I. Introduction

Augmented reality (AR) in minimally invasive surgery has rapidly grown over the recent years [1]. Commonly, the surgical scene is augmented through the endoscopic view with a 3D model extracted from a preoperative acquisition. Nevertheless, due to the probable pneumoperitoneum and the patient displacement, the organs of interest often drastically change in shape and place [2]. Methods exist that attempt to recover or simulate the distortion between the pre- and intraoperative states. For example; surface reconstruction can be used to recover the organ surface from the endoscopic image, but the registration is reliable only for the visible part of the organ. Also, methods have been designed to recover the organ distortion through simulation using biomechanical models or through successive intraoperative 3D acquisitions but these still require yet another calibration process to locate the endoscope. The most popular solution is to resort to optical tracking; however this technique is costly, cumbersome and prohibits the surgeon from occluding the line of sight between the tracking cameras and the endoscope. This also requires many calibration steps that can slow down the surgical workflow.

The 3D rotational C-arm is one piece of equipment getting more and more popular as its relatively small bulkiness makes reasonable the use of intraoperative imaging. In this context, we propose a new paradigm to automatically register the referential frame of the intraoperative 3D imaging system with the one of the endoscopic camera, without any external tracking system or analysis of the endoscopic image. We propose to include the distal part of the endoscope in the C-arm acquisition (as shown in Fig.1) and estimate the position of the optical center from the scope position and orientation in the reconstructed volume.

Depending on the intervention stage, the camera is brought to move or remain still. For example, during the surgical stages of segmentectomy or thermal ablation, the endoscope may remain static. Then, our method allows an automatic registration that requires no additional equipment in the hybrid operating room or on the patient. If endoscopic navigation is required, we then could resort to a classic tracking technique but our approach would still remove any need of calibration for these systems.

*IHU, Institut Hospitalo-Universitaire, Strasbourg, FRANCE
†ICube (UMR 7357 CNRS), Université de Strasbourg, Strasbourg, FRANCE
‡IRCAD, Virtual Surg, Strasbourg, FRANCE

sylvain.bernhardt@ihu-strasbourg.eu
II. Methodology

In order to carry out AR, one needs to merge accurately the model rendering and the image from the endoscope. To achieve this, we must determine what is the exact view of the model that is corresponding to the endoscopic image content. This can be done by determining all the intrinsic and extrinsic parameters of the camera that are subject to varying from one intervention to another. Among the intrinsic parameters, we consider the position of the optical center, the (radial) distortion of the lens and the field of view. The extrinsic parameters are essentially the location and orientation of the camera in the model space (i.e. the 3D image space).

The intrinsic parameters are obtained through a classical camera calibration based on Zhang’s method [3] with filming of a chessboard, once the surgeon has fixed the zoom and optics as desired. Thereby, we can notably determine the focal length, the position of the optical center in the image plane, as well as the radial distortion of the lens, allowing to warp (undistort) the image. During experiments we have led, we usually got reprojection errors below 0.7 pixel.

To determine the extrinsic parameters, we use the video information about the endoscope tip included in the volume image. The metallic composition of the endoscope makes it highly reflective to x-rays and yields very large values in the CT image – much larger than surrounding artifacts or what any human tissue would produce. Thus, it is trivial to threshold the voxels representing the endoscope tip. Afterwards, knowing the tubular (right cylinder-shaped) shape of the endoscope, we retrieve the orientation and tip location of the endoscope in the volume image. It is worth observing that the resulting tip location does not necessarily correspond to the position of the optical center on the cylinder axis, as it depends on the zoom factor. Nonetheless, calibration experiments have shown that, with the FOV often set to a minimum value during interventions, the optical center is very close to the actual tip of the endoscope. In case of a zoom change, we could resort to a solution similar to [4,5] i.e. inferring the FOV from the size of the circular outline in the endoscopic image. So, we are able to determine all the extrinsic parameters but one; the tubular shape of the endoscope prevents from determining the actual roll angle around the axis of revolution. This estimation is fully automatic by using a camera-mounted high-end accelerometer which measures the angular position (pitch and roll) with respect to the gravity field.

III. Experiments

A dedicated chessboard has been made to assert our method’s accuracy. It has been painted with a mix of black paint and baryum sulfate, a common radio-opaque component used in radiology, which makes it visible from both the CT machine and the camera. Once the camera has been calibrated and the tiny chessboard has been acquired by both modes of imaging following the protocol previously described, the endoscope position in the CT volume has been estimated with our method. Then, we independently extract the chessboard corners from both the CT volume and the endoscopic image with OpenCV and VTK. Thus, once the image from the volume rendering is blended with the camera view, the registration error (in pixels) can be computed between the two sets of image points across the 1080p image. We performed this experiment three times with different positions for the chessboard and the camera and got no more than 7 pixels of average registration error. The largest registration error corresponds to approximately no more than 0.70mm in the scene. Therefore, these errors are significantly small enough to fulfill AR purposes for surgical procedures.

We also performed our method on in vivo pig data, as shown by Fig.2.
IV. Conclusion

In this article, we presented a novel way of registering the intraoperative model with the endoscopic camera, enabling AR without any external tracking device. The inclusion and analysis of the endoscope tip inside the C-arm field of acquisition allows a precise determination of almost all the camera extrinsic parameters. Rigorous evaluations with a calibration object and in vivo showed both quantitatively and qualitatively the benefits of our method. In the near future, more testing ought to be done to fully validate the technique and an integration to an operation room should also be achieved in order to experiment fully intraoperatively.

Acknowledgments

The authors would like to thank the IHU of Strasbourg for making the experiments possible, as well as P. Mewes and J. Fallert from Siemens and Storz for lending the necessary equipment. This work is part of the LASAR projet funded by IHU.

References


