Perspective Shape from Shading for Wide-FOV Near-Lighting Endoscopes

Diogo Roxo, Nuno Gonçalves, and João P. Barreto

Institute for Systems and Robotics,
Faculty of Sciences and Technology,
University of Coimbra,
3030 Coimbra, Portugal
(diogoroxo,nunogon,jpbar)@isr.uc.pt

Abstract. Near-Lighting Endoscopes are self-illuminated cameras often used in minimal invasive surgery. Since they have a wide field of view, their images are affected by high radial distortion and reduced resolution in the periphery of the image. Perspective Shape From Shading has been used for reconstruction but suffers from the problem of resolution and high radial distortion when endoscopes are used. We propose in this paper two improvements to the state of the art methods for PSFS in Near-Lighting Endoscopes. The first contribution is the introduction of the radial distortion model directly in PSFS equations and the second contribution is the compensation of the reduced resolution of the image in its periphery, due to wide FOV. Tests performed in real objects and in a knee bone show that by modeling these two effects our method highly enhances the accuracy of the estimation.

Keywords: Shape from Shading, Perspective, Near Point Lightning, Radial Distortion, Endoscopy

1 Introduction

The use of endoscopes in surgery is becoming more and more common. They provide the surgeon visual access to zones of the human body that are difficult to reach, assisting in minimal invasive procedures. However, usually the images of bones and organs are partial and illuminated directly by the endoscope probe, whose interpetation is not an easy task.

Shape from shading (SFS) has been lately used in this scenario, in order to provide the surgeon shape reconstructions of bones and organs. Since the light source is incorporated in the endoscope probe, the general orthographic shape from shading model is, however, not adequate to this case. Instead, the perspective shape from shading (PSFS) model is more appropriate. Furthermore, endoscopes are vision systems with wide field of view (FOV) and thus presenting high radial distortion. The existing SFS methods correct the radial distortion in a previous step, by computing an undistorted image. While this step presents no problems for images with small distortion, it is a big source of error for images with medium to high, hence for endoscopes.

In this paper we propose two modifications to the general formulation of the PSFS problem using near-lighting endoscopes. The first contribution is the modelling of the radial distortion directly in the shape from shading reflectance equations. The new reflectance equations allow us to use non interpolated data (original image) in the estimation and so reduce the estimation error. The second contribution is to compensate the effect of reduced resolution in the periphery of the image, a problem that arises for wide FOV cameras.

1.1 Related Work

Shape from shading has been introduced to computer vision since the early works by authors like [1], [2]. The major part of papers are focused on Lambertian surfaces with orthographic projection and distant light sources ([1],[3]), however some authors have been considering more complex and realistic environments like non-Lambertian surfaces, perspective projection ([4], [2]). There are some relevant works for near-lightning and taking into account $1/r^2$ attenuation factor (fall-off law of isotropic point sources). Namely [5] and [4] have considered the particular case of the endoscope. In the former case they assume that the light source is coincident with the projection center and in the latter, they assume two sources of light very close (and symmetric) to the camera center of projection.

2 Perspective Shape from Shading for a Near Point Light Source

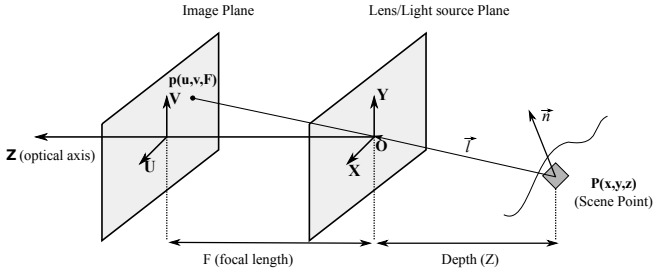


Fig. 1. Perspective projection model for a near light source located at the projection center **O**. The camera plane **XYZ** is centred at **O**.

Due to the light source position on the tip of the endoscope, we consider a perspective projection and formulate the shape from shading problem for a near point light source located at the projection center, as seen in (1). Assuming a Lambertian reflectance of the surface and the inverse square distance fall-off for the light intensity, the scene radiance can be recovered by:

$$R = I_0 \rho \, \frac{(\hat{\boldsymbol{n}} \, . \, \hat{\boldsymbol{l}})}{r^2} \tag{1}$$

where I_0 and ρ are the light source intensity and the surface albedo, respectively. The $Z - axis$ of the camera plane (XYZ) corresponds to the depth. The unit vector $\hat{\boldsymbol{l}}$ represents the incident direction of a light ray on the surface point P while r is the distance from O to P. As proposed by [1], the surface normal $\hat{\boldsymbol{n}}$ is obtained in terms of partial derivatives of the scene depth z:

$$\hat{\boldsymbol{n}} = \left[-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, -1\right] / \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1} \tag{2}$$

When this problem is formulated under orthographic projection these derivatives are applied directly in the image (u, v) coordinates. However under perspective projection we have:

$$x = \frac{uz}{F} \ , \ y = \frac{vz}{F} \tag{3}$$

where F is the focal length. So we should represent the partial derivatives of z in terms of image coordinates as [4]:

$$\begin{aligned} \frac{\partial z}{\partial x} &= \frac{Fp}{z + up}, \quad p = \frac{\partial z}{\partial u} \\ \frac{\partial z}{\partial y} &= \frac{Fq}{z + vq}, \quad q = \frac{\partial z}{\partial v} \end{aligned} \tag{4}$$

As the light source is located at the projection center $r = \overrightarrow{\boldsymbol{l}}$ and we can explicitly write the light rays direction vectors as:

$$\overrightarrow{\boldsymbol{l}} = \left[-\frac{uz}{F}, -\frac{vz}{F}, -z\right] \qquad \hat{\boldsymbol{l}} = \frac{\overrightarrow{\boldsymbol{l}}}{\| \overrightarrow{\boldsymbol{l}} \|} \tag{5}$$

From equations (2-5) the Reflectance equation is then rewritten as function of u, v, z, p, q:

$$R(u, v, z, p, q) = I_0 \rho \, \frac{\hat{\boldsymbol{n}}(u, v, z, p, q) \, . \, \hat{\boldsymbol{l}}(u, v, z)}{r(u, v, z)^2} \tag{6}$$

Estimating the Reflectance map: The mapping of the surface reflectance is done taking into account that we have 3 unknown parameters to estimate, the depth z and the corresponding partial derivatives p and q. From the Image Irradiance Equation [1] we have $R(u, v, z, p, q) = E(u, v)$, where E represents irradiance transmitted by the surface radiance R. Different optimization methods have been proposed ([6],[7], [4]). We use the minimization method proposed by [4]. The Error Function is thus computed as:

$$e(z, p, q) = \lambda e_i(z, p, q) + (1 - \lambda) e_s(z, p, q) \tag{7}$$

where e_i is the integrability error and e_s is the smoothness constraint:

$$e_i(z,p,q) = \int\int_{image} (E(u,v) - R(u,v,z,p,q))^2 \,\mathrm{d}u\mathrm{d}v \tag{8}$$

$$e_s(z,p,q) = \int\int_{image} z_u^2 + z_v^2 + p_u^2 + p_v^2 + q_u^2 + q_v^2 \mathrm{d}u\mathrm{d}v \tag{9}$$

The unknowns z, p, q are computed in an iterative process where each one is defined by discretizing and minimizing the error function. So for each pixel position, $\boldsymbol{\omega} = [z, p, q]$ the update equations are:

$$\boldsymbol{\omega}^{n+1} = \boldsymbol{\omega}_m^n + \frac{\lambda}{4(1-\lambda)}[E(u,v) - R(u,v,\omega^n)]\frac{\partial R}{\partial \omega}|_{\omega^n} \tag{10}$$

where λ is initialized with a small value (0.005), being increased by a step of 0.02 as the error is reduced by 1% and ω_m corresponds to the four-neighbourhood average for each pixel.

2.1 Improvements to the Estimation model

As already stated, we aim for a better reconstruction when using wide-view lenses with high radial distortion, so we introduce two enhancements to [4] estimation model. First we insert the *Radial Distortion Correction* directly in the Reflectance map equation, then we introduce the so called *Field of View Compensation* in order to improve the reconstruction results in the image periphery.

Division model for Radial Distortion Correction: According to [8] the image radial distortion (RD) can be described using the 1st order division model where the level of distortion is quantified by only one parameter ξ (typically $\xi < 0$). Let $\mathbf{u} = (u, v)^\mathsf{T}$ and $\mathbf{u}_u = (u_u, v_u)^\mathsf{T}$ be the corresponding distorted and undistorted points, expressed with respect to a reference frame with origin in the principal point of the image [8]. $\mathbf{f}$ is a vector function that maps points from distorted, $\mathtt{I}_d$, to undistorted, $\mathtt{I}_u$, images:

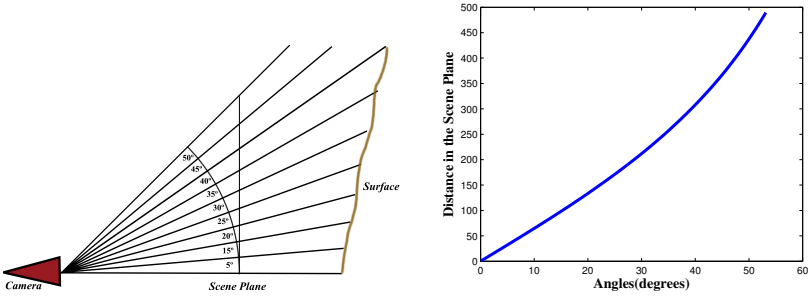
$$\mathbf{u}_u = \mathbf{f}(\mathbf{u}) = (1 + \xi \mathbf{u}^\mathsf{T}\mathbf{u})^{-1}\mathbf{u}. \tag{11}$$

The radius of $\mathbf{u}$ is $r = \sqrt{\mathbf{u}^\mathsf{T}\mathbf{u}}$, and the corresponding undistorted radius is $r_u = (1 + \xi r_d^2)^{-1} r_d$.

Instead of obtaining an undistorted image we integrate (11) directly in the Reflectance equation which becomes a function of u_u, v_u, z, p, q:

$$R(u_u, v_u, z, p, q) = I_0 \rho \, \frac{\hat{\boldsymbol{n}}(u_u, v_u, z, p, q) \,.\, \hat{\boldsymbol{l}}(u_u, v_u, z)}{r(u_u, v_u, z)^2} \tag{12}$$

By introducing the radial distortion directly in the reflectance equation, and so eliminating the previous step of creating an undistorted image, we eliminate an interpolation step which is one source of error. This is specially important for medium to high radially distorted images, as endoscopes are.



(a) (b)

Fig. 2. (a) shows the effect of equal spacing in angles and its effect on the resolution and (b) the scene distance variation with respect to the angle. We can observe that to an isotropic variation in the angle between the optical axis and the light ray irradiated by the surface the distances in the scene plane increase while in the arc they are constant.

Field of View Compensation: A typical camera presents a FOV between 40° and 70°. Which is a rather low value when compared to the 140° of our endoscope. We can thus observe in the images a gradual loss of quality from the center to the periphery. Figure 2 shows that for each 5° interval, the distances between the light rays increase along the *scene plane*. This means that for equal distances in the *scene plane* as the viewing angle increases, the light rays spacing also increases, and hence there is a loss of resolution which can be observed in the *image plane* by backprojection of the light rays.

Regarding this we modified the smoothness constraint in the estimation model, to compensate the loss of resolution in the reconstruction. By discretizing we can represent (9) for each pixel (i, j) as:

$$\begin{aligned} e_s(z,p,q) = &(z_{i+1,j} - z_{i,j})^2 + (z_{i,j+1} - z_{i,j})^2 \\ &+(p_{i+1,j} - p_{i,j})^2 + (p_{i,j+1} - p_{i,j})^2 \\ &+(q_{i+1,j} - q_{i,j})^2 + (q_{i,j+1} - q_{i,j})^2 \end{aligned} \tag{13}$$

This constraint is based in a pixel by pixel difference which becomes larger in the image periphery due to loss of resolution. We propose a new *smoothness constraint* where the differences are compensated according to the viewing angle. In 2 we can see that the rays concentration in the arc tangent to the scene plane is always the same for equal distances. Using the arc length we calculated the image coordinates in the arc. Considering $\mathbf{x} = (x, y)^\mathsf{T}$

$$\mathbf{x}^{arc} = \mathbf{x} \left[\frac{F \tan\left(\frac{\sqrt{x^2+y^2}}{F} \right)}{\sqrt{x^2+y^2}} \right] \tag{14}$$

where $\mathbf{x}^{arc} = (x^{arc}, y^{arc})^{\mathsf{T}}$ are the coordinates in the arc tangent to the *image plane*. We thus redefine the smoothness constraint as:

$$\begin{aligned} e_s(z,p,q) = &((x^{arc}_{i+1,j} - x^{arc}_{i,j})(z_{i+1,j} - z_{i,j}))^2 + ((y^{arc}_{i,j+1} - y^{arc}_{i,j})(z_{i,j+1} - z_{i,j}))^2 \\ &+((x^{arc}_{i+1,j} - x^{arc}_{i,j})(p_{i+1,j} - p_{i,j}))^2 + (y^{arc}_{i,j+1} - y^{arc}_{i,j})(p_{i,j+1} - p_{i,j}))^2 \\ &+((x^{arc}_{i+1,j} - x^{arc}_{i,j})(q_{i+1,j} - q_{i,j}))^2 + (y^{arc}_{i,j+1} - y^{arc}_{i,j})(q_{i,j+1} - q_{i,j}))^2 \end{aligned} \tag{15}$$

In this way the pixel differences are attenuated according to the loss of resolution. This also implies new equations to the estimation $\boldsymbol{\omega} = [z, p, q]$:

$$\boldsymbol{\omega}^{n+1} = \overline{\boldsymbol{\omega}}^n_m + \frac{\lambda}{\overline{dist}_{arc}(1-\lambda)}[E(u,v) - R(u,v,\omega)]\frac{\partial R}{\partial \omega}|_{\omega} \tag{16}$$

$$\overline{dist}_{arc} = (x^{arc}_{i+1,j} - x^{arc}_{i,j})^2 + (x^{arc}_{i-1,j} - x^{arc}_{i,})^2 + (y^{arc}_{i,j+1} - y^{arc}_{i,j})^2 + (y^{arc}_{i,j-1} - y^{arc}_{i,j})^2 \tag{17}$$

$$\begin{aligned} \overline{\boldsymbol{\omega}}^n_m = \frac{1}{\overline{dist}_{arc}}[&\omega_{i+1,j}(x^{arc}_{i+1,j} - x^{arc}_{i,j})^2 + \omega_{i-1,j}(x^{arc}_{i-1,j} - x^{arc}_{i,j})^2 + \\ &\omega_{i,j+1}(y^{arc}_{i,j+1} - y^{arc}_{i,j})^2 + \omega_{i,j-1}(y^{arc}_{i,j-1} - y^{arc}_{i,j})^2] \end{aligned} \tag{18}$$

3 Calibration of the Endoscopic System

In figure 3 we present the framework of the aquisiction which mainly consists in a geometric and radiometric calibration of the images. The use of a perspective projection requires the computation of the camera intrinsics respecting the assumptions made in the calculus of the reflectance function.

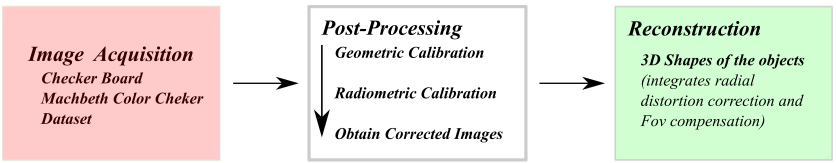


Fig. 3. Overview of the system framework.

3.1 Geometric Calibration

The aim in this step is to determine the intrinsic calibration K_0 (as defined in [9]) and the radial distortion ξ when the lens probe is at a fixed position.

Several authors addressed the specific problem of intrinsic calibration or RD correction in medical endoscopes [10] [11] [9]. We used a recent toolbox, EasyCamCalib [9], which only requires an image of a checker-board to fully calibrate a camera specifically developed to lens presenting moderate to high radial distortion. In [9] it can be found a detailed explanation of the algorithm.

3.2 Radiometric Calibration

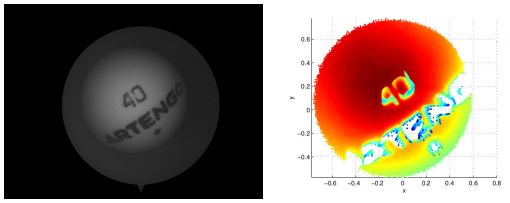
The radiometric calibration is frequently addressed in literature as an essential step to SFS [4]. As stated earlier, we must guarantee a linear photometric response of the CCD and an isotropic distribution of the light rays, by which mean we define the image irradiance perceived by the camera as $E(u,v) = \Gamma^{-1}(I(u,v))U^{-1}(u,v)$, where Γ^{-1} corresponds to the Photometric Response Function, $I(u,v)$ to the Image Brightness at each pixel and $U^{-1}(u,v)$ to the anisotropy of the light source. From equations (1) and assuming that the image irradiance equals the scene radiance we can rewrite the irradiance as:

$$\Gamma^{-1}(I(u,v)) = \rho\, I_0\, U(u,v) \left(\frac{\hat{\boldsymbol{n}} \,.\, \hat{\boldsymbol{l}}}{r^2} \right) \tag{19}$$

In order to estimate Γ^{-1},U^{-1} and $(\hat{\boldsymbol{n}} \,.\, \hat{\boldsymbol{l}}) \,/\, r^2$ all the other variables ρ,I_0, $\hat{\boldsymbol{n}}$, $\hat{\boldsymbol{l}}$ and r must be known, so for calibration purposes a MachBeth Colour Chart composed by 24 patches with known albedo is used. As formulated by [4] by applying a logarithm on both sides of (19) we can reformulate the problem as:

$$\log(\Gamma^{-1}(I(u,v))) = \log(\rho) + \log(I_0) \;\; + \log\left(U(u,v) \left(\frac{\hat{\boldsymbol{n}} \,.\, \hat{\boldsymbol{l}}}{r^2} \right) \right) \tag{20}$$

Then we found separate solutions to the 3 unknown parameters by using several images of the Colour Chart patches. For the Response Function Γ^{-1} we make comparisons at the same pixel position for different images while for the anisotropy, U^{-1}, we use different pixel positions on the same image.



(a) Ball Calib Image (b) Ball Reconstruction

Fig. 4. Reconstruction for one of the spheres used (in this case a ping-pong ball). We can see that the reconstruction is quite good, and the surface obtained is very smooth.

4 Experimental Evaluation

For the evaluation we used an oblique-viewing endoscope with a single light source located at the end of the scope tip which has a 4 millimeters diameter. The light source has a sickle shape and a thickness smaller than 1 millimeter. The

endoscope distortion is about 40% and the FOV near 140°. We compared our estimation method with [4] using 3 geometric objects: 2 spheres (Ball and YBall) and a cylinder (Roll) with a radius of 43, 18 and 26 millimeters, respectively. Figure 4 shows an example of a reconstruction using the YBall. 30 images, 10 for each, were used in the evaluation. We also perform reconstructions for a knee model in order to test our estimation with bone-like surfaces.

4.1 Results with Geometric Shaped Images

From the obtained images we performed the corresponding geometric approximations of a sphere and a cylinder. These estimations were improved using RANdom SAmple Consensus (RANSAC)[12]. The metrics used to test the approximations were the *Mean Error* from the distance of each pixel to the sphere surface, its *Standard Deviation* and the *percentage of Inliers*.

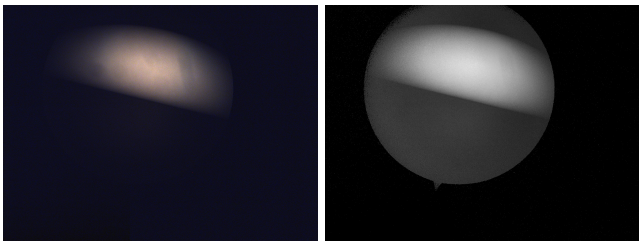
Table 1. Comparison of our method incorporating the radial distortion correction and the FOV compensation against the original estimation method proposed by [4]

Algorithm	**Mean Error**	**Std Deviation**	**Inliers**(%)	**Object (N° Images)**
Ours	0.25	0.19	99	Ball(10)
	0.26	0.25	99	YBall(10)
	1.05	0.75	97	Roll(10)
Wu [4]	0.30	0.24	99	Ball(10)
	0.31	0.25	98	YBall(10)
	1.89	1.35	89	Roll(10)

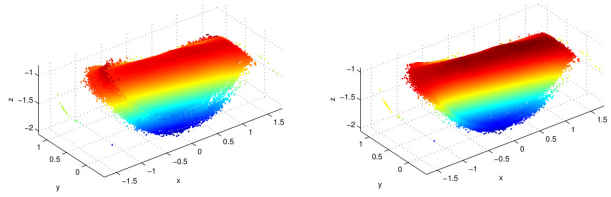
In table 1 we compare the performance of our algorithm with [4]. For all datasets used we obtained inferior values for Mean Error, Standard Deviations and better values for the % of inliers. The results improved 20% in both spheres and 80% in the cylinder. They are quite satisfactory. The improvement for the cylinder is probably related with the FOV compensation as we can see in 5 where we compare our test by the difference introduced in the reconstruction by the FOV compensation. There is a noticeable quality loss in the periphery of the cylinder reconstruction when not using the FOV compensation.

4.2 Results from Images of the Knee Model

In figure 6, although the irregularities presented along the reconstructed surface resulting from the non-smoothness of the bone surface, its shape is inferred quite accurately. The goal of this kind of reconstructions is to access the viability of using it in endoscopic interventions were a surgeon could access in real time to a robust 3D shape of the surface in study.



(a) Original (b) Calibrated Image



(c) No FOV compensation (d) With FOV comp.

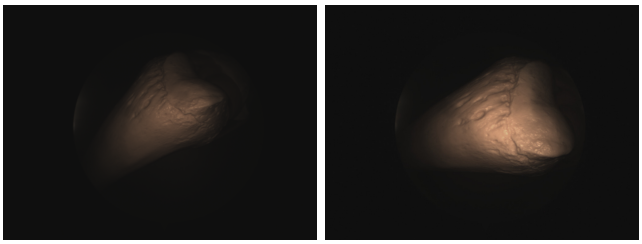
Fig. 5. Influence of the FOV compensation in the reconstructions: (a) is the image acquired with the endoscope. (b) is the corresponding calibrated image, we can see the difference in the radiometric response and in the anisotropy of the light source. (c) presents our reconstruction method integrating only the radial distortion correction. In (d) the radial distortion correction and the FOV compensation were both used.

5 Conclusions

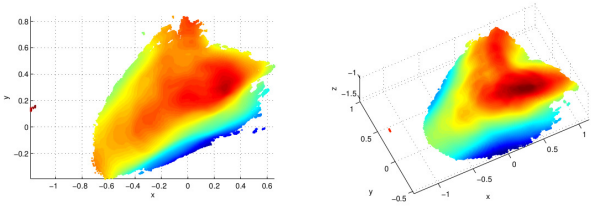
This paper presented an estimation method for near-lightning Perspective Shape from Shading for Endoscopes and by taking into account the effect of high radial distortion, typical in Endoscopes, and also the reduced resolution in the periphery of the image due to a wide FOV. The improvements to the PSFS state of the art are the integration of radial distortion estimation directly in the radiance equation and a new smoothness constraint for wide view cameras. The framework used allowed us to compare with good accuracy the reconstructions obtained. The experimental results shown in general very small errors with our algorithm having a better behaviour than the original proposed by [4]. By visually inspecting the reconstruction results, our method shows a more robust reconstruction in the image periphery as intended by the FOV compensation.

References

1. Horn, B.K.P., Brooks, M.J.: The variational approach to shape from shading. Computer Vision, Graphics, and Image Processing **33**(2) (1986) 174–208
2. Penna, M.A.: Local and semi-local shape from shading for a single perspective image of a smooth object. Computer Vision, Graphics, and Image Processing **46**(3) (1989) 346–366



(a) Knee 1 Original (b) Knee 2 Original



(c) Knee 1 Reconstruction (d) Knee 2 Reconstruction

Fig. 6. Reconstructions of a Knee model:(a) and (b) are the original images acquired with the endoscope. The corresponding reconstructions are shown in (c) and (d).

3. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from shading: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **21** (1999) 690–706
4. Wu, C., Narasimhan, S.G., Jaramaz, B.: A multi-image shape-from-shading framework for near-lighting perspective endoscopes. International Journal of Computer Vision (February 2009)
5. Tankus, A., Sochen, N., Yeshurun, Y.: A new perspective [on] shape-from-shading. In: Ninth International Conference on Computer Vision (ICCV). (2003) 862–869
6. Ikeuchi, K., Horn, B.K.P.: Numerical shape from shading and occluding boundaries. In Horn, B.K.P., Brooks, M.J., eds.: Shape from Shading. MIT Press, Cambridge, MA (1989) 245–299
7. Leclerc, Y.G., Bobick, A.F.: The direct computation of height from shading. In: Conference on Computer Vision and Pattern Recognition, Hawaii (1991)
8. Willson, R.G., Shafer, S.A.: What is the center of the image? J. Opt. Soc. Am. A **11**(11) (1994) 2946–2955
9. Barreto, J.P., Roquette, J., Sturm, P., Fonseca, F.: Automatic Camera Calibration Applied to Medical Endoscopy. In: BMVC. (2009)
10. Asari, K., Kumar, S., Radhakrishnan, D.: A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation. IEEE Trans. Med. Imag. **18** (1999) 345–354
11. Wengert, C., Cattin, P.C., Duff, J.M., Baur, C., Székely, G.: Markerless endoscopic registration and referencing. In: MICCAI. (2006)
12. Fischler, M., Bolles, R.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Commun. ACM **24**(6) (1981) 381–395