$.

Furthermore, there exists another main advantage to our formulation, compared to other methods of resolution based upon optimization. If $I\sim$ denotes an estimate of the greylevel, related to the computed shape \tilde{u} through Eqs. (1) or (2), then E (which designates E^o or E^p) is straightforwardly linked to the root mean square error on I over $\overline{\Omega}_u$, denoted $|\Delta I|_2$, since:

$$|\Delta I|_2 = \sqrt{\frac{1}{\text{card}\overline{\Omega}_u} \sum_{(x,y) \in \overline{\Omega}_u} [I(x, y) - \tilde{I}(x, y)]^2}. \quad (8)$$

Comparing (8) with (5) and (6), and given that \tilde{I} is computed through (1) or (2), it comes:

$$|\Delta I|_2 \approx \sqrt{\frac{E(P_{i,j})}{\text{card}\overline{\Omega}_u}}, \quad (9)$$

this approximation becoming an equality when $\Omega_u = \overline{\Omega}_u$ *i.e.*, when the mean is computed on the maximal useful domain $\overline{\Omega}_u$. Thus, minimizing E^o or E^p necessarily decreases the difference between I and $I\sim$. On the other hand, this is not the case for energy functions containing terms of regularization like “smoothness constraint” or “integrability constraint”, as for example in [13,15,19,16].

4. Experiments

Regarding the implementation of the proposed method, we used Bernstein’s polynomials [44] as functions $B_{i,m}$ and $B_{j,n}$ and the Minpack¹ function `lmdif` as optimization method (non-linear least squares), with iterations of the Levenberg–Marquardt type. The control parameters are initialized in such a way that the 3D-spline value is higher in the neighbourhood of each (manually detected) singular point than everywhere else. It seems that in practice the method converges in a few iterations.

In all the experiments, we compute $|\Delta u|_2$, which is the error estimator generally used (cf. *e.g.*, [1–3,29,34]), except when no ground truth is available, in which case $|\Delta I|_2$ is used:

$$|\Delta u|_2 = \sqrt{\frac{1}{\text{card}\overline{\Omega}_u} \sum_{(x,y) \in \overline{\Omega}_u} [u(x, y) - \tilde{u}(x, y)]^2}. \quad (10)$$

¹ <http://www.netlib.org/minpack>

4.1. Choice of the useful domain

Since our approach enables us to choose the useful domain, we have to evaluate the quality of the results using various useful domains.

4.1.1. Density of the selected pixels

Let us consider in Fig. 3-a1 the synthetic greylevel image of a vase under weak perspective projection, and in Fig. 3-b1 the corresponding surface. The various useful domains we tested are represented in black in Figs. 3-a2, 3-a3 and

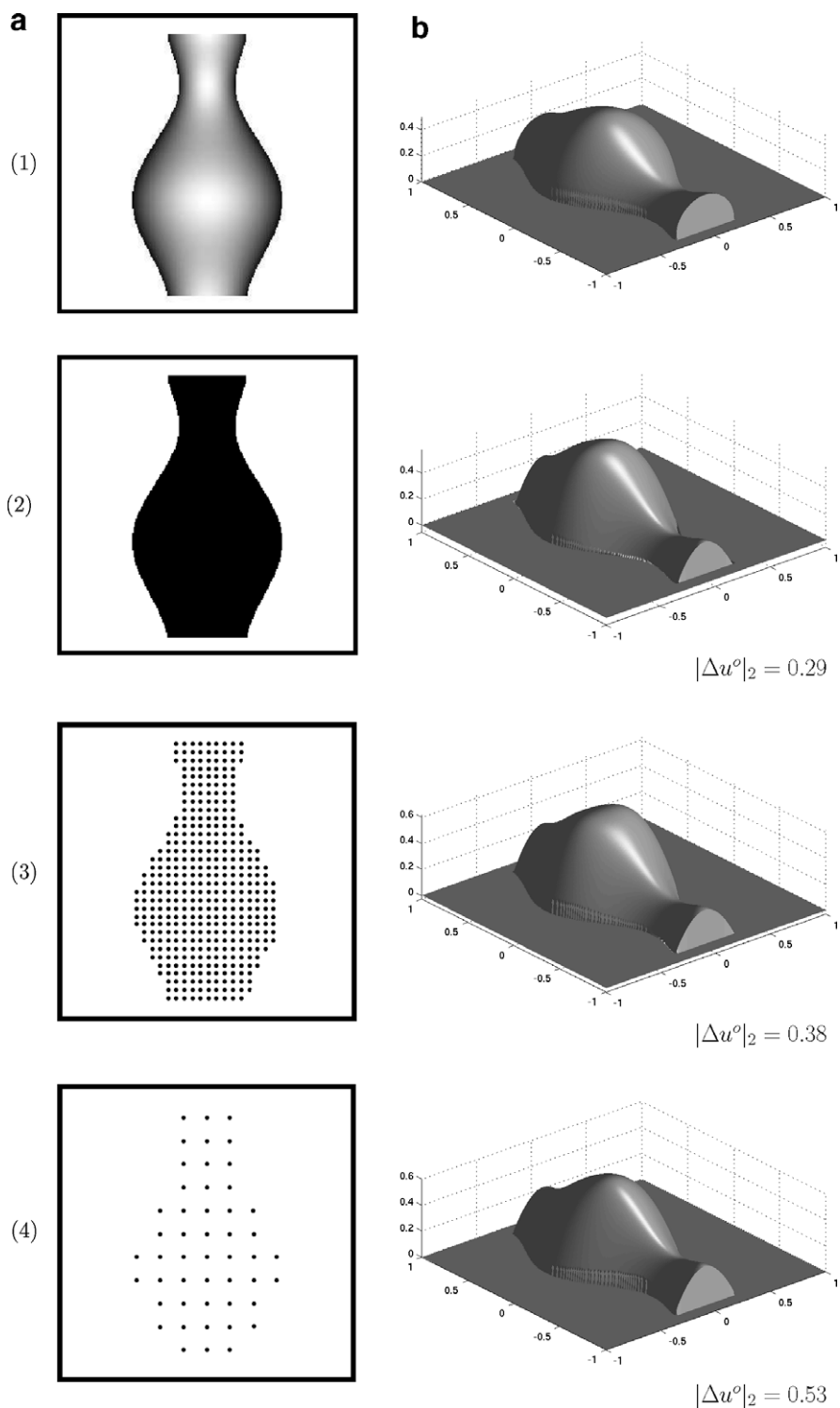


Fig. 3. Influence of the density of the selected pixels on the reconstructed surface: (a1) synthetic image of a vase; (b1) true surface; (a2–a4) three useful domains representing different densities of the selected pixels; (b2–b4) corresponding reconstructed surfaces and values of the error estimator $|\Delta u^o|_2$.

3-a4. The corresponding reconstructed surfaces are represented in Figs. 3-b2, 3-b3 and 3-b4. First, we can notice that for the first useful domain (Fig. 3-a2), which is in fact the maximal useful domain $\bar{\Omega}_u$, the quality of the reconstructed surface is quite good (cf. Fig. 3-b2). However, the computing time comes to 3 min on a P4, 2.4 GHz.

The useful domain represented in Fig. 3-a3, made up of 3% only of the maximal useful domain, allows a quasi-identical quality of reconstruction (cf. Fig. 3-b3) in less than 10 s. The useful domain of Fig. 3-a4, which represents only 1% of the maximal useful domain, allows to obtain a lower quality surface (cf. Fig. 3-b4), but in 1 s. It is note-

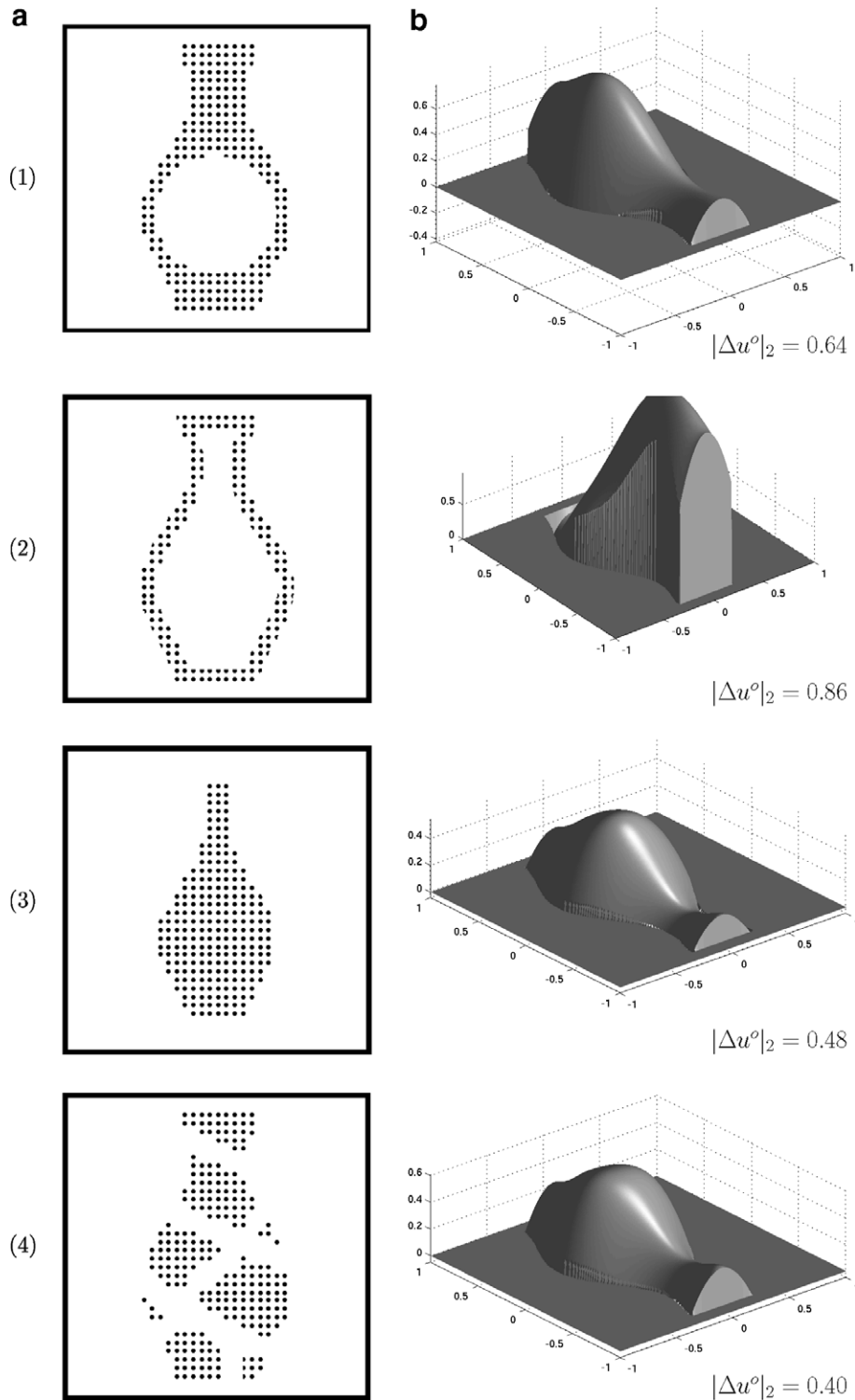


Fig. 4. Influence of the spatial distribution of the selected pixels on the reconstructed surface: (a1–a4) four useful domains representing different spatial distributions of the selected pixels; (b1–b4) corresponding reconstructed surfaces and values of the error estimator $|\Delta u^\circ|_2$.

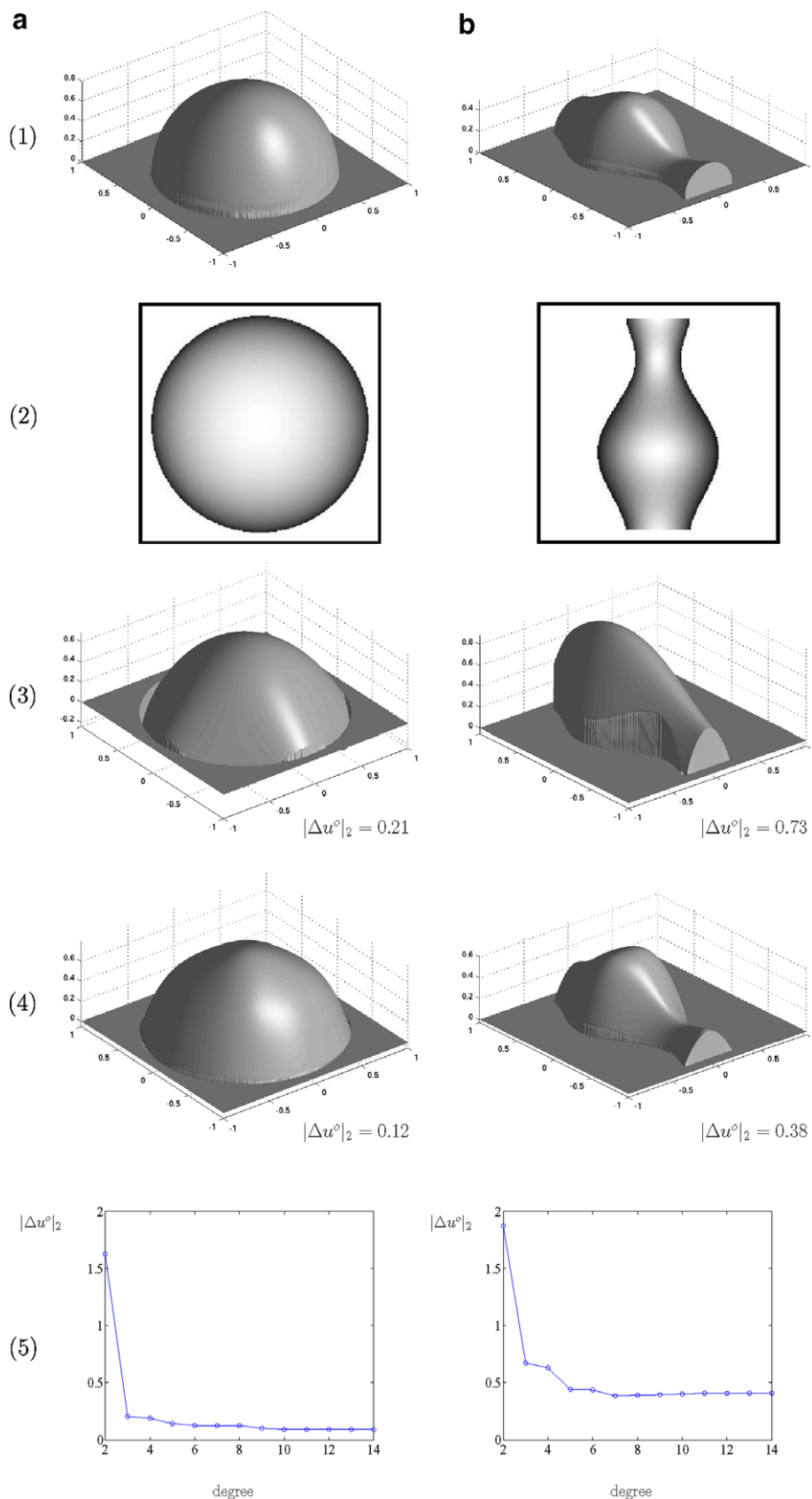


Fig. 5. Examples of surfaces reconstructed from synthetic images generated under weak perspective projection: (1) true surfaces; (2) corresponding synthetic images; (3) reconstructed surfaces using splines of degrees (3,3) and (4) using splines of degrees (9,9); (5) evolution of $|\Delta u^\circ|_2$ w.r.t. the degree n (meaning that the splines are of degrees (n,n)).

worthy that the “classical methods”, as for example those described in [13], could also easily discard the unreliable

pixels by setting the corresponding intensity error terms to zero: the unknowns at such pixels would then be

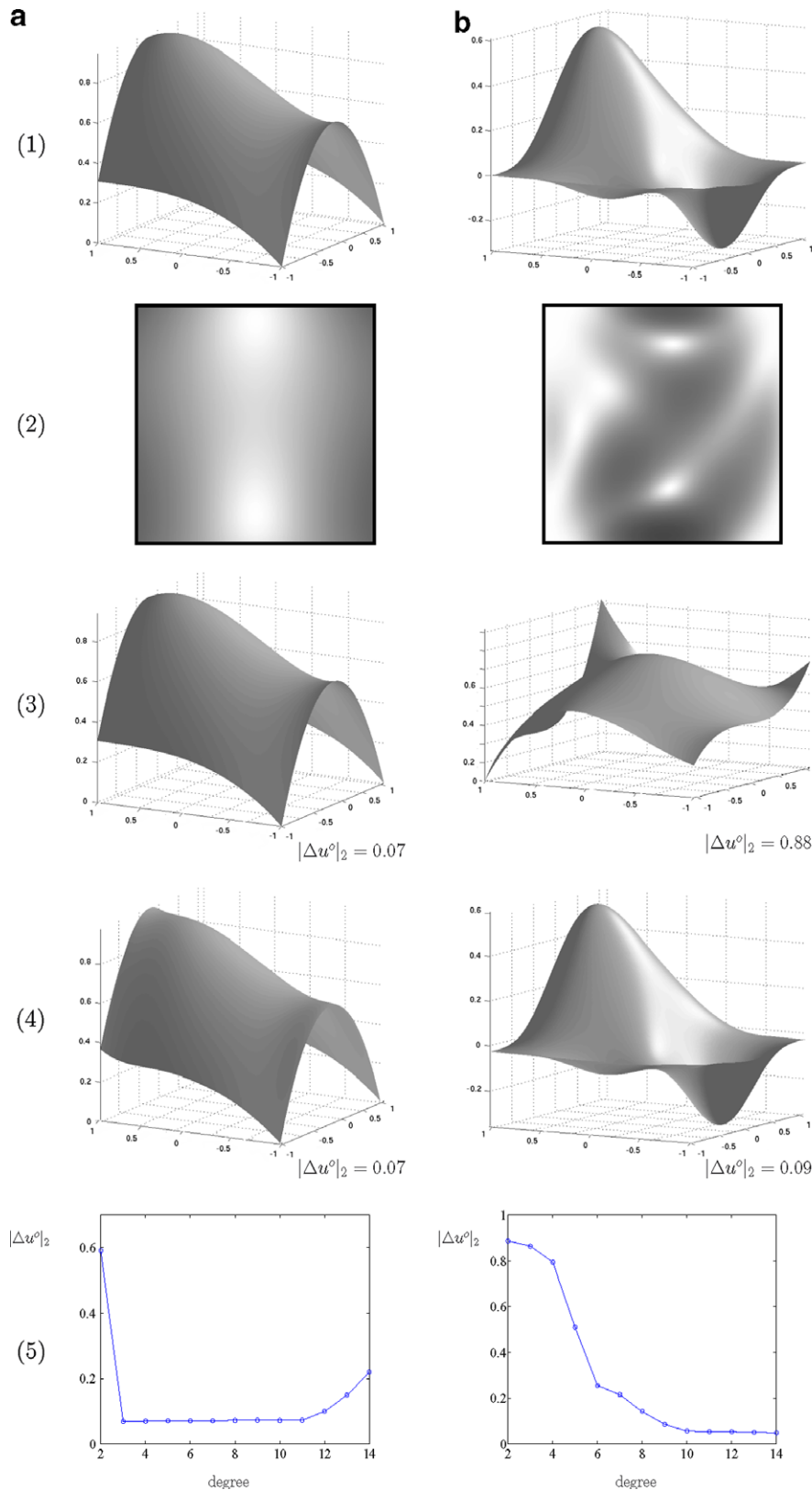


Fig. 6. Examples of surfaces reconstructed from synthetic images generated under weak perspective projection: (1) true surfaces; (2) corresponding synthetic images; (3) reconstructed surfaces using splines of degrees (3,3) and (4) using splines of degrees (9,9); (5) evolution of $|\Delta u^o|_2$ w.r.t. the degree n (meaning that the splines are of degrees (n,n)).

updated by a local average *i.e.*, by explicit interpolation, but this would not decrease the CPU time much, since all the pixels of the reconstruction domain would still be visited at each step. In our approach, the CPU time possibly much decreases because interpolation is implicit.

4.1.2. Spatial distribution of the selected pixels

Four different useful domains are represented in Figs. 4-a1, 4-a2, 4-a3 and 4-a4. The corresponding reconstructed surfaces are represented in Figs. 4-b1, 4-b2, 4-b3 and 4-b4. This allows us to evaluate how important the spatial distribution of the selected pixels in the image is. Indeed, if one chooses useful domains like 4-a1 or 4-a2, which correspond in the scene to raised slopes, the reconstructed surfaces 4-b1 or 4-b2 are of poor quality, whereas if the useful domain is made up of areas corresponding to weak slopes of the surface, like Fig. 4-a3, then the reconstructed surface Fig. 4-b3 is of much better quality. The last useful domain, represented in Fig. 4-a4, allows us to evaluate the behaviour of our method when a significant part of the reconstruction domain is unusable, for example if parts of the object to be reconstructed are occluded by other objects. We can obviously notice, observing the reconstructed surface Fig. 4-b4 as well as the value of the error estimator $|\Delta u^0|_2$, that the lack of reliable pixels does not prevent from a reconstruction of very good quality, provided that reliable information is sufficiently well spatially distributed. Indeed, the quality of the result represented in Fig. 4-b4 is very similar to that of Fig. 4-b2, obtained using the whole of the reliable pixels.

4.2. Synthetic images

According to the previous discussion, in the next experiments, we select an useful domain uniformly distributed in the maximal useful domain, made up of approximately 3% of the pixels.

4.2.1. Weak perspective projection

We now focus on the minimization of energy function E^p given in (5). We generated, under the hypothesis of weak perspective projection, four synthetic images of size 256×256 which are represented in Figs. 5-a2, 5-b2, 6-a2 and 6-b2. The corresponding surfaces are represented in Figs. 5-a1, 5-b1, 6-a1 and 6-b1, and are, respectively, a sphere, the same vase as before, a spline of degrees (3,3) and a spline of degrees (9,9). The remaining figures represent the results using our method: surfaces Figs. 5-a3, 5-b3, 6-a3 and 6-b3 are obtained using splines of degrees (3,3) *i.e.*, 16 unknowns; surfaces Figs. 5-a4, 5-b4, 6-a4 and 6-b4 are obtained using splines of degrees (9,9) *i.e.*, 100 unknowns. Referring to Figs. 5-b3 and 6-b3, we can notice that if the degrees of the spline are too low, then the quality of the reconstructed surface is not good, because a spline of degrees (3,3) does not always have sufficient degrees of freedom to estimate precisely the surface. On the other hand, as we can see in Figs. 5-b4 and 6-b4, a spline of degrees

(9,9) ensures a good quality of reconstruction. Furthermore, the surface of Fig. 6-a4, which is reconstructed using a spline of degrees (9,9), is of a slightly inferior quality to that of the surface of Fig. 6-a3, reconstructed using a spline of degrees (3,3).

The evolution of $|\Delta u^0|_2$ w.r.t. the degree n (meaning that the splines are of degrees (n,n)) is shown in Figs. 5-a5, 5-b5, 6-a5 and 6-b5. Empirically, it is generally better to choose n rather too large than too small. When the degrees are well chosen, we can notice, within sight of the results presented in Figs. 5-a4, 5-b4, 6-a4 and 6-b4, that the quality of the reconstructed surfaces is very satisfactory, particularly on the edges of the reconstruction domains, without knowledge on the boundary and without *a priori* information on the scene. In Fig. 6-a5, one can notice that $|\Delta u|_2$ increases for $n \geq 10$: this is due to an overfitting effect, knowing that the error is computed on the whole reconstruction domain, whereas the useful domain is made up of 3% only of the reconstruction domain.

4.2.2. Perspective projection

Let us now consider the perspective projection modeling. Energy function E^p given in (6), corresponding to the perspective eikonal equation, is the new energy to be minimized. Contrary to the first approach, the fact of taking perspective into account adds three parameters which are the focal length and the coordinates in the image of the

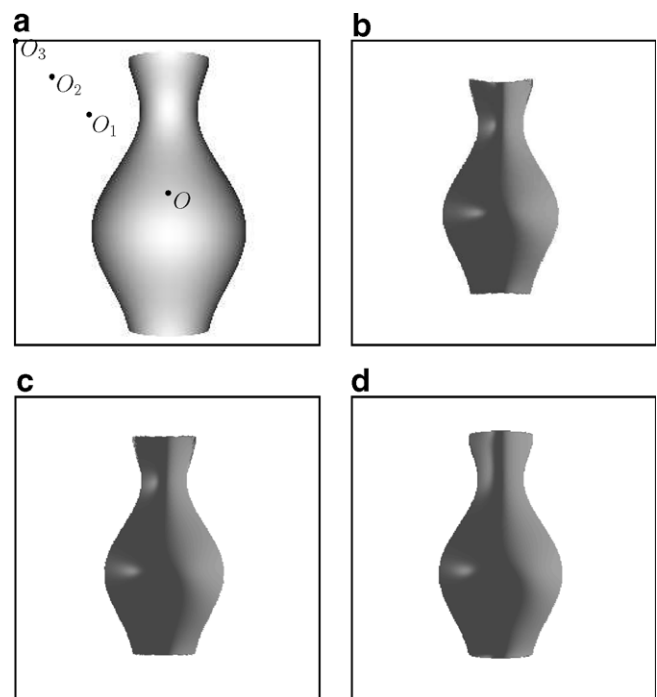


Fig. 7. Qualitative effect of an erroneous estimate of the focal length: (a) synthetic image of a vase under perspective projection with the true position O and three erroneous positions O_1 , O_2 and O_3 of the principal point; orthogonal projection views of the reconstructed surfaces, for three values of f : (b) relative error of -35% , (c) true value and (d) relative error of $+50\%$.

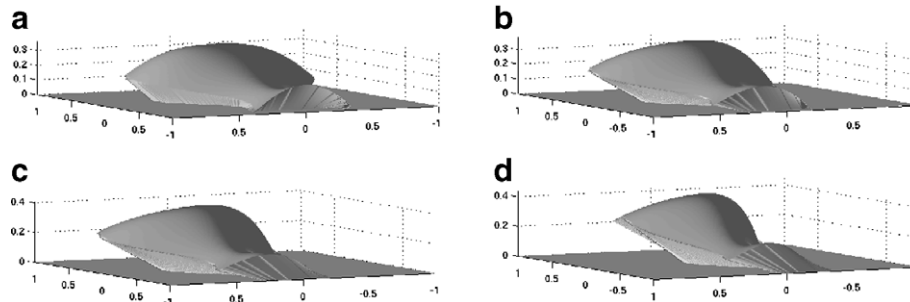


Fig. 8. Qualitative effect of an erroneous estimate of the principal point: reconstructed surfaces using (a) O , (b) O_1 , (c) O_2 and (d) O_3 as estimates of the principal point (cf. Fig. 7a).

principal point. Thus, it is interesting to evaluate the sensitivity of our method to possible errors on these parameters. For that, we are interested in the synthetic image of a vase under perspective projection, generated with a focal length $f = 1.95$ (remind that we deal with normalized pixel coordinates, varying within $[0, 1]$) and for which the principal point O is located at the center of the image. While simulating an error on each of these parameters, we will evaluate the robustness of our method.

The influence of an error on the estimate of f is illustrated in Fig. 7: (a) represents the synthetic image of the vase under perspective projection; (b) and (d), respectively, (c), represent the images (under orthogonal projection view) of the reconstructed surfaces for erroneous, respectively, true values of f . This vase being a symmetric surface of revolution, the top and the bottom of its silhouette should logically be, for an orthogonal projection view, segments of straight lines, as is the case in Fig. 5-b2. Only Fig. 7-c satisfies this property, since in Fig. 7-b, the top and the bottom of the vase are concave, and in Fig. 7-d, they are convex. With the aim of clearly identifying the effects of a “badly” estimated f , we deliberately chose high relative errors on f , equal to -35% and $+500\%$, respectively, in Figs. 7-b and 7-d.

Let us examine now the effect of a bad estimate of the principal point. In Fig. 7-a, O represents the true location and O_1 , O_2 , O_3 three erroneous locations of the principal point. Figs. 8-a, 8-b, 8-c and 8-d represent the reconstructed surfaces using O , O_1 , O_2 and O_3 as estimates of the principal point. Of course, the more the estimate of the principal point is erroneous, the more the reconstructed surface is dissymmetric.

These qualitative statements are quantitatively confirmed by the curves of Fig. 9, in which the error $|\Delta u^p|_2$ is plotted in function of: (a) a relative error on the focal length; (b) an error on the location of the principal point, divided by the image size (in fact, this last error is the mean of eight errors obtained in eight directions around the true location of the principal point). The useful domain is made up of 3% of the maximal useful domain. For both curves, the minimum corresponds to $|\Delta u^p|_2 = 0.44$. From a practical point of view, it is commonly admitted that succeeding calibration methods obtain a relative error on the focal length less than

10%, and an error on the location of the principal point less than 20% of the image size. Hence, clearly, we can conclude that the robustness of our method to errors on the internal parameters is quite sufficient.

4.3. Real images

We now assess the “realism” of the obtained reconstructed surfaces when processing real images. In order to limit the invalidation from the SFS hypotheses, the light source we use is an annular flash (circular flash fixed around the objective). The focal length and the coordinates of the principal point are estimated using Bouguet’s calibration toolbox.²

The first column of Fig. 10 represents the photographs of three real scenes. The second column represents the reconstruction domains and the maximal useful domains, which both have been manually designed. Some pixels of these images are unreliable, when the reconstruction domain is partially occluded by other objects composing the scene, as in Fig. 10-a2, or partially unusable, because the greylevels are not only due to shading, but also to a non-uniform albedo, as in Figs. 10-a1 and 10-a3. Note also that the specular spots, which are clearly visible in Figs. 10-a1 and 10-a2, have not been removed from the corresponding useful domains (cf. Figs. 10-b1 and 10-b2), since they do not affect the reconstructions much.

The surfaces reconstructed from images Figs. 10-a1 and 10-a2 using our method are represented in Figs. 11-a1 and 11-a2. We can notice that the surface of the mouse is correctly reconstructed, in particular on the boundary of the reconstruction domain, without any *a priori* knowledge on the scene. The last image (cf. Fig. 10-a3) represents a warped paper sheet, on which a regular grid has been printed. The associated maximal useful domain (cf. Fig. 10-b3) is not connex. Thus, without use of boundary conditions, our approach calculates a shape which seems quite satisfactory at least, visually speaking (cf. Fig. 11-a3).

Reconstructions from the same images using Tsai and Shah’s method [45], denoted TS, are represented in Figs.

² <http://www.vision.caltech.edu/bouguetj>

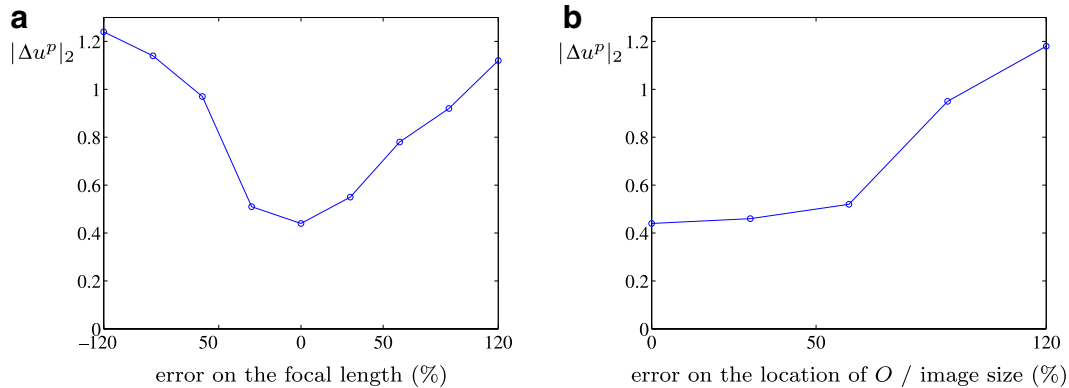


Fig. 9. Quantitative effect of an erroneous estimate of (a) the focal length; (b) the principal point.

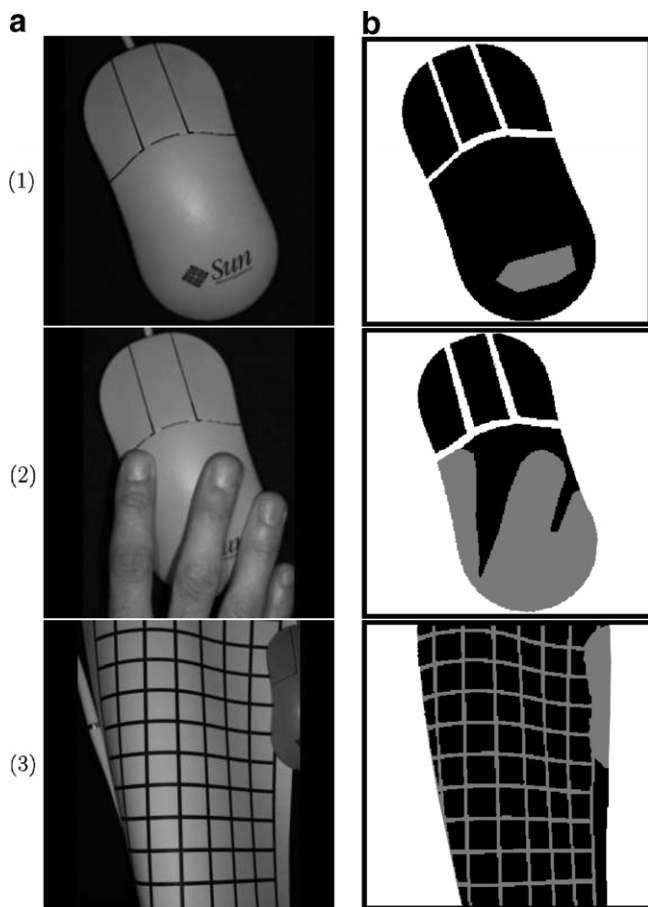


Fig. 10. (a) Three real images; (b) maximal useful domains in black, reconstruction domains in black or in grey.

11-b1, 11-b2 and 11-b3. The choice of TS by way of comparison was motivated by two factors: on the one hand, its code is freely available³; on the other hand, it is one of the few methods tested in [1] which do not require additional knowledge on the boundary. It seems logical to run TS on the maximal useful domains only, like all the classical SFS methods. There are therefore holes in the recon-

structed shapes, which could be *a posteriori* filled via interpolation, contrary to the shapes obtained using our method, which intrinsically solves the problem of interpolation. Apart from these holes, the reconstructed surfaces shown in Fig. 11-b display a wrinkled structure, an artifact typical of TS which was already noticed in a recent survey on SFS.⁴

The reconstructions using TS are qualitatively clearly worse than using our method, particularly on the boundaries. The use of the error estimator $|\Delta^P|_2$ shows that this qualitative statement is quantitatively confirmed: $|\Delta^P|_2$ is about 10 times greater using TS than our method. Finally, notice that the shapes Figs. 11-a1 and 11-a2 were reconstructed using splines of degrees (3,3), whereas the shape Figs. 11-a3 was reconstructed using a spline of degrees (9,9). Concerning the computing times, our method takes about 1 s in the case of degrees (3,3) and less than 5 s in the case of degrees (9,9), for images of 256×256 pixels, among which 3% are taken into account. For each of these three images, TS computing time is about 1 s only.

5. Conclusion and perspectives

We propose in this work a new SFS method, which combines the use of a 3D-spline to model the scene surface and the concept of useful domain. The tests carried out on synthetic as well as on real images make it possible to validate this new method.

The proposed method suffers nevertheless from two weaknesses. On the one hand, the choice of the degrees (m,n) of the spline, which is significant, is a non-trivial problem and is still under investigation. On the other hand, the useful domain has been manually designed in all our experiments. It would be worthy of interest to consider the unreliable pixels as well, but to replace the classical least-square optimization with robust estimation, since unreliable pixels are nothing else than outliers.

³ <http://www.cs.ucf.edu/~vision/shadsrc.html>

⁴ <http://www.irrit.fr/sfs>

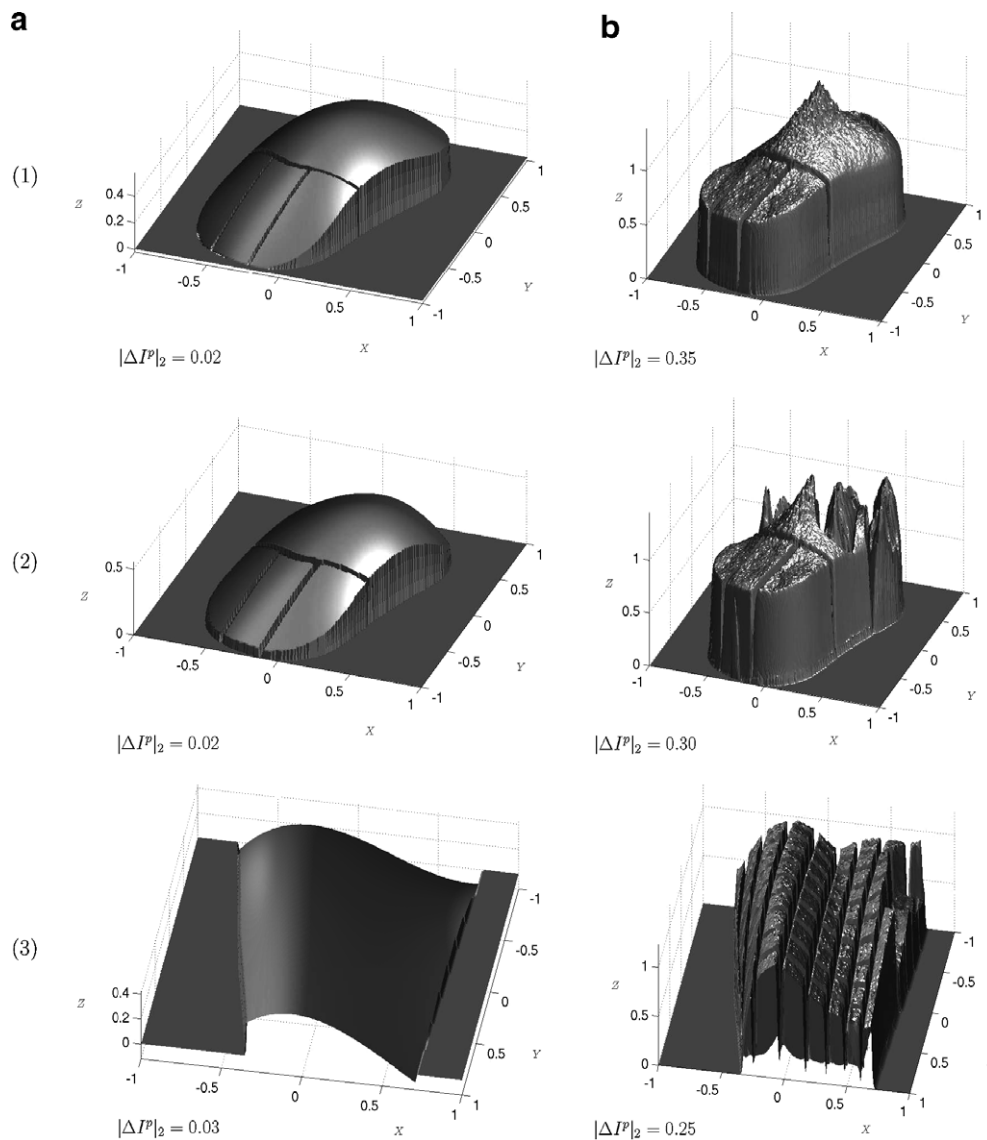


Fig. 11. Reconstructions from the images of Fig. 10 using (a) our method; (b) TS.

Many prospects can be considered. The use of B-splines, instead of Bézier-splines, would allow us to consider more complex scenes. The use of a non-deterministic optimization method, like simulated annealing, could allow to partially solve problem Pb5, avoiding local minima.

The restriction to parallel illumination and viewing directions is a severe one. In many practical applications such as the generation of digital elevation maps, this condition is even very unfavourable. Moreover, it could also appear that our framework cannot be applied to non-Lambertian surfaces. Hence, it would be interesting to generalize the method to oblique illumination or non-Lambertian materials.

References

- [1] R. Zhang, P.-S. Tsai, J.E. Cryer, M. Shah, Shape from shading: a survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (8) (1999) 690–706.
- [2] E. Prados, O. Faugeras, Perspective shape from shading and viscosity solutions, in: *Proceedings of the 9th IEEE International Conference on Computer Vision*, vol. II, Nice, France, 2003, pp. 826–831.
- [3] A. Tankus, N. Sochen, Y. Yeshurun, A new perspective [on] shape-from-shading, in: *Proceedings of the 9th IEEE International Conference on Computer Vision*, vol. II, Nice, France, 2003, pp. 862–869.
- [4] F. Courteille, A. Crouzil, J.-D. Durou, P. Gurdjos, Towards shape from shading under realistic photographic conditions, in: *Proceedings of the 17th International Conference on Pattern Recognition*, vol. II, Cambridge, UK, 2004, pp. 277–280.
- [5] E. Prados, O. Faugeras, Shape from shading: a well-posed problem?, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. II, San Diego, CA, USA, 2005, pp. 870–877.
- [6] A. Tankus, N. Sochen, Y. Yeshurun, Perspective shape-from-shading by fast marching, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. I, Washington, DC, USA, 2004, pp. 43–49.
- [7] B.K.P. Horn, Obtaining shape from shading information, in: P.H. Winston (Ed.), *The Psychology of Computer Vision*, McGraw-Hill, NY, 1975, pp. 115–155 (Chapter 4).

- [8] R. Kimmel, A.M. Bruckstein, Tracking level sets by level sets: a method for solving the shape from shading problem, *Computer Vision and Image Understanding* 62 (1) (1995) 47–58.
- [9] P.-L. Lions, E. Rouy, A. Tourin, Shape-from-shading, viscosity solutions and edges, *Numerische Mathematik* 64 (3) (1993) 323–353.
- [10] M. Falcone, M. Sagona, An algorithm for the global solution of the shape-from-shading model, in: *Proceedings of the 9th International Conference on Image Analysis and Processing*, vol. I. From: Lecture Notes in Computer Science, vol. 1310, Florence, Italy, 1997, pp. 596–603.
- [11] M. Bichsel, A.P. Pentland, A simple algorithm for shape from shading, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Champaign, IL, USA, 1992, pp. 459–465.
- [12] J. Oliensis, P. Dupuis, A global algorithm for shape from shading, in: *Proceedings of the 4th IEEE International Conference on Computer Vision*, Berlin, Germany, 1993, pp. 692–701.
- [13] B.K.P. Horn, M.J. Brooks, The variational approach to shape from shading, *Computer Vision, Graphics, and Image Processing* 33 (2) (1986) 174–208.
- [14] J.-D. Durou, D. Piau, Ambiguous shape from shading with critical points, *Journal of Mathematical Imaging and Vision* 12 (2) (2000) 99–108.
- [15] B.K.P. Horn, Height and gradient from shading, *International Journal of Computer Vision* 5 (1) (1990) 37–75.
- [16] R. Szeliski, Fast shape from shading, *Computer Vision, Graphics, and Image Processing: Image Understanding* 53 (2) (1991) 129–153.
- [17] M. Sagona, A. Seghini, An adaptive scheme on unstructured grids for the shape-from-shading problem, in: M. Falcone, C. Makridakis (Eds.), *Numerical Methods for Viscosity Solutions and Applications*, Series on Advances in Mathematics for Applied Sciences, vol. 59, World Scientific, Singapore, 2001, pp. 197–219 (Chapter 11).
- [18] D. Terzopoulos, Image analysis using multigrid relaxation methods, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (2) (1986) 129–139.
- [19] A. Crouzil, X. Descombes, J.-D. Durou, A multiresolution approach for shape from shading coupling deterministic and stochastic optimization, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (11) (2003) 1416–1421.
- [20] M. Oren, S.K. Nayar, Generalization of the Lambertian model and implications for machine vision, *International Journal of Computer Vision* 14 (3) (1995) 227–251.
- [21] P.N. Belhumeur, D.J. Kriegman, A.L. Yuille, The Bas-Relief Ambiguity, *International Journal of Computer Vision* 35 (1) (1999) 33–44.
- [22] T. Okatani, K. Deguchi, Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center, *Computer Vision and Image Understanding* 66 (2) (1997) 119–131.
- [23] T.C. Pong, R.M. Haralick, L.G. Shapiro, Shape from shading using the Facet model, *Pattern Recognition* 22 (6) (1989) 683–695.
- [24] P.K. Bora, Y.V. Venkatesh, K.R. Ramakrishnan, Shape from shading using discrete polynomials, in: *Proceedings of the XVI Annual Convention and Exhibition of the IEEE in India*, Bangalore, India, 1990, pp. 260–264.
- [25] K.M. Lee, C.-C.J. Kuo, Shape from shading with a linear triangular element surface model, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (8) (1993) 815–822.
- [26] K.M. Lee, C.-C.J. Kuo, Shape from shading with perspective projection, *Computer Vision, Graphics, and Image Processing: Image Understanding* 59 (2) (1994) 202–212.
- [27] D. Samaras, D.N. Metaxas, Incorporating illumination constraints in deformable models, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, USA, 1998, pp. 322–329.
- [28] D. Samaras, D.N. Metaxas, Coupled lighting direction and shape estimation from single images, in: *Proceedings of the 7th IEEE International Conference on Computer Vision*, vol. II, Kerkyra, Greece, 1999, pp. 868–874.
- [29] D. Samaras, D.N. Metaxas, Incorporating illumination constraints in deformable models for shape from shading and light direction estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (2) (2003) 247–264.
- [30] I. Bakhadyrov, M. Jafari, Snake-based deformable surface model for shape-from-shading, in: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Hammamet, Tunisia, 2002, pp. 218–222.
- [31] H. Saito, N. Tsunashima, Estimation of 3-D parametric models from shading image using genetic algorithms, in: *Proceedings of the 12th International Conference on Pattern Recognition*, vol. I, Jerusalem, Israel, 1994, pp. 668–670.
- [32] J.J. Atick, P.A. Griffin, A.N. Redlich, Statistical approach to shape from shading: reconstruction of three-dimensional face surfaces from single two-dimensional images, *Neural Computation* 8 (6) (1996) 1321–1340.
- [33] B.-H. Kim, R.-H. Park, Shape from shading and photometric stereo using surface approximation by Legendre polynomials, *Computer Vision and Image Understanding* 66 (3) (1997) 255–270.
- [34] A. Robles-Kelly, E.R. Hancock, A graph-spectral approach to shape-from-shading, *IEEE Transactions on Image Processing* 13 (7) (2004) 912–926.
- [35] H. Ragheb, E.R. Hancock, Separating Lambertian and specular reflectance components using iterated conditional modes, in: *Proceedings of the 12th British Machine Vision Conference*, Manchester, UK, 2001, pp. 541–542.
- [36] P.-S. Tsai, M. Shah, Shape from shading with variable albedo, *Optical Engineering* 37 (4) (1998) 1212–1220.
- [37] P. Breton, S.W. Zucker, Shadows and shading flow fields, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 1996, pp. 782–789.
- [38] S.K. Nayar, K. Ikeuchi, T. Kanade, Shape from inter-reflections, *International Journal of Computer Vision* 6 (3) (1991) 173–195.
- [39] R. March, Visual reconstruction with discontinuities using variational methods, *Image and Vision Computing* 10 (1) (1992) 30–38.
- [40] P.L. Worthington, E.R. Hancock, Needle map recovery using robust regularizers, *Image and Vision Computing* 17 (8) (1999) 545–557.
- [41] T. Rindfleisch, Photometric method for lunar topography, *Photometric Engineering* 32 (2) (1966) 262–277.
- [42] M.A. Penna, A shape from shading analysis for a single perspective image of a polyhedron, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (6) (1989) 545–554.
- [43] J.K. Hasegawa, C.L. Tozzi, Shape from shading with perspective projection and camera calibration, *Computers and Graphics* 20 (3) (1996) 351–364.
- [44] G.E. Farin, *Curves and Surfaces for CAGD: A Practical Guide*, fifth ed., Morgan Kaufman Publishers, Los Altos, 2001.
- [45] P.-S. Tsai, M. Shah, Shape from shading using linear approximation, *Image and Vision Computing* 12 (8) (1994) 487–498.