A Comparison of Regularization Methods for Near-Light-Source Perspective Shape-from-Shading

Bilal Ahmad, Pål Anders Floor, Ivar Farup
Department of Computer Science,
Norwegian University of Science & Technology, 2815 Gjøvik, Norway.
Email: {bilal.ahmad, paal.anders.floor, ivar.farup}@ntnu.no

***Abstract*—3D shape reconstruction from images is an active topic in computer vision. Shape-from-Shading is an important approach which requires the surface properties and light source position to infer the 3D shape. A L2 regularizer is typically used to penalize the irradiance equation. In this article, anisotropic diffusion (AD) is introduced as a regularizer to solve the image irradiance equation. The method is then compared with L1 and L2 regularization methods, where all of the three techniques are formulated using gradient descent. Results shows that with AD, edges can be better preserved. AD shows lower depth error and higher correlation when compared with L1 and L2 regularization methods.**

***Index Terms*—Shape-from-Shading, near-light-source, anisotropic diffusion, L2 regularization, L1 regularization**

I. Introduction

3D shape reconstruction from images is a challenging problem in the field of computer vision. There are several techniques for shape reconstruction, such as stereo vision, structured light, shape-from-X (X = Shading, texture, focus/defocus, motion). Shape-from-shading (SfS) is an important approach when it comes to recover the three dimensional shape of an object from a single image. In some imaging applications [1], [2], only one shot of the scene may be available. In such situations, SfS could be a preferable choice for 3D reconstruction.

SfS was first discussed by Horn and Brooks [3], who formulated a non-linear partial differential equation (PDE) by relating intensity variation in the image to the 3D shape of the objects. SfS consists of two steps. In the first step, a reflection model is developed by reflectance properties of the surface, position of the camera and light source. In the second step, a numerical scheme is designed to minimize the difference between image irradiance and reflection function, the so-called image irradiance equation (IIE), which typically involves solving a non-linear PDE.

Most of the SfS approaches assume that the surface follows Lambertian reflection while solving the numerical scheme [4]. Roy and Tourin [5], the pioneers of viscosity based solution, solve IIE as a Hamilton Jacobi PDE provided exact boundary data. Several other schemes are also formulated based on viscosity solution theory [6], [7]. Tankus *et al.* [8] derived the IIE equation and obtained the approximate solution by first locally recover the surface by paraboloids and then obtain a global solution within a single scale factor. Queau *et al.* [9] presented a solution built on augmented Lagrangian approach

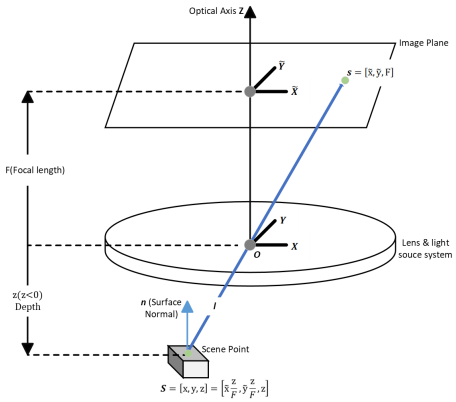


Fig. 1: Perspective Shape-from-Shading model with Light source at the camera projection center **O**. Camera coordinate system (x, y, z) is centered at **O**. z-axis is parallel to optical axis pointing towards image plane.

for solving PDE. Wu *et al.* [10] also formulated SfS as a minimization problem but modeled the reflection function with perspective projection and multiple light sources around camera center.

Numerical schemes are mostly implemented with L2 regularization to penalize the irradiance equation. With L2, sharp edges in 3D models tend to be somewhat smoothened because of linear diffusion. However, the effect of using different regularizers in SfS has not yet been thoroughly investigated.

In this paper, a non-linear, space-variant technique, called an-isotropic diffusion (AD) [11], is introduced as a smoothness constraint to solve the IIE. The aim of AD is in our application is to reduce the noise from the surface without smoothing edges, lines or other details which are important to interpret the surface. IIE is also solved with L1 and L2 regularization methods and the results of all three methods are compared with each other.

The remainder of this article is organized as follows. Section II explains perspective SfS model along with different numerical schemes. Results are compared and discussed in Section III and Section IV concludes the article.

II. Point Light Source Perspective SFS Model

In this section, a brief description of SfS under point lighting and perspective projection is provided with light source at camera center as shown in Fig. 1. Radiance emitted by the surface point $\mathbf{S}$ can be computed according to Lambertian cosine law and inverse square distance fall of law of point light source [10],

$$R(\widetilde{x},\widetilde{y},z,p,q) = I_o\rho\left(\frac{\mathbf{n}(\widetilde{x},\widetilde{y},z,p,q)\cdot\mathbf{l}(\widetilde{x},\widetilde{y},z)}{r(\widetilde{x},\widetilde{y},z)^2}\right), \quad (1)$$

where I_o is the light intensity and ρ is the surface albedo. $p = \frac{\partial z}{\partial \widetilde{x}}$ and $q = \frac{\partial z}{\partial \widetilde{y}}$ are the surface gradients. $\mathbf{n}$ denotes the surface unit normal and $\mathbf{l}$ is a unit vector representing the direction of the light ray incident at $\mathbf{S}$. r^2 is inverse square distance fall-off law of isotropic point light. The light source is considered at the camera center, but can easily be extended to multiple point light source not necessarily at the center [10].

The surface normal $\mathbf{n}$ can be represented in terms of partial derivatives of the depth z with respect to x and y [3]:

$$\mathbf{n} = \frac{[-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, 1]}{\sqrt{(\frac{\partial z}{\partial x})^2 + (\frac{\partial z}{\partial y})^2 + 1}}, \quad (2)$$

where (x, y, z) are camera coordinates. Under perspective projection we have,

$$x = \widetilde{x}\frac{z}{F} \qquad y = \widetilde{y}\frac{z}{F}, \quad (3)$$

where F is the focal length and $(\widetilde{x},\widetilde{y})$ are image coordinates, and camera is pointing in the negative z-direction as depicted in Fig. 1.

According to Horn and Brooks [3], the image irradiance equation is,

$$R(\widetilde{x},\widetilde{y},z,p,q) = I(\widetilde{x},\widetilde{y}). \quad (4)$$

Equation (4) is solved to estimate z by minimizing the difference between image irradiance $I(\widetilde{x},\widetilde{y})$ and reflectance map $R(\widetilde{x},\widetilde{y},z,p,q)$ given in (4). Different from the previous optimization method [10] only depth z is optimized. p and q are updated by taking the gradient of updated z as both p and q are derived from z. By updating them independently one may loose its connection with z. The error $E(z)$ can be computed as,

$$\arg\min_z E(z) = \lambda e_i(z) + (1-\lambda)e_s(z), \quad (5)$$

where e_i is the irradiance error and e_s represents smoothness constraint. λ is the weighting factor between e_i and e_s.

$e_i(z)$ can be computed over the image domain ($\Omega \subset \mathbb{R}$) as,

$$e_i(z) = \int_\Omega (I(\widetilde{x},\widetilde{y}) - R(\widetilde{x},\widetilde{y},z,p,q))^2 d\Omega. \quad (6)$$

$e_s(z)$ is solved by all three methods (L2, L1, AD). It is then combined with $e_i(z)$ according to Equation (5) which is then solved using gradient descent. A small time step Δt is introduced to ensure stability with higher values of λ.

A. L2 Regularization

$e_s(z)$ can be solved with L2 regularization method such as,

$$e_{s(L2)}(z) = \int_\Omega \mid \nabla z \mid^2 d\Omega. \quad (7)$$

Equation (6) and (7) are solved and the results are incorporated in Equation (5) which is then solved using gradient decent,

$$\frac{\partial z}{\partial t} = \nabla^2 z + \frac{\lambda}{1-\lambda}(I-R)\frac{\partial R}{\partial z}. \quad (8)$$

B. L1 Regularization

With L1 regularization method $e_s(z)$ becomes,

$$e_{s(L1)}(z) = \int_\Omega \mid \nabla z \mid d\Omega. \quad (9)$$

Combining the results of Equation (6) and (9) in Equation (5) and then solving yields,

$$\frac{\partial z}{\partial t} = \nabla\cdot\left(\frac{\nabla z}{\mid \nabla z \mid}\right) + \frac{\lambda}{1-\lambda}(I-R)\frac{\partial R}{\partial z}. \quad (10)$$

C. Anisotropic Diffusion

A 2×2 structure tensor is derived from the gradient of the depth z which is given as,

$$S_{i,j} = \frac{\partial z}{\partial x^i}\frac{\partial z}{\partial y^j}. \quad (11)$$

Afterwards, corresponding eigenvalues (λ_+, λ_-) and eigenvectors (θ_+, θ_-) are derived similar to [12]. In terms of eigenvalues, langrangian density $\psi(\lambda_+, \lambda_-)$ can be written as [13],

$$e_{s(AD)}(z) = \int_\Omega \psi(\lambda_+, \lambda_-) d\Omega. \quad (12)$$

The corresponding Euler-Langrange equation of (5), after putting the solution of Equations (6) and (12) in it, gives,

$$\frac{\partial z}{\partial t} = \nabla\cdot(D\nabla z) + \frac{\lambda}{1-\lambda}(I-R)\frac{\partial R}{\partial z}, \quad (13)$$

where D is the diffusion tensor which is computed as,

$$D = \frac{\partial \psi}{\partial \lambda_+}\theta_+\theta_+^T + \frac{\partial \psi}{\partial \lambda_-}\theta_-\theta_-^T. \quad (14)$$

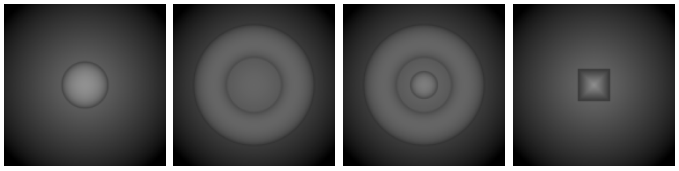
For simplicity and linearizing the terms with diffusion tensor to compute it once and for all, structure tensor in Equation (11) is derived from the gray scale image $\mu(\widetilde{x},\widetilde{y})$.

III. Results & Discussions

A. Ground Truth Models

The methods are tested on four different objects: Sphere, Torus, Sphere inside a Torus (Sphus) and Pyramid, which are modeled in Blender. These objects are used as a ground truth when compared with recovered 3D shapes. Blender is mainly chosen to construct a ground truth scenario by controlling different parameters such as light intensity (I_o), focal length (F) etc.

An environment is designed similar to Fig. (1). The camera is placed at $(0, 0, 0)$. F of the camera is set to 25mm. A point light source is also placed at camera center and power (P) is



(a) Sphere (b) Torus (c) Sphus (d) Pyramid

Fig. 2: Rendered images from Blender

set to 200mW where $I_o = P/4\pi$ in Equation (1). A flat surface is placed below the camera center at negative z-axis. The four different objects are also placed on the flat surface (one model at a time) such that the it is cutting the three shapes (sphere, torus, sphus) from the center while the pyramid is placed on top of flat surface. Material properties of all the models are chosen to be Diffuse BSDF with a constant albedo $\rho = 1$. Ground truth models are shown in Fig. (3).

B. Image Irradiance from Blender Images

200×200 size images are rendered for all the models. Images are shown in Fig. (2). $I(\widetilde{x}, \widetilde{y})$ falling on the camera sensor is related to gray scale image $\mu(\widetilde{x}, \widetilde{y})$ via camera response function $r(\cdot)$ [10],

$$I(\widetilde{x}, \widetilde{y}) = \frac{r^{-1}[\mu(\widetilde{x}, \widetilde{y})]}{M(\widetilde{x}, \widetilde{y})}, \tag{15}$$

where $M(\widetilde{x}, \widetilde{y})$ is anisotropy of the light source. Point lights are perfectly isotropic by defination and so $M(\widetilde{x}, \widetilde{y}) = 1$. Images are saved in Portable Network Graphics (PNG) file format and therefore, image irradiance is just the gamma correction $\gamma = 2.2$ of the gray scale image i.e.,

$$I(\widetilde{x}, \widetilde{y}) = \mu^{\gamma}(\widetilde{x}, \widetilde{y}). \tag{16}$$

$I(\widetilde{x}, \widetilde{y})$ is also converted from pixel units to physical units in order to have correspondence between $I(\widetilde{x}, \widetilde{y})$ and R. Conversion to physical units is given by,

$$\begin{aligned} I_p(\widetilde{x}, \widetilde{y}) = & \frac{I(\widetilde{x}, \widetilde{y}) - \min I(\widetilde{x}, \widetilde{y})}{\max I(\widetilde{x}, \widetilde{y}) - \min I(\widetilde{x}, \widetilde{y})} \\ & \times \left(\frac{I_o \cos\theta_1}{r_1^2} - \frac{I_o \cos\theta_2}{r_2^2} \right) + \frac{I_o \cos\theta_2}{r_2^2}, \end{aligned} \tag{17}$$

where $I_p(\widetilde{x}, \widetilde{y})$ represents the physical value of the image irradiance and (θ_1, r_1) and (θ_2, r_2) decides the upper and lower bound of $I_p(\widetilde{x}, \widetilde{y})$. (θ_1, θ_2) are the angle between surface normal and light ray at the maximum and minimum point on the surface respectively. (r_1, r_2) are the distance from the light source to the maximum and minimum point on the surface respectively. These points are experimental in our case and chosen in order to have the same scale between I_p and reflection function R, which is computed in the first step.

C. 3D Reconstruction

A flat surface is given as an initial condition for all objects in order to test the robustness of the algorithm. p and q of initial z are computed by taking the gradient in x and y directions. Initial reflectance map is then computed using equation (4). Updated z values are calculated by solving equations (8), (10) and (13) for L2, L1, AD respectively. p and q are updated as the gradient of z. For L2 regularization, λ is kept to 0.9997. The value of λ is empirical in our experiment. Recovered 3D shape with L2 regularization are shown in Fig. (4). Although, results look nice but L2 has smoothened those areas where the plane is intersecting the shapes. This is also one of the reason to compute the structure tensor in Equation (11) from gray scale image $\mu(\widetilde{x}, \widetilde{y})$ instead of depth z for better edge information.

Both L1 and AD are sensitive to the initial condition. Therefore, the reconstructed surface from L2 regularization is fed as an initial condition to both of these methods. λ is then tuned a bit while the number of iterations are kept the same for all three methods. Reconstructed surfaces with L1 and AD are shown in Fig. (5) and (6) respectively.

In order to evaluate the robustness of the different methods, the reconstructed surfaces are compared with ground truth by measuring correlation and depth error. These methods are chosen to assess different features of the reconstructed surfaces. Correlation is chosen because it evaluates the shape of the reconstructed surface independent of scale and position. Depth error (e_d) is chosen because it is scale dependent and it will correctly evaluate the geometric deformation of the reconstructed shape. e_d is given by,

$$e_d = \frac{1}{\Omega} \sum \sum_{i,j \in \Omega} \left| \frac{\widehat{D}_{i,j} - D_{i,j}}{D_{i,j}} \right|, \tag{18}$$

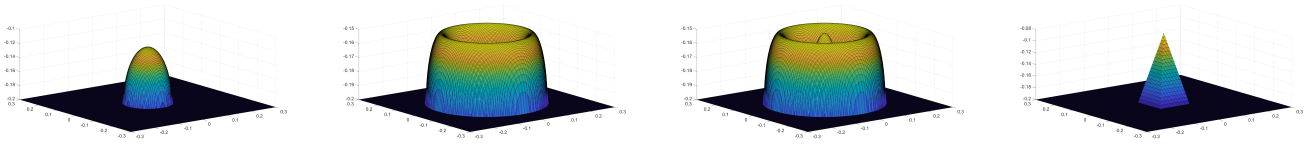
where D is the ground truth and $\widehat{D}$ is recovered 3D shape. Ω represents the region of the 3D model considered for error estimation.

TABLE I shows that AD performs well with lower depth error for most of the cases and higher correlation for all the cases. AD also manages to preserve edges better than L2. Although L1 preserves edges to some degree, it has also flattened areas which are curved in the ground truth models.

SfS algorithms are mostly tested without the availability of ground truth data and hence, it is very challenging to predict the quality of 3D reconstruction. Irradiance error e_i is therefore compared with depth error e_d for all four shapes. Fig. (7) shows that both errors correlates well with each other. e_i

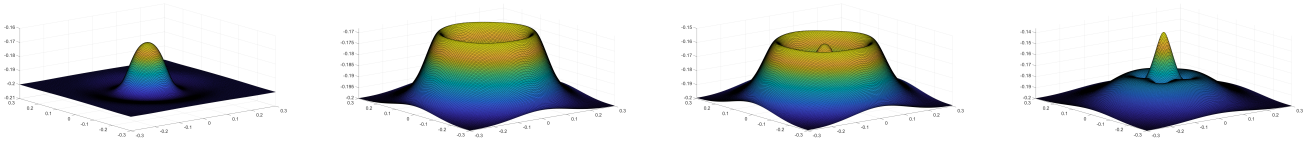
Shape	Method	Depth Error	Correlation
Sphere	L2	0.0171	0.9718
	L1	0.0163	0.9854
	AD	**0.0155**	**0.9937**
Torus	L2	0.0480	0.8694
	L1	0.0491	0.8720
	AD	**0.0410**	**0.9098**
Sphus	L2	0.0540	0.8292
	L1	0.0593	0.8083
	AD	**0.0539**	**0.8729**
Pyramid	L2	**0.0532**	0.5241
	L1	0.0540	0.5201
	AD	0.0536	**0.6207**

TABLE I: Comparison of L2, L1 and AD methods



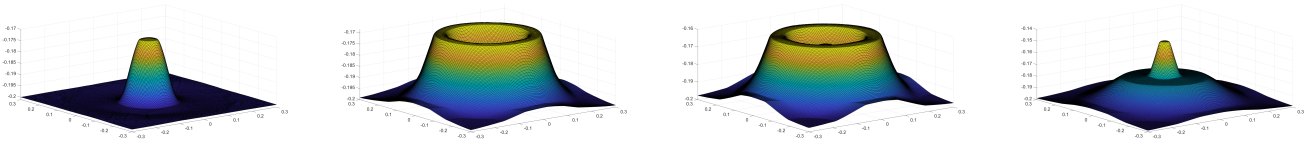
(a) Sphere (b) Torus (c) Sphus (d) Pyramid

Fig. 3: Simple objects for ground truth



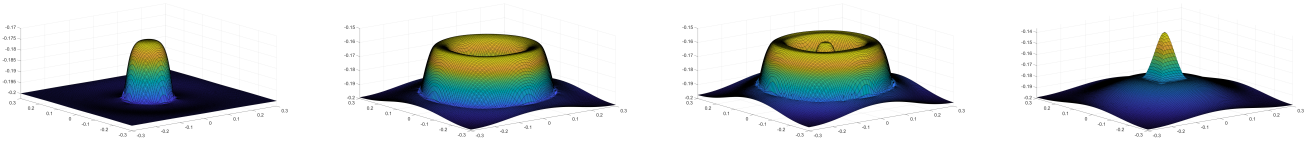
(a) Sphere (b) Torus (c) Sphus (d) Pyramid

Fig. 4: 3D shapes recovered with L2 regularization



(a) Sphere (b) Torus (c) Sphus (d) Pyramid

Fig. 5: 3D shapes recovered with L1 regularization



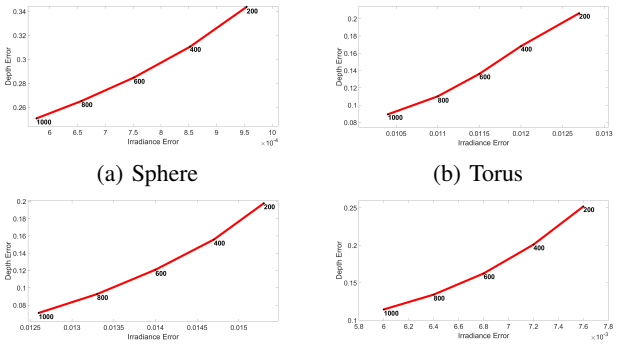
(a) Sphere (b) Torus (c) Sphus (d) Pyramid

Fig. 6: 3D shapes recovered with Anisotropic Diffusion

reduces with e_d as number of iterations increases for all four objects. This confirms that when SfS techniques is applied in situations where no ground truth is available, e_i can be considered an indicator of the quality of 3D reconstruction.

IV. Conclusion

In this article, near light source perspective SfS method with different regularizers is discussed to recover 3D shapes. Given a reflection model, numerical schemes are formulated with three different regularizers (L2, L1, AD). The resulting shape for each scheme is then compared by measuring the average depth error and correlation with respect to ground truth models. Results shows that the shapes recovered with AD has less error and preserve edges better than the other two schemes. The irradiance error e_i is also compared with the depth error e_d to confirm their consistency so that when these methods will be applied in conditions where no ground truth is available, irradiance error can be considered as an indicator of the quality of 3D reconstruction.



(a) Sphere (b) Torus

(c) Sphus (d) Pyramid

Fig. 7: Irradiance VS Depth Error for different shapes.

Acknowledgment

Funding was provided by the Research Council of Norway under the project CAPSULE no. 300031.

REFERENCES

[1] W. C. Liu, B. Wu, and C. Wöhler, "Effects of illumination differences on photometric stereo shape-and-albedo-from-shading for precision lunar surface reconstruction," *ISPRS journal of photogrammetry and remote sensing*, vol. 136, pp. 58–72, 2018.

[2] M. Bas and D. Spinczyk, "Evaluation of shape from shading surface reconstruction quality for liver phantom," in *Information Technology in Biomedicine*. Springer, 2021, pp. 81–93.

[3] B. K. Horn and M. J. Brooks, "The variational approach to shape from shading," *Computer Vision, Graphics, and Image Processing*, vol. 33, no. 2, pp. 174–208, 1986.

[4] J.-D. Durou, M. Falcone, and M. Sagona, "Numerical methods for shape-from-shading: A new survey with benchmarks," *Computer Vision and Image Understanding*, vol. 109, no. 1, pp. 22–43, 2008.

[5] E. Rouy and A. Tourin, "A viscosity solutions approach to shape-from-shading," *SIAM Journal on Numerical Analysis*, vol. 29, no. 3, pp. 867–884, 1992.

[6] E. Prados and O. D. Faugeras, "" perspective shape from shading" and viscosity solutions." in *ICCV*, vol. 3, 2003, p. 826.

[7] G. Wang and Y. Chu, "A new oren–nayar shape-from-shading approach for 3d reconstruction using high-order godunov-based scheme," *Algorithms*, vol. 11, no. 5, p. 75, 2018.

[8] A. Tankus, N. Sochen, and Y. Yeshurun, "A new perspective [on] shape-from-shading," in *null*. IEEE, 2003, p. 862.

[9] Y. Quéau, J. Mélou, F. Castan, D. Cremers, and J.-D. Durou, "A variational approach to shape-from-shading under natural illumination," in *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer, 2017, pp. 342–357.

[10] C. Wu, S. G. Narasimhan, and B. 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.

[11] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[12] G. Sapiro and D. L. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering," *IEEE transactions on image processing*, vol. 5, no. 11, pp. 1582–1586, 1996.

[13] D. Tschumperlé and R. Deriche, "Vector-valued image regularization with pdes: A common framework for different applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 4, pp. 506–517, 2005.