Overcoming shadows in 3-source photometric stereo

Carlos Hernández, *Member, IEEE,* George Vogiatzis, *Member, IEEE,* and Roberto Cipolla, *Member, IEEE*

Abstract—Light occlusions are one of the most significant difficulties of photometric stereo methods. When three or more images are available without occlusion, the local surface orientation is overdetermined so shape can be computed and the shadowed pixels can be discarded. In this paper we look at the challenging case when only two images are available without occlusion, leading to a 1 degree of freedom ambiguity per pixel in the local orientation. We show that, in the presence of noise, integrability alone cannot resolve this ambiguity and reconstruct the geometry in the shadowed regions. As the problem is ill-posed in the presence of noise, we describe two regularization schemes that improve the numerical performance of the algorithm while preserving the data. Finally the paper describes how this theory applies in the framework of color photometric stereo where one is restricted to only three images and light occlusions are common. Experiments on synthetic and real image sequences are presented.

Index Terms—photometric stereo, shadows.

I. INTRODUCTION

Photometric stereo is a well established 3d reconstruction technique based on the powerful shading cue. A sequence of images (typically three or more) of a 3d scene are obtained from the same viewpoint and under varying illumination. Assuming a Lambertian reflectance model, one can estimate the local orientation of the surface that projects onto that pixel from the intensity variation in each pixel. By integrating all these surface orientations, a very detailed estimate of the surface geometry can be obtained. As with any other reconstruction method, photometric stereo faces several difficulties when dealing with real images. One of the most important of these difficulties is the frequent presence of shadows in an image. No matter how careful the arrangement of the light sources, shadows are an almost unavoidable phenomenon, especially in objects with complex geometries. This paper investigates the phenomenon of shadows in photometric stereo with three light sources.

Shadows in photometric stereo have been the topic of a number of papers [1], [2], [3]. Most papers assume we are given four or more images under four different illuminations. This over-determines the local surface orientation and albedo (3 degrees of freedom) which implies that we can use the residual of some least squares solution, to determine whether shadowing has occurred. However when we are only given three images there are no *spare* constraints against which to test our hypothesis. Therefore the problem of detecting shadows becomes more difficult. Furthermore, when a pixel

C. Hernández and G. Vogiatzis are with Toshiba Resrearch Europe Ltd, 208 Cambridge Science Park, Cambridge, CB4 0GZ, UK.
E-mail: carlos.hernandez@crl.toshiba.co.uk
E-mail: george.vogiatzis@crl.toshiba.co.uk
R. Cipolla is with Cambridge University Engineering Department.
E-mail: cipolla@eng.cam.ac.uk

is in shadow in one of the three images most methods simply discard it. In this work we show how one can use the remaining two image intensity measurements to estimate the surface geometry inside the shadow region. Using an argument based purely on counting degrees of freedom and equations, this is theoretically possible since we need to estimate 2 DOF per pixel (depth and albedo) and we have two independent measurements per pixel. The solution we propose is based on enforcing (1) integrability of the gradient field, as well as (2) smoothness in the recovered shape.

Using photometric stereo on just three images may seem like an unreasonably hard restriction. There is however a particular situation when only three images are available. This technique is known as color photometric stereo [4] and it uses three light sources with different light spectra. When the scene is photographed with a color camera, the three color channels capture three different photometric stereo images. Because shape acquisition is performed on each frame independently, the method can be used on video sequences without having to change illumination between frames [5]. In this way we can capture the 3D shape of deforming objects such as cloth, or human faces. Since the method is constrained to operate only on three images, it is an ideal application of the theory we present here. Summarizing, the main contributions of this paper are the following:

1) We show how to exploit image regions in photometric stereo where one of the three images is in shadow. A geometric formulation of the problem is given where a set of point-to-point and point-to-line distances are minimized under the integrability condition. The integrability condition is implicitly enforced by parametrizing the surface as a height field, as opposed to our previous work [6] where the surface was parametrized by a gradient field and thus integrability had to be explicitly enforced during optimization.
2) We develop two regularization schemes that make the optimization problem well posed while not suppressing the data.
3) We apply the technique to color photometric stereo. The regularization schemes have been validated in a practical capture setup by running them on thousands of frames of captured video.

A. Previous work

A vast literature exists on the topic of photometric stereo. Its applications range from 3D reconstruction [7], medical imaging [8] or cloth modeling [5]. One way of characterizing photometric stereo methods is based on the number of different lights required and how they cope with highlights or shadows.

A minimum of 3 lights is required to perform photometric stereo with no extra assumptions [7], and only 2 lights with

the additional assumption of constant albedo [9]. Whenever more lights are available, the light visibility problem becomes a labeling problem where each point on the surface has to be assigned to the correct set of lights in order to successfully reconstruct the surface.

For objects with constant albedo, [3] used a Rank-2 constraint to detect surfaces illuminated by only 2 lights. In the case of general albedo, every point on the surface has to be visible in at least 3 images. A 4-light photometric stereo setup was proposed in [10], where light occlusion was detected by checking the consistency of all the possible triplets of lights. The work by [11] was able to detect light occlusions in a 4-light setup and simply treat them as outliers. In [1] a similar algorithm to [10] is presented using a 4-light colored photometric stereo approach.

In the recent work by [2], an iterative MRF formulation is proposed for detecting light occlusion and exploiting it as a surface integration constraint. However, the algorithm also requires a minimum of 4 lights and is targeted for setups with a large number of lights. Similarly, previous work on shape-from-shadows [12], [13] exploit the shadows for 3d reconstruction. However, they need large quantities of images in order to get acceptable results since the shadows are the cue, while in this work the shadows are considered an artifact that we need to correct for.

In this paper we propose a novel solution for 3-light photometric stereo with shadows and varying albedo. We are able to detect and exploit photometric stereo constraints with only two lights while imposing smooth shape priors that are specific to our problem. Compared to our previous work [6], we have two new contributions. The first one is that we implicitly enforce the integrability constraint by parametrizing the surface as a height field. As a second contribution, we use the shape regularization scheme described in section IV-B for the first time in a least squares framework to perform two source photometric stereo and three-source photometric stereo in the presence of shadows, both with varying albedo.

II. THREE-SOURCE PHOTOMETRIC STEREO WITH SHADOWS

In classic three-source photometric stereo we are given three images of a scene, taken from the same viewpoint, and illuminated by three distant light sources. The light sources emit the same light frequency spectrum from three different non-coplanar directions. We will assume an orthographic camera (with infinite focal length) for simplicity, even though the extension to the more realistic projective case is straightforward [14]. In the case of orthographic projection one can align the world coordinate system so that the xy plane coincides with the image plane while the z axis corresponds to the viewing direction. The surface in front of the camera can then be parametrized as a height function $z(x, y)$. If ∇z is the gradient of the function wit respect to x and y, one can define the vector

$$\mathbf{n} = \frac{1}{\sqrt{1 + |\nabla z|^2}} \begin{pmatrix} \nabla z \\ -1 \end{pmatrix},$$

that is locally normal to the surface at (x, y). We can also define a 2d projection operator $\mathcal{P}[\mathbf{x}] = (x_1/x_3,\ x_2/x_3)$ so

that it follows that $\nabla z = \mathcal{P}[\mathbf{n}]$.

Now for $i = 1 \ldots 3$ let $c_i(x, y)$ denote the pixel intensity of pixel (x, y) in the i-th image. We assume that, in the i-th image, the surface point $(x, y, z(x, y))^\top$ is illuminated by a distant light source whose direction is denoted by the vector $\mathbf{l}_i$ and whose spectral distribution is $E_i(\lambda)$. We also assume that the surface point absorbs incoming light of various wavelengths according to the reflectance function $R(x, y, \lambda)$. Finally, let the response of the camera sensor at each wavelength be given by $S(\lambda)$. Then the pixel intensity $c_i(x, y)$ is given by [4]

$$c_i(x, y) = \left(\mathbf{l}_i^\top \mathbf{n}\right) \int E(\lambda) R(x, y, \lambda) S(\lambda)\, d\lambda. \quad (1)$$

The value of this integral is known as the surface *albedo* ρ so that (1) becomes a simple dot product

$$c_i = \mathbf{l}_i^\top \rho \mathbf{n}. \quad (2)$$

Photometric stereo methods use the linear constraints of (2) to solve for $\rho\mathbf{n}$ in a least squares sense. From this they obtain the gradient of the height function $\nabla z = \mathcal{P}[\rho\mathbf{n}]$ which is then integrated to produce the function z itself. In three-source photometric stereo, when the point is not in shadow with respect to all three lights, we measure three positive intensities c_i, each of which gives a constraint on $\rho\mathbf{n}$. If we write $L = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 & \mathbf{l}_3 \end{bmatrix}^\top$ and $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}^\top$ then the system has exactly one solution which is given by

$$\rho\mathbf{n} = L^{-1}\mathbf{c}. \quad (3)$$

If the point however is in shadow, say in the 3-rd image, then the measurement of c_3 cannot be used as a constraint. Since each equation (2) describes a 3d plane, the intersection of the two remaining constraints is a 3d line given by

$$(c_2\mathbf{l}_1 - c_1\mathbf{l}_2)^\top \mathbf{n} = 0. \quad (4)$$

More generally, if the point is in shadow in the i-th image, (4) can be rearranged as

$$[\mathbf{c}]_\times^i L\mathbf{n} = 0, \quad (5)$$

or equivalently

$$\mathcal{P}[[\mathbf{c}]_\times^i L]\nabla z = 1, \quad (6)$$

where $[\mathbf{c}]_\times^i$ is the i-th row of the cross product matrix $[\mathbf{c}]_\times$, *i.e.* $[\mathbf{c}]_\times^1 = [0, c_3, -c_2]^\top$, $[\mathbf{c}]_\times^2 = [-c_3, 0, c_1]^\top$ and $[\mathbf{c}]_\times^3 = [c_2, -c_1, 0]^\top$. Equation 6 was derived by [15] and used for stereo matching in a two-view photometric stereo setup, and subsequently used by [16] to perform uncalibrated photometric stereo and by [17] in their proof of nonexistence of a general illumination invariant. In this paper we use this equation for the first time in a least squares framework to perform three-source photometric stereo in the presence of shadows.

III. INTEGRATING IN THE SHADOWED REGIONS

According to the image constraints and assuming no noise in the data, we can have one of the following three cases (see Fig. 1):

1) The surface point is in shadow in two or more images. In this case there is no constraint in ∇z from the images.
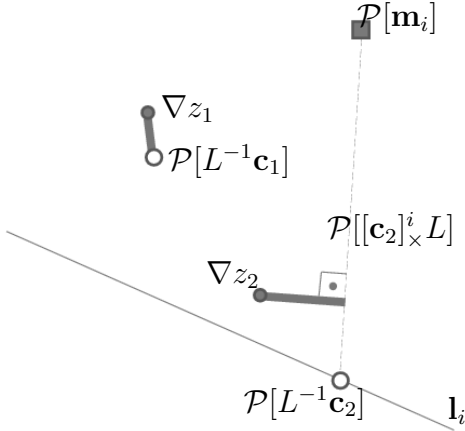
Fig. 1. **Geometry of shadowed pixels**. The points ∇z_j (dark dots) represent the partial derivatives of the height function at pixel j. For each point ∇z_j there is a corresponding data point $\mathcal{P}[L^{-1}\mathbf{c}]$ (white dot). Pixel 1 is unoccluded and hence ∇z_1 must be as close as possible to its data point $\mathcal{P}[L^{-1}\mathbf{c}_1]$. Pixel 2 however is occluded so ∇z_2 must be as close as possible to its shadow line $\mathcal{P}[[\mathbf{c}_2]^i_\times L]$. Note that all the shadow lines cross at a single point $\mathcal{P}[\mathbf{m}_i]$.

2) The surface point is not in shadow in any of the three images. In this case ∇z coincides with $\mathcal{P}\left[L^{-1}\mathbf{c}\right]$.
3) The surface point is in shadow in exactly one image, say the i-th. In this case ∇z must lie on the line $\mathcal{P}[[\mathbf{c}]^i_\times L]\nabla z = 1$. We call this line the *shadow line* of the shaded pixel.

Now in the presence of noise in the data $\mathbf{c}$, cases 2 and 3 above do not hold exactly as $\mathcal{P}\left[L^{-1}\mathbf{c}\right]$ and $\mathcal{P}[[\mathbf{c}]^i_\times L]$ are corrupted. The estimation of the unknown height function z becomes a least squares problem with two different data terms, one for pixels under shadow and another one for pixels seen in all three images. For non-shadowed pixels, the difference between model and data can be measured by the *point-to-point* square difference term

$$\mathcal{E} = |\nabla z - \mathcal{P}\left[L^{-1}\mathbf{c}\right]|^2. \tag{7}$$

In the case of the shadowed pixels we have a *point-to-line* square difference term

$$\overline{\mathcal{E}}^{(i)} = (\mathcal{P}[[\mathbf{c}]^i_\times L]\nabla z - 1)^2, \tag{8}$$

where $\overline{\mathcal{E}}^{(i)}$ denotes the error term for pixels shaded in the i-th image.

Assume we are given a labeling of pixels into all the possible types of shadow. Let $\mathcal{S}$ contain all non-shadowed pixels while $\mathcal{S}_i$ contains pixels shaded in the i-th image. Our cost function becomes

$$\sum_{j \in \mathcal{S}} \mathcal{E}_j + \sum_{j \in \mathcal{S}_1} \overline{\mathcal{E}}^{(1)}_j + \sum_{j \in \mathcal{S}_2} \overline{\mathcal{E}}^{(2)}_j + \sum_{j \in \mathcal{S}_3} \overline{\mathcal{E}}^{(3)}_j,$$

which is a set of quadratic terms in ∇z and thus z. Finding the minimum of this quantity is a simple unconstrained linear least squares problem that can be solved using a sparse linear solver such as UMFPACK [18].

Figure 2 shows this idea applied in practice on synthetic data. It provides evidence that in its present form the problem is ill-conditioned, especially in larger shadowed regions (see Fig. 2c). The following section sheds more light on this and describes our proposed remedy (see figures 2d and 2e).

IV. REGULARIZATION SCHEMES

The linear least squares optimization framework described in section II when executed in practice shows signs of ill-posedness in the presence of noise. This is demonstrated in the synthetic case of figure 2 where three images of a sphere have been generated. Three shadow regions corresponding to each of the three lights have been introduced. Even though the overall shape of the object is accurately captured, some characteristic 'scratch' artifacts are observed. These are caused by the point-to-line distances which do not introduce enough constraints in the cost function. The point ∇z can move significantly in a direction parallel to the corresponding shadow line only to gain a slight decrease in the overall cost. This results in violent perturbations in the resulting height function that manifest themselves as deep scratches that follow the 2d flow $\mathcal{P}[[\mathbf{c}]^i_\times L]$.

If we push the analysis even further and have one of the images completely shadowed, we then fall back to the two-source photometric stereo setup shown in Fig. 3. When only two images are available without shadow (see Fig. 3 top), after factoring out the albedo (5) we can only determine the depth gradient along specific directions for each pixel $\mathcal{P}[[\mathbf{c}]^i_\times L]$. If we look at these directions as a vector field, then depth can be computed independently along each streamline or "characteristic curve" (see Fig. 3b). In other words, there is no constraint between the depth of two characteristic curves and one pixel can only belong to a single characteristic curve. After integrating every characteristic curve independently (see Fig. 3c), we obtain a possible reconstruction that is different from the original true shape, but that perfectly agrees with the given constraints. In order to choose one among the possible solutions, some type of regularization is needed [19](see Fig. 3d and 3e).

The regularization can be seen as a simple prior on the type of solutions we expect. In order to understand the types of solution we can expect, we describe our problem once more. We have a three source photometric stereo with varying albedo setup, *and* one of the lights is occluded, *i.e.* we locally have a two source photometric stereo setup with varying albedo. From the theory we know that in the photometric stereo setup, the albedo and the geometry are coupled, and if there is not enough data available, both are indistinguishable. This coupling exactly indicates what two types of priors one would expect to use: either a shape smoothness prior favoring smooth shapes or an albedo smoothness prior favoring smooth albedo. In this work we choose a shape smoothness prior since we will be dealing with smooth objects with high frequency albedo such as faces or cloth. We also have two main requirements on the choice of regularizing criterion:

- The scheme must be consistent with the linear least squares framework. No non-linear constraints can be enforced.
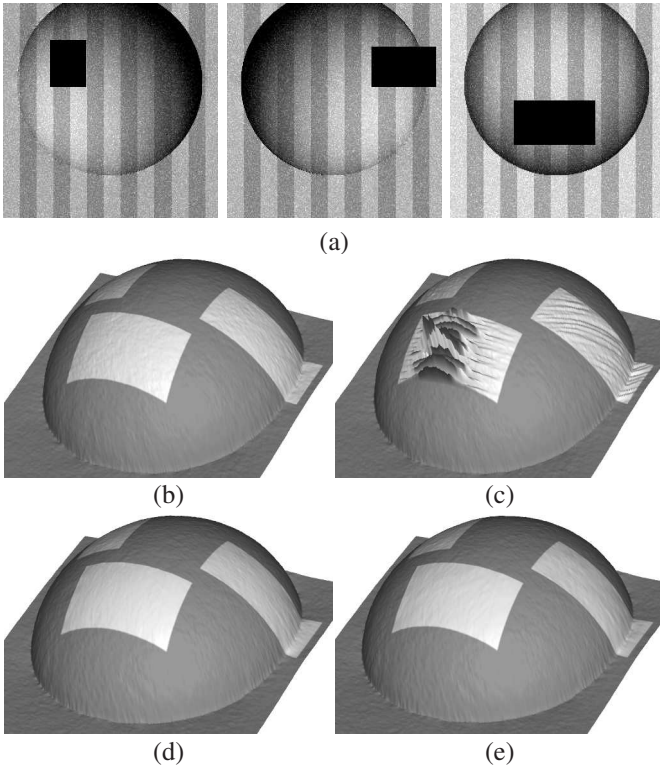- It must suppress noise while preserving as much of the data as possible.

(a)

(b)   (c)

(d)   (e)

Fig. 2. **Regularization schemes**. This is an experiment on a synthetic sphere designed to validate the proposed regularization constraints. **(a)** shows the input images where the black rectangles correspond to occluded regions.This object is illuminated from three directions and the three white regions are occluded in the corresponding images. Middle row shows the photometric stereo solution without shadows **(b)** and the effect of optimizing the surface with no regularization at all,*i.e.* just using integrability **(c)**. Note the char- acteristic 'scratch' artifacts. **(d)** shows the resulting surface after adding a shading regularization term with default values $\alpha = 0.2, \beta = 0$. **(e)** shows the resulting surface after adding a shape regularization term with default values $\alpha = 0.15, \beta = 1$. See Section IV for a description of the algorithms. The artifacts have been suppressed while the data has been preserved unsmoothed. Note how both regularization schemes give almost identical results.

In the following we describe two different regularization schemes that favor smooth shapes while preserving the data as much as possible. Their main difference is that one favors shapes with a *smooth shading* under the occluded light, while the second one favors *smooth shapes*. The second scheme can be used in a two-source photometric stereo setup as it is independent of the occluded third light (see Fig. 3).

A. Shading regularization

In this approach we want to impose regularization on the collected shading intensities, thereby "inpainting" [20] the shadowed regions in order to recover the intensities we would collect had the light not been occluded and the albedo been constant. From equation (3) we can parametrize the shadow line as a function of the missing shading μ

$$\nabla z = \mathcal{P}\left[L^{-1}\begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix} + \mu\rho L^{-1}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right]. \quad (9)$$

This parameter represents the value $\mathbf{l}_3^\top \mathbf{n}$ would have, had the point not been in shadow in the 3-rd image. In order to
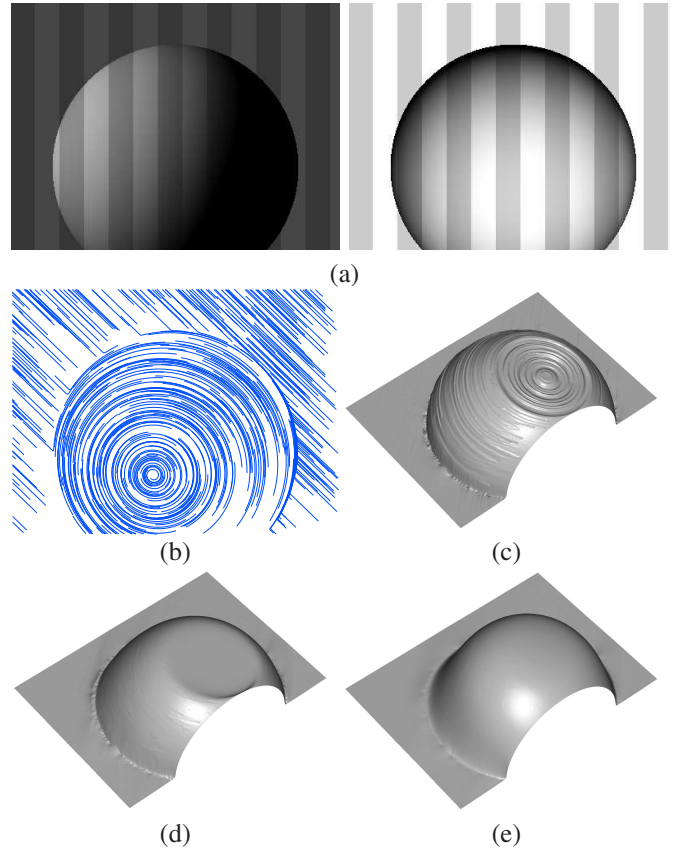

(a)

(b)   (c)

(d)   (e)

Fig. 3. **Two-source varying albedo photometric stereo setup**. In this experiment we show a two-source photometric stereo with varying albedo. **(a)** shows the two input images. **(b)** shows the characteristic curves obtained by plotting seeds following the 2d flow $\mathcal{P}[[\mathbf{c}]_\times^3 L]$. **(c)** shows one possible reconstruction of the characteristic curves. Note how each characteristic curve is reconstructed independently as there is no constraint "across" the curves. Bottom row shows how a successful reconstruction can be achieved when using the proposed shape regularization scheme with first order regularization $\alpha = 0.1, \beta = 0$ **(d)** and second order regularization $\alpha = 0, \beta = 0.5$ **(e)**.

simplify the notation of (9) we define matrix $M = L^{-1}$ where vector $\mathbf{m}_i$ is the ith column of matrix M, giving

$$\nabla z = \mathcal{P}[c_1\mathbf{m}_1 + c_2\mathbf{m}_2 + \mu\rho\mathbf{m}_3]. \quad (10)$$

We observe that, because c_1 and c_2 already encode the albedo ρ in, equation (10) is in fact independent of ρ due to the projection operator. We also note that ∇z is not a linear function of μ meaning that we cannot directly regularize the missing shading μ in a linear least squares framework. However, we can perform a change of variables and introduce a new variable w per shaded pixel

$$w(\mu) = \frac{\mathbf{e}_3^\top(c_1\mathbf{m}_1 + c_2\mathbf{m}_2)}{\mathbf{e}_3^\top(c_1\mathbf{m}_1 + c_2\mathbf{m}_2) + \mu\rho\mathbf{e}_3^\top\mathbf{m}_3}, \quad (11)$$

with $\mathbf{e}_3 = (0,0,1)^\top$. The new variable w still specifies a location along the shadow line of that pixel so equation (10) simply becomes

$$\nabla z = w\mathcal{P}[c_1\mathbf{m}_1 + c_2\mathbf{m}_2] + (1-w)\mathcal{P}[\mathbf{m}_3]. \quad (12)$$

The term is now quadratic with respect to ∇z and w, allowing us to regularize the solution in a meaningful way by using first order $|\nabla w|$ and second order $|\nabla^2 w|$ regularization terms on

w. The point-to-line distance of (8) can now be replaced with the following point-to-point distance

$$\overline{\mathcal{E}}^{(3)} = \ \ |\nabla z - w\mathcal{P}[c_1\mathbf{m}_1 + c_2\mathbf{m}_2] - (1-w)\mathcal{P}[\mathbf{m}_3]|^2 + \\ \alpha|\nabla w|^2 + \beta|\nabla^2 w|^2, \tag{13}$$

where α and β are regularization weights. As w is a proxy for μ, this corresponds to introducing smoothness in the product $\mathbf{l}_3^\top \mathbf{n}$. Since both α and β compare two quantities with different scales, w is in the range $[0,1]$ while ∇z is in the range of the size of the image, we scale both regularization terms by the size of the images in order to define normalized regularization parameters that are independent of the image size.

B. Shape regularization

The most common way of regularizing shape is using first-order and second-order regularization terms. In the context of a height field, this is achieved by minimizing the norm of the gradient of the height field $|\nabla z|$ or minimizing the Laplacian of the height field $|\nabla^2 z|$. The latter is known to have good noise reduction properties and to produce smooth well behaved surfaces with low curvature. However, both the gradient and the Laplacian are isotropic so they tend to indiscriminately smooth along all possible directions. See [21] for a good discussion of anisotropic alternatives to Laplacian filtering in the context of gradient field integration. In the context of our problem, there is an efficient way of achieving anisotropic versions of both the first-order and the second-order regularization terms. From equation (6), we observe that the shape is totally unconstrained along perpendicular directions to $\mathcal{P}[[\mathbf{c}]_\times^i L]$. The directions $\mathcal{P}[[\mathbf{c}]_\times^i L]$ define characteristic curves, visually showing the constraint induced by the two lights (see Fig. 3b). Therefore a good way of regularizing the shape is along directions $\mathbf{u}$ that are perpendicular to the characteristic curves, *i.e.* $\mathbf{u}^\top \mathcal{P}[[\mathbf{c}]_\times^i L] = 0$. This regularization will effectively *tie* together the characteristic curves while minimizing the distortion of the curves themselves. The point-to-line distance term (8) is therefore extended with anisotropic first and second order regularization terms

$$\overline{\mathcal{E}}^{(i)} = (\mathcal{P}[[\mathbf{c}]_\times^i L]\nabla z - 1)^2 + \alpha|\mathbf{u}^\top \nabla z|^2 + \beta|\mathbf{u}^\top H(z)\mathbf{u}|^2, \tag{14}$$

α and β being the regularization weights and $H(z)$ the Hessian matrix.

V. COLOR PHOTOMETRIC-STEREO

It may seem that a photometric stereo scheme with three images is unnecessarily restrictive. The overall cost in practice of acquiring one more image is small compared to the rest of the process (calibration, darkening the environment, changing the illumination etc). In this section we examine color photometric stereo [4]. This is a setup where it is not possible to obtain more than three images. The key observation is that in an environment where red, green, and blue light is simultaneously emitted from different directions, a Lambertian surface will reflect each of those colors simultaneously without any mixing of the frequencies. The quantities of red, green and

blue light reflected are a linear function of the surface normal direction. A color camera can measure these quantities from a single RGB image. Recently [5] it was shown how this idea can be used to obtain a reconstruction of a deforming object. Because color photometric stereo is applied on a single image, one can use it on a video sequence without having to change illumination between frames. In [5] shadowed pixels were detected and discarded. Here we show how to improve that method by incorporating shadow regions into the reconstruction. In color photometric stereo each of the three camera sensors can be seen as one of the three images of classic photometric stereo. The pixel intensity of pixel (x, y) for the i-th sensor is given by

$$c_i(x, y) = \sum_j \left(\mathbf{l}_j^\top \mathbf{n}\right) \int E_j(\lambda) R(x, y, \lambda) S_i(\lambda) d\lambda. \tag{15}$$

Note that now the sensor sensitivity S_i and spectral distribution E_j are different per sensor and per light source respectively. To be able to determine a unique mapping between RGB values and normal orientation we need to assume a monochromatic surface. We therefore require that $R(x, y, \lambda) = \rho(x, y)\alpha(\lambda)$. Where $\rho(x, y)$ is the monochromatic albedo of the surface point and $\alpha(\lambda)$ is the characteristic chromaticity of the material. Let

$$v_{ij} = \int E_j(\lambda)\alpha(\lambda) S_i(\lambda) d\lambda,$$

and

$$\mathbf{v}_j = \left(\begin{array}{ccc} v_{1j} & v_{2j} & v_{3j} \end{array} \right)^\top.$$

Then the vector of the three sensor responses at a pixel is given by

$$\mathbf{c} = \left[\begin{array}{ccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{array} \right] \left[\begin{array}{ccc} \mathbf{l}_1 & \mathbf{l}_2 & \mathbf{l}_3 \end{array} \right]^\top \rho\mathbf{n}.$$

Essentially each vector $\mathbf{v}_j$ provides the response measured by the three sensors when a unit of light from source j is received by the camera. If matrix $\left[\begin{array}{ccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{array} \right]$ is known, then we can compute

$$\hat{\mathbf{c}} = \left[\begin{array}{ccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{array} \right]^{-1} \mathbf{c}.$$

The values of $\hat{\mathbf{c}}$ can be treated in exactly the same way as the three gray-scale images of section (II).

VI. EXPERIMENTS

We present two synthetic experiments and four real experiments that validate our approach. In figure 2 we study the effect of the proposed framework to automatically correct light occlusions on a half sphere with varying albedo and 10% Gaussian noise. Figure 2b shows the ground truth photometric stereo reconstruction of the sphere in the absence of shadows. As soon as we introduce the shadows, the recovered shape in Fig. 2c shows some characteristic artifacts due to an almost unconstrained variation of ∇z along the shadow lines in (6). These artifacts basically show that the recovered shape and albedo are coupled and integrability constraints on their own are not enough to separate them when one intensity constraint is missing. Introducing the regularization schemes of section
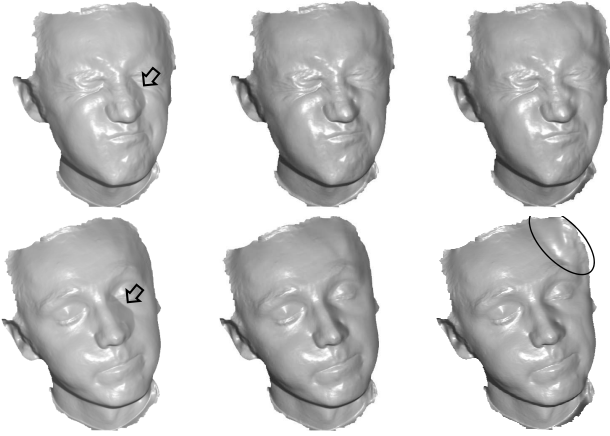
Fig. 5. **Failure case of the shape regularization scheme**. The figures correspond to the bottom face in Fig.4. Left shows characteristic curves describing the light occlusion on the right-side of the face. Middle and right show the rendering of the shape under the occluded light using the shading regularization scheme (middle) and the shape regularization scheme (right). The failure of the shape regularization scheme is clearly visible at the top right of the image.

Fig. 4. **Face sequence**. Two different frames out of a 1000 frame face video sequence. The left column shows the reconstruction when shadows are ignored. Middle and right columns show the corresponding reconstructions after detecting and compensating for the shadow regions using the shading regularization scheme (middle) and shape regularization scheme (right). Note the improvement in the regions around the nose reconstruction where strong cast shadows appear (see arrows). Note also how the shape regularization scheme fails to reconstruct some boundary regions (see circle). This behavior is further explained in Fig. 5.

IV adds priors on the expected shading (Fig. 2d) or on the expected shape Fig. 2e). This helps recovering the correct shape while minimizing the loss of information. In order to quantify the shape error against ground truth, we compute the Root Mean Square Error between the normals of the two surfaces in degrees. Note that both solutions give very similar results, with the shading regularization giving an RMSE of 3.23 degrees while the shape regularization gives an RMSE of 3.17 degrees.

This experiment shows a surprising result concerning the parameters of the shading regularization scheme that is further confirmed in Fig. 7. Namely that the second order regularization parameter β plays no role in improving the RMSE, leading to the optimizer always selecting $\beta = 0$ as the best solution. This has also been verified in two real experiments with ground truth data in the T-shirt and Carpet real sequences. It basically means that a prior of constant values for w, controlled by α in (13), is always better than a prior of constant gradient, controlled by β.

Figure 3 shows a two-source synthetic experiment giving further insight on the role of the parameters α and β controlling the shape regularization scheme. It shows that second order regularization is always preferable, but it can cause some problems due to a concave-convex ambiguity, as shown in Fig. 5 and discussed below.

It is worth mentioning that the shape regularization scheme always gives better RMSE results when comparing against ground truth. This is expected since it directly regularizes the shape (where we measure the RMS error), as opposed to the shading regularization scheme that regularizes the shape "rendered" with the occluded light. This does not mean that the shading regularization scheme should be discarded as both schemes behave in slightly different ways.

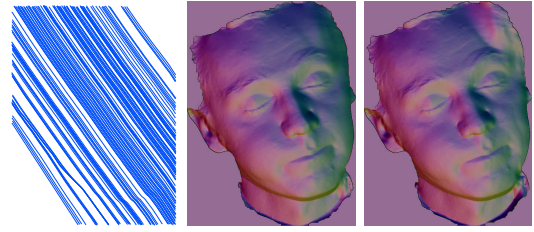We have performed a first experiment with video data of a white-painted face illuminated by three colored lights in a similar way as in [5]. Both the shadow segmentation and the setup calibrated are performed as described in [6]. Figure 4 shows two different frames of the video sequence without taking the shadows into account (left) and after detecting and adding the shading constraints (middle) and the shape constraints (right). We can appreciate how the nose reconstruction is dramatically improved when correctly processing the shadows (see arrows), even though only two lights are visible in the shadowed regions. We also note that the shape regularization scheme fails in some boundary regions (see circle in right column) leading to an incorrect reconstruction of the side of the face. This is caused by the Laplacian regularization term. The term suffers from an ambiguity of two possible solutions, concave or convex, both solutions having similar energy and the data term being unable to disambiguate them.

Figure 5 shows a more detailed analysis of the bottom face in figure 4. The solution of the shape regularization scheme agrees with the constraints (Fig. 5 left) even though it picks the incorrect "concave" solution instead of the convex solution. This is confirmed by looking at the shade rendering of the face under the occluded light (see Fig. 5 middle and right). The shading regularization scheme shows a smooth surface (Fig. 5 middle) while the shape regularization scheme (Fig. 5 right) shows a clear artifact. This is expected since the shading regularization does exactly that, it finds the surface that minimizes the variation of the shading when rendering the shape with the occluded light. The extra knowledge of the missing light is exactly what the shape regularization scheme is missing in order to make the right decision and choose the convex solution.

A second facial performance capture using [5] is shown in figure 6. This time the face is not painted, which implies an assumption of constant albedo chromaticity. In order to cope with shadows, the shading regularization scheme is used. We observe that, despite the constant albedo deviations, *e.g.* the lips, the system successfully captures fine details such as skin wrinkles.

Two more real experiments with ground truth are presented in figures 7, 8 and 9. A classic photometric stereo setup is used, where three photographs are acquired under three different illuminations assuming a distance point light source
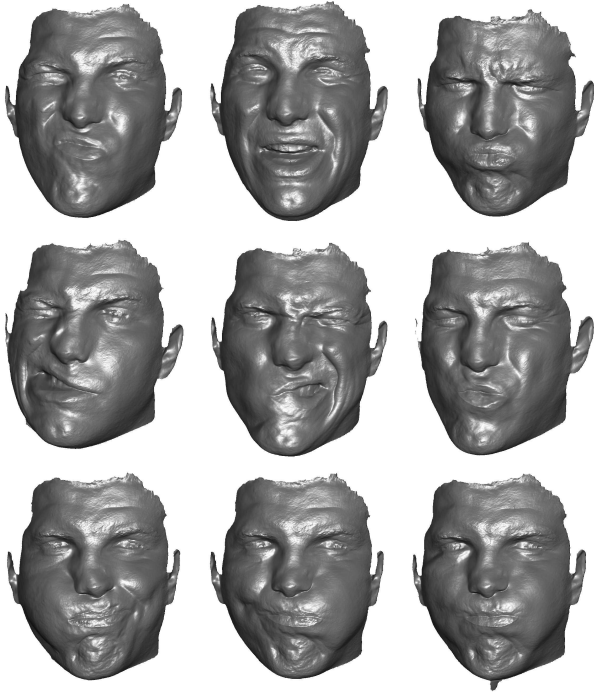
Fig. 6. **Face sequence**. Acquisition of 3d facial expressions using [5] and the shadow processing technique described in this paper. The shadows are processed with the shading regularization scheme. The full video sequence has more than a 1000 frames reconstructed.

Fig. 7. **Stability of regularization parameters for the Carpet (dashed red) and T-shirt (solid blue) sequences.**. The figures show RMSE curves around optimal values. Top row shows RMSE curves for the shading regularization scheme while bottom row shows RMSE curves for the shape regularization scheme. Left column shows RMSE a a function of α while β is set to its optimal value. Right column shows RMSE as a function of β when α is set to its optimal value. Both schemes exhibit good properties with a large range of regularization parameters producing low RMSE. Based on these curves, default values are set to $\alpha = 0.2$ $\beta = 0$ for the shading scheme, and $\alpha = 0.15$, $\beta = 1.0$ for the shape scheme. These values are the default ones used in the sphere, carpet and t-shirt sequences.

and calibrated following [6]. One of the images is repeated with and without a light occlusion. The three images without shadows are used to compute a photometric stereo solution, which will be used as ground truth (see figures 8.b and 9.b). The solutions obtained by not using any regularization are shown in figures 8.c and 9.c. A first experiment is performed to study how the regularization parameters affect the overall RMSE (see fig. 7). The RMSE curves show that a large range of parameters produce acceptable results, allowing us to use a default set of parameters for all the sequences. The shading regularization scheme solutions with default parameters are shown in figures 8.d and 9.d while the shape regularization scheme solutions with default parameters are shown in figures 8.e and 9.e. The RMSE of both schemes are low, the shape regularization scheme always performing slightly better than the shading one. Moreover, the shape regularization scheme is able to perform two-source photometric stereo by completely discarding the image with the shadow (see figures 8.g and 9.g). The RMSE approximately doubles, but it demonstrates a working two-source photometric stereo capture system with varying albedo.

Note that the Carpet's RMSE is more than three times the T-shirt's RMSE. This is mainly due to the presence of higher frequency detail on the Carpet. This detail is much harder to preserve by the regularization scheme while still providing enough regularization to disambiguate the shape.

VII. DISCUSSION

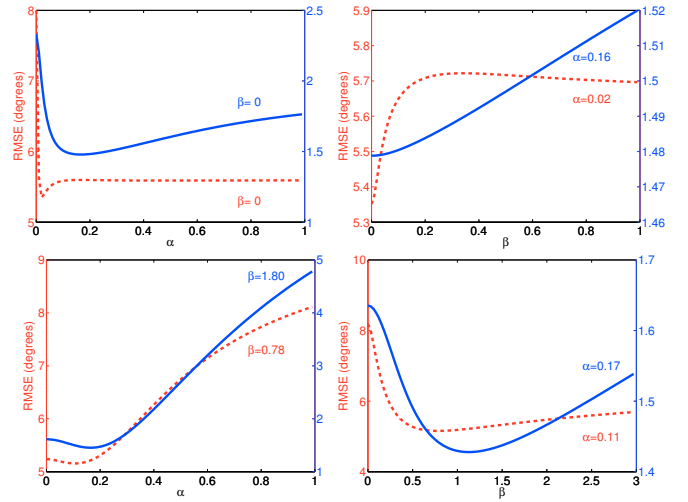This paper investigated the problem of shadows in the context of three-source photometric stereo with an extension to two-source photometric stereo. This is a particularly challenging setup because the surface orientation is under-determined. In its pure form the problem is ill posed even in the presence of no noise in the data. We provided a remedy to this in the form of two regularization schemes that do not suppress the data of the problem. Finally we showed how the ideas in this paper can be applied to the interesting acquisition setup of color photometric stereo. As future work we would like to investigate ways of exploiting the shadows for 3d reconstruction [2] and making the algorithm faster in order to include it in a real-time capture system.

REFERENCES

[1] Barsky, S., Petrou, M.: The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. IEEE Trans. Pattern Anal. Mach. Intell. **25** (2003) 1239–1252
[2] Chandraker, M., Agarwal, S., Kriegman, D.: Shadowcuts: Photometric stereo with shadows. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). (2007)
[3] Drew, M.: Reduction of rank-reduced orientation-from-color problem with many unknown lights to two-image known-illuminant photometric stereo. In: IEEE International Symposium on Computer Vision. (1995) 419–424
[4] Petrov, A.: Light, color and shape. Cognitive Processes and their Simulation (in Russian) (1987) 350–358
[5] Hernández, C., Vogiatzis, G., Brostow, G., Stenger, B., Cipolla, R.: Non-rigid photometric stereo with colored lights. In: Proc. 11th Intl. Conf. on Computer Vision (ICCV). (2007)
[6] Hernández, C., Vogiatzis, G., Cipolla, R.: Shadows in three-source photometric stereo. In: Proc. 10th Europ. Conf. on Computer Vision (ECCV). (2008)
[7] Woodham, R.: Photometric method for determining surface orientation from multiple images. Optical Engineering **19** (1980) 139–144
[8] Lee, S., Brady, M.: Integrating stereo and photometric stereo to monitor the development of glaucoma. Image and Vision Computing **9** (1991) 39–44
[9] Onn, R., Bruckstein, A.: Integrability disambiguates surface recovery in two-image photometric stereo. Intl. Journal of Computer Vision **5** (1990) 105–113
[10] North Coleman, Jr., E., Jain, R.: Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. In: Shape recovery. Jones and Bartlett Publishers, Inc., , USA (1992) 180–199
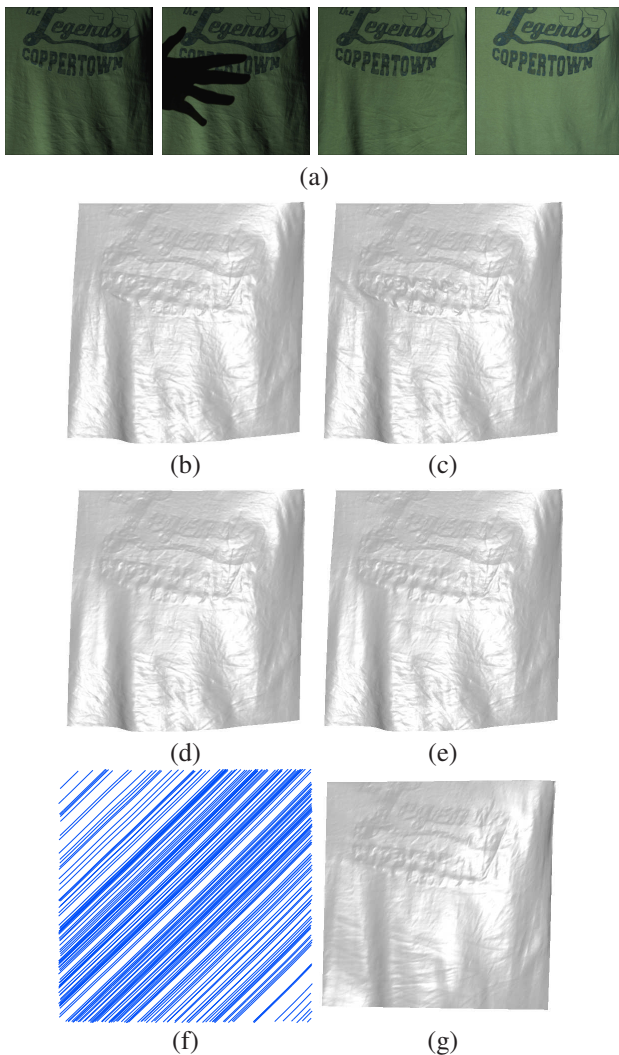
(a)

(b)   (c)

(d)   (e)

(f)   (g)

Fig. 8.   **T-shirt sequence**. **(a)** Set of four 1024x1024 images acquired under three different illuminations. First and second images share the same illumination. Ground truth sequence (b) uses first, third and fourth images, while (c), (d) and (e) use second, third and fourth, and the two-source solution (f) and (g) uses only third and fourth images. **(b)** Ground truth photometric stereo solution. **(c)** Result obtained using just integrability, *i.e.* no regularization is used. RMSE is 2.34 degrees. **(d)** Result obtained using the shading regularization scheme with default parameters. RMSE is 1.49 degrees. **(e)** Result obtained using the shape regularization scheme with default parameters. RMSE is 1.49 degrees. **(f)** Characteristic curves describing the two-source photometric stereo problem using only third and fourth images. **(g)** Two-source photometric stereo result obtained with the shape regularization scheme and parameters. RMSE is 7.9 degrees.
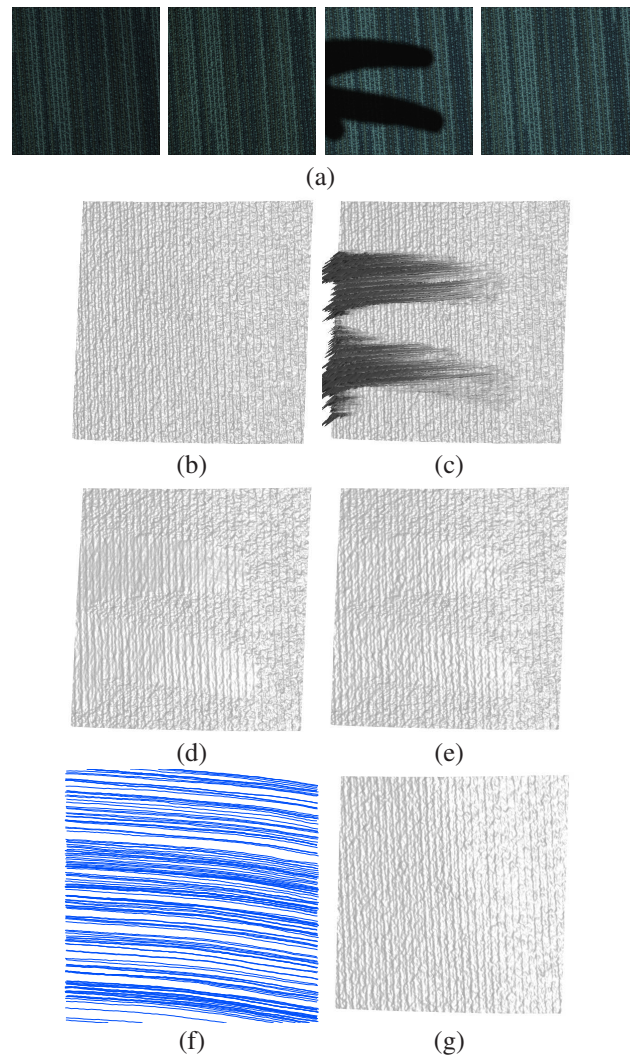

(a)

(b)   (c)

(d)   (e)

(f)   (g)

Fig. 9.   **Carpet sequence**. **(a)** Set of four 1024x1024 images acquired under three different illuminations. Third and fourth images share the same illumination. Ground truth sequence (b) uses first, second and fourth images, while (c), (d) and (e) use first, second, and third, and the two-source solution (f) and (g) uses only first and second images. **(b)** Ground truth photometric stereo solution. **(c)** Result obtained using just integrability, *i.e.* no regularization is used. RMSE is 37.1 degrees. **(d)** Result obtained using the shading regularization scheme with default parameters. RMSE is 5.61 degrees. **(e)** Result obtained using the shape regularization scheme with default parameters. RMSE is 5.26 degrees. **(f)** Characteristic curves describing the two-source photometric stereo problem using only first and second images. **(g)** Two-source photometric stereo result obtained with the shape regularization scheme and default parameters. RMSE is 8.65 degrees.

[11] Yuille, A., Snow, D.: Shape and albedo from multiple images using integrability. In: CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), Washington, DC, USA, IEEE Computer Society (1997) 158

[12] Daum, M., Dudek, G.: On 3-d surface reconstruction using shape from shadows. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on **0** (1998) 461

[13] Yu, Y., Chang, J.T.: Shadow graphs and 3d texture reconstruction. Int. J. Comput. Vision **62** (2005) 35–60

[14] Tankus, A., Kiryati, N.: Photometric stereo under perspective projection. In: Proc. 10th Intl. Conf. on Computer Vision (ICCV), Washington, DC, USA, IEEE Computer Society (2005) 611–616

[15] Wolff, L.B., Angelopoulou, E.: 3d stereo using photometric ratios. In: ECCV '94: Proceedings of the Third European Conference-Volume II on Computer Vision, London, UK, Springer-Verlag (1994) 247–258

[16] Fan, J., Wolff, L.B.: Surface curvature and shape reconstruction from unknown multiple illumination and integrability. Comput. Vis. Image Underst. **65** (1997) 347–359

[17] Chen, H., Belhumeur, P., Jacobs, D.: In search of illumination invariants. In: Proc. IEEE Conf. CVPR. Volume 2. (2000) 254–261

[18] Davis, T.A.: Algorithm 832: Umfpack, an unsymmetric-pattern multifrontal method. ACM Transactions on Mathematical Software **30** (2004) 196–199

[19] Ikeuchi, K., Horn, B.: Numerical shape from shading and occluding boundaries. Artificial intelligence **17** (1981) 141–184

[20] Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: SIGGRAPH '00, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co. (2000) 417–424

[21] Agrawal, A., Raskar, R., Chellappa, R.: What is the range of surface reconstructions from a gradient field? In: Proc. 9th Europ. Conf. on Computer Vision (ECCV), Springer (2006) 578–591