Direct Shape-from-Shading with
Adaptive Higher Order Regularisation

Oliver Vogel, Andrés Bruhn, Joachim Weickert, and Stephan Didas

Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science, Building E1.1
Saarland University, 66041 Saarbrücken, Germany
{vogel,bruhn,weickert,didas}@mia.uni-saarland.de

Abstract. Although variational methods are popular techniques in the context of shape-from-shading, they are in general restricted to indirect approaches that only estimate the gradient of the surface depth. Such methods suffer from two drawbacks: (i) They need additional constraints to enforce the integrability of the solution. (ii) They require the application of depth-from-gradient algorithms to obtain the actual surface. In this paper we present three novel approaches that avoid the aforementioned drawbacks by construction: (i) First, we present a method that is based on homogeneous higher order regularisation. Thus it becomes possible to estimate the surface depth directly by solving a single partial differential equation. (ii) Secondly, we develop a refined technique that adapts this higher order regularisation to semantically important structures in the original image. This addresses another drawback of existing variational methods: the blurring of the results due to the regularisation. (iii) Thirdly, we present an even further improved approach, in which the smoothness process is steered directly by the evolving depth map. This in turn allows to tackle the well-known problem of spontaneous concave-convex switches in the solution. In our experimental section both qualitative and quantitative experiments on standard shape-from-shading data sets are performed. A comparison to the popular variational method of Frankot and Chellappa shows the superiority of all three approaches.

Keywords: computer vision, shape-from-shading, variational methods, partial differential equations.

1 Introduction

The recovery of the 3-D shape of a surface from a single shaded image is one of the classical reconstruction problems in computer vision. Since the first prototypical approach of Horn three decades ago [5], a variety of algorithms have been developed; see e.g. [9, 16]. In particular, two classes of shape-from-shading methods are frequently used in the literature: propagation techniques that recover the shape by propagating information from a set of known surface points (critical points) to the whole image [5, 8, 11, 15, 13], and variational methods that compute the solution as minimiser of a suitable energy functional [7, 2, 4]. In this paper
we focus on the latter class of techniques. Although variational methods have been very popular in the 80s, they have stopped being considered already one decade later, when propagation approaches started to dominate. In this context, they have been criticised to suffer from a number of drawbacks [11]:

- **Indirect Strategy.** In contrast to many other techniques that allow a direct estimation of the depth field [10, 12], variational methods have the reputation to be only applicable for the recovery of the gradient field of the surface depth. Such indirect variational approaches are e.g. the methods of Horn and Brooks [2] and the algorithm of Frankot and Chellappa [4]. These methods compute first the gradient field of the depth and then have to rely on the subsequent application of a depth-from-gradient technique [4, 1].

- **No Intrinsic Integrability.** Moreover, indirect techniques also require the use of integrability constraints to prevent impossible solutions [2, 6, 4]: If the two gradient functions in x- and y-direction – let them be given by \( p(x, y) \) and \( q(x, y) \), respectively – are computed independently from each other, there is no guarantee that there exists a common depth map \( z(x, y) \) for which holds \( p(x, y) = z_x(x, y) \) and \( q(x, y) = z_y(x, y) \). Such integrability constraints have been first introduced by Horn and Brooks in [6]. However, since these constraints only encourage the integrability of the solution, but not enforce it, still intermediate steps are necessary that backproject the estimated gradient field into the range of admitted solutions [4].

- **Over-regularisation.** Furthermore, since variational methods are based on a regularising smoothness assumption, it has often been remarked that this regularisation introduces a strong blurring in the solution that may deteriorate the quality of the reconstruction [11]. This issue has recently been researched by Agrawal et al [1], however, only with respect to variational depth-from-gradient techniques. For direct variational shape-from-shading approaches such an investigation of adaptive regularisers is missing.

- **Spontaneous Concave-Convex Switches.** Finally, it is well-known that shape-from-shading in its original formulation (orthographic projection, Lambertian surface) is an ill-posed problem [5]. Due to the related concave-convex ambiguity, spontaneous switches in the solution may occur if the strength of the regularisation is too low. In this context one should note by incorporating knowledge on the surface at critical points [11] or assuming other conditions such as perspective instead of a orthographic projection [14], shape-from-shading can be turned into a partially or even fully well-posed problem. In this case the problem of spontaneous switches does not occur. However, since we restrict ourselves to the classical problem formulation without prior information, this issue remains relevant for us.

In our paper all four aspects are addressed. We show by the example of the methods of Horn and Brooks [2] and of Frankot and Chellappa [4] how the
use of higher order smoothness terms allows a direct computation of the surface depth. Consequently, no shape-from-gradient algorithms are required and all integrability constraints become obsolete by construction. Moreover, we investigate how the smoothness term of our novel method can be adapted. This also includes a strategy to cope with spontaneous concave-convex switches in the solution.

Our paper is organised as follows. In Section 2, we give a short review on the shape-from-shading problem, while in Section 3 we discuss the indirect approaches of Horn and Brooks and Frankot and Chellappa. Based on the outcome of this discussion, we then develop a novel direct variational approach for shape-from-shading in Section 4. In Section 5 we extend this approach by adaptive smoothing strategies. We propose variants that are based on image- and depth-driven smoothness terms. In Section 6, experiments for all three novel algorithms are presented, while the summary in Section 7 concludes this paper.

2 The Shape-from-Shading Problem

Let us consider a shaded image as a function \( f(x, y) \) where \((x, y)\) denotes the location within a rectangular image domain \( \Omega \). Furthermore, let us assume that the surface \( z(x, y) \) depicted in this image has been illuminated by a single light source only and that its reflectance properties can be expressed in terms of a reflectance map \( R(z_x(x, y), z_y(x, y)) \). Such a reflectance map is a function that describes the amount of light reflected by the surface in viewing direction depending on its gradient \((z_x(x, y), z_y(x, y))\). Then, solving the shape-from-shading problem means to find a suitable surface \( z(x, y) \) such that the amount of light reflected by it equals the grey value that is observed at the corresponding pixel of the image:

\[
\mathbf{f} = R(z_x, z_y). \tag{1}
\]

In the literature this equation is known as the image irradiance equation [5,6]. Further assumptions in the classical shape-from-shading problem [5] are that the shaded image was obtained by an orthographic projection and that the surface is a Lambertian surface, i.e. the reflectance map reads

\[
R(z_x, z_y) = \rho \cdot \langle s, n \rangle \tag{2}
\]

where \( n \) denotes the unit surface normal

\[
n = \frac{1}{\sqrt{1 + z_x^2 + z_y^2}} \begin{pmatrix} -z_x \\ -z_y \\ 1 \end{pmatrix}, \tag{3}
\]

\( s \) stands for the light source direction, and \( \rho \) is the albedo of the surface—a constant that specifies the ratio of scattered to incident light. Evidently, this problem is ill-posed: For surfaces that are illuminated from above, i.e. \( s = (0, 0, 1)^T \), both the surface \( z \) and its negative counterpart \(-z\) are solutions of the image irradiance equation. As we will see later, this ambiguity yields the well-known concave-convex switches in the solution.
3 Variational Shape-from-Shading

In order to solve the classical shape-from-shading problem, numerous variational approaches have been proposed in the literature. All these approaches, however, are based on an indirect two-step strategy: First the surface gradient is computed, then the actual surface is determined. Two typical representatives the method of Horn and Brooks [2] and its improved variant by Frankot and Chellappa [4]. In the following both approaches are discussed in detail.

3.1 Horn and Brooks

One of the first variational methods for shape-from-shading that still enjoys great popularity is the approach by Horn and Brooks [2]. This approach computes the surface gradient \( \nabla z = (z_x, z_y) =: (p, q) \) as minimiser of the following energy functional:

\[
E(p, q) = \int_{\Omega} \left( (f - R(p, q))^2 + \alpha \left( |\nabla p|^2 + |\nabla q|^2 \right) \right) \, dx \, dy .
\]  

(4)

While the first term (the data term) penalises deviations from the image irradiance equation, the second term (the smoothness term) assumes the recovered surface derivatives \( p \) and \( q \) to be smooth. The degree of smoothness of the solution is steered by a regularisation parameter \( \alpha > 0 \).

Obviously, this method suffers from two drawbacks. First of all, it is not considered explicitly in the formulation of the energy functional that \( p \) and \( q \) are derivatives of a common surface \( z \). Thus it is not surprising that in most cases there might not even exist a surface \( z \) with \( \nabla z = (p, q) \) for the computed solution. Secondly, even if the integrability would be enforced by means of other constraints during the solution process of (4) as proposed in [6,4], it is not trivial to actually obtain the desired surface \( z \). Additional depth-from-gradient techniques [4,1] are necessary anyway to compute the actual surface.

3.2 Frankot and Chellappa

Frankot and Chellappa [4] proposed a solution to both the integrability problem and the problem of recovering the depth from the gradient field. After each iteration step of the Horn and Brooks algorithm, they project the computed gradient \((p, q)\) onto the closest integrable pair of functions \((\tilde{p}, \tilde{q})\) by minimising

\[
\int_{\Omega} \left( |p - \tilde{p}|^2 + |q - \tilde{q}|^2 \right) \, dx \, dy .
\]  

(5)

To this end, they compute the Discrete Fourier Transforms (DFT) of \( p \) and \( q \) and perform the projection step in the frequency domain. Moreover, they also propose a way to integrate \((p, q)\) in the frequency domain. However, from a modelling point of view such an alternating approach is neither desirable nor is its convergence mathematically understood. Hence, it would be much more natural to estimate the surface depth \( z \) directly.
4 Who Dares Wins: Higher Order Regularisation

For solving both previously discussed problems more reliably, we propose the following strategy: By considering the actual approach of Horn and Brooks in (4) and replacing $p$ by $z_x$ and $q$ by $z_y$, we obtain a direct variational shape-from-shading method that overcomes all integrability problems by construction. The corresponding energy functional of this novel approach is given by

$$E(z) = \int_\Omega \left((f - R(z_x, z_y))^2 + \alpha \left(\frac{\partial^2}{\partial x^2} z + 2z_{xx} + z_{yy}\right)\right) \, dx \, dy.$$  \hspace{1cm} (6)

Please note that all integrability problems only vanish if this energy functional is solved for $z$ and for $z_x$ and $z_y$ as proposed in [4]. Such a direct computation of $z$ also offers another advantage: Since it guarantees that $z_{xy} = z_{yx}$, we obtain a second order smoothness term in a natural way (the smoothness term in (4) is only a first order regulariser). This new smoothness term can be identified as the squared Frobenius norm of the Hessian.

Following the calculus of variations [3], we know that a minimiser of an energy functional must satisfy its Euler-Lagrange equation(s). For our novel approach this partial differential equation is given by

$$0 = \frac{\partial}{\partial x} \left((f - R(z_x, z_y))R_{x_x}(z_x, z_y)\right) + \frac{\partial}{\partial y} \left((f - R(z_x, z_y))R_{x_y}(z_x, z_y)\right)$$

$$+ \alpha \left(\frac{\partial^2}{\partial x^2} z + 2z_{xx} + z_{yy}\right).$$  \hspace{1cm} (7)

As one can see, our new regulariser results in a homogeneous fourth order diffusion term. Evidently, this makes its discretisation more complicated than the one of a standard second order diffusion expression. However, a common strategy proved to be very successful in this case: By discretising the continuous functional in (6) and computing its derivatives a suitable discretisation for (7) can be derived that even contains the correct boundary conditions.

In order to solve the resulting nonlinear system of equations, we apply a Jacobi-like method as proposed in [6]. Thereby the complete data term is taken from the old iteration step. In this context, one should note that the solution of this equation system may not be unique, since our energy functional is not strictly convex. This means that the solution we find might not be the global minimum of the functional, but only a local one. Although one can avoid this problem by manually providing the algorithm with the correct shape at the image boundary and occluding boundaries (and thus turn the classical shape-from-shading task into a well-posed problem), we want to tackle the original problem and thus intentionally refrain from providing this this prior knowledge.

5 Adaptive Higher Order Regularisation

It turns out that the shape of simple images can be reconstructed well using small weights $\alpha$ for the smoothness term. However, for complex images with many
edges and occluding boundaries, it is necessary to use a stronger smoothness term to obtain a convergent iteration. Evidently this decreases the accuracy of the reconstruction. In the following we present two additional approaches that tackle this problem: By adapting the smoothness term either to the input image or to the depth map they reduce the amount of unnecessary regularisation and thus allow for a more precise reconstruction of the final surface.

5.1 Image-Driven Regularisation

So far, our variational approach from Section 4 only uses a homogeneous smoothness term, i.e. a regulariser that does not distinguish between different locations in the image. However, it is well known that at discontinuities the information provided by the input image is very poor. Consequently, the regularisation at such locations should be much stronger than in flat areas, where the shape of the surface is easy to reconstruct. In other words: One can improve the reconstruction quality in flat areas by regularising less.

In order to model this observation, we propose a modified approach that adapts its regularisation to the local image structure. It is given by the energy functional

$$E(z) = \int_{\Omega} \left( (f - R(z_x, z_y))^2 + \alpha g(|\nabla f|^2) \|\text{Hess}(z)\|_F^2 \right) dx \, dy.$$  

(8)

that makes use of a weighting function $g(|\nabla f|^2)$ in the smoothness term. Considering $|\nabla f|$ as an edge indicator, this weighting function should be large where $|\nabla f| \gg 0$, and small where $|\nabla f| \approx 0$. Thus, any positive, monotonically increasing function $g$ can be chosen for this task. In the context of our image-driven approach, we use

$$g(s^2) = \sqrt{s^2 + \epsilon^2}$$  

(9)

where $\epsilon > 0$ is a small positive constant that ensures at least a small amount of regularisation. As one can easily verify, this function attains its minimum value $\epsilon$ at $s^2 = 0$ and approaches the identity function for large $s^2$.

As for our homogeneous method, the minimisation of (8) requires to solve its Euler-Lagrange equations. In the case of image-driven regularisation they are given by

$$0 = \frac{\partial}{\partial x} \left( (f - R(z_x, z_y))R_{z_x}(z_x, z_y) \right) + \frac{\partial}{\partial y} \left( (f - R(z_x, z_y))R_{z_y}(z_x, z_y) \right) + \alpha \left( \frac{\partial^2}{\partial x^2} (g(|\nabla f|^2)z_{xx}) + 2\frac{\partial^2}{\partial x \partial y} (g(|\nabla f|^2)z_{xy}) + \frac{\partial^2}{\partial y^2} (g(|\nabla f|^2)z_{yy}) \right).$$  

(10)

Compared to the Euler-Lagrange equation 7 of our shape-from-shading method with homogeneous higher order regularisation, the adaptive smoothness term induces a linear fourth order diffusion process, where $g$ plays the role of a diffusivity function. Again, we suggest to derive a suitable discretisation for this equation by discretising the continuous functional and computing its derivatives. The obtained equation system is then solved once more by a Jacobi-like iteration step.
5.2 Depth-Driven Regularisation

In our second approach, we proposed to use edge information of the input image \( f \) to control the strength of the regularisation. However, in particular with respect to situations where frequent convex-concave switching artifacts occur, it may make much more sense to adapt the weight of the induced diffusion process to the edge information of the evolving depth map instead. This has the following reason: Since such artifacts manifest themselves in large gradients in the recovered surface, they can be tackled by increasing the smoothness at the corresponding locations.

Therefore, we propose to replace the image gradient \(|\nabla f|\) in the Euler-Lagrange equations of the image-driven case by the depth gradient \(|\nabla z|\). Thus, equation (10) turns into the following PDE:

\[
0 = \frac{\partial}{\partial x} \left( (f - R(z_x, z_y)) R_{xx} (z_x, z_y) \right) + \frac{\partial}{\partial y} \left( (f - R(z_x, z_y)) R_{yy} (z_x, z_y) \right) + \alpha \left( \frac{\partial^2}{\partial x^2} \left( g(|\nabla z|^2) z_{xx} \right) + \frac{\partial^2}{\partial y^2} \left( g(|\nabla z|^2) z_{yy} \right) \right).
\]

One should note that in contrast to the image-driven approach, this approach is now based on a fourth order diffusion process which is nonlinear. However, essentially the same strategy as in the previous two cases can be applied to solve this equation numerically.

6 Experiments

Let us now investigate the reconstruction quality of all of our three novel techniques. To this end, we perform experiments using two popular shape-from-shading test images: the penny and the Mozart face [16]. Moreover, we compared the results of our techniques to the reconstruction of the algorithm of Frankot and Chellappa. This allows to analyse the advantages of the presented methods in a systematic way. With respect to the presented results for the method of Frankot and Chellappa, one should note that we also refrained in their case from providing information about singular points or occluding boundaries to the algorithm (this was also done in their original paper). In this way a fair comparison can be guaranteed. The parameter \( \alpha \) was optimised manually to minimise the average \( L_1 \)-error (12).

6.1 The Penny

In our first experiment, we used the penny image of size \( 128 \times 128 \) pixels depicted in Figure 1. It shows the surface of a coin with Lambertian reflectance properties. This data set is challenging due to several reasons: One on hand the surface of the contain is a large-scale structure with a sharp edge at the top. Since we assume the surface to be smooth, this part should be difficult to reconstruct. On the other hand the coin is mainly concave, with convex engravings. The
algorithm does not incorporate this shape information. So if the general shape of the coin was convex, this would be a good reconstruction as well. Moreover, the engravings on the coin are a problem themselves: While the head of Abraham Lincoln is pretty large, there are also writings on the coin, which are small-scale structures that are very difficult to recover.

In the left column of Figure 1, the computed reconstructions for the penny image are shown. Using the Frankot-Chellappa algorithm, we obtain a very curvy surface: Some parts of the surface are convex, some concave. Moreover, the edge of the coin is hardly recovered at all. Our homogeneously regularised algorithm improves this result: Lincoln is reconstructed relatively well, and the coin edges are also recovered much better. The surface at the edge, however, is partly estimated to be convex, partly to be concave. The image-driven algorithm has the same problem at the coin edge, however, the detailed structures are recovered much shaper. Nevertheless, we can observe that the surface tends to switch between concave and convex shape spontaneously – this can be observed very well at Lincoln’s hair. Using depth-driven regularisation, these artifacts disappear almost completely. The reconstruction now looks somewhat like a coin (although its general shape is convex instead of concave). In the middle column, the back-projections of the reconstructions are compared. As one can see, the reconstructions using the Frankot-Chellappa algorithm and our homogeneously regularised algorithm are quite blurry, the writings are hardly readable and details of Lincoln are lost. With our algorithms based on adaptive higher order regularisation this blurring is reduced significantly: Much more details are preserved and in the case of the image-driven smoothness term even the engravings are readable very well. This is also confirmed by our error plots in the right column. There, the absolute differences between the ground truth and the backprojections of the reconstructions are shown (scaled by a factor 12 to improve visibility of the error).

6.2 The Mozart Face

In our second experiment we computed the reconstructions for the Mozart face of size 256 × 256. In this case, the task is even more difficult than in the case of the penny, since the original surface is very complex. As a consequence, the image contains many edges and singular points which may result in numerous concave-convex switchings in the reconstruction.

The computed surfaces presented in the first column of Figure 2 show similar tendencies as for the penny image. While the result for the Frankot-Chellappa algorithm is curvy and suffers significantly from a higher number of concave-convex switchings, the reconstruction using our homogeneously regularised algorithm is already slightly better. However, once again the approaches based on adaptive higher order regularisation yield the most detailed reconstructions: While the image-driven approach gives the sharpest results, it still suffers from several concave-convex switches. In the case of the depth-driven approach these switches are almost not existent. The backprojected images in the central column look once again very reasonable for all approaches. However, the error plots
Fig. 1. The Penny. Left to right: Depth map, orthographic projection from above, difference image (scaled by factor 12). Top to bottom: Ground truth, Frankot-Chellappa, homogeneous regularisation, image-driven regularisation, depth-driven regularisation.
in the right column reveal the superiority of the proposed approaches. Again, for the Frankot-Chellappa algorithm, the error is spread all over the image, while it is mainly concentrated to edges in the adaptively regularised reconstructions. Thereby, the concave-convex switching artifacts lead to scar-like artifacts. These artifacts are greatly reduced in the case of our depth-driven algorithm.

For a quantitative inspection of our qualitative results, we used the average $L_1$-error of the image irradiance equation [16] (the error in the backprojection of the reconstructed surface), which is given by

$$ E = \frac{1}{|D|} \int_D \left| f(x, y) - R(z_x(x, y), z_y(x, y)) \right| \, dx \, dy. \quad (12) $$

The error rates in 1 confirm our qualitative observations from Figure 2. As one can see, the reconstruction quality of the new algorithms is clearly better than the one of Frankot-Chellappa: Improvements up to 33% with the depth-driven approach are possible.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Error $E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frankot-Chellappa</td>
<td>0.0192</td>
</tr>
<tr>
<td>Homogeneous Regularisation</td>
<td>0.0176</td>
</tr>
<tr>
<td>Image-Driven Regularisation</td>
<td>0.0138</td>
</tr>
<tr>
<td>Depth-Driven Regularisation</td>
<td>0.0127</td>
</tr>
</tbody>
</table>

7 Summary and Conclusions

In this paper we introduced three novel variational approaches for solving the classical shape-from-shading problem. Unlike existing techniques that first compute the surface gradient and then recover the actual surface, all these approaches recover the desired surface directly within a single partial differential equation. Thus additional constraints to enforce the integrability of the solution and the subsequent application of depth-from-gradient algorithms became obsolete. Within this three methods, we developed one approach based on homogeneous regularisation, while the other two adapt their smoothing behaviour either to edges in the original image or edges the evolving depth map. Results show a clear advantage of all three concepts, whereby the image-driven approach yielded the sharpest results, while the depth-driven approach gave the best reconstructions itself.

We hope that by deriving such direct approaches, we can revive the research in the field of variational shape-from-shading methods. As other areas in computer vision show – such as stereo reconstruction or optical flow estimation – there is a large potential in variational methods. This potential has evidently not yet been exploited for the purpose of shape-from-shading.
Fig. 2. Mozart's face. Left to right: Depth map, orthographic projection from above, difference image (scaled by factor 12). Top to bottom: Ground truth, Frankot-Chellappa, homogeneous regularisation, image-driven regularisation, depth-driven regularisation.
References


