sciendo

Int. J. Appl. Math. Comput. Sci., 2021, Vol. 31, No. 3, 501–513
DOI: 10.34768/amcs-2021-0034

amcs

A LINEARIZATION–BASED HYBRID APPROACH FOR 3D RECONSTRUCTION OF OBJECTS IN A SINGLE IMAGE

MUHAMMED KOTAN [a,*], CEMIL ÖZ [b], ABDULGANI KAHRAMAN [c]

[a]Department of Information Systems Engineering
Sakarya University
Esentepe Campus, Serdivan, Sakarya, 54050, Turkey
e-mail: mkotan@sakarya.edu.tr

[b]Department of Computer Engineering
Sakarya University
Esentepe Campus, Serdivan, Sakarya, 54050, Turkey
e-mail: coz@sakarya.edu.tr

[c]Department of Computer Engineering and Computer Science
University of Louisville
Belknap Campus, Louisville, KY 40292, USA

The shape-from-shading (SFS) technique uses the pattern of shading in images in order to obtain 3D view information. By virtue of their ease of implementation, linearization-based SFS algorithms are frequently used in the literature. In this study, Fourier coefficients of central differences obtained from gray-level images are employed, and two basic linearization-based algorithms are combined. By using the functionally generated surfaces and 3D reconstruction datasets, the hybrid algorithm is compared with linearization-based approaches. Five different evaluation metrics are applied on recovered depth maps and the corresponding gray-level images. The results on defective sample surfaces are also included to show the effect of the algorithm on surface reconstruction. The proposed method can prevent erroneous estimates on object boundaries and produce satisfactory 3D reconstruction results in a low number of iterations.

Keywords: 3D reconstruction, shape from shading, shape recovery, image comparison, surface inspection.

1. Introduction

Shape-from-shading (SFS) uses shading variations in images and creates corresponding 3D surfaces. In addition to the mathematical and algorithmic development of the problem, the 3D depth information obtained is used in many different fields such as surface topography and terrain analyses (e.g., Worthington and Hancock, 2001; Wu *et al*., 2020), biometric studies (e.g., Cadavid and Abdel-Mottaleb, 2008; Kemelmacher-Shlizerman and Basri, 2010), industrial quality control (e.g., Kong, 2008; Kotan and Öz, 2017), and medical diagnosis and treatment (e.g., Yamany and Farag, 1998; Ciaccio *et al*., 2017).

In some survey studies (Zhang *et al*., 1999; Durou *et al*., 2008), the Tsai-Shah (1994) and Pentland (1989) algorithms, which are the basis of this paper, were examined under the class of linearization-based methods that look for a solution by linearizing the image irradiance equation. Pentland applies the Fourier transform in order to obtain a closed-form solution for the depth at each point. Tsai-Shah initially employs the discrete approach of the gradient and then addresses the linear approach of the reflection function directly in terms of depth. The comparison of both methods, Pentland's and Tsai-Shah's, is also presented as a separate heading in the study of Tsai-Shah.

We were inspired by the Tsai-Shah instance, which concluded that "In our case, without any prior knowledge about the input image, the best initial estimation of depth, $Z(x, y)$, for each pixel is zero" (Ping-Sing and Shah, 1994). Tsai-Shah uses Newton's method to

*Corresponding author

solve quadratic equations. When the initial values are close to the exact solution, the Tsai-Shah algorithm converges very quickly. However, the Newton method does not always guarantee convergence. Therefore the Tsai-Shah algorithm is more susceptible to initial estimation compared with the Pentland one (Wilson and Laxminarayan, 2006).

In this study, initial linear information is provided with the use of different spatial coefficients of numerical gradients of gray-level images. Four functionally generated images and two basic 3D reconstruction data sets are used to compare the Pentland and Tsai-Shah algorithms, and the proposed method. Performing the comparisons of synthetic images enables us to make more quantitative and reliable comparisons, as well as visual interpretations of the surfaces by the present ground truth data. Five different metrics are used to evaluate the performances of the algorithms: the mean absolute deviation (MAD), the mean squared error (MSE), Euclidean norm (L2), the peak signal-to-noise ratio (PSNR), and the structural similarity (SSIM) index. In addition, a comparison of the required time and CPU time is also included. The results show that the hybrid method does not need a high-value smoothing process and provides a satisfying 3D recovery for most synthetic images. Since defect detection is one of the most preferred areas of SFS algorithms, sample defective images from the NEU surface detect data set (Song and Yan, 2013) and the corresponding 3D reconstruction results are also included.

The rest of the paper is structured as follows. Section 2 is dedicated to a literature review of SFS methods. Section 3 presents the description of the mathematical and algorithmic structure of the proposed method. The synthetic images generated, 3D reconstruction data sets used, and the performance comparisons of the algorithms with five evaluation metrics are given in detail in Section 4. Finally, the results are discussed in Section 5.

2. Literature review

SFS is one of the techniques under the general heading of shape from X and it is an inverse problem from 2D images to 3D surfaces. In the beginning, Horn (1970) suggested calling this problem "shape-from-shading" and then, with the advent of new methods, based on the search for a solution, the algorithms were listed under different classes. An extensive pool of methods that suggested different approaches in algorithmic terms was formed. We begin our literature review by focusing on the general shape from shading and then discuss the most recent studies. Finally, we cite the papers that used linearization methods.

Zhang *et al.* (1999) discussed SFS algorithms under four classes and analyzed six well-known SFS algorithms comparatively over the mean and standard deviation errors of surface information, surface gradient (p, q), and CPU time parameters on synthetic and real images. Durou *et al.* (2008) made some updates on the study of Zhang *et al.* (1999) and divided SFS methods into three classes. They provided a short description of the methods selected from each class, and explained their basic assumptions and the mathematical approach. They also applied tests on efficiency and accuracy. For detailed information on the early development of the algorithms and the differences, we direct our readers to the survey papers by Zhang *et al.* (1999) and Durou *et al.* (2008).

SFS has inspired many researchers because of the abundance of its parameters such as the reflectivity coefficient (e.g., Barron and Malik, 2011), surface reflection (e.g., Wang *et al.*, 2020), constraints in minimization approaches (e.g., Frankot and Chellappa, 1988), solutions to partial differential equations (e.g., Quéau *et al.*, 2017), the illumination position (e.g., Zheng and Chellappa, 1991), projection (e.g., Breuß and Yarahmadi, 2020), and ambiguity (e.g., Abada and Aouat, 2015).

Combining the strengths of the SFS technique with other 3D reconstruction methods has also been found interesting. Sakarya and Erkmen (2003) proposed a new method for improving the photometric stereo method's efficiency by combining it with a local SFS method. Within each homogeneous field, SFS provided more detailed information. Maurer *et al.* (2018) combined a Lambertian SFS method with a stereo model and supplemented a second-order smoothness term to the resulting energy functional. They also expanded the resulting model to estimate depth, albedo, and illumination. Haefner *et al.* (2018) utilized heterogeneous depth and color data to simultaneously tackle the ill-posed depth super-resolution and shape-from-shading problems. Furthermore, optimization (e.g., Chen *et al.*, 2010) and machine learning-based methods (e.g., Abada and Aouat, 2016) have marked their place in the literature. Recently, studies on 3D reconstruction with deep neural networks have rapidly increased. Yang and Deng (2018) addressed the SFS problem by training deep networks with synthetic images. Unlike traditional methods combining deep learning and synthetic imaging, they proposed an approach that does not require any external data set to render synthetic images. Using a data-driven approach, Bednarik *et al.* (2018) developed a framework for reconstructing the 3D shape of a textureless, deformable surface from a single image and performing shape-from-shading.

In addition to the aforementioned solutions, depth maps obtained from 3D reconstruction have been used successfully in many different fields of study such as industry, medicine, surface inspection, biometry, and 3D modeling (e.g., Gallen *et al.*, 2015; Han and Zhu, 2005;

Kazmi *et al*., 2016).

The Pentland and Tsai-Shah algorithms have been present for 25 years and are still being preferred in the studies by virtue of their simple structures and ease of implementation. Fanany *et al.* (2002) presented a neural network learning scheme called the smooth projected polygon representation (SPPR) neural network for facial reconstruction. They considered depth maps as the partial face shape to be reconstructed and used the Tsai-Shah algorithm for obtaining depth maps. Fanany and Kumazawa (2004) proposed a new artificial neural network with the use of a backpropagation model for reconstructing a 3D shape, and their model relied on the depth maps obtained from the Tsai-Shah algorithm. Ghayourmanesh and Zahng (2007) developed an iterative technique which was a further improvement of Pentland's linear shape-from-shading method. The equations and geometric modeling were extracted for SAR imagery.

Cadavid and Abdel-Mottaleb (2008) used the algorithm developed by Tsai-Shah for modeling ear objects in each video frame. Kong (2008) proposed a method based on the Tsai-Shah algorithm and reconstructed the solder joint's surface in printed circuit boards. Pradhan *et al.* (2010) reconstructed the 3D shape of satellite imagery with the use of Pentland and the other two SFS methods. Barron and Malik (2011) considered the assumption that the albedo value should be uniform or known and introduced the problem of shape and albedo from shading (SAfS). The "single scale" model in their work was an improved version of Tsai-Shah's method. Wang *et al.* (2017) proposed a hybrid method for computing and re-animating the fluid surface, whose height field, which corresponds to each video frame, was estimated with the use of Tsai-Shah's shape-from-shading method. Salary *et al.* (2017) studied the online monitoring of functional electrical properties. They tested the Horn, Pentland, and Tsai-Shah methods to obtain the cross-sectional profiles from the online images. Yang *et al.* (2018) established a novel welding detection system for the arc welding robot and used Tsai-Shah for reconstructing the 3D shapes of the welding seam.

Zhang *et al.* (2018) obtained height and surface normals with the use of the Tsai-Shah algorithm on four types of fruit and proposed a method for correcting the adverse effects generated by the curvature of fruit objects in images acquired by the cameras. Medical imaging and processing require efficient tools such as shape approximation (Franchini *et al*., 2020), and the structures may need to be defined very precisely in 3D space throughout the surgeries (Ciecierski, 2020). Turan *et al.* (2017) suggested the use of the RGB-depth simultaneous localization and mapping (SLAM) method, which was developed especially for endoscopic capsule robots, and assisted the doctors in their more intuitive and accurate diagnoses. The Tsai-Shah method was used to generate depth estimates of successive endoscopic images. In addition, the Pentland or Tsai-Shah algorithms have been used as a comparison method in some papers (e.g., Kotan and Öz, 2017; Zhang *et al*., 1999; Durou *et al*., 2008; Salary *et al*., 2017)

3. Linearization-based 3D reconstruction

Linear approaches simplify the non-linear problem by linearizing the reflection equation. These methods usually linearize the reflection map in terms of the gradient or depth. The shape of the objects is obtained by solving the linearized models. The Pentland algorithm gives a non-iterative, closed-form solution with the use of the Fourier transform as

$$F_Z(w_x, w_y) = \frac{F_I}{-jw_x \cos\tau \sin\sigma - jw_y \sin\tau \sin\sigma}, \quad (1)$$

where F_I is the Fourier transform of the given image and F_Z is the Fourier transform of the depth map $Z(x, y)$. Then, with the help of the inverse Fourier transform, the depth $Z(x, y)$ is calculated via (Wilson and Laxminarayan, 2006)

$$Z(x, y) = F^{-1}\{F_Z(w_x, w_y)\}. \quad (2)$$

The problem lies in the linear approximation of the reflectance map, which causes difficulties when the non-linear terms are large. "Frequency doubling" occurs when the quadratic terms dominate, and thus the recovered shape would not be consistent with the illumination conditions (Zhang *et al*., 1999; Pentland, 1989).

In the approximation of brightness gradients p and q, Tsai and Shah employed the finite differences method first, and then used the surface equation $Z(x, y)$ to linearize the reflection diagram. Their method obtained depth at every point with the use of the Jacobi iterative scheme (Zhang *et al*., 1999; Durou *et al*., 2008; Ping-Sing and Shah, 1994; Hu *et al*., 2019).

The Tsai-Shah method applies discrete approaches to finite differences of p and q in order to linearize the reflection map in Z terms. If we consider the reflection function for a Lambert surface, Tsai-Shah linearizes the function $f = I - R = 0$ with Z terms around Z_{n-1} which is the surface reconstructed in the $(n-1)$-th iteration. For a Lambertian surface, the reflection function is

$$\begin{aligned} I(x, y) &= R(p, q) \\ &= \frac{p_s p + q q_s + 1}{\sqrt{p_s^2 + q_s^2 + 1}\sqrt{p^2 + q^2 + 1}}. \end{aligned} \quad (3)$$

Equation (3) can also be written as follows in terms of slant and tilt angles of the source:

$$R(p, q) = \frac{\cos\sigma + p\cos\tau \sin\sigma + q \sin\tau \sin\sigma}{\sqrt{1 + p^2 + q^2}}, \quad (4)$$

where $p = \partial Z/\partial x$ and $q = \partial Z/\partial y$, τ is the tilt, σ is the slant of the illuminant, p and q can be approximated discretely:

$$\begin{aligned} p &= \frac{\partial Z}{\partial x} = Z(x,y) - Z(x-1,y), \\ q &= \frac{\partial Z}{\partial y} = Z(x,y) - Z(x,y-1). \end{aligned} \tag{5}$$

The image irradiance equation can be rewritten as

$$\begin{aligned} 0 &= f[I(x,y), Z(x,y), Z(x-1,y), Z(x,y-1) \\ &= I(x,y) - R(Z(x,y) - Z(x-1,y), \\ &\quad Z(x,y) - Z(x,y-1)). \end{aligned} \tag{6}$$

If Eqn. (6) is extended by Taylor's expansion, the new linear system can be solved by the Jacobi iterative method. For a given initial approximation of Z^0, $Z(x,y)$ at the n-th iteration can be solved with the use of the previous estimates. Therefore, the Taylor series expansion of Eqn. (6) is simplified to the form

$$\begin{aligned} 0 &= f(Z(x,y)) \\ &\approx f(Z^{n-1}(x,y)) \\ &\quad + (Z(x,y) - Z^{n-1}(x,y))\frac{\mathrm{d}}{\mathrm{d}Z(x,y)} f(Z^{n-1}(x,y)). \end{aligned} \tag{7}$$

For more information on Jacobi expansions and simplification of equations, see the work of Tsai-Shah (1994). The depth map at the n-th iteration is directly solved by

$$Z^n(x,y) = Z^{n-1}(x,y) + \frac{-f(Z^{n-1}(x,y))}{\frac{\mathrm{d}}{\mathrm{d}Z(x,y)} f(Z^{n-1}(x,y))}, \tag{8}$$

where

$$\begin{aligned} \frac{\mathrm{d}f(Z^{n-1}(x,y)}{\mathrm{d}Z(x,y)} = -\Big(&\frac{(p_s + q_s)}{\sqrt{p^2+q^2+1}\sqrt{p_s^2+q_s^2+1}} \\ &- \frac{(p+q)(pp_s+qq_s+1)}{\sqrt{(p^2+q^2+1)^3}\sqrt{p_s^2+q_s^2+1}}\Big). \end{aligned} \tag{9}$$

Then, assuming $Z^0 = 0$, the resulting depth $Z(x,y)$ is obtained iteratively by Eqn. (8). The Tsai-Shah's iterative method sometimes deviates from the solution, but this can be avoided by interrupting the iteration after a fixed number of steps as suggested by Durou *et al*. (2008) or Ping-Sing and Shah (1994).

Inspired by Pentland's assumption, we provided the Fourier transform coefficients of the central differences (calculated on the gray-level image) as initial information for the Tsai-Shah algorithm and examined the results. Before the implementation of the SFS method, we first calculated the numerical gradients of the image matrix. The numerical gradient of a function is a method of estimating partial derivative values for each dimension with the use of known values of the function at specific points.

For a function $F(x,y)$ depending on two variables, $p_x = \partial F/\partial x$ characterizes differences in the horizontal (x) direction, and $q_y = \partial F/\partial y$ shows differences in the vertical (y) direction. They form a collection of vectors pointing the direction of the steepest ascent of $F(x,y)$. The following equation can be used to calculate the directional difference matrices of p_x and q_y for the $M \times N$ gray-level image:

$$\begin{aligned} p_x(x,y) &= \frac{I(x,y+1) - I(x,y-1)}{2}, \\ q_y(x,y) &= \frac{I(x+1,y) - I(x-1,y)}{2}. \end{aligned} \tag{10}$$

The single-sided differences can be used for values along the edges of the image. For example,

$$\begin{aligned} p_x(x,1) &= \frac{I(x,2) - I(x,1)}{2}, \\ p_x(x,N) &= \frac{I(x,N) - I(x,N-1)}{2}. \end{aligned} \tag{11}$$

Then, the Fourier transforms of the numerical gradients are calculated:

$$p_{fx} = \mathrm{FFT}(p_x), \quad q_{fy} = \mathrm{FFT}(q_y) \tag{12}$$

The surface Z^0 is initialized by means of Eqns. (3) and (6). Gray-level numerical gradients and the commenced depth are used for the calculation of Z^1 as indicated in Eqn. (8). Below, the algorithm steps of the proposed method are presented.

4. Experimental results

Sample implementations can be found at `https://github.com/muhammedkotan/hybridSFS`. Two different implementations were carried out for analyzing and comparing the method performances under different conditions. Four synthetic surfaces were generated functionally, and the depth maps and gray level images produced by the algorithms were compared. Then, various images were selected from two present datasets used for 3D reconstruction in the literature, and the gray-level images, provided by the algorithms, were compared. The MAD, MSE, L2, PSNR, and SSIM metrics were used to compare the errors and similarities of the produced surfaces and images with ground truth data. The performances of the Pentland, the Tsai-Shah and the developed method on the synthetic surfaces and for

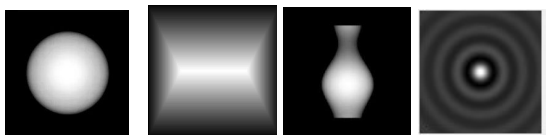
Algorithm 1. Hybrid algorithm.

Require: $I(x,y)$ = gray-level input image
l_s = spacing coefficient
f_c = filtering coefficient
$maxIter$ = max iteration number
$L(x,y,z)$ = illumination direction (can be estimated if it is not known)

1: $[M,N] = size(I)$;
2: $p_x(i,j) = (I(i,j+1) - I(i,j-1))/2$
3: $q_y(i,j) = (I(i+1,j) - I(i-1,j))/2$
4: $p_f x = fft(p_x), q_f y = fft(q_y)$
5: $p^0 = p_f x, q^0 = q_f y$
6: $D = \sqrt{x^2+y^2+z^2}, sx = x/D, sy = y/D, sz = z/D$
7: $Z^0(M,N) = f(I(x,y), p^0, q^0, p_s, q_s)$
8: **for** $n = 1 : maxIter$ **do**
9: $R = (1 + pp_s + qq_s) / \sqrt{(p^2+q^2+1)(p_s^2+q_s^2+1)}$
10: $f = I - R$
11: $dfz = -1.\{(p_s+q_s) / \sqrt{(p^2+q^2+1)(p_s^2+q_s^2+1)} - (p+q)(pp_s + qq_s + 1)/\sqrt{(p^2+q^2+1)^3(p_s^2+q_s^2+1)}\}$
12: $Z(i,j) = Z(i,j) + -f(Z(i,j))/dfz$
13: $p = Z(i,j) - Z(i-1,j), q = Z(i,j) - Z(i,j-1)$
14: **end for**
15: $Smooth\ Surface(Z(x,y), f_c)$
16: $Normalize\ Depth(Z(x,y), Z_{\max}, Z_{\min})$
17: **return** $Z(x,y)$

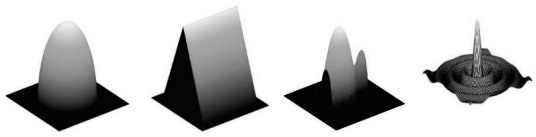
selected images from the current datasets were examined in detail.

The Pentland and Tsai-Shah algorithms are among the best quantitative methods on synthetic images (Zhang *et al*., 1999). In the comparison of the algorithms, we experientially determined the number of iterations as 25 at first. Convergence could also be achieved in a lower number of iterations. It should be noted that the Pentland algorithm does not have an iterative structure. Images were given to algorithms along with their backgrounds. No smoothing filter was applied for the Pentland method, and the results of Fourier transforms were directly visualized. While coding the Tsai-Shah algorithm and the proposed method, the results were visualized by applying a median filter to the obtained outcomes. Each output pixel contained the average value of the corresponding pixels in its 3×3 neighborhood. Some of the results could be improved or decreased with the use of different smoothing or filtering values, but only a 3×3 neighborhood filtering value was used for providing a pure comparison of the algorithms and for showing the effect of the proposed method without any filtering. Illumination directions directly affect the performance of the algorithms. Different depth maps were obtained in



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

Fig. 1. Generated synthetic images: Sphere (a), Tent (b), Vase (c) and Sinc (d).



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

Fig. 2. Ground-truth surfaces of the synthetic images: Sphere (a), Tent (b), Vase (c) and Sinc (d).

different directions and the performances of algorithms differed. Therefore, it was preferred to determine the unknown illumination values with the use of one of the estimation methods present in the literature studies. In the cases where the lighting position was unknown, the illumination directions were estimated with the use of a method similar to the one presented by Zheng and Chellappa (1991).The known directions provided by the authors were used for Data Set 1. The same directions were applied to all three methods. In the following parts of the study, we abbreviated the Pentland algorithm as P, the Tsai-Shah algorithm as TS, and the proposed algorithm as PM.

4.1. Synthetic image generation and test results. First, performance tests were performed on four synthetic images, some of which are commonly used in SFS methods and are functionally generable. Surface formulations and some reference studies using the selected synthetic images are given in Table 1.

The synthetic images generated according to formulations of Table 1 are shown in Fig. 1, and the ground truth surface figures of these images are given in Fig. 2.

Reconstructions, by the algorithms, of each synthetic image in Fig. 1 are given in Fig. 3. Furthermore, gray-scale images of these recovered depth maps are shown in Fig. 4. For the algorithms not to affect the performances of each other, each image was presented to the algorithms under the same conditions.

Because it is hard to obtain a unique solution in SFS algorithms, an algorithm that works very well for one image may not provide the same performance for another. We also observed that our Fourier-based approach smoothed the fluctuations in the occluding boundaries. These improvements are visible at the

Fig. 3. 3D reconstructions of selected methods for the synthetic images in Fig. 1.

Fig. 4. Regenerated gray-scale images of recovered depth maps shown in Fig. 3.

top of the sphere and on the edges of the tent in Fig. 3. The difference can be more clearly observed in Fig. 4. These surfaces appear to be more uniform with Fourier-initialized pre-processing, even with lower smoothing parameters and iteration numbers.

Experientially, we noticed that the performance improvement of the presented algorithm usually occurs after several iterations. In Figs. 5 and 6, the reconstructions of TS and PM in the 10th and 15th iterations are shown for the sphere and tent images. Since P was not an iterative solution, the results did not change. It should also be noted that increasing the number of iterations does not always improve the result. Algorithms

Table 1. Formulation of synthetic images.

Synthetic image	General formulation	Sample usage in SFS literature
Sphere (1501×1501)	$Z(x,y) = \sqrt{(r^2 - y^2 - x^2)}$, r : radius	Ping-Sing and Shah, 1994, Sakarya and Erkmen, 2003, Yuen *et al.*, 2007
Tent (256×256)	$Z(x,y) = \min\{-2\lvert x\rvert + 2\alpha, -\lvert y\rvert + \alpha\}$	Durou *et al.*, 2008, Tozza and Falcone, 2016, Quéau *et al.*, 2018
Vase (360×360)	Formulas proposed by Durou *et al.* (2008) and Ascher and Carter (1993): $Z(x,y) = \sqrt{f(x)^2 - y^2}$ $f(x) = \alpha_1 x^6 + \alpha_2 x^5 + \alpha_3 x^4 + \alpha_4 x^3 + \alpha_5 x^2 + \alpha_6 x + \alpha_7$	Zhang *et al.*, 1999, Durou *et al.*, 2008, Breuß *et al.*, 2010
Sinc Surface (81×81)	A classical sinc function expressed as $Z(x,y) = \sin(\sqrt{x^2+y^2})/\sqrt{(x^2+y^2)}$	We could not find any usage of this surface in our SFS literature reviews. It was chosen to create a more challenging surface.

may show deviations. In such cases, the number of iterations can be fixed.

In addition to the above qualitative comparisons, the results were evaluated quantitatively. Since ground truth depth matrices of synthetic images were available, five different methods were used to evaluate the recovered depth maps and generated gray-level images. The MAD, MSE and L2 metrics are error-based, so they were applied to the recovered and reference depth maps. The reference and recovered shapes are shown in Figs. 2 and 3, respectively. Here, the smaller the value, the lower the error rate. The error-based results are presented in Table 2. The PSNR and SSIM are considered as the second group of performance evaluators because they calculate the similarity on the gray-level image. The reference and obtained gray-level images are shown in Figs. 1 and 4, respectively. Here, the higher the value, the higher the similarity ratio. The similarity-based results are shown in Table 3. For each image in Tables 2 and 3, the best approximation values obtained from the P, TS, and PM are indicated in bold.

The summary table, created for each image according to the above evaluations, and the performance ratio of the presented method are shown in Table 4.

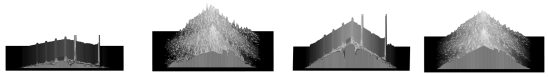
Table 4 proves the inconsistency of SFS algorithms. For the reconstruction performances of P and TS under different illumination directions, see the work of Zhang *et al.* (1999). Better performance values may be achieved with the use of different iterations and filtering coefficients. However, in order to maintain the stability of the parameters, we applied the same constant values for all synthetic images.

Since Fourier-based approaches are time-consuming, the required clock and CPU times are also provided in Table 5 (system: i7-9700K CPU@3.60 GHz [8 CPUs], 32 GB DDR4-3200 MHz). In order to measure the required time, we used the MATLAB `timeit` function, which calls the specified function multiple times and returns the median of the measurements. A handle to each SFS method is taken, and it returns the typical execution time in seconds. We also utilized the MATLAB `cputime` function for CPU time measurement. The `cputime` function measures the total CPU time, but it may be misleading if the measured function uses the processing cores equally. By virtue of parallel programming, the required clock and CPU time, in Table 5, may be shortened.

The reconstruction time of surfaces is the longest in PM due to its large size, gradient calculations, and Fourier transforms at the initialization. Although Zhang *et al.* (1999) mentioned in their examples that TS was one of the fastest methods, the iteration value was not indicated clearly. When a lower iteration value is specified, the duration of the TS and PM methods will be significantly reduced. The purpose of Table 5 is not to make an exact comparison of algorithms in terms of time but to give an idea of the execution for various surfaces of different sizes. Therefore, the time values were obtained with the use of 25 iterations of TS and PM.

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

Fig. 5. Convergence of the sphere shape: TS, 10th iteration (a), PM, 10th iteration (b), TS, 15th iteration (c), PM, 15th iteration (d).



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

Fig. 6. Convergence of the tent shape: TS, 10th iteration (a), PM, 10th iteration (b), TS, 15th iteration (c), PM, 15th iteration (d).

4.2. Comparisons with the use of 3D reconstruction data sets. Many studies have been conducted in the literature on the reconstruction of a 3D surface, and different data sets have emerged to assist researchers in their evaluations. These data sets include data such as various input images, image masks, depth maps, and illumination directions that provide ground truth data at the evaluation stages and more reliable results during the comparison stages. In the comparison of P, TS and PM, two different data sets present in the literature were used:

- Data Set 1: From Shading to Local Shape (Xiong *et al.*, 2014).
- Data Set 2: MIT Intrinsic Images (Grosse *et al.*, 2009).

The 3D depth maps and estimated gray-level images, generated by the methods for the selected gray-level images, were compared with the intonation images provided by the data sets, and the 3D reconstruction performance was measured. Measuring the reconstruction error by re-rendering the estimated normal map into a shading image and comparing that with the actually captured one is one of the preferred methods in 3D reconstruction studies (Xiong *et al.*, 2014). Information about the selected data sets, images, and illumination directions are given in Table 6. The images selected from both datasets are shown in Figs. 7 and 8, respectively. The directions known were used for Data Set 1. For the images selected from Data Set 2, the directions were estimated with the use of a similar method proposed by Zheng and Chellappa (1991). In addition, the Cat image was provided to the algorithms in two different ways, with black and white backgrounds. Thereby, the

Table 2. MAD, MSE and L2 results.

Synthetic image	MAD			MSE			L2		
	P	*TS*	*PM*	*P*	*TS*	*PM*	*P*	*TS*	*PM*
Sphere	**0.0011**	0.0625	0.0101	0.0626	0.0089	**0.0004**	366.93	135.03	**22.70**
Tent	0.0457	0.0512	**0.0301**	0.0379	0.0091	**0.0030**	46.62	24.34	**13.06**
Vase	0.0629	0.0500	**0.0158**	0.0344	0.0057	**0.0006**	64.64	26.31	**8.31**
Sinc	0.0033	0.0018	**0.0001**	0.0106	**0.0001**	0.0002	7.6791	0.8218	**0.6905**

Table 3. PSNR and SSIM results.

Synthetic image	PSNR			SSIM		
	P	*TS*	*PM*	*P*	*TS*	*PM*
Sphere	12.0322	20.4860	**33.2230**	0.5909	0.9397	**0.9864**
Tent	14.2093	20.3976	**25.2539**	0.5862	0.8933	**0.9359**
Vase	14.6315	22.4488	**32.0632**	0.7609	0.9688	**0.9802**
Sinc	19.7447	**37.3641**	35.6676	0.7689	**0.9724**	0.9719

Table 4. Tests summary.

Synthetic image	MAD	MSE	L2	PSNR	SSIM	PM-Score
Sphere	P	**PM**	**PM**	**PM**	**PM**	**4/5**
Tent	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Vase	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Sinc	**PM**	TS	**PM**	TS	TS	**2/5**
	75.0%	**75.0%**	**100.0%**	**75.0%**	**75.0%**	**80.0%**

Table 5. Required clock and CPU time.

Synthetic image	Clock time (sec.)			CPU time (sec.)		
	P	*TS(it:25)*	*PM(it:25)*	*P*	*TS(it:25)*	*PM(it:25)*
Sphere (1501×1501)	0.0723	1.7151	3.5450	0.468	10.2969	19.4375
Tent (256×256)	0.0017	0.0351	0.0972	0.031	0.2344	0.7656
Vase (360×360)	0.0027	0.0431	0.1983	0.015	0.1406	1.1719
Sinc (81×81)	0.0003	0.0057	0.0104	≈0.0001	≈0.1	≈0.1

performances of the algorithms were evaluated under different backgrounds and situations.

Depth maps created by algorithms for input images in Data Sets 1 and 2 are shown in Figs. 9. and 10, respectively. Gray-level images corresponding to the estimated depth maps are shown in Figs. 11 and 12. Gray-level images showing the differences of the PM and TS results are also presented in Fig. 13.

The MAD, MSE and L2 values, relevant to the error rates between the images provided by the data sets and the images produced by the algorithms, are presented in Table 7. The PSNR and SSIM values, indicating similarity rates, are given in Table 8. Moreover, all evaluation results and the performance of the proposed method are summarized in Table 9.

The similarity between the MSE and PSNR results is due to the relationship between the MSE-PSNR formulations and proves the consistency of the results. Another advantage of applying PM is that the depth information can be obtained without applying any filtering processes. TS usually requires an appropriate smoothing and filtering step in order to converge to the 3D shape. The depth information of the Cat image, obtained without the applying any filters after 25 iterations, is visualized in Fig. 13.

4.3. 3D reconstruction on defect images. The fact that the details on the object can be obtained from a single image under suitable scene conditions has made SFS algorithms one of the preferred methods in defect

Table 6. Selected datasets and image properties.

Dataset	Input	Illumination	Background
From Shading	Cat	[0,1949; 0,2419; 0.9270]	Black
to Local	Cat	[0,1949; 0,2419; 0.9270]	White
Image	Turtle	[0,1590; 0,2789; 0.9431]	Black
MIT Intrinsic*	Dino	[0,8551; 0,4618; 0.2353]	Black
	Frog	[0,8644; 0,3968; 0,3086]	Black
	Pear	[0,8639; 0,3909; 0,3174]	Black

*Estimated values were used for the illumination directions.

Table 7. MAD, MSE and L2 results.

Image	MAD			MSE			L2		
	P	*TS*	*PM*	*P*	*TS*	*PM*	*P*	*TS*	*PM*
Cat (black)	0.053	0.009	**0.004**	0.104	0.0003	**0.0002**	161.06	7.60	**5.55**
Cat (white)	0.041	**0.021**	0.034	0.104	0.325	**0.004**	154.38	274.62	**32.09**
Turtle	0.079	0.038	**0.00005**	0.059	0.003	**0.0002**	113.27	34.91	**5.93**
Dino	0.018	0.006	**0.001**	0.047	0.0001	$\approx$ **0.00001**	86.13	3.58	**1.37**
Frog	0.025	0.005	**0.0020**	0.12106	0.0010	**0.0009**	121.63	2.50	**2.16**
Pear	0.024	0.004	**0.0016**	0.0594	0.0001	$\approx$ **0.00001**	64.65	**1.90**	2.07

Table 8. PSNR and SSIM results.

Image	PSNR			SSIM		
	P	*TS*	*PM*	*P*	*TS*	*PM*
Cat (black)	9.81	34.94	**35.37**	0.386	0.949	**0.952**
Cat (white)	9.801	4.871	**23.062**	0.661	0.535	**0.886**
Turtle	12.259	24.156	**35.805**	0.511	0.929	**0.940**
Dino	13.272	37.709	**42.735**	0.115	0.982	**0.991**
Frog	9.171	29.991	**30.345**	0.1065	0.921	**0.928**
Pear	12.259	38.845	**40.348**	0.280	0.963	**0.984**

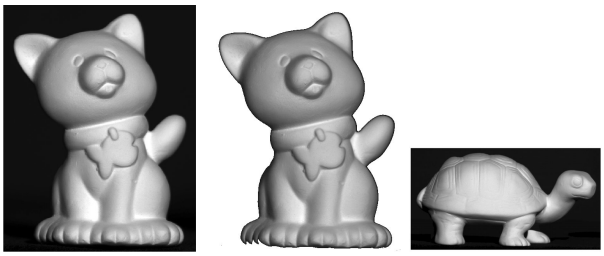
Table 9. Tests summary.

Image	MAD	MSE	L2	PSNR	SSIM	PM-Score
Cat (black)	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Cat (white)	TS	**PM**	**PM**	**PM**	**PM**	**4/5**
Turtle	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Dino	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Frog	**PM**	**PM**	**PM**	**PM**	**PM**	**5/5**
Pear	**PM**	**PM**	TS	**PM**	**PM**	**4/5**
	83.3%	**100.0%**	**83.3%**	**100.0%**	**100.0%**	**93.3%**

detection studies. In addition to literature studies, some companies are also developing their own patented SFS algorithms. To show the effect of the developed algorithm on the defect images, we presented the 3D examples of the images obtained from the NEU Surface Detect Dataset (Song and Yan, 2013) in Fig. 14.

5. Conclusion

A hybrid shape-from-shading method for linearization-based 3D reconstruction of a single image was presented. We studied whether Fourier-based central differences of the image would improve the performance of the reconstruction and whether there is good preliminary information for the estimation of shape. The presented method was compared with the Tsai-Shah and Pentland methods, which are the basis of this study, with the use of four functionally generated images, and two 3D reconstruction data sets being present. Recovered depth maps and generated gray-level images of test images were evaluated with five commonly used metrics: mean absolute deviation (MAD), mean squared error



(a) (b) (c)

Fig. 7. Images selected from Data Set 1: Cat (black background) (a), Cat (white background) (b), Turtle (c).

(a) (b) (c)

Fig. 8. Images selected from Data Set 2: Dino (a), Frog (b), Pear (c).

(MSE), Euclidean norm (L2), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM). The first three methods were applied to calculate the error between recovered and ground truth data. The latter two methods were used to calculate the similarity ratio of the obtained gray-level images to the original images.

The response of the algorithms against different background colors was also studied. The results are listed in detail in the tables. In addition, the required time and CPU time were included to guide the researchers.

The surface estimates that the developed method can produce for six different types of defects generally encountered on metallic surfaces were also added. After a few iterations, our experiments showed that the proposed hybrid SFS algorithm, which was initialized by the Fourier coefficients of the gradient, can better converge the 3D shape and the gray-level image, does not require a higher filtering process, is reliable regarding sudden changes on the surfaces, and is usable for different backgrounds.

Acknowledgment

This study has been supported by the Sakarya University Scientific Research Projects Coordination Unit (project no. 2020-7-25-2).

References

Abada, L. and Aouat, S. (2015). Tabu search to solve the shape from shading ambiguity, *International Journal on Artificial Intelligence Tools* **24**(05): 1550035.

Abada, L. and Aouat, S. (2016). A machine learning approach for shape from shading, *arXiv* 1607.03284.

Ascher, U.M. and Carter, P.M. (1993). A multigrid method for shape from shading, *SIAM Journal on Numerical Analysis* **30**(1): 102–115.

Barron, J.T. and Malik, J. (2011). High-frequency shape and albedo from shading using natural image statistics, *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 11), Colorado Springs, USA*, pp. 2521–2528.

Bednarık, J., Fua, P. and Salzmann, M. (2018). Learning shape-from-shading for deformable surfaces, *arXiv* 1803.08908.

Breuß, M., Cristiani, E., Durou, J.-D., Falcone, M. and Vogel, O. (2010). Numerical algorithms for perspective shape from shading, *Kybernetika* **46**(2): 207–225.

Breuß, M. and Yarahmadi, A.M. (2020). Perspective shape from shading, *in* J.D. Durou *et al.* (Eds), *Advances in Photometric 3D-Reconstruction*, Springer, Cham, pp. 31–72.

Cadavid, S. and Abdel-Mottaleb, M. (2008). 3-D ear modeling and recognition from video sequences using shape from shading, *IEEE Transactions on Information Forensics and Security* **3**(4): 709–718.

Chen, Z.-M., Cao, J.-Z. and Huang, J.-Q. (2010). A novel 3D reconstruction algorithm based on hybrid immune particle swarm optimization, *Proceedings of the 29th Chinese Control Conference, Beijing, China*, pp. 5228–5231.

Ciaccio, E.J., Bhagat, G., Lewis, S.K. and Green, P.H. (2017). Use of shape-from-shading to characterize mucosal topography in celiac disease videocapsule images, *World Journal of Gastrointestinal Endoscopy* **9**(7): 310.

Ciecierski, K.A. (2020). Mathematical methods of signal analysis applied in medical diagnostic, *International Journal of Applied Mathematics and Computer Science* **30**(3): 449–462, DOI: 10.34768/amcs-2020-0033.

Durou, J.-D., Falcone, M. and Sagona, M. (2008). Numerical methods for shape-from-shading: A new survey with benchmarks, *Computer Vision and Image Understanding* **109**(1): 22–43.

Fanany, M.I. and Kumazawa, I. (2004). A neural network for recovering 3D shape from erroneous and few depth maps of shaded images, *Pattern Recognition Letters* **25**(4): 377–389.

Fanany, M.I., Ohno, M. and Kumazawa, I. (2002). A scheme for reconstructing face from shading using smooth projected polygon representation NN, *International Conference on Image Processing, Rochester, USA*, Vol. 2, pp. II–II.

Franchini, S., Gentile, A., Vassallo, G. and Vitabile, S. (2020). Implementation and evaluation of medical imaging techniques based on conformal geometric algebra, *International Journal of Applied Mathematics and Computer Science* **30**(3): 415–433, DOI: 10.34768/amcs-2020-0031.

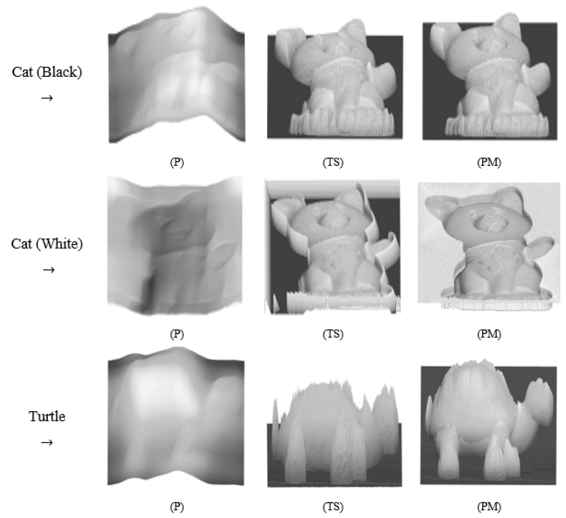


Fig. 9. Created depth maps for Data Set 1.

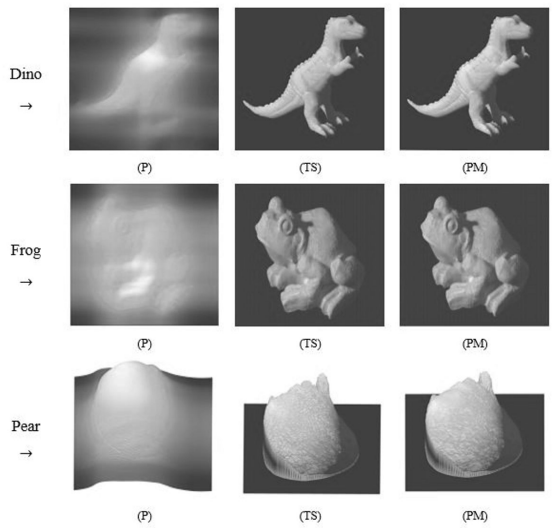


Fig. 10. Created depth maps for Data Set 2.

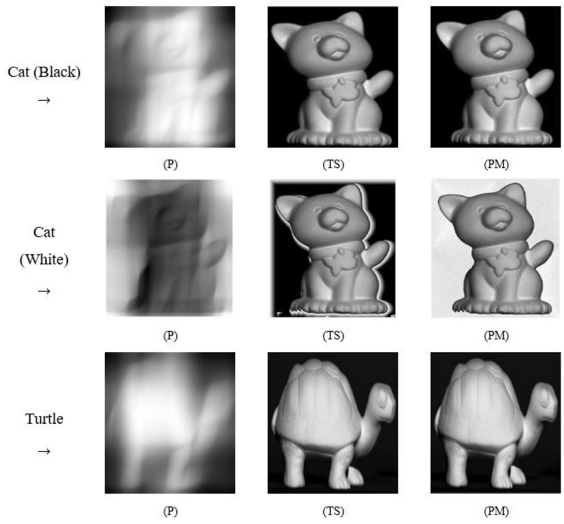


Fig. 11. Gray-level images produced by algorithms for Data Set 1.

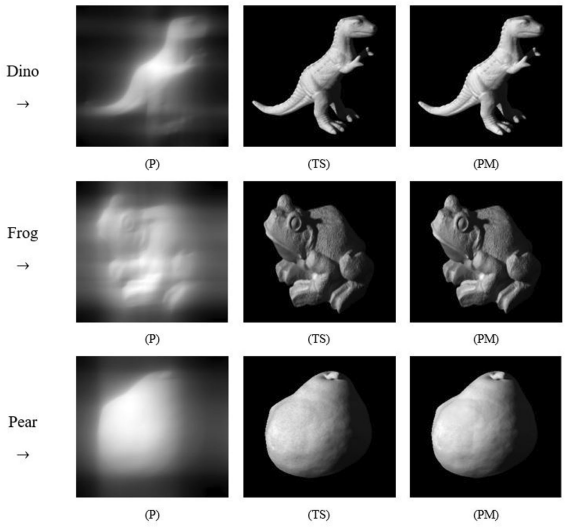


Fig. 12. Gray-level images produced by algorithms for Data Set 2.

Frankot, R.T. and Chellappa, R. (1988). A method for enforcing integrability in shape from shading algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**(4): 439–451.

Gallen, R., Eastop, D., Bozia, E. and Barmpoutis, A. (2015). Digital imaging: The application of shape-from-shading to lace, seals and metal objects, *Journal of the Institute of Conservation* **38**(1): 41–53.

Ghayourmanesh, S. and Zahng, Y. (2007). Shape from shading of SAR imagery in Fourier space, *2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain*, pp. 835–837.

Grosse, R., Johnson, M.K., Adelson, E.H. and Freeman, W.T. (2009). Ground truth dataset and baseline evaluations for intrinsic image algorithms, *2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan*, pp. 2335–2342.

Haefner, B., Quéau, Y., Möllenhoff, T. and Cremers, D. (2018). Fight ill-posedness with ill-posedness: Single-shot variational depth super-resolution from shading, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA*, pp. 164–174.

Han, F. and Zhu, S.-C. (2005). Cloth representation by shape from shading with shading primitives, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, USA*, Vol. 1, pp. 1203–1210.

Horn, B.K. (1970). *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One*

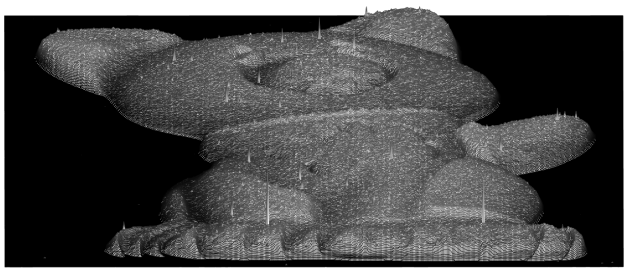


Fig. 13. Depth image obtained without applying any filters to the calculated Cat data.

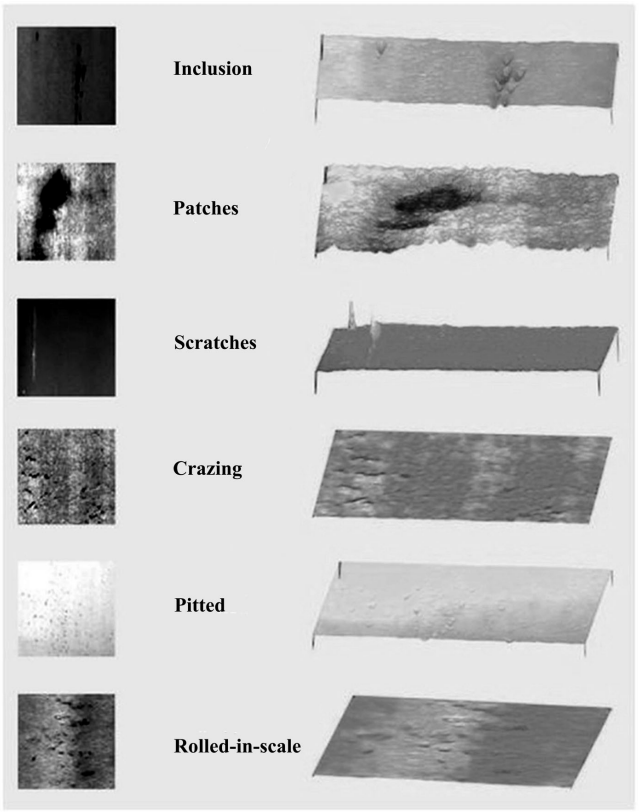


Fig. 14. Detecting defective surfaces by using the hybrid algorithm (Inclusion, Patch, Scratch, Crazing, Pitted, Rolled-in-scale).

View, PhD thesis, Massachusetts Institute of Technology, Cambridge.

Hu, Q., Shou, Z., He, L., Cai, Q., Qu, M. and Zhang, Y. (2019). Three-dimensional characterization method of pile-rock interface roughness based on fractal geometry, *Arabian Journal of Geosciences* **12**(18): 599.

Kazmi, I.K., You, L. and Zhang, J.J. (2016). A hybrid approach for character modeling using geometric primitives and shape-from-shading algorithm, *Journal of Computational Design and Engineering* **3**(2): 121–131.

Kemelmacher-Shlizerman, I. and Basri, R. (2010). 3D face reconstruction from a single image using a single reference face shape, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(2): 394–405.

Kong, F.-H. (2008). A new method of inspection based on shape from shading, *2008 Congress on Image and Signal Processing, Sanya, China*, Vol. 2, pp. 291–294.

Kotan, M. and Öz, C. (2017). Surface inspection system for industrial components based on shape from shading minimization approach, *Optical Engineering* **56**(12): 123105.

Lu, J., Zhang, S., Shi, L., Hou, D. and Wang, X. (2018). Automatic correction of the adverse effects of light on fruit surfaces using the shape-from-shading method, *Czech Journal of Food Sciences* **36**(1): 37–43.

Maurer, D., Ju, Y.C., Breuß, M. and Bruhn, A. (2018). Combining shape from shading and stereo: A joint variational method for estimating depth, illumination and albedo, *International Journal of Computer Vision* **126**(12): 1342–1366.

Pentland, A. (1989). Shape information from shading: A theory about human perception, *Spatial Vision* **4**(2–3): 165–182.

Ping-Sing, T. and Shah, M. (1994). Shape from shading using linear approximation, *Image and Vision Computing* **12**(8): 487–498.

Pradhan, R., Ghose, M. and Jeyaram, A. (2010). Extraction of depth elevation model (DEM) from high resolution satellite imagery using shape from shading approach, *International Journal of Computer Applications* **7**(12): 40–46.

Quéau, Y., Durou, J.-D. and Aujol, J.-F. (2018). Variational methods for normal integration, *Journal of Mathematical Imaging and Vision* **60**(4): 609–632.

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

Sakarya, U. and Erkmen, İ. (2003). An improved method of photometric stereo using local shape from shading, *Image and Vision Computing* **21**(11): 941–954.

Salary, R.R., Lombardi, J.P., Rao, P.K. and Poliks, M.D. (2017). Online monitoring of functional electrical properties in aerosol jet printing additive manufacturing process using shape-from-shading image analysis, *Journal of Manufacturing Science and Engineering* **139**(10): 101010.

Song, K. and Yan, Y. (2013). A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects, *Applied Surface Science* **285**(B): 858–864.

Tozza, S. and Falcone, M. (2016). Analysis and approximation of some shape-from-shading models for non-Lambertian surfaces, *Journal of Mathematical Imaging and Vision* **55**(2): 153–178.

Turan, M., Almalioglu, Y., Araujo, H., Konukoglu, E. and Sitti, M. (2017). A non-rigid map fusion-based direct SLAM method for endoscopic capsule robots, *International Journal of Intelligent Robotics and Applications* **1**(4): 399–409.

Wang, C., Wang, C., Qin, H. and Zhang, T.-y. (2017). Video-based fluid reconstruction and its coupling with SPH simulation, *The Visual Computer* **33**(9): 1211–1224.

Wang, G., Zhang, X. and Cheng, J. (2020). A unified shape-from-shading approach for 3D surface reconstruction using fast eikonal solvers, *International Journal of Optics* **2020**(8): 1–12.

Wilson, D. and Laxminarayan, S. (2006). *Handbook of Biomedical Image Analysis: Volume 1: Segmentation Models*, Springer Science & Business Media, New York.

Worthington, P.L. and Hancock, E.R. (2001). Surface topography using shape-from-shading, *Pattern Recognition* **34**(4): 823–840.

Wu, B., Li, F., Hu, H., Zhao, Y., Wang, Y., Xiao, P., Li, Y., Liu, W.C., Chen, L., Ge, X., Yang, M., Xu, Y., Ye, Q., Wu, X. and Zhang, H. (2020). Topographic and geomorphological mapping and analysis of the Chang'e-4 landing site on the far side of the moon, *Photogrammetric Engineering & Remote Sensing* **86**(4): 247–258.

Xiong, Y., Chakrabarti, A., Basri, R., Gortler, S.J., Jacobs, D.W. and Zickler, T. (2014). From shading to local shape, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **37**(1): 67–79.

Yamany, S.M. and Farag, A.A. (1998). A system for human jaw modeling using intra-oral images, *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, China*, Vol. 2, pp. 563–566.

Yang, D. and Deng, J. (2018). Shape from shading through shape evolution, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA*, pp. 3781–3790.

Yang, L., Li, E., Long, T., Fan, J., Mao, Y., Fang, Z. and Liang, Z. (2018). A welding quality detection method for arc welding robot based on 3D reconstruction with SFS algorithm, *International Journal of Advanced Manufacturing Technology* **94**(1–4): 1209–1220.

Yuen, S.Y., Tsui, Y.Y. and Chow, C.K. (2007). A fast marching formulation of perspective shape from shading under frontal illumination, *Pattern Recognition Letters* **28**(7): 806–824.

Zhang, R., Tsai, P.-S., Cryer, J.E. and Shah, M. (1999). Shape-from-shading: A survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(8): 690–706.

Zheng, Q. and Chellappa, R. (1991). Estimation of illuminant direction, albedo, and shape from shading, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(7): 680–702.

Muhammed Kotan received his MS and PhD degrees in computer and information engineering from Sakarya University, Turkey, in 2014 and 2020, respectively. He works at the Information Systems Engineering Department there. His present research interests include image processing, computer vision, pattern recognition and virtual reality.

Cemil Öz received his MS degree in electronics and computer education from Marmara University, Istanbul, in 1993. During his studies, he worked as a lecturer at Istanbul Technical University. In 1998, he completed his PhD studies in electronics engineering at Sakarya University. He then worked as a research fellow at the University of Missouri at Rolla, USA. He has since been working as a professor at the Computer and Information Sciences Faculty, Department of Computer Engineering, Sakarya University. His research interests include robotics, vision, artificial intelligence, virtual reality, and pattern recognition.

Abdulgani Kahraman received his MS degree in computer engineering and science from the University of Louisville, and currently he is pursuing a PhD there. His present research interests include data mining, machine learning, green energy, and image processing.

Received: 5 November 2020
Revised: 31 March 2021
Re-revised: 7 June 2021
Accepted: 8 June 2021