ArthroNav:Computer Assisted Navigation System for Orthopedic Surgery using Endoscopic Images

João. P. Barreto, R. Melo, M. Antunes, M. Lourenço and F. Vasconcelos
Institute of Systems and Robotics
Faculty of Science and Technology, University of Coimbra
Coimbra, Portugal
jpbar@{deec.uc.pt, isr.uc.pt}

***Abstract*—In the future, medicine will increasingly rely in Computer Aided Surgery and Diagnosis, hence techniques to extract information from endoscopic images of human-body cavities will be crucial. The ArthroNav project main goal is to process the endoscopic video in order to improve practioner's perception and navigation skill inside the knee joint.**

Some navigation systems for CAS (Computer Aided Surgery) based on image processing techniques have been already proposed in the literature. However, most of them employ methods designed for conventional cameras and do not account with the special optical arrangement of the endoscopic lens. We intend to adapt/design new computer vision techniques invariant to the numerous challenges that endoscopic video places.

The usefulness of developing algorithms for endoscopic images is not limited to Computer Assisted Orthopedic Surgery. Many other medical fields can benefit from the research herein proposed. The processing of endoscopic video can be equally used in other minimally invasive procedures like nose surgery, neurosurgery, and cardiac surgery.

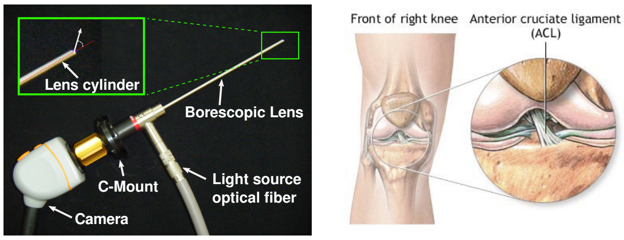
***Index Terms*—CAS navigation; Endoscopy; Computer Vision; Orthopedic Surgery**

I. Context

The 3D reconstruction of scenes using monocular video sequences is a classical problem in computer vision. The literature covers, not only 3D reconstruction, but also related subjects as camera calibration and 3D registration. However, most of the described methodologies were developed for perspective images of everyday scenes. The goal of this project is to extend this framework to endoscopic video sequences in the context of computer aided orthopedic surgery (CAOS) [1]. Our target application is the visualization and navigation during the reconstruction of the Anterior Cruciate Ligament (ACL) using arthroscopy (endoscopic observation of articulations). Since we aim to develop methods to extract 3D information from endoscopic images, the project outcome will be useful for many other minimally invasive medical procedures.

The ACL runs from the femur to the tibia, and is the major stabilizing ligament of the knee [2] (see Fig. 1(b)). The tearing of the ACL is a common injury, and its treatment usually requires minimally invasive surgery to replace the torn ligament by a tendon graft. During the procedure, the injured ligament is removed, and drill holes are open on the tibia plate and femur notch at the attachment sites of the original ligament. The graft is pulled through the open tunnel into the knee joint, and screws are employed to held it in place. This procedure must be performed by well trained surgeons. The practitioner must be able to navigate in the knee joint, recognize anatomical landmarks, and open the graft tunnel using only the visual feedback provided by the arthroscope. The accurate positioning of the drill holes from the outside is the most critical and difficult surgical step. Small errors in the graft placement can lead to abnormal tensions during motion, cause pain and new injuries, and ultimately force to a corrective surgery [3].

João. P. Barreto is the PI of the project.



(a) Rigid endoscopic camera (b) Anterior Cruciate Ligament

Fig. 1. The rigid medical endoscope combines a boroscope with a camera (Fig. 1(a)). The system enables the visualization of human-body cavities with difficult or limited access, being broadly used for both surgery and diagnosis in different medical fields. Fig. 1(b) is an anatomical image of the interior of a knee joint.

Studies show that the rate of clinical success is 85%, and that 50% of the failures are due to graft misplacement. Certain authors claim that significant errors in positioning the tunnel occur in 10 to 40% of the ACL reconstructions [4], [5]. This is a scenario where CAOS can have strong social and economical impacts. Computer systems for enhancement of surgeons perception and precise navigation in the knee joint can improve the clinical success rate and decrease the practitioner training requirements. Since the tearing of the ACL is a common injury (over 75000 cases per year just in the US) [6], the development of such systems can bring great benefits in improving the life quality of young patients and diminishing health care costs.

A surgical navigator is a visualization system providing real-time positional information about instruments and tools with respect to a target organ (bones). The position and orientation of external rigid objects, like the arthroscope and surgical tools, can be easily computed using a commercial optical tracker. A stereo head with infrared cameras (the optical tracker) tracks a set of LED markers that are rigidly attached to each instrument. Simple triangulation is used to determine the 3D pose of the different tools in a common world reference frame. The main difficulty in developing a CAOS system for ACL reconstruction is the estimation of the position and orientation of the femur/tibia. The patient's leg moves during the surgery, and optical tracking can not be used because the bones are not visible from the outside. We propose to solve the problem by reconstructing the knee joint from the arthroscopic video stream. The partial reconstructions can be registered with pre-operative volumetric images (typically CT scan) in order to estimate the 3D pose of the femur/tibia with respect to the camera. Since the rigid motion of the arthroscope is tracked by an external stereo head, we can compute the position and orientation of the bones in the world coordinate system where the surgical tools are referenced.

II. State-of-the-Art Overview

The recent developments in computer vision [7], and the fact that endoscopes will always be present in minimally invasive procedures, make the use of endoscopic video for CAS an increasingly attractive proposition. In [8], pre-operative CT models are registered to the 2D endoscope image during bronchoscopy. They optically detect the pulmonary branches to navigate inside the bronchi limiting its applicability to bronchial surgery. Another navigation aid using photogrammetry during endoscopic surgery has been studied in [9]. They use the structural information to prevent the endoscope image from flipping upside-down while rotating the camera. Wengert et al. use a tracked monocular endoscope to sparsely reconstruct the surgical scene in 3D [10]. They detect and track image interest points in the images, discard outliers using the epipolar constraint [7], and reconstruct the points by triangulation. The algorithm runs near to real-time thanks to an efficient data management scheme. They claim an accuracy of 0.1mm for the reconstruction of a vertebra, and 0.5mm for registration using a standard ICP approach [12]. In [11], Burschka et al. go one step further and reconstruct a scaled 3D model of the surgical scene from unknown camera motion. The method is applied for the localization during the sinus surgery and the claimed average error is around 0.65mm.

III. ArthroNav Goals

To the best of our knowledge, navigation using endoscopic video has never been attempted for the arthroscopic reconstruction of the ACL. The development of such system implies going beyond the state-of-the-art in the following topics:

1) Geometry of Image Formation - The arthroscope is a non-conventional vision sensor and can not be described by the perspective model. We intend to derive a suitable projection model, study the image geometry, and develop calibration methods.
2) 3D Reconstruction - The reconstruction of the knee joint is difficult due to the nature of the scenes, the image distortion and the lightning conditions. We need to investigate features types that can be reliably tracked, deal with textureless regions and high image distortion.
3) 3D Registration - The recovered local structure must be registered with a pre-operative CT scan. The 3D registration algorithm must be accurate despite of the leg's movements, sparse local 3D information, and real-time requirements.

IV. Team and Institutions

The development of a successful CAOS system requires a multidisciplinary team, where the exchange of knowledge is fundamental. Our team joins the Coimbra University Hospital with engineering research units (including Institute of Systems and Robotics and Faculty of Science and Technology, University of Coimbra). We want to promote the research in medical engineering and form the highly skilled human-resources required in the biotechnology industry.

V. Project Outcome

In this section we present the main features of our system. We will divide this section in 3 subsections, each one related with our research topics introduced in section III

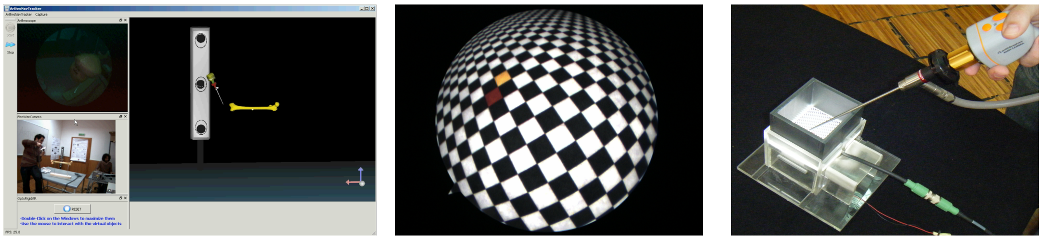
A. Infrastructure to be used in the OR

1) Optical tracking system for determination of the position and orientation of medical tools: The outcome of this task was a system capable of the synchronous acquisition of optical tracking data, and image data from the arthroscope and a FireWire camera [18], [19]. In addition, we implemented modules for visualization and storing that can be easily inserted in the software application (Fig. 2(a)). Extensive tests were performed to confirm the synchronism and real-time processing [18], [19].

2) Automatic Camera Calibration in Medical Endoscopy: We developed a method that calibrates a camera with lens distortion using a single image of a planar chessboard pattern. The radial distortion (RD) is modeled using the first order division model and the method provides a closed form estimation of the intrinsic parameters and distortion coefficient. The fact that the distortion follows a known model provides additional geometric cues for achieving calibration from a single image [14]. The experiments show that our single image calibration has good repeatability and provides results comparable to the ones obtained with the state-of-the-art Bouguet toolbox [14], [15].

The method is available as EasyCamCalib Toolbox[1]. The software requires minimum user intervention (no need of clicking corners), which makes it specially suited for usage in the OR.

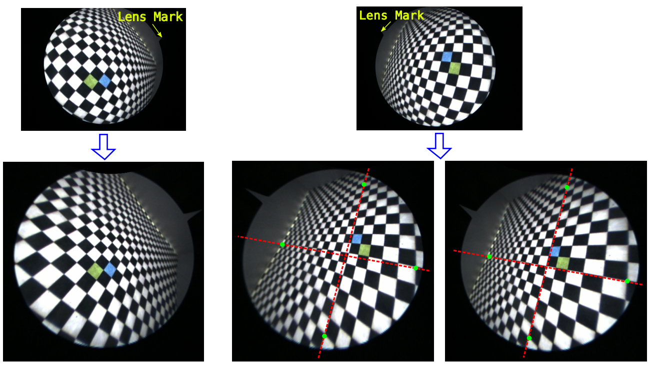
[1]The software was made publicly available for usage among the scientific community at http://arthronav.isr.uc.pt/autocalibration/index.html



(a) ArthroSync (b) Calibration image (c) Calibration setup

Fig. 2. ArthronNav main application (Fig. 2(a)) and calibration of the endoscopic camera from a single image of a planar checker board pattern (Fig. 2(b), 2(c)). The calibration image of Fig. 2(b) is acquired in an unconstrained manner (no assumptions about the relative pose between camera and plane) using the setup shown in Fig.2(c). The setup consist in a simple acrylic box with the purpose of controlling the lightning conditions. The planar grid is back-illuminated in order to enable robust and accurate detection of the image corners with no user intervention.

3) Improving Surgeon Perception of Depth: The visualization module uses camera calibration information to correct the image distortion and improve the surgeons depth perception of the knee joint. However, in rigid endoscopes with exchangeable optics, the flexibility to perform rotations between the camera sensor and the lens system preclude the use of a single static camera calibration. AS so, we propose a new intrinsic camera model that takes into account the rotation of the lens cylinder in the image formation [20]. The final system presents a completely distortion free image, independent of the lens cylinder rotation (Fig. 3). The application is designed to work with any type of cameras, and is able to handle a wide range of image resolutions. As most of the workload is distributed across the graphics processing unit (GPU), we are able to correct high definition images in real time [18], [20], precluding its usage in the ACL reconstruction.

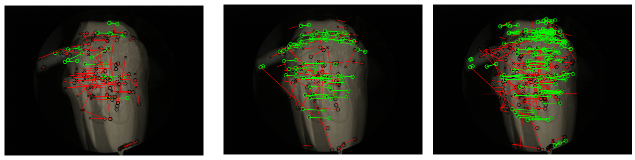


(a) Calibration image (b) Rotation of the lens cylinder

Fig. 3. Radial distortion correction under different environments and lens cylinder rotations. The distorted image (top row) is corrected independently of the lens cylinder rotation (bottom row). (a) The endoscope is calibrated for the lens position indicated in the top image. (b) Comparison between radial distortion correction without (left) and with(right) lens rotation compensation after a rotation of the lens.

B. 3D Reconstruction from Arthroscopic images

Structure from Motion (SfM) is a classic computer vision problem that has already been intensively studied and shown successful practical results. Almost every SfM method heavily relies on a very high number of image features to provide good results, on the other hand surgical video sequences are poorly informative and it is difficult to retrieve many reliable image features. Moreover, state-of-the-art keypoint detectors and descriptors were designed for images captured by pinhole cameras. However, cameras with unconventional optical arrangements such as medical endoscopes can not be described by this standard camera model. Despite of this fact, the state-of-the-art Scale Invariant Feature Transform has been applied in the past to images with significant distortion. While ones simply ignore the pernicious effects of RD and directly apply the original SIFT algorithm over distorted images [11], others perform a preliminary correction of distortion through image rectification and then apply SIFT (rectSIFT) [13].



(a) SIFT (b) rectSIFT (c) sRD-SIFT

	# Matches	# of Inliers	% of Inliers	Reproj. Error
SIFT	110	13	11%	—
rectSIFT	130	52	40%	0.987
sRDSIFT	306	134	42%	0.457

Fig. 4. SfM in endoscopic stereo images with low texture. The two images are overlaid and the point correspondences for each method are marked (Fig. 4(a) to 4(c)). The table shows the number of input matches for the RANSAC, the inlier selection (green matches), and the final re-projection error. The relative motion between views is refined by minimizing the re-projection error using iterative bundle adjustment [7]. For the case of SIFT the re-projection error is not shown because the RANSAC was unable to provide a plausible initialization for the camera motion.

We developed a SIFT-based RD covariant method [16], [17] (coined sRD-SIFT[2]) for feature detection and matching that provides more accurate correspondences across views (some results can be seen in Fig. 4). The modifications consist in using adaptive filtering for the keypoint detection stage and local modeling of the image gradients to account for the radial distortion. Extensive experiments, using both synthetic and real imagery, show the superiority of our method in establishing point correspondence across images with strong

[2]The software was made publicly available for usage among the scientific community at http://arthronav.isr.uc.pt/~mlourenco/srdsift/

distortion [16], [17].

C. Registration of a pre-operative 3D model

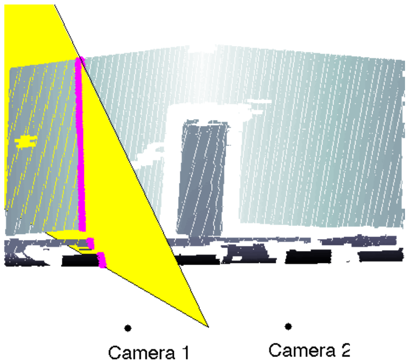


Fig. 5. A virtual passes between the cameras and intersects the scene in a discontinuous curve.

As for the reconstruction task, the main difficulty stems from the lack of available intrinsic features for registration, the unavoidable movement of the leg during surgery, and the required accuracy and real-time performance. At this points we are conducting research towards developing a new technique for stereo reconstruction. This technique consists in back-projecting the stereo views onto an arbitrary virtual plane that passes between the cameras (Fig. 5) . The back-projections give rise to two image signals that are respectively symmetric and anti-symmetric at the points of intersection between the virtual plane and scene surfaces, or the "profile cuts" (see Fig. 6) [21].

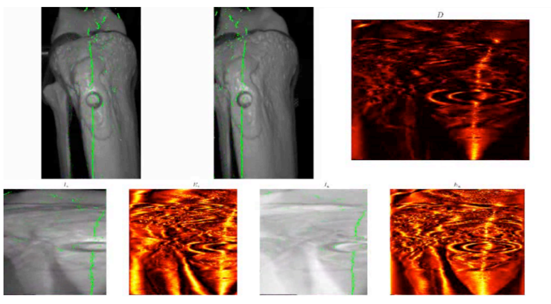


Fig. 6. Obtaining the "profile cut" cut for a virtual plane position using the virtual laser reconstruction technique. This profile cuts will be used to register pre-operative images. On the first row we can see the stereo pair with the symmetry detected. On the second row we presented the symmetric and antisymmetric image signals (1^{st} and 3^{rd} images) with the corresponding energy signals (2^{nd} and 4^{th} images).

VI. Conclusion

In this paper the ArthroNav-CAS system is presented. The main features of our system are the following: (i) Optical tracking of the medical tools synchronized with endoscopic video; (ii) Camera calibration is performed in a fully automatic manner. The surgeon only need to acquire an image for a planar chess grid, which makes the simple usage specially suited for the OR; (iii) Real time distortion correction for improved perception of depth. The RD correction runs real time and the camera model takes into account the rotation of arthroscopic lenses; (iv) 3D sparse reconstruction of the endoscopic scene using RD covariant features.

As future directions we plan to perform 3D registration using the symmetric profile cuts obtained using the new cue to stereo and, as final step, we plan to test the full system working in the OR not only to have surgeon feedback but also to access the commercialization potential of the ArthroNav system.

Acknowledgment

The authors acknowledge the Portuguese Science Foundation, that generously funded this work through grants PTDC/EEA-ACR/68887/2006

References

[1] N. Sugano, "Computer Assisted Orthopedic Surgery", Journal of Orthopedic Science, 2003.

[2] Anatomical Chart Compan, *Knee Injuries Anatomical Chart*.http://www.anatomical.com/product.asp_Q_pn_E_9781587797569.

[3] J. Kartus, L. Magnusson, S. Stener, S. Brandsson, B. Eriksson, *Complications following arthroscopic anterior cruciate ligament reconstruction: A 25-year follow-up of 604 patients with special emphasis on anterior knee pain*, Springer Berlin / Heidelberg 1999.

[4] Hospital for Special Surgery, http://www.hss.edu/conditions_knee-arthroscopy-acl-meniscal-other-problems.asp.

[5] Medical News Impact, http://www.medicalnewstoday.com/articles/165951.php

[6] E-medicine, http://emedicine.medscape.com/article/89442-overview

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[8] M. Tjoa, M. Krishnan, S Zheng. "A novel endoscopic image analysis approach using deformable region model to aid in clinical diagnosis". Proc of the Int Conf of the IEEE Engineering in Medicine and Biology Society, 2003.

[9] D. Koppel, Y. Wang and H. Lee, "Image-based rendering and modeling in video-endoscopy", IEEE Int. Symp. on Biomedical Imaging, 2004.

[10] C. Wengert, P. Cattin, J. Duff, C. Baur and G. Szekely, "Markerless Endoscopic Registration and Referencing", Proc. of Medical Imaging Computing and Computer Assisted Intervention 2006.

[11] D. Burschka, M. Li, R. Taylor, G. Hager, "Scale-Invariant Registration of Monocular Endoscopic Images to CT-scans for Sinus Surgery" , Proc. of Medical Imaging Computing and Computer Assisted Intervention 2004.

[12] J. Maintz, M. Viergever, "A survey on medical image registration", Medical Image Analysis, Volume 2, Issue 1 , 1998.

[13] R. Castle, D. Gawley, G. Klein, and D Murray, "Towards simultaneous recognition, localization and mapping for hand-held and wearable cameras," in IEEE Int. Conf. on Robotics and Automation, April 2007.

[14] J. Barreto, J Roquette, P. Sturm, and F. Fonseca, Automatic calibration of medical endoscopes using a single image of a planar grid, Proc. British Machine Vision Conference, 2009.

[15] J. Y. Bouguet. "Camera calibration toolbox for Matlab", http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[16] M. Lourenco, J. P. Barreto, and F. Vasconcelos, "sRD-SIFT: Keypoint Detection and Matching in Images with Radial Distortion," IEEE Transactions on Robotics. (under submission).

[17] M. Lourenco, J. P. Barreto, and A. Malti, "Feature Detection and Matching in Images with Radial Distortion," IEEE International Conference on Robotics and Automation, 2010.

[18] R. Melo, Interfaces and Visualization in Clinical Endoscopy Masters thesis, FCTUC, Portugal, June 2009.

[19] M. Antunes, Integration of optical tracking system for determination of the position and orientation of instruments durin the surgery, Masters thesis, FCTUC, Portugal, August 2008.

[20] R Melo, J. P. Barreto, and G. Falcão, "Improved Visualization in Medical Endoscopy by Real-Time Radial Distortion Correction," IEEE Transactions on Medical Imaging (under submission).

[21] M. Antunes,"3D Reconstruction and Registration in Medical Endoscopy," PhD thesis project, FCTUC, Portugal, September 2010.