Reconstruction of Archaeological Finds using Shape from Stereo and Shape from Shading

Hannes Fassold[1], Reinhard Danzl[2], Konrad Schindler[1], Horst Bischof[1]

[1]Institute for Computer Graphics and Vision, Graz University of Technology

[2] VRVis Research Center for Virtual Reality and Visualization

8010 Graz, Austria

e-mail: {fassold, danzl, schindl, bischof}@icg.tu-graz.ac.at

Abstract

We propose a novel algorithmic framework for the refinement of sparse 3D models using shape from shading. Starting from an initial model obtained by shape from stereo we use a global optimization scheme in order to refine the surface. The constraints we use are based on the shading in the image, the initial 3D points obtained by stereo and the smoothness of the surface. In contrast to many previous approaches that assume that the photometric properties of the scene are known we iteratively update the light source direction and several parameters of the reflectance map. A careful choice of several boundary conditions on the reflectance map, the use of a multigrid scheme, the discretization of the object on a pseudo-hexagonal grid and the frequent switching between different refinement modules allows us to optimize the highly underconstrained problem. We provide results on synthetic data and reconstruction examples of frieze parts from an excavation site in Sagalassos (Turkey).

1 Introduction

The reconstruction of archaeological finds has become an important part of recording and conserving cultural heritage [10] [7] [15]. If the objects can be transported to a laboratory they are usually reconstructed using structured light [16] or laser scanning methods [15] that lead to results of high accuracy. If the objects can or shall not be removed from an archaeological site stereoscopic methods such as shape from stereo [7] and video [11] are usually preferred. These methods however have the disadvantage that the accuracy is often limited since many archaeological finds are rather textureless and the correspondence problem becomes tough.

In order to overcome the problems of stereoscopic methods we propose to combine them with shape from shading (SfS) [8] which is a reconstruction method that estimates the surface normals and hence the shape of an object from the shading information in a single image. Although this method is too under-constrained to be used for a stand-alone reconstruction, it is

well-suited for the refinement of rough initial models, especially in regions where the albedo of the object is almost constant (those regions where stereo has the most problems).

Since the invention of SfS three decades ago a large amount of different single SfS methods and several integrations of shape from shading and shape from stereo have been proposed. Although most of them work quite well for synthetic objects the amount of real world applications is rather limited which is mainly due to the fact that the reflectance properties and the light source directions have to be perfectly known in order to lead to good results. In order to apply our method to real images of archaeological objects we thus propose an approach that solves for shape, light source direction and various parameters of a reflectance map simultaneously.

As proposed in [6] we formulate the problem as a global minimization problem and build a framework that frequently switches between the refinement of the shape and the refinement of the photometric properties. This is done within a multigrid scheme that discretizes the model on a pseudo-hexagonal grid. Additionally we reproject the perspective images to orthographic images using the initial 3D models in order to simplify the energy minimization.

We first give an overview of our framework in Sect. 2. Afterwards we discuss the estimation of the reflectance map and the light direction in Sect. 3. Sect. 4 contains the details of our integrated reconstruction scheme based on stereoscopic and shading constraints. We continue with results on synthetic images and real datasets from the excavation site of Sagalassos (Turkey) in Sect. 5 and conclude with ideas for further research.

2 Overview

In this section we give an overview of our algorithmic framework and compare it with existing approaches that use an integration of stereo and shading in order to reconstruct light conditions, reflectance properties and shape of a scene.

Early approaches that reconstruct shape, light direction and the constant albedo of the surface are found in [20, 13, 6], which use only an initial stereo model and the image to calculate the light direction. A method that does not consider the estimation of the light direction but searches for shape and a two-dimensional polynomial model of the reflectance map has been proposed in [9]. Lange [12] has proposed an integrated framework using various loosely coupled modules and models the light direction and various parameters of Phong's reflection model. His experiments however use stereo pairs under different light conditions that are not available in the traditional shape from shading problem. Recently lots of work in this field has been done by Samaras and Metaxas [17, 18], who proposed an algorithmic framework for the integration of shape from stereo and shape from shading using deformable models [17] and presented a shading algorithm that iteratively refines the light source direction and the shape of the surface. Additionally, they correctly model perspective projection which has also been addressed by [6, 12, 14].

As proposed in [18] we also iteratively refine the reflectance map and the shape of the surface since the rough initial light and reflectance map estimate that can be obtained from the initial stereo model is not accurate enough. The main differences between their and our approach are different optimization methods, a different energy function and a different dis-

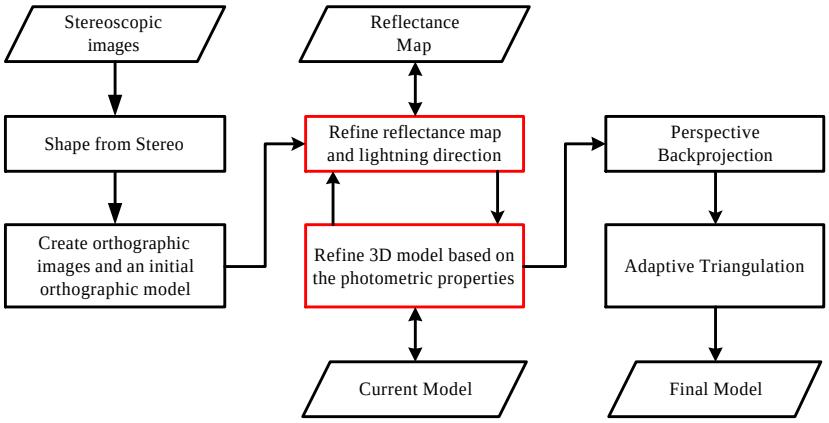

Figure 1: The workflow of our integrated algorithm

cretization of the object. An additional contribution is our introduction of boundary conditions for the optimization of the reflectance map estimation that make the estimation more robust.

Our algorithmic framework is depicted in Fig. 1. We start with several stereoscopic images and reconstruct an initial model using a standard shape from stereo method [10]. In order to simplify the integrated reconstruction equations we orthographically retransform the input images and the initial models. We are aware that this leads to errors in the input images since the rough stereo model is used for the transformation. However we currently work on a modification that uses correct perspective projection between the surface model and the images as proposed by Fua [6]. Afterwards the main reconstruction starts which iteratively switches between the refinement of the reflectance map and the light direction and the refinement of the surface shape. Although shape, reflectance and light direction could also be optimized at the same time, we chose to separate the processes as in [18]. This allows to separate between the optimization of global features such as the reflectance model and local features such as the shape and makes it possible to use appropriate optimization methods for each subproblem. In order to reduce the amount of data for visualization we perform adaptive triangulation after the final shape has been estimated by removing triangles in flat regions. Finally we retransform the model based on the perspective projection model.

3 Estimating the Reflectance Map and the Light Source Direction

In this section we derive a robust model for the nonlinear problem of estimating the light direction l and the parameters of the reflectance map $R(\alpha)$ where α is the angle between surface normal and light direction. The light direction is modelled via spherical coordinates as $l = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$ which ensures the unit length condition $\|l\| = 1$ automatically. We use the Lambertian reflectance map because it is a reasonable assumption for the archaeological objects we are interested in, however the algorithm could be easily extended to

more complex rotationally symmetric reflectance maps $R(\alpha)$.

Our experiments have shown that the shape from shading problem with unknown light direction and reflectance properties is too ill-posed and therefore tends to produce reflectance maps that are too flat. Thus we propose the following novel reflectance model that imposes two additional boundary conditions on the reflectance map and prevents the optimization from converging to the wrong solutions:

$$\begin{aligned} R(\alpha) &= \begin{cases} \widetilde{R}(\alpha) & \alpha \leq \frac{\pi}{2} \\ \widetilde{R}(\frac{\pi}{2}) - \frac{1}{10}(\alpha - \frac{\pi}{2}) & \alpha > \frac{\pi}{2} \end{cases} \\ \widetilde{R}(\alpha) &= c_1 + c_2^2 \cos\alpha \\ SDC - Constraint &: R(\frac{\pi}{2}) \leq \frac{1}{2} R(0) \\ RUB - Constraint &: R(0) \leq 255 \end{aligned}$$

By means of the quadratic coefficient c_2^2 we ensure that $R(\alpha)$ is a monotonic decreasing function in the interval $0 \leq \alpha \leq \frac{\pi}{2}$, while the SDC-Constraint assures that the function decreases at least to the half of $R(0)$ within the interval. Additionally, the RUB-Constraint ensures that $R(a)$ is always less than the maximum intensity value 255 and cannot leave the valid range. For angles $\alpha > \frac{\pi}{2}$ which belong to shadowed areas the reflectance map $R(\alpha)$ is almost constant. However, we introduce a small linearly decreasing term which improves the convergence properties and avoids that the algorithm gets stuck in a saddle point. Although some constraints are rather specific to our application (e.g. the SDC constraint) they can be easily adapted to other situations and allow convergence of the optimization for shape, reflectance map and light source that is often not possible otherwise.

In order to use the inequalities within our optimization framework we first transform them to equations. An inequality $g(x) \leq 0$ is transformed into an equation $g(x) + s^2 = 0$ by introducing a slack variable s that has to be additionally optimized. Afterwards the equations are traditionally transformed into additional constraints of the energy function.

Finally we arrive at a nonlinear optimization problem for the unknown parameters of the reflectance model c_1, c_2, ϕ, θ plus slack variables s_1, s_2. We solve it using the BFGS-method [4] which is a fast and robust Quasi-Newton-method for nonlinear optimization and achieves good convergence rates by approximating the inverse of the Hessian matrix (matrix of second derivatives) of the energy function.

4 Integrated Reconstruction

In the past various methods have been proposed to integrate stereo and shading clues. Several approaches use the stereo information only as an initial condition [13], some reconstruct two models separately and fuse them afterwards using various filter techniques [1] and others enforce the low-frequency information of an initial stereo model during each step of an iterative shading algorithm [5] by operation in the Fourier domain.

Based on a recent survey of different shape from shading methods [19] we have chosen to use a variational approach that tries to minimize a global energy function containing stereo,

shading and regularization constraints. This approach avoids artifacts that appear when using filtering methods in the frequency domain while enforcing low-frequency information during the refinement which is usually lost by pure SfS algorithms.

The reconstruction is done by minimizing a global energy function C which is the weighted sum of three terms

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

where C_B is the brigthness term from the image irradiance equation [8], C_{ST} is a stereo term that prevents the loss of low frequency information and C_S is a smoothness term for the stabilization of the algorithm. The first term 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 stereo term is of the form

$$C_{ST} = \sum_{k} (z_k - d_k)^2. \tag{3}$$

where d_k represents the height value of the model obtained by shape from stereo and z_k is the corresponding height value of the current model.

In order to avoid artifacts during the calculation of surface normals and smoothness terms we use a pseudo-hexagonal grid that is not pixel centered and is basically obtained by shifting each second row of the grid by half a pixel. Consequently our smoothness does not depend on four but on six neighboring gridpoints $z_a..z_f$

$$C_S = \sum_{k} (2z_k - z_a - z_d)^2 + (2z_k - z_b - z_e)^2 + (2z_k - z_c - z_f)^2. \tag{4}$$

where z_k is the height of the current gridpoint and the points $z_a..z_f$ are clockwise ordered. The term serves a dual purpose. First it 'convexifies' the energy functional which improves the convergence properties of the optimization procedure. Second, in the presence of noise, some amount of smoothing is required to prevent the mesh from over-fitting the data and excessively wrinkling the surface.

The optimization of the energy function is done using the conjugate gradient method [3] [6] which has proven to be suitable for large scale optimization problems such as shape from shading. In order to assure faster convergence and to avoid getting stuck in local minima too early we use a multigrid scheme that starts at a low resolution and increases the resolution till the final resolution of one model point per pixel is reached. Additionally, we use the continuation method as proposed by [13] which means that at each gridlevel we start with a high smoothness weight in order to avoid immediate overfitting of the data and decrease the smoothness weight as the computation goes on to allow a good reconstruction of discontinuity edges. The weights w_B, w_{ST} and w_S in the energy function are normalized with respect to

gradients of the corresponding constraints which makes them rather independent of the size and contents of the images.

This is similar to [6] also using a hexagonal grid and conjugate gradient descent for minimization. In contrast to them we do not use a stereo correlation term but a term that directly enforces the values of the initial stereo model which leaded to more robust results for our datasets.

5 Results

In the following we first apply our algorithm to a synthetic dataset with known ground truth and reflectance properties in order to show that the algorithm is able to estimate light source direction, reflectance parameters and 3D shape. Afterwards we provide an example using a real data set and show the applicability of our algorithm to real scenes of archaeological finds with almost uniform albedo.

5.1 Synthetic Data

For our synthetic experiment we used the well-known Mozart statue provided by the University of Central Florida [19]. The model was illuminated from the direction $(1, 1, 2)$ assuming Lambertian reflectance and rendered to an image of 256×256 pixels as shown in Fig. 2a. In order to simulate the result of a shape from stereo - algorithm we created our intial 3D model (Fig. 2c and e) by tiling the provided ground truth into triangles of 4×4 pixel and adding evenly distributed z-noise of $\pm 5\%$ of the total object height to the triangle nodes. We applied our SfS algorithm three times to this initial model using different initial light positions which were $35°$, $73°$ and $105°$ away from the correct light position. The initial Lambertian reflectance model is shown in the grey curve at the top of Fig. 2b.

In Table 1 different initial and final errors are summarized for the three experimental runs. The almost identical results show that our approach is rather insensitive to the initial position of the light source. Each time the difference EA_{final} between the correct and final estimated angle of the light source was about $2.7°$. Additionally, we have measured the mean error E_R between the correct and the final reflectance map which always is below 5 grey levels. The mean brightness error E_B between the input image and an image generated using the final model and the final reflectance properties was also only about 5 grey levels for each run.

In Fig. 2d and f the results of the algorithm for the initial light position error of $73°$ are visualized which show that the combined shading/stereo algorithm achieves a significant improvement of the quality of the 3D model. The relative z-error decreased from 1.7% to 0.5% and the 3D models are much closer to the ground truth than the initial models. Fig 2b shows two histograms where the x-axis corresponds to the angle $\alpha \in [0, \pi]$ between the surface normal of an object point and the light source and the y-axis to the grey value $g \in [0, 255]$ in the input image. The top histogram corresponds to the initial model and the bottom histogram to the final model. It shows the good correlation between angles and grey values for the final model. The estimated reflectance map is drawn into the histogram images as a grey curve.

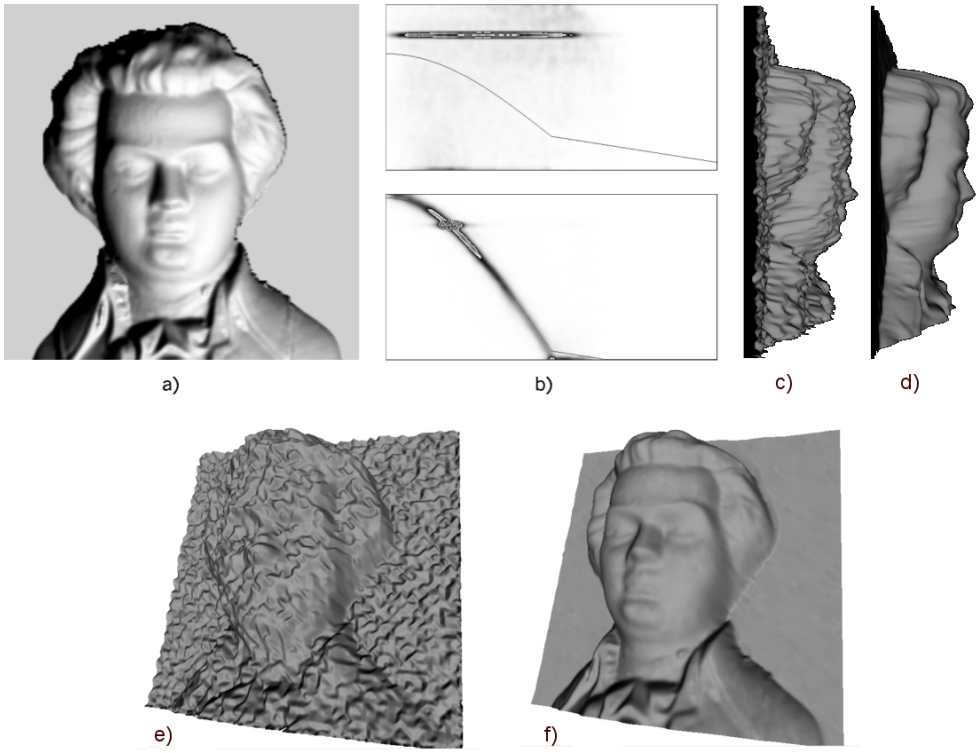

Figure 2: Reconstruction results for the synthetic object 'Mozart': a) Input image. b) Top: initial reflectance map. Bottom: final estimated reflectance map (details are given in the text) c) e) Initial 3D model obtained by subsampling the ground truth and adding 5% height noise. d) f) Final reconstructed 3D model.

EA_I	$EZ_I(\%)$	$ER_I(GV)$	$EB_I(GV)$	$EA_F(\%)$	$EZ_F(\%)$	$ER_F(GV)$	$EB_F(GV)$
$35°$	1.7	49.8	59	$2.7°$	0.5	4.7	5.1
$73°$	1.7	49.8	90.5	$2.7°$	0.5	5	5.2
$105°$	1.7	49.8	112.9	$2.8°$	0.6	4.8	5

Table 1: Reconstruction results for the synthetic object 'Mozart'. EA is the angular difference between the correct light source direction and the initial (I) and final (F) light source direction, respectively. EZ is the mean relative height error in $\%$. ER is the mean error in greyvalues between the correct and the initial/final reflectance map and EB is the mean error in greylevels between the input image and the image generated using the initial/final model and the corresponding reflectance map.

5.2 Real Data

We evaluated our algorithm on various real data sets from the archaeological excavation site in Sagalassos (Turkey) and present the reconstruction results for a stone frieze shown in Fig. 3a. The initial 3D model (Fig. 3e) was obtained using a standard stereo-matching algorithm on four input images [10] and shows a good approximation of the overall shape while lacking

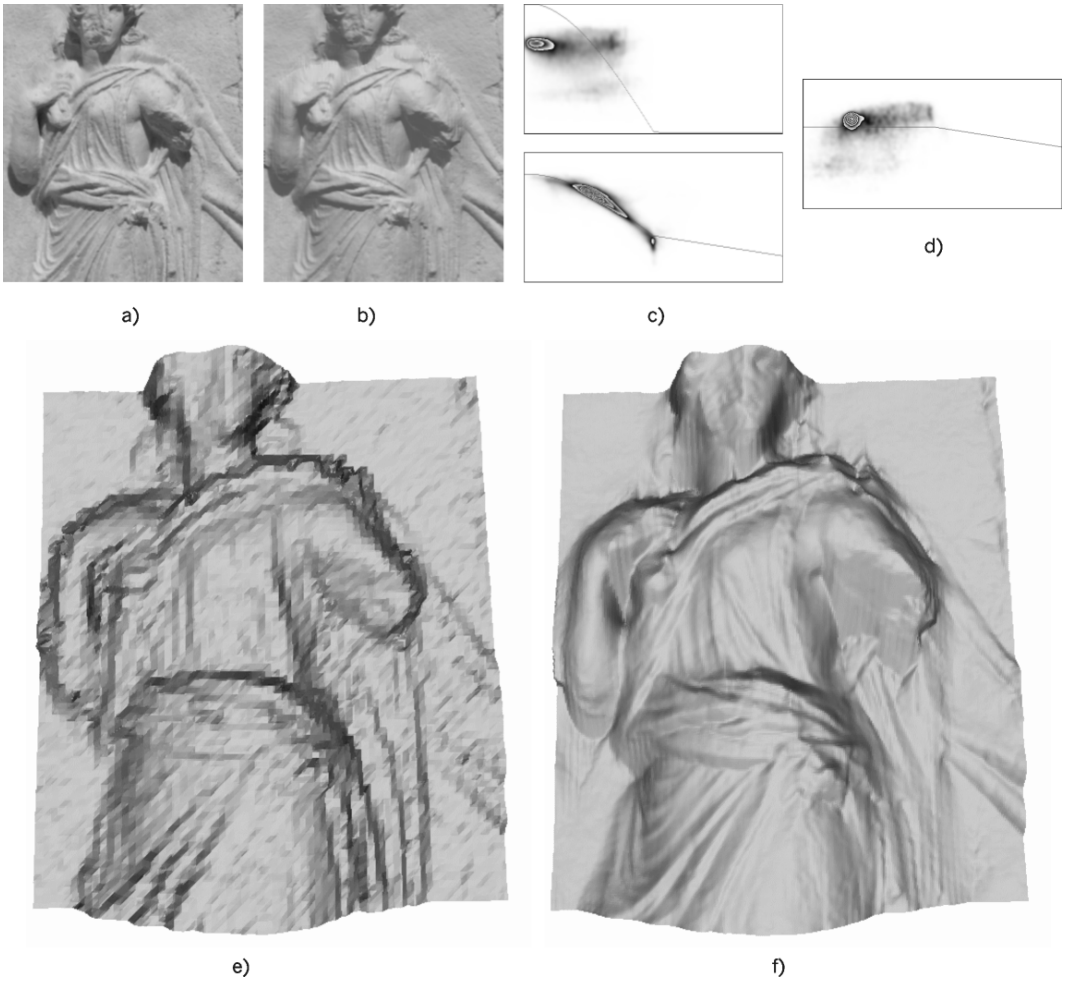

Figure 3: Reconstruction results for the stone frieze: a) Input image. b) Synthetic image generated using the final 3D model, the estimated light source and the estimated reflectance map. c) The initial reflectance map (top) and the final reflectance map (bottom). d) The final reflectance map if the SDC constraint is not used. e) The initial 3D model obtained by shape from stereo. f) The final refined 3D model.

small surface details. The final model after the refinement with SfS is visualized in Fig. 3f and shows the ability of the algorithm to reconstruct the high frequency components of the object. A synthetic image generated using the final model and the estimated reflectance parameters is shown in Fig. 3b, which is very close to the input image $(512 \times 592 pixel)$ and shows that the algorithm is able to decrease the brightness error sufficiently. On top of Fig. 3c the initial reflectance map and the (angle,greyvalue)-histogram for the initial model and reflectance parameters is visualized. The final reflectance map at the bottom of Fig. 3c approximates the histogram much better and shows that the assumption of a Lambertian reflectance model is justified. When we turn off the SDC-Constraint of our reflectance map the reflectance map does not converge to the correct solution as visualized in Fig. 3d which demonstrates the

necessity of this constraint.

In total the method has been evaluated on six different statues with the same set of parameters and for all datasets converged to visually correct results.

6 Conclusions and Future Work

We have proposed a new algorithm that allows to refine initial sparse and noisy 3D models using shading information of a single image. As shown in Section 5 the algorithm is able to reconstruct not only the shape but also the source of the light direction and the parameters of the Lambertian reflectance model. This use of shading information is especially valuable for various archaeological finds that usually have quite uniform albedo and allows a significant qualitative and quantitative improvement of an initial 3D model. During our experimental evaluation we have seen that some aspects of the algorithm are especially important. These include the use of boundary conditions for the reflectance map, a very frequent iteration between the surface model and the reflectance map and the use of a multigrid scheme during the optimization. A slightly modified version of the algorithm has also been applied to images from a scanning electron microscope [2], which shows its applicability to problems with very different reflectance characteristics.

There are two main limitations we want to address in the future. One is to proper model perspective projection during the reconstruction process which is currently done by initially rectifying the input images based on the initial 3D models. The second important aspect is to think about ways how to handle shadowed areas more specifically.

References

[1] J.E. Cryer, P.S. Tsai and M. Shah, Integration of shape from shading and stereo, *Pattern Recognition* Vol. 28, No. 7, pp.1033-1043, July 1995.

[2] R. Danzl and K. Karner, Integrating Shape from Shading and Shape from Stereo for Variable Reflectance Surface Reconstruction from SEM Images, *26th Workshop of the Austrian Association for Pattern Recognition*, pp.281-288, 2002

[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. Fletcher, *Practical Methods of Optimization. 2. ed.*, Wiley-Interscience Publication, Chichester, 1987.

[5] R.T. Frankot and R. Chellappa, A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, pp.439-451, July 1998.

[6] P. Fua and Y.G. Leclerc. Object-centered surface reonstruction: Combining multi-image stereo and shading, *International Journal of Computer Vision*, Vol. 16, No. 1, pp.35-56, Sept. 1995.

[7] A. Gruen, F. Remondino and Li Zhang, Image-based reconstruction of the great Buddha of Bamiyan, Afghanistan, *Proc. of the Int. Soc. for Optical Engineering*, pp.129-136, 2003.

[8] B.K.P. Horn and M.J. Brooks (eds.), *Shape from Shading*, Cambridge, Mass.: MIT Press, 1989.

[9] D.R Hougen and N. Ahuja, Adaptive polynomial modelling of the reflectance map for shape estimation from stereo and shading, *Proc. Computer Vision and Pattern Recognition*, pp.991-994, 1994.

[10] K. Schindler, M. Grabner and F. Leberl, Fast on-site reconstruction and visualization of archaeological finds, *Proc. 19th International CIPA Symposium*, pp.463-468, 2003.

[11] R. Koch, M. Pollefeys and L. Van Gool, Multi viewpoint stereo from uncalibrated sequences, *European Conference on Computer Vision*, pp.55-71, 1998.

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

[13] Y.G. Leclerc and A.F. Bobick. The direct comuptation of height from shading. *Proc. of Computer Vision and Pattern Recognition*, pp. 552-558, 1991.

[14] K.M. Lee and C.C.J. Kuo, Shape from shading with perspective projection, *CVGIP - Image Understanding*, Vol. 59, No. 2, pp.202-212, March 1994.

[15] M. Levoy, The digital Michelangelo project, *Sec. Int. Conf. on 3D Digital Imaging and Modeling*, pp.2-11, 1999.

[16] M. Proesmans and L. Van Gool, Reading between the lines - a method for etracting dynamic 3D with texture, *Proc of the Int. Conf. on Computer Vision*, pp.1081-1086, 1998.

[17] D. Samaras, D. Metaxas, P. Fua and Y.G. Leclerc, Variable albedo surface reconstruction from stereo and shape from shading, *Proc. of Computer Vision and Pattern Recognition*, pp.480-487, 2000.

[18] D. Samaras and D. Metaxas, Incorporating illumination constraints in deformable models from shape from shading and light direction estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 2, pp.247-264, 2003.

[19] R.Zhang, P.S. Tsai, J.E. Cryer and M. Shah, Shape from shading: a survey, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, pp. 690-705, August 1999.

[20] Q. Zheng and R. Chellappa, Estimation of lluminant direction, albedo and shape from shading, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 680-702, July 1991.