A Global Solution to the SFS Problem Using B-spline Surface and Simulated Annealing

Frédéric Courteille Jean-Denis Durou
IRIT-UPS, Toulouse, France
{courteille,durou}@irit.fr

Géraldine Morin
IRIT-INPT, Toulouse, France
morin@enseeiht.fr

Abstract

This paper restates the shape from shading problem regarding both surface modeling and optimization. We combine the use of a B-spline as 3D model for the scene surface and the use of stochastic optimization, through simulated annealing. The proposed method overcomes successfully the main difficulties usually encountered in shape from shading. Experiments are presented on both synthetic and real images.

1. Introduction

The shape from shading (SFS) problem consists in recovering the shape of a scene from a single greylevel image. Even if this technique is sometimes considered as a "stylistic exercise", several recent works [2, 9, 13] have proposed more realistic modelings, which allow to consider practical applications. However, three remaining problems often occur: (i) the necessity of boundary conditions, (ii) the very large number of unknowns, and (iii) the concave/convex ambiguity. Our first contribution aims at overcoming problems (i) and (ii) simulteaneously, using a B-spline to approximate the scene shape. Our second contribution consists in solving problem (iii), using the simulated annealing (SA) algorithm. To our knowledge, no existing SFS method provides a solution addressing these three points.

2. State of the Art

3D Models for SFS. In order to render the problem well-posed, rather than adding equations (for instance, imposing boundary conditions), it seems natural to reduce the number of unknowns by choosing a model to approximate the scene shape. This idea has been introduced in [8] through the use of a quadratic model with five parameters. Another approach, based on finite elements methods, has been proposed in [5]. The results are of rather good quality but, since all the pixels are considered as nodes, a major drawback is slowness. This problem has been partially solved in [12], where the same model is implemented, but using a non-regular grid and a multiresolution process. In [11], a superquadric shape with ten parameters is used, but no convincing result is shown. On the other hand, the results reported in [1] are of good quality, even if the method is intrinsically dedicated to face reconstruction. Finally, let us mention some papers [4, 10] 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. To our knowledge, the splines were not really exploited within the framework of SFS, contrary to neighbouring domains like photometric stereo [6] or binocular stereo [7].

Stochastic Optimization for SFS. Probably because of the large number of unknowns, stochastic optimization has not been often used for SFS. In [3], a multiresolution approach coupling SA and deterministic optimization has been successfully designed, but is very slow, due to the huge number of unknowns. In [11], a genetic algorithm is used, but no experiment on real images is proposed.

3. B-spline from Shading

Mathematical Background. We propose to use polynomial uniform splines to solve the SFS problem. We define a 2D function $S_{m,n}$ as the tensor product B-spline:

$$S_{m,n}(s,t) = \sum_{(i,j)\in\mathbb{Z}^2} P_{ij} N_{i,m}(s) N_{j,n}(t), \qquad (1)$$

where $N_{i,m}$ and $N_{j,n}$ are 1D B-spline basis functions of degrees m and n. The real values P_{ij} are the control points of the spline, and here, the parameters of the model. The B-spline basis function $N_{i,m}$ is a piecewise polynomial defined over a knot sequence $\{s_i\}_{i\in\mathbb{Z}}$ recursively. Here, we choose to work with uniform B-splines, that is, $s_i = i$ (resp. $i + 1/2$) if m is odd (resp. even), and therefore

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

$N_{i,m}(s) = N_{0,m}(s-i)$. The function $N_{0,m}$ is characterized by being of degree m on each interval $[i, i+1]$ (resp. $[i-1/2, i+1/2]$)), of class C^{m-1}, and of minimal support, that is, $[-\frac{m+1}{2}, \frac{m+1}{2}]$. We shall consider control points only over a bounded domain. Note that $S_{m,n}$ and its gradient have the same parameters P_{ij}.

B-spline Energies for SFS. We describe the scene in a 3D basis $(Cxyz)$ related to the camera, whose origin C is the optical center (*cf.* Fig 1). A 2D basis (Oxy) is used in the image, whose origin O is the principal point.

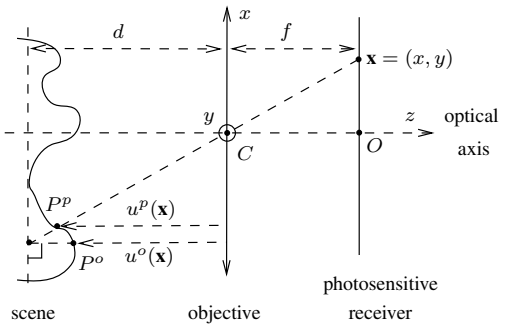


Figure 1. Orthographic projection ($\mathbf{x}$ conjugate with P^o) and perspective projection ($\mathbf{x}$ conjugate with P^p).

Under orthographic projection, a pixel $\mathbf{x} = (x, y)$ is conjugate with object point P^o. The datum is the greylevel $I(\mathbf{x})$. The unknown is the height $u^o(\mathbf{x})$ of P^o. Using the usual SFS hypotheses [2] and assuming the lighting to be parallel, uniform and frontal, the eikonal equation is:

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

where $g = -f/d$ and $I_{\max}$ is the maximal value of I. Therefore, if $S^o_{m,n}$ is the B-spline of degrees (m, n) which characterizes the scene shape, SFS amounts to minimizing the following energy:

$$E^o(P^o_{ij}) = \sum_{\mathbf{x}\in\Omega_u} \left[I(\mathbf{x}) - \frac{I_{\max}}{\sqrt{g^2 \left\|\nabla S^o_{m,n}(\mathbf{x})\right\|^2 + 1}} \right]^2, \tag{3}$$

where Ω_u is a subset of the image containing the pixels which are in accordance with the SFS hypotheses (in our tests, Ω_u is manually designed).

Under perspective projection, $\mathbf{x}$ is conjugate with object point P^p. The new unknown is the height $u^p(\mathbf{x})$ of P^p. Under the same assumptions than before, the perspective eikonal equation is [2, 9, 13]:

$$\hat{g}(\mathbf{x})^2 \left\|\nabla u^p(\mathbf{x})\right\|^2 = \frac{{I_{\max}}^2}{I(\mathbf{x})^2} - 1, \tag{4}$$

where $\hat{g}(\mathbf{x}) = f/(u^p(\mathbf{x}) + \mathbf{x} \cdot \nabla u^p(\mathbf{x}))$. So, if $S^p_{m,n}$ is the B-spline of degrees (m, n) which characterizes the scene shape, SFS amounts to minimizing the following energy:

$$E^p(P^p_{ij}) = \sum_{\mathbf{x}\in\Omega_u} \left[I(\mathbf{x}) - \frac{I_{\max}}{\sqrt{\hat{g}(\mathbf{x})^2 \left\|\nabla S^p_{m,n}(\mathbf{x})\right\|^2 + 1}} \right]^2. \tag{5}$$

One can notice a strong similarity between (3) and (5). Both equations could easily be generalized to non-frontal lighting. These two energies depend on the control points of the respective B-splines, but only E^p depends on the intrinsic parameters of the camera (the focal length f and the coordinates of the principal point O). Note that minimizing these energies can be considered for complicated splines (many control points), and more important, the number of unknowns does not depend on the size of the image, but only on the number of parameters of the 3D model.

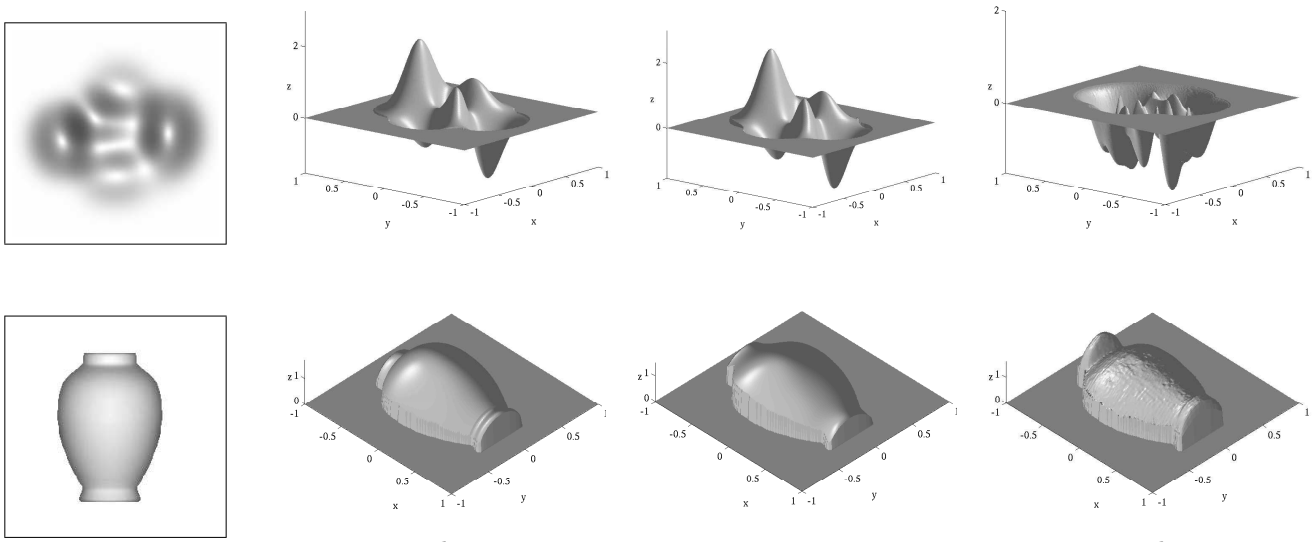
Also, note that only pixels in Ω_u are taken into account to check the accordance of the 3D model with the image. This allows to decrease the CPU time (in our tests, Ω_u usually contains 1% only of the pixels) and to reconstruct the scene by interpolation on parts of the scene outside Ω_u.

4. Experimentations

Now, we shall solve SFS by minimizing the energies given by (3) and (5), with $(m, n) = (3, 3)$, using the SA algorithm. As it is often used to speed up convergence, we choose a geometric series $T_k = \alpha^k T_0$ as cooling scheme, where T_k is the temperature at iteration k, and T_0 the initial temperature. The tests have revealed that $T_0 = 10$ and $\alpha = 0.9998$ are good values to converge to the global minimum, which is reached after about 5.10^4 iterations (at each iteration, all the P_{ij} are sequentially recomputed).

Orthographic Projection. Experiments on two 256×256 images are shown in Fig. 2: in column (a), a DEM[1] image simulated under orthographic projection, and a real image of a vase are shown; column (b) shows the corresponding shapes. The computed surfaces are represented in column (c) using our method (denoted as BS-SA), and in column (d) using Tsai and Shah's method [14] (denoted as TS). We use 16×16 control points for the DEM, and 9×9 for the vase, which are the best compromises between a good reconstructed surface and a low computing time. Visually, BS-SA produces more satisfactory results than TS. For DEM, even if the surface contains several convex and concave areas, the reconstructed surface using BS-SA is very close to the real surface, compared to that obtained using TS. The example of the vase shows that our method is robust to noise:

[1] Function `peaks` of `Matlab`.



(a) (b) (c) (d)

Figure 2. Results under the hypothesis of orthographic projection: (a) two images; (b) the corresponding surfaces; (c) the reconstructed surfaces using BS-SA (Ω_u contains 1% only of the pixels); and (d) the reconstructed surfaces using TS.

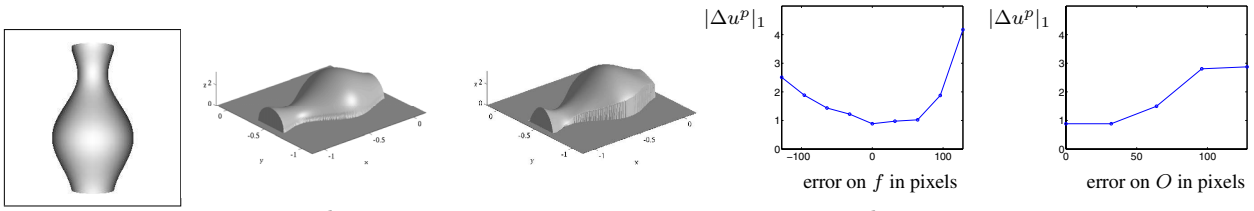
	BS-SA	TS
$\|\Delta u\|_1$	$0.77e-01$	$3.61e-01$
$\|\Delta u\|_\infty$	$0.72e+00$	$2.73e+00$

(a)

	BS-SA	TS
$\|\Delta u\|_1$	$1.90e-01$	$3.30e-01$
$\|\Delta u\|_\infty$	$1.28e+00$	$1.49e+00$

(b)

Table 1. Errors on the computed shapes using BS-SA and using TS: (a) DEM; (b) vase.



(a) (b) (c) (d) (e)

Figure 3. (a) Synthetic image of a vase simulated under perspective projection; (b) corresponding surface; (c) reconstructed surface using BS-SA; (d) $|\Delta u^p|_1$ in function of an error on f; (e) $|\Delta u^p|_1$ in function of an error on the location of O.

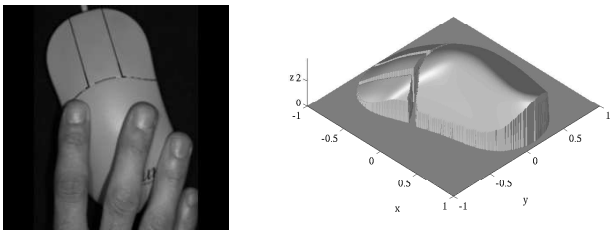


Figure 4. Real image of a mouse, and reconstructed surface using BS-SA.

whereas TS produces a noisy surface, the vase shape computed using BS-SA is smooth. Also notice that even if TS and BS-SA do not suppose any boundary condition, BS-SA produces a more accurate height on the boundary of the vase.

Let us numerically analyze these results, in accordance with the two following error estimators:

$$\begin{cases} |\Delta u|_1 = \frac{1}{\text{card}(\Omega_r)} \sum_{\mathbf{x}\in\Omega_r} |\tilde{u}(\mathbf{x}) - u(\mathbf{x})|, \\ |\Delta u|_\infty = \max_{\mathbf{x}\in\Omega_r} \{|\tilde{u}(\mathbf{x}) - u(\mathbf{x})|\} \end{cases} \quad (6)$$

where Ω_r is the reconstruction domain in the image, $\tilde{u}$ denotes the computed height and u the real height. Table 1 gives the values of these two estimators for the surfaces computed from the DEM and the vase images. We note that, even if the CPU time of our method is much higher (about 1400 sec. against about 1 sec. for TS), the computed surfaces are much more accurate.

Perspective Projection. Now, let us consider the influence of an error concerning the focal length or the location of the principal point. Fig. 3(a) represents the synthetic image of a vase simulated under perspective projection. The corresponding shape is represented in Fig. 3(b), and the computed shape using BS-SA, without errors on the intrinsic parameters, is represented in Fig. 3(c). Fig. 3(d) and Fig. 3(e) represent the graphs of error $|\Delta u^p|_1$ in function of an error on f and in function of an error on the location of O. According to these tests, we can conclude that the reconstruction is acceptable when the error on the estimates does not exceed 20% for the focal length, and 25% of the size of the image for the location of the principal point.

Finally, the last result allows to appreciate another capability of our method: Fig. 4 represents a real image of the mouse, and the computed surface using BS-SA[2]. Whereas the mouse is partially occluded by the hand and moreover, the hand causes shadows, so that some greylevel information is not usable, BS-SA can reconstruct the whole mouse shape rather well, using 9×9 control points.

[2]The intrinsic parameters have been estimated by a classical calibration method.

5. Conclusion and Perspectives

We propose in this work a new SFS method which combines the modeling of the scene shape using a B-spline and stochastic minimization through simulated annealing. According to the results obtained on both synthetic and real images, the proposed approach sounds to be completely convincing.

Many further investigations may be considered, as for example: the use of other 3D models that allow discontinuities in the scene; the automatic choice of the best number of control points; the automatic design of Ω_u.

References

[1] J. J. Atick, P. A. Griffin, and A. N. Redlich. Statistical Approach to SFS: Reconstruction of 3D Face Surfaces from Single 2D Images. *Neural Computation*, 8(6):1321–1340, Aug. 1996.

[2] F. Courteille, A. Crouzil, J.-D. Durou, and P. Gurdjos. Towards SFS under realistic photographic conditions. In *ICPR'04*, volume 2, pages 277–280, Aug. 2004.

[3] A. Crouzil, X. Descombes, and J.-D. Durou. A Multiresolution Approach for SFS Coupling Deterministic and Stochastic Optimization. *PAMI*, 25(11):1416–1421, Nov. 2003.

[4] B.-H. Kim and R.-H. Park. SFS and Photometric Stereo Using Surface Approximation by Legendre Polynomials. *CVIU*, 66(3):255–270, June 1997.

[5] K. M. Lee and C.-C. J. Kuo. SFS with Perspective Projection. *CVGIP: IU*, 59(2):202–212, Mar. 1994.

[6] K. M. Lee and C.-C. J. Kuo. Shape from Photometric Ratio and Stereo. *Journal of Visual Communication and Image Representation*, 7(2):155–162, June 1996.

[7] M. H. Lin and C. Tomasi. Surfaces with occlusions from layered stereo. *PAMI*, 26(8):1073–1078, Aug. 2004.

[8] T. C. Pong, R. M. Haralick, and L. G. Shapiro. Shape from Shading Using the Facet Model. *PR*, 22(6):683–695, 1989.

[9] E. Prados and O. Faugeras. A Generic and Provably Convergent SFS Method for Orthographic and Pinhole Cameras. *IJCV*, 65(1/2):97–125, Nov. 2005.

[10] A. Robles-Kelly and E. R. Hancock. A Graph-Spectral Approach to SFS. *IP*, 13(7):912–926, July 2004.

[11] H. Saito and N. Tsunashima. Estimation of 3-D Parametric Models from Shading Image Using Genetic Algorithms. In *ICPR'94*, volume 1, pages 668–670, Oct. 1994.

[12] D. Samaras and D. N. Metaxas. Incorporating Illumination Constraints in Deformable Models. In *CVPR'98*, pages 322–329, Santa Barbara, California, USA, June 1998.

[13] A. Tankus, N. Sochen, and Y. Yeshurun. SFS Under Perspective Projection. *IJCV*, 63(1):21–43, June 2005.

[14] P.-S. Tsai and M. Shah. SFS Using Linear Approximation. *IVC*, 12(8):487–498, Oct. 1994.

0-7695-2521-0/06 $20.00 © 2006 IEEE