

# Ray-based Calibration of Rigid Medical Endoscopes

J. P. Barreto<sup>1</sup>, J. M. Santos<sup>1</sup>, P. Menezes<sup>1</sup> and F. Fonseca<sup>2</sup>

<sup>1</sup> Institute for Systems and Robotics, Faculty of Science and Technology

<sup>2</sup> Hospitais da Universidade de Coimbra, Faculty of Medicine  
University of Coimbra, 3000 Coimbra, Portugal

**Abstract.** The medical endoscope consists in a rigid tube with optical elements inside that transfer the projection center of the objective lens at one end, to a virtual point at the other end. The endoscope is a non-conventional camera that presents severe non-linear distortions. Moreover it does not necessarily follow perspective projection, specially when mechanical torsion is applied to the probe. In order to study these aspects we describe a ray-based calibration procedure, with no prior assumptions about centrality or specific distortion models. The accuracy of the achieved calibration is slightly more than one tenth of millimeter. In addition, we quantify the errors in approximating the projection by a central model, and test the accuracy of the Bouguet's toolbox in calibrating this specific system.

## 1 Introduction

The borescope is an optical device used in industry and medicine for the observation of cavities with difficult or limited access (Fig. 1). The device consists of a rigid tube, with an eyepiece on one end and an objective lens on the other, linked by a relay optical system. The relay works as an optical repeater, moving the projection center from the objective lens to a virtual point at the back end of the rigid tube. The eyepiece is often fitted with a coupler lens to enable the use of a CCD camera for digital image acquisition. The optical relay system is usually one of two types: (i) achromatic doublets, consisting of pairs of lenses correctly spaced to drive the light rays; (ii) or gradient index rod lenses with variable index of refraction [1]. The former solution works well on large diameter borescopes, while the latter is better fitted for small diameters. In medicine the borescope is usually referred as *rigid endoscope*. Henceforth, we will use the name endoscope instead of borescope.

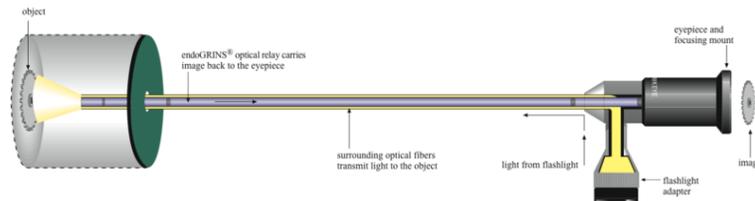
Endoscopic cameras are employed in different medical fields, ranging from orthopedics (arthroscopy) to abdominal surgery (laparoscopy). Endoscopes enable minimally invasive procedures with little or no injury to healthy organs and tissues. Most of these procedures are very difficult to execute, and even the best trained professionals make mistakes with inevitable consequences for the patient. In the last few years efforts have been done towards developing systems for computer aided surgery [2]. The idea is to assist the practitioner during the

intervention in order to minimize human error and improve clinical results. Many of these systems rely in the processing of endoscopic video to increase surgeon's perception and provide guidance inside the human body.

Camera calibration can hardly be avoided in the context of image-based computer aided surgery. However, developing an effective method for geometric calibration of a medical endoscope is a very challenging task. The reasons are the following: (i) the calibration result must be extremely accurate, because many medical procedures require positioning accuracies of the order of tenth of millimeter; (ii) the endoscopic optics introduces different types of aberrations, both geometric (non-linear distortions) and radiometric (e.g. vignetting); (iii) the configuration of the endoscope changes during operation, namely in oblique endoscopes where the scope cylinder can rotate around a symmetry axis; (iv) the endoscopic probe is not completely rigid and torsion during operation can misalign optical components making the projection non-central; (v) and finally the calibration procedure has to be performed by a non-expert in the Operation Room (OR), which requires the method to be simple, fast and robust.

Previous literature in endoscope calibration give substantial attention to the radial distortion issue. Smith et al. describe the endoscope distortion using orthogonal Chebyshev polynomials [3], while in [4], [5] and [6] the relation between distorted and undistorted distances to the center is provided by an  $N$  order polynomial in accordance to the standard model initially proposed by Conrady [7][8]. In all these works the expansion coefficients are computed using data from images of a calibration grid. The results of distortion compensation are evaluated qualitatively, by measuring the straightness of lines projected in the image [4] [5] [6], or the constancy of circle areas imaged at different locations [3].

In [9], Shahidi et al. describe a sophisticated calibration apparatus with opto-tracking to determine the 3D pose of camera and calibration target. The calibration is carried using Tsai's method [10] assuming a one parameter distortion model. They report mean projection errors of less that 0.5mm and 1mm at depths up to 25mm and 45mm respectively. In oblique viewing endoscopes the viewing direction can be changed by simply rotating the scope cylinder without having to move it. In [11] this rotation is determined using an encoder and opto-tracking, and Tsai's model is extended to take into account the changes in the viewing direction. A similar approach is followed in [12] with opto-tracking to measure the rotation and using Zhang's method for the calibration [13]. All



**Fig. 1.** Schematic view of a borescope (taken from wikipedia)

these works assume a perspective camera with radial distortion described by the standard Conrady's polynomial model. More recently Barreto et al. [14] carried calibration experiments using the parameter-free distortion correction proposed in [15]. Their results shadow some doubts about using the standard polynomial model [7][8] for describing radial distortion in medical endoscopes.

From the above discussion it seems unclear which calibration approach leads to better results, and how non-linear distortions should be described to obtain an accurate projection model. Moreover it is arguable if endoscopes are always central projection cameras, specially when the scope cylinder suffers significant mechanical torsions. We try to address these issues by performing the endoscope calibration using the ray-based approach proposed in [16]. The calibration consists in establishing the mapping between photosensitive elements on the image detector - the *pixels* - and incoming scene rays - the *raxels*. In [16] the raxels include geometric, radiometric and optical properties. In this article we will only consider the geometric aspects, which results in incident light rays being simply described by 3D lines. The main advantage of the ray-based calibration is its generality, without prior assumptions about central projection or specific distortion models. With this approach we manage to obtain a very accurate camera description, with RMS projection errors of 0.12mm at depths up to 45mm. This description is employed as ground truth for evaluating the errors in modeling the camera as being central. We also compare our calibration results with the ones obtained using Bouguet's implementation of Zhang's algorithm [13]. It is important to make clear that we are not advocating that non-parametric raxels should be used as the ultimate solution for the calibration of medical endoscopes. We simply want to obtain the best possible characterization of camera projection as a key step towards future calibration procedures that can provide the required accuracy, be able to cope with changes of configuration, and be used by a non-expert in the OR.

## 2 Experimental Apparatus for Ray-based Calibration

Ray based calibration maps image pixels into incoming light rays (the *raxels*). For each image pixel we must find at least two scene points that are projected at that exact image location. These points define a 3D line modeling the incident light ray.

Our experimental setup employs an active display in conjunction with an opto-tracker. Tab. 1 gives an account of the relevant equipment specifications. The display is instrumented with a marker of LEDs that can be tracked by the opto-tracker. The opto-tracker computes the pose (rotation and translation) of the screen plane, allowing points on the display to be referenced with respect to a system of coordinates in the 3D world. Moreover, the active display enables the projection of patterns suitable to establish a dense mapping between pixels in the image and pixels in the screen [17]. This dense mapping, complemented by accurate localization of the screen plane, provides a set of 3D data points with known image projection. By moving the display and repeating the procedure it is

possible to obtain new data sets. The final ray-based calibration is accomplished by estimating the 3D lines joining corresponding points in each data set (see Fig. 3). This section discusses in detail the dense mapping between the image plane and the display plane.

**Table 1.** Technical specifications of the equipment used in the experimental apparatus.

<b>Endoscopic Camera</b> (Smith & Nephew 460H)		<b>Opto-tracker</b> (NDI CERTUS)		<b>Active Display</b>	
Img. Size (pix.)	576 × 720	Res. (mm)	0.01	Res. (pix)	1050 × 1680
Img. Area (%)	46	Acc. (mm)	0.15	Pix. Size (μm)	258 × 258
Vid. Out.	DV-25	Op. Range (m)	2.5 ~ 6	Resp. (ms)	2

## 2.1 Mapping between Image and Active Display

Let each point on the screen plane  $\mathbf{p}$  be projected into an image point  $\mathbf{q}$ , and consider the following back projection function

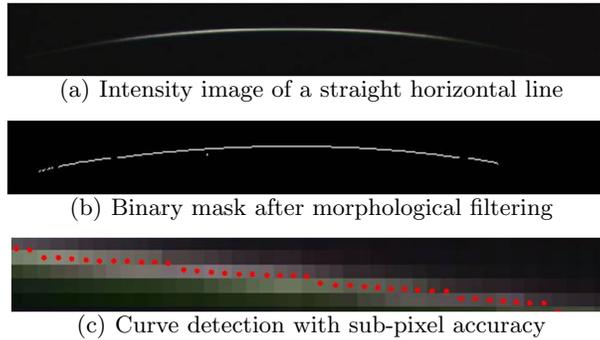
$$\mathbf{p} = \mathbf{f}(\mathbf{q}) = (f_x(\mathbf{q}), f_y(\mathbf{q}))^T.$$

$\mathbf{f}$  maps image points into points on the display. Our goal is to make a dense evaluation of the function at every discrete image pixel location.

The discrimination between pixels can be accomplished by using the active display to project patterns with distinguishable features. Salvi et al. compare different codification strategies for generating these patterns [17]. The most popular approach uses simple gray coding. Given a display with  $K$  locations, we can make each point distinct by using a sequence of  $\log K$  bit coded images/patterns. However, and since the endoscope operates at close depth ranges (1 to 5 cm), an accurate dense mapping would only be possible with a very high resolution display. The problem of lack of resolution gets even worse if we take into account the compression caused by the strong barrel distortion at image boundaries. Thus, we propose an alternative approach that, despite of time consuming, provides results with sub-pixel accuracy.

In our method the scalar functions  $f_x$  (horizontal axis) and  $f_y$  (vertical axis) are evaluated separately. For sampling  $f_x$  we display a vertical white line over a dark background at a known screen location. The line is projected into a curve, and image points  $\mathbf{q}$  on the curve are detected (see algorithm below). The line is moved to the next position and the image detection repeated till sweeping the entire camera field of view. At the end of the procedure we obtain a very dense sampling of function  $f_x$ . The function is evaluated at discrete pixel location using bilinear interpolation over a mesh of delaunay triangles. The method for  $f_y$  is similar, but displaying horizontal lines instead of vertical ones.

Fig. 2 illustrates the detection procedure for the case of displaying an horizontal line. The straight line is projected into a curve because of non-linear lens



**Fig. 2.** Detecting the curve corresponding to the projection of a straight line

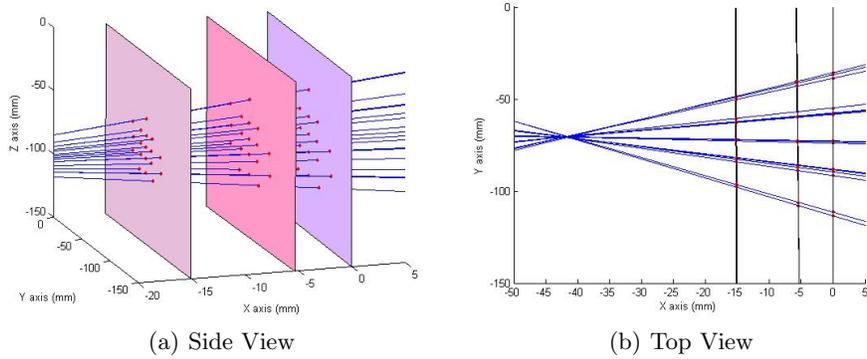
distortion. The vignetting effect causes the curve to vanish towards the image boundaries. The endoscopic camera system performs automatic white balance, but the compensation is never perfect. The outline of the detection algorithm is as follows:

1. For each display input acquire 20 images, convert to gray-scale, and use the average to increase the robustness to noise (Fig. 2(a)).
2. Apply a top-hat operator to enhance curve at image boundaries, and use a median filter to remove spurious artifacts.
3. Convert to binary using adaptive thresholding, and apply a skeleton operator to obtain a mask with 1 pixel thickness (Fig. 2(b)).
4. Compute the intensity gradient in the initial frame and keep the orientation information.
5. For each pixel in the mask, fit a gaussian curve to the intensity values along the gradient direction.
6. The points detected with sub-pixel accuracy are the centers of the computed gaussians (Fig. 2(c)).

## 2.2 Acquired Data

The described line sweeping approach is used to back-project image pixels  $\mathbf{q}$  into points  $\mathbf{p}$  in the display plane. Display pixel coordinates are converted into metric units (see pixel size in Tab. 1), and referenced in the world reference frame using the opto-tracker readings. For each run of the procedure, image pixels  $\mathbf{q}$  are mapped into a set of co-planar 3D points  $\mathbf{Q}$ . Since a 3D line is uniquely defined by two points, then two runs with different test planes are enough for achieving the ray-based calibration. However, we will be using  $N = 3$  test planes in order to have redundant data for evaluating accuracy and detect possible experimental errors.

We ran two distinct experiments, henceforth referred as NORMAL and TORSION. In the former the endoscope is on its natural stand, and in the latter a



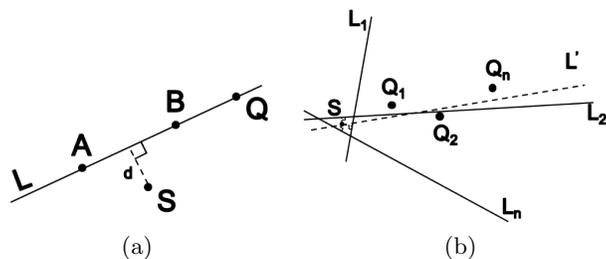
**Fig. 3.** Ray-based calibration (NORMAL). Each test plane provides 3D points indexed by image pixel coordinates. Raxels are determined by fitting 3D lines to sets of corresponding points across planes.

weight is hanged from the lens tip to cause mechanical torsion of the cylindrical scope. In both cases we collected data from three different plane positions. The test planes were approximately front-parallel to the camera, with the closest one at less than 30 mm from the endoscope tip, and the farthest at no more than 45 mm (see Fig. 3). The next section introduces the mathematical background required to process the acquired data, while section 4 discusses the final results of the ray-based calibration for both NORMAL and TORSION.

### 3 Mathematical Background for Data Processing

With the setup described above, we find for each image pixel a set of  $N$  points, with  $N$  being the number of test planes. These points define a line in 3D that is the raxel incident on that particular image location (remember that we will be neglecting the radiometric and optical properties of the raxel [16]). A 3D line is a geometrical entity that can be represented in different manners [18]. In this paper lines will be parametrized using Plücker coordinates. This section reviews basic concepts about Plücker coordinates and discusses the estimation methods used to process the experimental data.

The notation is as follows: matrices are represented by symbols in sans serif font, e.g.  $\mathbf{M}$  and vectors by bold symbols, e.g.  $\mathbf{Q}$ . Equality of matrices or vectors up to a scalar factor is written as  $\sim$ .  $[\mathbf{A}]_{\times}$  denotes the skew-symmetric matrix associated with the cross product of 3-vectors, e.g.  $\mathbf{A} \times \mathbf{B} = [\mathbf{A}]_{\times} \mathbf{B}$ .



**Fig. 4.** The left figure shows a 3D line  $\mathbf{L}$ , defined by points  $\mathbf{A}$  and  $\mathbf{B}$ , with euclidean/orthogonal distance  $d$  to  $\mathbf{S}$ . In the right,  $\mathbf{S}$  is the point that minimizes the sum of the square of the euclidean distances to lines  $\mathbf{L}_j$ .  $\mathbf{L}'$  is the line going through  $\mathbf{S}$ , that minimizes the sum of the square of the algebraic distances to points  $\mathbf{Q}_i$ .

### 3.1 Plücker Coordinates

Consider two points in the 3D space with non-homogeneous coordinates  $\mathbf{A}$  and  $\mathbf{B}$  (Fig. 4(a)). These two points define a line with Plücker parametrization

$$\mathbf{L} \sim \begin{pmatrix} \mathbf{D} \\ \mathbf{M} \end{pmatrix} \sim \begin{pmatrix} \mathbf{A} - \mathbf{B} \\ \mathbf{A} \times \mathbf{B} \end{pmatrix}.$$

Since  $\mathbf{L}$  is an homogeneous vector, the equality holds up to a scalar factor. Remark that  $\mathbf{D}$  is the line direction and  $\mathbf{M}$  is always orthogonal to the plane defined by line and the origin of the reference frame. As a matter of fact, not every  $6 \times 1$  vector represents a line in 3D space.  $\mathbf{L}$  is a Plücker vector *iff* it belongs to the Klein quadric in the 5D projective space [19], which is defined by

$$\mathbf{L}^\top \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} \mathbf{L} = 0 \quad (1)$$

with  $\mathbf{I}$  being the  $3 \times 3$  identity matrix.

Let  $\mathbf{Q}$  and  $\mathbf{S}$  be also points in 3D as shown in Fig. 4(a). If the line  $\mathbf{L}$  goes through  $\mathbf{Q}$ , then the following equality holds

$$\mathbf{Q} \times \mathbf{D} + \mathbf{M} = 0 \quad (2)$$

It can also be shown that the euclidean distance  $d$  between the line and a generic point  $\mathbf{S}$  is given by

$$d = \sqrt{\mathbf{S}^\top \mathbf{S} + \frac{\mathbf{M}^\top \mathbf{M} - (\mathbf{S}^\top \mathbf{D})^2 - 2\mathbf{S}^\top (\mathbf{M} \times \mathbf{D})}{\mathbf{D}^\top \mathbf{D}}} \quad (3)$$

### 3.2 Estimation of the Line going Through $N$ points

From the experiment, for each image point  $\mathbf{q}$  we obtain a set of 3D points  $\mathbf{Q}_i$ , with  $i = 1, 2, \dots, N$ . The raxel incident in  $\mathbf{q}$  is the 3D line  $\mathbf{L}$  going through points

$\mathbf{Q}_i$ . From equation 1 comes that  $\mathbf{L}$  must satisfy

$$\underbrace{\begin{pmatrix} [\mathbf{Q}_1]_{\times} & \mathbf{I} \\ [\mathbf{Q}_2]_{\times} & \mathbf{I} \\ \vdots & \vdots \\ [\mathbf{Q}_N]_{\times} & \mathbf{I} \end{pmatrix}}_{\mathbf{A}} \mathbf{L} = 0 \quad (4)$$

In general the linear system does not have a non trivial solution. Matrix  $\mathbf{A}$  is full rank because the noise in measurements causes points  $\mathbf{Q}_i$  not to be perfectly collinear. We can obtain a solution for  $\mathbf{L}$  in the least squares sense, by considering the SVD decomposition

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T,$$

and selecting the column of  $\mathbf{V}$  corresponding to the smallest singular value. The problem is that for  $\mathbf{L}$  to be a valid Plücker vector it must satisfy equation 1. Fortunately, the particular structure of matrix  $\mathbf{A}$  guarantees that every column of  $\mathbf{V}$  is a  $6 \times 1$  vector lying in the Klein quadric.

*Proof:* From the SVD decomposition of  $\mathbf{A}$ , and taking into account that  $\mathbf{U}$  and  $\mathbf{V}$  are orthonormal matrices, it comes that

$$\mathbf{A}^T \mathbf{A} \mathbf{V} = \mathbf{V} \mathbf{S}^2$$

Let  $\mathbf{v}_j$  be the columns of matrix  $\mathbf{V}$  and  $s_j$  the corresponding singular values. From the previous equation it is true that

$$\mathbf{A}^T \mathbf{A} \mathbf{v}_j = s_j^2 \mathbf{v}_j, \forall j=1,2,\dots,6 \quad (5)$$

Moreover, the structure of the symmetric matrix  $\mathbf{A}^T \mathbf{A}$  is

$$\mathbf{A}^T \mathbf{A} = \begin{pmatrix} -\sum_{i=1}^N [\mathbf{Q}_i]_{\times}^2 & \sum_{i=1}^N [\mathbf{Q}_i]_{\times}^T \\ \sum_{i=1}^N [\mathbf{Q}_i]_{\times} & \mathbf{I} \end{pmatrix},$$

Taking into account the properties of the vector cross product, it is easy to see that the equality of equation 5 holds *iff*

$$\mathbf{v}_{u,j}^T \mathbf{v}_{l,j} = 0, \forall j=1,2,\dots,6$$

with  $\mathbf{v}_{u,j}$  and  $\mathbf{v}_{l,j}$  being the upper and lower 3-vectors of  $\mathbf{v}_j$ . This means that every column of matrix  $\mathbf{V}$  satisfies equation 1 and is a valid Plücker representation.

□

### 3.3 Approximating a General Set of Lines by a Pencil

Consider the set of  $K$  lines  $\mathbf{L}_j$  shown in Fig. 4(b). We aim to compute the 3D point  $\mathbf{S}$  that minimizes the sum of the square of the orthogonal distances  $d_j$  given by equation 3:

$$\min f = \sum_{j=1}^K d_j^2$$

Computing the zero of the gradient it comes that the minimum is

$$\mathbf{S} = \left( \sum_{j=1}^K \mathbf{I} - \frac{\mathbf{D}_j \mathbf{D}_j^\top}{\mathbf{D}_j^\top \mathbf{D}_j} \right)^{-1} \left( \sum_{j=1}^K \frac{\mathbf{M}_j \times \mathbf{D}_j}{\mathbf{D}_j^\top \mathbf{D}_j} \right)$$

with  $\mathbf{D}_j$  and  $\mathbf{M}_j$  the upper and lower 3-vectors of  $\mathbf{L}_j$ , and  $\mathbf{I}$  the  $3 \times 3$  identity matrix.

Let's now estimate the line

$$\mathbf{L}' \sim \begin{pmatrix} \mathbf{D}' \\ \mathbf{M}' \end{pmatrix}$$

going through points  $\mathbf{Q}_i$ , with  $i = 1, 2, \dots, N$ , and satisfying

$$\mathbf{S} \times \mathbf{D}' + \mathbf{M}' = 0$$

Remark that this is a different problem from the one discussed in section 3.2, because we are enforcing point  $\mathbf{S}$  to be in the line. From the previous equation it comes that

$$\mathbf{M}' = -\mathbf{S} \times \mathbf{D}'.$$

Replacing the lower 3-vector of  $\mathbf{L}'$ , the matrix equation 4 becomes

$$\underbrace{\begin{pmatrix} [\mathbf{Q}_1]_\times - [\mathbf{S}]_\times \\ [\mathbf{Q}_2]_\times - [\mathbf{S}]_\times \\ \vdots \\ [\mathbf{Q}_N]_\times - [\mathbf{S}]_\times \end{pmatrix}}_{\mathbf{B}} \mathbf{D}' = 0$$

Thus, the direction  $\mathbf{D}'$  can be estimated from the SVD decomposition of matrix  $\mathbf{B}$ , and  $\mathbf{M}'$  is subsequently computed from the cross product between  $\mathbf{D}'$  and  $\mathbf{S}$ . Remark that, while the estimation of  $\mathbf{S}$  minimizes euclidean distances, the estimation of  $\mathbf{L}$  uses algebraic distances. This is explained by the non existence of a closed form solution for the line minimizing euclidean distances.

### 3.4 Intersection between a Line and a Plane

The measurement of errors in raxel estimation will require the computation of the points where a certain line  $\mathbf{L}$  intersects the test planes. A 3D line can be

**Table 2.** Global errors for the entire image. The back-projection error is measured as the distance between data points and points where the raxel intersects the test planes. The distance to center is the distance from the original rays to the hypothetical projection center (point  $\mathbf{S}$  in central enforcing and camera axis origin on Bouguet).

	NORMAL			TORSION	
	Ray-based	Enforce Central	Bouguet	Ray-based	Enforce Central
<b>RMS back-projection error (mm)</b>	0.1280	0.1315	0.3499	0.1157	0.1202
<b>RMS distance between raxels and hypothetical center (mm)</b>	-	0.1435	0.4171	-	0.2680

represented by, either a Plücker vector  $\mathbf{L}$ , or a  $4 \times 4$  matrix known as the Plücker matrix [18]. The Plücker matrix  $\mathbf{L}$  and the Plücker vector  $\mathbf{L}$  relate as follows

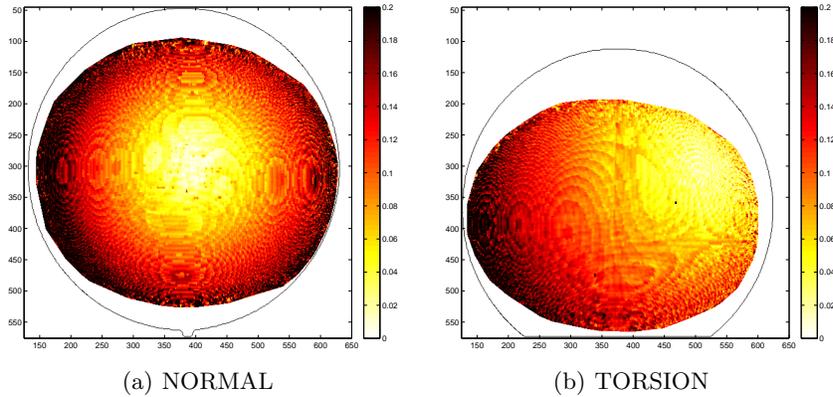
$$\mathbf{L} \sim \begin{pmatrix} -[\mathbf{M}]_{\times} & \mathbf{D} \\ -\mathbf{D}^{\top} & 0 \end{pmatrix}$$

Let  $\mathbf{\Pi}$  be an homogeneous 4-vector representing a 3D plane. The non-homogeneous coordinates  $\mathbf{R}$  of the intersection point between the line  $\mathbf{L}$  and the plane  $\mathbf{\Pi}$ , can be computed in a straightforward manner using the following formula [18]:

$$\begin{pmatrix} \mathbf{R} \\ 1 \end{pmatrix} \sim \mathbf{L} \mathbf{\Pi}$$

## 4 Calibration Results

This section discusses the results of the ray-based calibration for the endoscope under NORMAL and TORSION situations (c.f. section 2.2). For each case, data was collected from  $N = 3$  test planes, and lines were fitted to corresponding points across planes using the estimation algorithm described in section 3.2. Fig. 3 shows the results for the NORMAL experiment, where it can be seen that raxel paths go very close to a single 3D location. This suggests that projection is close to be central. In order to quantify the errors in assuming a central model, we computed the point  $\mathbf{S}$  minimizing the sum of square distances to raxels and fitted a pencil of lines to the original data points (c.f. section 3.3). A final study was made about using Bouguet’s toolbox to calibrate the camera from a set of images of a planar calibration grid. The calibration images were emulated assuming the initial non-central raxel model. The virtual planar grids were randomly placed in 3D space, and the calibration images synthesized using interpolation over the intersections between raxels and plane. Tab. 2 shows global Root Mean Square (RMS) errors for the entire image, while Fig. 5, 6 and 7 show how the error behaves across image.

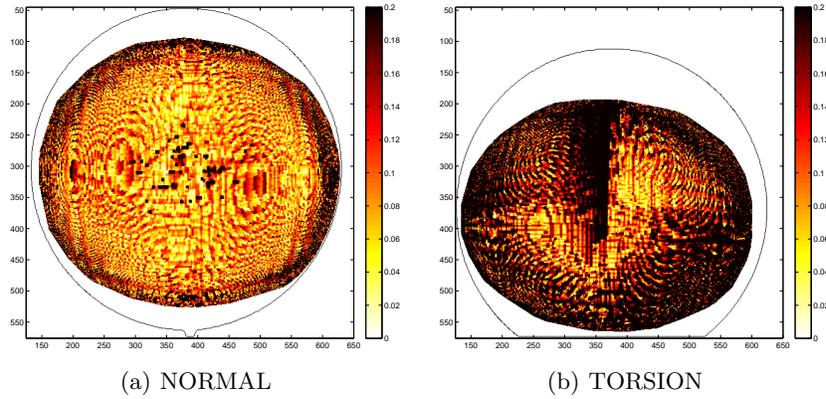


**Fig. 5.** RMS back-projection error for different image regions (the color scale ranges from 0 to 0.2 mm). Each calibrated pixel is assigned with the corresponding RMS of the error measured across the 3 test planes. The left side concerns the NORMAL situation, and the right side the TORSION case. In the former data was obtained for 83% of the image area, while in the latter the calibration was achieved for only 74% of the pixels. The camera field of view is almost  $100^\circ$ . The closed black curve represents the image boundary, and enables to evaluate the amount of torsion applied.

#### 4.1 Ray-based Calibration

The ray-based calibration is achieved by fitting lines to data points lying in  $N = 3$  test planes using the estimation algorithm described in section 3.2. The 3 data points across planes are not perfectly aligned because of noise. We test the errors in alignment, henceforth referred as back-projection errors, in order to evaluate the accuracy of the experimental procedure. This is done by intersecting the raxel with each test plane (cf. Section 3.4), and measuring the distance to the corresponding data points used for the estimation. This metric is used instead of simple orthogonal distances  $d$  (equation 3), because it was observed that the latter benefits projection rays in the image periphery, with larger incidence angles with respect to the planes. Tab. 2 shows the global Root Mean Square (RMS) error for the ray-based calibration, measured over the entire image under NORMAL and TORSION situations. Fig. 5 shows the errors for each calibrated image pixel.

Fig. 5(a) shows the error increasing from the image center to the periphery. We believe this is due to imperfect vignetting compensation and the compression caused by radial distortion. During the line sweeping procedure the decrease in intensity makes the point detection more noisy towards the image periphery. The distortion aggravates the problem causing the lines to be closer and less distinguishable. As expected the values for the global RMS back-projection error are similar for NORMAL and TORSION (the difference in the order of  $\mu m$  in Tab. 2 is negligible). However Fig. 5 shows that the spatial distribution across



**Fig. 6.** Orthogonal distance between the estimated projection center  $\mathbf{S}$  and the non-central raxels (the color scale ranges from 0 to 0.2 mm). Apparently the central projection model fits better the NORMAL case than the TORSION situation.

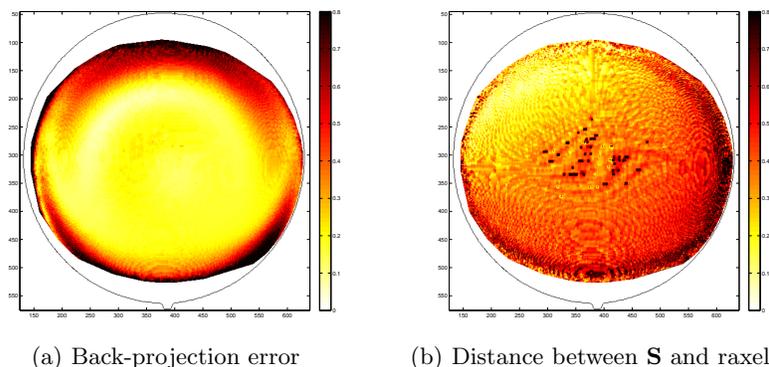
image is quite different, indicating that mechanical torsion causes important changes in terms of distortion and radiometric behaviors.

#### 4.2 Approximation by a Central Projection Model

In this experiment we used the estimation procedure outlined in section 3.3 to enforce a central projection model. Fig. 6 shows the orthogonal distances between the estimated center  $\mathbf{S}$  and each raxel for both NORMAL and TORSION. For NORMAL the distance is quite uniform across pixels in the image, suggesting that the projection is nearly central. This assumption is confirmed by the almost negligible increase in the global RMS back-projection error (Tab. 2). For TORSION the distance to the center is in average twice the one observed for the NORMAL situation. Apparently the mechanical torsion affects the optical relay causing a decentering effect. Surprisingly the global RMS back-projection error is still quite good, which means that in practice the central assumption can be used even under heavy mechanical torsion of the cylindrical scope.

### 5 Calibration using the Bouguet Toolbox

Finally we tried to calibrate the endoscope in NORMAL situation using the standard Bouguet toolbox. The calibration was performed assuming a  $7^{th}$  order polynomial for modeling radial distortion (3 distortion parameters). We used as input 10 images of a planar calibration grid, randomly placed in 3D at depths ranging from 20 to 50 mm. The correspondences were synthetically generated by projecting the grid assuming the ray-based calibration obtained in section 4.1. Each calibration plane provided 100 data points spread uniformly in the image including the periphery. The results are shown in Fig. 7 and Tab. 2.



**Fig. 7.** Results for Bouguet calibration under NORMAL situation. The left side shows the distribution of the back-projection errors across image, and the right side the distances between the estimated projection center and the raxels. The color scale ranges from 0 to 0.8 mmm. It is clear that Bouguet calibration, even using a 7th order polynomial model for distortion, is far from achieving the accuracies observed when the central projection model is directly fitted to the ray-based calibration (c.f. section 4.2)

## 6 Conclusions

This article describes the ray-based calibration of a medical rigid endoscope [16]. We suggest an experimental procedure for acquisition of 3D points, as well as the mathematical background required to process the data and estimate the incident light rays. It is important to emphasize that the endoscope is designed to operate at very close depths, precluding the use of binary patterns generated by standard resolution active displays. We overcome the problem using the line sweeping approach described in section 2.1, and report back-projection accuracies of 0.12 mm for the achieved ray-based calibration.

Moreover, the projection in medical endoscopes seems to be well approximated by a central model, even under strong mechanical torsion of the cylindrical scope. The decrease in accuracy is negligible, and images can still be used for computer aided surgery applications with high positioning requirements. Finally, Bouguet’s toolbox is used to calibrate the camera. The observed accuracy in back-projection is almost 3 times worst when compared to the one obtained by directly fitting a pencil of lines to the incoming rays. This suggests that in the future it is important to develop a method that takes into account the specificities of the calibration problem herein enunciated. Our work is an important step towards this direction.

## Acknowledgements

The authors are grateful to the Portuguese Science Foundation by generous funding through grant PTDC/EEA-ACR/68887/2006.

## References

1. Charles Y. Liu, M.Y.W., Apuzzo, M.L.J.: The physics of image formation in the neuroendoscope. *Child's Nervous System* **20** (2004) 777–782
2. Berlinger, N.T.: Robotic surgery. squeezing into tight places. *New England Journal of Medicine* (2006)
3. Smith, W.E., Vakil, N., Maislin, S.A.: Correction of distortion in endoscope images. *IEEE Trans. on Medical Imaging* (1992)
4. Haneishi, H., Yagihashi, Y., Miyake, Y.: A new method for distortion correction of electronic endoscope images. *IEEE Trans. on Medical Imaging* (1995)
5. Asari, K.V., Kumar, S., Radhakrishnan, D.: A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation. *IEEE Trans. on Medical Imaging* (1999)
6. Helferty, J.P., Zhang, C., McLennan, G., Higgings, W.E.: Videoendoscopic distortion correction and its application to virtual guidance of endoscopy. *IEEE Trans. on Medical Imaging* (2001)
7. Conrady, A.: Decentering lens systems. *Monthly notices of the Royal Astronomical Society* **79** (1919) 384–390
8. Brown, D.: Decentering distortion of lenses. *Photogrammetric Engineering* **32** (1966) 444–462
9. Shahidi, R., Bas, M., Maurer, C., Johnson, J., Wilkinson, E., Wang, B., West, J., Citardi, M., Manwaring, K., Khadem, R.: Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Trans. on Medical Imaging* (2002)
10. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation* (1987)
11. Yamaguchi, T., Nakamoto, M., Sato, Y., Konishi, K., Hashizume, M., Sugano, N., Yoshikawa, H., Tamura, S.: Development of a camera model and calibration procedure for oblique-viewing endoscopes. In: *Computer Aided Surgery*. (2004)
12. Wuy, C., Narasimhan, S.G., Jaramazy, B.: Shape-from-shading under near point lighting and partial views for orthopedic endoscopy. In: *Workshop on Photometric Analysis For Computer Vision*. (2007)
13. Zhang, Z.: Flexible camera calibration by viewing a plane from unknown orientations. In: *International Conference in Computer Vision*. (1999)
14. Barreto, J.P., Swaminathan, R., Roquette, J.: Non parametric distortion correction in endoscopic medical images. In: *3DTV - Conference. The True Vision Capture, Transimission and Display of 3D Video*. (2007)
15. Hartley, R.I., Kang, S.B.: Parameter-free radial distortion correction with centre of distortion estimation. In: *International Conference in Computer Vision*. (2005)
16. Grossberg, M.D., Nayar, S.K.: The raxel imaging model and ray-based calibration. *Int. J. Comput. Vision* **61** (2005) 119–137
17. Salvi, J., Pages, J., Batlle, J.: Pattern codification strategies in structured light systems. *Pattern Recognition* **37** (2004) 827–849
18. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*. 2nd edn. Cambridge University Press (2004)
19. Semple, J., Kneebone, G.: *Algebraic Projective Geometry*. Clarendon Press (1998)