

3D Surface Reconstruction by Combination of Photopolarimetry and Depth from Defocus

Pablo d' Angelo and Christian Wöhler

DaimlerChrysler AG Research and Technology, Machine Perception,
P. O. Box 2360, D-89013 Ulm, Germany
{pablo.d.angelo, christian.woehler}@daimlerchrysler.com

Abstract. In this paper we present a novel image-based 3D surface reconstruction technique that incorporates reflectance, polarisation, and defocus information into a variational framework. Our technique is especially suited for the difficult task of 3D reconstruction of rough metallic surfaces. An error functional composed of several error terms related to the measured reflectance and polarisation properties is minimised by means of an iterative scheme in order to obtain a 3D reconstruction of the surface. This iterative algorithm is initialised with the result of depth from defocus analysis. By evaluating the algorithm on synthetic ground truth data, we show that the combined approach strongly improves the accuracy of the surface reconstruction result compared to techniques based on either reflectance or polarisation alone. Furthermore, we report 3D reconstruction results for a raw forged iron surface. A comparison of our method to independently measured ground truth data yields an accuracy of about one third of the pixel resolution.

1 Introduction

A well-known image-based surface reconstruction method is *shape from shading*. This approach aims at deriving the orientation of the surface at each pixel by using a model of the reflectance properties of the surface and knowledge about the illumination conditions – for a detailed survey, cf. [2]. The integration of shadow information into the shape from shading formalism and applications of such methods in the context of industrial quality inspection have been demonstrated in [10].

A further approach to reveal the 3D shape of a surface is to utilise polarisation data. Most current literature concentrates on dielectric surfaces, as for smooth dielectric surfaces, the angle and degree of polarisation as a function of surface orientation are governed by elementary physical laws as detailed in [6]. There, information about the polarisation state of light reflected from the surface is utilised to reconstruct the 3D shape of transparent objects, involving multiple light sources equally distributed over a hemisphere and a large number of images acquired through a linear polarisation filter at many different orientations. This method is not straightforward to use in practice because it requires a somewhat elaborate setting of light sources. For smooth dielectric surfaces, in

[7] a 3D surface reconstruction framework is proposed relying on the analysis of the polarisation state of reflected light, the surface texture, and the locations of specular reflections. In [11] reflectance and polarisation properties of metallic surfaces are examined, but no physically motivated polarisation model is derived. In [9] polarisation information is used to determine surface orientation.

The depth of a scene point can also be estimated by analysis of defocus information (*depth from defocus*, cf. [1] for a detailed survey). This approach makes use of the fact that the point spread function (PSF), i. e. the “blur” at a certain image location, depends on the distance of the corresponding scene point from the camera. It requires a calibration procedure which yields a relation between PSF width and distance.

The three described image-based 3D surface reconstruction methods have in common that corresponding applications to real-world scenarios are rarely described in the literature. A possible reason is the fact that each method alone will not reveal the surface shape in an accurate and unique manner. Shape from shading tends to produce ambiguous results for strongly specular non-Lambertian surfaces even when several images are used. The shape from polarisation approach suffers from the same ambiguity problem and additionally from the lack of a valid polarisation model for non-dielectric, especially metallic, surfaces. Depth from defocus produces a dense solution for the surface shape, which, however, tends to be very noisy.

Hence, we will present in this paper an image-based method for 3D surface reconstruction based on the simultaneous evaluation of reflectance, polarisation, and defocus data. All extracted information will be integrated into a unified variational framework. This method is systematically evaluated based on a synthetically generated surface to examine its accuracy and is applied to the sophisticated real-world example of a raw forged iron surface.

2 Combination of Photopolarimetric and Defocus Information for 3D Surface Reconstruction

In our scenario, we will assume that the surface $z(x, y)$ to be reconstructed is illuminated by a point light source and viewed by a camera, both situated at infinite distance in the directions \mathbf{s} and \mathbf{v} , respectively. Parallel incident light and an orthographic projection model can thus be assumed. For each pixel location (u, v) of the image we intend to derive a depth value $z(u, v)$. The surface normal is given in *gradient space* by the vector $\mathbf{n} = (-p, -q, 1)^T$ with $p = \partial z / \partial x$ and $q = \partial z / \partial y$. In an analogous manner, we set for the illumination direction $\mathbf{s} = (-p_s, -q_s, 1)^T$ and for the viewing direction $\mathbf{v} = (-p_v, -q_v, 1)^T$. The *incidence angle* θ_i is defined as the angle between \mathbf{n} and \mathbf{s} , the *emission angle* θ_e as the angle between \mathbf{n} and \mathbf{v} , and the *phase angle* α as the (constant) angle between the vectors \mathbf{s} and \mathbf{v} .

Estimation of depth from defocus relies on the fact that the image of a scene point situated at a distance z from the camera becomes more and more blurred with increasing depth offset $|z - z_0|$ between the scene point at distance z and

the plane at distance z_0 on which the camera is focussed. The blur is described by the *point spread function* (PSF) of the camera lens, which is assumed to be Gaussian with a width parameter σ .

The calibration procedure for estimating depth from defocus involves the determination of the lens-specific characteristic curve $\sigma(|z - z_0|)$ [1]. For this purpose we acquire two pixel-synchronous images of a rough, uniformly textured plane surface consisting of forged iron, inclined by 45° with respect to the optical axis. The image part in which the intensity displays maximum standard deviation (i. e. most pronounced high spatial frequencies) is sharp and thus situated at distance z_0 . A given difference in pixel coordinates with respect to that image location directly yields the corresponding depth offset $|z - z_0|$. The first image is taken with small aperture, e. g. $f/16$, resulting in virtually absent image blur, while the second image is taken with the aperture that will be used later on for 3D reconstruction, e. g. $f/4$, resulting in a perceivable image blur that depends on the depth offset $|z - z_0|$. In practice, it is desirable but often unfeasible to use the sharp image acquired with small aperture for 3D reconstruction – the image brightness then tends to become too low for obtaining reasonably accurate polarisation data. Hence, the surface reconstruction algorithm must take into account the position-dependent PSF.

The images are partitioned into windows of $n_w \times n_w$ pixels size, for each of which the depth offset $|z - z_0|$ is known. After Tukey windowing, the PSF width parameter σ in frequency space is computed by fitting a Gaussian to the quotient of the amplitude spectra of the corresponding windows of the first and the second image, respectively. Only the range of intermediate spatial frequencies is regarded in order to reduce the influence of noise on the resulting value for σ . This technique and alternative methods are described in [1].

Once the characteristic curve $\sigma(|z - z_0|)$ is known, it is possible to extract a dense depth map from a pixel-synchronous pair of images of a surface of unknown shape, provided that the images are acquired at the same focus position and with the same apertures as the calibration images. The resulting depth map $z_{\text{dfd}}(u, v)$, however, tends to be very noisy. We therefore extract large-scale information about the surface orientation from it by fitting a plane of the form $\tilde{z}_{\text{dfd}}(u, v) = p_{\text{dfd}}u + q_{\text{dfd}}v + c$ to the data. The sign of either p_{dfd} or q_{dfd} must be provided a-priori.

The 3D surface reconstruction formalism we utilise throughout this paper is related to the shape from shading scheme described in detail in [2,3,4]. It relies on the global minimisation of an error function that consists of a weighted sum of several individual error terms. The observed image I is a convolution $G * I_0$ of the “true” image I_0 with the PSF G . The *intensity error* term

$$e_i = \sum_{u,v} [I(u, v) - G * R(\rho(u, v), p(u, v), q(u, v))]^2 \quad (1)$$

describes the mean square deviation between the observed pixel intensity $I(u, v)$ and the modelled reflectance R convolved with the PSF G [5] extracted from the image as described above. It depends on the *surface albedo* $\rho(u, v)$ and the surface gradients $p(u, v)$ and $q(u, v)$, which in turn depend on the pixel coordinates (u, v) .

The polarisation state of light reflected from the surface is determined based on several images of the surface acquired through a linear polarisation filter. The transmitted radiance of the reflected light oscillates sinusoidally with the orientation of the polarisation filter between a maximum I_{\max} and a minimum I_{\min} . The *polarisation angle* $\Phi \in [0^\circ, 180^\circ]$ describes the orientation under which maximum transmitted radiance I_{\max} is observed. The *polarisation degree* D is given by $D = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min}) \in [0, 1]$. To integrate polarisation information into the 3D surface reconstruction framework we define the polarisation angle error term

$$e_\Phi = \sum_{u,v} [\Phi(u,v) - G * R_\Phi(p(u,v), q(u,v), \alpha)]^2 \quad (2)$$

and the polarisation degree error term

$$e_D = \sum_{u,v} [D(u,v) - G * R_D(p(u,v), q(u,v), \alpha)]^2, \quad (3)$$

describing the mean square deviation between the observations Φ and D and the corresponding modelled values R_Φ and R_D , respectively. Again the estimated PSF G is taken into account. As in this paper we will primarily examine rough metallic surfaces of industrial parts (cf. Sect. 4) in the context of industrial quality inspection, for which accurate physically motivated models are neither available for the surface reflectance nor for its polarisation properties, we will empirically determine R , R_Φ , and R_D as described in Sect. 3.

As the intensity and polarisation information alone is not necessarily sufficient to provide an unambiguous solution for the surface gradients $p(u,v)$ and $q(u,v)$, a regularisation constraint e_s is introduced as a further error term, requiring smoothness of the surface, for example small absolute values of the directional derivatives of the surface gradients. We will therefore make use of the error term (cf. also [2,4])

$$e_s = \sum_{u,v} [p_x^2 + p_y^2 + q_x^2 + q_y^2]. \quad (4)$$

The smoothness constraint is acceptable for the surfaces regarded in this paper. For strongly wrinkled surfaces, too smooth results are yielded by (4); it might then be replaced by the *departure from integrability* error term discussed in detail in [3].

The surface gradients $p(u,v)$ and $q(u,v)$ are obtained by minimising the overall error functional $e = e_s + \lambda e_i + \mu e_\Phi + \nu e_D$, where the Lagrange parameters λ , μ , and ν denote the relative weights of the error terms. As described in detail in [4], setting the derivatives of e with respect to p and q to zero yields an iterative update rule for p ,

$$p_{n+1} = \bar{p}_n + \lambda(I - G * R)G * \frac{\partial R}{\partial p} + \mu(\Phi - G * R_\Phi)G * \frac{\partial R_\Phi}{\partial p} + \nu(D - G * R_D)G * \frac{\partial R_D}{\partial p}, \quad (5)$$

with \bar{p} as a local average. An analogous expression is obtained for q . Numerical integration of the gradient field, employing e. g. the algorithm described in [4], yields the surface profile $z(u, v)$.

Despite regularisation constraint (4) there is usually no unique solution for the surface gradients $p(u, v)$ and $q(u, v)$, especially for highly specular reflectance functions, such that the obtained solution tends to depend strongly on the initial values $p_0(u, v)$ and $q_0(u, v)$. As no a-priori information about the surface is available, we make use of the large-scale surface gradients obtained by depth from defocus analysis and set the initial values of the surface gradients to $p_0(u, v) = p_{\text{dfd}}$ and $q_0(u, v) = q_{\text{dfd}}$.

3 Determination of Empirical Photopolarimetric Models

For the purpose of determination of empirical reflectance and polarisation models for the surface material the surface normal \mathbf{n} of a flat sample is adjusted by means of a goniometer, while the illumination direction \mathbf{s} and the viewing direction \mathbf{v} are constant over the image. Intensity I , polarisation angle Φ , and polarisation degree D are determined over a wide range of surface normals \mathbf{n} .

According to [8], the reflectance of a typical rough metallic surface consists of three components: a diffuse (Lambertian) component, the *specular lobe*, and the *specular spike*. We model these components by the phenomenological approach

$$R(\theta_i, \theta_e, \alpha) = \rho \left[\cos \theta_i + \sum_{n=1}^N \sigma_n \cdot (2 \cos \theta_i \cos \theta_e - \cos \alpha)^{m_n} \right] \quad (6)$$

with $2 \cos \theta_i \cos \theta_e - \cos \alpha \equiv \cos \theta_r$ describing the angle between the specular direction \mathbf{r} and the viewing direction \mathbf{v} . For $\theta_r > 90^\circ$ only the diffuse component proportional to $\cos \theta_i$ is considered. The albedo ρ is assumed to be constant over the image. The shapes of the two specular components are expressed by $N = 2$ terms proportional to powers of $\cos \theta_r$, where the coefficients $\{\sigma_n\}$ denote the strength of the specular components relative to the diffuse component and the parameters $\{m_n\}$ their widths.

The polarisation angle Φ is phenomenologically modelled by an incomplete third-degree polynomial in p and q according to

$$R_\Phi(p, q) = a_\Phi p q + b_\Phi q + c_\Phi p^2 q + d_\Phi q^3. \quad (7)$$

Without loss of generality we assume illumination from the right hand side ($p_s < 0, q_s = 0$) and view along the z axis ($p_v = q_v = 0$). Equation (7) is antisymmetric in q , and $R_\Phi(p, q) = 0$ for $q = 0$, i. e. coplanar vectors \mathbf{n} , \mathbf{s} , and \mathbf{v} . These properties are required for geometrical symmetry reasons as long as an isotropic interaction between the incident light and the surface material can be assumed. The polarisation degree D is modelled by an incomplete second-degree polynomial in p and q according to

$$R_D(p, q) = a_D + b_D p + c_D p^2 + d_D q^2. \quad (8)$$

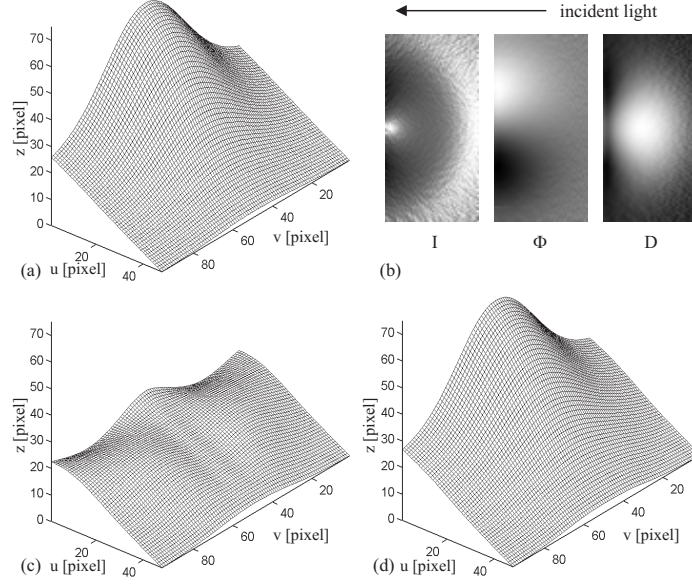


Fig. 1. 3D reconstruction of synthetic scene. (a) Ground truth. (b) Intensity I , polarisation angle Φ , and polarisation degree D images. The 3D reconstruction result was obtained based on photopolarimetric analysis (c) without and (d) with depth from defocus information. For details refer to Sect. 4.

For rough metallic surfaces, R_D is maximum near the direction of specular reflection. Symmetry in q is imposed to account for isotropic light-surface interaction.

4 Experimental Results

To examine the accuracy of 3D reconstruction we applied the algorithm described in Sect. 2 to the synthetically generated surface shown in Fig. 1a. For the PSF width parameter we assumed $\sigma \propto |z - z_0|^{-1}$ [1]. To generate a visible surface texture in the photopolarimetric images, we introduced random fluctuations of $z(u, v)$ of the order 0.1 pixels. We assumed a perpendicular view on the surface along the z axis with $\mathbf{v} = (0, 0, 1)^T$. The scene was illuminated from the right hand side at a phase angle of $\alpha = 75^\circ$. The surface albedo ρ was computed from (6) based on the specular reflections, which appear as regions of maximum intensity and for which we have $\theta_r = 0^\circ$ and $\theta_i = \alpha/2$. We have set $p_0(u, v) = p_{\text{dfd}}$ and $q_0(u, v) = q_{\text{dfd}}$ in (5) when depth from defocus information was used, while otherwise, the PSF G was set to unity and the surface gradients were initialised with zero values due to the lack of prior information. We found that for the sort of surface regarded here, the influence of G on $z(u, v)$ is negligible and that the main benefit of depth from defocus analysis comes from the improved initialisation. The best results are obtained by utilising a combination of polarisation angle and degree, of reflectance and polarisation angle, or a combination of all three features. Detailed evaluation results are reported in Table 1.

Table 1. Evaluation results on synthetic ground truth data

Utilised information	RMS error (without dfd)			RMS error (with dfd)		
	z	p	q	z	p	q
Reflectance	11.6	0.620	0.514	9.62	0.551	0.514
Pol. angle	17.0	0.956	0.141	6.62	0.342	0.069
Pol. degree	4.72	0.138	0.514	4.73	0.135	0.514
Pol. angle and degree	1.83	0.121	0.057	1.71	0.119	0.056
Reflectance and pol. angle	12.0	0.528	0.099	2.52	0.280	0.055
Reflectance and pol. degree	10.9	0.575	0.514	8.46	0.418	0.517
Reflectance and polarisation (angle and degree)	10.2	0.277	0.072	0.91	0.091	0.050

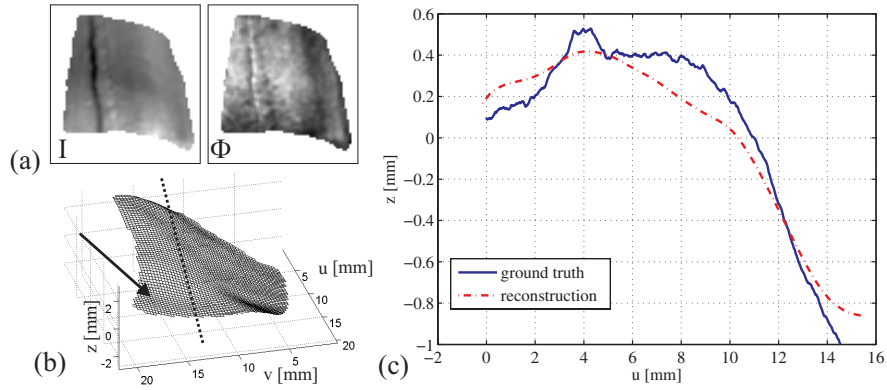


Fig. 2. 3D reconstruction of a forged iron surface. (a) Intensity and polarisation angle image. (b) 3D reconstruction result. The somewhat malformed, asymmetric surface shape is due to a fault caused during the forging process (arrow). The cross-section used in (c) is indicated by a dashed line. (c) Comparison to ground truth.

We furthermore applied our 3D surface reconstruction method to a raw forged iron part. According to Sect. 2 we empirically determined $\sigma \propto |z - z_0|^{-0.66}$ as the relation between PSF width parameter and depth offset. We found that the polarisation degree for specular reflection tends to vary over the surface by up to 20 percent, depending on the locally variable microscopic roughness. In contrast, the behaviour of the polarisation angle turned out to be very stable over the surface. We thus performed 3D reconstruction based on a combination of reflectance, polarisation angle, and depth from defocus (Fig. 2a and b). Image resolution was 0.30 mm per pixel. A cross-section of the same surface was measured with a laser focus profilometer at an accuracy of better than $1 \mu\text{m}$ (Fig. 2c), thus serving as ground truth. The RMS deviation between ground truth and the corresponding cross-section extracted from the 3D profile reconstructed with our approach amounts to 0.11 mm or about one third of the pixel resolution.

5 Summary and Conclusion

In this paper we have presented an image-based 3D surface reconstruction technique that incorporates reflectance, polarisation, and defocus information into a variational framework. Our method relies on an error functional composed of several error terms related to the measured reflectance and polarisation properties, which is minimised by means of an iterative scheme in order to obtain a 3D reconstruction of the surface. The iteration is initialised with the result of depth from defocus analysis. By evaluating the algorithm on synthetic ground truth data we show that the combined approach strongly increases the accuracy of the surface reconstruction result, compared to 3D reconstruction based on either reflectance or polarisation alone. Furthermore, we report 3D reconstruction results for a raw forged iron surface, demonstrating a reconstruction accuracy of our approach of about one third of the pixel resolution.

References

1. S. Chaudhuri, A. N. Rajagopalan. *Depth From Defocus: A Real Aperture Imaging Approach*. Springer-Verlag, New York, 1999.
2. B. K. P. Horn, M. J. Brooks. *Shape from Shading*. MIT Press, Cambridge, Massachusetts, 1989.
3. B. K. P. Horn. *Height and Gradient from Shading*. MIT technical report 1105A. <http://people.csail.mit.edu/people/bkph/AIM/AIM-1105A-TEX.pdf>
4. X. Jiang, H. Bunke. *Dreidimensionales Computersehen*. Springer-Verlag, Berlin, 1997.
5. M. V. Joshi, S. Chaudhuri. *Photometric Stereo Under Blurred Observations*. *Int. Conf. on Pattern Recognition*, Cambridge, UK, 2004.
6. D. Miyazaki, M. Kagesawa, K. Ikeuchi. *Transparent Surface Modeling from a Pair of Polarization Images*. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 73-82, 2004.
7. D. Miyazaki, R. T. Tan, K. Hara, K. Ikeuchi. *Polarization-based Inverse Rendering from a Single View*. *IEEE Int. Conf. on Computer Vision*, vol. II, pp. 982-987, Nice, France, 2003.
8. S. K. Nayar, K. Ikeuchi, T. Kanade. *Surface Reflection: Physical and Geometrical Perspectives*. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, 1991.
9. S. Rahmann, N. Canterakis. *Reconstruction of Specular Surfaces using Polarization Imaging*. *Int. Conf. on Computer Vision and Pattern Recognition*, vol. I, pp. 149-155, Kauai, USA, 2001.
10. C. Wöhler, K. Hafezi. *A general framework for three-dimensional surface reconstruction by self-consistent fusion of shading and shadow features*. *Pattern Recognition*, vol. 38, no. 7, pp. 965-983, 2005.
11. L. B. Wolff. *Constraining Object Features Using a Polarization Reflectance Model*. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, 1991.