

A full geometric and photometric calibration method for oblique-viewing endoscopes

C. Wu, B. Jaramaz & S.G. Narasimhan

To cite this article: C. Wu, B. Jaramaz & S.G. Narasimhan (2010) A full geometric and photometric calibration method for oblique-viewing endoscopes, Computer Aided Surgery, 15:1-3, 19-31, DOI: [10.3109/10929081003718758](https://doi.org/10.3109/10929081003718758)

To link to this article: <https://doi.org/10.3109/10929081003718758>



Published online: 30 Apr 2010.



Submit your article to this journal [↗](#)



Article views: 804



View related articles [↗](#)



Citing articles: 3 View citing articles [↗](#)

BIOMEDICAL PAPER

A full geometric and photometric calibration method for oblique-viewing endoscopes

C. WU, B. JARAMAZ, & S.G. NARASIMHAN

Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh

(Received 28 January 2008; accepted 20 September 2009)

Abstract

Oblique-viewing endoscopes (oblique scopes) are widely used in computer assisted surgeries. The viewing direction of an oblique scope can be changed by rotating the scope cylinder; this extends the field of view, but also makes the geometric calibration process more difficult. Although few calibration methods have yet been developed, calibration is critical for the application of augmented-reality technologies such as stereo vision to procedures involving oblique scopes. Moreover, to our knowledge, no photometric calibration method has yet been introduced for endoscopes, even though such calibration is important for illumination-based visualization techniques such as shape-from-shading. In this paper, we present a complete calibration process for oblique-viewing endoscopes, estimating both geometric and photometric properties. Experimental results demonstrate that our methods are practical and accurate.

Keywords: *Calibration, oblique-viewing endoscope, geometric, photometric*

Introduction

One of the main goals of computer assisted orthopedic surgery is to enable true minimally invasive surgery (MIS). As a key MIS tool, the endoscope is attracting increasing attention for its potential role in computer assisted surgery, especially in conjunction with a surgical navigation system. By tracking the endoscope in space using a position localizer, its role can be significantly augmented. For example, it can be used to create MIS augmented-reality systems, to merge virtual and real endoscopic images [1], to overlay real endoscopic images with 3D surfaces derived from CT images [2], or to recover the 3D shape from a single endoscopic image [3, 4] or multiple images [5–7].

Geometric calibration of the camera, an important step in endoscope-related applications and mostly based on Tsai's model [8], has been addressed in several publications [2, 9–11]. However, with the exception of that by Yamaguchi

et al. [11], most of these reports concern the forward-viewing endoscope, in which the viewing direction is aligned with the axis of the scope. The range of movement of such a tool is constrained by the small incision, so the field of view is very small. To provide a sideways view, the oblique scope has been designed with a tilted viewing direction, and a wider viewing field can be obtained by rotating the scope cylinder. Figure 1 illustrates an oblique-viewing endoscope. Rotation occurs between the scope cylinder and the camera head. Yamaguchi et al. first modeled and calibrated the oblique scope [11, 12]. They formulated the rotation parameter of the scope cylinder as another external parameter in Tsai's camera model, and used two extra transformations to compensate for the rotation θ of the lens system and stillness of the camera head. While their camera model successfully compensates for the rotation effect, the method requires five additional parameters and the model is complicated. In this paper we propose an alternative approach to simplify the calibration in which an

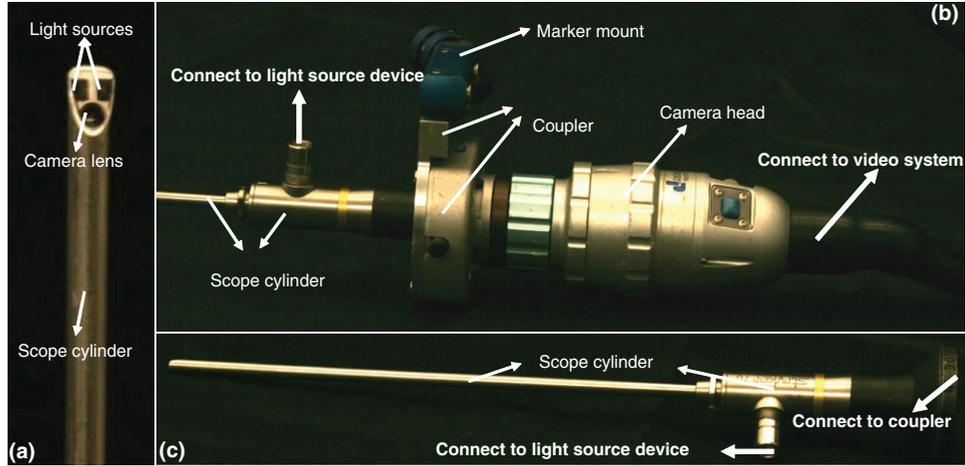


Figure 1. Stryker 344-71 arthroscope Vista (70 degree, 4 mm). An oblique endoscope consists of a scope cylinder with a lens and point light sources at the tip (which is tilted at an angle from the scope cylinder axis), a camera head that captures video images, and a light source device that supports the illumination. The scope cylinder is connected to the camera head via a coupler. This connection is flexible to enable rotation of the scope cylinder and camera head separately or together.

optical marker is attached to the scope cylinder, instead of to the camera head, with a newly designed coupler (as shown in Figure 1b). As a result, our camera model is simpler and we only need to estimate one additional parameter.

Photometric calibration of the camera, another important process in illumination-related applications, is performed to find the relationship between the image irradiance and image intensity for the camera. This relationship is called the camera response function. Traditional photometric calibration recovers the camera response function by changing the camera's exposure time. Compared with regular cameras, it is hard to control the exposure time for an endoscope, and the light spatial distribution can be anisotropic. We have therefore developed a method to calibrate all of these unknown parameters simultaneously.

Geometric calibration

The common orthopedic endoscope has a single camera and one or more point light sources at the tip of the scope. For this work, we used two oblique endoscopes as examples (shown in Figures 1 and 6).

Model for oblique-viewing endoscope

The camera model of Yamaguchi et al. is based on that of Tsai [8, 9]:

$$\begin{aligned} \lambda p_i &= A \cdot {}^cT_m(\theta) \cdot {}^mT_w \cdot P_w \\ {}^cT_m(\theta) &= T_R(-\theta; l_h(\theta))T_R(\theta; l_s) {}^cT_m(0) \end{aligned} \quad (1)$$

where P_w is a 3D point in the world coordinate system, p_i is the corresponding 2D image pixel, mT_w is a rigid transformation from the world coordinates to the optical marker coordinates, and ${}^cT_m(\theta)$ is a rigid transformation from the marker (camera head) to the camera coordinates. ${}^cT_m(\theta)$ is dependent on the rotation angle θ . By considering the marker coordinates (camera head) as a reference, only the lens system rotates while the camera head, i.e., the image plane, remains fixed irrespective of the rotation. Yamaguchi et al. [11, 12] describe such a transformation due to the rotation by decomposing the single physical rotation into two mathematical rotations: $T_R(\theta; l_s)$ is a rotation of both the scope cylinder and the camera head (image plane) around the axis of the cylinder l_s ; $T_R(-\theta; l_h(\theta))$ is an inverse rotation of the image plane around the z-axis of the lens system l_h . Both l_s and l_h have two unknown parameters. Although this model works well, it is very complicated.

Figure 2a shows the geometric model modified from that of Yamaguchi et al. by attaching an optical marker to the scope cylinder instead of to the camera head.

$$\begin{aligned} \lambda p'_i &= A \cdot {}^cT_m \cdot {}^mT_w \cdot P_w \\ p_i &= R(\theta) \cdot (p'_i - cc) + cc \end{aligned} \quad (2)$$

where P_w is a 3D point in the world coordinates, p'_i is the corresponding 2D image pixel without rotation, p_i is the image pixel with rotation θ , mT_w is a rigid transformation from the world coordinates to the optical marker coordinates, cT_m is a rigid transformation from the marker (scope cylinder) to

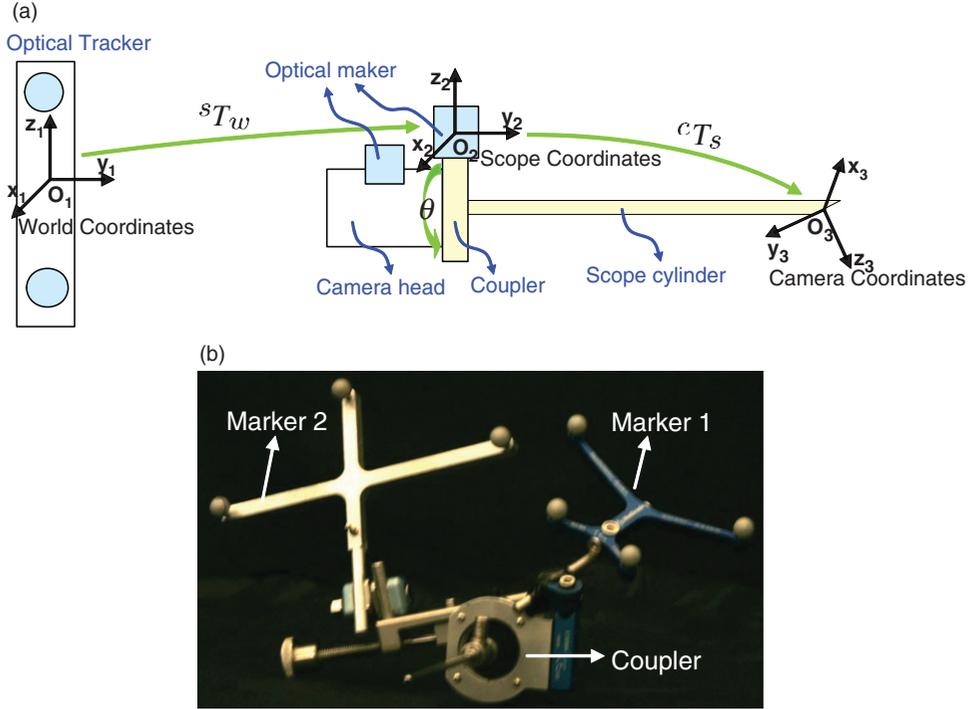


Figure 2. The geometric model of an endoscope based on a tracking system. A new coupler (see Figure 1b) has been designed to enable mounting of an optical marker on the scope cylinder, ensuring that the transformation from the scope (marker) coordinates O_2 to the lens system (camera) coordinates O_3 is fixed. The optical tracker defines the world coordinates O_1 . Two optical markers are attached to the coupler and camera head separately to compute the rotation θ between them.

the camera coordinates and independent of θ , cc is the principal point which is an intrinsic parameter, and $R(\theta)$ represents a rotation of the image plane around cc by θ . Thus, the camera intrinsic matrix A and external matrix cT_m can be calibrated by using Zhang's method [9] and mT_w can be obtained directly from the tracking system. In our model we only need to estimate the rotation angle.

The rotation angle can be estimated by using a rotary encoder, as done by Yamaguchi et al. [12]. When this is absent, the rotation angle can be estimated by using two optical markers; one attached to the scope cylinder and the other mounted on the rod (the camera head).

A comparison between our model and that of Yamaguchi et al. is presented in Figure 3. Yamaguchi et al. used the camera head as a reference coordinate in their hand-eye calibration system. As surgeons rotate the scope cylinder with respect to the camera head in order to obtain a sideways view, it is natural to consider the camera head as a reference. However, this makes the camera model very complex. To approach this from the opposite direction, no matter how the surgeons rotate the scope cylinder, if the reference coordinate is on the cylinder itself, the lens system is fixed with

respect to the cylinder but the camera head rotates around θ . Thus, the external parameters are no longer affected by the rotation. Since the image plane is in the camera head, the rotation only affects the image plane. Our method is therefore developed based on the above observations. The model of Yamaguchi et al. requires five more parameters, whereas we only need one. Likewise, their method uses two optical markers and one rotary encoder, whereas our method only requires two optical markers.

Estimation of rotation angle using two optical markers

Let the marker attached to the scope cylinder be Marker 1 and that attached to the rod (camera head) be Marker 2 (Figure 2b). As shown in Figure 4, when we rotate the camera head around the scope cylinder by θ , point P_r in Marker 2's coordinates O_2 will move along a circle with respect to a point O on the axis of the scope cylinder, in Marker 1's coordinates O_1 . Thus, we can estimate the center O of the circle first and compute θ as

$$\theta = \arccos \frac{\|O\vec{P}_r^A\|^2 + \|O\vec{P}_r^B\|^2 - \|P_r^A P_r^B\|^2}{2\|O\vec{P}_r^A\| \cdot \|O\vec{P}_r^B\|} \quad (3)$$

| | Yamaguchi et al.'s system | Our system |
|-----------------------------|--|--|
| System | | |
| Reference part | Rod (Camera head) | Scope cylinder |
| Rotated part | Scope cylinder | Rod (Camera head) |
| Fact | <ol style="list-style-type: none"> 1. Lens system is rotated around the scope cylinder by θ 2. Image plane is fixed | <ol style="list-style-type: none"> 1. Image plane is rotated around the principal point by θ 2. Lens system is fixed |
| Extra Transformation | <ol style="list-style-type: none"> 1. Rotate the scope cylinder around its axis by θ 2. Inversely rotate the image plane around z-axis of the lens system by θ | <ol style="list-style-type: none"> 1. Rotate the image plane around the principal point by θ |
| Unknown parameters | <ol style="list-style-type: none"> 1. θ 2. Axis of the scope cylinder 3. Axis of the lens system | <ol style="list-style-type: none"> 1. θ 2. Principal point |

Figure 3. A comparison between the system of Yamaguchi et al. and our own system. In the former, the camera head is tracked such that the transformation from the marker to the lens system is not fixed but depends on the rotation angle θ . Using the marker coordinates as a reference, the lens system is rotated around the scope cylinder through θ , but the image plane (that is in the camera head) remains the same. Yamaguchi et al. use two additional transformations to describe the effect of rotation, so their model becomes complicated. Moreover, they must calibrate the axis of both the scope cylinder and the lens system by using another optical marker attached to the scope cylinder. Based on our observation, it is possible to simplify the model if we fix the transformation between the marker and the lens system. We designed a coupler that enables the optical marker to be mounted on the scope cylinder. We then set the marker coordinates as a reference; the lens system remains fixed. The rotation only affects the image plane since the camera head is rotated around the cylinder (reference), and the image plane only rotates around the principal point. Since the principal point is an intrinsic parameter, we need only estimate the rotation angle. As a result, we have a very simple model (see details in the text).

where A is the position of O_2 when $\theta = 0$ and B is the position of O_2 given a rotation θ . The center of the circle can be represented in terms of the transformation from the world coordinates O_w to Marker 1's coordinates O_1 and Marker 2's coordinates O_2 , and at least 3 different positions of Marker 2 (O_2) (with different values of θ) are necessary.

Estimation of the center of circle in 3D

We rotate the camera head around the cylinder to acquire 3 different positions of Marker 2. Let the transformation matrix from the world coordinates O_w to both Marker 1's coordinates O_1 and Marker 2's coordinates O_2 for position i be $({}^{o_1}T_{o_w}^i, {}^{o_2}T_{o_w}^i)$ ($i = 1, 2, 3$). Given any point \vec{P}_r in O_2 , we first

compute the position \vec{P}_i in O_1 corresponding to different rotations as follows:

$$\vec{P}_i = {}^{o_1}T_{o_w}^i \cdot ({}^{o_2}T_{o_w}^i)^T \cdot \vec{P}_r, \quad i = 1, 2, 3. \quad (4)$$

Therefore, O is the center of the circumcircle of the triangle $(\vec{P}_1, \vec{P}_2$ and $\vec{P}_3)$.

If $\vec{R}_1 = \vec{P}_1 - \vec{P}_3$, $\vec{R}_2 = \vec{P}_2 - \vec{P}_3$, the normal of the triangle is $\vec{n} = \vec{R}_1 \times \vec{R}_2$. The perpendicular bisector \vec{L}_1 of \vec{R}_1 and \vec{L}_2 of \vec{R}_2 can be computed as follows:

$$\begin{aligned} \vec{L}_1 &= \vec{P}_3 + \vec{R}_1/2 + \lambda_1 \cdot \vec{n} \times \vec{R}_1 \\ \vec{L}_2 &= \vec{P}_3 + \vec{R}_2/2 + \lambda_2 \cdot \vec{n} \times \vec{R}_2 \end{aligned} \quad (5)$$

where λ_1 and λ_2 are parameters of the line \vec{L}_1 and \vec{L}_2 . The intersection of these two lines is the center

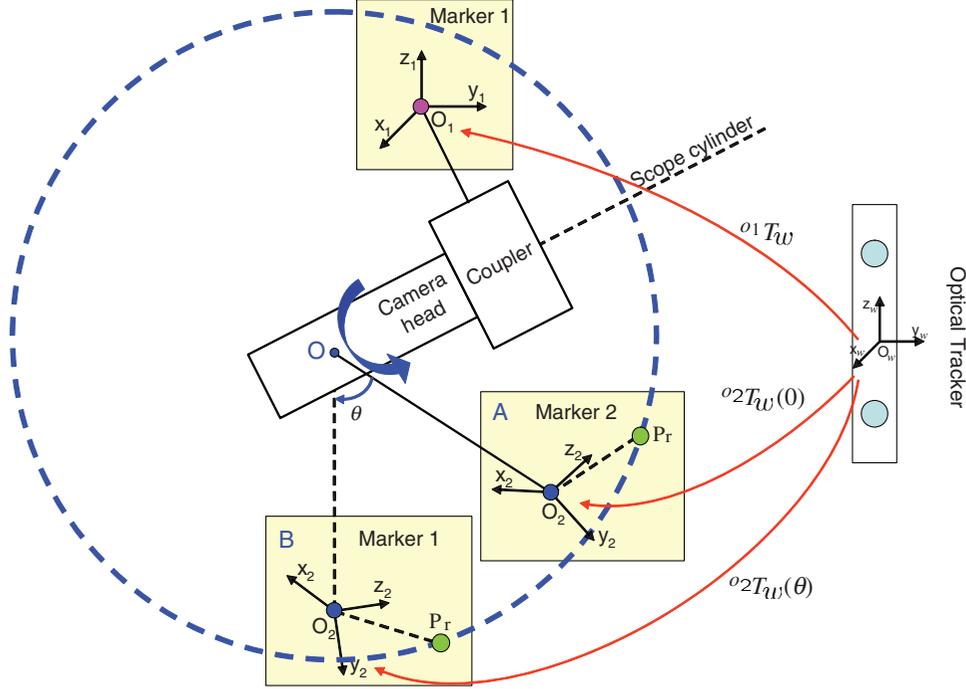


Figure 4. Relationship between the rotation angle θ and two marker coordinates. O_1 is attached to the scope cylinder and O_2 is attached to the camera head. A indicates the position of O_2 when $\theta=0$ and B indicates the position of O_2 given a rotation θ . Given any point P_r in O_2 , its trace following the rotation of the camera head is a circle in Marker 1's coordinates O_1 . It moves from position P_r^A to P_r^B in Marker 1's coordinates O_1 . This circle is also on the plane perpendicular to the axis of the scope cylinder. O is the center of the circle.

Table I. Pseudo code of RANSAC for estimating the center of the circle

Loop k=1:K (K=2000)
 Generate a random point P_r from 3D space
 Generate random number x, y, z between $[1, N]$
 Compute P_x, P_y, P_z using Eq. 4
 Compute O_k using Eq. 6
 Compute $|\overline{O_k P_j}|, j \in [1, N], j \neq x, y, z$
 Compute v_k
 Save O_k, v_k
End loop
 Return $O_q, q = \arg \min(v_k)$

of the circle. From Equation 5 we can derive the center of the circle as follows:

$$\vec{O} = \frac{(\vec{R}_2 - \vec{R}_1) \cdot \vec{R}_1 / 2}{|\vec{R}_1 \times \vec{R}_2|^2} \cdot (\vec{R}_1 \times \vec{R}_2) \times \vec{R}_2 + \vec{R}_2 / 2 + \vec{P}_3 \quad (6)$$

It can easily be proven that O does not depend on the selection of \vec{P}_r . Since at least 3 different positions are necessary, we rotate the camera head around the scope cylinder by N different angles. We then apply a RANSAC algorithm to estimate \vec{O} using random positions, and select the \vec{O} which corresponds to the smallest variance as the center of

Table II. Pseudo code of RANSAC for estimating the rotation angle

Loop k=1:K (K=1000)
 Generate a random point P_r from 3D space
 Compute P_A and P_B using Eq. 4
 Compute θ_k using Eq. 3
End loop
 Return $\theta = \frac{1}{K} \sum_k \theta_k$

the circle. The pseudo code of RANSAC is listed in Table I. It can be also proven that θ does not depend on the selection of P_r either. A similar RANSAC algorithm as shown in Table II is then used to compute θ . Figure 5 shows the estimated rotation angle using the RANSAC algorithm for two different endoscopes. The red curves represent the output angles from different RANSAC iterations, and the black curve is the average angle. As can be seen, the variance of the estimation is very small (less than 0.2 degrees).

Experimental results

We tested our algorithm using two different systems. The first test was conducted in our laboratory, using a Stryker 344-71 arthroscope Vista (70 degree, 4mm) oblique-viewing endoscope,

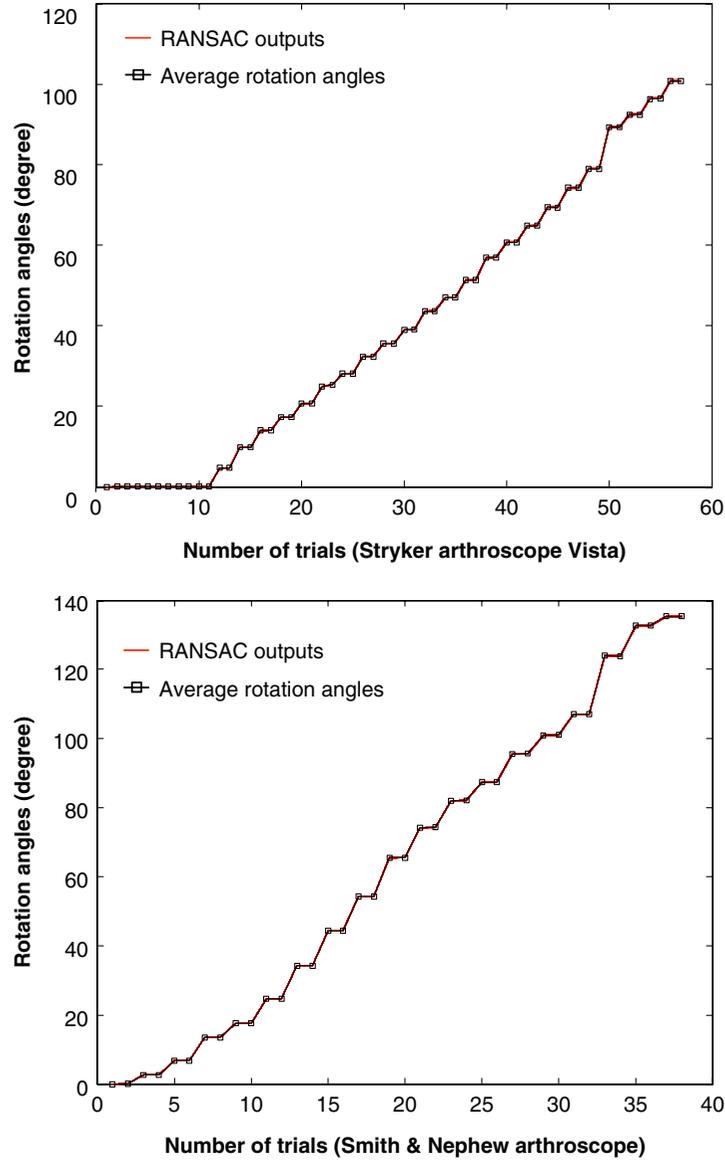


Figure 5. Estimated rotation angles for the two endoscopes. In each trial we rotated the camera head with respect to the scope cylinder and captured an image. We captured a few images for the initial position, then acquired two images for each rotation angle. The red curves are the estimated rotation angles from different RANSAC iterations; the black curve is the average rotation angle.

a DYONICS DyoCamTM 750 video camera, a DYONICS DyoBrite 3000 light source, and a Polaris optical tracker (Northern Digital Inc., Waterloo, Ontario, Canada). Next, we tested the algorithm in conjunction with our standard operating room equipment: a Smith & Nephew SN-OH 272589 video arthroscope (autoclavable; 30 degree, 4 mm), a DYONICS video camera and light source, and an OPTOTRAK optical tracker (Northern Digital Inc.). Figure 6 shows the different endoscopes and optical trackers used.

The endoscope was first fixed in place and the calibration pattern was rotated on the table to enable

the capture of images. A set of images were captured without any rotation between the scope cylinder and camera head. These were used to estimate both the intrinsic matrix A (including focal length and radial distortion coefficients) and the extrinsic matrix cT_m using Zhang’s method [9] (implemented using OpenCV functions). Another set of images were captured when there was rotation between the camera head and the scope cylinder, and the center of the circle could be computed using Equation 6. Next, we fixed the calibration pattern, with two optical markers attached to the scope cylinder and the camera head, and captured a set of

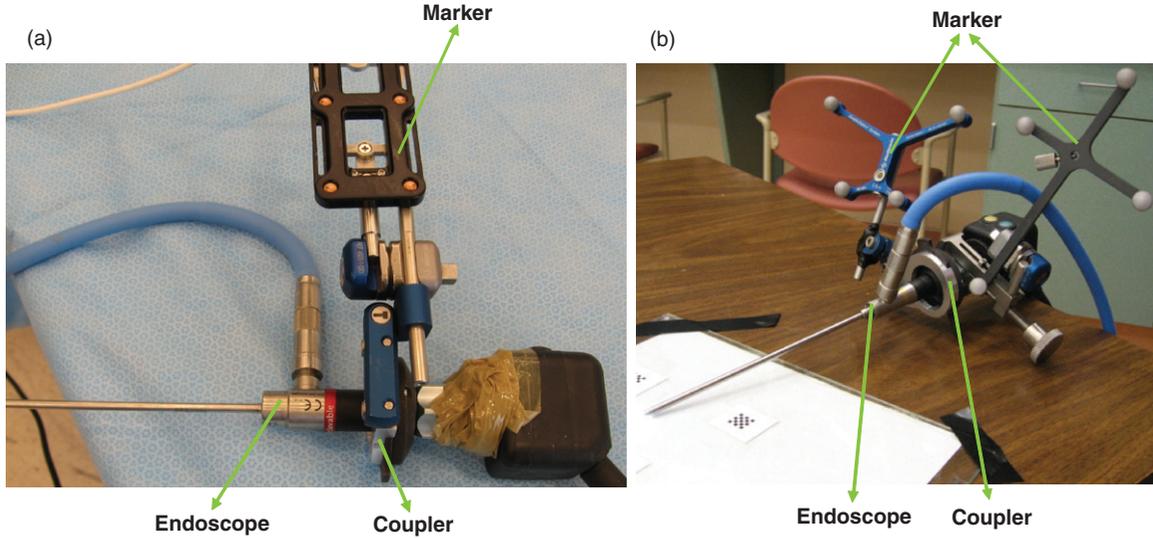


Figure 6. The endoscopes used in the experiments. (a) Smith & Nephew video arthroscope – autoclavable SN-OH 272589 (30 degree, 4 mm). (b) Stryker 344-71 arthroscope Vista (70 degree, 4 mm).

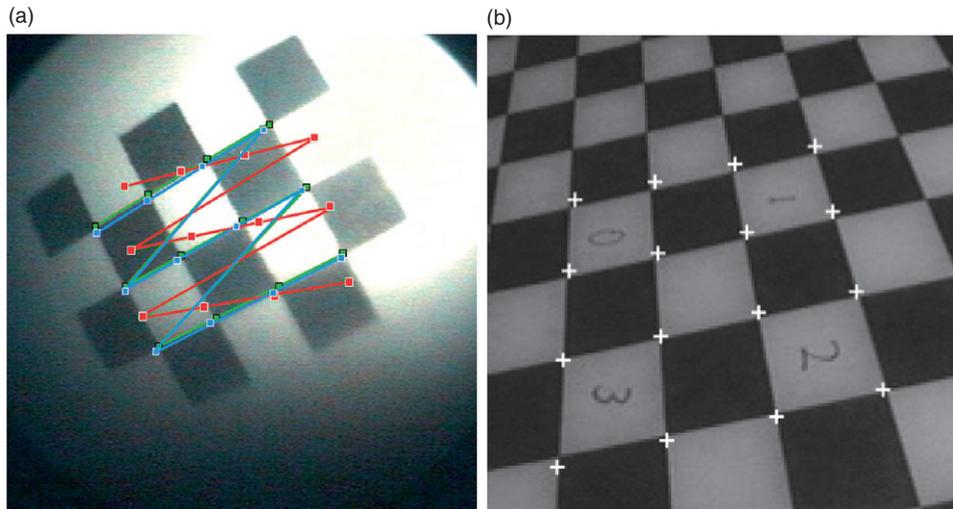


Figure 7. (a) The back projection with and without rotation compensation. The green points are the ground truth – 2D corner pixels on the image of the calibration pattern. The red points are the back projection of the 3D world positions of the corners using the first equation of Equation 2, which has no rotation compensation. The blue points are the back projection using both equations of Equation 2. Since the rotation is included in the camera model, the back-projected pixels are much closer to the ground truth than the red points. (b) An image used in the paper by Yamaguchi et al. [11, 12]. This image has higher resolution, better lighting and less distortion than our own.

images by applying general motions of the endoscope (moving the whole scope body or rotating the camera head with respect to the scope cylinder [better described as rotating the scope cylinder with respect to the camera head]). This set of images was used to estimate the rotation angles. The initial position of the camera head was considered as the reference position A seen in Figure 4. Figure 7 illustrates the back projection of the 3D corners of the calibration pattern with (blue) and without (red)

any rotation compensation. The green points are the ground truth. For each rotation angle of the endoscope, we computed the average back-projection error for this angle as follows:

$$\epsilon(\theta) = \frac{1}{M} \sum_{i=1}^M |p_i - p(P_i, \theta)| \quad (7)$$

where P_i is a 3D point in the world coordinates, p_i is the corresponding 2D image pixel, $p(P_i, \theta)$ is the

back-projected 2D image pixel of P_i , and M is the number of corners on the calibration pattern. We used different grid patterns (3×4 as shown in Figure 7, 4×5 and 5×6 ; the size of each check is $2 \text{ mm} \times 2 \text{ mm}$). In order to obtain enough light on the grid pattern, the endoscope must be placed very close to the target (usually at a distance of 5–15 mm). The smaller grid could not capture the radial distortion, but the bigger grid exceeded the field of view. The 5×6 grid gave the best results.

Finally, we conducted many trials by moving and rotating the endoscope randomly and estimating θ simultaneously. The average back-projection error with respect to the different rotation angles is shown in Figure 8. Figure 8a is the result obtained using the Stryker 344-71 arthroscope Vista (70 degree, 4 mm) and the Polaris optical tracker. Figure 8b is the result obtained using the Smith & Nephew SN-OH 272589 video arthroscope (30 degree, 4 mm) and OPTOTRAK optical tracker. The red curve represents the back-projection error without taking into account the rotation angle, and the blue curve shows the error considering the rotation angle. The results indicate that including the rotation angle in the camera model significantly improves the accuracy of the calibration.

Figure 8 shows that different endoscopes have different accuracy. The reason for this is that endoscopes have different magnifications and optical trackers have differing accuracies (according to the manufacturer, the RMS error is 0.1 mm for OPTOTRAK and 0.3 mm for Polaris). Yamaguchi et al. [11, 12] used an OTV-S5C laparoscope (Olympus Optical Co. Ltd., Tokyo, Japan) and Polaris optical tracker. They achieved a high accuracy of less than 5 mm back-projection error when the rotation angle was within 140 degrees. Our results show that we can achieve the same level of accuracy when the rotation angle is within 75 degrees. Beyond this range, due to the greater magnification, larger radial distortion and poorer lighting (a comparison between images used in our experiment and those of Yamaguchi et al. is shown in Figure 7), the back-projection error is increased to 13 mm when the rotation angle is 100 degrees. Given endoscopes of the same quality, we should be able to achieve the same level of accuracy.

Photometric calibration

To our knowledge, photometric calibration for the endoscope has not been investigated in the literature. In this paper we propose a method for computing the camera response function, light source intensity and light spatial distribution

function simultaneously, inspired by the work of Litvinov and Schechner [13].

Reflectance model under near-field lighting and projective projection

Assuming the bone surface is Lambertian, the scene radiance can be computed according to Lambertian cosine law as follows:

$$R(x, y, z) = G_0 \rho \left(\frac{\mathbf{n} \cdot \mathbf{l}_1}{r_1^2} + \frac{\mathbf{n} \cdot \mathbf{l}_2}{r_2^2} \right) \quad (8)$$

where G_0 is the intensity of two light sources s_1 and s_2 , ρ is the surface albedo, \mathbf{n} is the surface normal, \mathbf{l}_1 and \mathbf{l}_2 are two light rays arriving at the surface, and r_1 and r_2 are the distances from each light source to the surface. (x, y, z) indicates the 3D location of the scene point P (see Figure 9).

On the other hand, the image irradiance E can be estimated from the image intensity v given the camera response function $H(\cdot)$:

$$E(x, y) = \frac{H^{-1}(v(x, y))}{M(x, y)} \quad (9)$$

where $M(x, y)$ is the spatial distribution of the source intensity that is assumed to be equal for both light sources. Assuming there is no loss of energy when light rays propagate from the scene to the camera lens, the radiance remains the same and thus $R = E$. Combining Equations 8 and 9, we have

$$\begin{aligned} H^{-1}[v(x, y)] &= \rho \cdot G_0 \cdot \tilde{M}(x, y) \\ \tilde{M}(x, y) &= M(x, y) \cdot \left(\frac{\mathbf{n} \cdot \mathbf{l}_1}{r_1^2} + \frac{\mathbf{n} \cdot \mathbf{l}_2}{r_2^2} \right) \end{aligned} \quad (10)$$

For calibration, we use a Macbeth color chart with known albedo for each patch. We capture a set of images by varying light-source intensity for each patch. We apply log to both sides of Equation 10 to obtain a linear system of equations:

$$h[v_i^j(x, y)] = p_i + g_j + \tilde{m}(x, y) \quad (11)$$

where i indicates the different albedos and j indexes the light intensity. $h[v_i^j(x, y)] = \log\{H^{-1}[v_i^j(x, y)]\}$, $p_i = \log(\rho_i)$, $g_j = \log(G_0)$ and $\tilde{m}(x, y) = \log[\tilde{M}(x, y)]$. The unknowns ($h(\cdot)$, g_j , $\tilde{m}(x, y)$) can be obtained by solving this linear system of equations. The term $M(x, y)$ is then estimated from $\tilde{M}(x, y)$ by measuring the distance to the chart from the scope tip. An image of a square patch with uniform albedo is captured and distortion is then removed. We measure the distance between the scope tip and four corners, and then compute the term $\frac{(n \cdot \mathbf{l}_1)}{r_1^2} + \frac{(n \cdot \mathbf{l}_2)}{r_2^2}$ for each image pixel. Finally, $M(x, y)$ is estimated by removing the above term from $\tilde{m}(x, y)$.

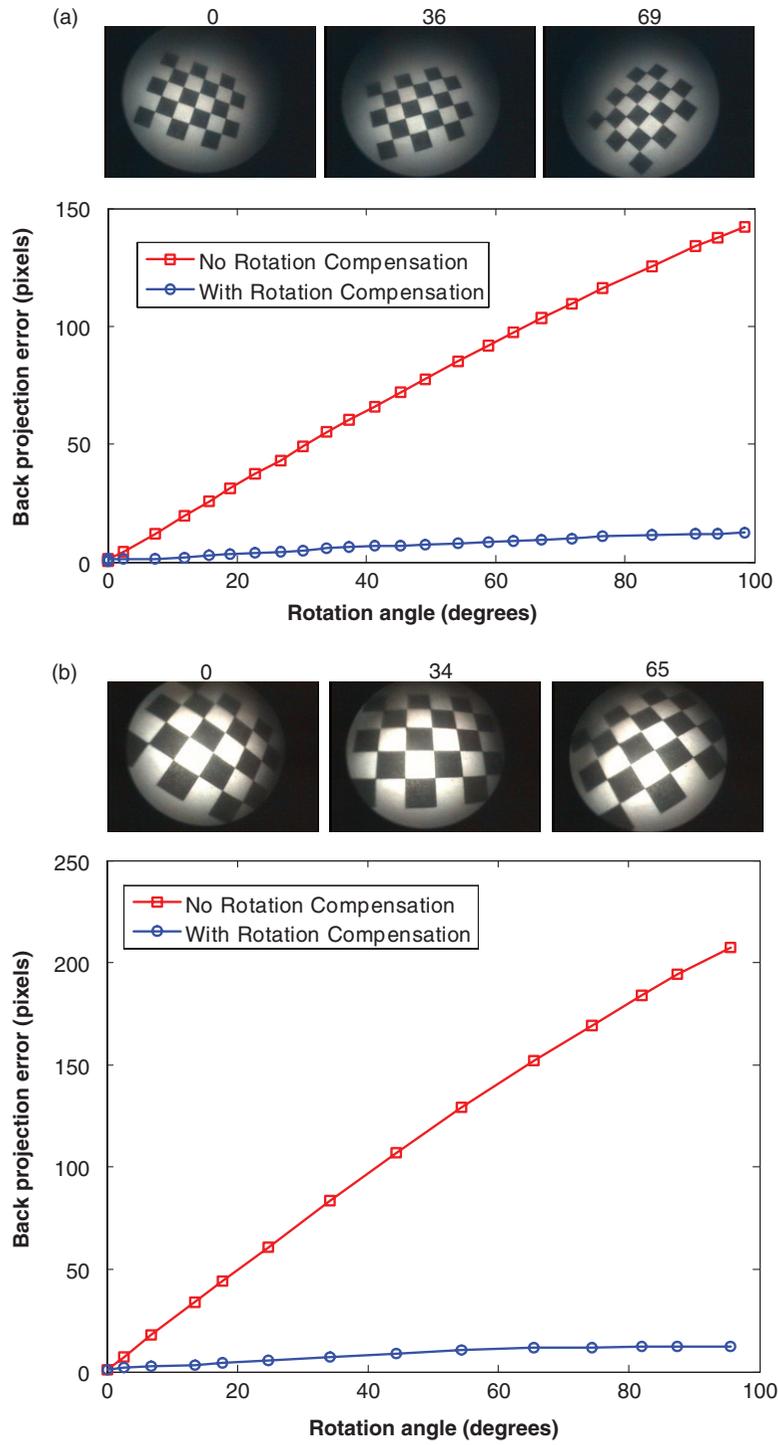


Figure 8. Back-projection errors with respect to the rotation angles for the two systems. (a) Stryker 344-71 arthroscope Vista and Polaris optical tracker in our laboratory. (b) Smith & Nephew video arthroscope and OPTOTRAK optical tracker in the operating room. The three images above the graphs correspond to different rotation angles (specified above each image). The red curves represent the errors without rotation compensation; the blue curves are errors with rotation compensation.

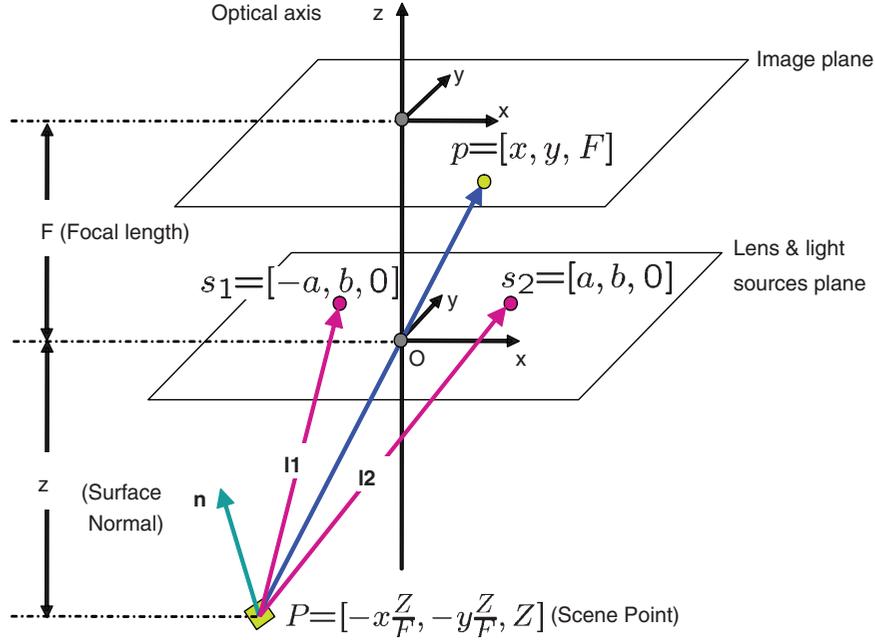


Figure 9. The perspective projection model for an endoscope imaging system with two near point light sources: O is the camera projection center; s_1 and s_2 indicate two light sources. We assume the plane consisting of O , s_1 and s_2 is parallel to the image plane. The coordinate system is centered at O and Z is parallel to the optical axis and pointing toward the image plane. X and Y are parallel to the image plane, F is the focal length, and a and b are two parameters related to the position of the light sources. Given a scene point P , the corresponding image pixel is p . Assuming a Lambertian surface, the surface illumination thus depends on the surface albedo, light source intensity and fall-off, and the angle between the normal and light rays.

Solution to $h(\cdot)$

Given the same light intensity g_j and pixel value $v(x, y)$ but two different albedos ρ_{i_1} and ρ_{i_2} , we have

$$\begin{cases} h[v_{i_1}^j(x, y)] - g_j - \tilde{m}(x, y) - p_{i_1} = 0 \\ h[v_{i_2}^j(x, y)] - g_j - \tilde{m}(x, y) - p_{i_2} = 0. \end{cases} \quad (12)$$

Subtracting the first line from the second line of Equation 12, we obtain

$$h[v_{i_2}^j(x, y)] - h[v_{i_1}^j(x, y)] = p_{i_2} - p_{i_1} \quad (13)$$

We can use different pixels in the same image (albedo) or different images (albedos) to make as many equations like Equation 13 as required, as long as we fix the light intensity for each pair of albedos. Since $v_{i_1}^j(x, y)$ changes from 0 to 255 (image intensity), we only need 256 such equations and stack them as follows:

$$\begin{bmatrix} \dots & 1^* & -1^* & \dots \\ \dots & -1^\# & 1^\# & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \cdot \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(255) \end{bmatrix} = \begin{bmatrix} p_{i_2} - p_{i_1} \\ p_{i_4} - p_{i_3} \\ \vdots \end{bmatrix} \quad (14)$$

where 1^* and -1^* correspond to the columns $v_{i_2}^j(x, y) + 1$ and $v_{i_1}^j(x, y) + 1$, respectively, and $1^\#$ and $-1^\#$ correspond to the columns $v_{i_4}^j(x, y) + 1$ and

$v_{i_3}^j(x, y) + 1$, respectively. Therefore $h(v)$ is solved from Equation 14 and $H(v) = \exp(h(v))$.

Solution to g_j

Given the same albedo p_i and pixel value $v(x, y)$ but two different light intensities g_{j_1} and g_{j_2} , we have

$$\begin{cases} h[v_i^{j_1}(x, y)] - g_{j_1} - \tilde{m}(x, y) - p_i = 0 \\ h[v_i^{j_2}(x, y)] - g_{j_2} - \tilde{m}(x, y) - p_i = 0. \end{cases} \quad (15)$$

We then subtract the first line from the second line of Equation 15:

$$h[v_i^{j_2}(x, y)] - h[v_i^{j_1}(x, y)] = g_{j_2} - g_{j_1} \quad (16)$$

We use the minimum light intensity g_1 as a reference; for other light intensities g_j , $j = 2, \dots, N_{lights}$ we have

$$g_j = g_1 + h[v_i^j(x, y)] - h[v_i^1(x, y)] \quad (17)$$

Given the estimated $h(v(x, y))$ and by changing the albedos and pixels, we compute the average value for each g_j as follows and $\bar{G}_j = \exp(g_1) \cdot \exp(\bar{g}_j)$.

$$\bar{g}_j = g_1 + \frac{1}{N_{albedo}} \cdot \frac{1}{N_{pixels}} \sum_i \sum_{x, y} \{h[v_i^j(x, y)] - h[v_i^1(x, y)]\} \quad (18)$$

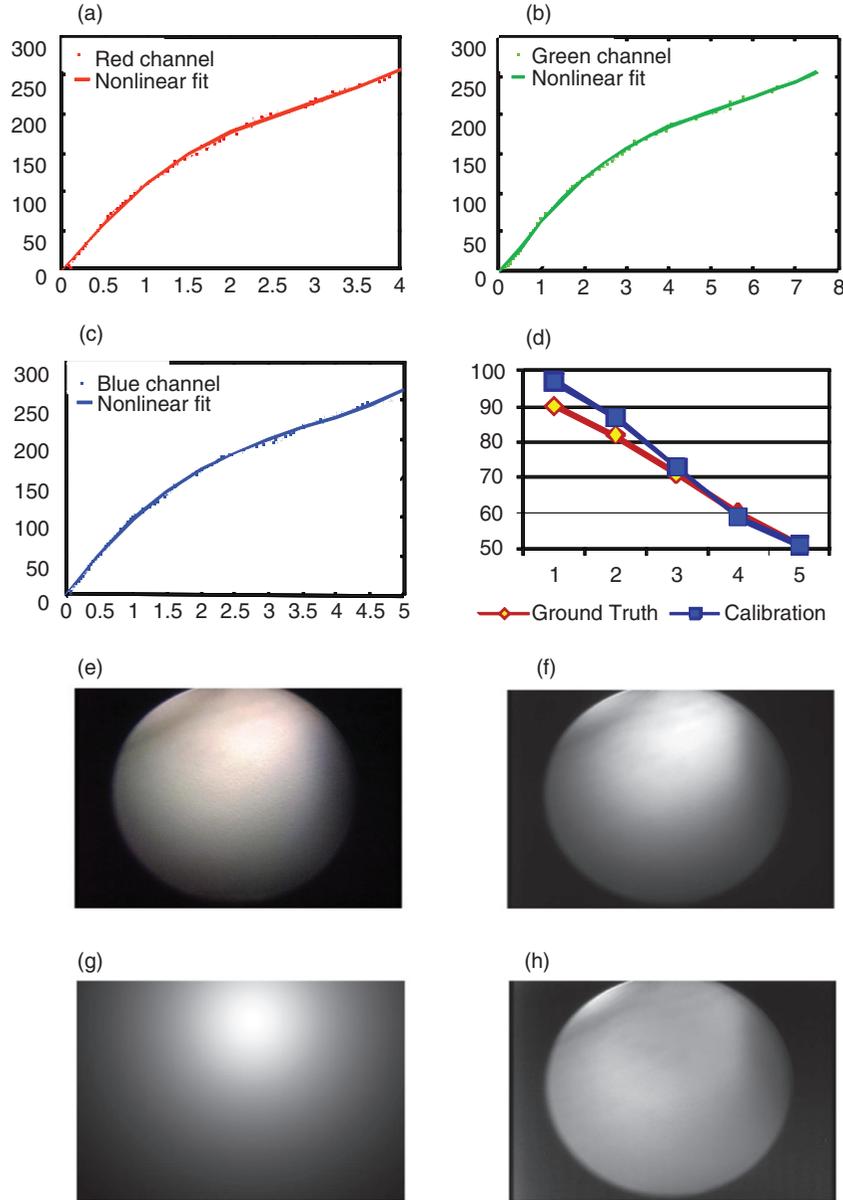


Figure 10. Results of photometric calibration. (a) Camera response function in the red channel. The red dots represent the data points and the magenta line represents the nonlinear fit. (b) Camera response function in the green channel. The green dots represent the data points and the magenta line represents the nonlinear fit. (c) Camera response function in the blue channel. The blue dots represent the data points and the magenta line represents the nonlinear fit. (d) Calibrated light intensity at different levels (blue) and the ground truth (green). We use level 6 as a reference and plot levels 1-5 with a small level corresponding to high light intensity. A bit variation in range for the high intensities may be caused by saturation. (e) Original image on color chart. (f) \tilde{m} . (g) The cosine term $\frac{(n-1)}{r_1^2} + \frac{(n-1)}{r_2^2}$. (h) The spatial distribution function $m(x, y)$.

Solution to $\tilde{m}(x, y)$

Again, Given the same albedo p_i and light intensity g_j but two different pixels, (x_p, y_p) and x_q, y_q , we have

$$\begin{cases} h[v_i^j(x_p, y_p)] - g_j - \tilde{m}(x_p, y_p) - p_i = 0 \\ h[v_i^j(x_q, y_q)] - g_j - \tilde{m}(x_q, y_q) - p_i = 0. \end{cases} \quad (19)$$

We subtract the first line from the second line of Equation 19:

$$h[v_i^j(x_p, y_p)] - h[v_i^j(x_q, y_q)] = \tilde{m}(x_p, y_p) - \tilde{m}(x_q, y_q) \quad (20)$$

For each g_j and ρ_i , instead of using $C_{N_{pixels}}^2$ different pairs of pixels, we choose only $N_{pixels} = 720 \times 480$

pairs and stack the equations thus:

$$\underbrace{\begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & & \ddots & \\ & & & & 1 \\ -1 & & & & & 1 \end{bmatrix}}_A \cdot \begin{bmatrix} \tilde{m}(x_1, y_1) \\ \tilde{m}(x_2, y_2) \\ \vdots \\ \tilde{m}(x_{N_{pixels}}, y_{N_{pixels}}) \end{bmatrix} = \begin{bmatrix} h[v_i^j(x_1, y_1)] - h[v_i^j(x_2, y_2)] \\ h[v_i^j(x_2, y_2)] - h[v_i^j(x_3, y_3)] \\ \vdots \\ h[v_i^j(x_{N_{pixels}}, y_{N_{pixels}})] - h[v_i^j(x_1, y_1)] \end{bmatrix} \quad (21)$$

It is not practical to solve Equation 21 using singular value decomposition (SVD) directly since matrix A requires a huge memory. However, we noticed that the matrix A in Equation 21 is a special N_{pixels} by N_{pixels} matrix such that we could get the inverse directly by using Gauss-Jordan Elimination:

$$A^{-1} = \begin{bmatrix} 0 & & 0 & -1 \\ -1 & 0 & & 0 & -1 \\ -1 & -1 & 0 & & 0 & -1 \\ \vdots & \vdots & \vdots & \dots & 0 & -1 \\ -1 & -1 & -1 & \dots & -1 & -1 \end{bmatrix} \quad (22)$$

Thus we successfully compute each element of $\tilde{m}(x, y)$ independently and again, $M(\tilde{x}, y) = \exp(\tilde{m})$.

Experimental results

A series of color chart images are used for photometric calibration. We use 6 different levels of image intensity. Figures 10a–c show the camera response function with red/green/blue channels. Figure 10d shows the recovered light intensity in different levels and compared to the ground truth. Smaller values on the x-axis correspond to higher intensity. We see a bit variance when light intensity is high that may be caused by the saturation. Figure 10e shows the original image and 10f shows \tilde{m} . Figure 10g shows the cosine term $\frac{(n) \cdot l_1}{r_1^2} + \frac{(n) \cdot l_2}{r_2^2}$, and 10h shows the spatial distribution function $m(x, y)$.

Discussion and conclusions

In this paper we have described the development of a full calibration process for both geometric and photometric properties of endoscopes. Our geometric calibration method simplifies previous work by using a coupler attached to the scope cylinder.

It is easy to implement and practical to apply with standard operating room equipment such as the surgical navigation system. The only drawback of this method is the necessity for two markers to be visible to the optical trackers at all times.

To our knowledge, photometric calibration has not yet been applied to endoscopes by other groups. Most previous work in this field did not rely on the physical model of the endoscope light sources, or restricted the changing of light sources during the operation. Recently, shape-from-shading has been applied to endoscopic images based on a simplified light-source model without calibrating the model parameters. However, to reconstruct an accurate shape from endoscopic images, knowledge of the light sources is necessary and important.

Both geometrical and photometrical parameters are very useful for 3D visualization and reconstruction of anatomical structures such as bones. We have already applied the calibrated endoscopes to artificial reconstruction of the spine from individual endoscopic images based on a calibrated photometric model [14]. With the calibrated geometric parameters of the endoscope, we were able to transfer all individual reconstructed shapes to the world coordinates and thus find the global occluding boundaries, which helped to later re-grow the shape-from-shading. The results demonstrate that calibrated endoscopes can yield very good reconstruction results, which is promising for real surgical applications.

References

1. Clarkson MJ, Rueckert D, King AP, Edwards PJ, Hill DLG, Hawkes DJ. Registration of video images to tomographic images by optimising mutual information using texture mapping. In: Taylor C, Colchester A, editors. Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '99), Cambridge, England, September 1999. Lecture Notes in Computer Science 1679. Berlin: Springer; 1999. pp 579–588.
2. Dey D, Gobbi DG, Slomka PJ, Surry KJ, Peters TM. Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: Creating stereoscopic panoramas. IEEE Trans Med Imaging 2002;21(1):23–30.
3. Deguchi K, Okatani T. Shape reconstruction from an endoscope image by shape-from-shading technique for a point light source at the projection center. Proceedings of IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 1996), San Francisco, CA, June 1996. pp 290–298.
4. Forster CHQ, Tozzi CL. Towards 3D reconstruction of endoscope images using shape from shading. Proceedings of XIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2000), Gramados, Brazil, October 2000. pp 90–96.

5. Mourgues F, Coste-Manière E. Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery. In: Dohi T, Kikinis R, editors. Proceedings of 5th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2002), Tokyo, Japan, September 2002. Part I. Lecture Notes in Computer Science 2488. Berlin: Springer; 2002. pp. 25–34.
6. Stoyanov D, Darzi A, Yang GZ. A practical approach towards accurate dense 3D depth recovery for robotic laparoscopic surgery. *Comput Aided Surg* 2005;10(4): 199–208.
7. Wengert C, Reeff M, Cattin PC, Székely G. Fully automatic endoscope calibration for intraoperative use. In: Handels H, Ehrhardt J, Horsch A, Meinzer H-P, Tolxdorff T, editors. Proceedings of Bildverarbeitung für die Medizin 2006, Algorithmen, Systeme, Anwendungen, March 2006, Hamburg, Germany. Berlin: Springer, 2006. pp 419–423.
8. Tsai RY. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J Robotics Automation* 1987;RA-3:323–344.
9. Zhang Z. A flexible new technique for camera calibration. Microsoft Research MSR-TR-98-71, 1998.
10. Shahidi R, Bax MR, Maurer CR, Johnson JA, Wilkinson EP, Wang B, West JB, Citardi MJ, Manwaring KH, Khadem R. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Trans Medical Imaging* 2002;21(12):1524–1535.
11. Yamaguchi T, Nakamoto M, Sato Y, Nakajima Y, Konishi K, Hashizume M, Nishii T, Sugano N, Yoshikawa N, Yonenobu K, Tamura S. Camera model and calibration procedure for oblique-viewing endoscope. In: Ellis RE, Peters TM, editors. Proceedings of 6th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2003), Montreal, Canada, November 2003. Part II. Lecture Notes in Computer Science 2879. Berlin: Springer; 2003. pp. 373–381.
12. 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. *Comput Aided Surg* 2004;9(5):203–214.
13. Litvinov A, Schechner YY. Addressing radiometric non-idealities: A unified framework. In: Proceedings of IEEE Computer Society International Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, June 2005. Vol. 2. pp. 52–59.
14. Wu C, Narasimhan S, Jaramaz B. A multi-image Shape-from-Shading framework for near-lighting perspective endoscopes. *Int J Comput Vision* 2009;86: 211–228.