

# METRIC DEPTH RECOVERY FROM MONOCULAR IMAGES USING SHAPE-FROM-SHADING AND SPECULARITIES

*Marco Visentini-Scarzanella*

Communications and  
Signal Processing Group  
Imperial College London  
marcovs@imperial.ac.uk

*Danail Stoyanov*

Centre for Medical Image  
Computing  
University College London  
danail.stoyanov@ucl.ac.uk

*Guang-Zhong Yang*

The Hamlyn Centre for  
Robotic Surgery  
Imperial College London  
g.z.yang@imperial.ac.uk

## ABSTRACT

Despite recent advances in modeling the Shape-from-Shading (SFS) problem and its numerical solution, practical applications have been limited. This is primarily due to the lack of perspective SFS models without the assumption of a light source at the camera centre and the non-metric spatial localisation of the reconstructed shape. In this work, we propose a novel formulation of the SFS problem that allows the reconstruction of surfaces lit by a near point light source away from the camera centre. We also show how knowledge of the light source position can enable the recovery of depth information in a metric space by triangulating specular highlights. Validation of the proposed technique is reported on synthetic and endoscopic images.

## 1. INTRODUCTION

Shape-from-Shading (SFS) as a technique for 3D reconstruction from single frames presents some obvious advantages. A single monocular view is all that is needed, in contrast with other vision-based methods requiring extra cameras or presence of trackable features. This is important where hardware size and image quality are an issue, such as endoscopy.

In their paper, Prados and Faugeras [1] proposed an SFS formulation for the case where the light source is at the optical centre that does not neglect an attenuation term proportional to the square of the distance from the light source  $\frac{1}{r^2}$ . The resulting PDE can be proven to be well-posed and theoretical ambiguities to the solution are removed.

However, several limitations still have to be addressed in SFS systems. Currently, depth can only be recovered up to an unknown scale factor determined by the surface albedo. Moreover, current state-of-the-art formulations do not realistically model the relative spatial configuration of the camera and light source. Furthermore, only purely Lambertian objects with uniform albedo can be considered. Finally, numerical methods for solving PDEs tend to be sensitive to initial conditions and noisy data.

We propose a novel approach for overcoming most of the limitations of conventional SFS methods. First, we present

a novel formulation of the SFS problem with a perspective camera and a light source close to the surface and away from the optical centre. The attenuation term  $\frac{1}{r^2}$  is not neglected, resolving the concave/convex ambiguities. In our case, the resulting Hamiltonian is solved with a Lax-Friedrichs sweeping technique [2] and the impact of different boundary conditions (b.c.) is investigated. Other approaches to SFS using the Lax-Friedrichs scheme have been proposed in the literature [3, 4], however they model the light source as either at the optical centre or infinitely far away.

A formulation similar to ours was developed by Wu et al. [5], however, their mathematical modelling differed from ours in their derivation of the surface normal. In addition, they introduced additional smoothness regularization terms, solving the SFS PDE within a variational framework using Lagrangian multipliers. No spatial localisation was performed.

In contrast, we present an explicit Hamilton-Jacobi PDE. To position the derived surface within metrically accurate coordinates, we solve for the unknown albedo by using spatial triangulation of image specularities. This is possible because we consider a configuration where the camera and light source are not coincident. To our knowledge, this is the first attempt to metrically lock a surface reconstructed with SFS with specular highlights.

Given surface normals obtained from SFS and a known spatial transformation between the camera centre and the light source, it is possible to find the exact 3D position in space of specular points. Since albedo is assumed to be uniform throughout the surface, the actual depth of the entire surface can then be recovered from a single monocular frame.

## 2. SFS WITH LIGHT SOURCE AWAY FROM OPTICAL CENTRE

Compared with the traditional setup of perspective camera and light source at the optical centre [1], considering the more realistic camera-light source setup shown in Fig. 1, a new parametrisation of the surface normal  $\mathbf{n}$  and of the light vector  $\mathbf{l}$  is required. In this setup, the camera  $C$  is at position  $(\alpha, \beta, \gamma)$ , which can be pre-determined with illumination po-

sition estimation methods [6]. Given the set of all  $\mathbf{x} = (x, y)$  in the image domain  $\Omega$ , the surface normal and light vector at the 3D point  $M$  corresponding to image point  $m$  can be represented as:

$$\mathbf{n} = \left( u_x, u_y, -\frac{(x + \alpha)u_x + (y + \beta)u_y + u(\mathbf{x})}{(f + \gamma)} \right) \quad (1)$$

$$\mathbf{l} = (x + \alpha, y + \beta, f + \gamma)$$

Where  $u(\mathbf{x})$  is the depth at point  $\mathbf{x}$  and  $u_x, u_y$  its spatial derivatives. Hence, the image irradiance equation can be expressed in terms of the proposed parametrisations of  $\mathbf{l}$  and  $\mathbf{n}$  without ignoring the distance attenuation term between the light source and the surface to solve the traditional Lambertian SFS scenario where  $I(\mathbf{x}) = \rho \frac{\mathbf{l} \cdot \mathbf{n}}{r^2}$  and  $\rho$  is the surface albedo. By performing the substitution  $v = \ln u$  we obtain the Hamiltonian:

$$H(\mathbf{x}, \nabla v) = I(\mathbf{x}) \frac{1}{\rho} \sqrt{(v_x^2 + v_y^2) + J(\mathbf{x}, \nabla v)^2} \cdot Q(\mathbf{x})^{\frac{3}{2}} \quad (2)$$

where:

$$\begin{cases} J(\mathbf{x}, \nabla v) = \frac{v_x(x + \alpha) + v_y(y + \beta) + 1}{(f + \gamma)} \\ Q(\mathbf{x}) = (x + \alpha)^2 + (y + \beta)^2 + (f + \gamma)^2 \end{cases} \quad (3)$$

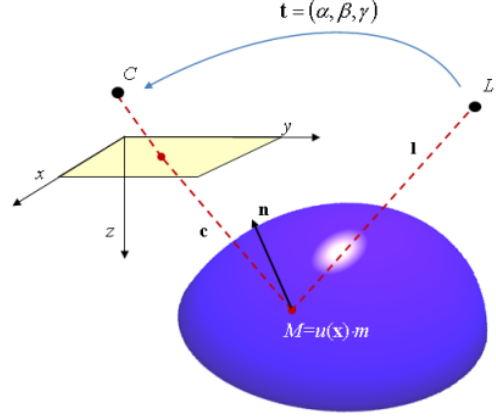
It is important to highlight that, excluding the method proposed in [5], all other approaches to SFS, including [1, 3, 4, 7], consider the case of perspective cameras with the light source either at the optical centre or infinitely far away. The proposed methods results not only in a physically less realistic model, but also in the impossibility to triangulate specularities in order to solve for the unknown albedo and locate the reconstructed shape in a metric space.

### 3. NUMERICAL ALGORITHM

To solve equation (2) up to the unknown scale factor, the 2D Lax-Friedrichs sweeping scheme proposed by Kao *et al.* [2] was applied. The scheme is an iterative numerical method able to find solutions to non-convex Hamiltonians of arbitrary complexity. In the Lax-Friedrichs scheme formulation, our SFS problem can be stated as:

$$\begin{cases} H(\mathbf{x}, \nabla v) - e^{-2v(\mathbf{x})} = R, & \forall \mathbf{x} \in \Omega \\ v(\mathbf{x}) = \phi(\mathbf{x}), & \forall \mathbf{x} \in \partial\Omega \end{cases} \quad (4)$$

where  $R$  is a small positive constant set in our experiments to  $10e^{-5}$  and  $\phi(\mathbf{x})$  are the boundary conditions. In this work we investigate Dirichlet b.c. where  $\phi(\mathbf{x}) = u(\mathbf{x})$  and Neumann b.c. where  $u_x = u_y = 0$ .

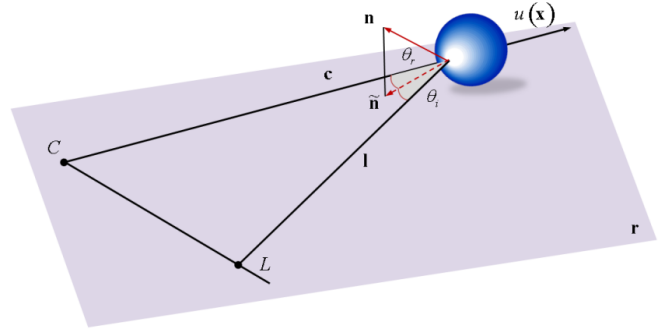


**Fig. 1.** Perspective projection with light source away from the optical centre. A 3D point  $M$  with surface normal  $\mathbf{n}$  is illuminated by a light source  $L$  through the illumination vector  $\mathbf{l}$  and projected to a point  $m$  on the retinal plane by a camera  $C$ .

A simple iterative procedure for the solution of (4) exists and for further details we refer the reader to the work in [2] and its application to SFS in [7]. The full implementation of the SFS algorithm is available for download on the author's page<sup>1</sup>.

### 4. RECOVERING METRIC DEPTH

The albedo is then recovered by treating specular highlights as constraints allowing the recovery of unscaled metric depth at specular points.



**Fig. 2.** Geometry of specular highlights. The peak of the highlight occurs whenever the angle of light incidence corresponds to the viewing angle.

Given the setup in Fig. 2, the peak of the specular highlight will occur when the angle of light incidence  $\theta_i$  between the light vector  $\mathbf{l}$  and the normal  $\mathbf{n}$  is the same as the angle of reflection  $\theta_r$  between the viewing vector  $\mathbf{c}$  and the normal.

<sup>1</sup><http://www.commsp.ee.ic.ac.uk/~marcovs/shading/>

This corresponds to the condition of the normal and the half-angle vector  $\mathbf{h}$  between  $\mathbf{c}$  and  $\mathbf{l}$  being perfectly aligned and can be expressed as:

$$\frac{(\mathbf{l} + \mathbf{c}) \cdot \mathbf{n}}{\|\mathbf{l} + \mathbf{c}\| \|\mathbf{n}\|} = \cos \theta_{hn} = 1 \quad (5)$$

As information about the normal is obtained directly from the SFS estimate and the  $(x, y)$  image coordinates of the peak of the specular highlight can be easily localised, (5) can be expressed as a quadratic constraint in a single unknown  $u(\mathbf{x})$ :

$$u(\mathbf{x})^2 (k_0 \|\mathbf{n}\|^2 - k_3^2) + u(\mathbf{x}) (k_1 \|\mathbf{n}\|^2 - 2k_3 k_4) + (k_2 \|\mathbf{n}\|^2 - k_4^2) = 0 \quad (6)$$

where:

$$\begin{cases} k_0 = \frac{4}{(f + \gamma)^2} Q(\mathbf{x}) \\ k_1 = \frac{4}{(f + \gamma)} [(x + \alpha)\alpha + (y + \beta)\beta + (f + \gamma)\gamma] \\ k_2 = \alpha^2 + \beta^2 + \gamma^2 \\ k_3 = -\frac{2[(x + \alpha)n_0 + (y + \beta)n_1 + (f + \gamma)n_2]}{(f + \gamma)} \\ k_4 = \alpha n_0 + \beta n_1 + \gamma n_2 \end{cases} \quad (7)$$

If the surface albedo and irradiance model are uniform throughout the visualised surface, the ratio between the surface depth calculated through (6) and the one recovered by our SFS method represents the unknown parameter  $\rho$ .

The solution of (6) assumes the existence of a perfect specular highlight detected, i.e.  $\theta_{hn} = 0$ . However, the irradiance is sampled at intervals related to the camera resolution, and bright highlights will also arise when  $\mathbf{h} \approx \mathbf{n}$  causing (6) not to have any real roots. The problem can be regularised by projecting the estimated normal to a new normal  $\tilde{\mathbf{n}}$  lying on the plane with normal  $\hat{\mathbf{r}}$  including  $\mathbf{c}$  and the light source  $L$ :

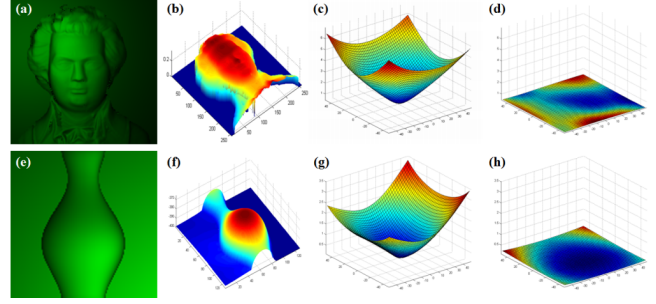
$$\tilde{\mathbf{n}} = \mathbf{n} - (\mathbf{n} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}} \quad (8)$$

## 5. RESULTS

First, the performance of the new SFS formulation for reconstructing object shape was tested against state-of-the-art methods presented in [1] and [8] to quantify the impact of the new formulation with the off-axis light source. Second, the recovery of the unknown albedo via specular highlight triangulation to localise the reconstructed shape in metric space was also separately validated.

### 5.1. Synthetic data

We used the standard SFS datasets of Mozart and Vase for comparison during the experiments. For both phases of the experiment, surfaces were reconstructed for a volume of light source positions  $(\alpha, \beta, \gamma)$  of side-length 48mm. Images were



**Fig. 3.** (a) Example of test image 'Mozart' lit from top left. (b) SFS reconstruction error for combinations of  $(\alpha, \beta)$  light positions are shown for (c) the methods in [1, 8] and (d) our method. The bottom row shows a similar analysis for 'Vase'.

generated with a Blinn-Phong reflectance model with an exponent of 100. Qualitatively, the reconstructed shapes are indistinguishable from the originals for most light configurations, as shown in Fig. 3. For both datasets, the error is generally uniform for all spatial configurations, while in formulations where the light source is at the optical centre the error becomes predictably progressively greater for large baseline light configurations. Significantly, there is little difference between Dirichlet and Neumann boundary conditions, suggesting that acceptable results can be obtained with no prior knowledge of the surface at the image boundaries.

The performance of the triangulation scheme was investigated in three phases: using the ground truth surface normal data to provide a baseline for the algorithm's accuracy, then with the surface normal estimates obtained from our method, and finally with the SFS estimates after automatic removal of areas with high estimates of surface curvature.

Results in table 1 suggest that when using ground truth normal data the average error is around 9.5% for both Mozart and Vase. This is due to the fact that the normals at the strongest specular highlights present are not exactly coincident with the vector  $\mathbf{h}$ . However, the scheme proves to be robust when using SFS surface normal estimates, with Vase localised with an error of 2.7% above the ground truth baseline. For Mozart, the error using raw SFS surface normal estimates is 19.2%, or 9.4% above the baseline, due to the inability of the Lax-Friedrichs scheme to correctly reconstruct sharp edges. By automatically excluding areas with high estimated surface curvature the error is reduced to 2.6% above the baseline. It is important to stress that due to the coincident configuration of camera and light source, it is impossible to perform triangulation using the formulation in [1, 8].

### 5.2. Endoscopic images

Behind our method, there is an assumption about the existence of a system able to separate diffuse and specular components from the input. However, during endoscopy the highly

focused lighting causes most specular pixels to be completely saturated and devoid of any chromatic information, causing the state-of-the-art algorithms such as [9, 10] to fail. Because of this, only the SFS stage of the algorithm was applied, with qualitative results shown in Fig. 4.

## 6. CONCLUSIONS

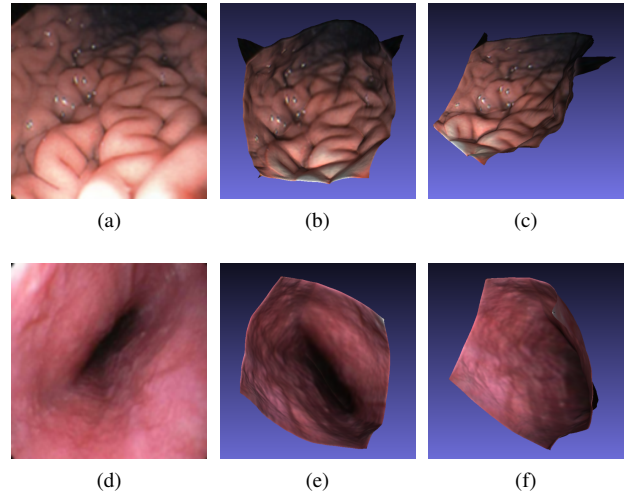
We have proposed a novel formulation of SFS which is a further step towards realistic modeling where the light source is located close to the surface and away from the optical centre. We showed how this configuration allows for spatial metric localisation of the displayed object through triangulation of specular highlights using surface normals estimated from the SFS procedure. Further work will concentrate on reliably separating diffuse and specular image components and on modelling surfaces with locally variable albedo.

Shape reconstruction error		
	Vase	Mozart
Proposed with Dirichlet b.c.	0.11%	0.47%
Proposed with Neumann b.c.	0.20%	0.59%
SFS [1, 8]	1.46%	3.37%
Specular highlight triangulation error		
	Vase	Mozart
Using ground truth	9.22%	9.80%
Using SFS estimate	11.96%	19.23%
SFS estimate + edge filtering	11.96%	12.42%

**Table 1.** Surface reconstruction and localisation results.

## 7. REFERENCES

- [1] E. Prados and O. Faugeras, “Shape from shading: a well-posed problem?,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR ’05)*, San Diego, CA, 2005, vol. 2, pp. 870–877.
- [2] Chiu-Yen Kao, Stanley Osher, and Yen-Hsi Richard Tsai, “Fast sweeping methods for static hamilton-jacobi equations,” *SIAM J. Numerical Analysis*, vol. 42, no. 6, pp. 2612–2632, 2005.
- [3] Li Zhang, A. M. Yip, M. S. Brown, and Chew Lim Tan, “A unified framework for document restoration using inpainting and shape-from-shading,” *Pattern Recogn.*, vol. 42, no. 11, pp. 2961–2978, 2009.
- [4] Lei Yang and Jiu-Qiang Han, “A perspective shape-from-shading method using fast sweeping numerical scheme,” *Optica Applicata*, vol. 38, no. 2, pp. 387–398, 2008.



**Fig. 4.** Results for *in vivo* data courtesy of www.gastrolab.net. (a) Stomach lining and (b), (c) renderings of its reprojected reconstruction. (d) Oesophagus and (e), (f) renderings of its reprojected reconstruction.

- [5] Chenyu Wu, Srinivasa G. Narasimhan, and Branislav Jaramaz, “A multi-image shape-from-shading framework for near-lighting perspective endoscopes,” *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 211–228, 2010.
- [6] Danail Stoyanov, Daniel S. Elson, and Guang-Zhong Yang, “Illumination position estimation for 3d soft-tissue reconstruction in robotic minimally invasive surgery,” in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2009, pp. 2628–2633, IEEE.
- [7] A. H. Ahmed and A. A. Farag, “Shape from shading under various imaging conditions,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR ’07)*, Minneapolis, MN, 2007, vol. 0, pp. 1–8.
- [8] A. H. Ahmed and A. A. Farag, “Shape from shading for hybrid surfaces,” in *IEEE International Conference on Image Processing (ICIP 2007)*, Atlanta, GA, 2007, vol. 2, pp. 525–528.
- [9] Robby T. Tan and Katsushi Ikeuchi, “Separating reflection components of textured surfaces using a single image,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 178–193, 2005.
- [10] Qingxiong Yang, Shengnan Wang, and Narendra Ahuja, “Real-time specular highlight removal using bilateral filtering,” in *ECCV (4)*, 2010, pp. 87–100.