A shape-from-shading framework for satisfying data-closeness and structure-preserving smoothness constraints

Rui Huang
rui@cs.york.ac.uk
William A.P. Smith
wsmith@cs.york.ac.uk

Department of Computer Science
University of York
UK

Abstract

A key problem in shape-from-shading is how to simultaneously satisfy data-closeness and regularisation (such as surface smoothness) constraints. This paper makes two contributions towards solving this problem. The first is to describe a smoothness constraint which preserves surface structure by adaptively smoothing according to the intensity gradient magnitude. The second is to derive a framework which seeks to strictly satisfy this constraint while maintaining zero brightness error. Experimental results on both synthetic and real world imagery demonstrate that our method is both robust and accurate and outperforms a number of existing techniques.

1 Introduction

Shape-from-shading is a classical problem in computer vision which has attracted over four decades of research [1, 9]. The problem is underconstrained and proposed solutions have, in general, made strong assumptions in order to make the problem tractable. However, even when these assumptions are satisfied (for example in a synthetically produced image) existing shape-from-shading algorithms still fail to recover accurate surface shape from images of complex objects.

Minimization methods are a traditional and robust way to solve the shape from shading problem, first proposed by Horn [5]. These methods try to optimize the brightness error subject to additional regularisation constraints. Worthington and Hancock [8] treated the image irradiance equation as a hard constraint. Their idea was to use robust regularisers to optimise the solution within the space of solutions which strictly minimise the brightness error. Prados and Faugeras [7] used viscosity solutions to solve the partial differential equation which arises from the shape-from-shading problem. Their method accounts for perspective projection effects but assumes frontal illumination. The problem with these methods is that they trade off either strict minimisation of the brightness error versus satisfaction of the additional regularisation constraint. The result is that the recovered surfaces are either oversmooth or highly susceptible to noise.

BMVC 2009 doi:10.5244/C.23.106

More recently, several authors have posed shape-from-shading in terms of pairwise Markov Random Fields [3, 4, 6]. Haines and Wilson [3] describe surface normal direction probabilistically in terms of a Fisher-Bingham distribution for each pixel. Belief propagation is used to solve for the undetermined degree of freedom for each surface normal. Although their model provides an elegant formulation of the problem, in practice the results obtained are unconvincing. Potetz [6] also uses belief propagation but solves for the surface gradient at each pixel. The framework uses factor nodes to capture irradiance, smoothness and integrability constraints. Solving the resulting model is extremely difficult and results are only shown for a single synthetic image.

In this paper we present a new shape-from-shading algorithm which obtains a smooth solution with zero brightness error while seeking to preserve surface structure. The novelty in our framework lies in allowing the regularisation constraint to be run to convergence. This is possible because our constraint preserves surface structure as opposed to being trivially optimised by a planar surface. In other words, our smoothness constraint is also, in a sense, subject to constraints derived from the image intensity. In practice, our algorithm provides improved results over existing methods on a wide range of imagery of complex objects.

2 Shape from Shading

The aim of computational shape-from-shading is to make estimates of surface shape from the intensity measurements in a single image. Since the amount of light reflected by a point on a surface is related to the surface orientation at that point, in general the shape is estimated in the form of a field of surface normals(a needle-map). Assuming a normalised and linear camera response, the image intensity predicted by the simplest Lambertian reflectance model is given by

$$\mathbf{E}(\mathbf{N},\mathbf{L},\rho_{\mathbf{d}}) = \rho_{\mathbf{d}}\mathbf{N}\cdot\mathbf{L}, \tag{1}$$

where $\mathbf{N}$ is the local surface normal, $\mathbf{L}$ is a vector in the light source direction and ρ_d is the diffuse albedo which describes the intrinsic reflectivity of the surface.

For an image in which the viewer and light source directions are fixed, the radiance function reduces to a function of the surface normal. The problem is an ill-posed problem because the surface normal has two independent variables. Some additional constraint are needed as the regularizing term. Two popular regularizing terms are smoothness constraint and integrability constraint. Solving the shape from shading problem is converted into a minimization problem which tries to minimize

$$\mathscr{E}(\mathbf{n}) = \int\int (\mathbf{I}(\mathbf{x},\mathbf{y}) - \mathbf{E}(\mathbf{N}(\mathbf{x},\mathbf{y}),\mathbf{L},\rho_{\mathbf{d}}))^2 + \lambda_{\mathbf{s}}\mathbf{S}(\mathbf{N}(\mathbf{x},\mathbf{y})) + \lambda_{\mathbf{i}}\mathbf{Int}(\mathbf{N}(\mathbf{x},\mathbf{y}))\mathbf{dxdy} \tag{2}$$

where $\mathbf{S}(\mathbf{N}(x,y))$ and $\mathbf{Int}(\mathbf{N}(x,y))$ are smoothness and integrability constraints respectively. λ_s and λ_i are the coefficients of each term.

The minimization of the above function was done through variational calculus. However, the major problem for the above method is the brightness error will not be zero even if we provide the ground truth normal as initial input. An approach which overcomes these deficiencies was proposed by Worthington and Hancock [8]. Their idea was to choose a solution which strictly satisfies the brightness constraint at every pixel but uses the regularisation constraint to help choose a solution from within this reduced solution space.If we make the assumption that the reflectance properties are homogenous across the surface (i.e. constant

unit albedo), we obtain a simple relationship between observed intensity and the angle of incidence, $\theta_i = \angle\mathbf{NL}$, between the light source and surface normal:

$$I(x,y) = \mathbf{N}(x,y) \cdot \mathbf{L} = \cos\theta_i. \tag{3}$$

Geometrically, this means that the surface normal must lie on a right circular cone whose axis is the light source direction and whose half angle is $\theta_i = \arccos(I)$. By constraining the surface normal to lie on the cone, we satisfy the image irradiance equation and hence ensure the fullest possible use of the input image.

To solve this minimisation, Worthington and Hancock use a two step iterative procedure which decouples application of the regularization constraint and projection onto the closest solution with zero brightness error:

$$\begin{aligned} &1.\ \mathbf{n}'_t = f_{\mathrm{Reg}}(\mathbf{n}_t) \\ &2.\ \mathbf{n}_{t+1} = \underset{\mathscr{E}_{\mathrm{Bright}}(\mathbf{n})=0}{\arg\min}\ d(\mathbf{n},\mathbf{n}'_t), \end{aligned}$$

where $d(.,.)$ is the arc distance between two unit vectors and $f_{\mathrm{Reg}}(\mathbf{n}_t)$ enforces a robust regularizing constraint. The second step of this process is implemented using

$$\mathbf{n}_{t+1} = \Theta(\mathbf{a},\alpha)\mathbf{n}'_t, \tag{4}$$

where Θ is a rotation matrix which rotates a unit vector about an axis $\mathbf{a}$ by an angle α. To restore a normal to the cone we set $\mathbf{a} = \mathbf{n}'_t \times \mathbf{L}$ and $\alpha = \theta_i - \arccos(\mathbf{n}'_t \cdot \mathbf{L})$. The result is the closest direction that satisfies $\theta_i = \arccos(I)$.

However, neither the conventional variational method nor the method proposed by Worthington and Hancock realize the important of the regularisation term. The smoothness and integrability constraints they used are both surface constraints which assure that the result is a continuous surface. But they fail to preserve the structure of the surface which manifested in the intensity.

3 A Structure-Preserving Constraint

Individual pixel intensities provide a partial constraint on the local surface normal direction (namely the opening angle of the cone). However, the change in intensity across an image conveys information about the structure of the surface. Our aim is to exploit this information in our regularisation term.

The accuracy of the variational approach is entirely dependent on the choice of regularisation constraint. Typical smoothness constraints depend on the second derivative of surface height which is described as

$$\int\int (p_x^2 + p_y^2 + q_x^2 + q_y^2)dxdy, \tag{5}$$

where p and q are the first order partial derivatives of the surface height. The smoothness constraint is an isotropic constraint which assumes the surface changes are identical in every direction. Although the intensity gradient constraint was introduced to solve this problem [10], which requires that the intensity gradient of the reconstructed image be close to the intensity gradient of the input image in both the x and y directions, there is no existing

framework for fully exploiting structural information from the local neighbourhood in every direction.

We propose an alternative regularisation constraint which employs information about the intensity gradient in all directions over a local neighbourhood. For a pixel (x,y), we define the local neighborhood as $\Omega(x,y) = \{(x+1,y),(x-1,y),(x,y+1),(x,y-1)\}$. We precompute the change in incident angle between all pairs of neighboring pixels:

$$S((x_1,y_1),(x_2,y_2)) = \frac{|\arccos(I(x_1,y_1)) - \arccos(I(x_2,y_2))|}{\Delta S_{max}}, \tag{6}$$

where ΔS_{max} is the largest change in incident angle over the image which is used to normalize the weight . We define a weight between adjacent pixels based on the magnitude of the change in incident angle: $W((x_1,y_1),(x_2,y_2)) = e^{KS((x_1,y_1),(x_2,y_2))}$, where the constant K determines the behavior of the constraint (we use $K = 10$). For small values, the constraint reduces to local smoothness, for large values more structure is preserved at the cost of increased sensitivity to noise.

The total of the weights between a pixel and its neighbors is given by:

$$Z(x,y) = \sum_{(i,j)\in\Omega(x,y)} W((x,y),(i,j)). \tag{7}$$

The surface normal at pixel (x,y) at iteration $t+1$ is given by the weighted average of its neighboring normals at iteration t:

$$\mathbf{N}^{(t+1)}(x,y) = \frac{\boldsymbol{\mu}^{(t+1)}(x,y)}{\|\boldsymbol{\mu}^{(t+1)}(x,y)\|}, \tag{8}$$

where

$$\boldsymbol{\mu}^{(t+1)}(x,y) = \sum_{(i,j)\in\Omega(x,y)} \mathbf{N}^{(t)}(i,j)\frac{W((x,y),(i,j))}{Z(x,y)}. \tag{9}$$

If we choose the function $S((x_1,y_1),(x_2,y_2))$ to be a constant equal to zero, the iterative process will simplify to the conventional smoothness constraint. If the change of the intensity in x and y directions alone are taken into consideration, the structural constraint here is equivalent to the intensity gradient constraint.

The regularisation constraint described above could be incorporated into the standard variational approach. However, this method would still suffer from poor data-closeness due to the treatment of the image irradiance equation as a soft constraint. Instead, we note that the application of this constraint retains surface structure and so iteratively applying the constraint until convergence does not result in an oversmoothed surface.

4 A New Framework for Shape-from-shading

The variational method results in oversmooth surface estimates since the minimisation procedure does not strictly minimise the brightness constraint. In addition, the choice of coefficients for each regularisation term dramatically effect the result and in fact must be tuned for each input image. Worthington and Hancock [8] treat the brightness error as a hard constraint and rotate the normal such that Lambert's law is strictly satisfied at each iteration. The problem here is that the regularisation constraint is never satisfied since only a single

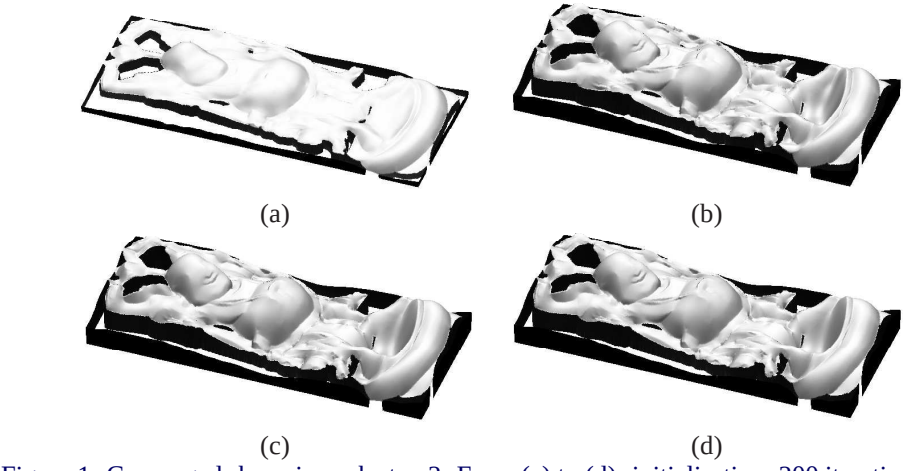


Figure 1: Converged shape in each step 2. From (a) to (d): initialisation, 200 iterations, 400 iterations, 600 iterations

application of the regularisation update is applied before the normal is rotated back to the cone. The effect is that surface normals flip between two positions: an on-cone position and a position representing a step towards reducing the error due to the regularisation constraint. The result does not improve significantly beyond the first iteration and decisions regarding the gross structure of the surface made at the stage of initialisation are simply reinforced.

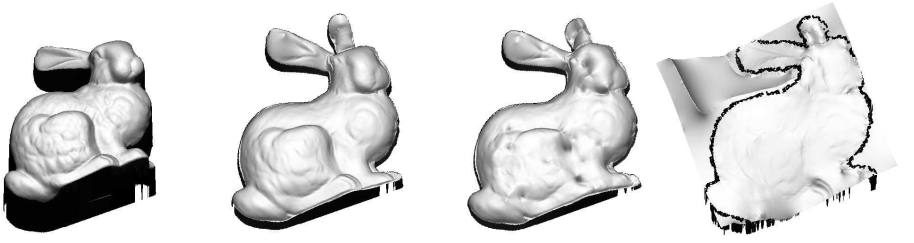
Our idea is to allow the regularisation update to be run to convergence before the surface normal is restored to the cone. This is only possible because our regularisation constraint seeks to preserve surface structure. Hence, even when run to convergence, the result still captures the gross structure of the surface. Our framework iteratively interleaves this process with rotating the normals back to their closest on-cone position. Since normals are allowed to move many times before being restored to the cone, changes to the gross structure of the surface are possible and information about surface structure is diffused over a wider region. We follow [8] and commence from an initialisation in which the surface normals are placed on their cones pointing in the direction of the negative image gradient.

A summary of our algorithm is as follows:

1. Obtain $\mathbf{N}^{(0)}(x,y)$ using negative gradient initialisation [8]
2. Repeatedly apply (8) until convergence
3. Rotate normals back to cone: $\mathbf{N}(x,y) = \Theta\mathbf{N}^{(\text{final})}(x,y)$
4. Stop if converged, otherwise iterate to step 2

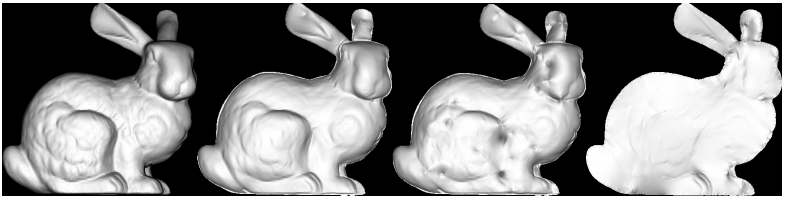
To obtain surface height estimates, we integrate the field of surface normals using the algorithm of Frankot and Chellappa [2].

Fig. 1 shows the converged shapes described in step 2. From left to right, top to bottom are the converged shape in subsequent statues before rotating back to the cone. The shape preserves the structure features, such as the the belly and the head, and notice that from the second recovered shape, the result is still reasonable before back to cone. The advantage of this framework will be seen in the experiment.



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

Figure 2: Surfaces recovered from the input image shown in the left panel of Fig. 3 From (a) to (d): ground truth, proposed algorithm, [8], [3].



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

Figure 3: Frontal View of Mesh. From (a) to (d): ground truth, proposed algorithm, [8], [3].

5 Experiment

We apply our method to both synthetic and real images. We compare our result with two previous methods: Worthington and Hancock's, and Haines and Wilson's. The unit albedo and frontal illumination is assumed for synthetic image.

Fig. 2 and Fig. 7 show the recovered surfaces on the Stanford bunny and Buddha using the proposed algorithm, Worthington and Hancock [8] and Haines and Wilson [3]. The corresponding view of the ground truth surface is shown in the first panel. The input image is shown in the top left panel of Fig. 3 The inaccuracy at the ear is because the ground truth data is pointed backwards and there is a discontinuity between the head and the ear. Note that the surface recovered by the proposed algorithm has better global structure whilst still containing much of the finescale surface detail. The advantage of our framework is also seen in this Fig. Notice that the holes in Worthington and Hancock's method [8] are the results of ruining the smooth constraint by rotating back to the cone in each iteration.

Fig. 3 and Fig. 8 show the frontal view of meshes of two objects. The corresponding view of the ground truth mesh is shown in the first panel. The meshes from left to right are proposed algorithm, Worthington and Hancock [8] and Haines and Wilson [3]. The surface recovered by the proposed algorithm is quite closed to the ground truth surface. The fine structure near the eyes and claws of the bunny and the belly hole of the Buddha are also preserved.

Fig. 4 and Fig. 9 compare the average angular error of normals in each iteration for both proposed method and Worthington and Hancock [8]. The proposed method performs better both at the error and converge speed. Note that from the use of data closeness constraint in the second time, the error increases rather than decreases. The reason for this is the concave/convex problem. The structure constraint smooths the surface to reduce the error while data closeness constraint increases the error.

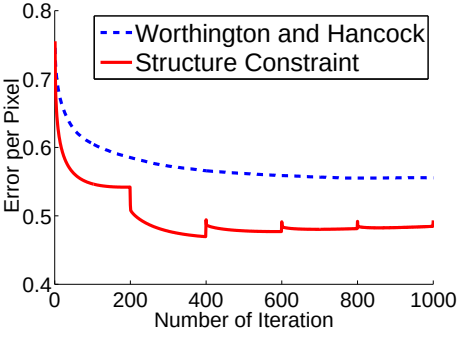
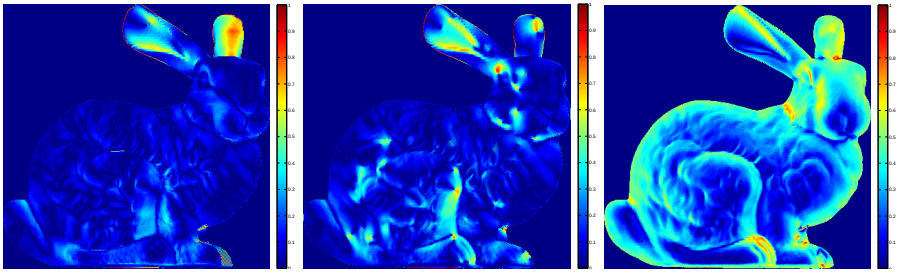
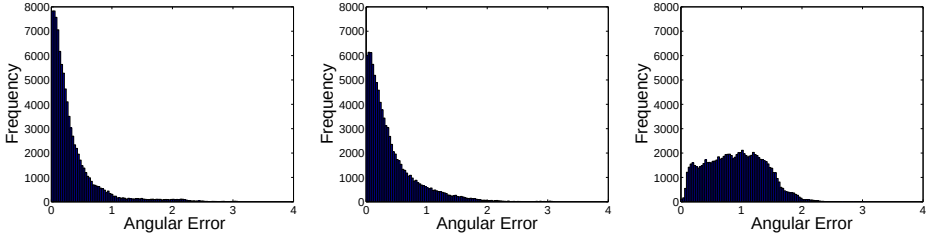


Figure 4: Mean angular Error in each iteration.



(a) (b) (c)

Figure 5: Error Map, From (a) to (c): proposed algorithm, [8], [3].



(a) (b) (c)

Figure 6: Histogram of Errors, From (a) to (c): proposed algorithm, [8], [3].

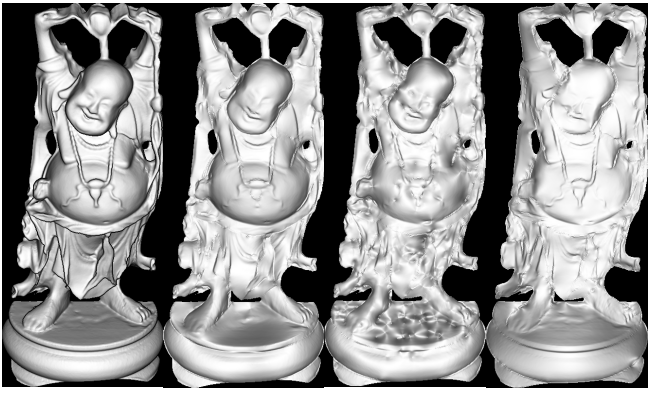
Fig. 5 and Fig. 10 show the location of angular errors in each recovered needlemap. The angular error has been mapped into the scale of 0 to 1. The dark blue represents the error is zero and the red represents the biggest error. Note that the Worthington and Hancock's method [8] performs well at the bunny's ear at the concave and convex decision. The proposed method has a better performance in whole for the global images are dark blue. The overall body and head of the bunny is blue and the belly of the Buddha is also better than those of other methods. The results illustrate that under the correct convex and concave situation, the proposed method does well in the detail structure.

Fig. 6 and Fig. 11 compare the histogram of the angular error. The propose method is much more accurate than other methods.

Finally, we apply our algorithm to a real captured at the mountain area. The sky has

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

Figure 7: Surfaces recovered from the input image shown in the left panel of Fig. 8. From (a) to (d): ground truth, proposed algorithm, [8], [3].



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

Figure 8: Frontal View of Mesh. From (a) to (d): ground truth, proposed algorithm, [8], [3].

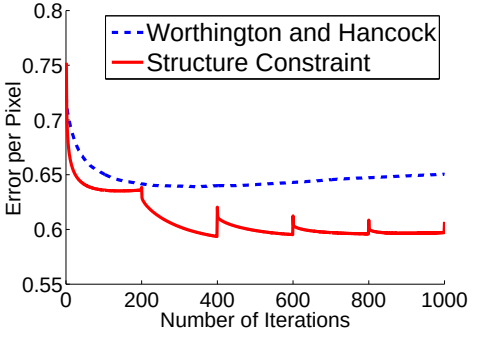
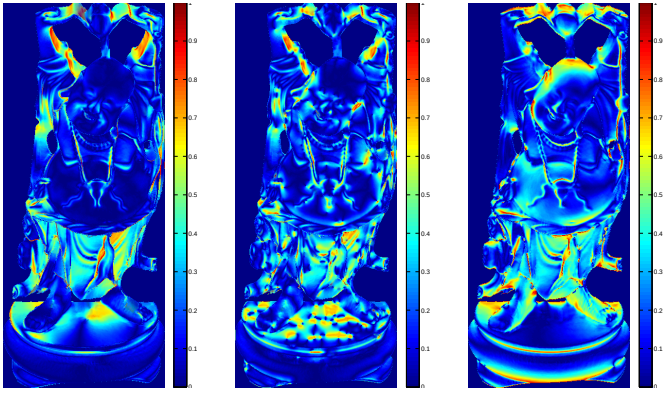


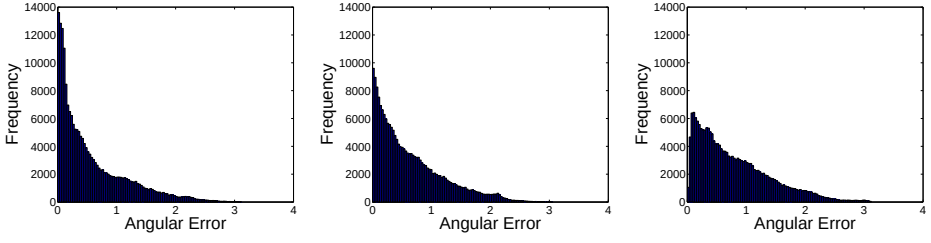
Figure 9: Mean angular error in each iteration.

been manually cut off. The light source direction is considered to be from the above. The proposed method also gives a plausible result under the non-frontal illumination case. Note that the parts with the cast shadow, the proposed method shows the potential in handling small change in albedo.



(a) (b) (c)

Figure 10: Error Map, From (a) to (c): proposed algorithm, [8], [3].



(a) (b) (c)

Figure 11: Histogram of Errors, From (a) to (c): proposed algorithm, [8], [3].

(a) (b) (c)

Figure 12: Real Image, (a) is input image, (b) and (c) are heightmap from different view point

6 Conclusions

We have presented a practical and robust framework for a shape-from-shading algorithm which recovers stable surface estimates from a wide range of real and synthetic imagery. The drawback of oversmoothed solution is overcome by applying the structure regularisation. The new framework incorporates both the structure constraint and data closeness constraint. In future work we intend to investigate alternative initialisations and explore how to incorporate integrability constraints within our framework.

References

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

[2] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10(4):439–451, 1988.

[3] Tom S. F. Haines and Richard C. Wilson. Belief propagation with directional statistics for solving the shape-from-shading problem. In *Proc. ECCV*, pages 780–791, 2008.

[4] Feng Han and Song-Chun Zhu. Cloth representation by shape from shading with shading primitives. In *Proc. CVPR*, volume 1, pages 1203–1210, 2005.

[5] B. K. P. Horn and M. J. Brooks. The variational approach to shape from shading. *Comput. Vis. Graph. Image Process.*, 33(2):174–208, 1986.

[6] Brian Potetz. Efficient belief propagation for vision using linear constraint nodes. In *Proc. CVPR*, pages 1–8, 2007.

[7] E. Prados and O. Faugeras. Perspective shape from shading and viscosity solutions. In *Proc. ICCV*, volume 2, pages 826–831, 2003.

[8] P. L. Worthington and E. R. Hancock. New constraints on data–closeness and needle map consistency for shape–from–shading. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21 (12):1250–1267, 1999.

[9] R. Zhang, P. S. Tsai, J. E. Cryer, and M. Shah. Shape–from–shading: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(8):690–706, 1999.

[10] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):680–702, 1991.