Neurocomputing 150 (2015) 136–146

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Perspective shape from shading for wide-FOV near-lighting endoscopes

Nuno Goncalves *, Diogo Roxo, Joao Barreto, Pedro Rodrigues

Institute of Systems and Robotics – University of Coimbra, Polo 2 – Pinhal de Marrocos, Coimbra, Portugal

ARTICLE INFO

Article history:
Received 31 January 2014
Received in revised form
5 July 2014
Accepted 12 August 2014
Available online 18 October 2014

Keywords:
Wide field of view
Perspective shape from shading
Endoscopy
Radial distortion

ABSTRACT

In minimally invasive medical procedures the surgeon uses as guidance the video acquired by near-lighting endoscopic cameras. Near-lighting endoscopes are small size self-illuminated cameras that typically have a wide or very wide field of view (FOV), and hence their images are affected by high radial distortion and small spatial resolution in the periphery of the image (an object of the same size at equal distance gets smaller in the periphery). To help the interpretation of the images by the surgeon, the 3D modeling of the observed cavities is highly interesting and shape-from-shading is a natural choice in the 3D shape reconstruction of bones and organs. However, the shape from shading technique has been traditionally performed using the perspective projection model (perspective shape from shading – PSFS) which is barely appropriate for modeling endoscopes since it does not account for the high distortion and reduced resolution in image periphery. Our aim is thus to adapt the shape from shading technique to the particular case of near-lighting endoscopes. 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 phantom show that by modeling these two effects our method highly improves the accuracy of the estimation.

1. Introduction

In the last two decades, surgical medicine has suffered great changes influenced by the growing of minimal invasive surgery techniques [1,2]. In minimally invasive surgery (MIS) the doctor executes the procedure guided by the images acquired by an endoscopic camera. Typically the endoscopic camera consists in autoclavable exchangeable optics that are mounted on a charge-coupled device (CCD) camera head just before the procedure starts (see Fig. 1). These optics can be forward-viewing, when they look up-front, or oblique-viewing if the tubular lens has a cut of 30°, 70° or 90° to enable a periscope-like view. The viewing direction of the oblique-viewing lenses can be changed without moving the camera head by simply rotating the endoscope around its symmetry axis [3–6]. Oblique-viewing endoscopes are specially useful in inspecting narrow cavities, such as the articulations (arthroscopy) or the sinus (rhinoscopy), where the space to maneuver the probe is very limited.

Although being highly beneficial for the patient in terms of recovery time and risk of infection, the stats show that MIS is used only in 25% of the procedures eligible for this technique. The reason for this low penetration is the fact that MIS procedures are much more difficult to execute than the equivalent open surgery counterparts. The access to the organs is very limited and surgeons must use intra-operative video as only guidance. This poses severe problems in terms of hand-eye coordination and some surgeons are simply unable to make the visual-spatial leap needed to master this technique [7].

In this context, the use of systems for computer aided surgery (CAS) can make a significant difference in the adoption and clinical outcomes of MIS. We envision that such systems will receive as input the intra-operative endoscopic video, eventually register the image data with pre-operative models of targeted organs, and use the information for assisting the doctor during the procedures. Such assistance can take multiple forms, ranging from providing a better visualization of the observed cavities to granting that the surgery is executed according to a pre-plan, and passing by helping the doctor to navigate inside the human-body.

Despite all the advantages referred, usually the images acquired in endoscopy are of a small part of the environment, imaging a partial view of bones and organs and are illuminated directly by the endoscope probe, whose interpretation is not an easy task. To overcome this issue, modeling the 3D shape of bones and organs is naturally important and able to give an actual help to the surgeon as usually reported by themselves. Since, however, few

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2014.08.077
0925-2312/

information about the scene is available and the illumination conditions are not ideal, shape from shading (SFS) has been lately used in the scenario of endoscopy, in order to provide the surgeon shape reconstructions of bones and organs. Similar to the human vision system SFS tries to infer shapes using only shading information. This software-based solution only needs a standard endoscopic system composed by a boroscope, a CCD camera and a light source. Due to the complexity of the problem many studies have adopted assumptions like Lambertian reflectance, known light source location and known surface albedo. 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. To formulate a robust PSFS it is then necessary to know the camera intrinsic parameters and the radiometric response which are obtained through geometric and radiometric calibration.

Another issue that arises from the use of endoscopes in PSFS is the amount of radial distortion that affects the images. The existing SFS methods correct the radial distortion in a previous step, by computing an undistorted image by interpolation. While this step presents no problems for images with small distortion, it can become a big source of error for images with medium to high distortion, hence for endoscopes. Actually, endoscopes are vision systems with generally wide fields of view (FOV) and thus usually presenting high radial distortion (RD). This is a persisting problem due to the very small size of the lenses and it causes a non-linear deformation in the image, increasing from the center to the periphery [8]. In Fig. 2 we show an endoscopic image of a regular grid where it is visible the effect caused by radial distortion, namely the fact that straight lines become curves as the distance to the image center increases. The visual perception becomes compromised and in some situations even a trained surgeon has a hard time when inferring size and depths of the object in study [9].

Fig. 1. Rigid oblique-viewing endoscope, used for image acquisition. The light source is connected to the boroscope through a fibre optic cable.

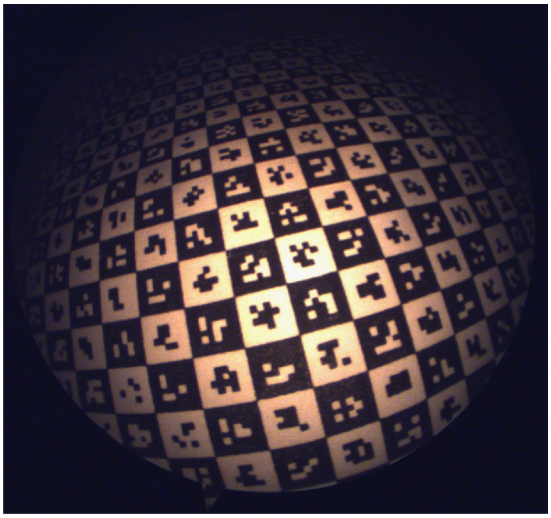


Fig. 2. Effect of radial distortion in endoscopic images.

In this paper we then propose two modifications to the general formulation of the PSFS problem using near-lighting endoscopes. The first contribution is the modeling 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 it reduces the estimation error. This new formulation also eliminates the burden of generating undistorted images reducing the computational time. 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. This paper is an extended version of our work published at the 6th Iberian Conference on Pattern Recognition and image Analysis [10], where the explicit equations of the reflectance with the radial distortion correction are explicitly formulated and new results with medical images are presented. The algorithm is also presented and discussed in more detail.

1.1. Related work

Shape from shading has been introduced to computer vision since the early works of Horn and Brooks [11] and Penna [12]. The majority of related works are focused on Lambertian surfaces with orthographic projection and distant light sources [11,13], however some authors have been considering more complex and realistic environments like non-Lambertian surfaces and perspective projection [14,12]. 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, [15,14] have among others 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.

Other studies using shape from motion were introduced in this scenario. By capturing image sequences while moving the endoscope it is possible to recover the shape information [16]. Applying it to medical field, two authors [17,18] presented reconstructions of different anatomical structures from motion cues like image features and correspondences which is only possible in a limited number of cases. Featureless organs like bones cannot be used in this context. Photometric stereo can also be tested in order to recover surface orientation [19], although it requires two light sources switched on and off at different times while the major part of the endoscopes were not designed considering this feature.

As mentioned in the Introduction section, typically the shapes recovered from endoscopic images correspond to a tiny part of the organ, which severally limits the reconstruction. Techniques such as endoscopic mosaicing from images [16] can be useful in order to visualize a larger shape of the object in observation and improve the reconstruction.

As for the problem of radial distortion, this problem has already been taken into consideration, and recently the manufacturer Stryker® presented an endoscopic system with reduced radial distortion by combining a new optical design with a cropping of the image periphery. However the images continue to present incorrect perspectives with a reduced Field of View (FOV) and this solution cannot be extended to small diameter lenses like the one tested in this paper. At this moment, computer vision techniques appear to be the only viable way to fully solve the RD issue. Using image warping it is then possible to render geometrically correct perspectives, modeling the camera intrinsic parameters. A fully

functional software-based system for correcting radial distortion running in real time on a standard computer has already been proposed by [20].

Although it may be possible to solve the RD issue, the visualization continues to be in 2D which means that the depth perception depends on the surgeon ability to precise location and sizes of anatomical structures. Sometimes to interpret the information supplied by the images can be rather difficult, even with an improved visualization, so several authors have already addressed improvements such as 3D visualization or surgical navigation systems. The most common approach has been the use of stereoscopic endoscopes such as the da Vinciâs system (Intuitive Surgical Inc.) [21] which supplies a true 3D visualization from the combination of two images acquired simultaneously. It consists in the use of two optical elements, corresponding to left and right eyes, and a 12 millimeter diameter stereo endoscope. However it cannot provide depth and sizes for use in computer analysis. An interesting approach was preformed by [22]. Using a laser with a known pattern and two different cameras, one sensitive to the laser, other for acquire the texture map. A 3D map is then generated from the images of the first camera and using a structured light technique it is thus possible to render the acquired images in a 3D scene.

Recent developments suggest that software based solutions using image processing techniques are a promising area to achieve a true 3D visualization in endoscopy.

2. Perspective shape from shading for a near point light source

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 Fig. 3. 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 the so-called reflectance equation:

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

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 $\hat{l}$ represents the incident direction of a light ray on the surface point P while r is the distance from the light source to P. As proposed in [11], the surface normal $\hat{n}$ is obtained in terms of the partial derivatives of the scene depth z:

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

When this problem is formulated assuming an orthographic projection, these derivatives are applied directly in the image coordinates (u,v). However under perspective projection it follows

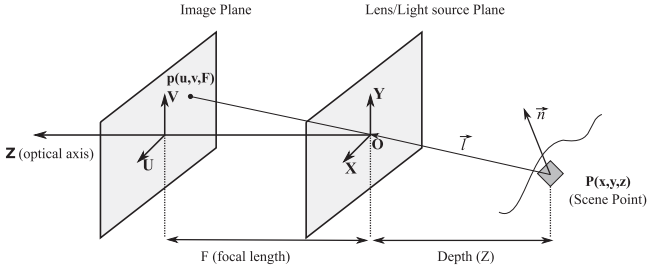


Fig. 3. Perspective projection model for a near light source located at the projection center **O**. The lens/light source plane **XYZ** is centered at **O**. A scene point **P** is projected into an image point **p**.

that

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

where F is the focal length. Similar to [14], we represent the partial derivatives of z in terms of image coordinates:

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

Since the light source is located at the projection center $r = \vec{l}$, we can explicitly write the light ray direction vectors as

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

Using Eqs. (2)–(5) the reflectance equation is then rewritten as function of u, v, z, p, q:

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

$$R(u,v,z,p,q) = I_0\rho\frac{F^2\cdot\left(1+\frac{up}{z+up}+\frac{vq}{z+vq}\right)}{z^2\cdot\sqrt{\left(\frac{p^2}{(z+up)^2}+\frac{q^2}{(z+vq)^2}+\frac{1}{F^2}\right)\cdot(u^2+v^2+F^2)^3}} \tag{7}$$

2.1. Estimating the reflectance map

The mapping of the surface reflectance is done taking into account that we have three unknown parameters to estimate: the depth z and the corresponding partial derivatives p and q. From the image irradiance equation (see [11]) we have $R(u,v,z,p,q) = E(u,v)$, where E represents irradiance transmitted by the surface and R the reflectance map. In [11] the problem is regarded as the estimation of a smooth surface in a specific image domain where some boundary conditions apply. Different optimization methods have been proposed [23,24,14]. In this paper we use the minimization method proposed by [14] since it conjugates a local quality assessment measure (integrability error constraint) with a regularization term (smoothness constraint) that helps in achieving a global optimization solution. The error function is thus computed as

$$e(z,p,q) = \lambda e_i(z,p,q) + (1-\lambda)e_s(z,p,q) \tag{8}$$

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

$$e_i(z,p,q) = \iint_{image}(E(u,v) - R(u,v,z,p,q))^2 \, du \, dv \tag{9}$$

$$e_s(z,p,q) = \iint_{image} z_u^2 + z_v^2 + p_u^2 + p_v^2 + q_u^2 + q_v^2 \, du \, dv \tag{10}$$

and λ is a Lagrange multiplier used to control the smoothness influence. The smoothness constraint is used to force the resulting surface to be as smooth as possible by integrating the first and second derivatives of the surface depth.

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

$$\vec{\omega}^{n+1} = \vec{\omega}_m^n + \frac{\lambda}{4(1-\lambda)}\left[E(u,v) - R(u,v,\vec{\omega}_m^n)\right]\frac{\partial R}{\partial \omega}\Big|_{\vec{\omega}_m^n} \tag{11}$$

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 p, q, z four-neighborhood average for each pixel.

3. Estimating SFS in endoscopic images

We aim for an accurate reconstruction when using wide-view lenses with a significant amount of radial distortion, so we introduce two enhancements to the 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.

3.1. Division model for radial distortion correction

The reflectance map equation (Eq. (7)) assumes that the image used as data source is undistorted. However, as explained above, the majority of endoscopes has medium to high radial distortion and the correction of this effect is usually performed in a previous step, by interpolating a new image. This new image, that is used as data source for the shape from shading algorithm, is thus undistorted. In this section we claim that instead of interpolating the original distorted image, as so introducing a severe source of error, it is possible to directly integrate the distortion equations in the reflectance map estimation.

According to [25] the image radial distortion (RD) can be described using the first order division model where the level of distortion is quantified by only one parameter ξ (typically $\xi < 0$). Let $\mathbf{u}_d = (u, v)^\mathsf{T}$ (the original image coordinates) 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 [25]. $\mathbf{f}$ is a vector function that maps points from distorted (I_d) to undistorted (I_u) image planes as

$$\mathbf{u}_u = \mathbf{f}(\mathbf{u}_d) = (1+\xi \mathbf{u}_d^\mathsf{T}\mathbf{u}_d)^{-1}\mathbf{u}_d. \tag{12}$$

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

Instead of obtaining an undistorted image we then integrate (12) directly in the reflectance equation which becomes a function of u_d, v_d, z, p, q and ξ:

$$R(u_u, v_u, z, p, q) = R(u_d, v_d, z, p, q, \xi)$$

$$= I_0 \rho \vec{n}(u_d, v_d, z, p, q, \xi). \frac{\hat{\vec{l}}(u_d, v_d, z, \xi)}{r(u_d, v_d, z, \xi)^2} \tag{13}$$

or, by expanding the final result:

$$R(u_d, v_d, z, p, q, \xi) = I_0 \rho \frac{F^2 \Delta^3 \cdot \left(\frac{u_d p}{\Delta z + u_d p} + \frac{v_d q}{\Delta z + v_d q} + 1\right)}{z^2 \cdot \sqrt{\left(\frac{p^2\Delta^2}{(z\Delta + u_d p)^2} + \frac{q^2\Delta^2}{(z\Delta + v_d q)^2} + \frac{1}{F^2}\right)(F^2\Delta^2 + u_d^2 + v_d^2)^3}} \tag{14}$$

where $\Delta = 1 + \xi(u_d^2 + v_d^2)$.

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 of the pixels values, which is a major source of error. This is specially important for medium to high radially distorted images as the ones obtained from endoscopic lenses.

3.2. Field of view compensation

A typical camera presents a FOV of approximately 40°, 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. Fig. 4 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 (10) we can represent the smoothness constraint for each pixel (u_d, v_d) as

$$\begin{aligned} e_s(z,p,q) \approx \sum_{image} \Big\{ & (z_{u_d+1,v_d} - z_{u_d,v_d})^2 + (z_{u_d,v_d+1} - z_{u_d,v_d})^2 \\ & + (p_{u_d+1,v_d} - p_{u_d,v_d})^2 + (p_{u_d,v_d+1} - p_{u_d,v_d})^2 \\ & + (q_{u_d+1,v_d} - q_{u_d,v_d})^2 + (q_{u_d,v_d+1} - q_{u_d,v_d})^2 \Big\} \end{aligned} \tag{15}$$

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 in the image resolution are compensated according to the viewing angle. In Fig. 4 we can see that the rays concentration in the arc tangent to the image plane is always the same for equal distances. Considering $\mathbf{u} = (u, v)^\mathsf{T}$ as the image coordinates of a pixel, we

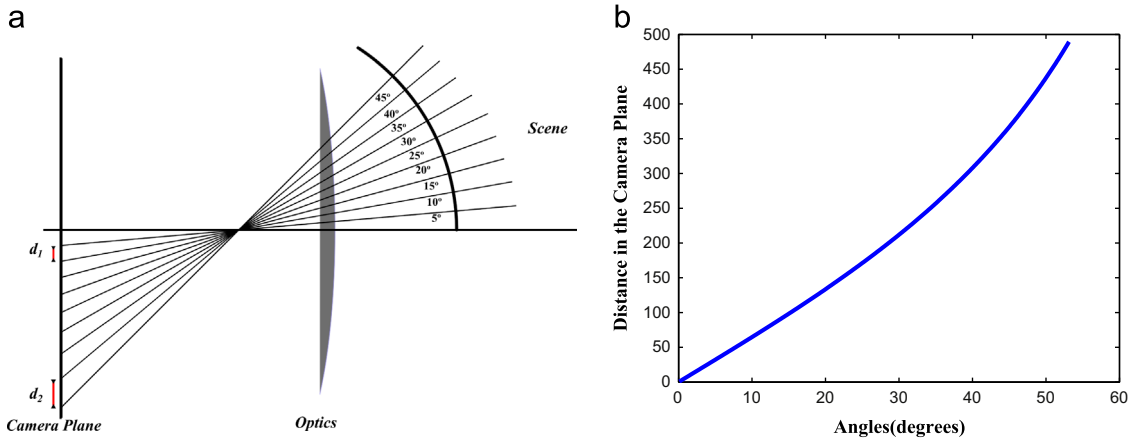


Fig. 4. (a) Light rays projection from the surface to the camera (image) plane, showing the effect of equal spacing in angles and its effect on the resolution and (b) the scene distance variation with respect to the angle. We can observe that for an isotropic variation in the angle between the optical axis and the light ray irradiated by the surface, the distances in the image plane increase ($d_1 > d_2$) while in the arc they are constant.

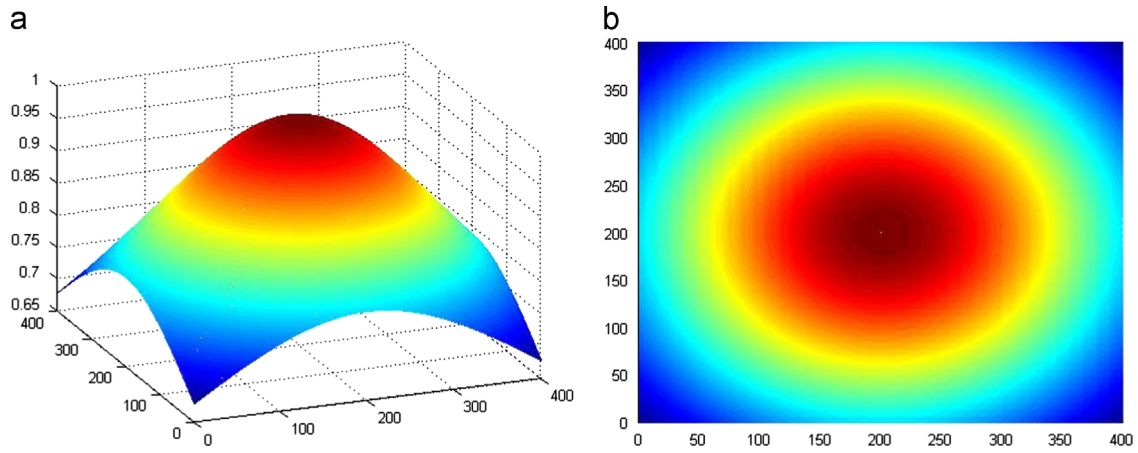


Fig. 5. Attenuation factor used to compensate for the loss of resolution in the image periphery. Example for an image of 400 × 400 pixels and a focal length of 200 pixels.

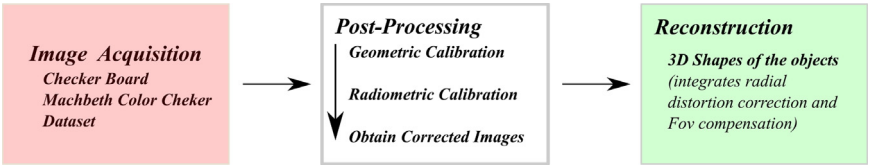


Fig. 6. Overview of the system framework.

propose to used the following scaled point:

$$\mathbf{u}^{arc} = \left[\frac{F\ \tan^{-1}\left(\frac{\sqrt{u^2+v^2}}{F}\right)}{\sqrt{u^2+v^2}}\right]\cdot\mathbf{u} = \frac{\theta}{\tan(\theta)}\cdot\mathbf{u} \tag{16}$$

where $\mathbf{u}^{arc}=(u^{arc},v^{arc})^{\mathsf{T}}$ are the coordinates in the arc tangent to the image plane, and θ is the angle between the projective ray and the principal axis.

In Fig. 5 we can see an example of the attenuation factor used in $\mathbf{u}^{arc}$ so that when applied to the smoothness constraint the differences at the periphery are attenuated as a function of the angle that the projective ray makes with the principal axis and as a function of the distance of the pixel to the principal point.

We thus redefine the smoothness constraint as

$$\begin{aligned} e_s(z,p,q) = \sum_{image}\Big\{ &\left(\left(u^{arc}_{u+1,v}-u^{arc}_{u,v}\right)(z_{u+1,v}-z_{u,v})\right)^2 \\ &+\left(\left(v^{arc}_{u,v+1}-v^{arc}_{u,v}\right)(z_{u,v+1}-z_{u,v})\right)^2 \\ &+\left(\left(u^{arc}_{u+1,v}-u^{arc}_{u,v}\right)(p_{u+1,v}-p_{u,v})\right)^2 \\ &+\left(\left(v^{arc}_{u,v+1}-v^{arc}_{u,v}\right)(p_{u,v+1}-p_{u,v})\right)^2 \\ &+\left(\left(u^{arc}_{u+1,v}-u^{arc}_{u,v}\right)(q_{u+1,v}-q_{u,v})\right)^2 \\ &+\left(\left(v^{arc}_{u,v+1}-v^{arc}_{u,v}\right)(q_{u,v+1}-q_{u,v})\right)^2\Big\} \end{aligned} \tag{17}$$

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

$$\vec{\omega}^{n+1}_m = \vec{\omega}^{n}_m + \frac{\lambda}{dist_{arc}\cdot(1-\lambda)}\left[E(u,v)-R\left(u,v,\vec{\omega}^{n}_m\right)\right]\frac{\partial R}{\partial\omega}\Big|_{\vec{\omega}^{n}_m} \tag{18}$$

where

$$dist_{arc} = \left(u^{arc}_{u+1,v}-u^{arc}_{u,v}\right)^2+\left(u^{arc}_{u-1,v}-u^{arc}_{u,v}\right)^2+\left(v^{arc}_{u,v+1}-v^{arc}_{u,v}\right)^2 + \left(v^{arc}_{u,v-1}-v^{arc}_{u,v}\right)^2 \tag{19}$$

and

$$\begin{aligned} \vec{\omega}^{n}_m = \frac{1}{dist_{arc}}\Big[&\omega_{u+1,v}\left(u^{arc}_{u+1,v}-u^{arc}_{u,v}\right)^2+\omega_{u-1,v}\left(u^{arc}_{u-1,v}-u^{arc}_{u,v}\right)^2 \\ &+\omega_{u,v+1}\left(v^{arc}_{u,v+1}-v^{arc}_{u,v}\right)^2+\omega_{u,v-1}\left(v^{arc}_{u,v-1}-v^{arc}_{u,v}\right)^2\Big] \end{aligned} \tag{20}$$

4. Calibration of the endoscopic system

To perform the shape from shading it is important to, before applying the appropriate algorithms, calibrate the system. Endoscopes are optical systems that need some attention and care in their calibration.

In Fig. 6 we present the proposed reconstruction framework which consists in a geometric and radiometric calibration of obtained images, followed by the SFS reconstruction algorithm. The use of a perspective projection requires the computation of the camera intrinsics respecting the assumptions made in the computation of the reflectance function.

4.1. Geometric calibration

The aim of this step is to determine the intrinsic calibration matrix K_0 and the radial distortion ξ. K_0 is defined by

$$K_0 \sim \begin{pmatrix} aF & sF & c_x \\ 0 & a^{-1}F & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{21}$$

where the aspect ratio a is assumed to be equal to 1 and the skew s is assumed to be zero. (c_x,c_y) represents the principal point where the optical axis intersects the image plane.

Planar checker-board patterns are widely used in geometric calibration because they are easily available and simplify greatly the goal of establishing point correspondences. Several authors addressed the specific problem of intrinsic calibration or RD correction in medical endoscopes [26–28]. Due to the usability and accuracy needed by our system, we used the EasyCamCalib toolbox

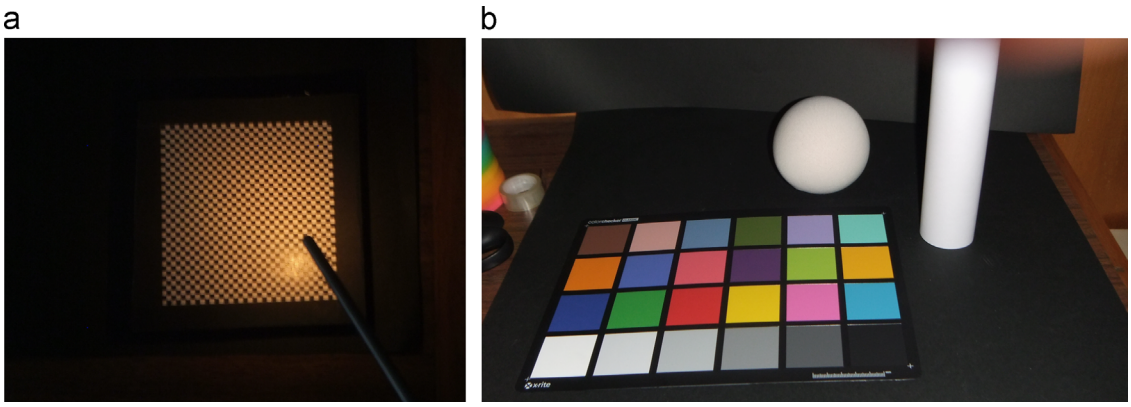


Fig. 7. (a) Planar checker-board used in geometric calibration and (b) MachBeth Color Chart for the radiometric calibration. The scope tip is positioned above the grid or the color patch as shown in the image.

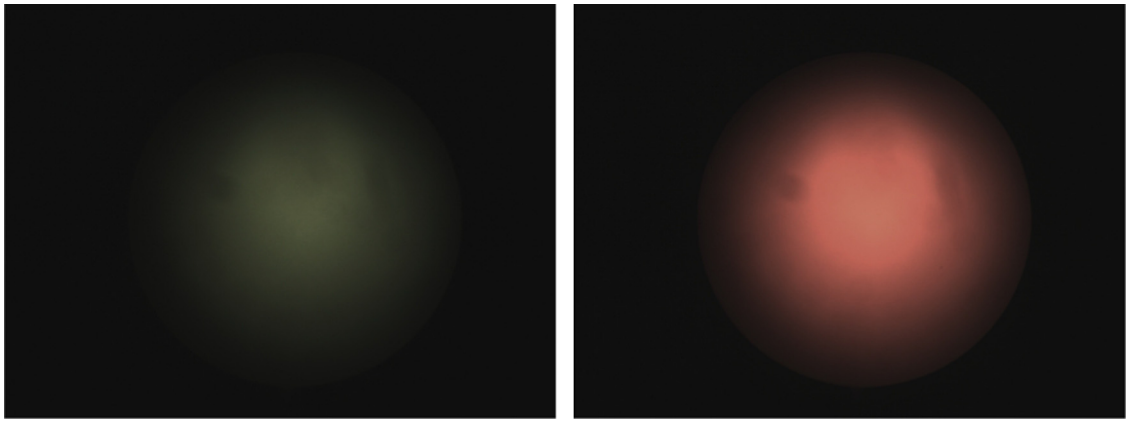


Fig. 8. Images of the color chart, each one corresponding to a different patch.

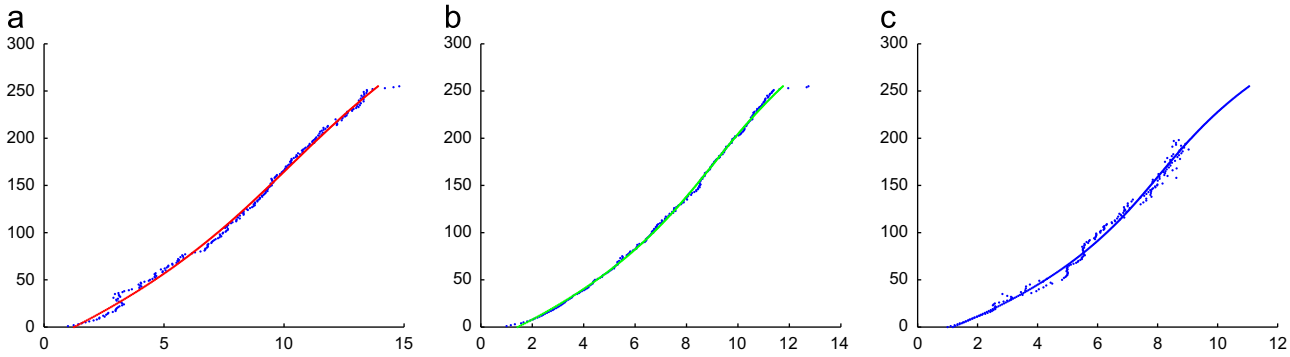


Fig. 9. Response curves, corresponding to the three color channels: (a) red, (b) green, (c) blue. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

[28], which only requires an image of a checker-board to fully calibrate a camera with medium to high radial distortion. In [28] a detailed explanation and evaluation of the algorithm can be found.

4.2. Radiometric calibration

The radiometric calibration is frequently addressed in literature as an essential step to SFS [14]. As stated earlier, we must guarantee a linear photometric response of the CCD and an isotropic distribution of the light rays, by which means we define the image irradiance perceived by the camera as

$$E(u,v) = \Gamma^{-1}(I(u,v))U^{-1}(u,v) \tag{22}$$

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 Eqs. (1) and (22) and assuming that the image irradiance equals the scene radiance we can rewrite Eq. (22) as

$$\Gamma^{-1}(I(u,v)) = \rho I_0 U(u,v)\left(\hat{\vec{n}} \cdot \frac{\hat{\vec{l}}}{r^2}\right) \tag{23}$$

In order to estimate Γ^{-1}, U^{-1} and $(\hat{\vec{n}}.\hat{\vec{l}})/r^2$ all the other variables ρ, I_0, $\vec{n}$, $\vec{l}$ and r must be known, so for calibration purposes a MachBeth Color Chart composed by 24 patches with known albedo is used (Fig. 7). As formulated in [14] by applying a logarithm on both sides of Eq. (23) we can reformulate the

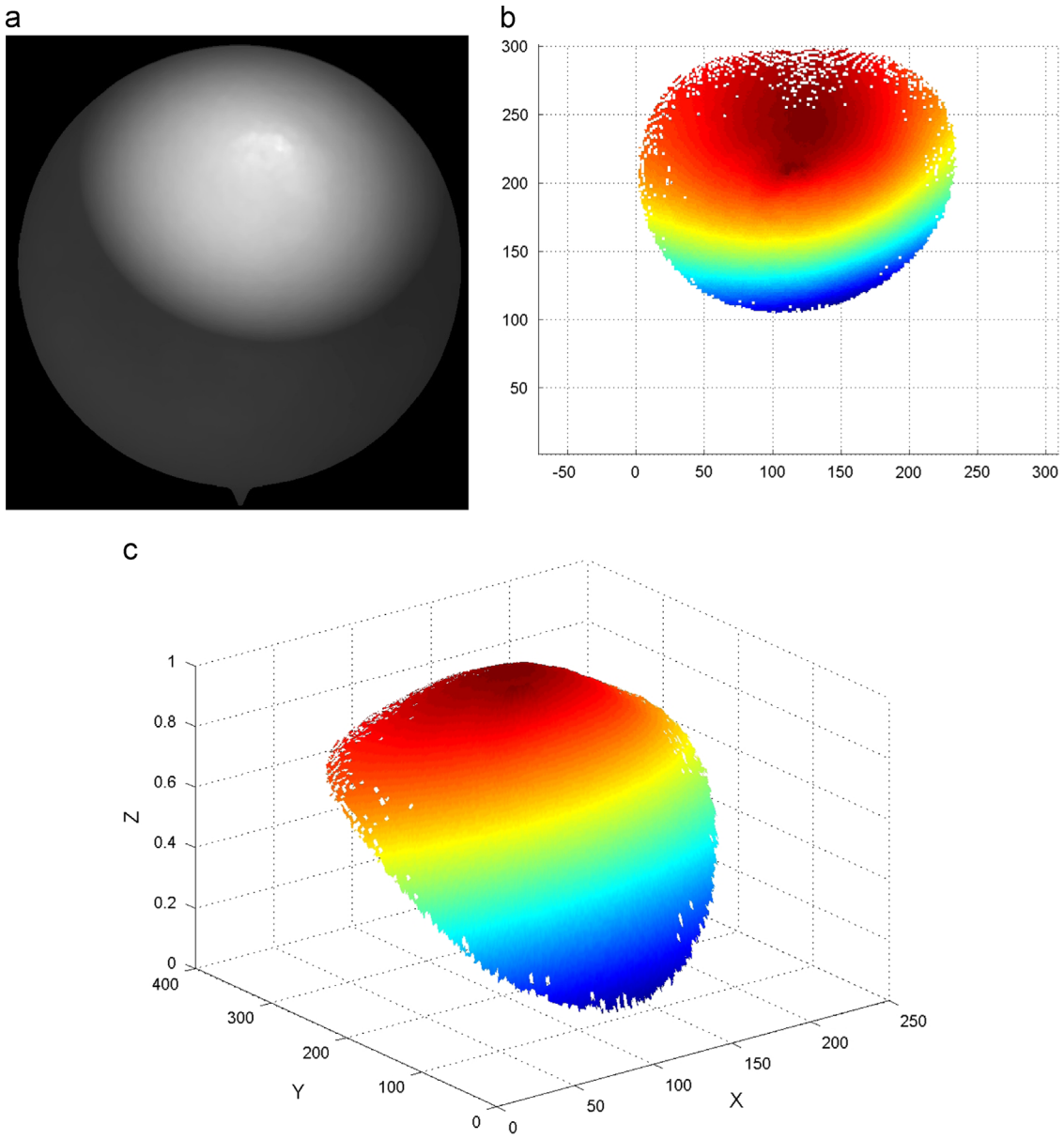


Fig. 10. Reconstruction (up to scale factor) for one of the spheres used (in this case a ping–pong ball). We can see that the reconstruction is quite good, and the surface obtained is very smooth. (a) Calibration ball image, (b) ball reconstruction – *XY* plot and (c) ball reconstruction – *XYZ* plot.

problem as

$$\log\left(\Gamma^{-1}(I(u,v))\right) = \log(\rho) + \log(I_0) + \log\left(U(u,v)\left(\hat{\vec{n}} \cdot \frac{\hat{\vec{l}}}{r^2}\right)\right) \quad (24)$$

Then we found separate solutions to the 3 unknown parameters by using several images of the Color Chart patches, as the ones in Fig. 8 and knowing their position and orientation. For the response function Γ^{-1} we perform comparisons at the same pixel position for different images while for the anisotropy, U^{-1}, we use different pixel positions on the same image. The obtained response functions are represented in Fig. 9.

5. Experimental evaluation

To evaluate the proposed method we used an oblique-viewing endoscope with a single light source located at the end of the scope tip which has a 4 mm diameter. The light source has a crescent shape and a thickness smaller than 1 mm. The endoscope distortion is close to 40% and the FOV near 140°. We compared our estimation method with [14] using 3 geometric objects: 2 spheres (Ball and YBall) and a cylinder (Roll) with a radius of 43, 18 and 26 mm, respectively. Fig. 10 shows an example of a reconstruction using the YBall. 10 images per object were used in the evaluation.

Additionally, we also performed reconstructions for a knee model in order to test our estimation with bone-like surfaces.

5.1. Results with geometric shaped images

In order to evaluate the accuracy of the reconstructed objects, we performed the corresponding geometric approximations of a sphere and a cylinder to the results obtained by the two methods being compared: ours and [14]. These estimations were obtained using RANdom SAmple Consensus (RANSAC) [29] to fit the reconstructed points to the geometric model of each object as

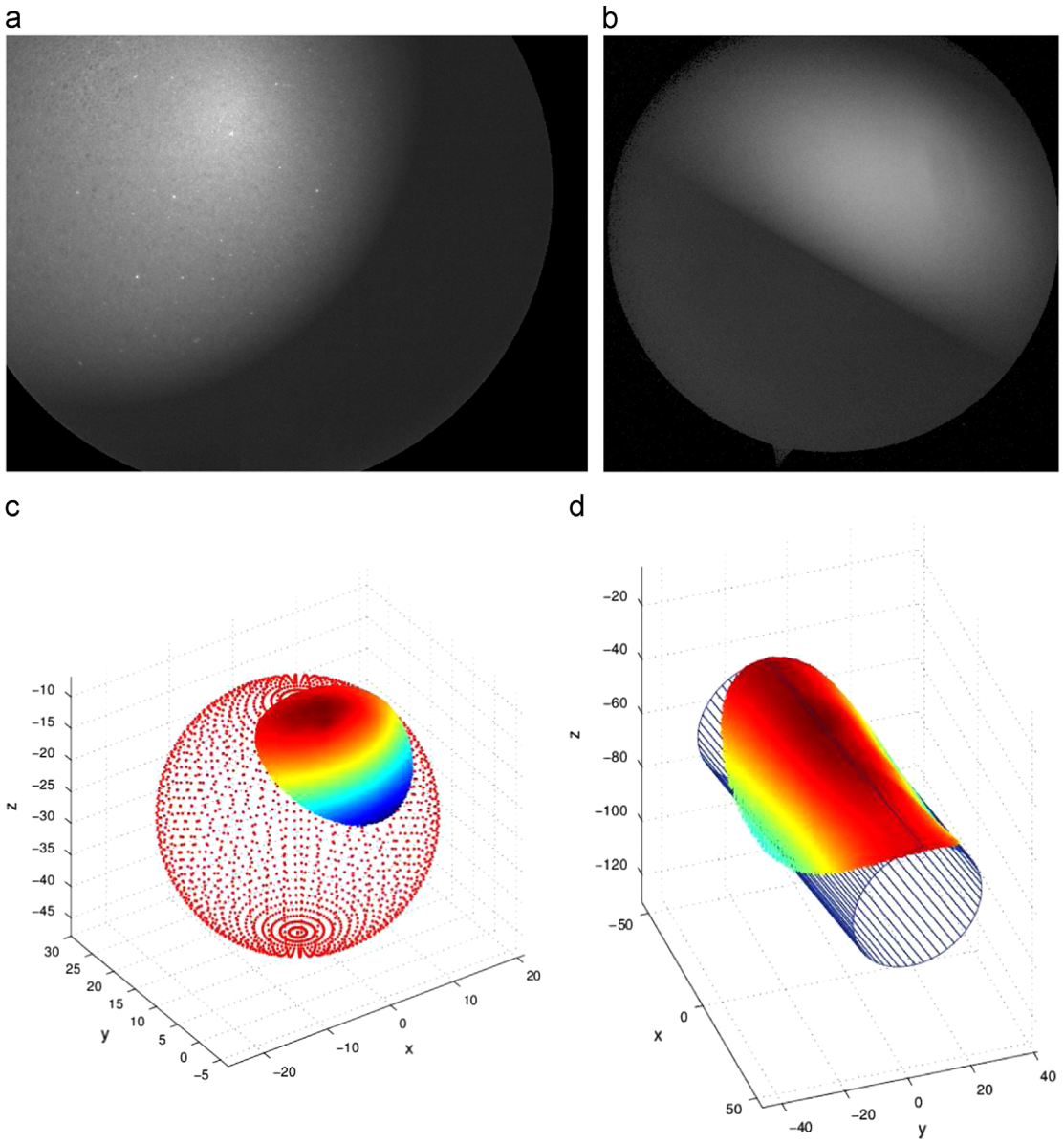


Fig. 11. Estimated geometric models for the reconstructed points of a sphere (a) and a cylinder (b) using RANSAC.

Table 1
Comparison of our method incorporating the radial distortion correction and the FOV compensation against the original estimation method proposed by [14].

Algorithm	Mean error	Std. dev.	Inliers (%)	Object (*N*)
Ours	0.25	0.19	99	Ball (10)
	0.26	0.25	99	YBall (10)
	1.05	0.75	97	Roll (10)
Wu [14]	0.30	0.24	99	Ball (10)
	0.31	0.25	98	YBall (10)
	1.89	1.35	89	Roll (10)

seen in Fig. 11. The metrics used to test the approximations were the Mean Error of each reconstructed point distance to the sphere surface, the corresponding Standard Deviation and the percentage of inliers of the fitted model.

In Table 1 we compare the performance of our algorithm with [14]. For all datasets used we obtained inferior values for mean error, standard deviations and larger percentage of inliers. The results improved 20% in both spheres and 80% in the cylinder. The improvement for the cylinder is due to the FOV compensation, as we can see in Fig. 12, where we compare 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.

5.2. Results from images of the knee model

In Fig. 13, despite the irregularities presented along the reconstructed surface resulting from the non-smoothness of the bone surface, its shape is inferred quite accurately. The images presented show a marker of one irregularity of the knee bone to ease the visual inspection. The goal of this kind of reconstructions is to assess the viability of using them in endoscopic interventions where a surgeon could visualize in real time a robust 3D shape of the surface being observed.

Additionally, we also performed the shape from shading reconstructions of some knee joint images in an arthroscopic surgery. As shown in Fig. 14, one can observe that the shape is correctly recovered and that the 3D images can help the surgeon to understand the shape of the bones.

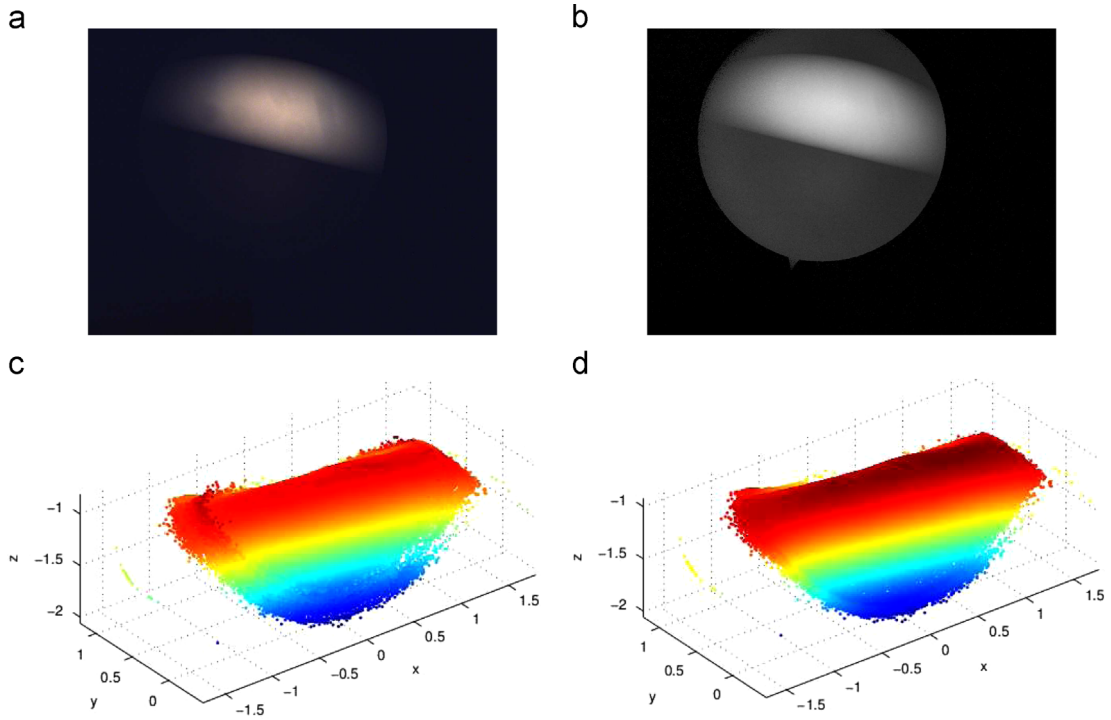


Fig. 12. Influence of the FOV compensation in the reconstructions: (a) is the image acquired with the endoscope. (b) is the corresponding calibrated image, we can see the difference in the radiometric response and in the anisotropy of the light source. (c) presents our reconstruction method integrating only the radial distortion correction. In (d) the radial distortion correction and the FOV compensation were both used.

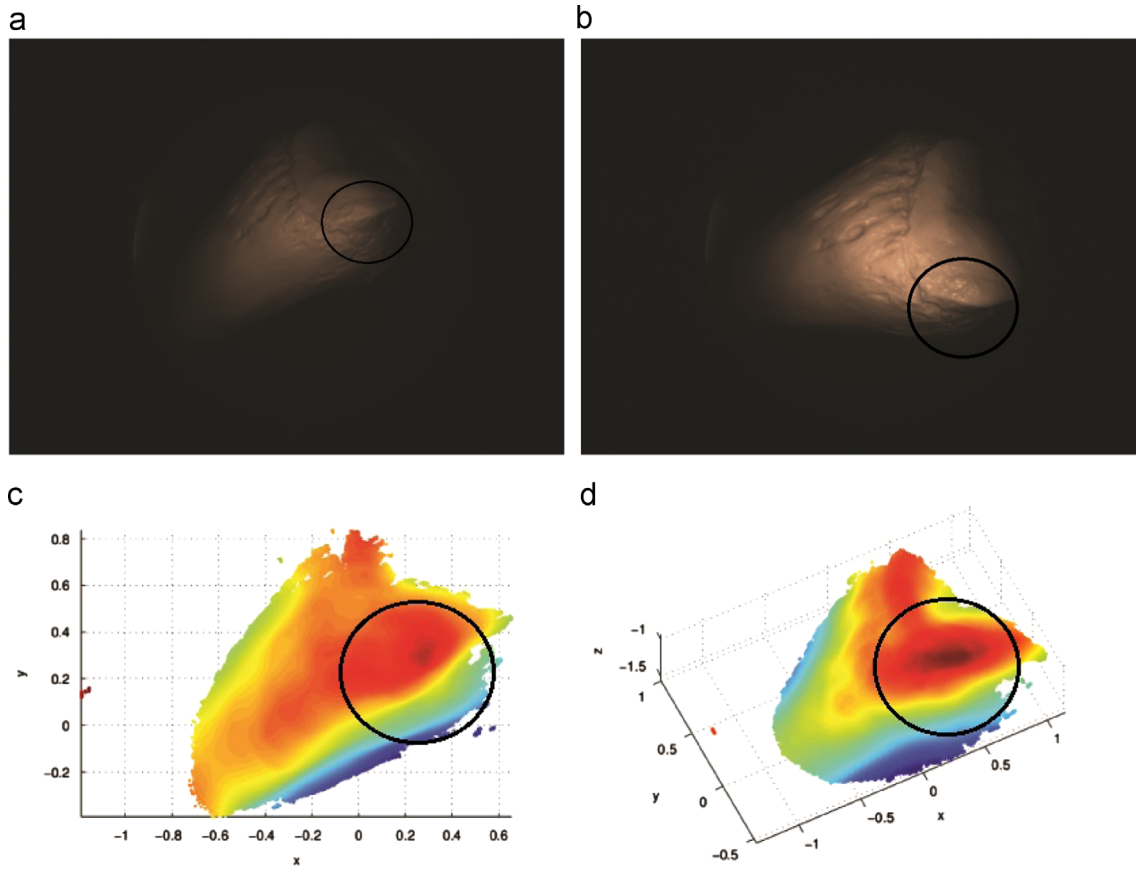


Fig. 13. Reconstructions of a knee model: (a) and (b) are the original images acquired with the endoscope. The corresponding reconstructions are shown in (c) and (d).

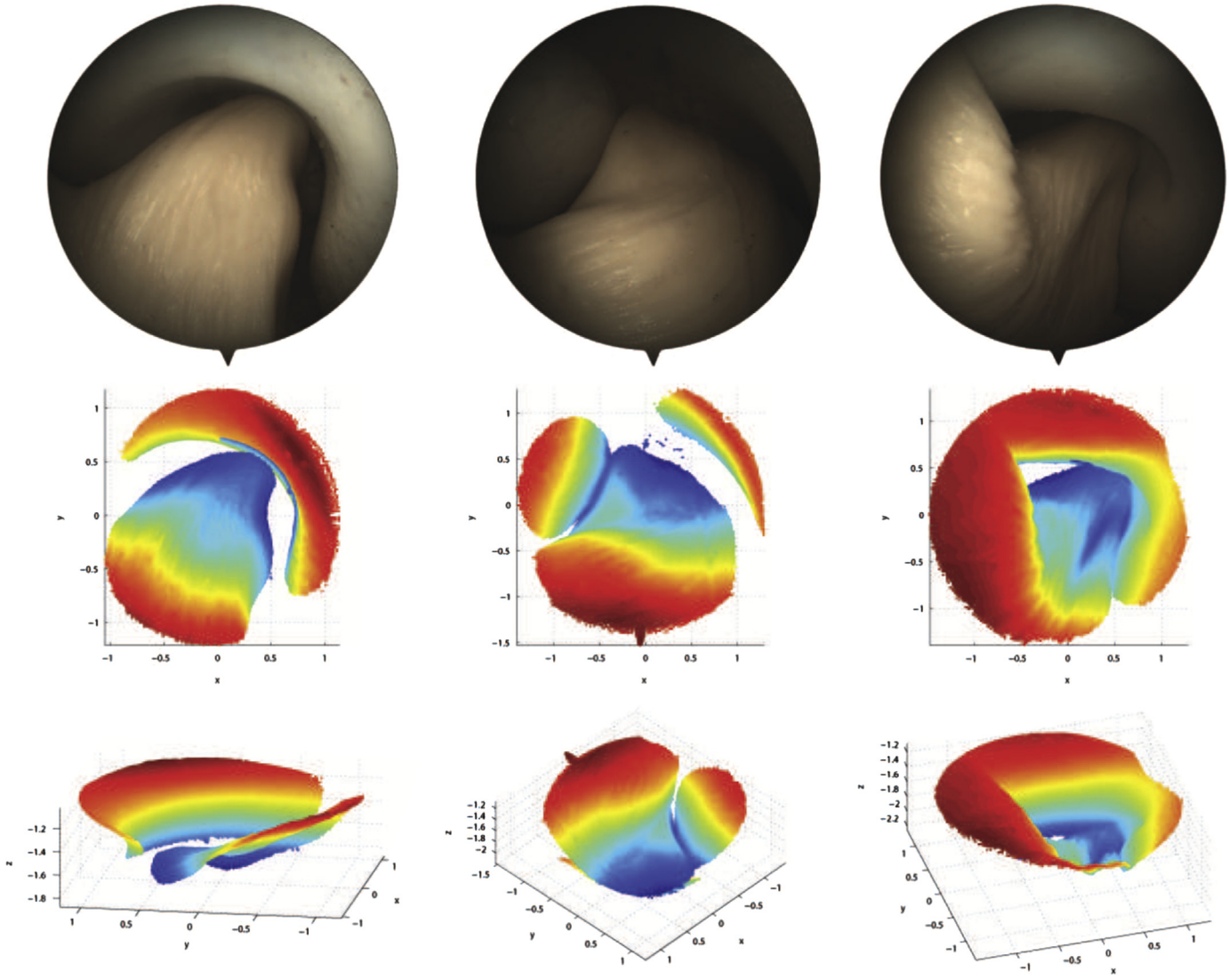


Fig. 14. Reconstructions of the knee joint in an arthroscopic surgery: the first row shows the original images acquired with the endoscope, the second row shows the *XY* plot of the recovered shape (the color represents the height) and the third row shows one view selected of the corresponding reconstruction. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

6. Conclusions

Different techniques to improve visualization in minimally invasive surgeries have been proposed in the last years. Successful 3D reconstructions from endoscopic images have already been seen using hardware based solutions by incorporating an extra camera or a laser device. However it has been proved to be rather difficult the development of software based solutions using computer vision methods as shape from shading and shape from motion due to the numerous variables that need to be considered. We focused in solving the problems related directly with the endoscopic system like light anisotropy, radial distortion, color response and loss of resolution due to a wide FOV.

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 angle lenses. The framework used allowed us to reconstruct surfaces with good accuracy, outperforming those obtained with state-of-the-art methods. The experimental results presented very small errors with our algorithm, thus having a better behavior than the original proposed by [14]. By visually inspecting the reconstruction results, our method shows a more robust reconstruction in the image periphery as result of the FOV compensation.

In terms of future directions, the reconstruction obtained by the shape from shading is always partial, representing just a small part of the object in study, for real scenarios. A shape from motion technique should be considered in order to fuse partial reconstructions and combining them with a navigation tracking system we could aim to a computer guided surgery supplying 3D shape information during the intervention [30]. Additionally, a complete study on performance and computational time issues is important to identify possible improvements to the algorithm in order to achieve real-time performance.

Acknowledgments

Joao Barreto and Pedro Rodrigues want to thank QREN-Mais Centro by funding through the project Novas Tecnologias para

Apoio à Saúde e Qualidade de Vida, Projecto Cirurgia e Diagnóstico Assistido por Computador Usando Imagem.

References

[1] K.H. Fuchs, Minimally invasive surgery, Endoscopy 34 (2) (2002) 154–159. http://dx.doi.org/10.1055/s-2002-19857.

[2] N. ZuKer, A. Gumbs, B. Gayet, Laparoscopic liver resection, in: N. Soper (Ed.), Mastery of Endoscopic and Laparoscopic Surgery, 4th ed.,Lippincott, Philadelphia, 2012.

[3] T. Yamaguchi, M. Nakamoto, Y. Sato, Y. Nakajima, K. Konishi, M. Hashizume, T. Nishii, N. Sugano, H. Yoshikawa, K. Yonenobu, et al., Camera model and calibration procedure for oblique-viewing endoscope, in: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003, Springer, Berlin Heidelberg, 2003, pp. 373–381.

[4] C. Wu, B. Jaramaz, S. Narasimhan, A full geometric and photometric calibration method for oblique-viewing endoscopes, Comput. Aided Surg. 15 (1–3) (2010) 19–31.

[5] N. Fukuda, Y.W. Chen, M. Nakamoto, T. Okada, Y. Sato, A scope cylinder rotation tracking method for oblique-viewing endoscopes without attached sensing device, in: 2010 2nd International Conference on Software Engineering and Data Mining (SEDM), IEEE, 2010, pp. 684–687. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5542836.

[6] R. Melo, João P. Barreto, G. Falcão, Superior Visualization in Endoscopy, Technical Report, URL: ⟨http://arthronav.isr.uc.pt/reports/visualizationreport.zip⟩, March 2011.

[7] Surgeons don 3d glasses, get 'remarkable' depth perception for keyhole surgery ⟨http://www.huffingtonpost.ca/2012/01/24/surgeons-don-3d-glasses-n1229173.html⟩, January 2012.

[8] R. Hartley, S.B. Kang, Parameter-free radial distortion correction with centre of distortion estimation, in: ICCV, 2005, pp. 1834–1841.

[9] J. Helferty, C. Zhang, G. McLennan, W. Higgins, Videoendoscopic distortion correction and its application to virtual guidance of endoscopy, IEEE Trans. Med. Imaging 20 (2001) 605–617.

[10] D. Roxo, N. Goncalves, J. Barreto, Perspective shape from shading for wide-fov near-lighting endoscopes, in: J. Sanches, L. Mico, J. Cardoso (Eds.), Pattern Recognition and Image Analysis, Lecture Notes in Computer Science, vol. 7887, Springer, Berlin, Heidelberg, 2013, pp. 21–30. http://dx.doi.org/10.1007/978-3-642-38628-2_3.

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

[12] M.A. Penna, Local and semi-local shape from shading for a single perspective image of a smooth object, Comput. Vis. Graph. Image Process. 46 (3) (1989) 346–366.

[13] R. Zhang, P.-S. Tsai, J.E. Cryer, M. Shah, Shape from shading: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 21 (1999) 690–706 http://doi.ieeecomputersociety.org/10.1109/34.784284.

[14] C. Wu, S.G. Narasimhan, B. Jaramaz, A multi-image shape-from-shading framework for near-lighting perspective endoscopes, Int. J. Comput. Vis. 86 (2-3) (2010) 211–228. http://dx.doi.org/10.1007/s11263-009-0207-3.

[15] A. Tankus, N. Sochen, Y. Yeshurun, A new perspective [on] shape-from-shading, in: Ninth International Conference on Computer Vision (ICCV), 2003, pp. 862–869.

[16] S. Seshamani, M.D. Smith, J.J. Corso, M.O. Filipovich, A. Natarajan, G.D. Hager, Direct global adjustment methods for endoscopic mosaicking, in: Proceedings of SPIE Conference on Medical Imaging, 2009, p. 72611D.

[17] F. Mourgues, F. Devernay, È. Coste-Manière, 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery, in: IEEE and ACM International Symposium on Augmented Reality (ISAR'01), New York, United States, 2001, p. 191. http://dx.doi.org/10.1109/ISAR.2001.970537. URL: ⟨http://hal.inria.fr/inria-00267255⟩.

[18] D. Stoyanov, M.A. ElHelw, B.P.L. Lo, A.J. Chung, F. Bello, G.-Z. Yang, Current issues of photorealistic rendering for virtual and augmented reality in minimally invasive surgery, in: IV, 2003, pp. 350–359.

[19] R. Kozera, On shape recovery from two shading patterns, Int. J. Pattern Recognit. Artif. Intell. 6 (4) (1992) 673–692.

[20] R. Melo, J. Barreto, G. Falcão, A new solution for camera calibration and real-time image distortion correction in medical endoscopy-initial technical evaluation, IEEE Trans. Biomed. Eng. 59 (3) (2012) 634–644.

[21] J. Byrn, S. Schluender, C. Divino, J. Conrad, B. Gurland, E. Shlasko, A. Szold, Three-dimensional imaging improves surgical performance for both novice and experienced operators using the da Vinci Robot System, Am. J. Surg. 193 (4) (2007) 519–522. http://dx.doi.org/10.1016/j.amjsurg.2006.06.042.

[22] K. Keller, J.D. Ackerman, H. Fuchs, Switched pattern laser projection for real-time depth extraction and visualization through endoscopes, in: International Symposium on Optical Science and Technology, International Society for Optics and Photonics, 2002, pp. 101–108.

[23] K. Ikeuchi, B.K.P. Horn, Numerical shape from shading and occluding boundaries, in: B.K.P. Horn, M.J. Brooks (Eds.), Shape from Shading, MIT Press, Cambridge, MA, 1989, pp. 245–299.

[24] Y.G. Leclerc, A.F. Bobick, The direct computation of height from shading, in: Conference on Computer Vision and Pattern Recognition, Hawaii, 1991.

[25] R.G. Willson, S.A. Shafer, What is the center of the image? J. Opt. Soc. Am. A 11 (11) (1994) 2946–2955.

[26] K. Asari, S. Kumar, D. Radhakrishnan, A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation, IEEE Trans. Med. Imaging 18 (1999) 345–354.

[27] C. Wengert, P.C. Cattin, J.M. Duff, C. Baur, G. Székely, Markerless endoscopic registration and referencing, in: MICCAI, 2006.

[28] J.P. Barreto, J. Roquette, P. Sturm, F. Fonseca, Automatic camera calibration applied to medical endoscopy, in: BMVC, 2009.

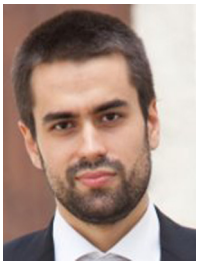
[29] M. Fischler, R. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.

[30] G. Ciuti, M. Visentini-Scarzanella, A. Dore, A. Menciassi, P. Dario, G.-Z. Yang, Intra-operative monocular 3d reconstruction for image-guided navigation in active locomotion capsule endoscopy, in: IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob) 2012, 2012.

Nuno Goncalves is a researcher at the Institute for Systems and Robotics and a Tenured Assistant Professor at the Department of Electrical and Computers Engineering of the University of Coimbra. He has received his M.Sc. and Ph.D. degrees in 2002 and 2008, respectively. His main research areas are computer vision and computer graphics with special emphasis to geometric problem in non-central projection vision systems and medical imaging. He has scientific publications in the following topics: omnidirectional vision, noncentral cameras, medical imaging, optics, camera models, motion estimation, pose estimation, reflections for image rendering, document classification, sports vision and legged robotics. He was a principal investigator of a finished project funded by the Portuguese Science and Technology Foundation in the area of non-central camera models for computer graphics and computer-aided surgery. He participates regularly as a reviewer for some of the most prestigious conferences and journals of the area of computer vision.

Diogo Roxo received the M.Sc. degree in biomedical engineering from the University of Coimbra, Portugal, in 2011. He was a research fellow in the Institute of Systems and Robotics, Coimbra, Portugal, since 2011 and he was working in camera calibration and 3D reconstruction using Endoscopic Images. His research interests include developing and testing computer vision algorithms in medical imaging.

Joao P. Barreto received the "Licenciatura" and Ph.D. degrees from the University of Coimbra, Coimbra, Portugal, in 1997 and 2004, respectively. From 2003 to 2004, he was a post- doctoral researcher with the University of Pennsylvania, Philadelphia. He has been an assistant professor with the University of Coimbra, since 2004, where he is also a senior researcher with the Institute for Systems and Robotics. His current research interests include different topics in computer vision, with a special emphasis in geometry problems and applications in robotics and medicine. He is the author of more than 30 peer-reviewed publications. He is also a regular reviewer for several conferences and journals, having received 4 Outstanding Reviewer Awards in the last few years.



Pedro Rodrigues received the M.Sc. degree in biomedical engineering from University of Coimbra, Portugal, in 2010. He was a research fellow in the AIBILI (Association for Innovation and Biomedical Research on Light and Image) from February 2011 to February 2014 and in March 2014 integrated the Institute of Systems and Robotics, Coimbra, Portugal, as a researcher. He was working in fully automatic early pathology detection, improved visualization, and diagnosis support with ocular fundus imaging data and recently started working in computer vision techniques for medical surgery. His research interests include computer vision, medical imaging and image processing.