Integrating Shape from Shading and Shape from Stereo for Variable Reflectance Surface Reconstruction from SEM Images

Reinhard Danzl[1] *and Stefan Scherer*[2]
[1]Institute for Computer Graphics and Vision,
Graz University of Technology
[2]Alicona Imaging GmbH, A-8074 Grambach, Graz
email: danzl@icg.tu-graz.ac.at,stefan.scherer@alicona.com

Abstract:
An algorithm is presented and evaluated that combines the methods shape from stereo and shape from shading in order to reconstruct the shape and reflectance properties of a surface from two stereoscopic images. Whereas elevation models obtained by shape from stereo are usually rather sparse and shape from shading is too under-constrained for complete reconstruction, the combination leads to dense elevation models that are consistent with the stereoscopic images. We concentrate on the reconstruction from images of the scanning electron microscope which have strongly varying reflectance properties. In contrast to previous approaches which assume that the reflectance properties of the scene are known or can be represented using simple mathematical models, the proposed algorithm provides an estimate by fitting a polynomial of degree four. It starts with an initial elevation model obtained by shape from stereo and then often iterates between the refinement of the reflectance map and the refinement of the surface. The algorithm is evaluated on synthetic images with known ground truths as well as on real stereoscopic images taken by a scanning electron microscope.

1 Introduction

We present an algorithm that combines the methods shape from shading and shape from stereo in order to reconstruct dense elevation models from two stereoscopic images. Since shape from shading works best in areas with homogenous texture and shape from stereo in regions where a lot of features are present this combination has recently attracted many researchers [9] [12] [11].
Whereas most work concentrates on images taken under normal light conditions we use images from the secondary electron detector of a scanning electron microscope (SEM) which have been rarely used in the context of shape from shading. Although the applicability of SEM images to photometric methods has soon been realized [6] [7] there have only been few further proposals

on the use of shape from shading [1] [8] or photometric stereo [13].
Since the reflectance properties in the SEM vary strongly from image to image the main problem lies in the search for a reflectance map that contains information on how object patches with a specific surface normal are mapped onto grey levels in the image. In contrast to previous approaches that assume that this function is known [1] [6] or can be represented by simple mathematical models [8] we provide an estimate by fitting a polynomial of degree four.
Whereas other researchers often decouple stereo and shading information during optimization of the surface [4] [2] we use the clues simultaneously by minimizing a global energy function that contains constraints on shading, on previously known height values and on the smoothness of the surface. For proper convergence the use of a multi-grid scheme on a pseudo-hexagonal grid has been implemented.
Section 2 contains a description of our algorithm, in Section 3 we provide experiments on synthetic and real data and conclude in Section 4 with proposals for further research.

2 An Iterative Refinement Scheme

We start our description with an overview of our iterative refinement scheme (Section 2.1) which is based on the minimization of global energy functions. The function used for the approximation of the shape is described in Section 2.2, those for refining the reflectance properties in Section 2.3. Section 2.4 contains details concerning our optimization algorithm as well as the grid design.

2.1 Overall Concept

The main idea of our algorithm is to frequently iterate between the refinement of the reflectance map and the refinement of the surface (Fig. 1). We start with two stereoscopic images and an initial digital elevation model (DEM) obtained by an area based stereo matching method based on adaptive windows [14]. The DEM is first sampled to a regular grid of low resolution since our algorithm works hierarchically on grids with different width to ensure correct reconstruction of low and high frequency components. Using the current DEM and the stereo images an initial estimate of the reflectance map is calculated. Next the DEM is refined using this reflectance map and the stereo images. In order to prevent the algorithm from converging to a wrong wrinkled surface this mutual refinement is repeated several times each time reducing the value of a smoothness weight in the energy function. When the smoothness weight is approximately zero the DEM and the reflectance map have been fully refined at the lowest resolution and the whole procedure is repeated on a denser grid. The algorithm ends when the DEM at the final resolution of one grid-point per pixel has been calculated.

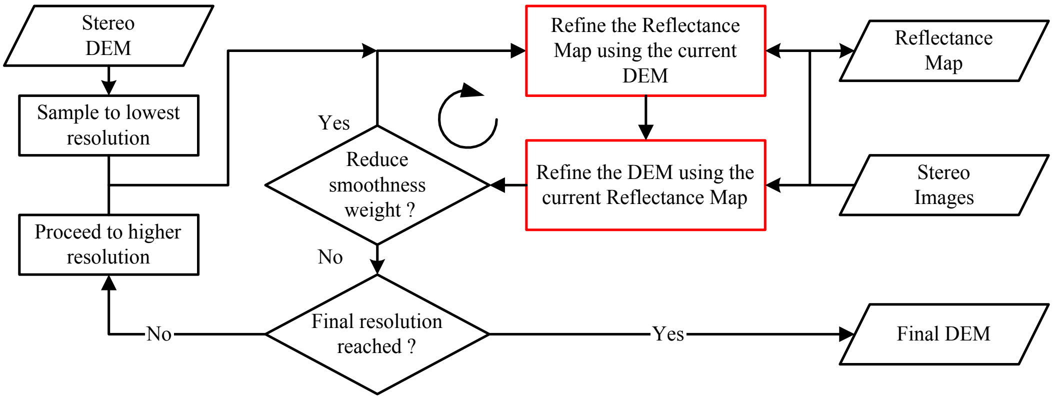

Figure 1: An iterative algorithm for the refinement of the reflectance map and the refinement of the DEM.

2.2 Approximating the Shape

The energy function for the optimization of the surface is the weighted sum of three terms

$$C = w_B C_B + w_S C_S + w_{ST} C_{ST} \tag{1}$$

where C_B is a brightness term, C_S is a smoothness term and C_{ST} is a stereo term. w_B, w_S and w_{ST} are the corresponding weights and are chosen as proposed in Section 2.4. The first term reduces the shading error and has the form

$$C_B = \sum_{(x,y)} (I(x,y) - R(\alpha(x,y)))^2 \tag{2}$$

where $R(\alpha(x,y))$ is the normalized reflectance map, $I(x,y)$ is the grey value of the image and $\alpha(x,y)$ is the angle between the direction of the light source and the normal vector of the surface at position (x,y). The second term contains the second derivatives of the height values $z(x,y)$

$$C_S = \sum_{(x,y)} (z_{xx}(x,y)^2 + z_{xy}(x,y)^2 + z_{yx}(x,y)^2 + z_{yy}(x,y)^2) \tag{3}$$

and is used for proper convergence especially when the reflectance map is not known a priori. The stereo term C_{ST} is of the form

$$C_{ST} = \sum_{k} (\tilde{z}_k - d_k)^2 \tag{4}$$

where d_k represents the height values of the DEM obtained by shape from stereo and $\tilde{z}$ is the corresponding height value of the current DEM. There are many other ways to include stereo information which happened to be less robust than the proposed approach. One method that adds the low-frequency component of the shading model and the high-frequency model of the stereo model [2] tends to smoothen sharp edges. Another approach by [5] that uses stereo correlation terms was not applicable since the orientation was not known exactly enough for our SEM images.

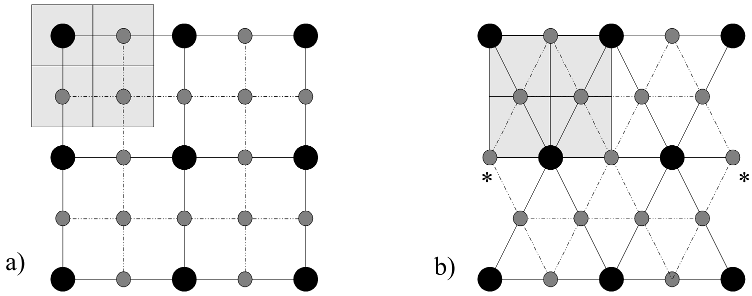

Figure 2: a) A conventional rectangular grid b) The non-pixel-centered hexagonal grid

2.3 Approximating the Reflectance Map

All studied publications on the use of shape from shading for SEM images model the reflectance map as a function that is only dependent on the angle α between the surface normal and the direction of the light source. For SEM images the light source corresponds to the electron beam of the microscope which we have taken as $(0, 0, 1)$ as it is usually assumed for this application. The special image formation process in the SEM produces images where flat regions of the object appear dark whereas steep edges are very bright which stays in contrast to images obtained using conventional point light sources. Although there have been some proposals to model the reflectance map of the SEM as a function proportional to $sec\ \alpha$ [8] this form was not sufficient for several of our images. Therefore the more general form

$$R(\alpha) = a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 \tag{5}$$

was chosen with $a, b, c, d, e \in \Re$ and $R(\alpha)$ being normalized between 0 and 255. The energy function used is

$$C = \sum_{(x,y)} (I(x,y) - R(\alpha(x,y)))^2 \tag{6}$$

which is equal to the brightness term C_B. Some additional constraints have to be put on the reflectance map to prevent the optimization procedure from converging to wrong solutions. This is done by adding virtual points to the above function which are very likely to lie on the reflectance map. These points include $R(0)$ to which the 5% quantile of grey levels is assigned and the point $R(255)$ whose corresponding grey level is obtained empirically.

2.4 The Optimization Procedure

The optimization method used is conjugate gradient descent which is well-suited for optimization problems in a large multi-dimensional space. In our case the number of dimensions is equal to the number of grid-points and soon reaches a number > 50000 (e.g. for images of the size 256×256). The line search algorithm we used as part of the conjugate gradient method is the backtracking method by Dennis and Schnabel [3].

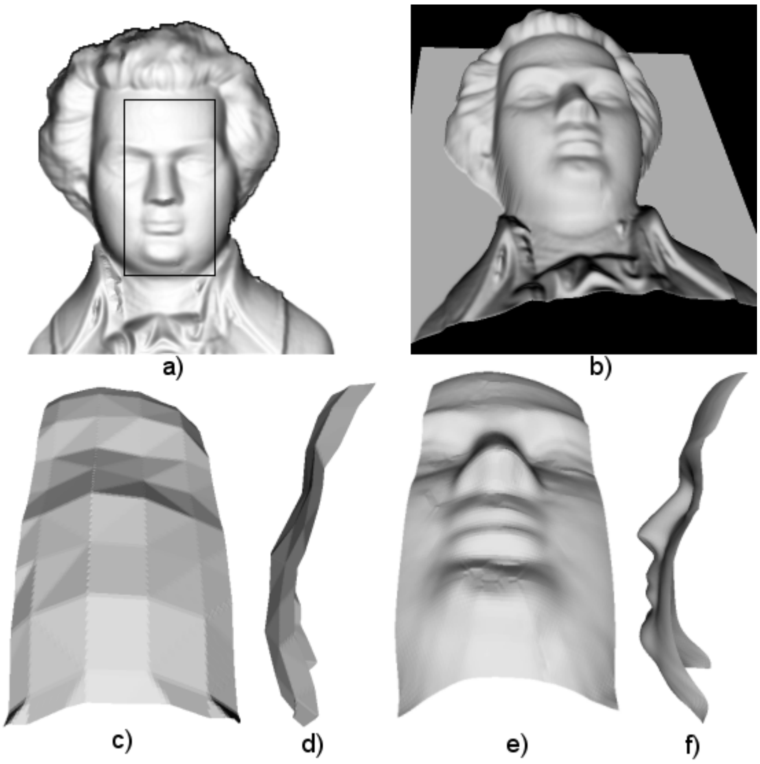

Figure 3: Experiment on synthetic data: a) input image b) ground truth c) d) initial DEM (sampling width 16 pixel) e) f) refined DEM

Unfortunately the best weights are strongly dependent on the brightness in the image, the shape of the object, the number of stereo-points and the grid-distance. A good method to normalize the weights and make them mainly independent of external factors is to set the weights in relation to the gradient of the corresponding term in the energy function [5]. While the ratio of the brightness and the stereo weight remains fixed, the smoothness weight is continuously decreasing during the iteration as described in Section 2.1.

Many previous approaches use a rectangular grid that approximates the surface normal using finite differences of neighboring grid-points and therefore tend to smoothen sharp edges of a specimen. Our approach uses a hexagonal grid similar to [5] which allows error-free calculation of the surface normals. When the grid points of a hexagonal grid are chosen to be pixel-centered the calculation of the grey value of each surface patch involves at least four pixels and results in a smoothing of the surface. We use a layout that is not pixel-centered (Fig. 2b) which does not only allow to calculate the surface normals analytically but the computation of the grey levels of each patch only involves one or at most two pixels. Additionally Fig. 2b shows how a grid of low resolution (thick dots) is up-sampled to a denser grid.

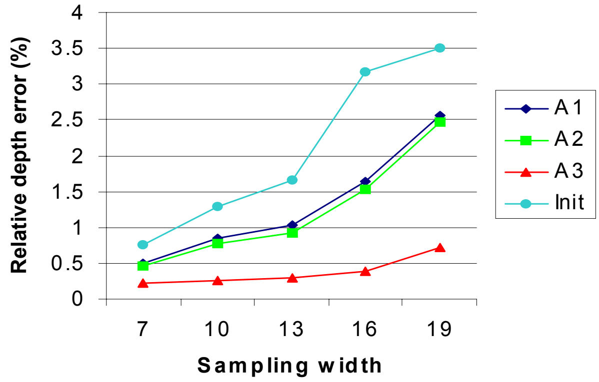

Figure 4: Relative depth error for three algorithms and different sampling widths of the initial DEM: Init: initial DEM, A1: algorithm by Leclerc and Bobick, A2: algorithm by Frankot and Chellappa, A3: new approach

3 Experiments

The first experiment was done using a synthetic image (Fig. 3a) generated from a Lambertian reflectance map and a known ground truth (Fig. 3b). Instead of using shape from stereo, an initial DEM (Fig. 3c,d) was generated by sub-sampling the ground truth with sampling widths between 7 and 19 pixels and the reflectance map was provided to the algorithm. The refined DEM in Fig. 3e,f shows the good performance of the system.

In Fig. 4 our algorithm is compared to two other algorithms. The first is based on Frankot and Chellappa's algorithm [4] that uses low-frequency enforcement for the inclusion of the stereo information. The second is based on Leclerc and Bobick's algorithm [10] and mainly differs from the proposed method by the use of a rectangular grid. The figure shows the relative depth error of the three algorithms for different sampling widths of the initial DEM and shows that the proposed algorithm is superior to the other two especially when the initial DEM contains few points.

The second experiment has been performed on two stereoscopic images from the SEM (left image shown in Fig. 5b) which have been obtained by tilting the specimen about an angle of 7 degree. The pixel size is $0.2\mu m$ and the number of points that have been found by shape from stereo is 1745. Fig. 5c shows the wrong initial reflectance map where the x-axis corresponds to the angle α between the light source and the surface normal and y-axis to the grey values from 0 to 255. Additionally a histogram is shown that contains points for each pixel in the input image where the y-coordinate is the grey value of the pixel and the x-coordinate is the angle α of the initial DEM at these points. Fig. 5a shows a synthetic image that has been generated using the initial reflectance map and the initial DEM (Fig. 5e). The refined reflectance map (Fig. 5d) approximates the pixels in the image and the corresponding surface elements much better and the final DEM in Fig. 5f provides a good impression of the surface. The synthetic image generated by the final DEM and the final reflectance map can visually not be distinguished from the left input image.

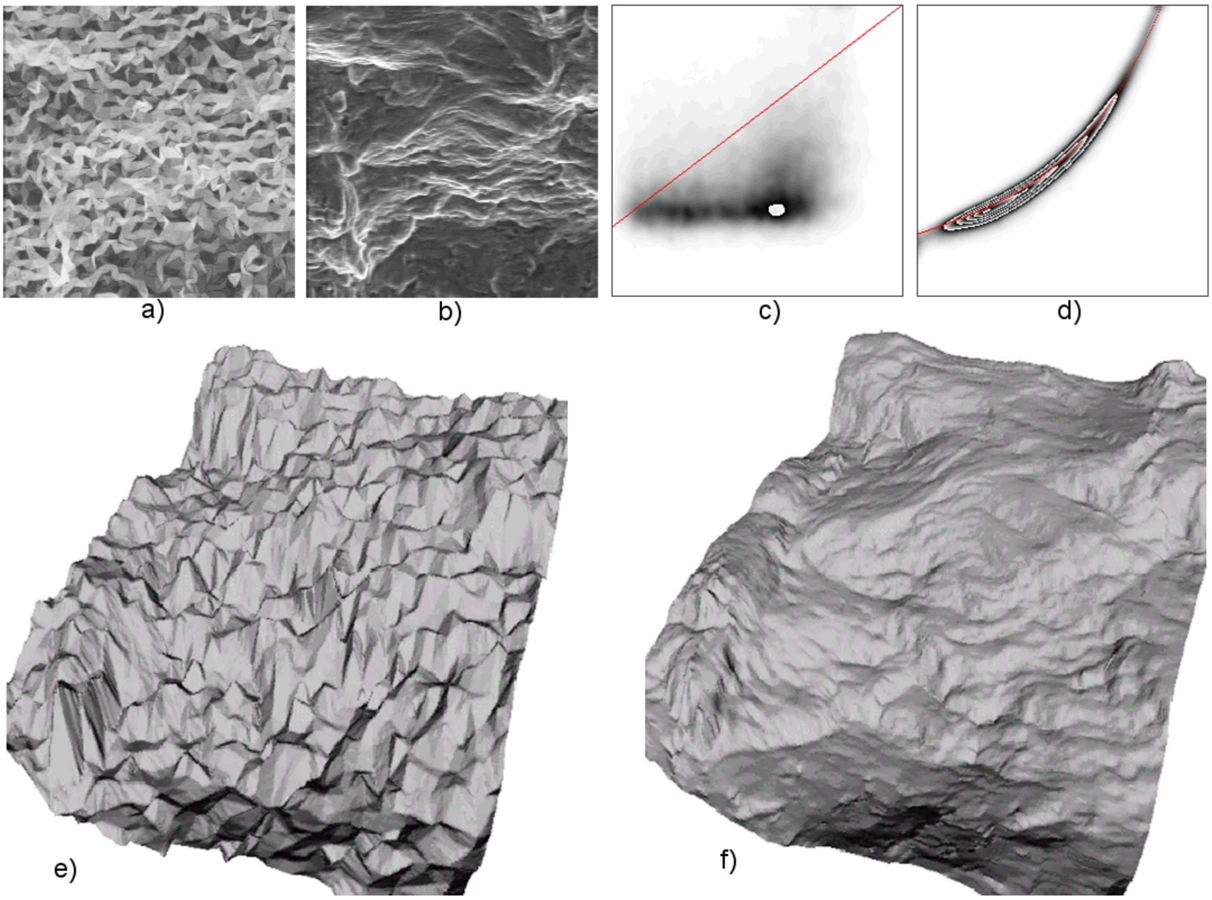

Figure 5: Experiment on real data: a) synthetic image generated using the initial reflectance map in c) and the initial DEM in e), b) input image which is identic to the synthetic image generated using the final DEM in f) and the refined reflectance map in d)

4 Conclusion

We have proposed a new integration for shape from shading and shape from stereo that allows to solve simultaneously for the shape of an object and its reflectance properties. The comparison with other algorithms shows that inclusion of stereo information during the whole reconstruction process and the use of a multigrid scheme on a hexagonal grid leads to particular improvements. Most important however is the proper modelling of a reflectance map. Further research will be done in applying the system to scenes with other reflectance properties or systems where the light source may not be assumed to be perpendicular to the object stage.

References

[1] W. Beil and I.C. Carlsen. Surface reconstruction from stereoscopy and 'shape from shading' in SEM images. *Machine Vision and Applications*, 4(4):271–285, Fall 1991.

[2] J.E. Cryer, P.S. Tsai, and M. Shah. Integration of shape from shading and stereo. *Pattern Recognition*, 28(7):1033–1043, July 1995.

[3] J.E. Dennis and R.B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations.* Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

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

[5] P. Fua and Y.G. Leclerc. Object-centered surface reconstruction: Combining multi-image stereo and shading. *International Journal of Computer Vision*, 16(1):35–56, Sept. 1995.

[6] B.K.P. Horn. Obtaining shape from shading information. *Chapter 4 in 'The Psychology of Computer Vision', Ed. P.H. Winston, McGraw-Hill Book Company, New York*, pages 115–155, 1975.

[7] K. Ikeuchi and B.K.P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17(1-3):141–184, Aug. 1981.

[8] A.G. Jones and C.J. Taylor. Scale space surface recovery using binocular shading and stereo information. In *BMVC '95 Proceedings of the 6th British Machine Vision Conference*, volume 722, pages 77–86, 1995.

[9] H. Lange. Advances in the cooperation of shape from shading and stereo vision. *Second International Conference on 3D Digital Imaging*, pages 46–58, 1999.

[10] Y.G. Leclerc and A.F. Bobick. The direct computation of height from shading. In *Proceedings of Computer Vision and Pattern Recognition*, volume 1991, pages 552–558, 1991.

[11] M.G.-H. Mostafa and S.M. Yamany. Integrating stereo and shape from shading. *Proceedings of the International Conference on Image Processing*, 3:130–134, 1999.

[12] D. Samaras, D. Metaxas, P. Fua, and Y.G. Leclerc. Variable albedo surface reconstruction from stereo and shape from shading. In *Proceedings of Computer Vision and Pattern Recognition*, pages 480–487, 2000.

[13] S. Scherer, A. Klaus, and A. Pinz. Surface reconstruction using backscatter electron images. *Proc. Workshop of the Austrian Association for Pattern Recognition*, pages 99–108, 1999.

[14] P. Werth and S. Scherer. A novel bidirectional framework for control and refinement of area based correlation techniques. *Proceedings of the International Conference on Pattern Recognition*, 3:730–733, 2000.