Automatic localization of endoscope in intraoperative CT image : a simple approach to augmented reality guidance in laparoscopic surgery

Sylvain Bernhardt

UNIVERSITÉ DE STRASBOURG

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE

ÉCOLE DOCTORALE ED 269 - Mathématiques, Sciences de l'Information et de l'Ingénieur

ICube, Équipe AVR (Automatique, Vision et Robotique)

THÈSE présentée par :

Sylvain Bernhardt

soutenue le : 25 Février 2016

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : **Informatique**

Automatic Localization of Endoscope in Intraoperative CT Image: A Simple Approach to Augmented Reality Guidance in Laparoscopic Surgery

THÈSE dirigée par :

Christophe DOIGNON	Professeur, ICube, Université de Strasbourg
Luc SOLER	Directeur de recherche, IRCAD Strasbourg

RAPPORTEURS :

Philippe CINQUIN	Professeur, TIMC-IMAG, Université Joseph Fourier, Grenoble
Pierre JANNIN	Directeur de recherche INSERM, Rennes

EXAMINATEURS :

Nassir NAVAB	Professeur, Technische Universität München
Stéphane NICOLAU	Directeur de recherche, IRCAD Strasbourg

Acknowledgements

First off, I would like to deeply thank Stéphane Nicolau for his trust and for giving me the opportunity to take part in this great project. I really enjoyed working under his supervision and I think the quality of our relationship was a prominent part of the success of this enterprise.

I would also like to thank Professor Christophe Doignon for his valuable guidance and feedback during these three years and Vincent Agnus for his assistance regarding numerous specific yet important technical details. Likewise, sincere thanks go to to Professor Adrien Bartoli for his paramount help regarding our work on shading analysis.

I am grateful towards Professor Jacques Marescaux and Professor Luc Soler for welcoming me into their institution and laboratory respectively. I am also really thankful towards the IHU of Strasbourg for funding my work.

Some parties have been of incredible help during the experimental phase of my work. I am very grateful to the technical team of the IHU of Strasbourg (Mourad, Gaël, Frank) and to our partners from the industry (Johannes Fallert from Karl Storz and Philip Mewes from Siemens Healthcare).

Moreover, I would like to thank my lab mates (especially Jordan, Fayçal, Valentin and Anant) for their support and friendship which have definitely smoothened my labor. Thank you also Caroline for the big book on radiology. Thank you Negar for your patience and aid during my seemingly endless quest for papers.

On a more personal note, I would like also to profoundly thank my family and especially my parents for their constant support during this journey. Similar thoughts go to Audrey for being there for me and for brightening the overwhelming days. Thank you also to Sullivan for providing the necessary amount of entertainment to balance the work load.

I am thankful towards the good people from the badminton club of UTA Strasbourg, many of whom have become close friends, for welcoming me and for providing so much fun during my leisure time.

Finally, I would like to thank Antoine and Camille for their precious friendship and for making me the godfather to their first child, Jeanne, who is coincidentally born on the very same day I finished writing this thesis.

Terminology and Abbreviations

Terminology clarification

in vivo: on living tissues

ex vivo: on dead and extracted tissues

in vitro: on a phantom

in silico: via a simulation

in situ: on the spot, at the operation place

ex situ: at a different location than the operation place

preoperative: before the operation

perioperative, peroperative, intraoperative: during the operation

postoperative: after the operation

Abbreviations and acronyms

AR : Augmented Reality

AV : Augmented Virtuality

CBCT : Cone-Beam CT

CT : Computed Tomography

DoF : Degree of Freedom

EM : electromagnetic

ENT : Ear, Nose, Throat

FLE : Fiducial Localization Error

fps : frames per second

FRE : Fiducial Registration Error

HMD : Head-Mounted Display

HU : Hounsfield Units

ICP : Iterative Closest Points

MIS: Minimally Invasive Surgery

MR(I) : Magnetic Resonance (Imaging)

NMI : Normalized Mutual Information

NOTES : Natural Orifice Translumenal Endoscopic Surgery

OR : Operating Room

PCA : Principal Component Analysis

RMS : Root Mean Square

ROI : Region Of Interest

SfM : Structure from Motion

SfS : Structure from Shading

SLAM : Simultaneous Localization And Mapping

SOI : Structure Of Interest

TF : Transfer Function

ToF : Time of Flight

TPS : Thin Plate Splines

TRE : Target Registration Error

TVE : Target Visualization Error

US : Ultrasound

VR : Virtual Reality

Publications

Journal article

Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. "Automatic Localization of Endoscope in Intraoperative CT Image: A Simple Approach to Augmented Reality Guidance in Laparoscopic Surgery." In Medical Image Analysis, pp. 130-143, 2016. – Available online

Conference papers

Sylvain Bernhardt, Stéphane A. Nicolau, Adrien Bartoli, Vincent Agnus, Luc Soler, and Christophe Doignon. "Using Shading to Register an Intraoperative CT Scan to a Laparoscopic Image." In Computer-Assisted and Robotic Endoscopy (CARE), 2015. MICCAI 2015 satellite event – Oral presentation, Best Paper Award (2nd place).

Stéphane A. Nicolau, Sylvain Bernhardt, Vincent Agnus, Christophe Doignon, Luc Soler, and Jacques Marescaux. "Validation Methodology For Augmented Reality In Laparoscopic Surgery." In Computer Assisted Radiology and Surgery (CARS), 29th International Congress on, 2015. – Oral presentation

Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. "Automatic detection of endoscope in intraoperative CT image: Application to AR guidance in laparoscopic surgery." In Biomedical Imaging (ISBI), 11th International Symposium on, pp. 563-567. IEEE, 2014. – Oral presentation

Contents

Chapter 1

Introduction

This chapter explains how the digital world has revolutionized surgery, especially through imaging. This is perfectly exemplified by the introduction of intraoperative 3D imaging systems in so-called hybrid operating rooms. Using this technology at our institution, we started a project to tackle some of the current challenges modern surgery faces. The goals of this project are listed as well as those of this thesis, which is part of the former. A description of all the sections also reveals their content. Finally, a list of all the contributions achieved during the course of my PhD is provided.

Source: Siemens Healthcare

1.1 General overview

Modern medicine is digital
The digitization of our world has not spared the medical field. For instance, although medical imaging technically exists since 1895 with Röntgen's discovery of X-ray, its true revolution occurred with the development of digital computing in the 1970s and the creation of Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). These techniques allow a three-dimensional acquisition of any part of a patient's anatomy or pathology which then becomes visible to physicians. For the past decades, tremendous progress has been achieved in order to improve the image quality and rendering thanks to the exponentially growing calculation power of computers. Indeed, major advancements in electronics and digital signal processing enable to run increasingly complex algorithms. Digitally processing medical images allows to automatically enhance, segment and register them, even if they originate from different types of imaging modality. Hence, it is possible to virtually model and render a patient's inside in great details and 3D (see Fig. 1.1). A few companies have even specialized in this area [1].

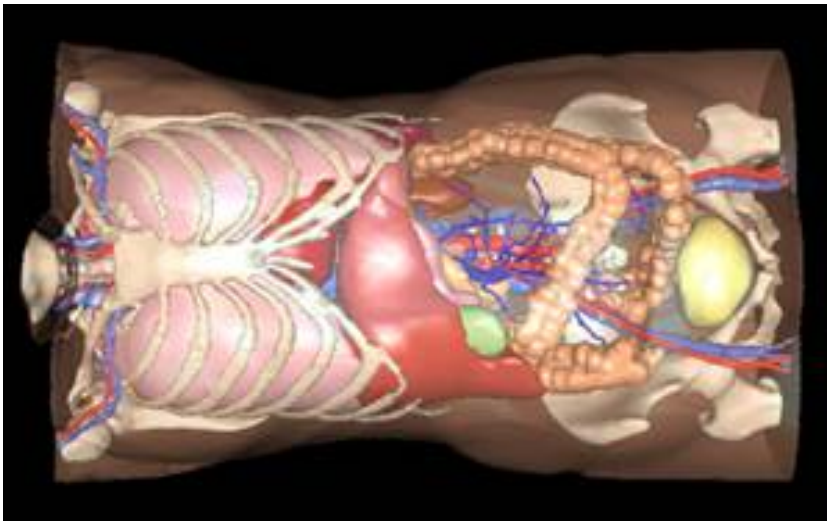


Figure 1.1: Three-dimensional model of a patient's organs.

Three dimensional medical imaging has so far mostly been used for diagnosis and surgical planning. Such data is usually acquired and processed before the intervention in order to assess the pathology and to plan the surgeon's course of actions. Typical visual aids include virtual markers, outlining of a target, resection trajectories and needle placements. The resulting information can be displayed on a screen during the intervention in the operating room (OR) alongside the data itself as a 3D model or 2D slices (see Fig. 1.2).

Minimally invasive surgery: the new surgical art
Surgery has also experienced major advancements for the past decades. Until the 90s, open surgery has been predominant for most interventions, during which the patient was cut open in order to access the target organs. Since then, Minimally Invasive Surgery (MIS) has progressively been the approach of choice. This kind of procedure is typically performed

[1] MEVIS, Visible Patient, EDDA Tech and Intrasense among others

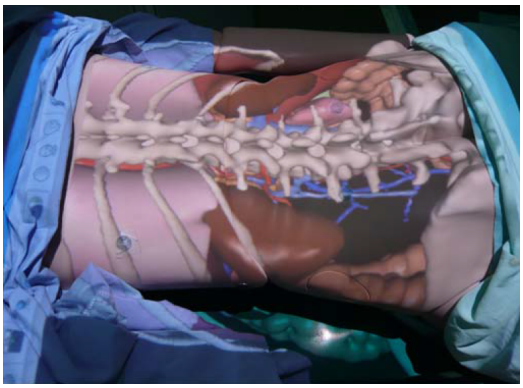
Figure 1.2: Various images and planning data displayed alongside the patient during surgery.

with long and thin instruments through little skin incisions or even natural orifices. Surgeons then resort to endoscopic cameras in order to guide their instruments and actions underneath the skin. The video feed is typically displayed on a dedicated screen next to the operating table. This separation impedes the surgeon's natural hand-eye coordination, which makes the learning curve of such interventions longer. However, MIS holds several advantages compared to its open counterpart, including smaller scars, reduced complication risks and faster recovery. On top of a constant cost reduction over the years, MIS has increasingly become popular.

From MIS to augmented reality

Performing the intervention via a digital display allows to easily enrich the surgeon's vision with valuable information. Those can be related to the patient's physiology (e.g. heartbeat), to the instruments (e.g. electrocauterization voltage) or else (e.g. intervention timing). Having those elements digitally overlayed onto the scene spares the surgeon from the necessity of constantly switching his gaze between the site and different monitors. Another very promising piece of information that can be superimposed to the laparoscopic video is the geometry and position of hidden anatomical structures, such as vessels, nerves, bones, organs and tumors. To achieve this, a standard approach consists in using preoperative data. Indeed, we previously mentioned that 3D models segmented from preoperative images and planning information were already commonly displayed for guidance on a screen in the OR. However, the surgeon has then to mentally match what he/she sees from the surgical scene with the preoperative information, often thanks to anatomical landmarks. This mental task may unnecessarily lengthen the intervention and can result in interpretation errors. Superimposing these anatomical models onto the surgical scene would thus help the surgeon focus as well as provide him/her see-through vision. This technique is known as Augmented Reality (AR).

For open surgery, AR can be achieved by projecting the image onto the patient, as shown in Fig. 1.3 (a). For MIS, this requires to digitally merge the rendering of the preoperative information with the endoscopic video, typically using transparency, as illustrated in Fig. 1.3 (b).


Source: IRCAD Taiwan [WWL+14]

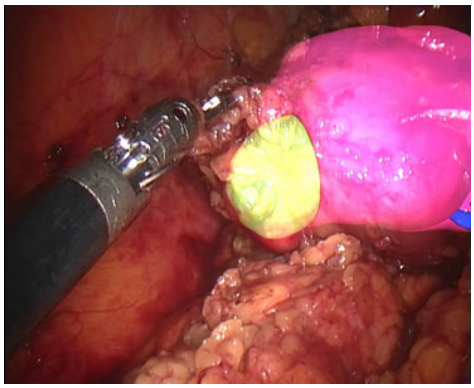

Source: [PMV+12]

Figure 1.3: Examples of AR in surgery by projection onto the patient (left) and by video see-through (right)

Interventions assisted by AR first occurred for neuro- and orthopedic surgery in the mid 80s [RSH+86], for which the accuracy of the alignment between the augmentation and the scene is easier to ensure due to the rigidity of the considered structures (brain and bones)[2]. As far as intra-abdominal surgery (also known as laparoscopic surgery) is concerned, AR-guided interventions are much more complicated due to the non-rigidity of organs. Between the preoperative and the intraoperative states, the patient's anatomy may change significantly due to a different patient's positioning. For example, a supine position is common for CT acquisitions, but a decubitus lateral position is preferred for various surgery with a folded table. Moreover, the numerous deformations that organs undergo during an operation (displacements, resections, etc) accentuate this discrepancy. Nonetheless, since the anatomy of a liver consists in many lobes, bile ducts and vessels and given that their position varies from one patient to another, imaging is a solid aid for interventions on the liver. As a result, two commercial software products have been released for hepatic surgery guidance [3], but they are dedicated to open surgery. As of now, these tools cannot be used in abdominal MIS, because no laparoscopic ultrasound (US) transducer with integrated electromagnetic tracking sensor has been commercialized yet.

To tackle this deformation problem, researchers turned towards the analysis of intraoperative data. Since endoscopic cameras provide only a 2D view of the visible surface, they also started to resort to imaging systems recently developed for interventional radiology (e.g. biopsies, percutaneous puncture). Indeed, CT or MRI scanners, usually designed for preoperative acquisitions, have been adapted to fit inside an OR, then designated as *hybrid* (see Fig. 1.4). The aforementioned deformations due to the pneumoperitoneum, the patient's different pose or the organ shifts are then directly taken into account in the 3D data as the acquisition is performed during the intervention. The surgeon is thus able to see an exact

[2] In reality, it has been proven since then that opening the skull causes a significant deformation of the brain [HHCS+03].

[3] CAS-ONE Liver™ from CAScination© and Explorer™ from PathFinder©

three-dimensional capture of patient's anatomy while he/she carries out the operation. If research shows that these interventional 3D imaging systems are also beneficial for surgical interventions, as it already is for radiology, new scanners will probably be designed for this specific purpose.

Figure 1.4: A hybrid operating room.

IHU of Strasbourg: a new institute to develop hybrid surgery
Showing the true potential of hybrid ORs is one of the objectives of the Institut Hospitalo Universitaire (IHU) of Strasbourg, France, a renowned medical and surgical center specialized in treating the pathologies of the digestive tract, notably the liver. Responsible for more than a hundred vital functions, the liver is a complex and indispensable organ whose pathologies are various and often lethal. According to the International Agency for Research on Cancer, liver cancer is the sixth most common form of cancer worldwide with more than 782,000 new cases in 2012. Moreover, it is the second most lethal one with more than 745,000 reported deaths in that same year. According to the American National Cancer Institute, the 5-year survival rate for patients with liver cancer is between 10% and 23% with non-surgical treatments such as chemotherapy. This rate rises to 27-70% with a liver resection and even 44-78% with a transplant. Hence, surgical treatments lead to a much better prognosis for liver cancer. AR aims at further improving this situation by assisting the surgeon during an hepatic intervention. By showing resection planes and critical structures, better margins can be achieved, hence sparing more healthy tissue, and the intervention time under anesthesia shall be reduced. Overall, more patients would be eligible to surgery and the risk of complications would be greatly reduced.

Presentation of the project
In light of this situation, the IHU of Strasbourg has created a research project called **LASAR**, which stands for Laparoscopic Assisted Surgery with Augmented Reality. This project aims at designing a practically feasible and yet effective solution in a hybrid OR to assist liver surgery thanks to augmented reality. By estimating the relationship between the intraoperative 3D data and the laparoscopic image, the former can augment the latter with relevant anatomical

structures, such as a vascular network or tumors. Such an innovation would shorten the procedure, decrease the risk of accidents and thus significantly increase the success rate. Our solution requires no preoperative image or tracking system, making the AR assistance immediate and accurate. Our method is based on an extraction of the endoscope purposefully showing in the intraoperative 3D data. As it is held by an arm and placed inside the patient towards the region of interest, the scanner acquires both at the same time (see Fig. 1.5). An analysis of the intraoperative 3D data thus provides its position and orientation with respect to the interventional 3D imaging system.

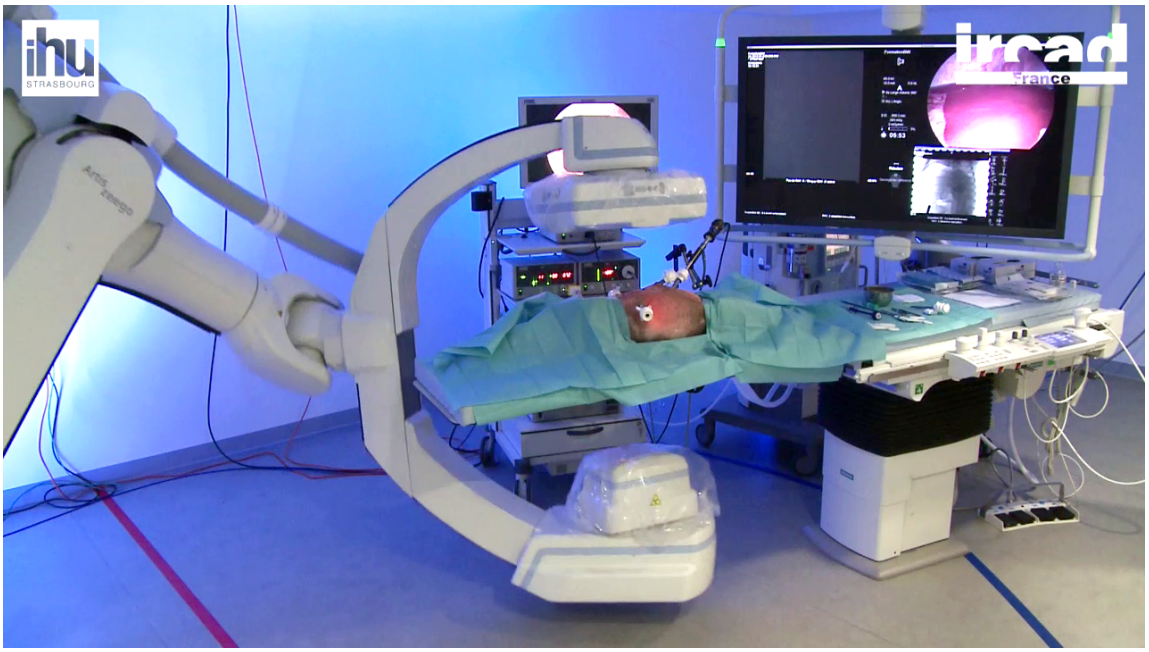


Figure 1.5: Example of intraoperative scan with the endoscope in the field of acquisition.

The LASAR project is led in close relation with the nearby Institut de Recherche sur le Cancer de l'Appareil Digestif[4] (IRCAD) in Strasbourg, France. There are also two industrial partners: Siemens, worldwide leader in healthcare solutions (especially through imaging), and Karl Storz, a major manufacturer of endoscopic equipment.

Objectives of this work

The work presented in this thesis is part of the LASAR project and focused on its core goal of static augmentation of the endoscopic image with information from the intraoperative 3D data. The objectives are the following:

- **Accuracy**. In order to truly assist the surgeon, our method shall provide an AR solution whose accuracy is well within the clinical requirements in MIS. The literature in surgical guidance for the liver typically reports an accuracy requirement of around 5 mm. Since the augmentation is performed in the image plane, the registration error may vary with respect to the depth in the scene due to perspective. Nonetheless, we aim at satisfying this accuracy requirement for the whole scene.

[4]Institute of Research about Cancer of the Digestive System

- **Compatibility**. Since our method is destined to be used in a hybrid OR during a surgical procedure, it ought to comply with strict constraints. Our AR solution shall naturally integrate in the surgical workflow and cannot lengthen the intervention beyond reason. Typically, we expect the augmentation to be provided within a minute after the intraoperative 3D scan.
- **Reproducibility**. The aim of the project is to go beyond the proof of concept and to provide a functional product. In order to truly advance the field of laparoscopic surgery, our AR solution must be usable in any hybrid OR by any surgical team. In other words, our method shall be compatible with common endoscopes and should not require additional impractical or expensive systems.

1.2 Organization of the report

This thesis is divided into seven chapters including the present introduction which constitutes Chapter 1.

Chapter 2 brushes some background in the different fields concerned by our work. It begins with a description of the four main medical imaging modalities: X-Ray, magnetic resonance, ultrasound and nuclear medicine. For each modality, we detail how the image is formed and what are its advantages, constraints and thus its typical applications. For instance, preoperative imaging is paramount to surgical planning. However, deformations occur between the scanning and intervention times. This compelled researchers to resort to intraoperative imaging systems to better capture the patient's anatomy during the operation. Originally used in interventional radiology, we will see how most modalities have been adapted to enter the OR and what are their potential for surgery. In particular, we will see to what extent such systems can be used for intra-abdominal MIS. Indeed, laparoscopy revolutionized surgery by providing a safer alternative to open procedures. However, this technique also introduces several constraints, which we will see by detailing the context and the different steps of a typical laparoscopic intervention. Based on these three outlines, the issue of fusion between images from different modalities naturally raises and the need of AR in MIS is revealed, which is the object of Chapter 3.

Chapter 3 establishes a comprehensive review of all the different methods of AR in intra-abdominal MIS proposed by the literature. They can be divided into two main categories. The direct methods of AR in MIS aim at directly aligning preoperative data to the endoscopic images by registering surfaces extracted from both. Their main disadvantage is the incapacity to handle deformations between the preoperative and the intraoperative states. On the contrary, the indirect methods of AR in MIS can compensate such deformations by using an intraoperative 3D imaging system as a middle step, but usually at the cost of requiring an extra apparatus such as a tracking system. All the various issues from these reviewed approaches are gathered and organized in order to outline the main remaining challenges of AR in MIS.

Chapter 4 describes the novel method that we propose to solve the previous issues that faces AR in MIS. In our project, we rely on an intraoperative 3D imaging system to provide accurate volumetric data about the patient's anatomy. Thus, based on the taxonomy used in Chapter 3, our AR approach is indirect. To provide the augmentation, the spatial relationship between the endoscopic camera and the interventional scanner must be determined, which

boils down to estimating six parameters (three in position and three in rotation). We propose to determine these six parameters without a tracking system, only by introducing the tip of the endoscope in the acquisition field of the 3D imaging system.
Concerning the extraction of the endoscope from the intraoperative 3D image, we first show that a simple threshold-based method is not stable, as the estimation of the six parameters varies with the value of the threshold. Thus, we introduce a matching method based on a cylindrical template, which is more robust but also more computationally expensive. Moreover, due to the tubular shape of the endoscope, the roll angle cannot be inferred from the intraoperative 3D data and we detail how including an accelerometer into the camera alleviates this issue.
Using a dedicated radio-opaque checkerboard, experiments have provided quantitative results that show a satisfactory accuracy but also demonstrate a constant bias for each different camera setup. We hypothesize that this phenomenon originates from the violation of the hypothesis of alignment between the optical axis and the rod's axis of revolution, even for straight endoscopes, as developed in Chapter 5.

Chapter 5 focuses on the actual structure of common straight endoscopes. Since we infer the position of the camera from the axis of the endoscope shaft, a tiny misalignment between the latter and the optical axis may result in significant registration errors. One way to compensate for this disparity by determining the projection of the endoscope axis in the image and translate the latter accordingly. This is equivalent to performing a perspective correction, but this approximation is acceptable since the misalignment is always small with respect to the dimensions of the scene.
Besides the purposefully designed decentering of certain endoscopes, this difference between the axes may have multiple factors such as the zoom, focus and position of the light cable, which are only set during the procedure. Therefore, this misalignment cannot be determined prior the intervention. Hence, we propose to easily determine this parameter during the surgery by making a pattern rotate around the endoscope. We show that the intuitive method of making the endoscope rotate on itself is not satisfactory both in terms of practicality and accuracy. Then, we introduce a novel solution that involves the rotation of a square tube resting at the end of the endoscope shaft.
We provide results from quantitative evaluations on the same radio-opaque checkerboard than in Chapter 4. The square tube approach proves to be the most practical and accurate one. This additional calibration induces a significant improvement in the registration accuracy compared to the results exposed in Chapter 4. Depending on the accuracy goal, we thus offer the possibility of taking into account impactful yet seemingly insignificant parameters. Our method then allows a registration between the intraoperative data and the endoscopic image with an average error below one millimeter at the nominal distance (5-10 cm).
Finally, we present an extensive statistical analysis of various experiments on checkerboard in order to evaluate the robustness of our method.

Chapter 6 presents *in vivo* results. We performed several sessions of experiments on pigs in order to demonstrate the clinical potential of our approach. We show instances of augmentation in various laparoscopic porcine interventions. A couple of examples show augmentations of the endoscopic view with liver lesions segmented from the 3D scan. For gynecological surgery, we used a volume rendering of the ureters being registered to the endoscopic image. In lung surgery, we showed how AR can avoid the surgeon's misinterpretation of the scene.
We also used these experimental data to perform an accurate texture mapping of the en-

doscopic image on the surface of the intra-abdominal cavity. We thus present instances of virtual photo-realistic models of surgical scenes.

Chapter 7 presents a novel Pose-from-Shading method. It has been designed as complementary to our registration method described in Chapter 4, which calculates the position and orientation of the endoscope. The remaining parameters, the roll angle and the translation in the image plane, are determined by an accelerometer and a dedicated calibration respectively. We propose to fix these same degrees of freedom by only considering shading constraints in the surgical scene. We also provide some background about the exploitation of shading in computer vision as found in the literature and its application to MIS.
Based on a simplified shading model, we propose a new piecewise constant albedo approach to shading, more appropriate to laparoscopic scenes. For a given pose of the camera, this allows us to measure the dissimilarity between the endoscopic image and the virtual view upon the intra-abdominal cavity surface extracted from the intraoperative 3D data. Minimizing this metric ensures that the endoscopic image and its augmentation are well aligned in the image plane. This method can thus complete our AR application in case the accelerometer and the square tube calibration are not available. In the opposite case, it may still serve as a refinement or control tool.
This Pose-from-Shading method has been tested on several porcine data. Results show an accuracy of 1.5° in rotation and less than 1 mm in translation.

In the **Conclusion** chapter, we first provide a summary of the main achievements of our project. Then, we detail the perspectives of our work, such as short-term plans to further improve our results, but also new related research topics that could complete our approach. Afterward, we discuss broader topics about AR such the potential of intraoperative scanners and the issue of depth perception. Finally, based on the content of this thesis, we present a complete and detailed concept of and ideal laparoscopic AR solution.

1.3 List of contributions

Here are the main contributions of the work presented in this thesis.

Introduction of a novel AR approach without external tracking of the camera

- We claim that our method is the first one to allow an automatic and accurate AR in laparoscopic surgery. All the data is acquired during the intervention, which maximizes the accuracy and circumvent the problem of deformation in preoperative data.

- Due to its simplicity, our method is very much importable to any hybrid OR with an intraoperative scanner. Thus, we provide a bronze standard to any research team who wishes to compare their laparoscopic AR method with another. This validation tool aspect of our project has been the object of a conference communication [NBA+15].

- Each augmentation is provided with very limited disturbance for the surgical workflow.

- The relationship between the endoscopic camera and the intraoperative 3D data is determined through a rapid extraction of the endoscope tip in the volume. We have drawn the attention to the difficulty of accurately segmenting a metallic object in the interventional CT data and present a robust template-based method to alleviate this issue.

- The augmentation is admittedly static and yet already useful for most laparoscopic procedures. It requires no external tracking system, which are often costly and cumbersome. If

dynamic augmentation is desired, our method would still provide an accurate initialization, which removes any need for calibration between the intraoperative imaging system and a possible tracking device.

- A preliminary version of our work has first been published in a conference paper presented orally [BNA+14]. A more detailed and advanced version has been submitted and accepted as a journal article [BNA+16].

Highlighting and solving geometrical relationship issues between the camera and its tubular container

- Our AR method relies on the unity of the endoscope and the camera. The position and orientation of the latter is provided by the extraction of the former from the intraoperative 3D data. However, we expose the impact of various factors upon their relationship, especially the misalignment between the axis of revolution of the shaft and the optical axis. This discrepancy, although often small, may still negatively affect the augmentation.

- The estimation of the geometrical relationship between a camera and its body is typically performed through an hand-eye calibration, but such a process cannot be performed during the intervention. Therefore, we propose a novel way to compensate intraoperatively for the induced perspective difference in the image plane thanks to a simple rotation of the scene with respect to the endoscope. This method has been described in our journal article [BNA+16].

Introduction of a novel Pose-from-Shading method

We propose a second solution, based on shading analysis, to complete the registration of our AR method without an accelerometer nor a supplementary calibration like the one with the square tube. Most shading methods aim at reconstructing the surface from shading constraints, while we only use it to refine the pose of the endoscopic camera. Thus, instead of Structure-from-Shading (SfS), we rather perform Pose-from-Shading.
Moreover, most shading methods assume that the albedo is constant for the whole scene. In a laparoscopic scene where various organs and tissues are often exposed, this hypothesis is unlikely to stand. To make our method applicable to such scenes, we relax this hypothesis by using a piecewise constant albedo approach.
In our method, the surface of the intra-abdominal cavity is extracted from the volume and observed by the virtual camera at a given pose. We then use the local formulation of the shading constraints to calculate a dissimilarity metric between the endoscopic image and the surface view. The pose of the camera is determined by the initial registration and three degrees of freedom composed of the roll angle and the translation in the image plane. By optimizing this criterion, we can obtain the pose of the camera, using only on the endoscopic image and the intraoperative 3D data. This method has been the object of a conference paper presented orally [BNB+15].

Porcine experimentation showing the feasibility in practice

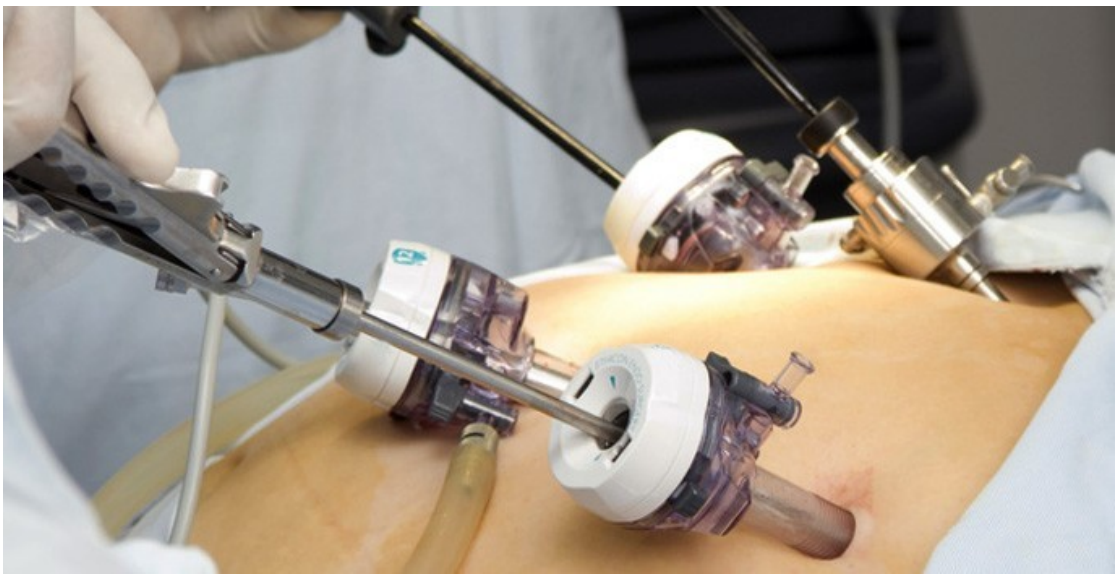
- We demonstrate the clinical feasibility of our AR method by reporting several experiments on pigs. While the method has been applied offline, the interventions did occur based on standard setups. Therefore, our method has been used in realistic conditions. The checkerboard calibration is typically performed within 30 seconds. The square tube calibration can be done within 20 seconds. Both require good lighting conditions without reflections, which can be achieved thanks the various potential light sources in an OR. Overall, our method has very limited interference with the standard surgical workflow.

- These experiments have also showed that introducing the endoscope in the intraoperative acquisition while looking at the liver is really not troublesome. Our method complies with most of the standard poses of laparoscopes. Such poses can be held using common articulated arms. The intra-abdominal anatomy of pig being very similar to ours, we can thus safely extend this remark to human interventions. Moreover, these *in vivo* revealed that the presence of the endoscope in the scan was not provoking artifacts in the region of interest.

Chapter 2

Medical Imaging and Laparoscopic Surgery

This chapter brushes some background in the different fields concerned by our work. It starts off with a description of the various medical imaging modalities currently available with details about their functioning and typical applications. An introduction to laparoscopic surgery is then provided, where common protocols and constraints are explained. The imaging modalities are often used preoperatively for planning, but also intraoperatively for guidance. Indeed, one particular challenge is the deformation of the patient's anatomy that can occur between the time of the preoperative scan and that of the operation. This compelled for the importation of various imaging modalities into the OR and we describe how these are adapted so.



Source: AdvancedSurgeonsPC.com

2.1 Medical imaging

Medical imaging allows physicians to observe a patient's inside without resorting to cutting. Hence, it constitutes an essential means of diagnosing pathologies, following their evolution and potentially planning a surgical intervention. Various modalities of imaging have been developed over the last 120 years, but they all have the same goals:
- **Contrast**: the potency of separating tissues of different natures;
- **Resolution**: mostly spatial, for the ability of revealing fine details, but also temporal in case of a dynamic acquisition, for the ability of exposing fast motions;
- **Safety**: the limitation of hazard for the patient, the operator and, if applicable, the surgeon.

Through these aspects, we will now review the four main imaging techniques used nowadays. The focus here is on their characteristics and their clinical applications, principally based on [BB11]. These modalities may be used in laparoscopic surgery for preoperative planning, but also intraoperatively for guidance, as detailed in Section 2.3.

X-ray imaging

X-rays refer to a certain kind of electromagnetic radiation whose wavelength ranges from 0.01 to 10 nm. If a pulse of X-radiation passes through a human body, its energy is more or less absorbed by tissues, according to the atomic number of their atoms. By collecting the remaining X-rays on a support as they exit the body, we are then able to perceive the various organs traversed by the beam (see Fig. 2.1 (a)). The contrast of the image thus originates from the composition and density of the tissues. For example, bones are rich in calcium, which has a high atomic number, and as such absorb more radiation than soft tissues, resulting in brighter representations in the image. Likewise, air, which has a low density, does not absorb much of the X-rays and thus appears very dark. Performed since the discovery of the X-ray in 1895 by Wilhelm Conrad Röntgen, this imaging technique is referred to as **projective radiography**. The receptive support was originally a film, but over the course of the 20th century, it has been progressively replaced with an array of electronic detectors, improving the image quality. Projective radiography is commonly used in orthopedics and dentistry due to the high contrast of bones and teeth. This imaging modality is able to reveal very small details with a spatial resolution of 0.08 mm, even 0.03 mm in mammography.

While a static image may be satisfactory, sometimes a continuous acquisition of several X-ray projections is performed. This dynamic radiography is referred to as **fluoroscopy** and is commonly used in angiography, but also as a guidance tool in interventional tasks such as the insertion of catheters. Its spatial resolution is 0.125 mm and its temporal resolution is usually real-time, i.e. 30 frames per second (fps), but can go up to 120 fps for cardiac observations.

Beside projections in which all tissues and organs are collapsed, radiography is also used to capture cross-sections of a patient's body into an image. By pulsing thin fan-beams of X-ray while both the emitter and the detector rotate around the subject, it is possible to capture hundreds of profiles of the anatomy in this planes (see Fig. 2.1 (b)). Then, using a mathematical tool called the Radon transform, these one-dimensional profiles are converted into a two-dimensional image. This principle, called tomography, has first been mathematically solved by Johann Radon in 1917. However, tomography has been concretized only in the 1970s by Godfrey Hounsfield thanks to the advent of computers, which is why this imaging technique is referred to as **Computed Tomography** (CT) or sometimes Computed

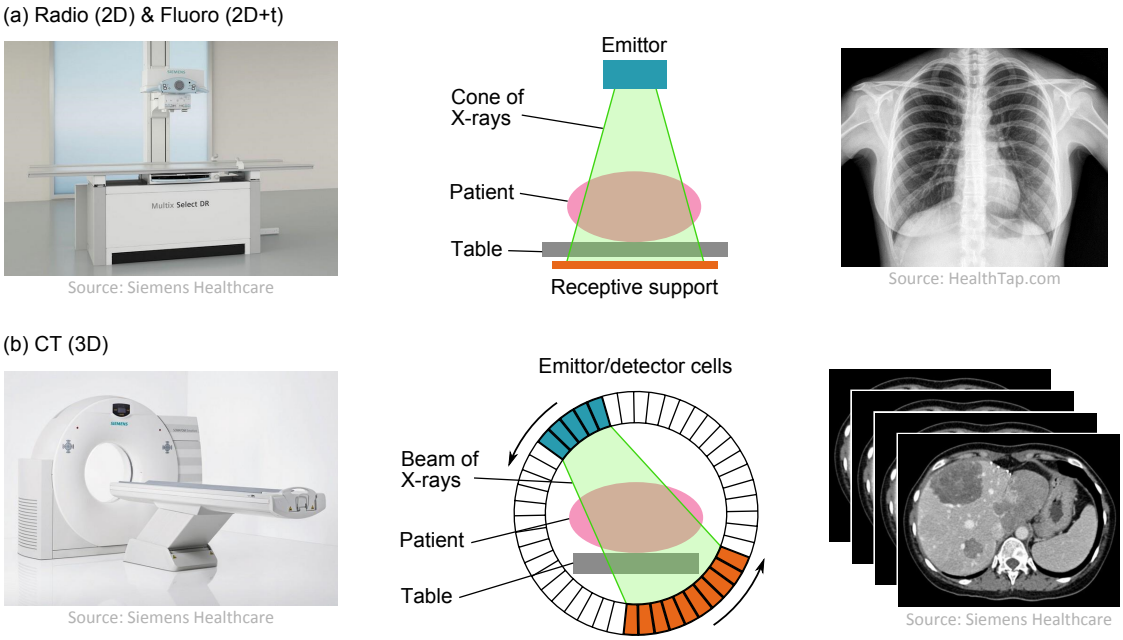


Figure 2.1: The main modalities based on X-radiation. For (a) Projective radiography/Fluoroscopy and (b) Computed Tomography, from left to right: picture of system, diagram of functioning and example of image.

Axial Tomography (CAT). Moreover, by acquiring successive cross-sections of a certain part of an anatomy and concatenating the resulting images, it is then possible to obtain a three-dimensional image of this whole area. Today's CT scanners feature a spatial resolution of 0.3 mm and perform an acquisition with hundred of slices in merely a few seconds. In a CT image, the density of the tissues is measured relatively and expressed in Hounsfield Units (HU). For example, water is assigned to 0 HU, while less dense components, like air or fat, are represented by negative values. On the contrary, materials denser than water, like muscles and especially bones, are designated by positive values. The intensity of the three-dimensional pixels, or voxels, is related to this measurement and thus compose a contrasted volumetric image of a whole body part.

However, despite the ability of current scanners to precisely measure the X-ray absorption of various tissues, sometimes the difference is to little to result in a contrasted and understandable image. Therefore, physicians often resort to the usage of contrast agents to artificially increase the attenuation discrepancy between a target and its surrounding structures, thus greatly emphasizing the former in the image. Indeed, these products are typically compounds with high atomic numbers, which make them much more opaque to X-ray than any organic matter. They can be administered orally or intravenously depending on the target.

So, X-rays are the base of various imaging techniques: projective radiography for 2D imaging, fluoroscopy for dynamic 2D imaging (2D+t) and CT for 3D imaging. An adaptation of CT towards dynamic 3D imaging (3D+t or 4D) is even under investigation [FKB+13], but is greatly limited by the amount of radiations required. Indeed, despite their common use, imaging techniques using X-rays are not harmless. Such radiations are ionizing for the human body and can provoke skin burns or even cancers. The risk increases with the dose

of radiations, related to both intensity and duration of exposure. Therefore, an equilibrium has to be reached between dosage and image quality. Due to the large amount of radiations necessary to perform a CT scan, physicians only resort to this imaging modality if they consider that the medical gain outweighs the risk, ever so little. This is often referred to as the ALARA principle (acronym for As Low As Reasonably Achievable). Fortunately, other less harmful imaging techniques are available, such as Magnetic Resonance Imaging.

Magnetic resonance imaging

Invented in the 1970s, Magnetic Resonance Imaging (MR or MRI) is performed on a large tore-shaped machine composed of a giant superconducting magnet (see Fig. 2.2). As the patient lies on a table inserted in the tube, the magnet generates a very strong magnetic field, measured at 0.5 to 2.0 T (Teslas), which is 10,000 to 40,000 times stronger than the Earth's magnetic field. A human body is predominantly composed of hydrogen atoms, especially present as a constituent of water. Since hydrogen nuclei contain a single proton, these atoms are dipoles, meaning they possess a north and south poles and are sensible to magnetic fields, as would a compass. Due to the strong magnetic field generated by the MR system, all hydrogen nuclei align with the field: half towards the north and half towards the south. However, this proportion is not perfect and there are always a few atoms more in either direction (approximately a few in a million). At the very moment of scanning, three additional smaller magnets, called gradient coils and oriented along the X, Y and Z-axes, locally perturb the field and allow to focus the acquisition in a certain plane across the patient. At the same time, a brief second electromagnetic wave, though of much smaller intensity, is then emitted by a third kind of magnet called the radio frequency coil. The frequency of these waves is precisely set in order to make hydrogen nuclei resonate and line up towards their opposite pole. Once this excitation stops, all these unstable protons return back to their original state, but emit back some of the excitation energy as a radio wave then captured by the system. Analyzing the frequency and phase of these waves allows to calculate their exact origin. Moreover, the delay between the excitation and the relaxation of these atoms depend on the type of tissue they are part of. Hence, given the location of each unstable protons and their discrepancy in relaxation, a computer can form an image representing a cross-section of the patient's anatomy. By performing several successive acquisitions and concatenating the resulting slices in a similar manner that of CT, one can thus obtain a three-dimensional image of the targeted body part.

Contrary to CT, MRI is capable of providing an excellent contrast in soft tissues. Therefore, it is widely used for neuroimaging, oncology and even to detect muscular lesions. Moreover, thanks to the gradient coils, an MR scanner is able to capture slices from the patient in any direction, whereas CT is mechanically limited to its rotation plane. However, an MR acquisition is much longer than a CT acquisition and typically requires tens of minutes depending on the desired size of the 3D image. Also, MRI spatial resolution is typically 1 mm, but could be refined with systems generating magnetic fields stronger than 2.0 T at the cost of much larger amounts of energy.

One last major difference of MRI compared to CT is the radiation safety. The magnetic field generated is considered harmless and the gradient magnets emit radiations which, given their wavelength, are not ionizing. The only risk of MRI is the strength of the magnetic field, turning any remote metallic object into a potential projectile attracted to center of the main magnet, where the patients body is at. Therefore, any MRI room is completely

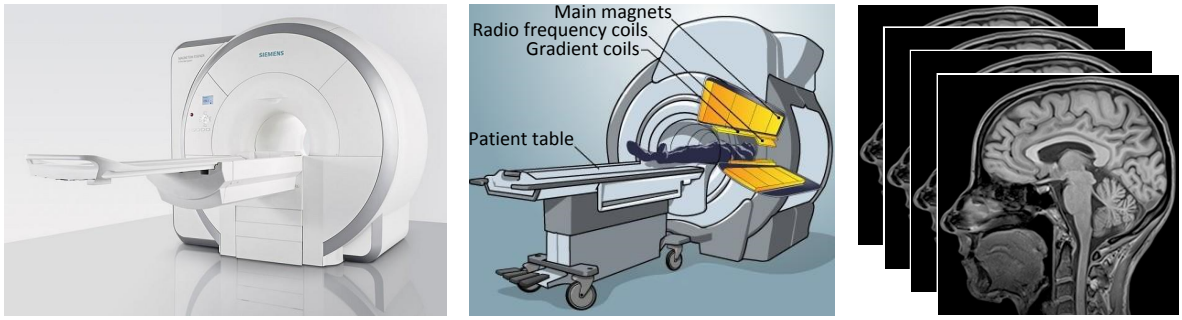


Figure 2.2: Magnetic Resonance Imaging. From left to right: picture of system, cutaway of structure and example of image.

empty of mobile metallic objects. Also, patients with a pacemaker cannot benefit from MRI as the strong magnetic field may lethally interfere with their device. Sometimes low-dose CT scans are performed beforehand to assess the presence of unknown metallic fragments in the patient's body, as these heat up from the radio waves and may cause burns.

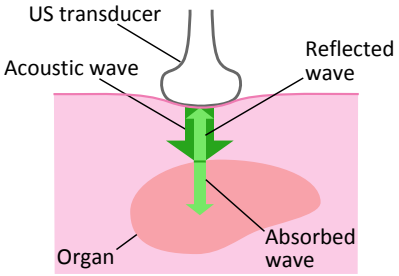
By resorting to special coils and reducing the resolution, the acquisition time of MRI can drop to a few seconds and can be also used to capture motion-prone parts of the anatomy such as the abdominal area, where breathing and heart beats would prevent to perform a standard MR acquisition. Such imaging technique is also able to capture blood motion, making it very useful for angiography. Moreover, by monitoring the blood flow in the brain, physicians are now able to observe the neural reactions to various stimuli. This imaging modality is referred to as functional MRI (fMRI).

Ultrasound

Contrary to the other modalities that rely on electromagnetic waves to reveal a patient's anatomy, ultrasound (US) imaging uses short high-frequency acoustic waves (1 to 5 MHz) that mechanically propagate in the tissues. A transducer, composed of an array of acoustic emitters and receptors, generates these waves in a plane above the region of interest (see Fig. 2.3). The sound, while traveling through the tissues, is partially reflected at the interface between two internal structures of different densities. Each reflected wave then propagates backwards to the transducer, which detects them. Given the delay and magnitude of an echo, a computer then calculates its depth and the density discrepancy that caused it. These are translated as a pixel location and intensity, respectively.

This modality of imaging is capable of a spatial resolution of 0.3 mm and is fast enough to allow real-time acquisitions of areas up to 10 cm deep. However, if the difference of density between two matters is too large, then the totality of the acoustic wave is reflected and what is beyond that interface remains invisible. This phenomenon typically occurs at the boundaries between soft tissues and solids like bones or between gases like air and soft tissues. As a result, for the sake of image quality, the operator has to maintain a good contact between the transducer and the surface, often aided by gel but also by applying a significant pressure, which tends to distort the anatomy below. Moreover, ultrasound is more prone to noise than other modalities, which tends to blur edges and limit image processing possibilities.


Source: PhcTorrance.org




Source: Radiopaedia.org

Figure 2.3: Ultrasound imaging. From left to right: picture of a US probe held against the abdomen, functioning of US imaging and example of US image of the liver.

Nonetheless, since ultrasonic waves cause no harm to the human body, ultrasound proves to be a safe, simple and cheap solution to medical imaging. It is mainly used in obstetrics and gynecology, cardiology, urology and oncology.

Acoustic waves reflected by moving structures have their frequency change. This phenomenon, known as the Doppler effect, can be used in ultrasound imaging to measure the velocity magnitude and direction of moving fluids in the body like blood.

Nuclear Medicine Imaging

In the previous modalities, an energy is provided by the system, reflected or transmitted by the patient's tissues and detected again by the machine in order to form an image. In nuclear medicine, the energy originates from radioactive markers introduced in the patient. Indeed, after being injected, these agents will naturally scatter in the body through the bloodstream. Besides, the radioactive decay of these isotopes emit gamma rays that are detected by the system and, upon calculation of their origins, an image is formed. Similarly to CT, this image represents only a slice of the patient and performing successive acquisitions is required to obtain a 3D reconstruction.

Since the concentration of radioactive markers depends on their targets and the patient's metabolism, nuclear medicine imaging (also known as molecular imaging) is rather oriented towards diagnosis and assessing the patient's physiology than his anatomy. One of its primary use is the detection of tumors and infections, as both generally provoke a high metabolic activity. There are two main techniques of nuclear medicine imaging: Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET). SPECT is less costly than PET but offers a spatial resolution of 7 mm, against 5 mm for PET. Moreover, the latter is less prone to artifacts and performs acquisitions faster. Given its poor resolution and its functional nature, nuclear medicine imaging is frequently combined with a CT acquisition in order to allow the fusion of anatomical and physiological data (SPECT-CT and PET-CT). Finally, as for X-rays, gamma radiations are also ionizing and thus should be considered with caution.

So, all these imaging modalities allow physicians to observe a patient's anatomy. In surgery, this proves particularly helpful for the planning of an intervention. The information revealed by the image greatly increases the accuracy of the "roadmap" and thus the chances of

positive outcome. Such preparation is particularly relevant in laparoscopy, where the surgeon has no direct vision on the scene.

2.2 Laparoscopic Surgery

In this section, we first consider the emergence and impact that laparoscopy has made in the field of abdominal surgery. Then, we describe a typical workflow of a laparoscopic intervention, thus emphasizing the surgical context of the LASAR project.

The advent of laparoscopic surgery

The content of this subsection is mainly based on [SW97] [HH05] [EC08]. Until the 20th century, surgeons used to gain visual and manual access to the abdominal organs only through laparotomy, whereby the patient's abdomen would be cut open. In 1902, George Kelling insufflated air into the abdominal cavity of a dog and peeped inside through an endoscope, thus introducing both the concepts of **pneumoperitoneum** and celioscopy (also later known as **laparoscopy**). The main benefit was to decrease blood loss, a major lethal factor back then, thanks to the smaller incision and the pressure inside the cavity. Many improvements occurred in the following decades. For example, carbon dioxide has become the gas of choice for the insufflation, because of its easy absorption by the tissues and its non-inflammability. Also, the use of optic fibers displaced the lighting source out of the patient while drastically increasing its brightness and quality. Besides, long thin instruments have been designed in order to facilitate manipulations in the cavity. However, laparoscopy has been mainly used for diagnosis up to the 1980s. Its first therapeutic use was an appendectomy in 1983, shortly followed by a cholecystectomy in 1985, but the impracticality of operating with one hand holding the endoscope remained. This issue has been lifted a few years after with the invention of electronic video sensors. Replacing the eye at the proximal end of the endoscope, such a camera would magnify and display the surgical scene onto a television screen for all the surgical team to see, while freeing the surgeon's hands.

Despite these constraints, laparoscopy has been applied to many types of abdominal surgery, including hernia repair [SSSW94], gynecologic procedures [JKB+04] [SWM97] and interventions on the gallbladder [MBA+94], adrenal glands [BDN+96], appendix [AHM+92], colon [oSTSG+04], stomach [HMS+05], esophagus [SPL+10], prostate [BBDK06], kidney [GKL+07] and, most importantly for our project, liver [BKM+02] [MMR+03] [BFD+07] [TFA+08] [CVR+09]. Many positive outcomes have been reported, especially reduced blood loss, shorter hospitalization, faster recovery, fewer complications and increased patient's comfort due to smaller scars, while maintaining or improving surgical success rates [NMT+11] [SSD+13] [SET+14]. In hindsight of these various advantages over laparotomy, laparoscopy has become a reliable solution for almost every type of intra-abdominal interventions, notably on the liver [BCG+09].

Nonetheless, the transition to MIS has also introduced inherent limitations [BCBC12]. Contrary to open surgery, the vision of the scene on a screen is two-dimensional, hence the surgeon loses both hand-eye coordination and sense of depth. Moreover, since the laparoscopic instruments are inserted through trocars, the surgeon's gestures are inverted and their range of motion is limited. Also, tactile feedback is removed since the surgeon cannot palpate directly the tissues and thus is no longer able to "feel" tumors or the pulse of a superficial vessel. These also constraints make the learning curve of laparoscopic procedures longer, but

solutions are possible in both hardware and software [BCBC12] and some are being developed. For example, the introduction of stereo endoscopes and 3D screens may restore the surgeon's perception of depth. Surgical robots, such as the Da Vinci from Intuitive Surgical, Inc. (see Fig. 2.4), allow the surgeon to perform the intervention on a remote console in the OR, with 3D vision, more natural and precise gestures and potentially artificial tactile feedback. Some recent reports even suggest that robotic surgery slightly improves the surgical outcomes [FES+10] compared to conventional laparoscopy. However, due to the lack of long-term data and especially because of its large cost, the number of robotically assisted interventions remain minor for most kinds of procedures.

Variations of laparoscopic surgery are under research. Single port surgery refers to interventions performed with both the laparoscope and the instruments inserted into a single 5 cm incision. Natural Orifice Translumenal Endoscopic Surgery (NOTES) refers to interventions that require no external incisions. Indeed, the laparoscope and the instruments are all inserted through the patient's natural orifices such as the mouth, rectum, urethra or vagina. However, due to their very limited range of action, both of these approaches are more difficult to operate and are less generic than standard multi-port laparoscopic surgery.

Figure 2.4: Da Vinci surgical system. The surgeon sits at a console and remotely drives a robot to perform the operation.

Typical workflow and current standards in laparoscopic surgery

The content of this subsection mostly relies on our surgeons' guidelines provided on the educational site Websurg[1]. The following is a general description of a laparoscopic intervention.

Preparation

After arrival in the OR, the patient is placed on the operating table, lying on his back in order to ease the access to the abdomen (see Fig. 2.5). The angle of inclination of the table may change from one intervention to another, as gravity can make the abdominal organs shift and

[1] www.websurg.com

facilitate their exposure during surgery. The surgeon's placement is also specific to the kind of procedure and he/she may stand between the patient's legs or by one his sides. Likewise, assistants and equipment are positioned in the most ergonomic way in order to maximize safety and comfort.

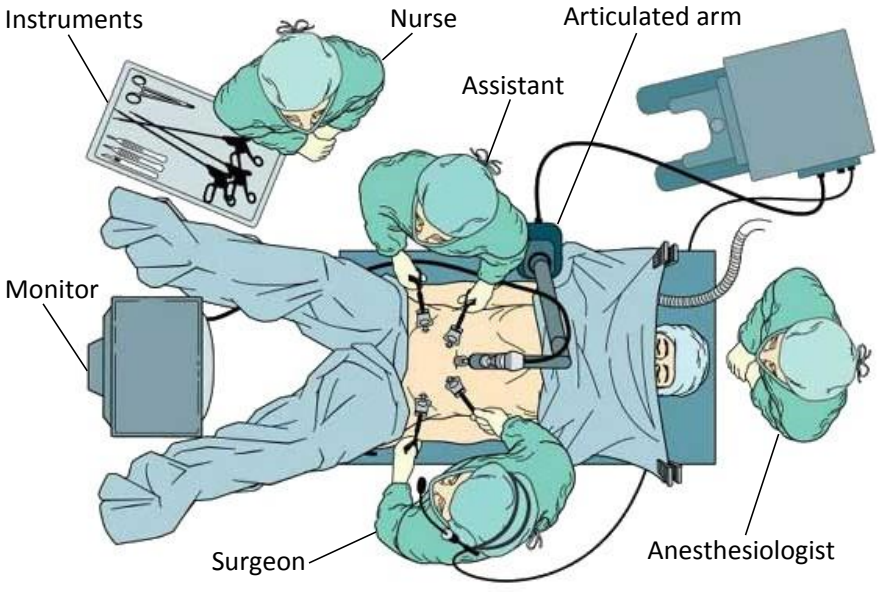


Figure 2.5: Typical OR organization for a laparoscopic intervention.

Incisions, pneumoperitoneum and exploration

Once the general anesthesia is performed, the surgeon operates a small incision of about 10 mm in the abdomen just below the navel, through which he inserts a port called a trocar. A pneumoperitoneum is then achieved by insufflating carbon dioxide into the abdomen, so that the region of interest can easily be visualized and accessed by instruments under the skin.Upon reaching a satisfactory intra-abdominal pressure, the insufflator is then replaced by the laparoscope. The latter consists of a digital camera and a long thin rod containing optical lenses which transmit the view from the surgical site to the imaging sensor (see Fig. 2.6). Some laparoscopes have a 30° angle between their optical axis and their rod, offering a better perspective in some cases. If applicable, a stereo laparoscope may also be used to provide three-dimensional vision to the surgeon.

After inspection of the surgical site, additional incisions are made so that instruments can be introduced. While these incisions may range from 5 to 10 mm, they usually are of 10 mm in order to allow all common instruments in. Since the action range of the instruments is limited by the trocars, their placement is performed so as to allow access to the organs of interest from specific directions (see Fig. 2.7), while preventing any contact. The location of the incisions is typically determined thanks to anatomical landmarks on the abdomen. The surgeon also often performs palpations at the skin surface while looking at the resulting bump in the surgical site with the laparoscope. Between two and five additional trocars may be inserted, their number depends on the kind of intervention.

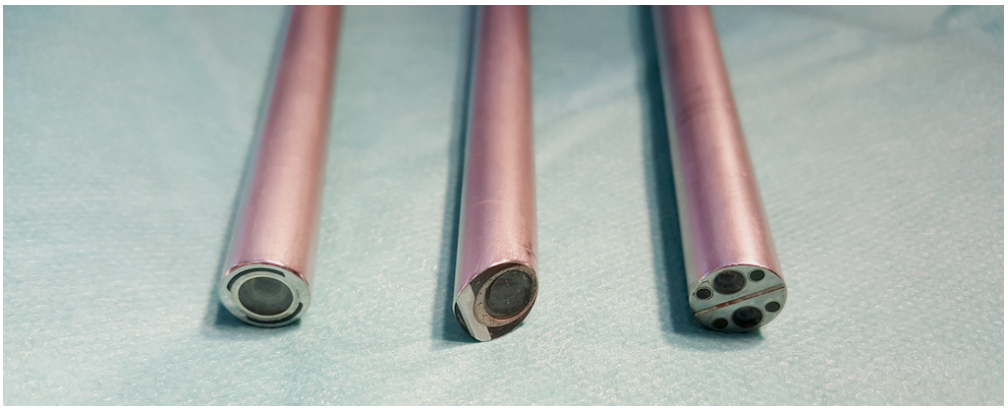


Figure 2.6: Different types of laparoscopes. From left to right: 0° angled, 30° angled and stereo.

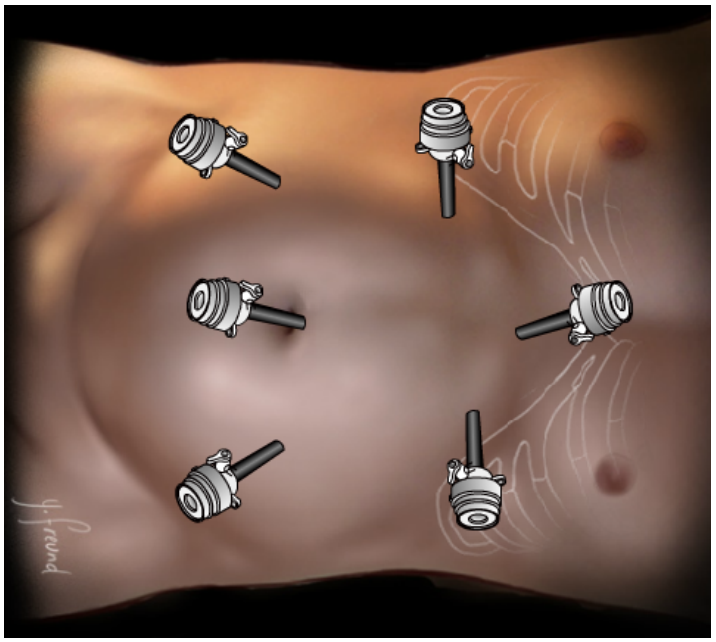


Figure 2.7: Typical port placement for an intervention on the liver, using 10/12 mm trocars.

Instruments

Laparoscopic instruments are various and often just adapted from those used in traditional open surgery. The challenge is designing such tools lies in the compromise between their efficiency and their thinness (see Fig. 2.8). Indeed, they must allow the surgeon to easily and reliably perform many types of actions such as grabbing, cutting, cleaning, suturing, while also being insertable into 10 mm-wide trocars.

Figure 2.8: Common laparoscopic instruments. From left to right: grasper, scissors, dissecting hook, bipolar scissors, needle holder and suction duct.

Performing the intervention

The intervention typically starts by exposing the region of interest, by moving organs using graspers and maintaining them in a particular position thanks to mechanical arms holding the instruments outside the patient. Depending on the type of intervention, the target area may be concealed by tissues or fat that may be cut with diathermy scissors or a dissecting hook. Fluids, especially blood, often appears and occludes some of the surgical site. A suction can then be used to remove this excess of liquid and maintain a clear view. Once the target has been identified and nicely exposed, the surgeon can then perform the goal of the intervention, like a resection for example (see Fig. 2.9). Sometimes, a larger incision and trocar are used in order to pull out resected tissues. An extraction bag is then inserted, deployed, closed and removed through a trocar. Once the procedure is completed, the laparoscope and instruments are removed, the abdomen is deflated and the incisions are closed with stitches, tape or staples. The remainder of carbon dioxide that has not been evacuated is naturally absorbed by tissues.

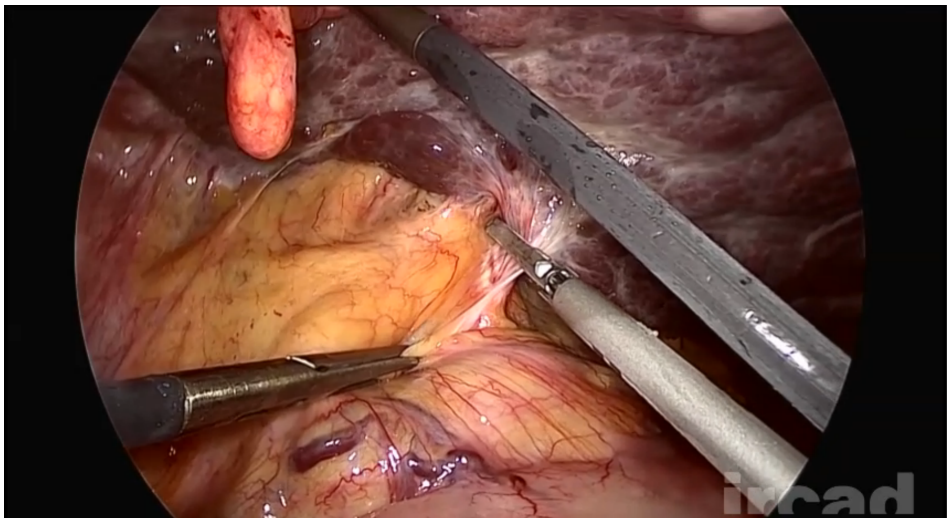


Figure 2.9: Laparoscopic view of the surgeon performing a sectionectomy for hepatocarcinoma on a cirrhotic liver.

The role of medical imaging in laparoscopic surgery
In the previous section, we have seen that the physician's inability of see through tissues has made them resort to medical imaging. In surgery, this is typically used before the operation for planning. However, the patient's anatomy may change between the preoperative and the intraoperative states, as explained in Section 3.4. Therefore, what the surgeon has planned from preoperative images may not fit the scene during the procedure. This led to the adaptation and importation of medical imaging systems into the operating room. As detailed in the next section, most common modalities have been the object of such an integration: radiography and CT, MRI and US.

2.3 Intraoperative imaging

In this section, we describe how the main imaging modalities have been adapted to be used intraoperatively. For each of them, we consider their advantages, drawbacks and typical applications.

Radiography, Fluoroscopy and CT in the OR

Being chronologically the first medical imaging modality, X-ray imaging was also the first to enter the OR, as early as 1955. To this end, the system has been made smaller and mobile in the form of an arch, thereby named a C-arm (see Fig. 2.10). The radiation source is at one end of the arm and the detector, composed of an image intensifier and a camera, at the other end. By sliding on a gantry, a C-arm thus allows to perform acquisitions from nearly any angle. This allows for instance to perform two orthogonal projections and get a better three-dimensional appreciation of the anatomy. Modern machines feature a digital flat-panel detector for an increased spatial resolution and image contrast. C-arms thus provide high-resolution images in real-time and can serve as a surgical guidance tool for many different types of procedure, including general surgery, orthopedics, angiology and cardiology [2,3,4]. Moreover, due to its presence in the OR, C-arms can come in handy for a rapid assessment of a patient's state in case of emergency. Finally, intraoperative radiography may help checking the outcome of a surgical procedure, while still in the OR.

Like regular X-ray systems, C-arms also provide fluoroscopy by performing continuous acquisitions. This mode is particularly useful to obtain real-time guidance and progression feedback in procedures involving motion, such as a catheter placement. However, this results in a high dosage for both the patient and the surgical staff, which remains an issue [Sho96] [MD97] [RFS+00], despite recent innovations [HWT02] [GTB+11] [SHR+13].

At last, volumetric imaging is possible by allowing a C-arm to rotate around the patient while performing successive acquisitions. In standard CT, profiles originating from the fan-beam X-rays are converted thanks to the Radon transform into slices, later concatenated into a volume. With an isocentric rotating C-arm, the cone-beam X-rays produces two-dimensional projections and applying a similar transform thus results directly in the volume image of the region being scanned (see Fig. 2.11). This modality, emerged in the late 1990s [DCF+99], is referred to as **Cone-Beam Computed Tomography (CBCT)** and sometimes as C-arm CT or flat-panel CT [SF08].

[2] `http://www.healthcare.siemens.com/surgical-c-arms-and-navigation/mobile-c-arms`
[3] `http://www3.gehealthcare.com/en/products/categories/surgical_imaging`
[4] `http://www.learningconnection.philips.com/en/catalog/product-group/interventional-x-ray`

Source: Siemens Healthcare

Figure 2.10: C-arm system, allowing X-ray imaging and fluoroscopy for intraoperative/interventional use, by projecting a cone-beam.

Source: Siemens Healthcare

Figure 2.11: The Artis Zeego, a CBCT imaging system by Siemens Healthcare.

Modern CT scanners no longer perform any mechanical motions, but rather use multiple sources and detectors activated in sequence to capture slices of the patient. This results in cumbersome machines that cannot be imported into a hybrid OR. On the contrary, CBCT systems are adapted from C-arms and can fit in a surgical environment. However, since the acquisition then relies on a mechanical rotation, CBCT has a lower temporal resolution

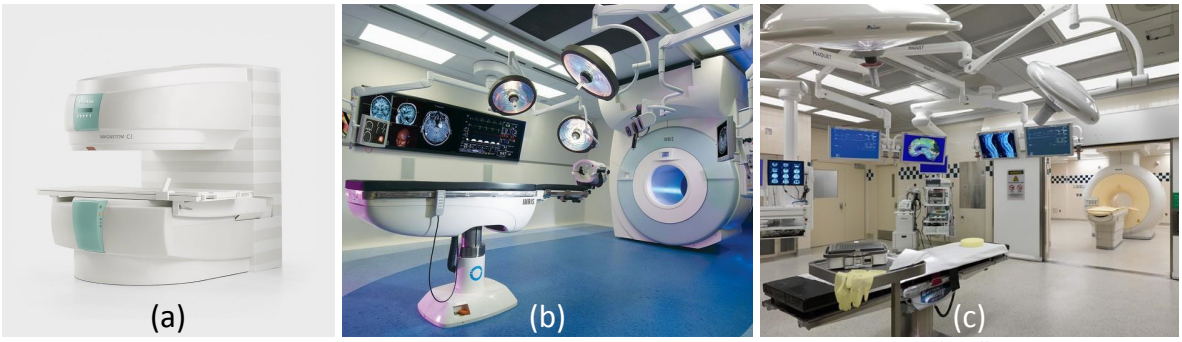
than standard CT. Also, as opposed to fan beams, the use of cone-shaped beams increase the scatter of X-rays, which, in turn, lowers the dynamic range and the Signal to Noise Ratio (SNR) [OWK+08]. The image quality of CBCT is lesser than conventional CT, but it has been shown that the radiation dose received by the patient is drastically reduced [HNY+06][BLM+12]. In depth details about CBCT are provided in [SF08].

Despite its poorer image quality, the ability of CBCT to provide a three-dimensional image of a region of interest after a single rotation convinced surgeons to use it intraoperatively. Applications for CBCT span over various kinds of procedures [WKG+08]. This imaging modality has been successfully applied to interventional neuroradiology [RP09], orthodontics [KRPH05], maxillofacial surgery [PBB+07] and spinal surgery [SMB+05]. CBCT is also a popular guidance tool in radiotherapy [JSWM02] [PBHC+05]. Moreover, it can be used for imaging vascular structures (angiography), potentially enhanced by a contrast agent, which can serve a guidance tool for intra-abdominal operations [NIM+12] [NFKF13] [KWG+14].

However, CBCT is not suitable to image all parts of the anatomy. For example, the skull hinders a good contrast for an image of the brain due to its high absorption of all incoming X-rays. As a result, in neurosurgery, this lack of image guidance has been fulfilled by the introduction of MRI in the OR.

MRI in the OR

As seen in Section 2.1, MRI provides superior contrast in soft tissues without ionizing radiations. Thus, it comes as no surprise that this modality has also been adapted for intraoperative usage. Starting in the 1990s, specific systems and hybrid ORs have been developed to this end. At first, such machines presented a magnetic field of relatively low strength, between 0.2 and 1.0 T, as it resulted in cheaper and smaller systems that could fit inside an OR without a designated suite [BOK+05]. Also, low-field MRI resorts to smaller magnets that open a large access to the patient for the surgeon (see Fig. 2.12 (a)). However, this drop in magnetic field strength involves also a decrease of image quality and lower SNR compared to conventional MRI. Therefore, over the last 15 years, high-field systems have also been developed with magnetic strength ranging from 1.0 up to 3.0 T [NMB+11] (see Fig. 2.12 (b)).



Source: Siemens Healthcare Source: IMRIS Source: Harley Ellis Devereaux

Figure 2.12: MRI system in a hybrid OR. (a) Low-field; (b) High-field ceiling-mounted; (c) High-field in separate room.

Compared to their low-field counterparts, they provide an increased spatial and temporal resolution, allowing even real-time imaging and thus intraoperative fMRI. Nonetheless, stronger magnetic fields require bigger magnets and shielding which make such machines larger and grant little access to the patient. So, depending on the purpose of intraoperative MRI, an equilibrium has to be reached between image quality and acquisition time on one side and cost, practicality and especially the surgeon's access to the patient on the other.

Moreover, as seen in Section 2.2, most standard surgical equipment are ferromagnetic and, as explained in Section 2.1, metallic items represent a hazard near MR systems. However, since the strength of the magnetic field decays with the distance from the machine, it is considered safe to use metallic instruments if submitted to a field of 5 Gauss (0.5 mT) or less. Therefore, every open MRI system has a certain restriction zone around it, called the "5G area". The size of the zone obviously depends on the magnetic strength deployed by the machine, which is why only low-field scanners have a sufficiently small 5G area to be directly located in the OR. Nevertheless, in this case, the patient must be moved in and out of the gantry between intervention and imaging stages and likewise all ferromagnetic instruments have to be inserted and removed each time. A solution might be for the surgeon could to resort to dedicated surgical instruments in titanium or ceramic, which have little to no interaction with the strong magnetic field but these are still not widely available and thus expensive.

For medium- and high-field scanners whose 5G area is too large to fit in the OR, a workaround could be to install the MRI system in a close but separate room next to the OR. However, displacing the patient this far is highly impractical and increases the risk of complications. Another solution may be to perform the reverse and bring the machine to the patient via rails attached to the ceiling, as illustrated in Fig. 2.12-c.

As stated at the end of the previous section, intraoperative MRI is particularly popular in neurosurgery [FZB09] [KtMS+11] [LHH+14] for accurate image guidance despite the infamous "brain shift" [MJHM+98] [CDT+05]. Other applications include spine surgery [TWWW08] and interventional procedures such as biopsies [WNL08], cryo-ablations [TAS+98] and laser therapy [VSZ+04]. Open MRI also offers a good perspective for laparoscopic liver resection [CRS+09].

Despite its superior image quality in soft tissues, intraoperative MRI is not widely used by surgeons because of the prohibitive financial cost of the machine and maintenance alike. As a result, many institutions resort to a much cheaper and less constraining intraoperative imaging modality: ultrasound.

Ultrasound in the OR

Ultrasonography is likely to be the most used intraoperative imaging modality. Indeed, US provides contrasted images in real-time, is safe for both the patient and the surgical team and requires only a simple cheap compact system that can easily fit in an OR. Despite its limited field of view and being prone to noise and shadowing, US is a reliable surgical tool in various situations. It may be used for detecting pathologies, assessing their extent, but may also serve to guide for needles for biopsies and instruments for resections. If set in Doppler mode, US imaging can even locate millimetric vessels. Therefore, ultrasonography is used intraoperatively in neurosurgery [URS+06] [MHT+07], cardiology (also known as echocardiography) [HLF+07] [KVNL07], urology [UOK+08] (with transrectal probes) and also for abdominal procedures [CJDL03] [GH11] [LVR+12].

In laparoscopic surgery, special US transducers have been developed so as to be insertable in 10 mm trocars. They are shaped as a thin long rods, the sensor being located at the distal tip (see Fig. 2.13). The shaft of intraoperative US probes is commonly flexible in order to allow the surgeon to maintain a full contact with the surface of the target, as the placement is limited by the trocars.

Figure 2.13: Laparoscopic US transducer with its flexible tip.

Since 3D laparoscopic transducers are only at the stage of prototypes, intraoperative US is limited to 2D imaging. However, a 3D reconstruction of an area is possible by resorting to freehand scanning with tracking of the transducer [SLT+07].

2.4 Conclusion

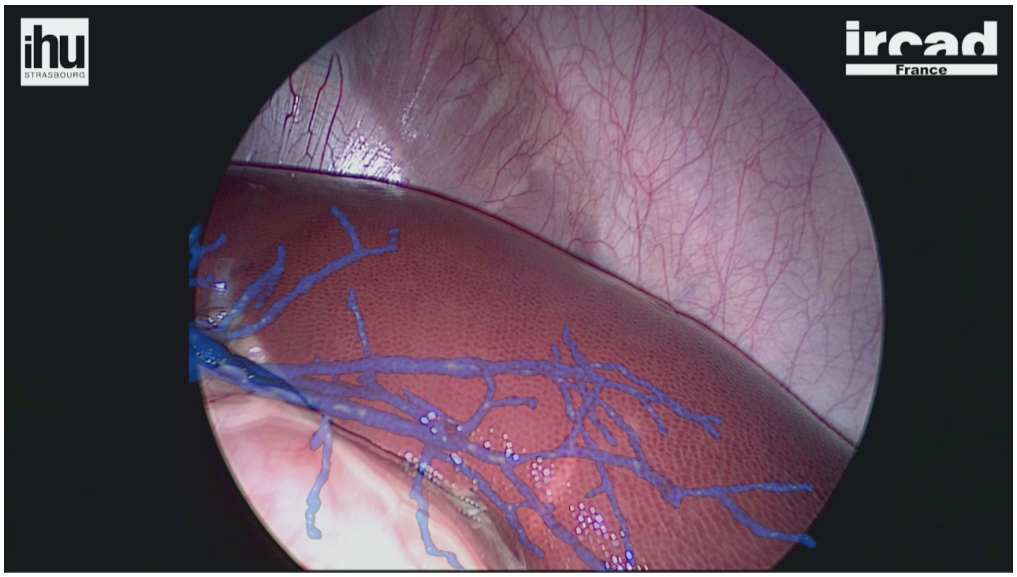
Imaging technology has revolutionized medicine over the last century. Granting physicians the ability to observe their patients' anatomy and pathologies have tremendously increased the accuracy of diagnosis and the potency of therapy. In surgery, imaging is an irreplaceable tool for thoroughly planning the intervention and thereby boosting the success rate. Moreover, most imaging modalities can also be used intraoperatively, whereby they may help to assess the progression of a procedure, guide the instruments or detect critical complications. Also, a preoperative acquisition may not represent accurately the anatomy of a patient during an intervention, due to a different positioning or simply surgical actions. Intraoperative imaging then helps updating the preoperative information and compensating for any deformation. This introduces the need of multi-modal image fusion, whereby acquisitions from different systems and times shall be aligned in a same referential frame. This task is known as registration. However, these images are often displayed on separate monitors in the OR and require the surgeon to mentally align them with his view on the surgical site, which is tiresome and inaccurate. In [NIM+12], the author concludes by stating the need for an automation of this task through a direct augmentation of the surgeon's view with the medical images.

This technology, called Augmented Reality (AR), is suitable for intra-abdominal keyhole surgery. As many other types of minimally invasive procedures, this one presents many advantages over traditional open surgery and has become widely accepted as a safe and reliable alternative. The surgeon then resorts to laparoscopy via an endoscopic camera and a digital display to see the surgical site. This offers a great opportunity to merge medical imaging with the surgeon's view, thus creating AR. The goal of the next Chapter is, through a literature review, to explain what are the challenges of AR in laparoscopy and to reveal the state-of-the-art for this topic.

Chapter 3

Augmented Reality in Laparoscopic Surgery

The topic of this thesis revolves around the static augmentation of laparoscopic images with information from intraoperative 3D data. In order to position our work with respect to the state-of-the-art, this chapter provides a comprehensive review of AR in laparoscopy. It begins with a brief historical summary of surgical AR, which helps settling the used terminology along with a few core definitions. Then, for the sake of being self-contained, we recall what are the common inputs and outputs of surgical AR as well as the challenges induced thereby, preventing its wide adoption. Finally, through a coherent classification, we review the current methods of laparoscopic AR and expose their limitations which, once aggregated, better outline the objectives of our work.



Source: IRCAD France / IHU Strasbourg

3.1 Review strategy

Our project, including this thesis, tackles AR in laparoscopic surgery and its application for abdominal interventions, with a slight focus on the liver. As a result, the following literature review will primarily deal with this topic, but we also included reports of laparoscopic AR on a different intra-abdominal organ. Indeed, these exceptions help better understand the main achievements and the realm of possibilities in this field. Overall, this chapter concerns all AR methods that have or could be applied to laparoscopic interventions on the liver.

Our focus also induces that we will consider only non-flexible straight endoscopes, monocular or binocular (stereoscopes). Since we mainly work with 0° optics, the latter will receive a particular attention, but oblique ones will also be considered. As for the mean of AR display, we will also focus on monitor augmentation as this is clearly the current trend, as well as ours. Finally, surgical AR applications dedicated to education and simulation will also be discarded as their goals and challenges differ from those of real interventions.

Despite these restrictions, the volume of related publications produced during the last decades remains large. For the sake of clarity, the following literature review will not be exhaustive but rather comprehensive, in an attempt to provide a thorough yet graspable state-of-the-art of AR in laparoscopy. It will include review papers, milestone reports, significant works absent from reviews and recent articles. For a broader perspective, a list of reviews about surgical image guidance is provided in Appendix A.

3.2 Purpose of Augmented Reality in MIS

Terminology of Mixed Reality technologies

Before delving into the subject, we believe it is pertinent to define what is augmented reality in distinction from similar technologies, thus lifting ambiguities. In 1994, a taxonomy of "Mixed Reality" technologies was proposed by [MK94], stating AR is part of the "reality-virtuality continuum", alongside Augmented Virtuality (AV) and Virtual Reality (VR). The concept of continuum is sound, since telling each technology apart may be sometimes difficult.

In this context, a real scene can be defined as the result of an optical projection of the real world onto an imaging sensor. On the contrary, a virtual scene is a purely digital environment which does not result from an optical acquisition. Then, AR can be roughly defined as a real scene to which one or several virtual elements are added. AV is a virtual scene tempered with measurements of the real world, such as an object tracking for example. VR is a purely virtual scene with no bounds to the real world. We believe a general rule of thumb could be the following: the augmentee occupies most of the scene and is behind the augmentation in terms of overlay. Fig. 3.1 illustrates these three concepts through a same scene.

In surgery, resorting to AR and/or AV depends on the kind of intervention considered. For example, since AV provides a better appreciation of depth, it is preferred for needle insertion guidance. In MIS, AR is predominant as the endoscopic view remains a standard for surgical guidance. Introducing AV in this case would require the surgeon to look at separate screens with two different points of view and hence would hinder his/her hand-eye coordination. Nonetheless, since depth perception is a problem in AR for MIS (as discussed in Section 8.3), AV may be complementary to AR for this kind of intervention. The choice between good hand-eye coordination and enhanced depth perception lies in the surgeon's discretion.

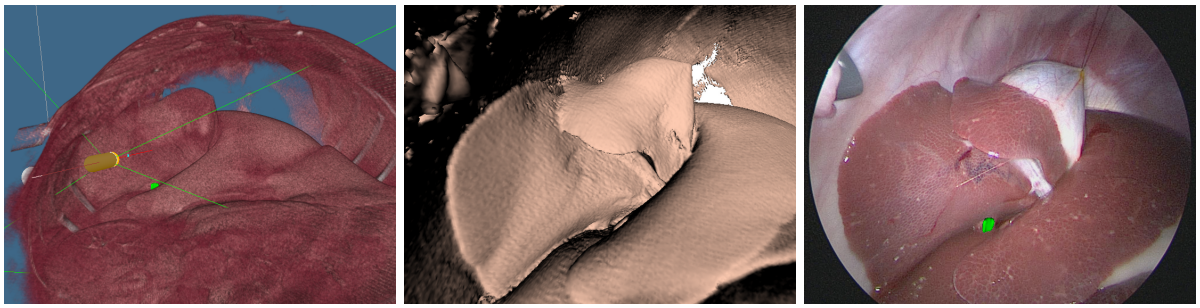


Figure 3.1: On the left, an example of AV showing the abdominal cavity displayed using volume rendering and augmented by a virtual endoscope and a segmented tumor (green). In the middle, an example of VR where the scene is rendered as a surface and observed from the actual point of view of the endoscope. On the right, an example of AR where the laparoscopic image (the reality) is augmented by the tumor.

Advantages of AR in MIS

The super power of seeing through solid objects is an old human fantasy that has been around for quite some time [Sup39]. Surgeons are far from exempt of desire for this ability. Indeed, we previously concluded the Chapter 2 by a clear need for seeing through tissues in laparoscopic surgery, to which AR might very well be the answer. Here are the main advantages of AR in laparoscopic surgery as a guidance tool:

- Laparoscopic AR may ease intraoperative planning by providing a rapid identification of subsurface targets (e.g. tumors, infections, foreign bodies) and critical structures (e.g. vessels, nerves, other organs). For example, this could prevent blood loss by unwanted perforation of unsuspected vessels (especially veins which have a weaker wall than arteries). Furthermore, faster intraoperative planning would shorten the intervention time, hence reducing the risk of complications.

- Laparoscopic AR would also spare the surgeon from having to mentally match information (mostly images) from different sources to the scene. By reducing the surgeons cognitive load during the intervention, laparoscopic AR should increase his/her focus and efficiency.

- Laparoscopic AR may guide resections by displaying cutting trajectories and margins planned beforehand on a virtual model. More accurate resections lead to negative margins (all cancerous cells removed) while sparing more healthy tissue, thus preserving more the functionality of the concerned organ and shortening the patient's recovery. Herrell et al. proved on a phantom that AR yields better resection ratios in robotic laparoscopic nephrectomy [HKMG09].

- Due to its diameter, the field of view of a laparoscope is narrow (typically 70° compared to 160° for humans). AR assistance in laparoscopy may prevent the surgeon's disorientation by lifting anatomical ambiguities and thus increase his spatial awareness. This is especially true for NOTES and arthroscopy where vision is also very short in depth and impedes landmark identification [VE06][ARS+12].

Pioneers in laparoscopic AR

The first intervention assisted by AR was performed in neurosurgery [RSH+86]. Ensued many reports of interventions again in neurosurgery [Pet06] [MPHH+15] [CSB15], but also in otolaryngology (also known as ENT for Ear, Nose, Throat) [WKS+11] [MUS+13] [DBWC14] [LRS+15], in maxillofacial surgery [ZSR+13] and in orthopedics [WCD+06]. The rapid development of AR assistance in these specific surgical fields is explained by the rigidity or semi-rigidity of the structures of interest thanks to their proximity to bones, which limits deformations between the preoperative acquisition/planning and the intraoperative state [Shu04] [MJB13] (see Section 3.4).

On the contrary, in digestive surgery, organs do not have a constant spatial relationship with rigid structures, hence significant deformations occur. As a result, AR-assistance has been little used in laparoscopic surgery compared to other surgical fields and only a few laparoscopic interventions were reported on humans. One of the first interventions assisted by AR in general surgery, an adrenalectomy, was performed at our institution in 2004 [MRA+04]. We have extended our experience since then [NDA+13][MPS+13][PDS+15]. In another group, Konishi et al. also reports 20 laparoscopic interventions assisted by AR between 2003 and 2004 [KHN+05]. The first AR-assisted intervention reported in urology was performed in 2006 on the kidney [UG08]. The same team successfully employed laparoscopic AR in 25 interventions [UNS+10]. In 2009, Teber et al. reported 10 successful laparoscopic partial nephrectomies under semi-automatic AR guidance, with no positive margins and no complications [TGS+09]. In 2010, Nakamura et al. reported two successful renal vessel treatments by novice surgeons, using a manual augmentation of the endoscopic view with volume-rendered preoperative CT data [NNZ+10].

Besides their pioneering value, all these reported interventions are only research works, essentially prototypes, and have not yet been transformed into a commercial product for widespread usage. Moreover, all augmentations were performed manually or semi-automatically, which requires extensive interaction from the surgeon. Finally, the non-rigidity of the surgical scenes were never taken into account, thus limiting the accuracy of the augmentations and, ultimately, the validation of the methods.

3.3 Means of Laparoscopic Augmented Reality

Typical data available for laparoscopic AR

Various data are available to augment the real scene, as listed below:

- **Preoperative data**. Usually, such data originate from conventional scanners (CT or MRI) and present a good resolution and image quality. Since they are acquired prior to the intervention, they are often digitally processed, manually or automatically, to enhance and/or isolate by segmentation structures of interest such as organs, tumors, bones or vessels. The vascular network can also be artificially emphasized by injecting a contrast agent. In this case, after the patient has been injected, two acquisitions are performed: one for the veinous phase and one for the arterial phase. Thus, both veins and arteries appear more contrasted in the image and are easier to segment. Classic methods of segmentation include thresholding, region growing, active contours, clustering and classifiers, as reviewed in [PXP00]. More advanced methods are Markov Random

Fields, Artificial Neural Network and Statistical Shape Models [HM09]. Based on a challenge workshop of the conference MICCAI 2007, [HVGS+09] provides comparison of various automatic and semi-automatic liver segmentation methods.

- **Intraoperative data** (non-optical). This kind of data is produced by imaging systems inside the OR. In a hybrid OR, these can be flat panel CBCT or open MR scanners. As explained in Section 2.3, the relative compactness and reduced invasiveness of such intraoperative scanners make them produce images with worse resolution and image quality than their conventional counterparts. Another popular intraoperative modality is ultrasound, but it has a small area of capture and is highly subject to speckle. The main advantage of intraoperative data is the fact that their representation of the patient's anatomy is much more up-to-date and hence more accurate than that of preoperative images. However, the poorer quality of intraoperative images often make them require some online digital image processing, as long as the latter does not significantly extend the intervention. Similarly to preoperative imaging, contrast agents may also used to increase the contrast of relevant structures. In a nutshell, if the Structures of Interest (SOI) are visible in the intraoperative data, intraoperative imaging may suffice for surgical guidance. In the opposite case, it can be viewed as a bridging modality between preoperative data and the endoscopic video.

- **Endoscopic image**. Current laparoscopic cameras commonly offer full HD vision (1080p) of the scene at a rate of 60 frames per second (fps). Some products provide 2K and even 4K resolutions (respectively 1440p and 2160p). Despite its incapacity to see beneath tissues, this modality can be self-augmenting. Although less common, this kind of augmentation results from real-time and automatic image analysis. For example, Amir et al. propose to detect subtle pulsatile motions at the surface of tissues to reveal hidden vasculature [AKHP+15].

- **Tracking data**. contrary to all the other items in this list, tracking data do not consist in images, but rather geometrical measurements in the world space. They are typically produced by tracking systems dedicated to locate specific objects in world space. These targets may be the patient's anatomy, surgical instruments or imaging systems (C-arm, laparoscopic camera or laparoscopic US transducer). They are tracked via attached markers detected by sensors. The two most popular tracking technologies are optical and magnetic (more details in Appendix B). Unlike for AV, tracking does not directly augment the surgical scene for AR, but instead aims at updating already placed augmentations with respect to the dynamics of the scene. However, such systems often necessitate extensive calibration procedures in order to meet accuracy requirements.

- **Planning data**. Such data are the only ones which commonly do not originate from machines, but from the surgeon. Usually, they consist of free-hand or assisted annotations used for intraoperative guidance. Popular annotations include labels, measurements or geometrical figures such as cutting lines and resection margins. Planning data can even be created during the intervention using telestration (drawing on live video) [ALF+08] and updated in real-time using image analysis to successfully follow a dynamic scene [KBCH12].

The quality of laparoscopic AR undeniably depends on that of the data it is based on. However, this quality is only passed onto the surgeon's eye through a good rendering. The next subsection describes the main two rendering techniques for volume visualization and what kind of content is appropriate for each of them.

Graphical rendering forms for laparoscopic AR

Augmentations can take several rendering forms, but since these mostly originate from 3D data, they are dominated by two [BP11]: Indirect Volume Rendering (IVR), more commonly known as surface rendering, and Direct Volume Rendering (DVR), often simply shortened to volume rendering.

- **Surface rendering**. This kind of rendering displays a surface representing the interface between two separate structures (lung/air, vessel/lumen or liver/fat for example). This commonly results from a segmentation, manual or semi-automatic, based on the gradient in voxel intensity. Fully automatic segmentation can be achieved by considering simple iso-value approaches such as marching cubes or more advanced methods such as statistical atlases, geometrical priors and random forest, but the result is not always satisfactory. Hence, due to the time requirement, this tedious task of interactive delineation is typically performed on preoperative data prior to the intervention rather than during its course. Surfaces can be rendered as a mesh composed of polygons and if important polygons are hidden, as wireframe (see Fig. 3.2 (a) & (b)). Current surface rendering techniques offer lighting effects (shading) based on the topology to enhance realism and ease interpretation. In case of poor data resolution, surfaces can also be smoothened by interpolation algorithms in order to avoid "staircasing" artifacts (see Fig. 3.2 (c)). Contrary to volume rendering, surface rendering does not display all the 3D data and thus is much less computationally demanding. Moreover, given the inherent clear delineation between structures, it is easier to interpret. Therefore, if available, surface rendering is the visualization of choice for most mixed reality applications in surgery, according to the literature [KOJC13].

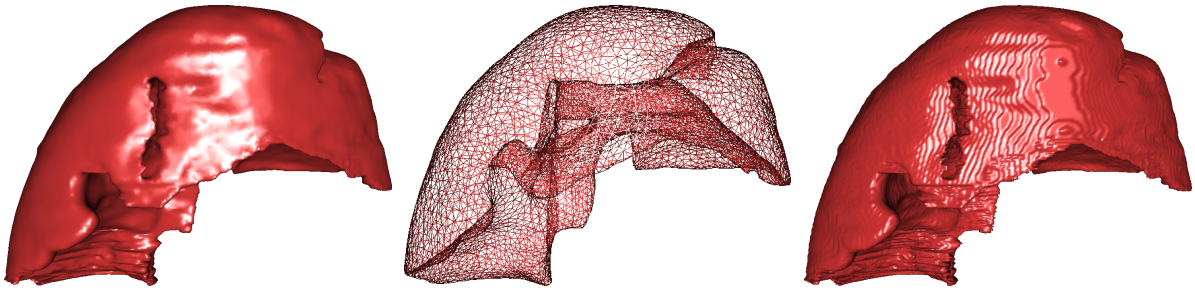
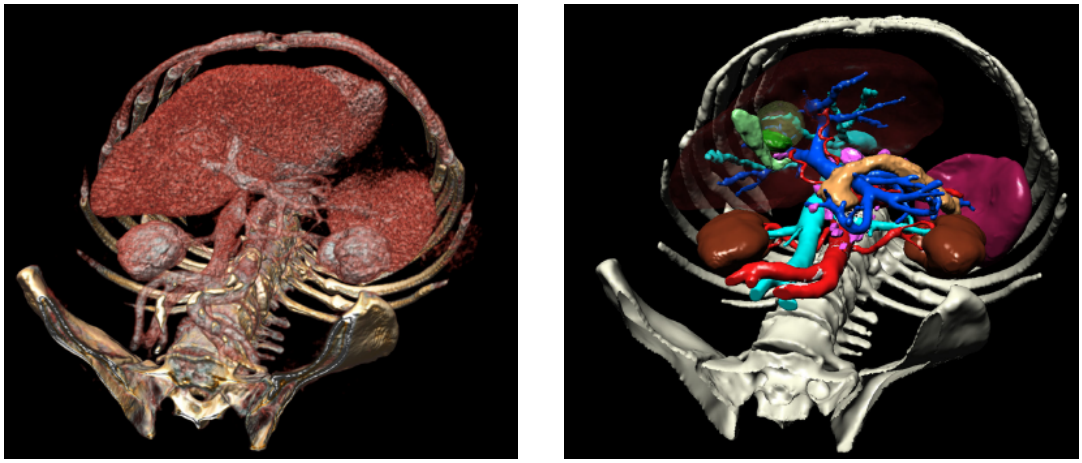


Figure 3.2: Surface rendering of a segmented liver. (a) Smooth surface with shading; (b) Wireframe visualization; (c) Staircasing artifacts.

- **Volume rendering**. Contrary to surface rendering, volume rendering visualizes all the data. In order to create the image, the most popular technique is ray casting, which relies on virtual rays traversing the 3D data. The intensity of the corresponding pixel on screen is given by a weighted summation of the color and opacity associated with

every voxel hit along the ray. The color and opacity a each voxel is based on their intensity value and determined by a predefined Transfer Function (TF). Therefore, volume rendering does not require the user to interact with the data to display structures, which makes it particularly suitable for the visualization of intraoperative 3D data. However, the absence of interaction also prevents from simulating resection and the absence of delineation removes the possibility of measuring volumes. Moreover, since all the 3D data is displayed, volume rendering is particularly computationally demanding. Finally, much like the quality of surface rendering depends on that of the segmentation, the quality of volume rendering is strongly related to the design of the TF (see Fig. 3.3). A TF being defined before the data acquisition, it may not be perfectly tailored and structures are sometimes hard to properly distinguish, even using contrast agent. Thus, adaptive initialization and refining of TFs is a popular topic of research [SG09] [CM11] [RBB+11]. One solution may lie in designing TFs with multiple dimensions: voxels are not just considered for their intensity, but also for their gradient or for their neighborhood [LLY06]. The reader can find more details about volume rendering in [ZEP11].



Source: [SNP+14]

Figure 3.3: On the left, volume rendering of an intra-abdominal cavity. On the right, the corresponding surface rendering.

One exception to surface and volume rendering in laparoscopic AR is the display of 2D US images in the surgical scene [KOW+14] [ONN+14] [PJHH+15]. However, this kind of augmentation is restrained to US-based navigated surgery and thus is out of the scope of our project. Therefore, there are only two main rendering techniques in laparoscopic AR. Concerning volume visualization, based on their advantages and limitations, surface rendering seems to be the approach of choice for preoperative 3D data, while volume rendering may be more appropriate for intraoperative 3D data. After the choice of rendering comes that of how the augmentation is displayed to the surgeon. Various technologies are available and are reviewed in the next subsection.

Display technologies of laparoscopic AR

Digital display technologies have multiplied for the past decades. Many of them have been adapted to enter the OR and have even been used for AR purposes, as listed below:

- **Projection onto patient**. This form of display makes AR the closest to its definition, in that this is the actual reality that is augmented. Since holograms are still very much in their infancy, the only current way to augment the actual reality is by projection. In surgery, the patient then becomes the display, and anatomical images can be projected onto his abdomen [NBG+08] [VPB+11] [WWL+14]. Resorting to such a display for AR has proven to be useful for port placement [SYK+10] or needle insertion [OGND10]. However, the surgeon's point of view differs from that of the projector and this discrepancy in perspective limits the AR accuracy. In laparoscopic surgery, the pneumoperitoneum worsens this effect by separating further the abdominal surface (the screen) and the organs (the scene) and therefore such a way of displaying AR is not recommended for this kind of surgery.

- **Optical see-through**. Sometimes referred to as "AR windows", this kind of display refers to a projection of the augmentation on semi-transparent surfaces like half-silvered mirrors, placed in front of the scene. In this case, the problem of perspective faced by directly projected AR is alleviated since the surgeon can place himself in alignment with the augmentation without occluding the projection. Several groups reported the use of semi-transparent mirrors for surgical AR [FDM+05] [MSLD08] [WMF+11]. However, much like AR projected onto the patient, this approach is geared towards instrument insertion and not suited for abdominal MIS. A variant of this technology makes use of integral videography for autostereoscopy [LISD10].

- **Video see-through**. This concerns all digital displays of a live video of the surgical scene:

  - **Head-Mounted Display** (HMD). Mounting a video headset on the surgeon's head was popular in research at the dawn of surgical AR [FLR+98] [BFM+03] [VKS06]. Since then, this approach has faded due to lack of comfort (heavy headsets), of graphics realism and, in the case of laparoscopy, unclear usefulness for the surgeon. Nonetheless, a recent regain of interest for VR headsets has occurred, mainly from the video game industry [1]. This resurgence of VR headsets may lead to lighter devices providing high quality video and thus a possible comeback in the OR. HMD may also be used for an optical see-through technology like the Google Glass project [MLZ+14].
  - **Hand-held display**. In merely a decade, tremendous technological advances in digital mobility have brought smartphones and tablets in every household. This display technology is now also being explored for surgery [SNP+14] [KWN+15], but is still at an early stage.
  - **Monitor**. This is the oldest form of digital display and still the most common used one. Most of AR approaches in MIS concern virtual data being overlayed on

[1] Various products are being developed by most major tech companies: Facebook-owned Oculus Rift, Valve-backed HTC Vive, Samsung Gear VR, Sony Morpheus and Microsoft Hololens

an endoscopic video displayed on a large monitor in the OR [BFMR08]. Recent monitors offer 4K resolution and 3D vision on Full HD display.

Various forms of AR display has been investigated for surgery. For more details about this topic, the reader can resort to [SFN08] [NSMM11]. As a side note, it is also worth mentioning that numerous software tools dedicated to or compatible with surgical AR are available, as listed in Appendix D.

Overall, there are many options to consider for the design of a surgical AR solution. However, not all choices are equivalent and an informed selection of data, rendering and display is paramount to the success of the AR method. Once these choices are made, there still remains the main challenge of surgical AR, which is an accurate registration between the various data, as explained in the next section. However, now that the core topic of this thesis is being touched, more specificity is sought and the focus is set onto laparoscopic AR.

3.4 Challenges of Laparoscopic Augmented Reality

This section explains what are the different challenges to achieve laparoscopic AR. The main challenge is registration accuracy between the imaging modalities in a static environment. A secondary challenge is then intraoperative dynamics. Other minor challenges are also mentioned.

The challenge of static registration

The most important criterion in surgical AR is accuracy. The margins depends on the intervention, but are strict for an AR application to enter the OR. The accuracy is ensured by a valid registration between the different imaging modalities. A registration consists in determining every degree of freedom (DoF) that parametrizes one model to match another one. As illustrated in Fig. 3.4, in laparoscopic surgery, the augmentation is usually a 3D model from preoperative data and the augmentee is commonly the laparoscopic view, as the surgeon's only view on the scene. Unless performed manually by an expert, registering these imaging modalities is far from being straight-forward because of two factors:

- The **deep difference between imaging modalities**. Preoperative data are three-dimensional and composed of voxels with a single intensity value, whereas the laparoscopic image is two-dimensional and composed of pixels with three values (RGB). These fundamental differences induce a profound lack of correspondences to work on. Therefore, a robust approach is to divide the registration into two different tasks (see Fig. 3.4). The first one is to register the preoperative data to the patient's anatomy in a referential such as the operating table or an intraoperative scanner. In case of a rigid registration, this consists in determining a rotation and a translation in world space, for a total of 6 DoFs. The second one is to localize the laparoscope in that same intraoperative referential. This pose determination is also equivalent to resolve a rotation and a translation, for another 6 DoFs and a total of 12. If non-rigid registration is considered for the first task, the total number of DoFs can then significantly build up, sometimes as far as thousands.

- The **anatomical deformation between acquisition times**. Between the time of the preoperative acquisition and that of the intervention, the patient's anatomy has been significantly deformed. First, his/her body is positioned differently either intentionally

as a requirement of surgery or non-intentionally due to subtle differences in posture comfort.

Second, the pneumoperitoneum introduces more pressure in the abdominal cavity that changes its mechanical equilibrium, making structures and organs shift and deform [SMMCL+11]. For the liver, the literature reports up to 28 mm of shift [ZLH+12]. For the kidney, similar shift of 46 mm has been observed [SNL+13].

If not taken into account, such deformations result in registration errors whose magnitude are not suitable for MIS, essentially because they are greater than standard resection margins (typically 5 mm according to our surgeons).

Despite these deformations, rigid registrations have been employed in laparoscopic AR [MRA+04] [KHN+05] [MLT+04] [NDA+13], but such proofs of concept cannot fulfill the accuracy requirement for surgical purposes. Moreover, not all organs are equal when it comes to deformation. In MIS, the kidney may shift to the same extent than the liver, but does not deform as much. This may explain the popularity of laparoscopic AR for renal interventions as a research topic [SVA+09] [BFMR08] [TGS+09]. For more details, we will review the various current means of overcoming the challenge of registration in Section 3.5.

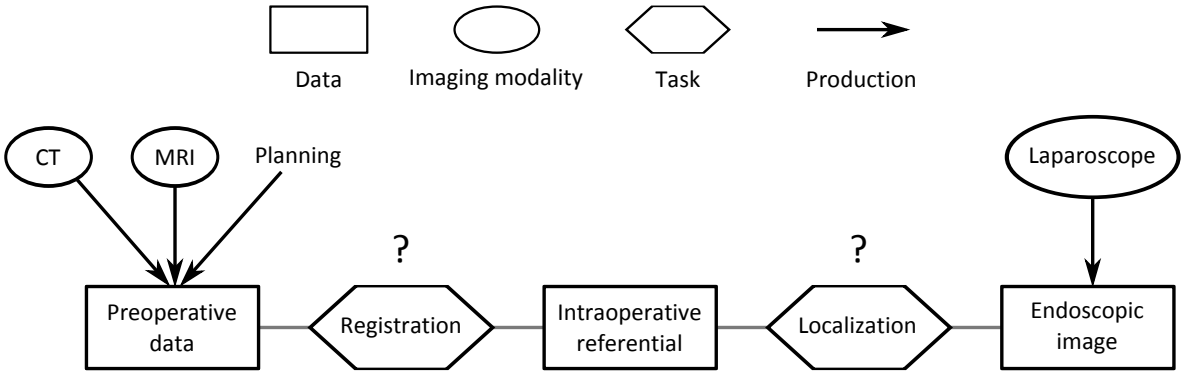


Figure 3.4: Challenges of laparoscopic AR. This diagram presents a generic AR pipeline in surgery. The first task to achieve is the registration between preoperative data and a chosen intraoperative referential. The second task is the localization of the laparoscope in that same referential.

The challenge of intraoperative dynamics

Reaching a satisfactory accuracy in laparoscopic AR is a major challenge. Maintaining it in real-time is another one. Indeed, the vision upon a surgical scene is highly dynamic due to various factors.

Besides camera and instruments motion, the surgical actions of mobilizing and resecting tissues introduce significant and sometimes permanent intraoperative deformations. Across 11 patients who underwent hepatic resection, Heizmann et al. observed liver deformations by more than 20 mm and volume variations from -13% to 24% [HZB+10].

Another dynamic factor is the patient's metabolism, which generates natural motions such as breathing and heart beat. Song et al. measured between 7.8 and 22.5 mm of respiration-induced displacement of the liver across 10 healthy humans [STJ+11]. Pneumoperitoneum-induced shift and breathing combined, the discrepancy can rise up to 44 mm [VRH+14].

However, both of these metabolic phenomena are periodical and can be estimated and compensated for [MVFCM03] [MY10], since respiratory motions are mechanically controlled.

Keeping the augmentation accurate requires to constantly update the registration, through a complete or partial reevaluation of all the DoFs. Moreover, most of the proposed methods in laparoscopic AR are computationally expensive. Only a few show true real-time capabilities, satisfying two essential criteria: no latency (up to 300 ms [MLG+01]) and a good frame rate (at least 10 fps, ideally 25 for continuous motion perception). Fortunately, advancements in parallel computing on CPU and GPU, combined with constant hardware improvements, steadily alleviate this issue. Overcoming the challenge of dynamics is out of the scope of this thesis but, for the sake of completeness, this subject is still explored in Appendix C.

Other challenges
Laparoscopic AR has to face other lesser yet important challenges. These are often overlooked, despite their paramount importance in the success of the AR method. The first one is the **robustness** of the laparoscopic AR approach and the lack of validation. A proposed method ought to be reproducible and show consistency in good results. Another challenge is the **relevance of visualization**. Any laparoscopic AR method shall guarantee that the scene is always augmented, not diminished. In other words, the surgical value of the visual information added must be greater than that of what it occludes. Finally, the developed technology has to be adapted in order to seamlessly integrate into the OR, which requires a solid cross-disciplinary collaboration. These supplementary challenges are discussed in the conclusion of this chapter.

3.5 Current Registration Methods for Laparoscopic AR

After more than a decade of research, proposed methods of laparoscopic AR can be categorized in two main approaches, in addition to the manual approach, as illustrated by Fig. 3.5.

The first approach (green) is **interactive**, as it mostly relies on manual inputs from an expert to register preoperative data to the laparoscopic image (task {0}).

The second approach (blue) is **surface-based**. It aims at recovering the intraoperative state of the scene via a surface reconstruction commonly based on laparoscopic images (task {1}). Then, the surface is registered rigidly or non-rigidly to the corresponding surface extracted from the preoperative 3D data (task {2}).

The third approach (orange) is **volume-based**. It performs a similar process than the first approach, except that the intraoperative state is acquired thanks to an intraoperative imaging modality such as a CBCT, open MRI or US. This approach requires two registrations: one between the intraoperative volume image and the preoperative 3D data (task {3}) and one between the laparoscope and the intraoperative volume image (task {4}). A variant of this approach consists in using only the intraoperative 3D data for the augmentation, provided that the relevant structures are visible thanks to contrast agent and/or a good TF. In this case, only task {4} has to be performed, as in [FMHN08] and in our project.

Each of these three approaches is reviewed into details in the next three subsections, with attention to expose their limitations.

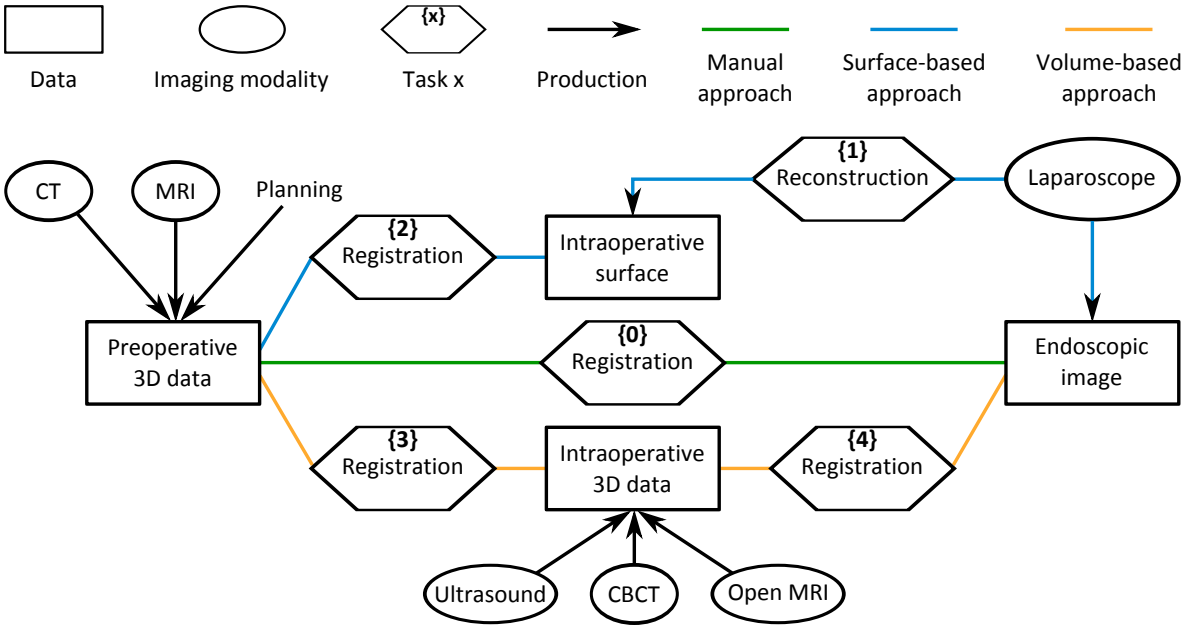


Figure 3.5: Current approaches of laparoscopic AR. The various tasks from the different approaches are numbered for easier reference throughout the review.

Interactive Approaches to Laparoscopic AR

As seen in Section 3.4, one of the main challenges of laparoscopic AR is the lack of cross-modality correspondences between the preoperative model and the laparoscopic image. Therefore, human interaction is often necessary in the registration process to ensure its accuracy. For example, a rigid registration may be initialized fully manually and updated with respect to camera motion. The latter can be determined via a dedicated tracking system (more details in Appendix B) or via the tracking of anatomical landmarks selected manually [NDA+13] or automatically [PSCM+14]. However, this manual initialization requires the operator to determine the correct values of six DoFs, which may be a lengthy task depending on his/her level of expertise and disturb the surgical intervention.

In an effort to accelerate this manual process, Pratt et al. propose to use stereovision to register a preoperative model of a kidney with its projection in both views [PMV+12]. After manually selecting a same landmark in one view and on the model, its 3D position in the scene is triangulated thanks to its correspondence in the other view. Then, the remaining rotation (three DoFs) is set by the operator using a trackball. The whole process reportedly lasts between 40 to 60 seconds.

In conclusion, while the manual process of interactive approaches may negatively affect the surgical workflow, the magnitude of its impact varies with the degree of interaction (fully manual or semi-automatic), the quality of its user interface (UI) and the operator's degree of expertise. A noteworthy advantage is their relative facility to get certified as a clinical product, especially if compared to automatic approaches. The latter are described next, starting with the methods based on surface extraction and followed by those relying on intraoperative volume acquisition.

Surface-based Approach

Surface-based methods of laparoscopic AR have a similar process:

1. Extraction of the surface of the organ of interest from the preoperative 3D data. This is performed before the intervention using manual or semi-automated segmentation tools.
2. Dense or sparse reconstruction of the surface of the laparoscopic scene once the organ of interest is partially visible (task {1}).
3. Initial rigid registration of the two surfaces (task {2.1}), performed manually, semi-automatically or fully automatically.
4. Automatic registration refinement (task {2.2}) using a rigid model or non-rigid model.

Surface Reconstruction in Laparoscopy

Surface reconstruction from image analysis is a well-known and well-documented topic in computer vision. A wide variety of methods have been developed over the past 30 years. Despite the particularity of laparoscopic scenes compared to others, most of these methods have been applied to such images with more or less success. Due to the large volume of literature about surface reconstruction from endoscopic images and since the work presented in this thesis does not rely on a surface-based approach, this subsection only briefly reviews the various methods. For more details, the author encourages to refer to several recent review papers on this topic. For a good introduction to the matter, Mountney et al. provide a tutorial paper about 3D tissue reconstruction and tracking in [MY10]. For a deeper understanding, Maier-Hein et al. reviewed all the state-of-the-art reconstruction algorithms based on optics in [MHMB$^+$13]. Leaning more towards hardware, Geng et al. consider the various endoscopic sensors used for surface reconstruction in [GX14]. By the time of writing this thesis, the latest survey article about endoscopic surface reconstruction and laparoscope localization is [LSQ$^+$15].

In the realm of endoscopic surface reconstruction (task {1} in Fig. 3.5), methods are of two kinds: passive and active.

Passive techniques
This kind of technique relies only on the analysis of visual cues in the endoscopic image to deduce the topology of the scene and sometimes the egomotion of the visual sensor all together.

- **Structure-from-Motion** (SfM). Such methods match elements across successive frames and infer their depth according to their displacement. SfM is compatible with any endoscope, but cannot perform without motion and can be computationally demanding for dense reconstruction. Moreover, it is mainly designed for non-deformable scenes [PB12], despite attempts with deformable ones using a template [MBC11] [HPF$^+$12].
- **Simultaneous Localization And Mapping** (SLAM). Such methods are similar to SfM, but uses a more sparse approach and relies on rigid points in the scene. SLAM is less computationally demanding and more robust than SfM, but requires a good initialization and the rigidity of most the scene. Variations of SLAM include Extended-Kalman-Filter-SLAM (EKF-SLAM), Parallel Tracking And Mapping

(PTAM) [GBC+14], SfM/SLAM hybrid [MNK+12] [CPB+13] and Motion-compensated SLAM [MY10]

- **Stereovision**. Binocular endoscopes aim at granting laparoscopic depth perception to surgeons and are nowadays common. A popular approach to laparoscopic surface reconstruction is to use such endoscopes and mimic human stereovision by matching features between the "left" and "right" images. Contrary to SfM and SLAM, stereo-based methods do not require motion to perform. In binocular endoscopes, two visual sensors are confined within a very limited radius, which imposes a narrow baseline and smaller lenses. While the former does not significantly affect stereo disparity given the high image resolution, the latter typically induces more noise and larger image distortion compared to monocular endoscopes. Much like SfM and SLAM, such methods heavily depend on the presence of numerous features to match, but laparoscopic scenes do not always offer such texture richness, especially on the liver surface. Reports of stereo-based laparoscopic surface reconstruction include [LCS+08] [HPR+09] [SSPY10] [CMC+10] [RBS+12] [HDP+13] [CHD+14] [TTS+14].

- **Shape-from-Focus** (SfF). Contrary to the previous methods, this approach does not use the parallax effect to infer depth, but rather the optical properties of the lens and the sharpness of elements in the laparoscopic image [TNK+09].

- **Shape-from-Shading** (SfS). Based on a lighting model, such methods exploit the shading effects on the elements of a scene to deduce their depths. Unlike all previous methods, SfS performs better in texture-less areas. Despite its good performance in simple scenes, it is a weakly constrained problem and its core hypotheses are often violated in laparoscopy (see more in Section 7.1). As concluded by Collins et al. in [CB12a], SfS is limited on its own, but provides an opportunity to multiply depth cues in hybrid methods of surface reconstruction, such as stereo/SfS [LCS+08] and SfM/SfS [MB14].

The clear advantage of passive surface reconstruction techniques lies in the compatibility with common endoscopes. Nonetheless, none of these methods have been specifically designed for laparoscopic scenes and are all adapted from algorithms in computer vision. Commonly, a lack of distinctive features to match across images compels to enforce local surface smoothness, via B-splines for instance, in order to avoid partially incomplete reconstructions. In laparoscopic scenes, the smoothness constraint limits the ability to handle the separation between organs and occlusions from surgical instruments. In other words, while passive techniques perform well in rigid non-complex scenes, they fall short in intricate, inhomogeneously illuminated, texture-less, highly dynamic and deformable scenes with frequent unpredictable occlusions by instruments, smoke or blood. Active techniques, although requiring specific equipment, may be better suited for such scenes.

Active techniques

These techniques are somewhat the inverse of passive ones: they are quite robust to laparoscopic scenes, but they are not clinically ready yet, as they require new or adapted endoscopic hardware.

- **Shape-from-Polarization** (SfP). Invisible to the human eye, polarization is a particular property of light that varies after reflection depending on the angle of incidence. By illuminating a laparoscopic scene with a source of polarized light and by including a polarizer in the endoscope, the surface of the site can be reconstructed [MHMMB13]. This approach is promising but the endoscope prototype remains too large to be inserted into standard trocars.

- **Structured light**. This technique attempts to tackle the lack of texture in laparoscopic scenes by optically projecting predefined patterns onto its surface. Then, the topology can be retrieved using any passive methods, let alone SfS. The main issue is that the surface resolution is limited by that of the projected pattern. Reports of use of structured light in laparoscopy include [NBG+08] [SFSA12] [MADdM12] [AHR13] [EPY+15] [FMM+15]. A variation of this technique is proposed by Collins et al. in [CB12a], as they perform SfS on a scene illuminated by separate light sources of different colors.

- **Time-of-Flight** (ToF). In the manner of a sonar, ToF cameras measure the topology of a site by sending and receiving pulsed signals at a known speed. However, in this case, the pulses are optical and emitted by a grid of lasers. This allows a much faster capture than that from a scanning from a single laser [HSN06]. The main issue with endoscopic ToF is that the limited radius of insertable instruments imposes a compromise between size and resolution. Moreover, its accuracy seems to be inferior to those of stereovision or structured light [MHGB+14]. Development and/or usage of endoscopic ToF cameras are reported in [PHS+09] [MMS+11] [BSH+13]. In an interesting approach, Köhler [KHB+13] proposes to merge a ToF camera with a standard imaging sensor in a single hybrid stereoscope of standard radius. The resolution of the resulting surface is virtually increased using a multi-frame scheme.

- **Physical tracing**. This type of surface reconstruction is performed by a tracked robotic arm which picks points at the surface of the scene. Tracing may be achieved using a *da Vinci*® robot [2] [AOG+11] or turning an optically tracked laparoscopic instrument into a stylus [HKMG09].

Unlike passive techniques, active ones handle the visual complexity of laparoscopic scenes and the lack of texture. However, this comes at a price of modified hardware in thin endoscopes, which often imposes a trade-off between surface resolution and compatibility with MIS standards. Furthermore, in a comparative study on phantoms of various abdominal organs, Groch et al. showed that surface reconstruction by stereovision was more accurate than with SfM or ToF [GSH+11]. In a separate but similar study, Maier-Hein et al. later confirmed this result by demonstrating the superior accuracy of stereovision over structured light and ToF [MHGB+14].

Nonetheless, for both active and passive techniques, once the surface is reconstructed, there still remains to register it to its counterpart extracted from the intraoperative 3D data in order to achieve AR.

[2] Intuitive Surgical, Inc., Sunnyvale, California, USA

Surface registration

In the surface-based approach to laparoscopic AR, the reconstructed surface of the intraoperative scene is then registered to the surface of the model segmented from preoperative 3D data. A wide variety of solutions to surface correspondence is contained in the literature. A more complete background is provided by Van Kaick et al. in their survey of shape representation and correspondence [VKZHCO11]. Also, Tam et al. recently published a comprehensive review of methods of rigid and non-rigid registrations between surfaces and point clouds [TCL+13]. For more focus on medical data, Heimann et al. examine solutions to shape-to-shape and shape-to-volume matching in [HM09].

In laparoscopic surgery, organs are often only very partially exposed. Therefore, the resulting reconstructed surface is itself very much incomplete. Even resorting to a full mobilization of an organ, only approximately half of its surface could be reconstructed at once. Moreover, organs often lack salient features which would lift ambiguity in the registration process. This issue varies depending on the considered organ. In a study on the kidney, Benincasa et al. reported that a minimum of 28% of visible surface is required to ensure an accurate registration [BCHG08].

Provided a sufficient exposure, registration in surface-based AR methods (task {2} in Fig. 3.5) is typically divided in two steps: initialization via a rigid registration and refinement, often non-rigid. While some have reported laparoscopic surface-based AR using only a rigid registration [PSCM+14] [TTS+15b], Simpson et al. experimentally demonstrated the importance of non-rigid registration on a realistic phantom of the liver [SDJM12].

Initial rigid registration (task {2.1}) The first step consists in a rigid registration, for a target organ, between its complete surface from the preoperative model and its intraoperative partial reconstruction. This task can be performed manually [CMC+10] [HDP+13]. However, even with a high level of expertise in anatomy, the operator would still have to perceive and manipulate three-dimensional geometry in scenes perceived from a single point of view and mostly in 2D, which may be a difficult task. As a result, such an approach is likely to be time consuming and is thus usually avoided in order not to lengthen the intervention and limit the risk of complication. In an effort to simplify the registration process, several groups have reported the use of semi-automatic methods [SVA+09] [AKNP+13] [CPB+14]. Provided an approximate yet close (manual) initialization, a popular registration algorithm is Iterative Closest Points (ICP) and its generalized variation [SHT09] [MHFdS+12]. Nonetheless, a fully automatic and robust approach to rigid initialization may be possible using shape feature descriptors, as proposed in [dSSK+14].

Registration refinement (task {2.2}) Refinement after the initial registration can be achieved using distance minimization on a rigid model [SVA+09], but a better accuracy is reachable with non-rigid methods such as B-splines, Thin Plate Splines (TPS), a spring-mass system [VSTH08], Coherent Point Drift (CPD) [MS10] or an adaptation of ICP with anisotropic scaling [CMBP15]. For a more realistic estimation of the deformation between the surfaces, the mechanics of the tissues can also be taken into account with the Finite Elements (FE) method. By incorporating FE into the refinement scheme, Cash et al. reported up to 4 mm of RMS error on a liver phantom [CMS+05] and more recently Suwelack et al. achieved less than 1 mm in near real-time *in silico* and on liver phantom [SRB+14]. By extension, the use of biomechanical models for the entire organ is also possible[AOG+11] [HDP+13]. In another approach to registration refinement, Wang et al. proposes to non-rigidly register sparse surface

reconstructions of an organ to its preoperative model using a learned structured dictionary [WT12]. Nonetheless, the creation of such a dictionary requires many manual segmentations of the structure of interest, preferably on different subjects.

In conclusion, the surface-based approach to laparoscopic AR bears clear advantages. Requiring only a preoperative model and endoscopic image analysis, this kind of method is already compatible with any current OR setup. No additional tracking system is needed, as the estimation of the camera pose is constant and inherent to the surface reconstruction, thus updating the augmentation according to the motion of the endoscope. Such a process can allegedly be executed in near real-time [SVA+09] [FRH+10] [RBS+12]. However, intraoperative modifications to the optical system (e.g. focus or zoom) induce changes in the intrinsic parameters. While methods exist to automatically compensate such variations for the linear intrinsic parameters, non-linear ones, such as lens distortions, require a new online camera calibration which affects the surgical workflow.

Summary of the limitations of the surface-based approach

The main disadvantage of the surface-based approach lies in the incompleteness of the surface reconstruction. Much like the tip of an iceberg, only a minor portion of the surface of the target organ is visible and hence reconstructible. This makes the registration to a complete preoperative model difficult and the result is only partially reliable. The deformation of both the invisible surface and inner critical structures thus remains relatively unknown, which is an issue for guidance. Moreover, if resection is let aside, there is a discrepancy in the magnitude of laparoscopic deformation between abdominal organs, due to their differences in size and composition. For example, as already mentioned in Section 3.4, the liver typically deforms more than the kidney in a laparoscopic setting. The use of biomechanical models may alleviate this issue, but only to a certain extent depending on the visible/invisible ratio.

Validation is also a concern for surface-based methods. Indeed, despite clear efforts to make phantoms as realistic as possible, *in vitro* experiments are still far from *in vivo* conditions, especially in terms of dynamism. Moreover, the lower anatomical complexity in phantoms unrealistically reduces the challenge in segmenting the intraoperative 3D data. Therefore, surface reconstruction methods still suffer a lack of intraoperative ground truth which prevents a proper accuracy validation for this type of AR approach.

In light of these drawbacks, the use of intraoperative 3D imaging in a volume-based approach to laparoscopic AR seems attractive, as explained in the next subsection.

Volume-based approach

This kind of approach involves an intraoperative 3D imaging system other than the laparoscope. Due to the popularization of hybrid rooms, these systems can be CBCT or open MR scanners. Sweeping a tracked 2D ultrasound transducer may also fulfill this role. The goal is to use the intraoperative volume data as an intermediary in the augmentation process, by splitting the registration into two distinct tasks (see Fig. 3.5).

Registering Preoperative 3D data to Intraoperative 3D/2D Acquisitions

The purpose of using an extra imaging system intraoperatively is to determine the complete deformation of the region of interest (ROI) between the time of the preoperative acquisition and the intraoperative state of that same area. This then allows a non-rigid registration of the preoperative 3D data to acquisitions performed intraoperatively (task {3} in Fig. 3.5).

Numerous medical imaging systems provide three-dimensional data, thus 3D-3D registration is a common task in this field. The literature is well furnished about this topic. A good introduction to medical image registration is presented in [RS11]. For more in-depth content, Heimann et al. review shape-to-volume and volume-to-volume approaches in [HVGS+09] and a comprehensive survey of deformable registration methods is given in [SDP13]. Point-based methods rely on the location of anatomical or artificial landmarks to perform the registration [BSMS+08]. Intensity-based methods make use of the value of all voxels, like the popular Demons algorithm and its variations [FVW12].

Besides the data interpretation, diversity in accuracy also occurs concerning the parametrization of deformation. A simple approach consists in adopting an affine deformable model [CPB+14] [PSCM+14]. Such a model can represent elongation and compression in every direction without volume constraint, but cannot encapsulate bending. For more complex deformations, models based on free-form B-splines or TPS but require more, well-scattered correspondences between the preoperative and the intraoperative 3D data.

Nonetheless, the resulting deformations may not be realistic. This can be alleviated by including biomechanical properties, such as Youngs modulus and Poissons ratio, into tetrahedron mesh modeling the preoperative 3D data [BHE+08] [HDP+13] [OZM+13] [BNH+13] [MFN+14]. The realism of the deformation is essentially bound to the accuracy of the fixed properties. Unfortunately, these properties are patient-specific and a dedicated model must be designed for each of them preoperatively, with a possible refinement using intraoperative US elastography. Real-time processing is difficult but possible depending on the realism of the model, the complexity of the simulated deformation and the potential collisions with other models.

So far, we only considered 3D-3D registration methods. However, a volume-based approach is also possible using several 2D acquisitions. For example, Markelj et al. reviewed the various ways to register preoperative 3D data to intraoperative fluoroscopies [MTLP12]. 3D-2D registration may also be performed thanks to US imaging. Despite its poor image quality and limited field of view (FOV), this low-priced and non-invasive imaging modality may be a popular alternative to CBCT and open MR in laparoscopic interventions, especially on liver [VRL+10]. Various methods exist for US/CT or US/MRI registration [SDJM12]. Ukimura et al. registered preoperative CT and MR data to the images of an optically-tracked sweeping laparoscopic transducer [UNS+10], based on a pre-calibration. Dagon et al. update a preoperative deformable model from detected vessels in intraoperative US images from a tracked external transducer [DBB08]. Other methods have been reported, that do not require tracking of the probe. For example, Lange et al. propose to register US and CT data based on anatomical landmarks and voxel intensity [LPH+09]. Nam et al. similarly registered US to CT data via edge matching [NKL+12]. Wein et al. resorted to simulated US from CT to perform a rigid or affine US to CT registration [WBK+08].

Overall, there are various ways of the preoperative 3D data to their intraoperative counterpart, but accurately determining the deformation still remains a challenging task. An alternative could be to discard all preoperative data and rely only on intraoperative ones [FMHN08] [SDB+10], despite lesser quality. Pros and cons of relying only on intraoperative imaging are discussed in Section 8.3. At any rate, the intraoperative acquisitions still need to be registered to the laparoscopic image, as seen in next.

Registering Intraoperative 3D Data to Laparoscopic Image

Despite being the representation of a same scene, intraoperative 3D data and laparoscopic images are very different in terms of information. Accurately registering both is a challenging task ({4} in Fig. 3.5).

Artificial fiducials may be used to alleviate the lack of cross modality landmarks. This technique was popular in the early days of laparoscopic AR where radio-opaque markers were stuck against the patient's skin and kept so until the intervention. The markers could then be used to ease a manual rigid registration with the laparoscopic image [MRA+04]. Another method is to point them with optically tracked instruments [MLT+04] [SIU+13][TTU+13], but this approach is too constrictive and approximate.

For a more accurate estimation of the position of the laparoscope with respect to the anatomy, Baumhauer et al. propose to move the fiducials inside the patient and to stick needles with colored heads into the organ [BSMS+08]. Then, a CBCT acquisition was performed and the fiducials were extracted semi-automatically. Finally, using "inside-out tracking" of these same markers in the laparoscopic image, the position of the camera with respect to the intraoperative 3D data was determined. The reported maximum Target Registration Error (TRE) is 0.89 mm during *in vivo* tests. However, their method still requires the unpractical insertion of fiducials into the organ. In a significantly less invasive manner, Cheng et al. propose to use photo-acoustic markers to register US images to laparoscopic ones [CKTB13]. As described in [CGK+14], the process involves the emission of laser pulses inside the abdomen via optic fibers. A sub-millimetric accuracy is reported, but their technique has been tested only *ex vivo*.

A similar approach to the one presented in this thesis is proposed in [MFN+14]. After registration of preoperative CT data to intraoperative CBCT one, Mountney et al. propose to include the tip of the endoscope into the field of acquisition of the CBCT scanner. Two fluoroscopies are then performed and analyzed to determine semi-automatically the position and direction of the laparoscope. These parameters are later refined using an approach similar to that of the surface-based methods described in Section 3.5. Indeed, the surface of the organ is reconstructed by stereovision and registered to the preoperative 3D model.

Probably the most popular approach to registering intraoperative 3D data to the laparoscopic image is to resort to external tracking, either optical as in [FMHN08] or magnetic as in [KHN+05]. These two technologies are explained in details in Appendix B.

In conclusion, volume-based approaches, despite their use of an additional imaging system, are capable to better capture the intraoperative anatomy. However, for both volume- and surface-based methods of laparoscopic AR, the augmentation needs to remain accurate while the surgeon needs it. Hence, camera motion and scene deformation have to be taken into account to update the AR. In the next section, we review the various techniques to handle dynamism in laparoscopic AR.

Summary of the limitations of the volume-based approach

The weakness of the volume-based approach is its requirement for additional systems. First, its usage is limited to hybrid ORs, as an intraoperative 3D scanner is necessary. Such machines are very expensive, but their usefulness has convinced more and more hospitals to invest. Second, most current volume-based methods rely on tracking systems for the registration, which are often impractical. Other methods rely on tracking natural or artificial landmarks

for triangulating the camera's position, but they are either invasive or lack robustness, as explained in Appendix C.

3.6 Conclusions

Laparoscopic AR is a promising technology that aims at providing pertinent guidance to the surgeon within the endoscopic image. Depending on the augmentation envisioned, various types of data, rendering and display are available. Nonetheless, laparoscopic AR is a challenging task, essentially two-fold. The first difficulty is to accurately register images from different modalities and acquired at different times, which respectively introduces an absence of texture correspondences and often significant deformations in between acquisitions. The second challenge is to maintain the augmentation accurate in a highly complex and dynamic scene, constantly moving and distorting because of metabolic motions or surgical actions.

For both challenges, this chapter has shown that a wide variety of methods has been proposed in the literature over the last decade, but also that they revolve around only two distinct approaches. The first one relies on the reconstruction of the surface of the laparoscopic scene and its registration to the surface of a preoperative 3D model. The main advantage of this approach is its inherent consideration for scene dynamics, especially camera motion. Its main drawback is that the registration is only reliable for the visible surface, which may be a problem for large organs like the liver. The second approach to laparoscopic AR is based on the intraoperative imaging of the patient's anatomy, often in 3D, and its registration to the preoperative data. On the contrary of the surface-based approach, this one takes into account deformations beneath the visible surface, but lacks the support of dynamics in the scene. As a result, volume-based AR methods often resort to additional and unpractical tracking systems to maintain an augmentation accurate.

Overall, no satisfactory method has yet been developed for laparoscopic AR. Despite many attempts during the past two decades, there is still no commercial AR solutions available for laparoscopy, contrary to other surgical fields. We identified a few remaining issues that need to be tackled in order to carry laparoscopic AR beyond the prototyping phase.

Accuracy and Validation

The main hurdle faced by laparoscopic AR at the moment is accuracy. Beside a qualitative visual appreciation, there is so far no mean of measuring the correctness of an augmentation and also no possible way of predicting the error. This situation is due to the usage of multiple complex systems which each adds to the uncertainty. For example, optical tracking is often used as a ground truth for camera motion and fiducial pointing in laparoscopic AR. However, as pointed out by Maier-Hein et al. in [MHMB+13], such systems accumulate errors from the tracking, the camera calibration and the hand-eye calibration.

This lack of accuracy introduces a clear need for validation means in surgical AR [CP10]. Currently, laparoscopic AR methods may be validated on phantoms or *ex vivo* data for rigid registrations, but only synthetic data are available for non-rigid methods, with a limited realism. As explained in [NBA+15], a surgical validation of AR would require to surgically expose the hidden structures showed by the augmentation, without deforming the scene nor moving the structures, which is unlikely.

Without a reliable validation protocol, a laparoscopic AR method does not comply with the strict requirements of certification and thus cannot translate into an actual product.

Clinical Integration and Evaluation
In [SFT+06], Sielhorst et al. outlines three essential criteria to ensure the quality of a medical augmented reality system:

reliability: the system offers by real-time accuracy control in any situation.

usability: a strict minimum of interaction by surgeon is required.

interoperability: generic data and protocols are used, which guarantees the largest compatibility with other equipment.

The previous paragraph already explained how current laparoscopic AR methods fail the first criterion. Unfortunately, the two other criteria are also often considered as secondary and thus overlooked. Therefore, proposed systems require complex protocols composed of tedious setups and lengthy calibrations. Moreover, its usage is typically complicated by an unintuitive user interface which leeches too much of the surgeon's attention. As a result, a seamless integration of current laparoscopic AR methods into the clinical workflow is nearly impossible and ultimately, with very limited existing trials, the actual benefit of laparoscopic AR has not yet been proven [OOK+14]. Indeed, the accuracy alone does not demonstrate the utility of a surgical system, especially since constraints like resection margins depend on the type of intervention [OOY+14] [NMSA13].

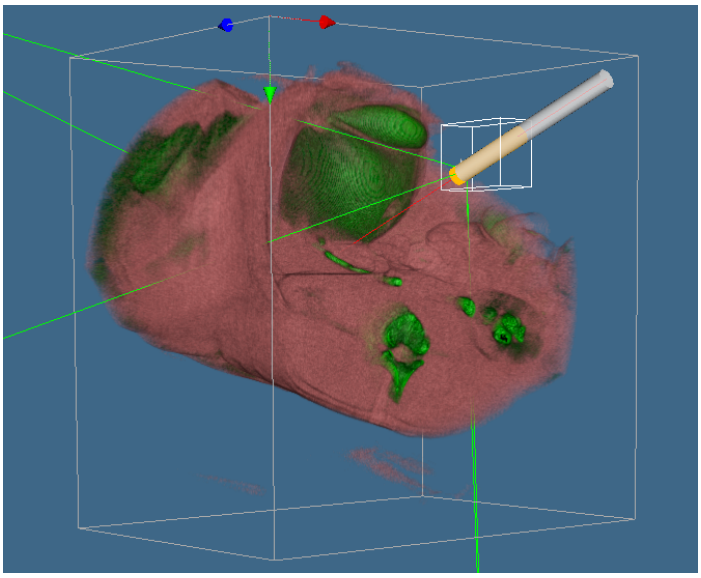
In [KOJC12], Kersten et al. insists on the distinction between **validation**, which is the fulfillment of accuracy goals, often demonstrated on phantoms, and **evaluation**, which consists in providing clinical proof of its utility via standard parameters such as blood loss, intervention time, complication rate and, of course, the surgeon's appreciation. This compels for more multidisciplinary research teams, as engineers often focus too much on validation and too little on evaluation.

Our approach to laparoscopic AR
In light of the previously detailed issues, matching the objectives stated in Chapter 1, we developed a new paradigm to static laparoscopic AR at our institution. We set our focus on accuracy, compatibility thanks to multidisciplinary work and reproducibility through a seamless integration to the OR. We have chosen a volume-based approach for its advantages and we alleviate its drawbacks through a novel way of registering the intraoperative 3D data with the laparoscopic image (task {4} in Fig. 3.5). Strong efforts have also been put into validation by experimenting on various *in vivo* data with clinically sound purposes. The following chapters describes in depth our methodology and major achievements.

Chapter 4

Linking Endoscope and Intraoperative 3D Data

This chapter details our original approach to laparoscopic AR based on the introduction of the endoscope in the acquisition field of an intraoperative 3D scanner. Thus, without a tracking system and a dedicated calibration, the relationship between the endoscopic camera and the intraoperative 3D data can be determined. We first show that a simple threshold-based extraction of the endoscope in the 3D data is not robust and a novel 3D template-based method is proposed. We then explain why the roll angle of the camera remains unknown, but we demonstrate that the integration of an accelerometer into the endoscope can lift this uncertainty. Finally, quantitative experimental results are provided to expose the potential of our approach, but these also reveal a structural problem of most endoscopes, which is the subject of Chapter 5.



4.1 A novel approach

Motivations

In light of the current issues in laparoscopic AR, as exposed in the previous chapter, we propose a new paradigm to automatically register the referential frame of the intraoperative model with that of the endoscopic camera. Our approach does not rely on an impractical or costly tracking system, requires limited calibration and the registration is valid even underneath the visible surface. The core idea is to include the distal part of the endoscope within the intraoperative 3D acquisition field and holding it with an articulated arm, thus allowing to directly determine the direction of the optical axis and the position of the optical center in the reconstructed volume. This approach allows us to determine directly the correspondence between the endoscopic camera and the 3D imaging system (in our case, an Artis Zeego by Siemens).

The idea of inserting the endoscope in the intraoperative 3D image has been introduced with our LASAR project. A parallel work, presented in [MFN+14], also exploits this idea by computing the position of the endoscope from fluoroscopies. Nonetheless, contrary to our work, their proposed method of endoscope extraction is not fully automatic and requires user's interaction to detect the endoscope in the images.

Applications

Obviously, this geometrical relationship holds only when the camera remains static, but this may occur several times during an intra-abdominal intervention such as a liver segmentectomy. First, static AR can be used to reveal hidden vessels during the outlining of the resection planes by electrocoagulation. Second, AR can be used during the resection stage itself to identify encountered veins as hepatic or portal ones, as the former requires electric scissors and the latter clipping. Third, overlays can help guide the needle during a thermal ablation. Static AR may also be used to reveal critical structures in other disciplines like ureters in gynecology.

Every time the endoscope is displaced, the static AR can be updated by performing another intraoperative acquisition. If continuous endoscopic navigation is required, then we could resort to a classic camera tracking technique such as SLAM or an external optical tracking apparatus, but our approach would still remove any need for calibration of this system. Finally, provided that a transfer function allows a good visualization of critical structures in the volume rendering mode of the intraoperative images, we are then able to directly achieve AR without preoperative imaging.

4.2 Description of the system and protocol

This section describes the typical intervention setup that we aim at with our method. The patient lies on the operation table near to a 3D rotational C-arm machine, as in Fig. 4.1.

After the pneumoperitoneum is performed, the surgeon inserts the endoscope inside the patient through a port and adjusts the zoom and focus of the camera. In this work, we only consider monocular 0° angled endoscopes as they are still widely used for interventions and their optics is easy to model. Nonetheless, we discuss the adaptation of our method to 30° endoscopes and stereoscopes in Section 8.2. Once the camera settings are fixed, the

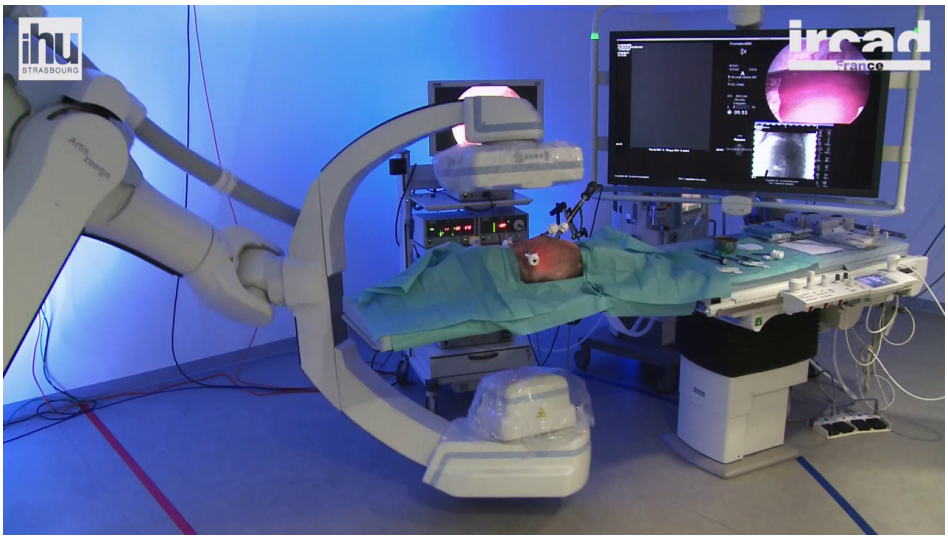
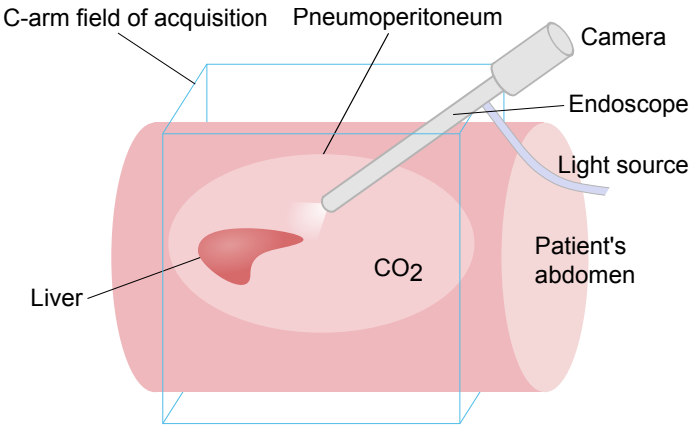


Figure 4.1: Picture of a typical setup in the hybrid operating room.

endoscope is pulled out of the patient and a camera calibration step is performed (as described later in this section), but this only involves an analysis of the camera output and is typically achieved within 30 seconds. The endoscope is then inserted back in the abdominal cavity and positioned so as to obtain an optimal view of the abdominal region of interest, but also so that it will appear in the intraoperative 3D image acquired by the C-arm (see Fig. 4.2-left). One can make sure it is sufficiently inserted by performing the common 2D fluoroscopic acquisitions that are parts of the clinical routine (see Fig. 4.2-right). An experimental study about the minimum depth of insertion is provided in Section 5.4.



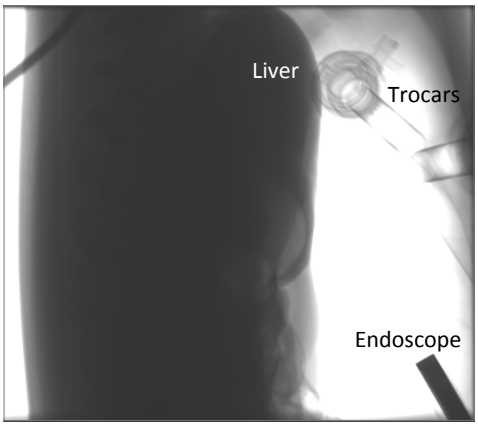


Figure 4.2: On the left, a simplified illustration of our typical setup for an abdominal intervention. On the right, a fluoroscopy showing both the anatomy and the endoscope tip.

One may argue that the presence of the metallic endoscope would introduce artifacts in the CT reconstruction. Nonetheless, at a nominal distance from the organs, the endoscope produces only minimal artifacts. Most importantly, these are located underneath the endo-

scope tip i.e. outside the ROI, as shown in Fig. 4.3. The contrary is unlikely, as the endoscope would have to be positioned right above the ROI, which is seldom in this kind of intervention.

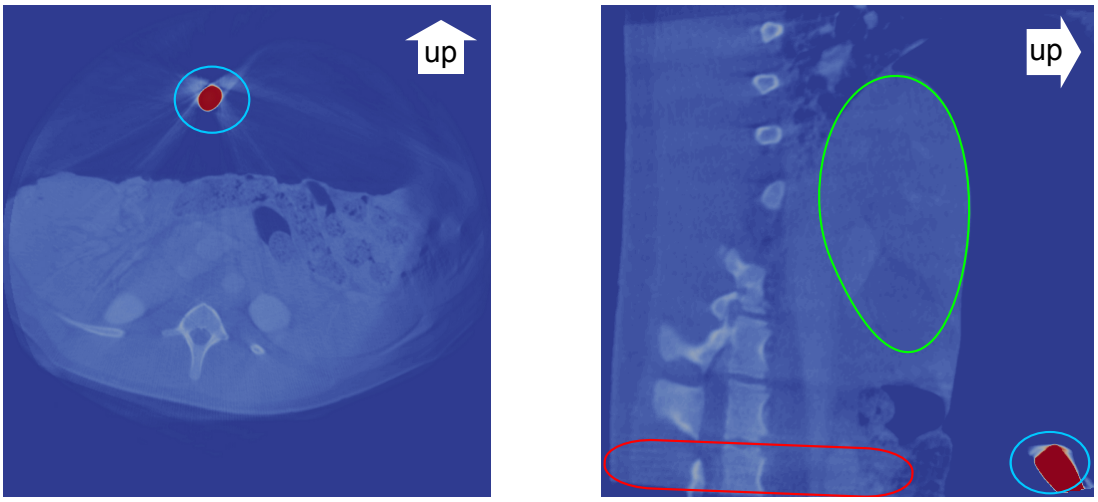


Figure 4.3: Illustration of endoscope-induced artifacts in pig image. The axial view (left) shows that the endoscope (outlined in blue) produces thin artifacts in the reconstructed volume (zoom in for details). However, the sagittal view (right) reveals that these (outlined in red) are only located underneath the inserted tip (in blue) and do not degrade the region of interest which is the liver (in green).

To accurately augment the endoscopic view with intraoperative data, we must create a virtual camera with the same view upon the ROI as the actual endoscope. Thus, the virtual view will match the endoscopic image content. Therefore, we have to determine all the intrinsic and extrinsic parameters of the endoscopic camera, as summarized in Fig. 4.4.

Intrinsic camera parameters
The intrinsic parameters are determined through a classic camera calibration based on Zhang's method [Zha00]. After the surgeon has set the optics as desired, a set of images of a generic checkerboard pattern is captured with several orientation angles. Thereby, we can estimate the focal length, from which the FOV for the virtual camera can be calculated. Additionally, the calibration provides the position of the optical center in the image plane and the radial distortion of the lens, which can be used to undistort the endoscopic image. The quality of the calibration is commonly measured by the error from reprojection. For our experiments, we typically got reprojection errors up to 0.7 pixel, which may seem large for such a calibration, but can be explained by the choice in checkerboard size[1].

4.3 Determination of the extrinsic parameters

The extrinsic parameters consist in the location and orientation of the camera in 3D space, which is here the referential of the intraoperative 3D data. There are six DoFs: three for

[1]Using a checkerboard of regular size sets it beyond the nominal distance of the endoscope and thus the image is not as sharp. On the contrary, if we resort to a small checkerboard, the image is much sharper but then the printing precision of the pattern can influence the outcome

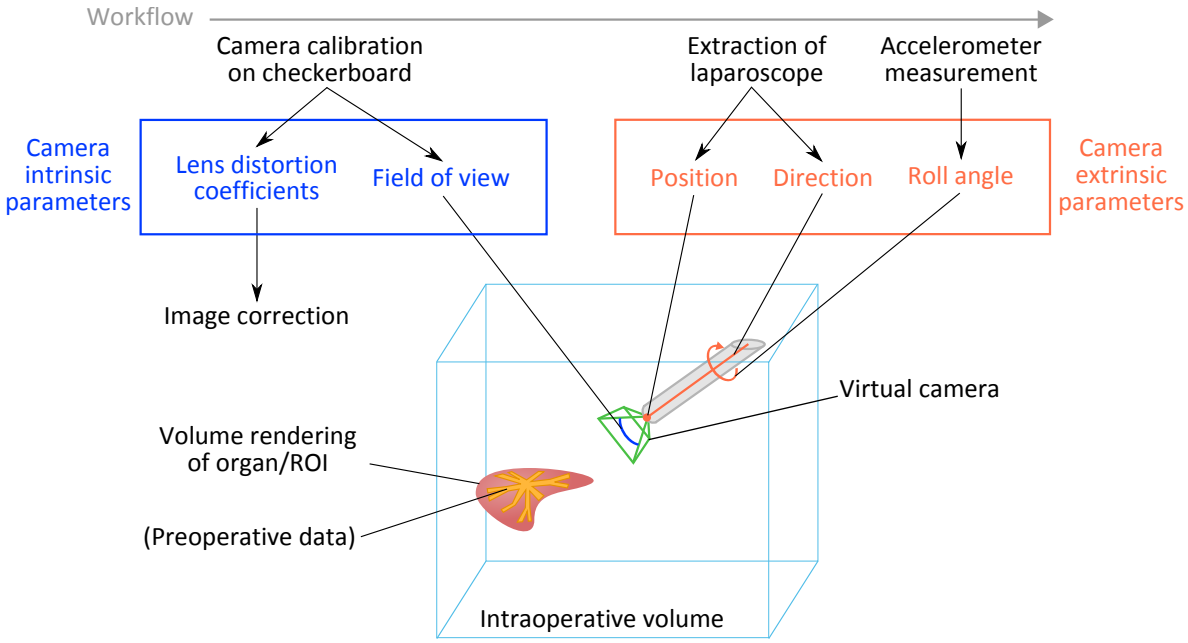


Figure 4.4: Summary of our approach. This illustration shows what are the various camera parameters to determine and their connection with our AR application.

the coordinates of a point denoting the position and three for the tilt, yaw and roll angles describing the orientation. The first five DoFs can be estimated by analysis of the intraoperative 3D data showing the tip of the endoscope. This may be achieved by a simple extraction based on a gray-level image threshold, but we show in Section 4.3 that this lacks robustness. Therefore, in Section 4.3, we propose a 3D template-based method to accurately determine these five DoFs. However, due to the cylindrical shape of the endoscope, the roll angle cannot be deduced from the intraoperative 3D data. Thus, we choose to resort to an integrated accelerometer for the resolution of this DoF, as explained in Section 4.3.

Endoscope extraction based on image thresholding

The endoscope is mainly composed of titanium and thus is highly opaque to x-rays. Therefore, it yields very high values in the CT image, much larger than those of surrounding artifacts or any human tissue, as shown in Fig. 4.5-left. Extracting the endoscope in the image then becomes trivial as we have a large range to set a threshold value for segmenting its voxels (orange stripe in Fig. 4.5-a on the right). This kind of approach has already been used to extract tubular yet thinner objects such as needles [HDH+02][CIN+03].

As explained in Fig. 4.6, once the endoscope voxels are segmented, we can successively select them (in plain blue) along the axis of entry into the volume (here the z-axis) and calculate the barycenter of the selected slice (in orange). The irregularity in the number of voxels in successive slices allows to discard incomplete ones near the entry plane. Likewise, this serves as an indicator as to when the end of the endoscope is reached (yellow slice). With knowledge of the tubular shape of the endoscope, all of these barycenters are supposedly aligned along its axis of revolution. Therefore, by applying a closed-form three-dimensional

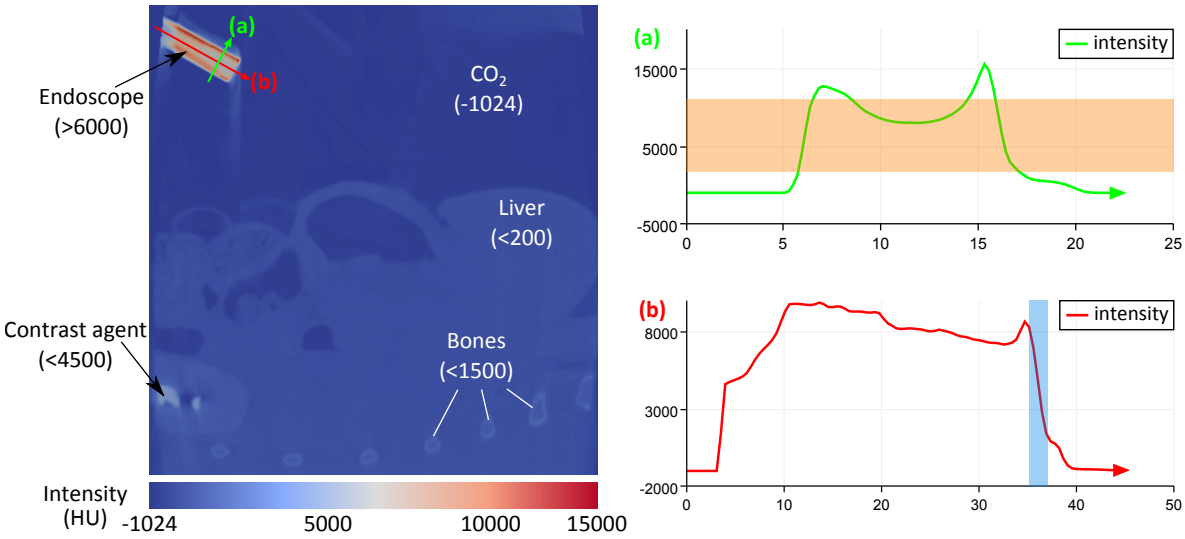


Figure 4.5: The image on the left is a slice of a CBCT image, going through both the endoscope and a pig's anatomy. On the upper right, the intensity profile across the endoscope (in Hounsfield units or HU), along the green line (a), illustrates how large the possible threshold range is (orange stripe). On the lower right, the intensity profile along the red line (b) shows the abrupt drop occurring at the tip location (blue stripe).

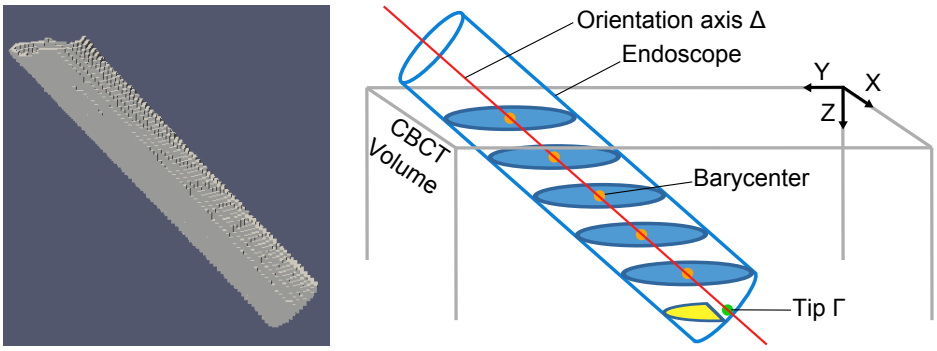


Figure 4.6: On the left, example of segmented voxels. On the right, determination of the axis of revolution Δ and the endoscope tip Γ.

linear regression over the barycenters, we then obtain the orientation axis Δ of the endoscope (red line in Fig. 4.6). This method of orientation estimation with slices prevents the bevel formed right at the volume boundary from falsifying the determination of Δ, as may occur with methods such as Principal Component Analysis (PCA). The Root Mean Square (RMS) error of each barycenter with respect to the obtained line is calculated as an indicator of quality of the estimation, but could also be used for weighted linear regression if need be. The position of the endoscope tip Γ (green dot in Fig. 4.6) is found by thresholding a large negative gradient in the intensity profile along the determined axis as shown by the graph in Fig. 4.5-b (blue stripe).

Dependency on threshold value

The choice in the threshold value depends on the content of the scene. However, tuning this parameter is difficult as the threshold value has a significant impact on the estimation of the orientation axis Δ, as illustrated by Fig. 4.7. This is due to the uneven absorption of the X-rays around the tip of the endoscope, caused by the anisotropic radiations from the CBCT.

Using a radio-opaque checkerboard (more details in Section 4.4), we have demonstrated this phenomenon experimentally by performing successive registrations while increasing the threshold value (see the bottom of Fig. 4.7). Thus, the accuracy of the registration is dependent on the value of the selected threshold.

In order to alleviate this issue and for the sake of providing a fully automatic AR approach whose results depend mostly on the data and not on the parameters, we propose another method based on matching a 3D template.

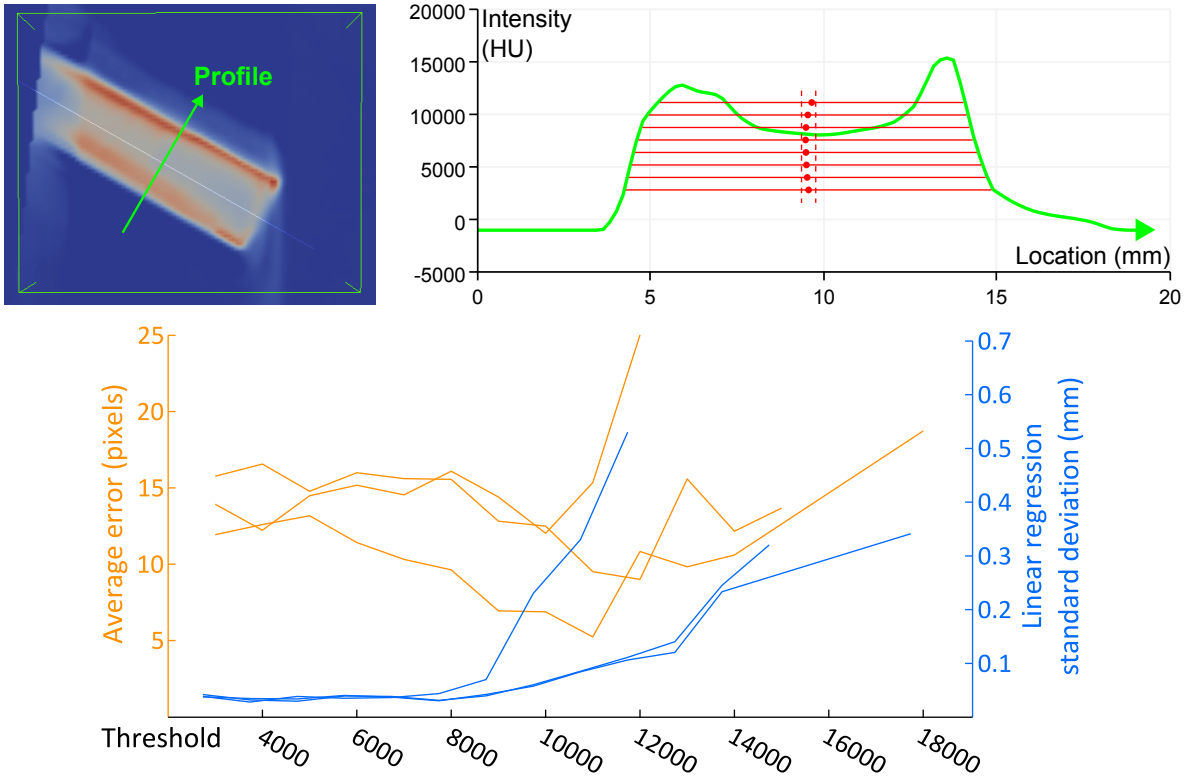


Figure 4.7: At the top, the intensity profile (in HU) across the endoscope section (green line) shows that the absorption is different from one side to another. As a result, determining the barycenter of a slice with various threshold values introduces a discrepancy in their location, shown between dashed lines. Although small, this inconsistency still impacts both the registration error and the linear regression fitness, whose variations with respect to the threshold value are showed at the bottom, in orange and blue respectively, in three different cases.

Endoscope extraction based on 3D template

Statistical shape modeling is a popular approach to find a specific structure of known geometry in 3D data [FLD02]. In our case, the tip of the endoscope presents a simple tubular shape which does not require such an advanced technique. Therefore, we can retrieve the position of the endoscope by fitting a 3D cylindrical template Ψ on its representation in the CBCT acquisition. The goal is to fit Ψ on the voxels of highest intensity, assumed to be at the surface of the endoscope, as illustrated in Fig. 4.8.

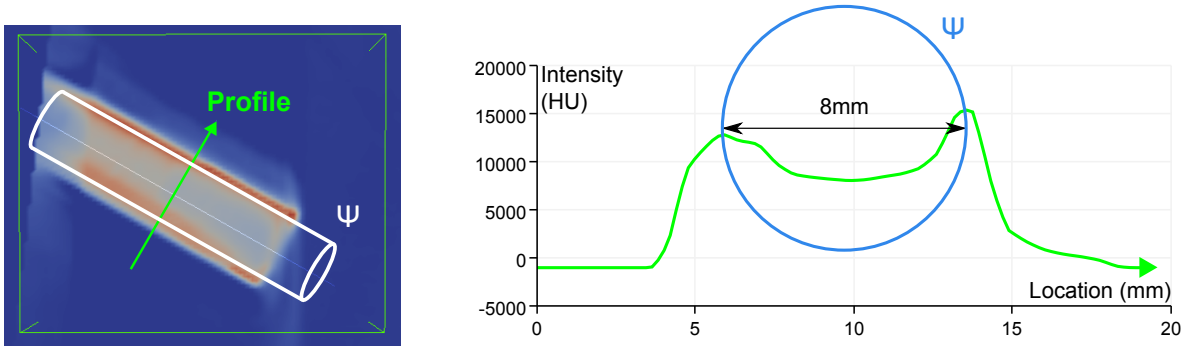


Figure 4.8: On the left, we introduce a cylindrical template Ψ (blue) to fit the surface of the endoscope (tip considered later). On the right, an intensity profile shows Ψ passing by the peaks corresponding to the tube.

The voxel intensity along Ψ is discretely measured by sampling the intraoperative 3D image I in circles, as shown in Fig. 4.9-a. As Ψ is centered at the origin O and aligned with the y-axis, the coordinates of the sampling points P are $\{R\cos(n\alpha), k\lambda - L/2, R\sin(n\alpha)\}$, with R the radius, $n \in \{0..N-1\}$, $k \in \{0..K\}$, the cylinder length being L and $\alpha = 2\pi/N$ and $\lambda = L/K$ the radial angular and longitudinal steps respectively. The choice in N and K is related to the spacing of I so as to sample every voxel along the template at least once. For example, our CBCT system provides a spacing of $s^3 = 0.49 \times 0.49 \times 0.49$ mm, so we can deduce that $N > 2\pi/\arcsin(s/R)$ and $K > L/s$. Despite knowing the actual radius of the endoscope, R is a parameter here due to the X-rays absorption. For example, the distance between the peaks in the intensity profile in Fig. 4.8) is measured at 8 mm for a 10-mm endoscope. The length L of the template Ψ is set as the longest diagonal of a bounding box based on the extracted voxels as in 4.3 and as illustrated by Fig. 4.9-b,c. These boundaries are also used to limit the search space during template fitting and sampling points outside the box are discarded.

As shown in Fig. 4.9-d, the sampling points P are displaced via a transformation T, consisting of a translation to a point C and two rotations based on the two first Euler angles (ϕ and θ). The average intensity of the voxels sampled at P may then be considered as an objective function F whose parameters are R, $C = \{C_x, C_y, C_z\}$, ϕ and θ. With $M \leq N(K+1)$ being the number of samples within the bounds, F can be expressed as:

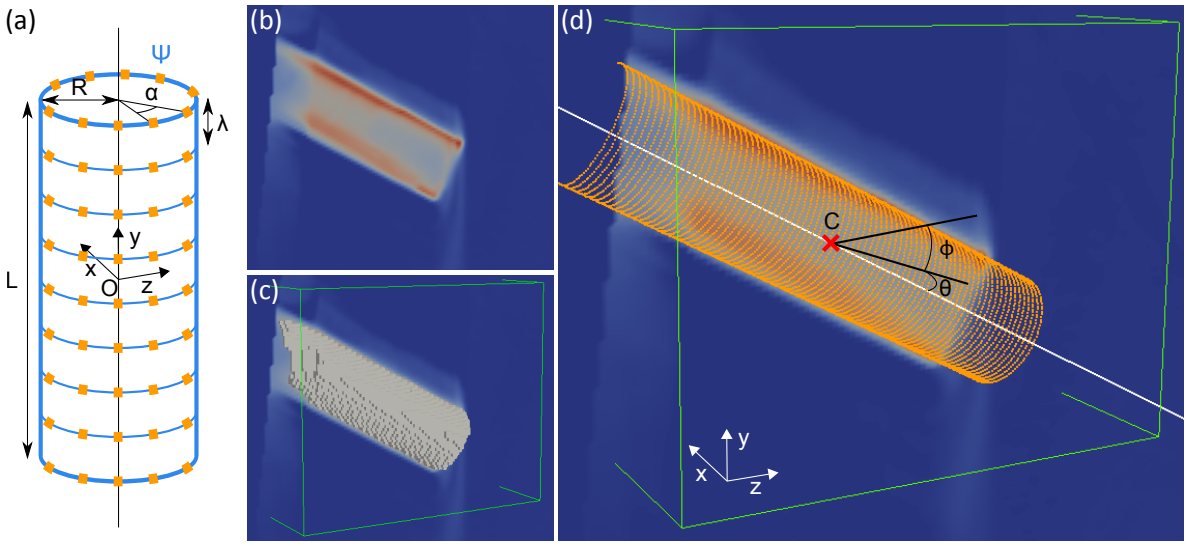


Figure 4.9: Sampling along cylindrical template. (a) The sampling points P (orange) are positioned along the template Ψ (blue) in circles of radius R. (b-c) The endoscope is shown in a slice from the CBCT data (top) and its voxels are extracted to determine a bounding box for the search (green, bottom). (d) The sampling points P are placed within this bounding box in the CBCT data, according to a transformation T applied to P. T consists of two rotations around the x- and y-axes (of respective angles ϕ and θ) and a translation to a point C on the axis of revolution.

$$
\begin{aligned}
F(R, C, \phi, \theta) &= \frac{1}{M} \sum_{p \in P} I(T(p)) \\
&= \frac{1}{M} \sum_{n=0}^{N-1} \sum_{k=0}^{K} I\Bigg(\begin{bmatrix} \cos(\theta) & \sin(\theta)\sin(\phi) & \sin(\theta)\cos(\phi) \\ 0 & \cos(\phi) & -\sin(\phi) \\ -\sin(\theta) & \cos(\theta)\sin(\phi) & \cos(\theta)\cos(\phi) \end{bmatrix} \begin{bmatrix} R\cos(n\alpha) \\ k\lambda - \frac{L}{2} \\ R\sin(n\alpha) \end{bmatrix} + \begin{bmatrix} C_x \\ C_y \\ C_z \end{bmatrix} \Bigg)
\end{aligned}
$$

The parameters of F are initialized with respect to the outputs of a PCA applied to the extracted voxels. C is first set as the barycenter and ϕ and θ are calculated from the eigenvector of the main component. The actual fitting is then obtained for $\hat{F} = \arg\max(F)$, using Powell's conjugate direction algorithm.

As with the threshold-based approach, the position of the endoscope tip is found by thresholding a large negative gradient in the intensity profile along the determined axis. It is worth noting that the resulting tip location will not actually correspond to the location of the virtual camera on the axis, notably because it depends on the zoom factor and the focus. Nonetheless, calibration experiments have shown that, for standard FOV and focus settings for such interventions, the correct location is very close to the actual tip of the endoscope and does not vary significantly across experiments. More details and experimental observations supporting this assertion are provided in Appendix E.

Also, given that we consider 0° endoscopes, there is a near perfect parallelism between the endoscope and its optical axis, as demonstrated in Appendix E and in Section 5.2. Therefore, we safely set the orientation of the virtual camera to that of the endoscope.

Moreover, for a perfect registration, the center of the virtual image ought to correspond the center of the real image. Due to limitations in manufacturing accuracy, such a perfect alignment is not guaranteed and shall be enforced. Since we assume that the optical axis and the axis of revolution of the endoscope coincide, the center of the virtual image is also the principal point. The latter is part of the intrinsic parameters of the camera and is commonly determined during its calibration. Therefore, enforcing alignment between the virtual and real images consists in a 2D shift in the image plane between the real image center and the principal point.

So, we are able to determine all the extrinsic parameters but one; the tubular shape of the endoscope prevents the determination of the roll angle around the axis of revolution.

Determination of the camera roll angle

Camera roll estimation - This angle is automatically estimated by using a high-end 3-axis accelerometer designed for low accelerations (model LIS331DLH from ST Microelectronics) included in the camera. It is aligned and soldered on the printed circuit board (PCB) together with the imaging sensor, thus ensuring their alignment within the typical reflow assembly margins (± 10 μm, resulting in a negligible ± 0.2° for this 3 mm-wide chip) (see Fig. 4.10). Moreover, this sensor is not considered as an external tracking device since it is integrated inside the endoscope.

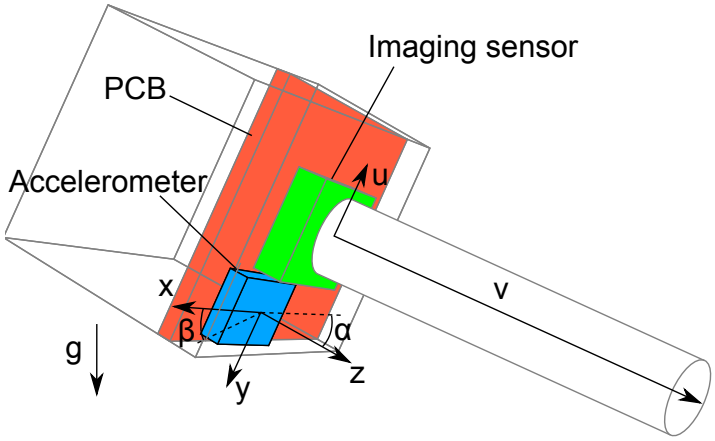


Figure 4.10: Simplified illustration of our endoscope and its sensing components. The scene is projected along the optical axis v onto the imaging sensor (in green) and forms an image whose up vector is u. The CCD chip is soldered on a PCB (in red) together with the accelerometer (in blue) which measures the tilt angles with respect to gravity denoted g. Therefore, its roll angle β and the pitch angle α are the same than those of the imaging sensor. The dashed lines represent horizontality.

So, the roll angle measured by the accelerometer is also that of the camera with respect to the gravity field. The interventional CT system is very precisely calibrated to provide

a volume with its vertical axis aligned with the Earth's gravity. Thus, the roll angle of the virtual camera with respect to the 3D CT data can be set to the value given by the accelerometer. The accuracy obviously decreases as the endoscope moves towards verticality. A cone of about 10° around the vertical axis is typically a satisfactory boundary. Although, given our setup (see Fig. 4.2), the endoscope is never placed vertically as this would prevent the interventional C-arm from rotating around the surgical site. Also, the manufacturer reports a variation of sensitivity with respect to the tilt angle outside of verticality. However, thanks to trigonometric formulae and a combination of the acceleration values along the three axes, the sensitivity can become independent from the actual tilt angle values (see Appendix F). Therefore, the roll angle estimation remains accurate in all cases.

We performed a quantitative evaluation of the roll angle estimation by comparing the value provided by the accelerometer and the one found by manually setting the rotation while trying to match visual cues in the scene. Across 14 experiments, we obtained an error of $0.15° \pm 0.24$, which is larger than the value of 0.1° reported by the manufacturer, but still sufficiently accurate for our purpose.

4.4 Experimental results

This section details the quantitatively validation approach for our registration method. First, we describe the setup and materials used. Second, the actual results and their analysis are provided.

Experimental setup - A dedicated checkerboard has been created to assess our method's accuracy (see Fig. 4.11). It has been printed in 3D in order to achieve a 0.1 mm accuracy for the pattern. Composed of 13×9 squares of 5×5 mm, it fills most of the field of view of an endoscopic camera at the nominal distance (around 70 mm). On one side, the regular checkerboard pattern is used by the camera to locate precisely each corner. On the other side, right underneath each corner, a 3 mm steel ball is inserted at a known depth of 5.8 mm. Since steel also highly attenuates x-rays, we are able to segment each of them using ITK and thus calculate the position of each corner in the volume image too. Hence, by placing the endoscope so that the checkerboard fills most of its view, it is possible to evaluate the precision of the registration over most of the camera field of view.

Following the protocol described in Section 4.2, once the camera has been tuned as desired and calibrated, the checkerboard is acquired by both modes of imaging (CBCT and endoscope). Using our template-based method, we locate the endoscope in the CT volume and place the virtual camera accordingly. Then, we independently determine the checkerboard corner locations from both the CT volume and the endoscopic image with respectively ITK (thresholding and labeling) and OpenCV (subpixel corner extraction). By performing the registration, the corner locations from the intraoperative 3D data are projected in the 2D space of the endoscopic image, as shown in Fig. 4.12. Therefore, we are able to calculate the registration error as the average distance in pixels between the two sets of points across the 1080p image.

Running time - On a mid-end machine with a Core i5-2500 processor (quad core at 3.30 GHz) and 8 GB of memory, our method takes between 7 and 13 seconds from launch to complete AR visualization, including volume rendering. Singled out, the actual extraction of the endoscope and its analysis range from 5 to 10 seconds. The camera calibration takes 7 seconds for 9 checkerboard images.

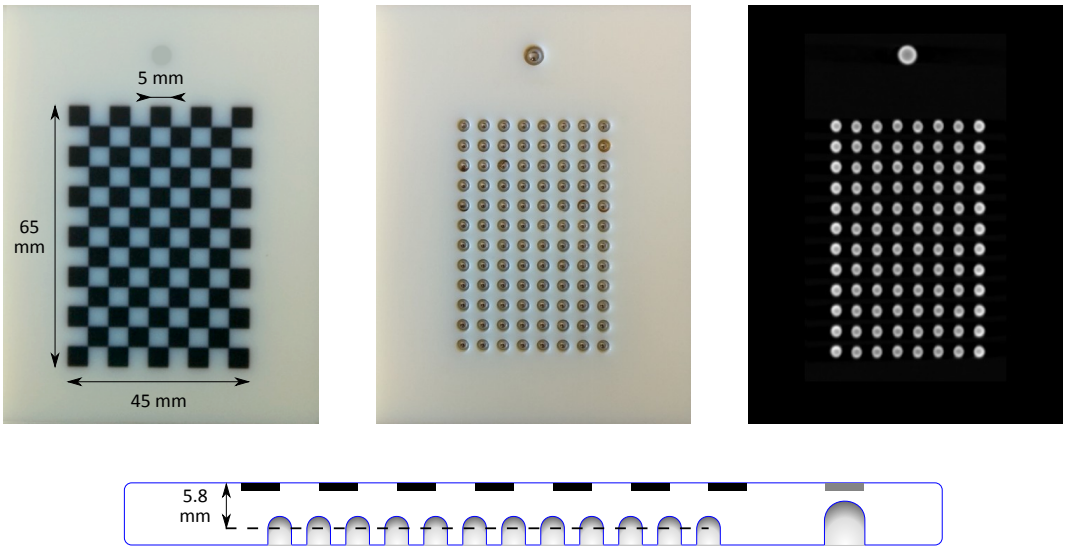


Figure 4.11: On the left is a front picture of the 3D-printed checkerboard. At the middle, a back picture shows the 3 mm steel balls inserted underneath the corners. On the right, the same balls seen from the CT image. On top of the checkerboard, a bigger 5 mm ball is inserted, observable from both CT and endoscope, which resolves ambiguity in the orientation. At the bottom, a cross-section shows the distance between the center of the balls and the surface.

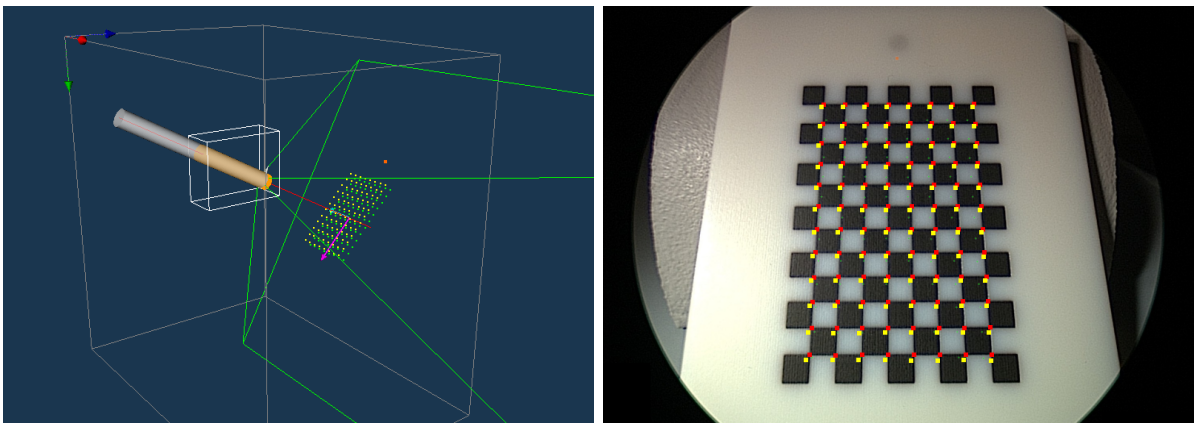


Figure 4.12: On the left, a rendering of the scene shows the outline of the C-arm volume (gray), the virtual camera placed based on our method and its frustum (green). The green dots are the centers of the segmented steel balls, which we use to infer the position of the actual checkerboard corners (yellow) by a known translation. On the right, these same points (yellow) are projected in the endoscopic view for comparison with the corners detected in the image (red). In this experiment, the registration error is $20.5 \pm 3.9p$.

Results - We performed this experiment during 17 sessions of acquisition. Before each session, the zoom and focus of the camera were arbitrarily set. Moreover, we used five endoscopes of the same model (Karl Storz Hopkins II), which may slightly vary in optical properties and

shape, due to manufacturing. For each session, up to three acquisitions were performed, each time with a different pose for both the endoscope and the radio-opaque checkerboard. Thus, a total of 34 experiments have been achieved and the registration errors are reported in Fig. 4.13.

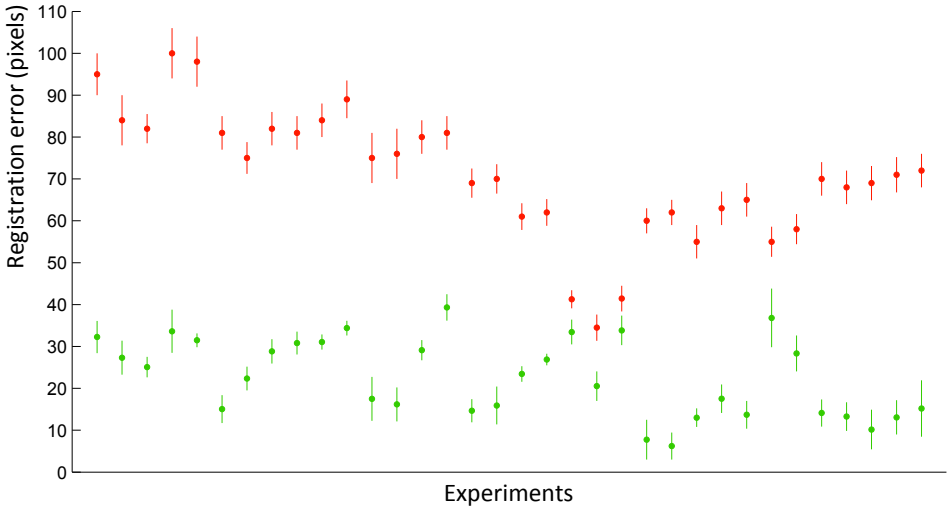


Figure 4.13: This graph shows the average registration error (average and standard deviation) across all 96 checkerboard corners for each of the 34 experiments. Results are shown without correction of the optical center shift (red) and with (green). One can distinguish error groups, which correspond to the different sessions.

According to the results, correction of the optical center shift brings a constant and significant improvement in accuracy. Across all 34 cases, registration errors then range between 6 and 40 pixels, with an average at around 23 pixels. The usage of pixels as a measurement unit is justified here as the registration occurs in the image plane, not in world space. An estimation of the accuracy in millimeters is feasible but obviously depends on the depth in the scene. Nonetheless, knowing the dimensions of the checkerboard pattern, we estimated that 1 mm was equivalent to 13 pixels at a nominal distance (around 70 mm), as in Fig. 4.12. As a result, we can extrapolate that the accuracy of our method is thus far comprised between 0.5 and 3.5 mm. This level of accuracy is already on par with most current laparoscopic AR methods (see Section 3.5) and may be satisfactory for surgical purposes depending on the considered margins.

4.5 Discussion

Based on the knowledge of the structure of the endoscope, we resort to a simple cylindrical template. The interface between the carbon dioxide and the metallic surface is sharply marked in the intraoperative 3D data but is still slightly smoothened. Moreover, inhomogeneous inner structures such as lenses may be visible to X-rays. Thus, a more advanced template that would also model these aspects could be used to increase robustness during fitting, but designing such a template is more complicated and would make the approach more device-dependent.

Regarding the ground truth accuracy provided by the radio-opaque checkerboard, it could be argued that we extract the steel balls from the volume image with a fixed threshold,

which may reduce the accuracy. Tuning its value yields slight changes in the position of the balls which in turn displaces the inferred checkerboard corners in the volume rendering. Nonetheless, since these uncertainties are of less than a fraction of a millimeter, they do not significantly impact the ground truth accuracy, even combined with the imprecision of the 3D printing (0.1 mm).

The requirement of an accelerometer to determine the roll angle of the endoscope may come as a constraint. Indeed, most laparoscopes are currently not equipped with such a feature. Nonetheless, building a prototype is more or less feasible depending on the level of integration and considered accuracy. Externally attaching the chip onto the body of the camera is simple, but at the cost of making the proximal part larger. Moreover, the unknown relationship between imaging sensor and accelerometer would require a one-time calibration. Our approach of internal integration is more seamless and does not require a calibration, but also requires more advanced manufacturing.

As shown by our quantitative experiments, image correction with respect to the optical center is necessary. The estimation of this intrinsic parameter is included in all calibration methods, but its accuracy is well-known for being limited and thus so is our registration method.

4.6 Conclusion

In this chapter, we have presented a new simple paradigm to register intraoperative 3D data with the endoscopic image, which relies on the inclusion of the endoscope tip inside the 3D field of acquisition. An analysis of the voxels allows to precisely determine all the camera extrinsic parameters with respect to the intraoperative 3D data, except the roll angle, given by an integrated accelerometer. For the endoscope localization, we have shown that a simple threshold-based approach makes the accuracy sensitive to the threshold value. Thus, we propose a 3D cylindrical template-based method, which is more stable.

During quantitative experiments, our method has shown a promising accuracy ranging from 0.5 to 3.5 mm throughout many evaluations using a dedicated calibration object. This magnitude of accuracy is already satisfactory for various kinds of laparoscopic interventions.

However, throughout the various sessions of experiments on the checkerboard, we realized that the misalignments in registration were similar within each session, in other words for a same camera calibration (see Fig. 4.13). We hypothesize this results from another parameter related to the camera configuration, that has not yet been taken into account. More specifically, we suspect that the optical axis and the axis of revolution may differ, even in straight centered endoscopes. For the sake of further improving the accuracy of our method, we investigate this idea in the next chapter and provide a solution to compensate this phenomenon.

Chapter 5

Registration refinement and robustness study

The hypothesis of coincidence between the endoscope's axis of revolution and the optical axis is often violated by structural misalignments, eventually leading to inaccuracy in our registration. As a result, we propose a novel, simple and yet effective way to compensate structural misalignments in the image plane. By analyzing the motion of a square tube rotating around the tip of the endoscope, we can better determine the relationship between the endoscope body and the image. This calibration step has been applied to the same validation data sets than in Chapter 4 and results show a large improvement in registration accuracy. Finally, we performed additional tests to identify the most unstable parts of our method and evaluate its overall robustness.

5.1 Realistic consideration of endoscope structure

In Chapter 4, we explained that in most cameras, endoscopes included, the optical axis Σ does not necessarily intersect the imaging sensor at its center C, but rather at a slightly off location called the principal point, henceforth called C_Σ. This offset between C_Σ and C is typically estimated during camera calibration as part of the intrinsic parameters. In our method, we have positioned the virtual camera on the axis of revolution Δ of the endoscope, which we hypothesized is equal to Σ. In this case, the offset can be compensated for by a 2D translation $\overrightarrow{CC_\Sigma}$.

However, it is important to note that this hypothesis may be violated and that often $\Sigma \neq \Delta$. As shown in Fig. 5.1 (a), such a misalignment can be large and intentional (on the left) or seemingly nonexistent, yet still present due to limited manufacturing accuracy (on the right). At any case, applying $\overrightarrow{CC_\Sigma}$ in the registration process is irrelevant if $\Sigma \neq \Delta$, as illustrated in Fig. 5.1 (b). Indeed, given the resolution of the sensor (1080p), even a few micrometers of misalignment between Σ and Δ would result in several pixels of registration error.

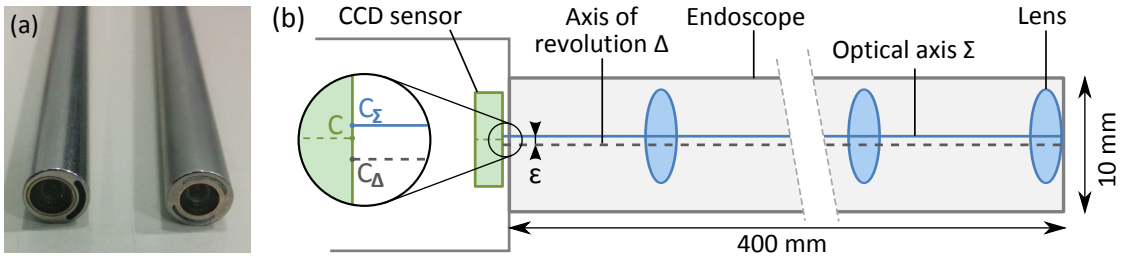


Figure 5.1: Anatomy of a laparoscope. (a) Picture of two endoscope tips: a decentered one (left) and a centered one (right). (b) A simplified cross-section of a laparoscope presents a misalignment ϵ between the optical axis Σ and the axis of revolution Δ, resulting in an registration error in translation in the image plane.

In light of this, in order to align both virtual and real image centers, we need to determine the geometrical relationship between the endoscope and the imaging sensor. We do not attempt to replicate the reality by actually moving the virtual camera away from Δ in world space, as this would be complex to estimate (micrometric scale). We rather measure the effect from the misalignment in the image plane, where the registration occurs. Hence, we aim at locating C_Δ in the image. Then, we can compensate for the offset by translating the virtual image by $\overrightarrow{C_\Delta C}$ in the final registration. However, this additional calibration step needs to be performed intraoperatively, as various parameters such as zoom and position of the light cable with respect to the camera influence the outcome (see Section 5.4).

Since the virtual and real cameras have different origins (the former is located on the endoscope axis and the latter on the optical axis), applying $\overrightarrow{C_\Delta C}$ can also be interpreted as a 2D correction of a perspective effect. However, the discrepancy between Σ and Δ is very small in both distance (a few millimeters at most) and parallelism compared to the dimensions of the scene (see Appendix E). Thus, the orthographic projection of the virtual camera is very similar to the perspective of the actual camera. Also, since this correction is operated from the center of the image, the approximation is all the more accurate in this area, which is where the region of interest usually lies.

5.2 Determining offset between endoscope and CCD sensor

Endoscope spinning method

A seemingly simple and intuitive way to determine the projection C_Δ of the axis of revolution Δ onto the image plane is to make the endoscope rotate on itself. Thanks to a holding arm, we make the endoscope rotate around Δ while recording a checkerboard pattern. Using the same corner detection technique that of the validation protocol, we extract the trajectory of each corner along the rotation of the endoscope. After performing a circular regression, we can then estimate the circularity of the trajectories and their centers, ideally located at the invariant point (see Fig. 5.2), which should also correspond to C_Δ.

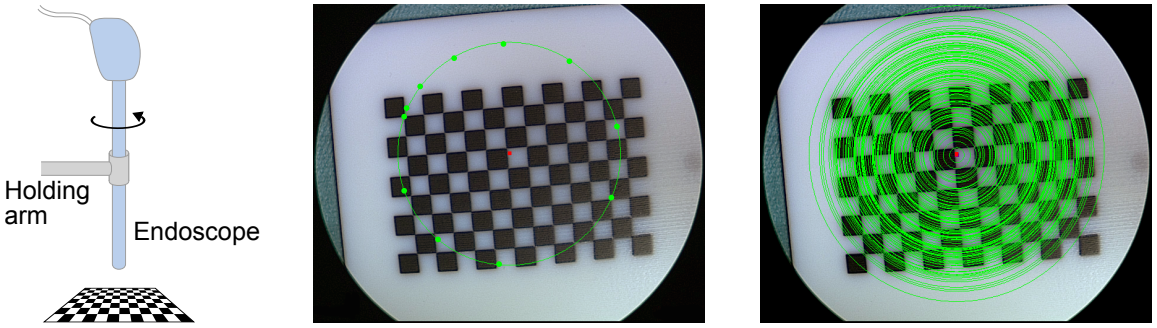


Figure 5.2: Endoscope spinning. The endoscope rotates around its axis of revolution Δ above a checkerboard pattern (left). The middle image here displays the locations of the top left inner corner in the pattern at successive times during the rotation (green points). The circle that fits these points best (green) is also shown as well as its center (red), which corresponds to the invariant point in the rotation. By performing so for each of the 96 corners, one can thus estimate the circularity of the trajectory during the complete spinning (right) and eventually the 96 centers of rotation.

We applied this protocol over four different sessions of experiments. Each session consisted of a single camera configuration and six or eight rotations. The results are provided in Fig. 5.3-left. One can notice the scatteredness that occurred over three out of the four sessions. Such variations in the determination of C_Δ is due to the physical difficulty of manually and yet accurately applying a rotation to the endoscope, while maintaining it at a constant height, with a simple holding arm. Such a problem may be alleviated by the introduction of a more complex arm, but this would defeat the purpose of providing a practical and affordable solution. This instability unfortunately makes it incompatible with the accuracy requirements of our registration, which aims at being refined below 15 pixels (around 1 mm in the scene).

Nonetheless, these experiments allowed us to prove again (in addition to Appendix E) the near perfect parallelism between the optical axis Σ and the axis of revolution Δ. Indeed, despite the rotations being relatively unstable, the resulting circles were very much concentric and their centers well aggregated. In Fig. 5.3-right, the standard deviations of these centers are shown to be quite small, with a majority below 1.5 pixels and a maximum at 4 pixels.

However, for the sake of robustness, we propose a more practical and stable method that could be considered as the reverse approach to the current endoscope spinning. Instead of making the endoscope rotate around the scene, we developed a very simple way of making a certain pattern rotate around the endoscope, with the help of a square tube.

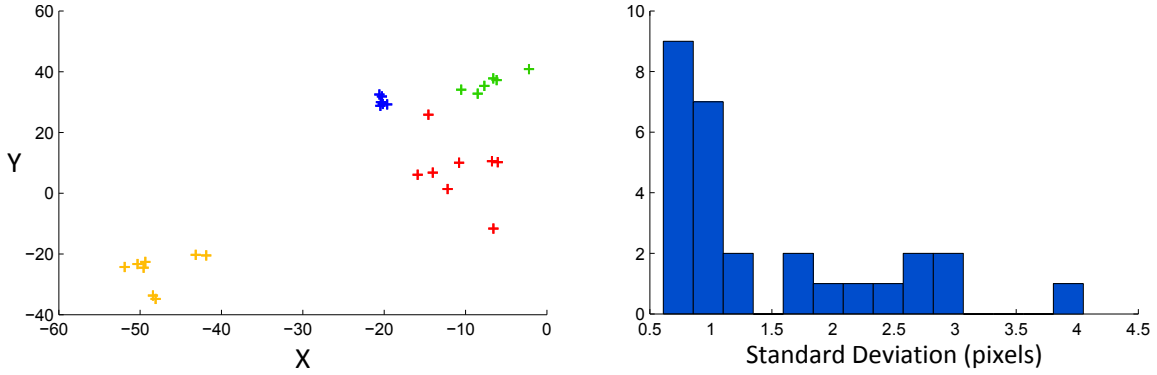


Figure 5.3: On the left, the various C_Δ obtained for different camera configurations (one color per configuration). On the right, the distribution of standard deviations of circle centers from C_Δ, across all rotations.

Square tube method

Provided that the rod of the endoscope is a right circular cylinder, all the diameters of the tube intersect on its axis of revolution Δ. Therefore, a practical way of estimating Δ is to determine several of these diameters in the endoscopic image and calculate their intersection. A simple technique to do so is to record images of a square-profiled tube rotating around the endoscope tip, as illustrated by Fig. 5.4. As long as the rod is maintained stuck in a corner, which is mechanically stable, the corresponding diagonal is also a diameter, that is then obtained by extracting the corners.

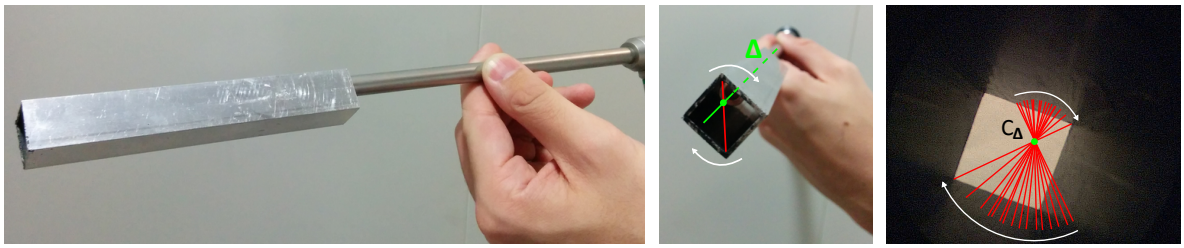


Figure 5.4: On the left, a square-profiled tube rests on the distal part of the endoscope. In the middle, the line passing by the top and bottom corners (red) intersects the sought axis of revolution Δ (green). On the right, during rotation of the square tube, these diagonals intersect at the sought point C_Δ in the image.

Corner extraction inside the square tube

To accurately determine the corners, the inside of the square tube is painted in matte black, so that the images obtained are very easily segmentable (see Fig. 5.5). First, a bilateral filter is applied to the image to reduce the noise while preserving the edges (Fig. 5.5-b). Second, we use an edge detector based on Canny filter to produce the binary edge map of the image (Fig. 5.5-c). Third, we apply the probabilistic Hough transform to extract the strongest segments from the map and cluster them by similarity of direction and location(Fig. 5.5-d). Finally, we average each cluster of segments by averaging the corresponding vectors and origins, in order to obtain the square edges and thus the corners and diagonals (Fig. 5.5-e).

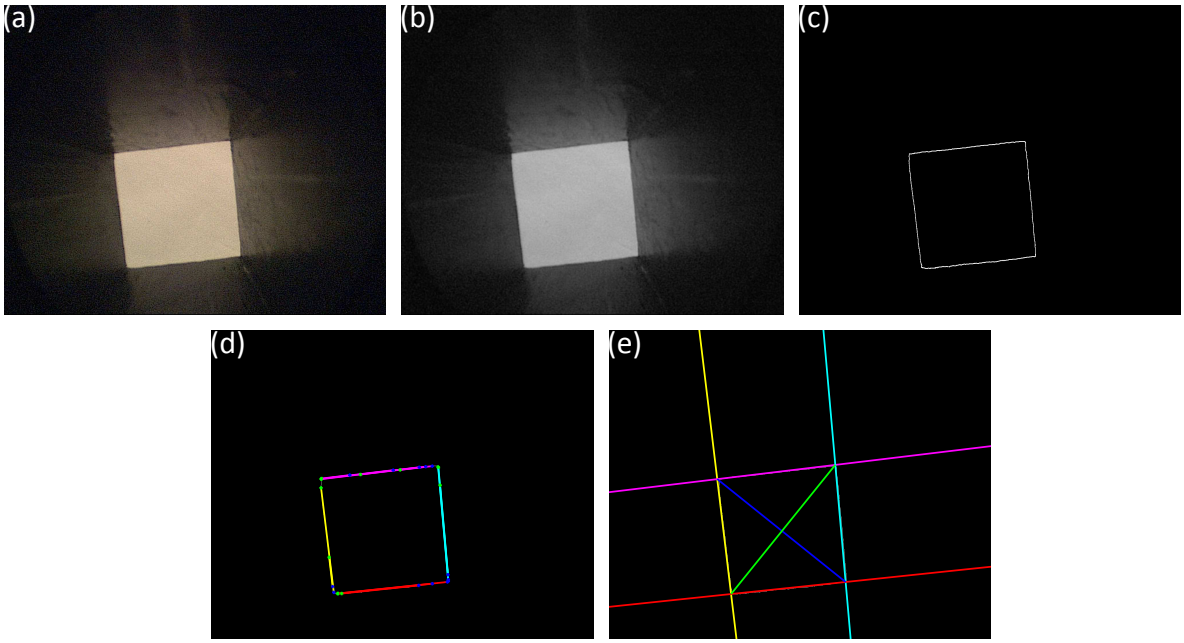


Figure 5.5: Our method of square edge extraction: (a) Original image; (b) Bilateral filtering; (c) Edge detection based on Canny filter; (d) Probabilistic Hough transform and separation of segments into four clusters; (e) Averaging of segments to obtain the edges and then the diagonals.

Determination of the intersection of diagonals

So, while the square tube slowly rotates on the endoscope, we extract the diagonals in the images and associate them from one frame to the next simply by considering their orientation similarity. This results in two sets of intersecting diagonals (see Fig. 5.6). To determine which set corresponds to the sought projection, we calculate the position of the point that minimizes its distance to all lines for each set. In the correct case, all the lines should intersect roughly at the same point and therefore the deviation should be the smallest.

Using a least-square formulation, we calculate the intersection of the diagonals as the point that minimizes its distance to each line. Let us consider a set of N seemingly intersecting straight lines, each defined by a point P_i and a normalized leading vector v_i. The squared distance of a point P to the i-th line is thereby:

$$d_i^2 = |P - P_i|^2 - ((P - P_i)^\top v_i)^2 \tag{5.1}$$

If we consider the vector derivatives of Eq.5.1, we get:

$$\begin{aligned} \frac{\partial d_i^2}{\partial P} &= 2(P - P_i - ((P - P_i)^\top v_i)v_i) \\ &= 2(I - v_i v_i^\top)(P - P_i) \end{aligned} \tag{5.2}$$

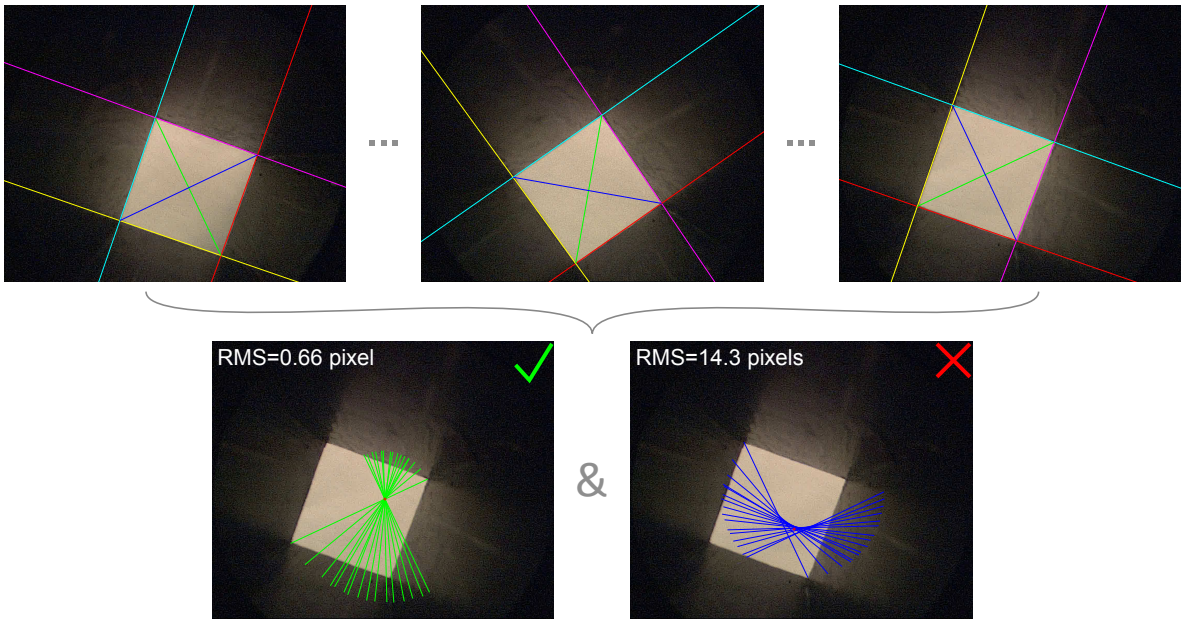


Figure 5.6: At the top, the edges and thus the diagonals of the squares are detected in a sequence of images, resulting in two sets of lines, the green ones and the blue ones, as illustrated at the bottom. Using a least squares method, we can calculate, for each set, the point that minimizes its distance from the lines. The set with the smallest discrepancy (here the left one) is the closest to a perfect intersection of all the lines at the same point, which is considered to be C_Δ.

The point C_Δ that best approximates the intersection according to the least-squares sense is also the null space of Eq.5.2 for every line i:

$$\sum_{i}^{N}(I - v_i v_i^\top)(C_\Delta - P_i) = 0 \tag{5.3}$$

Therefore, we can obtain C_Δ from Eq.5.3 as:

$$C_\Delta = [\sum_{i}^{N}(I - v_i v_i^\top)]^\dagger \sum_{i}^{N}(I - v_i v_i^\top)P_i \tag{5.4}$$

As a result, we are able to determine and compensate for the misalignment between the axis of revolution and the CCD sensor. Thus, the endoscopic view can be augmented with an accurately registered virtual image of the model. In the next section, we present results based on the same protocol and data sets than in Chapter 4. A side-by-side comparison of the two outcomes expose the accuracy improvement which validates our square tube calibration.

5.3 Experimental results

Using the same protocol and data than in Section 4.4, we performed our validation routine with the radio-opaque checkerboard, but this time completed by the square tube calibration. The actual rotation takes between 6 and 10 seconds to operate and the calculations to determine C_Δ last between 3 and 6 seconds depending on the number of frames captured during the motion. Accuracy results are displayed in Fig. 5.7.

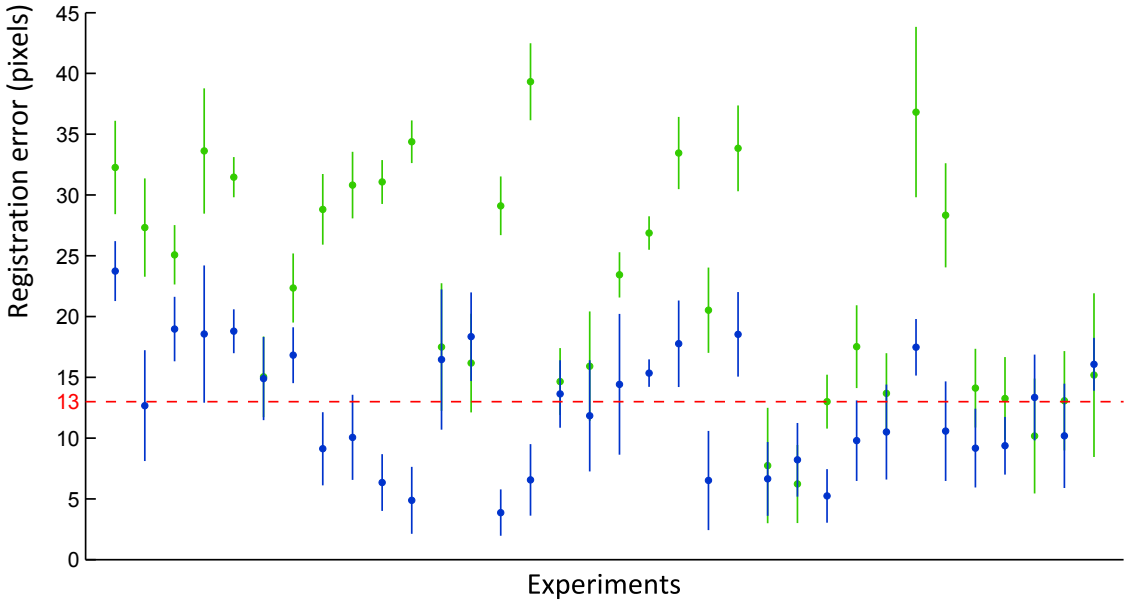


Figure 5.7: This graph shows the average registration error (average and standard deviation) across all 96 checkerboard corners for each of the 34 experiments. Results in green are those from Section 4.4, based on the hypothesis of coincidence between Δ and Σ. Results in blue are those obtained without this hypothesis and after square tube calibration.

The square tube calibration improved the registration accuracy in 30 out of 34 cases. Moreover, the average registration error across all cases has been decreased from 22.71 pixels to 12.50 pixels. Thus, if we again consider that 13 pixels corresponds to a millimeter in the scene at nominal distance (around 70 mm), our AR method has achieved a sub-millimetric accuracy in the majority of the cases. Likewise, the largest error obtained (24 pixels) represents only 1.8 mm in the scene. Therefore, according to the surgeons at our institution and current standards for intra-abdominal interventions, these errors are sufficiently small to fulfill AR purposes for surgical procedures. Nonetheless, clinical requirements also mandate a solid robustness and stability for this accuracy. This aspect of our method is considered and tested in the next section.

5.4 Precision and robustness

Our AR method requires several pieces of data and relies on various tasks. As a result, there are many parameters that can influence the registration such as the depth of insertion of the endoscope into the CBCT field and the square tube, the camera zoom, focus and light cable position. In order to better evaluate the precision of our method, in this section, we quantitatively estimate the stability of each parameter and their impact on the final outcome.

Depth of insertion into CBCT field

We have measured the relation between the registration error and the depth at which the endoscope is inserted into the volume, in order to provide some kind of recommended minimum insertion guideline. The results are shown in Fig. 5.8 and were performed on 12 various endoscope positions. One can see that all registration errors stabilize beyond 17 mm (34 pixels with a 0.5-mm spacing), which then becomes our recommendation for the tip insertion depth to ensure the accuracy of our method.

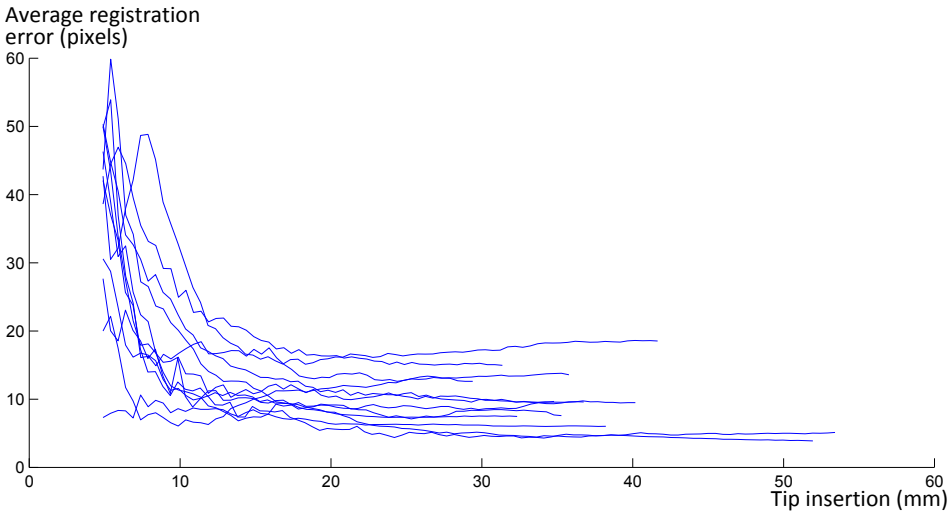


Figure 5.8: For each of 12 experiments, the average registration error in pixels is displayed with respect to the tip insertion depth in millimeters at which the endoscope is inserted in the volume image.

Square tube calibration

Concerning the square tube calibration, the determination of C_Δ may be influenced by the insertion depth in the square tube and the endoscope settings that are the zoom, the focus and the position of the light cable.

First, one may wonder by how much should the endoscope be inserted in the tube. Since the endoscope is calibrated to capture the scene at a certain distance, it is naturally inserted so that the square end of the tube appears sharp in the image. Thus, the output plane is at that same nominal distance than the scene to be imaged and the 2D correction of the perspective effect is accurate. Nonetheless, for the sake of robustness, we measured the variations of C_Δ at 10 different insertion depths around the depth of focus and we obtained a standard deviation of (0.9, 0.9) pixels along the image X- and Y-axes, which is negligible. This may also be considered as a proof that the perspective difference between the virtual and the actual camera is little, as stated in the last paragraph of Section 5.1.

Second, parameters bound to the endoscope and the camera such as the zoom, focus and the light cable position may have an impact on the registration (see Fig. 5.9). Therefore, we also measured the variations of C_Δ with respect to these parameters. For 11 different values across the focus field, the standard deviation of C_Δ is (2.0, 0.9) pixels along the image X- and Y-axes. For 16 different values of zoom between a medium zoom and a complete zoom-out, the standard deviation of C_Δ is (4.3, 7.5) pixels along the image X- and Y-axes. For 11 different positions of the light cable, the standard deviation of C_Δ is (16.1, 21.1) pixels along the image X- and Y-axes. Hence, while changing the camera focus has little impact on the 2D correction, variations of the zoom and the light cable position significantly influence the square tube calibration and thereby the augmentation. As a result, a new calibration ought to be performed each time one of these parameters is changed.

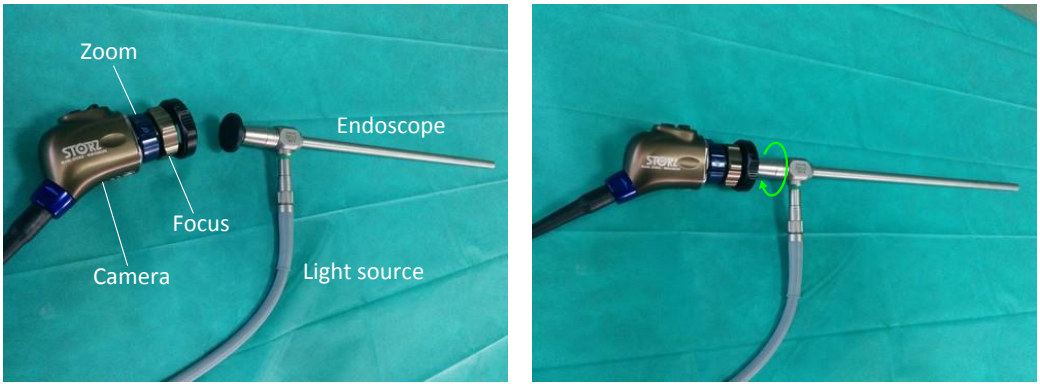


Figure 5.9: Pictures of a separable endoscope. The endoscope and the camera can be separated (left) and attached (right) through a joint that allows a spinning rotation of the endoscope (green arrow). The light cable can thus be placed according to the surgeon's desire during the intervention.

Precision per task

Now that the influence of the main parameters has been established, we focus on the precision and stability of each part of the process. To assess how much each part affects the final output, we performed each step of the complete process (camera and square tube calibrations, CBCT acquisition and endoscopic image capture) for three series of ten times. For each series, we aimed at introducing as much variation as possible in terms of scene conditions. To this end, the endoscope and the radio-opaque checkerboard have been placed differently, the lighting of the scene has been changed and checkerboards of various sizes have been used for the camera calibration. Table 5.1 reports the average errors and the standard deviations of registration in pixels measured as only one step of the process varies. This concerns the square tube calibration, the camera calibration, the endoscopic image (corners localization) and the CBCT volume (endoscope extraction and corners localization).

From the standard deviations reported in Table 5.1, one can notice that the camera calibration does not significantly affect the registration accuracy, probably due to the reliability of Zhang's method for the endoscopic camera calibration and the robustness of the corner detection method. The latter may also explain why variations in scene lighting captured by the endoscopic image does not impact the registration.

Successive acquisitions of a same scene by the 3D CT imaging system may yield non-

Acquisition series	1	2	3
Square tube calibration	16.25 ± **7.51**	15.11 ± **6.23**	13.34 ± **7.16**
Camera calibration	13.93 ± **0.28**	11.01 ± **1.01**	6.43 ± **0.36**
Endoscopic image	13.81 ± **0.07**	11.76 ± **0.09**	6.75 ± **0.11**
CBCT volume	14.76 ± **0.79**	14.80 ± **1.84**	6.76 ± **2.39**

Table 5.1: Averages and standard deviations of the registration error (in pixels) observed when only one part of the process changes ten times and across three different setups. While one step changes, all the others are assigned the values corresponding to the best result in terms of accuracy.

negligible fluctuations in the intensity of the voxels. Since our method relies on voxel intensity to estimate the endoscope pose in the volume image, we also logically observe that this part of the process affects the final output of our method. Nonetheless, so far, these variations represent at most only tenths of a millimeter in the augmented scene and may be ignored for the considered application.

On the contrary, various square tube calibrations may change the augmentation by up to 14 pixels which corresponds to almost a millimeter in the scene. Therefore, despite generally improving the accuracy of our AR method, this routine has also by far the most impact on the registration and ought to be performed with caution. More specifically, while making the rotation, the operator has to make sure that the square tube is correctly in contact with the endoscope along its edge.

Overall, we have shown with quantitative results that our AR method is quite precise and robust across its complete protocol. According to our experiments, the insertion depth of the endoscope into the square tube matters little, making such a manipulation more practical to perform than the spinning approach for example. Practicality is important as this calibration has to be performed intraoperatively. Indeed, the surgeon typically sets the camera parameters himself during the operation and parameters such as the zoom and the light cable position influence the registration. However, the square tube calibration has proven to be robust to focus change, so this parameter may be modified intraoperatively without impacting the augmentation.

5.5 Discussion

In a sense, the considered intrinsic parameter resulting from the misalignment between axes could also be determined using a common hand-eye calibration method. However, due to the absence of external tracking, the relationship between a calibration object (e.g. our radio-opaque checkerboard) and the camera could only be determined by means of an additional CBCT acquisition including the endoscope tip. If performed before the operation, this acquisition may not be useful as the zoom, focus and position of the light cable are not yet set as desired by the surgeon. If performed once the patient is on the table, a new acquisition is not possible without moving the patient. Either way, this would lengthen the surgical workflow significantly more than with the proposed method using a square tube.

In our work, we used a regular square tube, but our approach is also compatible with other polygonal shapes as long as the bisector of the corner in contact with the endoscope can be determined. Moreover, other features may be added to the tube in order to further improve

the method. For example, a magnet and/or a spring would help to maintain a permanent contact between the tube and the rod. Likewise, a dent inside the tube would help keep the tube in place during rotation and improve the mechanical stability. Finally, some small indicator painted inside next to the top corner would lift ambiguity between corners in the image. Thus, a more complex design of the square tube could improve the stability of the calibration, but would require more engineering and hence make it less available.

5.6 Conclusion

In this chapter, we have raised the issue of misalignment between the optical axis and the axis of revolution in a straight endoscope. Intentional or due to limited manufacturing accuracy, this phenomenon changes the approach to re-center the virtual camera with the center of the imaging sensor and compels for a novel calibration revolving around the axis of revolution and its projection onto the image plane. We have shown that making the endoscope spin on itself is too unstable and thus we resort to a simple original method based on a sterile square tube.

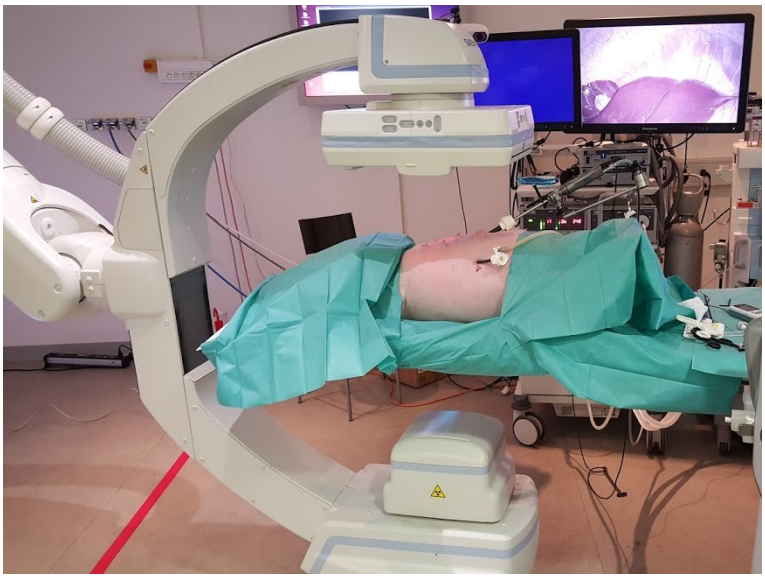
Quantitative experiments show that using this method significantly improves the accuracy of the registration well within the requirements for surgical purposes. We also demonstrated its robustness with respect to the various parameters that come into play. Its process is fast, simple and practical enough to be carried out intraoperatively.

The square tube calibration shall be performed with caution as its influence on the accuracy is prominent. If done so, one can then obtain precise augmentations like the ones performed on *in vivo* data and presented in the next chapter.

Chapter 6

Validation on *in vivo* data

This chapter reveals the true clinical potential of our method by applying it to in vivo *data acquired during laparoscopic interventions on pig. Various intra-abdominal scenes exposing the liver and the bladder have been augmented with volume rendering or manual segmentations. A qualitative appreciation of the accuracy of our method is allowed by considering the good alignment of salient anatomical clues. We also provide another means of visual validation by projecting the laparoscopic image onto the extracted surface of the abdominal cavity.*



6.1 Augmented Reality

Protocol

So as to assert the clinical feasibility and potential of our method, we performed six sessions of acquisition on pigs. For each session, the protocol followed is close to that of an intervention on human and is illustrated in Fig. 6.1. As soon as the pneumoperitoneum is done, a trocar is placed in order to allow intra-abdominal access to the endoscope. The latter is inserted, some navigation is performed and up to 3 additional trocars are placed for an easy access to the liver using laparoscopic instruments. Once the surgeon has set the camera parameters to his/her convenience, we performed a camera calibration using a classic checkerboard. Taking advantage of the endoscope being out of the cavity, the square tube calibration is also achieved. Lighting conditions in the OR are typically not good for both calibrations. In order to avoid specularities and reflections, the endoscope light is dimmed, but this makes the images noisy and poorly contrasted. Commonly, we address this issue by shedding light again onto the calibration screen, but from the ceiling-mounted scialytic lamp. After the calibrations, the endoscope is inserted back in, placed with a satisfying view of the ROI and maintained in that position with a holding arm. The CBCT scanner (Artis Zeego) is brought closer to the operating table and a fluoroscopy is performed to assess that the endoscope is well within the field of acquisition. Moreover, a trial rotation makes sure that no collision would occur, especially with the holding arm. Finally, the actual 3D acquisition is run, which typically lasts 6 seconds. During this lapse of time, the pig's breathing, which is controlled mechanically, is held in order to limit anatomical motions that would result in blurred acquisitions. At last, a capture from the laparoscopic view (the augmentee) is also taken.

So far, our experiments have been performed offline given the difficulty of aggregating online all the data from the various systems onto a single machine. Nonetheless, we are progressing toward a fully functional implementation of our solution for hybrid ORs, as is explained in Section 8.2. So, the data are gathered for post-processing on the computer described in Section 4.4. These data consist of the volume image in DICOM format (see Appendix D), a video sequence for the camera calibration (typically in MP4), a video sequence for the square tube calibration and the laparoscopic capture of the scene in either PNG or JPEG. For both calibrations, successive frames are extracted at short intervals from the video sequences and all these data are automatically processed by our method. Within 20 seconds, the laparoscopic view is shown, augmented with the volume rendering as seen by the virtual camera. If no accelerometer was included in the camera, the roll angle is manually adjusted, otherwise the angle is directly read from an encryption in the laparoscopic image.

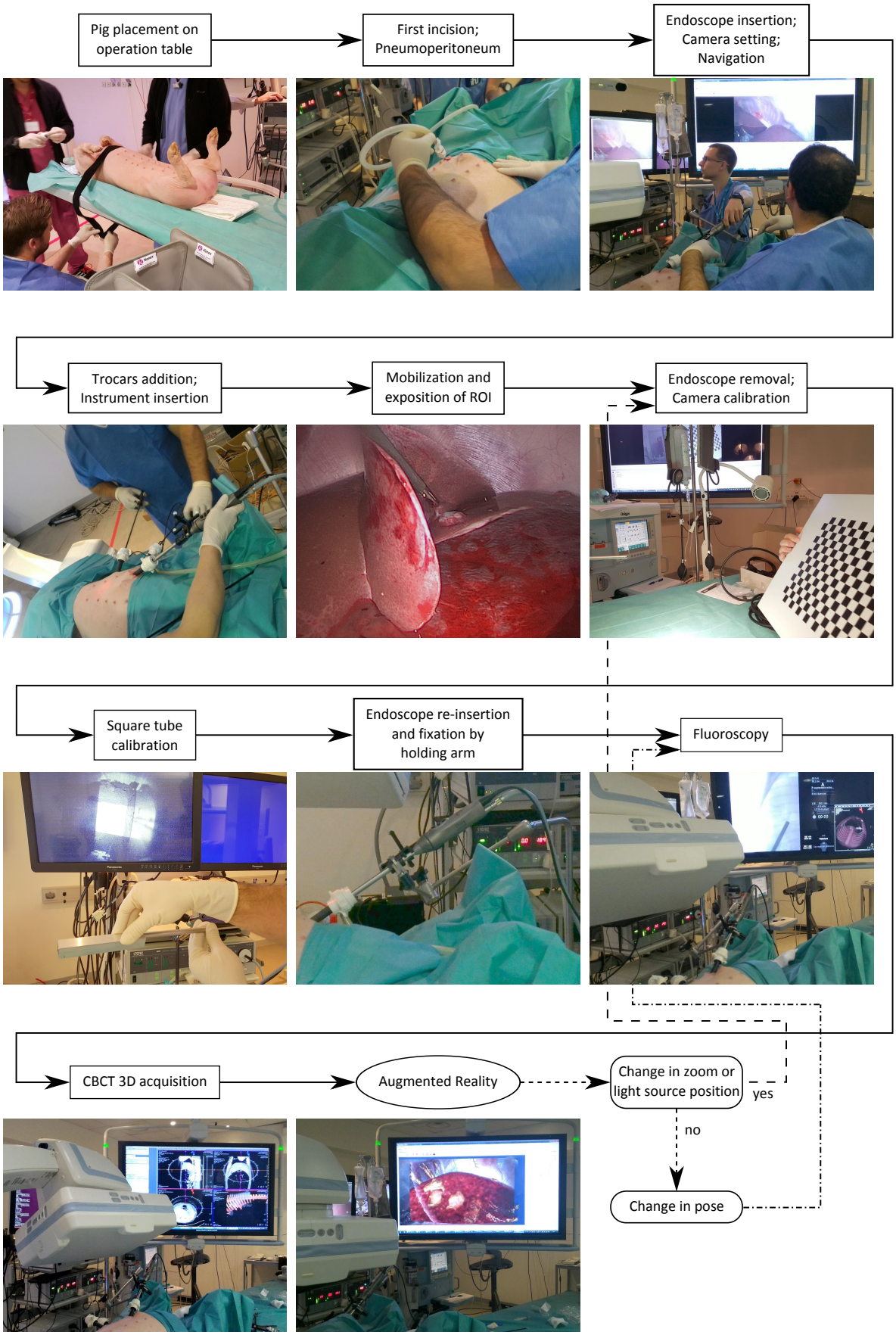


Figure 6.1: Workflow of our intervention on pig.

Results

Here follows three examples of thusly augmented scenes. One can qualitatively appreciate the accuracy of our method by examining the good correspondence between several visual cues such as contours, edges and landmarks in both the image and its augmentation.

Liver

In hepatic surgery, the complexity of an intervention commonly depends on the location of the tumor in the liver. Indeed, if the lesion is superficial and in the anterior part of the organ, vision and access are then simple since the endoscope is inserted through the navel. Such a case allows a good consideration for margins and limits the risk of damaging critical vasculature of the liver. On the other hand, if the tumor lies deeper in the organ, its localization is much more intricate via a mental alignment between the preoperative data and the laparoscopic view. Therefore, surgeons systematically resort to intraoperative ultrasound for assistance in spotting the tumor as well as surrounding vessels. Nonetheless, the relationship between the 2D US image and the preoperative 3D data has again to be mentally determined by the surgeon. Moreover, the estimated positions of the tumor are often faulted by the perspective from the endoscope. Hence, an AR solution would greatly benefit this kind of intervention by accurately displaying the position of the target in the laparoscopic view, thus also decreasing the surgeon's mental load.

In order to assert the efficacy of our AR method for hepatic interventions, adenomas have been artificially created on the liver of two pigs, as shown in Fig. 6.2. Porcine livers consisting of multiple relatively thin lobes, simulating a deep tumor therein is difficult. Nonetheless, the second tumor (right column in Fig. 6.2) has successfully been placed in the posterior part of a lobe, making it directly visible only after mobilization. After manual segmentation of the tumors, we performed our method and augmented the laparoscopic image with these lesions, thus serving as a visual validation for registration accuracy.

Bladder and ureters

In gynecologic surgery, interventions often consist in resecting tissues from the pelvic cavity (e.g. endometriosis, tumors, cysts). These operations are particularly difficult as numerous critical structures are hidden in the pelvic tissues, notably ureters, iliac arteries and nerves. Therefore, showing the position of such structures in the endoscopic view is appealing, especially since they essentially remain stationary before dissection. In order to prove the usefulness of our method in such context, we performed a few feasibility experiments on pigs.

As for human interventions, after the swine is placed on the operating table, an injection of contrast agent is performed, shortly followed by another one of Lasilix. The latter facilitates the flow of the agent from the kidneys to the bladder via the ureters, which then become highly visible in the 3D CBCT image. Provided a correct TF is used for the volume rendering, no further processing is necessary.

We performed our method as the endoscope was positioned in the pelvic cavity towards the bladder. Fig. 6.3 shows instances of augmentation of the bladder of two different pigs. Due to the contrast agent, the bladder and the ureters really stand out from the rest of the scene in the augmentation. The estimated position of the ureters have been verified by the surgeon using palpation, although they are slightly discernible in these laparoscopic images. While this may be the case in pigs, palpation is insufficient in humans as their ureters are typically covered by fatty tissues. Therefore, revealing the ureters in the laparoscopic view without any human interaction, as our method does, has a strong clinical potential.

Lungs

In thoracic MIS, the resection of lung tumors is a highly complex act, regardless to their location. Indeed, the intervention requires to deflate the concerned lung in order to create room for manipulating the instruments, much like the pneumoperitoneum for abdominal MIS. Since this artificial pulmonary collapse has to be performed under general anesthesia, it cannot also be achieved during preoperative scanning. Even so, intraoperatively replicating a preoperative deflation is impossible according to surgeons. Therefore, preoperative data are acquired while the lung is normally inflated, but the resulting anatomical information is no longer accurate during the intervention. Hence, the surgeon has to resort to endoscopy-controlled palpation to locate the tumor, which is a tedious and unreliable task as the lung moves and deforms under touch. In light of this context, an intraoperative guidance solution seems very promising for this kind of intervention.

After ethical approval, a first feasibility experiment has been postoperatively conducted on a patient's data. After he/she was placed on the operating table, the target lung has been deflated and an endoscopic camera has been inserted. Once care was taken that the tip of the endoscope was introduced into the scope of the CBCT scanner, an intraoperative 3D acquisition has been performed. Afterward, the laparoscopic view has been recorded for the rest of the intervention.

In post-processing, we have manually segmented the lesions in the two intraoperative 3D data, as illustrated in Fig. 6.4. The position of the endoscope with respect to this data is determined using our method, thus allowing the augmentation of the recorded laparoscopic sequence. We then asked the surgeon to outline, judging from the preoperative data, the position of the lesion in the video paused at the moment of the CBCT scan and without AR. After a comparison with the augmentation, a strong discrepancy has been found between the two estimations in both cases, as shown in Fig. 6.4. Watching the rest of the intervention recording, the surgeon eventually sided with our estimations. This proves how badly the deformation between preoperative data and the actual intraoperative state hinders the prediction of target location. Moreover, this experiment also showed the potential of our AR method in alleviating this problem in a hybrid OR. While manual segmentation was involved during this test, such lung lesions are typically visible in the scan images without contrast agents and provided a significant improvement in the volume rendering TF, this interaction might not be necessary any longer.

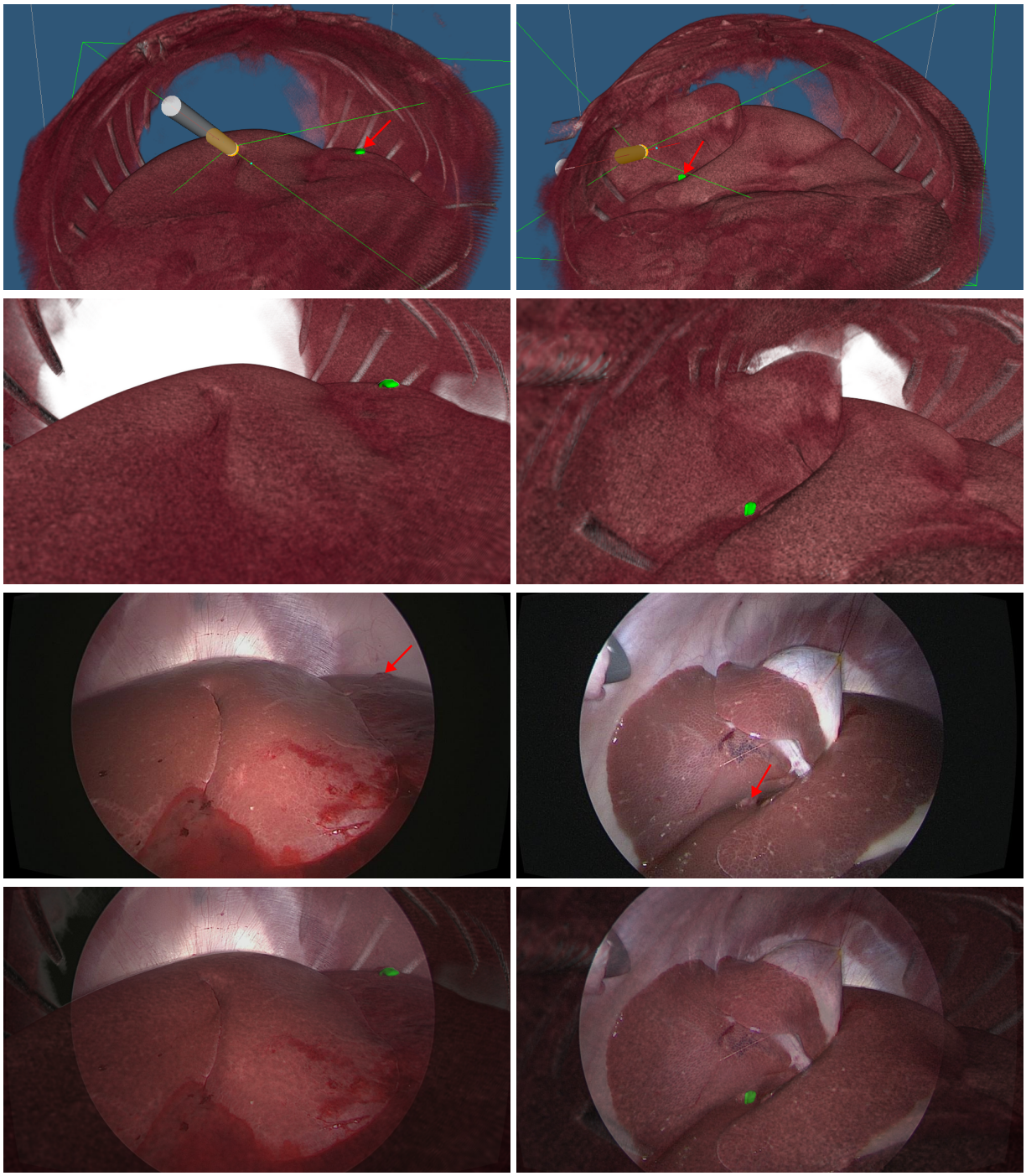


Figure 6.2: AR on liver. For two different scenes (one per column), displayed as volume rendering on top, the virtual image (mid-top) is projected onto the endoscopic image of the liver (mid-bottom) to augment it (bottom). In each case, an artificial adenoma has been manually segmented from the intraoperative volume image and is pointed by the red arrow.

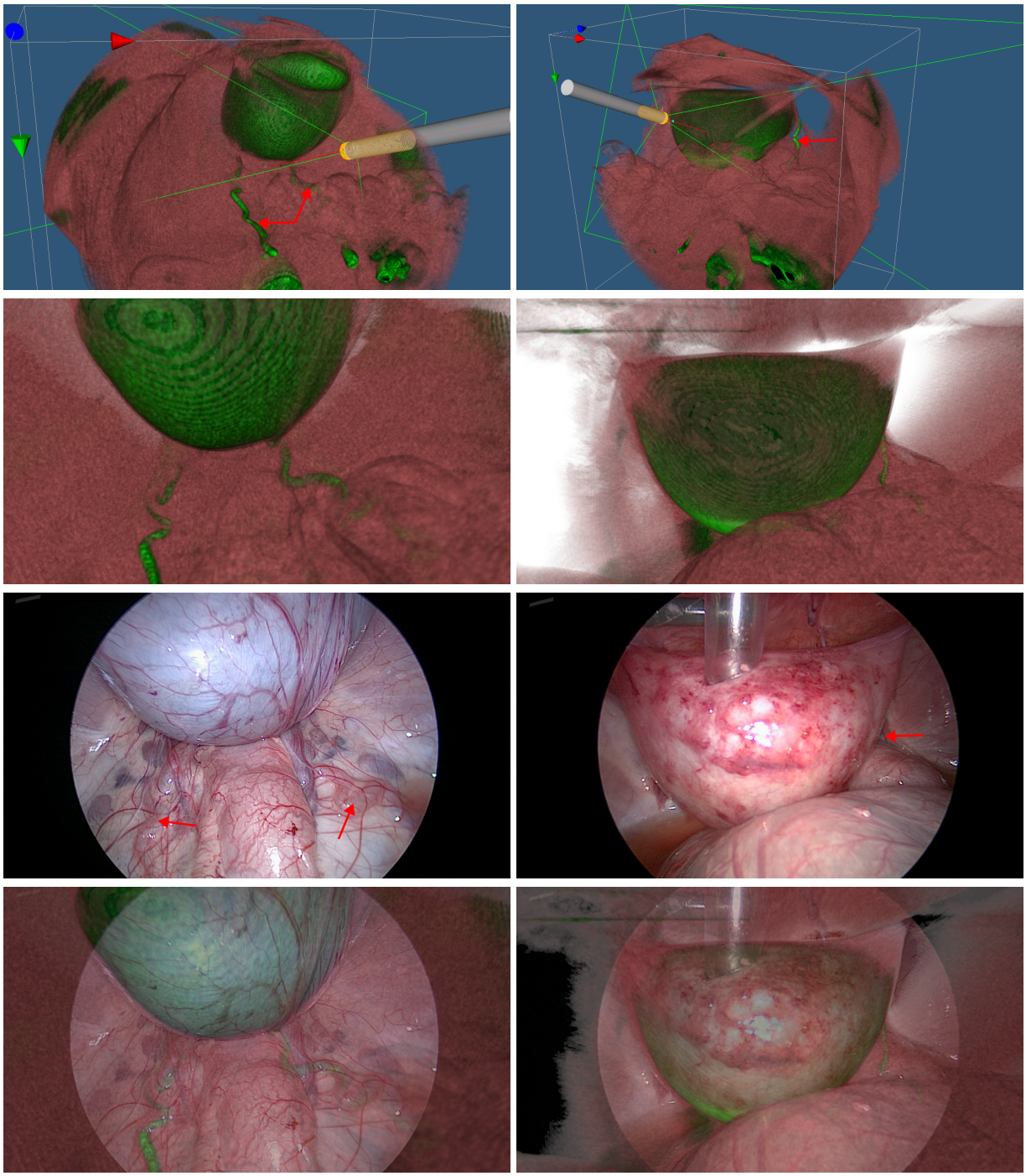


Figure 6.3: AR on ureters. For two different scenes (one per column), displayed as volume rendering on top, the virtual image (mid-top) is projected onto the endoscopic image of the bladder (mid-bottom) to augment it (bottom). In each case, a contrast agent enhances the contrast of the intraoperative data and, provided an appropriate TF, the ureters appear clearly in the augmentation, as pointed by the red arrows.

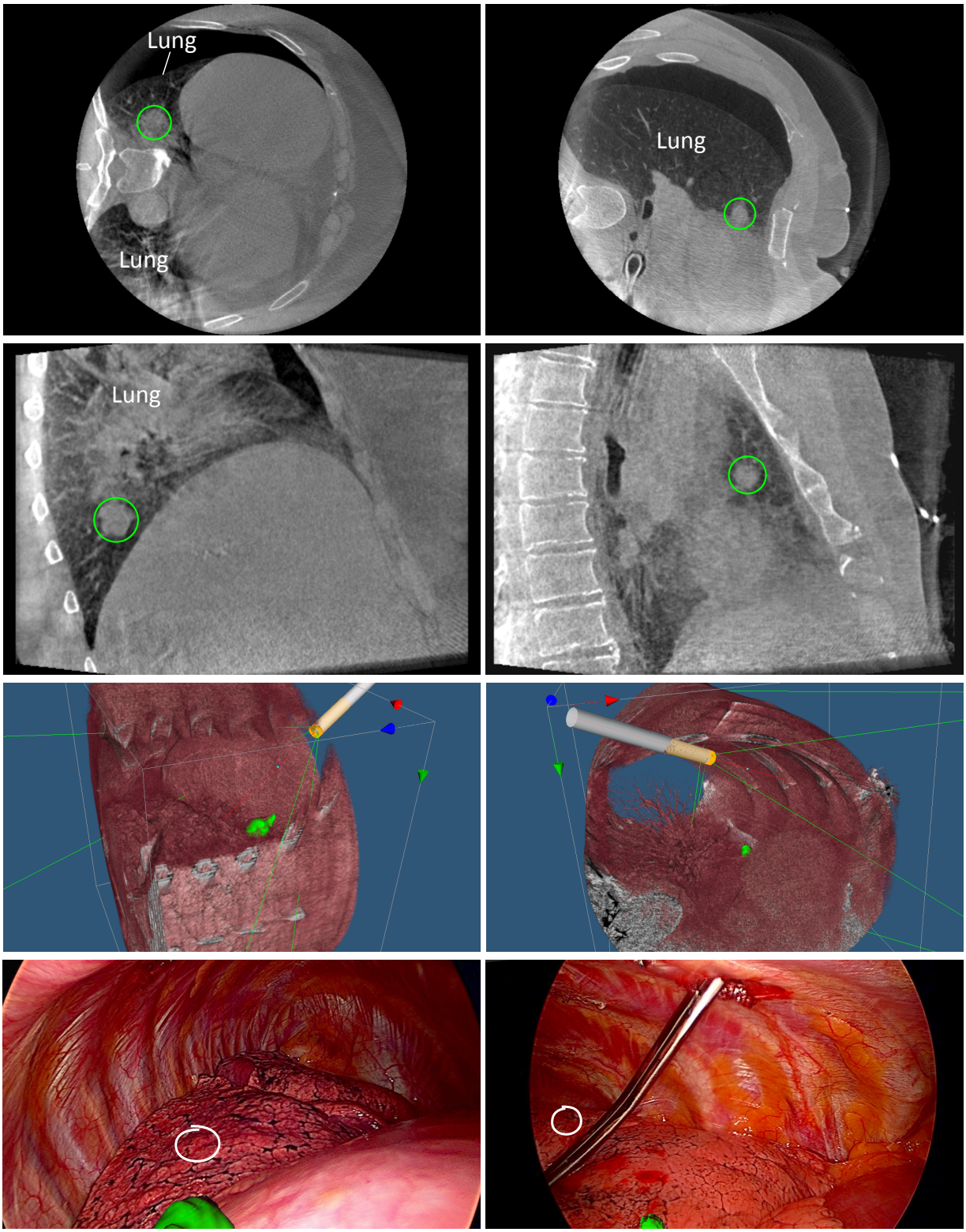


Figure 6.4: AR on lungs for two cases (one per column). The top and mid-top rows respectively show axial and frontal images of the lesions outlined in green. The mid-bottom row displays the segmented lesions (green) alongside the volume rendering of the scene and the virtual endoscopic camera determined by our method. At the bottom, one can observe the difference between the surgeon's prediction (white outline) and the augmentation (green).

Handling lens distortion

Due to the convex shape of their lenses, all endoscopes produce distorted images. A popular approach to approximate this deformation is to resort to the Brown-Conrady model, which encompasses both radial and tangential distortions. Assuming that both distortions are centered on the principal point (x_c, y_c), for a pixel (x_d, y_d) of the distorted image, its undistorted equivalent (x_u, y_u) is then given by the following infinite series:

$$\begin{aligned} x_u &= x_d(1 + k_1r^2 + k_2r^4 + k_3r^6 + ...) + (2p_1x_dy_d + p_2(r^2 + 2x_d^2))(1 + p_3r^2 + p_4r^4 + ...) \\ y_u &= y_d(1 + k_1r^2 + k_2r^4 + k_3r^6 + ...) + (p_1(r^2 + 2y_d^2) + 2p_2x_dy_d)(1 + p_3r^2 + p_4r^4 + ...) \\ &\text{with } r \text{ the distorted radius:} \quad r = \sqrt{(x_d + x_c)^2 + (y_d + y_c)^2} \end{aligned} \tag{6.1}$$

The radial and tangential distortions are parametrized by the coefficients k_i and p_i respectively. These coefficients are typically estimated during the camera calibration, for example following Zhang's method [Zha00] as in our case. Nonetheless, due to the polynomial form of the series, the influence of high-indexed coefficients is negligible. Therefore, for the sake of reducing the computation complexity, only $\{k_1, k_2, k_3\}$ and $\{p_1, p_2, p_3\}$ are considered in this model. Moreover, most lenses present a stronger radial distortion than a tangential one and thus, sometimes, only k_1 is determined. For our work, we made a compromise between accuracy and computational complexity by considering $\{k_1, k_2\}$ and $\{p_1, p_2\}$.

In the previous figures, all augmented scenes have been presented with the laparoscopic image undistorted using Eq. 6.1 and the virtual image unaltered. However, undistorting the laparoscopic image slightly decreases its quality and sharpness, possibly leading to a loss of the finest details such as tiny vessels. As a result, since the laparoscopic image is the surgeon's only means of vision onto the scene, such digital image deformation and its inherent interpolations preclude from obtaining FDA or CE certification for the proposed system.

So, in a effort to comply with this rule, we also adapted our method to perform the augmentation without undistorting the laparoscopic image. In turn, this requires to distort the virtual image to maintain the augmentation accuracy across all the image. This task is undertaken far less frequently than the undistortion process and thus is not as present in the literature. Reversing the resolution of Eq. 6.1 i.e. determining $\{x_d, y_d\} = f(x_u, y_u)$ is much more complex. As suggested in [DF01], there is a close-formed solution if only k_1 is considered, but this is incompatible with our choice of coefficients.

For the laparoscopic image undistortion, we use a warping function of OpenCV called "remap". Based on an input image (here the undistorted image I_u), this routine assigns new coordinates for every pixel in an output image (the distorted image I_d) using a map. This map is essentially a look-up table composed of two matrices (M_x and M_y) of the same dimension than the input I_u and thus, $I_u(x, y) = I_d(M_x(x, y), M_y(x, y))$. Using OpenCV's "undistortpoints" function, we build the map based on Eq. 6.1 as $\{M_x(x_d, y_d), M_y(x_d, y_d)\} = \{x_u, y_u\} = f(x_d, y_d)$ with $x_d \in \{1, ..., 1920\}$ and $y_d \in \{1, ..., 1080\}$ in Full HD resolution. Naturally, we also get $\{x_u, y_u\} \in \mathcal{R}^2$ and these non-integer pixel coordinates are automatically handled by linear interpolation. Building the map takes a fraction of a second and so does the warping of the input image. This lens distortion simulation feature of our method is illustrated in Fig. 6.5. To better gauge the distortion, we overlayed a virtual grid in the laparoscopic image.

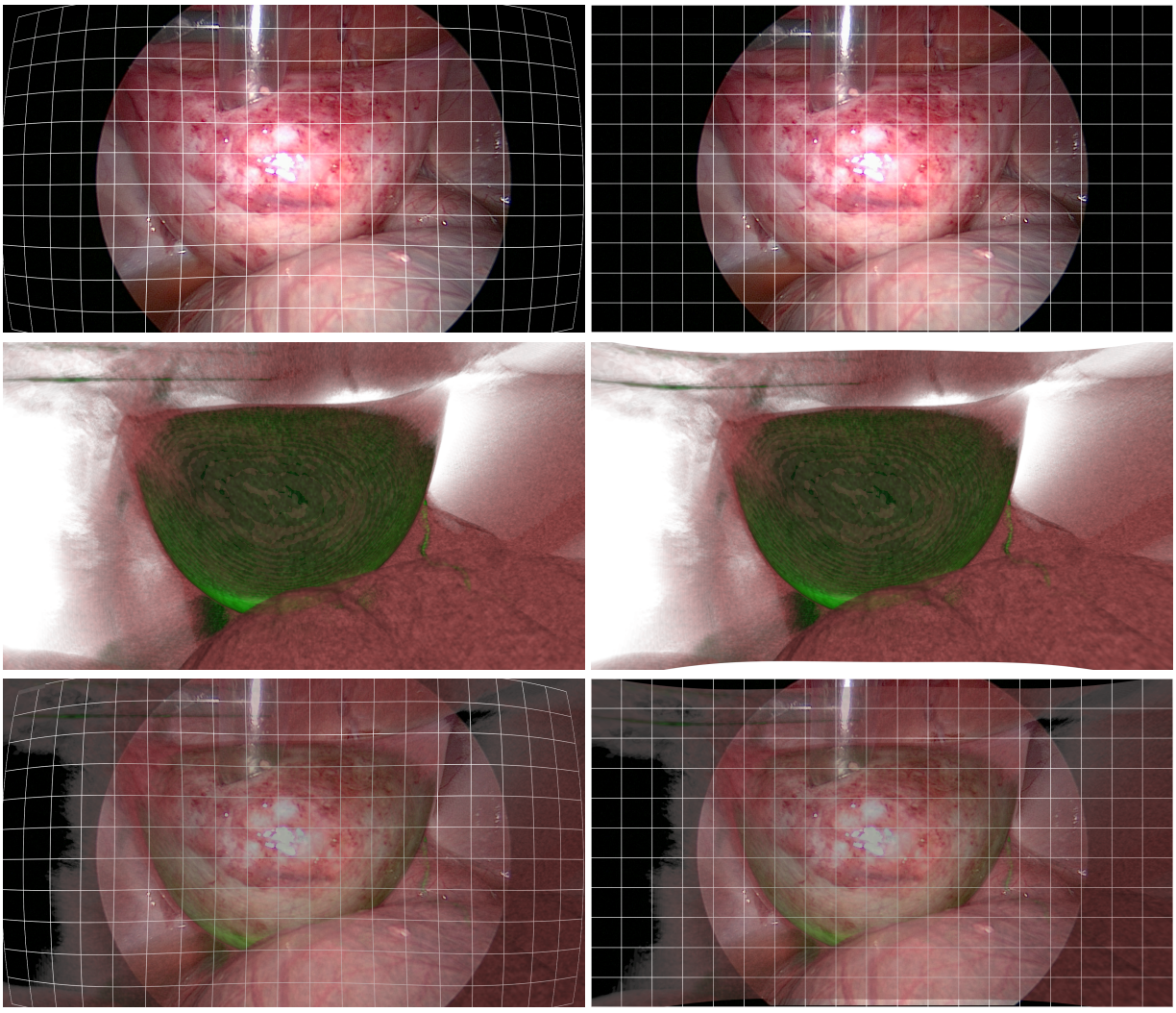


Figure 6.5: Lens distortion and undistortion. At the top, the undistortion of a laparoscopic image with a grid overlay (left) and its original unaltered version (right). In the middle, the corresponding virtual view of the scene (left) and its distorted version (right). At the bottom, the resulting augmented images.

6.2 Texture projection

As a secondary way of visually validating our registration method, we also implemented a way of projecting the endoscopic image back to the CBCT data. By applying a surface extraction method on the 3D data (in our case, marching cubes), we obtain the surface of the viscera as a mesh. Then, similarly to our AR approach, a virtual camera is placed with respect to the scene using our method. Besides, based on a camera frustrum, the graphics library VTK offers a ray casting tool that allows to project an image onto a mesh and assign the former as the texture of the latter. Resorting to this feature, we back-project the laparoscopic image onto the intra-abdominal cavity surface (see Fig. 6.6). For the sake of realism, we also display only the textured mesh cells visible from the camera i.e. within in the frustrum.

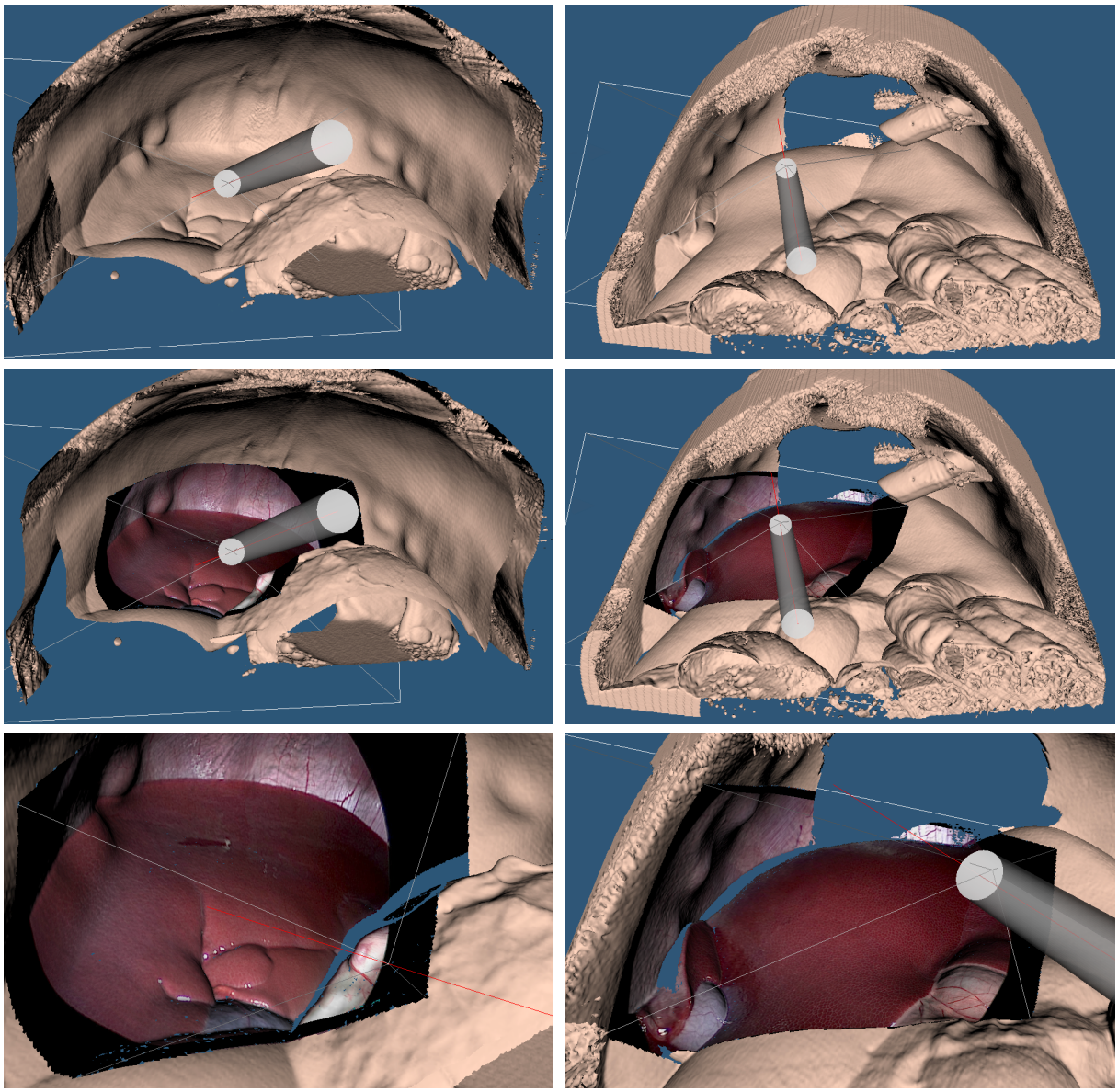


Figure 6.6: Surface texturing for two different scenes (one per column). After extraction from the intraoperative 3D data (top), the visible part of the cavity surface is textured with the laparoscopic image through back-projection from the virtual camera (mid, bottom).

One can assert the registration accuracy of our method by observing how well the texture contours match those of the surface. According to manual measurements on these renderings, the registration error is smaller than 1 mm in both cases, which confirms our quantitative results in Chapter 5. The endoscope used for the scene on the left has 0° optics, as per usual for our experiments. However, the scene on the right has been acquired through a 30° endoscope and demonstrates an early adapted version of our method for such optics. Nonetheless, the square tube calibration being so far inadequate for this kind of endoscope, the translation in the image plane has been performed manually for this instance.

6.3 Conclusion

In this chapter, we demonstrate the clinical potential of our method in realistic scenarios. First, a thorough description of the protocol of our experiments on pig has showed their realism and their close proximity to interventions on human. Second, we have presented various results of our AR method on *in vivo* data for three different surgical disciplines. For hepatic surgery, we have successfully augmented two different scenes with artificially grown adenomas on the liver. For gynecologic surgery, we have demonstrated the possibility of accurately augmenting the laparoscopic image with the ureters only through volume rendering, with no segmentation. For lung surgery, we have showed how our AR method can directly guide the surgeon to pulmonary tumors without tedious palpations and also that our approach can prevent misinterpretations of the scene.

For each set of data, the augmentation was accurate and demonstrated the clinical interest thereof. Naturally, only the quantitative experiments from the previous chapter truly assess the accuracy of our AR method. Nonetheless, these experiments on animal have also allowed to validate our approach by experts in terms of clinical usage. We have also showed our attention to conceive our AR method in line with common guidelines for certification. Indeed, the laparoscopic image may remain unaltered while the augmentation is virtually distorted by the endoscope lenses.

Beside the validation purpose, we have also showed that our method is capable of texturing laparoscopic surfaces, thus automatically producing highly photo-realistic textured 3D models and without manually identifying the correspondences between the surface and the endoscopic image. Such models may prove useful for education as well as for surgery training.

Chapter 7

Registration refinement using shading

Previously, we have seen that our AR method was determined up to a 1D rotation (roll angle) and a 2D translation (offset endoscope/CCD), both in the image plane. We proposed to estimate these remaining parameters thanks to an accelerometer for the former and a calibration with a square tube for the latter. In this chapter, we present another approach to estimate these parameters and thus complete the registration, purely based on image analysis. We propose to register the laparoscopic image of a scene with its surface by using local shading constraints with a piecewise constant albedo hypothesis. Results from experiments on in vivo data show a millimetric registration accuracy.

7.1 Introduction

A new complement for our AR method

In Chapter 4, we showed that it was possible to register a laparoscopic image with its corresponding intraoperative 3D data, by inserting the tip of the endoscope in the 3D acquisition field. However, due to the tubular shape of the laparoscope, the roll angle of the camera cannot be determined solely from analyzing the intraoperative 3D data. We proposed to estimate this degree of freedom thanks to an accelerometer included in the camera, but such a feature is *not* present in most laparoscopes.

Moreover, in Chapter 5, we explained that the hypothesis of coincidence between the optical axis of the laparoscope and its axis of revolution may be often violated. This discrepancy results in a 2D shift ϵ in the image plane, while attempting to align the CCD with the rod. Though small at the scale of the device, this difference can yet result in up to several tens of pixels of registration error in the augmentation. We have presented a simple calibration method using on a square tube, which allows to estimate ϵ. However, this step has to be performed during the intervention once the camera is set, as the zoom and the light cable position influence ϵ. While performing this supplementary calibration intraoperatively is feasible, the sake of preserving the workflow compels for a method purely based on image processing.

In this chapter, we present a novel method to complete our AR routine, which has been developed in partnership with Professor Adrien Bartoli from ALCoV-ISIT at Université d'Auvergne. It solves the previously mentioned registration issues using only information from the intraoperative volume and the laparoscopic image. As discussed, the parameters representing those three DoFs are to be determined – the roll angle and the translation ϵ along the image axes. We propose to obtain these by optimizing a dissimilarity metric between the laparoscopic image and the view from the virtual camera upon the content of the intraoperative volume (Fig. 7.1). Given the relatively poor contrast between the organs in an intraoperative CT image, the surface of the abdominal cavity is one of the most relevant information we could extract from the volume for the virtual camera. Since the cavity is insufflated with carbon dioxide, it presents a good contrast with the surrounding tissues and therefore extracting its surface is trivial, using for instance marching cubes.

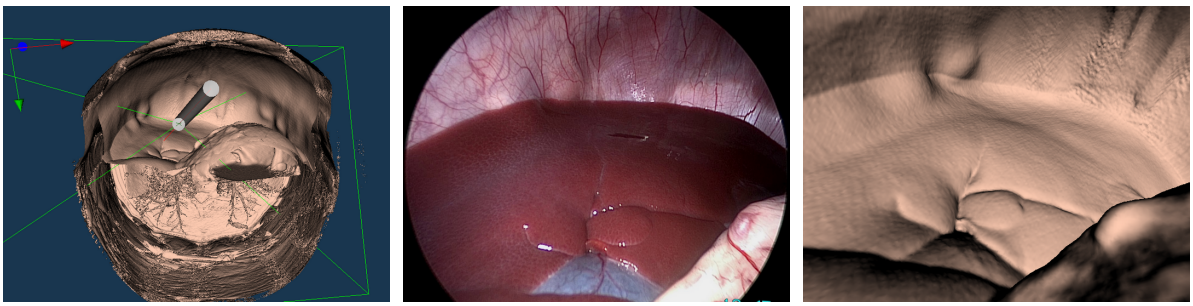


Figure 7.1: Registration refinement approach. Using our AR method, the virtual camera is placed with respect to the surface extracted from the intraoperative 3D data (left). The registration refinement consists of a 2D translation and a rotation in the image plane. Recovering these three DoF here consists in determining the relationship between the laparoscopic image (middle) and the virtual view of the surface (right).

Related work

There are various ways to register an intraoperative surface to the laparoscopic image. One way is to use Shape-from-Shading (SfS), which reconstructs a surface from a single image based on the pixels' intensity and the reflectance function [DFS08, MHMB$^+$13, LSQ$^+$15]. The reconstructed surface can then be registered to the surface extracted from the intraoperative volume using a method such as ICP. However, while SfS is solid for Lambertian surfaces [PCF06], it has been established that SfS should not be used on its own in laparoscopic surface reconstruction [CB12b], notably due to the falseness of the hypothesis of constant albedo throughout the scene. In our case, SfS would be overachieving since we do not need to reconstruct the surface from the laparoscopic image, but rather to design a dissimilarity metric between the intraoperative volume and the image. This enables us to use a local approach to shading and thus to alleviate the hypothesis of constant albedo (see Section 7.2), similarly to [DCY03].

Another means to relate a surface with its image is simply to perform a correlation between their luminance using Mutual Information or an equivalent. However, the surface extracted from the intraoperative volume is textureless. There is thus no color information and approaches based purely on luminance are likely to fail (see Section 7.3 for experimental results supporting this assertion).

We also cannot consider methods based on photo-consistency [CRHH01, JC04] which has been successfully applied to endoscopic scenes [FRH$^+$10], as two or more images are required. Similarly, Umeda et al. proposes to use gradient constraints [UGR04], but their method relies on the acquisition of "range intensity images", which merges intensity and topology.

Finally, a popular approach to registering a surface and an image is to rely on matching features and/or matching contours [KNZI02]. However, as pointed out in [WZH$^+$10], the former approach suffers from the difficulty of locating cross-rendering features and the latter requires a large number of correspondences or a good initial alignment to avoid local minima.

Our proposed method to complete the registration uses a local formulation of the shading constraints. In the next section, we present the shading model and the formulation of the dissimilarity metric between the two inputs.

7.2 Registration refinement using shading

This section describes the shading model used to determine the received light intensity. This model is simple because it is applied locally on the surface and uses piecewise constant albedo and piecewise constant light intensity hypotheses.

Shading Model

As illustrated by Fig. 7.2, the only light source inside the abdominal cavity is the one from the laparoscope, modeled as a point light source of position $S \in \mathbb{R}^3$ and intensity $l \in \mathbb{R}$ supposed constant locally. We consider Σ the surface extracted from the intraoperative volume and $\varphi \in C^2(\mathbb{R}^2, \mathbb{R}^3)$ the embedding of Σ which provides the surface point for each pixel $q \in \mathbb{R}^2$ in the laparoscopic image I. φ is known up to the sought pose of the virtual camera. The normal to Σ at φ is given by $\mathcal{N} \in C^2(\mathbb{R}^2, \mathbb{R}^3)$. In a typical laparoscopic image, there are often specularities and poorly lit areas. If we discard those (see Section 7.2), it is reasonable to assume that the camera response is linear and therefore a quantity of light k is converted

by the sensor into a pixel intensity given by $\tau(k) = ak, a > 0$. The albedo $\zeta \in C^0(\mathbb{R}^2, \mathbb{R})$, or surface reflection coefficient, is supposed constant on the surface locally for a same tissue and therefore $\zeta(q) = b, b > 0$. This is the classic limiting hypothesis in SfS, which we relax in Section 7.2.

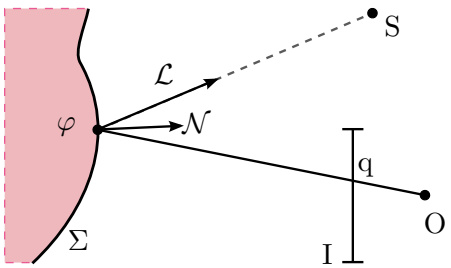


Figure 7.2: Shading model. The point light source S emits a ray $-\mathcal{L}$ that hits the surface at φ. Assuming a Lambertian surface, the light is reflected with respect to the normal $\mathcal{N}$ and the illumination vector $\mathcal{L}$. This reflection is projected onto the image plane I at q with O being the optical center and the origin of the world space.

In a laparoscopic setting, the effect of illumination fall-off may be strong. We model this by dividing the amount of received light by the squared surface-to-light source distance. Assuming S and the origin O coincide, the illumination vector $\mathcal{L} \in C(\mathbb{R}^2, \mathbb{R}^3)$ at φ is thus given by:

$$\mathcal{L} = l\frac{\overrightarrow{\varphi S}}{\|\overrightarrow{\varphi S}\|^2} = l\frac{S-\varphi}{\|S-\varphi\|^2} = -l\frac{\varphi}{\|\varphi\|^2} \tag{7.1}$$

Assuming that the surface is Lambertian, the reflectance $\mathcal{R} \in C^2(\mathbb{R}^2, \mathbb{R})$ is given by $\mathcal{R} = \mathcal{L} \cdot \mathcal{N}$. Finally, using the camera response function τ, the intensity I of a pixel q is predicted by:

$$I = \tau \circ (\zeta\mathcal{R}) = ab(\mathcal{L} \cdot \mathcal{N}) = cJ \qquad \text{with} \quad c = abl \quad \text{and} \quad J = -\frac{\varphi^\top \mathcal{N}}{\|\varphi\|^2} \tag{7.2}$$

Thus, based on reasonable assumptions about shading in the abdominal cavity, Eq. (7.2) is a simple solution to relating the surface to the luminance I in the laparoscopic image, via the calculation of the reflection potential J ($J > 0$ since $\overrightarrow{O\varphi} \cdot \mathcal{N} < 0$). The coefficient c would ideally be a function of space as both albedo and light intensity vary in the scene. Therefore, we assume c to be constant only locally. The next section explains how this piecewise relationship between the surface and the laparoscopic image can be used in order to determine the three unknown registration degrees of freedom.

Shading-based Surface-Image Dissimilarity

Eq. (7.2) is valid for areas in the scene that are not extremely lit (specularities), unlit and for which the albedo is approximately constant. Therefore, we first apply a simple large median filter (23×23) on the 1080p laparoscopic image in order to robustly remove high frequencies (texture and specularities) while preserving the edges. Dark areas are discarded with a simple

threshold on luminance. Satisfying the locally constant c requirement is equivalent to locally enforcing constancy for both albedo and intensity.

Therefore, we divide the image into a set $\mathcal{P}$ of homogeneous patches using the watershed algorithm (Fig. 7.3 a-c). The distance between the watershed seeds is related to the size of the image and the kind of its content. In a typical laparoscopic scene filmed at 1080p, the size of the different organs is commonly above 100 pixels, due to the close-up view. Setting the seeds too coarsely would result in missing small structures, while patches not large enough would not contain enough shading information and thus would fail at constraining the dissimilarity measurement. From our experience, a distance between the seeds of 150-200 pixels is satisfactory for 1080p laparoscopic images i.e. 15-20% of their size.

For each patch $p \in \mathcal{P}$, we use Eq. (7.2) at each pixel $q \in p$ to estimate c by linear regression (Fig. 7.3 d-f). The resulting residuals constitute a least-squares cost function f_p that measures how well the laparoscopic image and the virtual view of the cavity surface concur for a patch p. The variable is the camera pose ω, which affects both φ and $\mathcal{N}$ (and thus J) through the location of the coinciding points O and S.

$$f_p(\omega) = \arg\min_{c\in\mathbb{R}} \sum_{q\in p} \|I(q) - cJ_\omega(q)\|^2 = \arg\min_{c\in\mathbb{R}} \sum_{q\in p} \left\| I(q) + c\frac{\varphi_\omega(q)^\top \mathcal{N}_\omega(q)}{\|\varphi_\omega(q)\|^2} \right\|^2 \tag{7.3}$$

Finally, we obtain the transformation $\hat{\omega}$ composed of the three sought degrees of freedom by minimizing the residuals for each patch $p \in \mathcal{P}$ in the global cost function F:

$$F(\omega) = \sum_{p\in\mathcal{P}} f_p(\omega) = \sum_{p\in\mathcal{P}} \left(\arg\min_{c\in\mathbb{R}} \sum_{q\in p} \left\| I(q) + c\frac{\varphi_\omega(q)^\top \mathcal{N}_\omega(q)}{\|\varphi_\omega(q)\|^2} \right\|^2 \right) \tag{7.4}$$

We solve $\arg\min_{\omega\in\mathbb{R}^3} F(\omega)$ by using a continuous numerical optimization algorithm (Powell's conjugate direction search in our case). The registration between the laparoscopic image and its virtual equivalent can thus be completed in rotation and translation, allowing an accurate augmentation of the surgical scene.

Faster picking on a surface mesh

A simple approach to determine φ and $\mathcal{N}$ for a certain pixel in the image is to "pick" that same pixel in the window displaying the surface mesh (see Fig. 7.4-left). The VTK library then proceeds by casting a ray from this pixel and detecting the first intersecting cell in the mesh. Typically, all the normals are preemptively calculated and stored for all cells, so the one at the intersection then directly provides its position (φ) and its normal vector ($\mathcal{N}$).

However, solving $\arg\min F$ requires to determine φ and $\mathcal{N}$ for a very large number of pixels multiple times, which would require to perform this picking routine equally often. In practice, this approach is too computationally expensive, unless with a specific implementation on a powerful GPU.

We tackled this issue by resorting to a simple trick. Instead of displaying the mesh with lighting effects and no texture, as in Fig. 7.4-left, we perform the opposite and show the mesh where each cell is color-coded with its own identification number (id), as in Fig. 7.4-right. Thus, for any pixel in the window, its color directly indicates the corresponding cell and thus its position and normal, without ray casting. The cell id is encoded into RGB using the 256 modulus, so the maximum number of cells is 256^3 i.e. 16.7 million, which is definitely enough for this kind of mesh. For comparison, picking 1000 cells lasts 110 seconds with the common technique and a fraction of a second with ours.

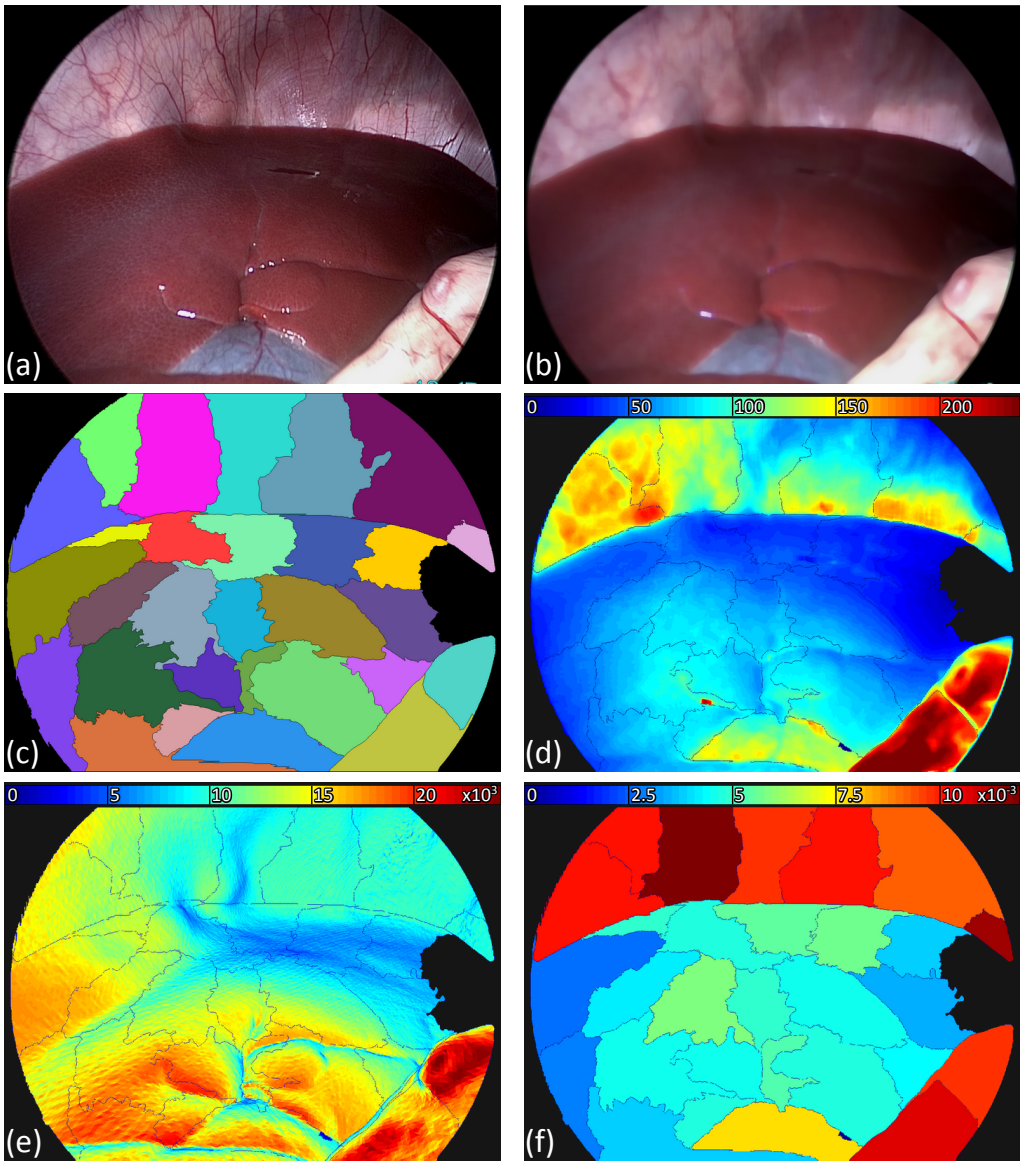


Figure 7.3: Illustration of each step of our method. (a) The input laparoscopic image is undistorted. (b) A median filter is applied. (c) The image is divided into homogeneous patches by watershed. Dark patches are discarded. Within each patch, we consider (d) the true luminance I and (e) the reflection potential J estimated from the surface to calculate the constant c. As shown in (f), c is rather constant across patches for a same kind of tissue.

7.3 Experiments and Results

In the previous section, we proposed to minimize the cost function (Eq. 7.4) in order to accurately register the laparoscopic image and the intraoperative volume. Hence, the success of our method also depends on the difficulty that optimization algorithms may have to find the global minimum in the search space. A couple of considerations ensure that an initialization at (0,0,0) is close to the global optimum. First, the surgeon is very unlikely to rotate the

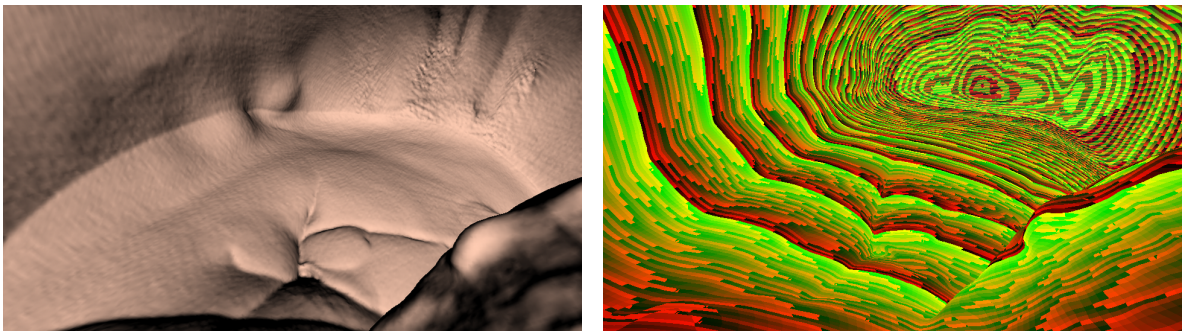


Figure 7.4: Surface picking. On the left, the mesh of a surface from a laparoscopic scene is displayed with lighting effects and no texture. On the right, the same mesh is displayed without lighting effects and each cell is uniquely colored based on its own id number.

laparoscope so much that the scene would be upside down. Second, the sensor cannot diverge too much from the laparoscope axis without hindering the completeness of the image captured out of the optics. An example of a clear global optimum in such a 3-dimensional search space around the initialization is illustrated by Fig. 7.5. These data originate from an *in vivo* acquisition of a pig's liver (see Chapter 6), for which we applied our method. A total of four different acquisitions on four different pigs were performed. Each time, the intraoperative images were taken during breathhold. Results are displayed in Fig. 7.6 and Fig. 7.7.

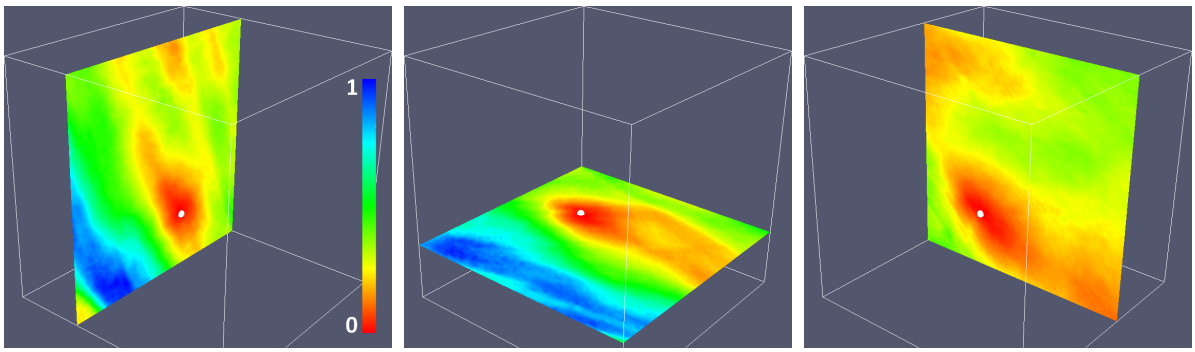


Figure 7.5: Example of search space typically ranging $\pm 30°$ and $(\pm 150)^2$ pixels. The cost issued by $F(\omega)$ is here normalized and colored from blue (high) to red (low). Sections are displayed along each of the three dimensions and passing by the global optimum (white dot).

For these experiments, one can notice the very good registration accuracy achieved by our method. Over the four data sets, we performed manual measurements of the Target Visualization Error (TVE) by pointing 15 visual cues such as organ outlines in both views (Tab. 7.1). Our AR method has become more than twice as accurate than in Chapter 4 and significantly more accurate than in Chapter 5, with an average TVE of 13 ± 5.1 pixels. This corresponds to less than a millimeter in the scene at nominal distance. Thus, the remaining three degrees of freedom are accurately determined and so is the complete relationship between the laparoscopic image and the intraoperative 3D data, without additional apparatus or calibration. Typical optimization computation times range from 15 to 30s on a standard PC. Added to the initialization, the complete augmentation process takes between 25 to 55s.

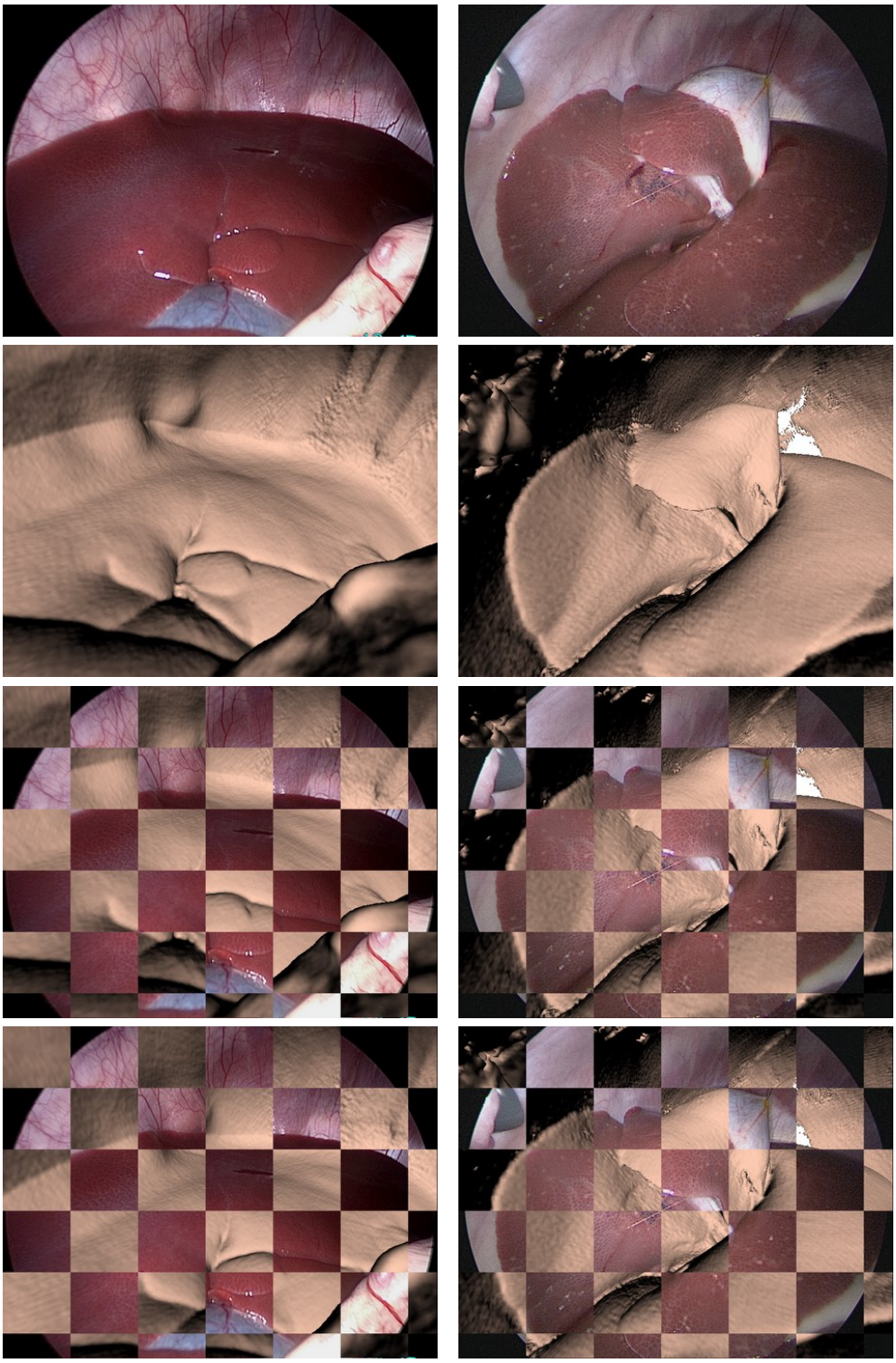


Figure 7.6: Tests on liver data from monocular endoscope (Cases 1 & 2). The laparoscopic image (top) is registered with the view from the virtual camera upon the surface extracted from the intraoperative 3D data and rendered in VTK (middle top). A mosaic of the two shows the alignment before the proposed optimization (middle bottom) and after (bottom).

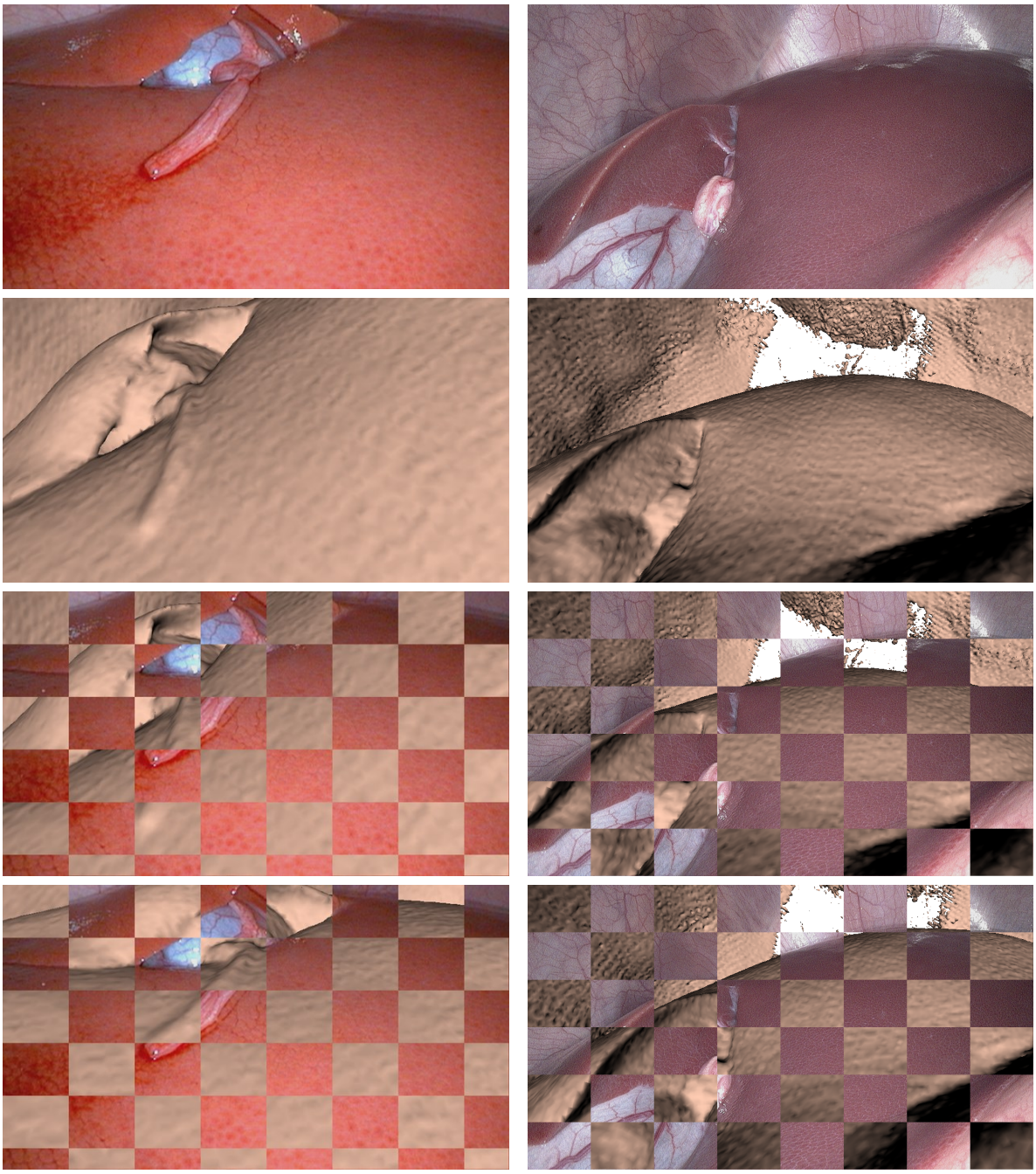


Figure 7.7: Tests on liver data from stereo endoscope – left channel only (Cases 3 & 4). The laparoscopic image (top) is registered with the view from the virtual camera upon the surface extracted from the intraoperative 3D data and rendered in VTK (middle top). A mosaic of the two shows the alignment before the proposed optimization (middle bottom) and after (bottom).

	Initialization	Method from Chap. 4	Method from Chap. 5	Proposed method
Case 1	123	13	-	6
Case 2	59	21	16	13
Case 3	>300	44	22	15
Case 4	160	49	20	18
Average	>160 ± 102	31.75 ± 17.5	19.3 ± 3	13 ± 5.1

Table 7.1: TVE (in pixels) manually measured for each case after initialization at (0,0,0), after performing our AR method as in Chapter 4, as in Chapter 5 and with the proposed method. Case 1 being part of an early experiment, it cannot be applied the square tube calibration, which had not yet been developed.

Finally, in the Introduction we asserted that classic 2D image-to-image registration methods such as Mutual Information would fail with such data. For the sake of verification, we calculated for each case the Normalized Mutual Information (NMI) between the endoscopic image and the surface view, while setting the translation to its correct value and varying only the angle. Similarly, to demonstrate the importance of a piecewise approach to shading, we calculated the proposed cost function $F(\omega)$ with globally constant c and piecewise constant c. These three metrics are compared against each other in Fig. 7.8.

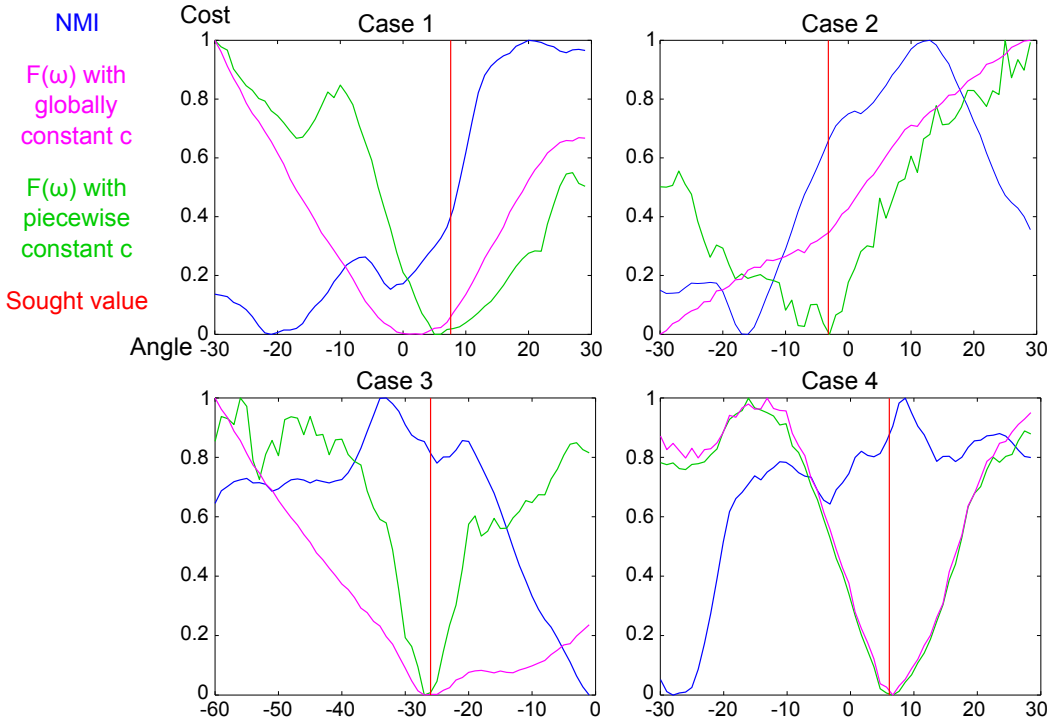


Figure 7.8: Display of values from the three metrics in rotation only, normalized between their minimum and maximum values. For NMI (blue), higher is better. For $F(\omega)$ with globally constant c (pink) and with piecewise constant c (green), lower is better. The graphs show that only our piecewise approach clearly displays a global optimum at the correct angle value (red) in all four cases.

One can notice that NMI does not reach its global maximum at the sought angle value for any of the four *in vivo* data sets, but is close in Case 4. Moreover, our method with a globally constant c performs well only in Cases 3 and 4. Since $c = abl$ (from Eq. 7.2), this suggests that, for these two instances, most of scene in the laparoscopic image presents similar albedo and lighting. The homogeneity in lighting is due to the close proximity from the camera, and thus the light source, to the surface. The homogeneity in albedo is quite clear in Case 3 where the image mostly consists of only one organ, but is less foreseeable in Case 4. At any rate, our method with the piecewise approach is the only metric to successfully converge to the correct angle value in all four cases.

7.4 Discussion

Our shading approach is not able to handle all types of scenes. Sometimes, the lighting conditions are too extreme, resulting in specularities too large. Some wet tissues may also reflect other parts of the scene and some organs may present highly inhomogeneous texture, not to mention the potential presence of smoke or blood. For all these cases, our shading refinement method is likely to fail, unlike the square tube calibration which is independent from the content of the scene. In good conditions though, our AR method seems to reach a better accuracy with the shading approach compared to that of the square tube. Both registration refinement methods provide a good visual accuracy with no significant misalignments. As a result, the choice in the refinement method depends on the complexity of the laparoscopic scene. More experiments on various *in vivo* data ought to be carried out in order to better discern the limitations of our shading approach in terms of scene content. These additional tests may also further validate this registration refinement method in terms of accuracy, possibly resorting to small fiducials placed on the cavity surface.

Besides, there is still room for improvement. First, the piecewise approach of our method makes it highly parallelizable and a GPU implementation would allow it to reach a shorter processing time. This would make our application more suitable for clinical applications, but also could compensate for breathing if real-time processing is achieved. Second, our approach obviously requires that the laparoscope tip has to appear in the intraoperative scan. Although various experiments with surgeons have proved that doing so is not problematic for them, we plan to investigate the possibility of extrapolating our work and determining all six DoFs of this rigid registration approach only from the shading constraints. If not feasible in real time, and for the sake of providing a dynamic augmented reality solution in the hybrid operating rooms, we could also look into updating the augmentation with laparoscope tracking techniques such as Vision SLAM or a robotic arm.

7.5 Conclusion

We have presented a novel method to complete a partial registration between a laparoscopic image and a surface extracted from intraoperative 3D data. When combined with our original AR method, we can provide a millimetric registration between the laparoscopic view and the intraoperative referential frame, using only standard hybrid operating room equipment and requiring no extra calibration process. This facilitates a fast and reliable augmentation of the scene with relevant information coming either from the intraoperative or the preoperative acquisitions.

So, while most shading methods aim at recovering the structure of the scene, we seek the camera pose, which we estimate by estimating shading of the scene from its surface. Thus, we do Pose-From-Shading-From-Shape rather than Shape-From-Shading. The concept of using shading to estimate the camera pose with respect to a known model is new. Moreover, most existing works on shading assumes a constant albedo over the whole image. This assumption is obviously false in a typical intra-abdominal scene where different organs and tissues have different albedo and reflectance. This is why we propose this novel piecewise approach to shading, making it compatible with such scenes.

Chapter 8

Conclusion

As a closure to this thesis, the main achievements of our work are listed and compared to its initial objectives. Then, a few perspectives are provided about our project, including new research topics that arose from our work. Finally, several broader topics about surgical AR are discussed, namely the place of intraoperative 3D imaging, the issue of depth perception and what could constitute an ideal laparoscopic AR system.

Source: IRCAD France

8.1 Achievements

Based on a comprehensive review of past and current methods of laparoscopic AR (Chapter 3), we identified major issues that still need to be addressed before attempting a translation to a commercial product. Those hurdles match the objectives of our project stated in Chapter 1, that are **accuracy**, **compatibility** and **reproducibility**. After a brief summary of our approach, we detail how these objectives were met by our laparoscopic AR solution.

A novel approach
For the purpose of AR, we developed a new approach to the registration of intraoperative 3D data to a corresponding laparoscopic image in a hybrid OR equipped with a CBCT system. By including the distal tip of the endoscope in the field of acquisition of the scanner, a rapid detection and analysis of the resulting voxels provides the geometrical relationship between the tip and the 3D image of the ROI. Based upon this relationship, we create a virtual camera with the same intrinsic and extrinsic parameters than those of the real endoscope. Thus, the intraoperative 3D data, rendered as a volume with enhancement of the structures of interest, is captured by the virtual camera from the same point of view than the real scene by the actual endoscope. As a result, the virtual image can accurately augment the laparoscopic one. Registration refinement in the image plane can be performed through a dedicated calibration with a square tube or purely by image analysis using our piecewise shading approach.

While our method so far provides only static AR, we have shown its usefulness for various kinds of laparoscopic procedures (Chapter 6). Besides, if the augmentation is required during endoscopic camera motion, our method could still provide an accurate initialization that would rid of any calibration procedure, often tedious.

Accuracy
Accuracy is the primary problem with laparoscopic AR and thus has been our main objective. Many aspects of our method are dedicated to reach this goal. First, we have implemented a template-based method of endoscope localization in the intraoperative 3D data, which is robust to artifacts and does not entirely rely on a threshold (Chapter 4). Second, we have taken into account the possible misalignment between the optical axis and the axis of revolution of the endoscope, which results in significant registration errors. To compensate for this phenomenon, we proposed two solutions: a simple calibration of the endoscope with a square tube (Chapter 5) and a purely image-based registration refinement with a novel piecewise approach to shading (Chapter 7). Third, since all data are acquired intraoperatively, there is much less deformation than with preoperative images and the anatomy is more accurately represented during the augmentation.

Extensive validation experiments on a dedicated radio-opaque checkerboard have yielded quantitative results proving that our method is robust and capable of an accuracy of 1 mm (Chapter 5), while the original objective was 5 mm. This accuracy has also been confirmed on multiple *in vivo* data sets from various realistic surgical scenarios and the utility of the augmentations were assessed by experts (Chapter 6).

Compatibility
Our method has been designed with a constant attention to compatibility with current standards of surgery in a hybrid OR. We have demonstrated the feasibility of our method during real interventions on pigs (Chapter 6). Despite the offline treatment of the data, the latter was acquired in realistic conditions during the intervention. The few extra steps required by

our method are typically achieved within a minute (Chapter 5 & 6), which represents only a brief interruption of the standard workflow and is thus acceptable by surgeons.

Our numerous animal experiments in a hybrid OR have also showed that introducing the endoscope in the intraoperative 3D acquisition is really not troublesome (Chapter 4). Our method is compatible with standard poses and we confirmed that, in such positions, the metallic tip does not provoke artifacts in the region of interest in the scan. Given the similarity between porcine and human abdomens, the previous observations can be extended to interventions on patient. Moreover, checking the correct insertion of the endoscope within the CBCT field of acquisition is possible during the fluoroscopy which is already routinely performed before the scan.

Overall, our laparoscopic AR method has limited interference with a typical surgical workflow and naturally integrates therein.

Reproducibility
For our project, we have also aimed at simplicity. Our method has a short learning curve. Indeed, performing the camera calibration and, if need be, the square tube calibration, both in good lighting conditions, takes only a few trials to comprehend and master.

Moreover, our approach does not require additional impractical or expensive systems, notably no external tracking device. We are aware that accelerometers are not typically mounted onto endoscopes, but such electronics is cheap and likely to integrated more and more endoscopes as image guidance gains in popularity. A square tube is also very easy and cheap to obtain. Moreover, our registration refinement based on piecewise shading removes the need for both the accelerometer and the square tube calibration (Chapter 7). As for holding the endoscope during the scan, the use of an articulated arm is already common in hybrid ORs. Other items such as the checkerboard and the square tube are inexpensive and widely available.

Therefore, our laparoscopic AR method is simple and very much importable to any hybrid OR with an intraoperative scanner. Thus, we also provide a bronze standard to any research team who wishes to compare their laparoscopic AR method with another.

8.2 Perspectives

Given the success of our project, notably the promising results on pigs, we expect human trials at the new IHU facility upon its completion in 2016 (see Fig. 8.1).

In the meantime, despite the achievements listed in the previous section, our method can be further improved by reaching a few goals:

- Designing and manufacturing a **more advanced square tube** would greatly increase the precision of the related calibration. Nonetheless, its simplicity shall remain so as to preserve the reproducibility of our method. For example, the introduction of a spring or an equivalent material inside the tube would help keep the tip of the endoscope pressed against the opposite corner during rotation. Also, closing the end of the tube with a white semi-opaque surface, like paper, would ensure a clearer and better contrasted square in the resulting images, thus helping the detection of its edges.

- More validation experiments ought to be done using **stereo endoscopes** to assess the compatibility of our method with such devices, whose popularity is rapidly growing. Indeed, contrary to the monocular endoscopes used for most of our tests, the two imaging

Figure 8.1: Project of the new IHU of Strasbourg, to open mid 2016.

sensor of stereo endoscopes are located at the distal tip, which further weakens the hypothesis of parallelism between the axis of revolution and the optical axis. Fortunately, recent tests on **in vivo** data showed promising results (Chapter 7).

- Our registration refinement method based on piecewise shading analysis would benefit from more experiments to better demonstrate its potential and better grasp its limitations. A new optimized implementation of the algorithm, for example on GPU, is also desirable for shorter runtimes. Finally, extending our method from 3 to 6 DoFs would allow **complete Pose-from-Shading** that could further refine the registration between the intraoperative 3D data and the laparoscopic image. This may be particularly helpful when the optical axis deviates from the axis of revolution, which could occur for stereo endoscopes (see above) but certainly concerns those with 30° optics.

Moreover, promising research topics stem from this work and could complete or extend our project:

Improving intraoperative 3D image
Our approach to laparoscopic AR relies on intraoperative 3D data to augment the laparoscopic view. If the structures of interest are not distinguishable in this data, then the purpose of AR is defeated and preoperative data shall be used again through a non-rigid registration, at the potential cost of losing accuracy. Thus, improving the quality and contrast of the intraoperative 3D image is a promising task that would allow to spread the usage of AR to more kinds of intervention. Progress may be made also in the design of TFs, which condition the volume rendering (more about his topic in Section 8.3.

Parallelly, developing innovative tools of annotation for the volume rendering of CBCT images is also a relevant topic, as the surgeon may need to interact with the augmentation

itself i.e. the volume rendering, during the intervention. These tools may be imported from those already used on preoperative images, but they probably require some adaptation for an online usage, in order to minimize the impact on the surgical workflow. Such annotations can be cutting planes, margins, measurements, trajectories, etc.

Extension to other types of endoscopes
In our project, we originally focused on straight monocular endoscopes with 0° optics, since their design would allegedly comply with our hypotheses. Given the promising results, we have also started to experiment with other kinds of endoscopes, like stereoscopes (Chapter 7) and 30° endoscopes (Chapter 6). The former does not require an adaptation of our method, since it is basically equivalent to two monocular endoscopes attached to each other. As mentioned previously in the perspectives, more quantitative experiments are expected in order to validate our method for such optics too.

Oblique endoscopes, on the other hand, require an adaptation of our laparoscopic AR method as the optical axis is then far from being aligned with the detected axis of revolution and so should be the virtual camera. A few solutions have been proposed in the literature to take this aspect into account. One approach is to use rotary encoders [YNS+04][IUK+12] or optical markers [DBMDS07][WJN10] to determine the angular position of the bevel with respect to the camera, but these methods require tedious calibrations and additional devices. A less invasive approach consists of a tiny alteration of the aperture visible in the image [FCN+10][MBF12], but this limits the zooming range. Given our approach to AR, a more practical solution may be available, that does not require additional devices or hardware modifications and that does not impact the workflow with extra calibration procedures. If the tip of an oblique endoscope is acquired by the intraoperative scanner, as per our method, then the bevel ought to be visible in the resulting 3D data. Early tests confirm this hypothesis and promising results were obtained, but more work needs to be sunk into this topic to reach more accuracy and robustness.

Finally, tracking flexible endoscopes is so far feasible only using an electromagnetic system, which is not ideal (see Appendix B). Since the tip of such devices is locally rigid, we may apply our method and thus provide the only sub-millimetric AR solution in flexible endoscopy.

Online calibration
So far, our laparoscopic AR method requires a calibration of the camera outside of abdomen and any subsequent change to the determined parameters would require to repeat this process, otherwise the augmentation becomes inaccurate. While camera calibration is relatively not tedious and can be performed within a minute, the intervention would surely benefit if any change to the camera settings could be automatically detected, estimated and applied to the augmentation.

The literature reports several methods to partially compensate for changes in one or more camera parameters. For instance, Melo et al. propose to infer the intrinsic parameters from circular outline shown in the endoscopic image [MBF12], which can be adapted in real-time [MFB12]. However, this method is applicable only to endoscopes that display such a circular outline (monocular) and also the zooming is restricted as most of this contour has to appear. In a similar manner, Lourenco et al. propose to determine the zoom by tracking salient points in the image [LBF+14]. An issue with both approaches is that, while they seek to remove the need for camera calibration, they still require one lengthy pre-calibration process per endoscope model used. Indeed, numerous successive measurements are performed to

accurately determine the relationship between the camera parameters and the considered features in the endoscopic image.

A more general approach to online calibration is known in computer vision as "self-calibration". The latter relies on calibration algorithms similar to the ones using calibration objects such as a checkerboard and applies them to scenes of unknown topology. As pointed out by Maier et al. in [MHMB+13], self-calibration of endoscopes remains very challenging, especially for monocular devices. However, Stoyanov et al. reported in 2005 some preliminary progress on self-calibration with stereo endoscopes, but their method was not robust to focus changes and it seems their work has not been continued ever since. Thus, Maier et al. concludes that the most promising approach so far would be to fall back to calibration objects, more specifically items of known dimensions in the surgical scenes, as in [SBK+09][PVH09][AOT+13].

We think an interesting approach to online calibration would be a mix of both pre-calibration and self-calibration, as in [PBDY14]. In this paper, Pratt et al. propose to automatically calibrate intraoperatively a stereo laparoscope by reducing the problem to a 1D search on the focus setting, based on pre-calibrated data and the insertion of a tiny checkerboard pattern in the cavity. This technique is more invasive but also more practical than a conventional calibration.

8.3 Discussions

Upon reaching the end of this thesis, we now take a step back and discuss a few broader topics. First, since our method relies only on intraoperative imaging, we discuss the pertinence of such a choice given current technologies. Second, we address a problem in surgical AR that has received, in our opinion, too little consideration so far: the issues of perception. Last, we share some reflexions of what would constitute an ideal laparoscopic AR system.

Is intraoperative 3D imaging the solution to surgical AR ?

As explained in Chapters 2 & 3, the main issue with preoperative imaging is the deformation occurring between of the acquisition and the intervention. This factor compelled to introduce intraoperative scanners in the OR and, for the past few years, there has been a steady growth in the creation of hybrid ORs. Obviously, such improved accuracy in the representation of the patient's anatomy is interesting for AR purposes. However, given the current systems, can an AR solution solely rely on intraoperative imaging ?

Typically, during an intervention in a hybrid OR, one can perform an intraoperative acquisition using, for example, a CBCT machine. Provided that a contrast agent has been injected beforehand, then the vasculature is visible in the 3D data and can be registered to the laparoscopic image using optical trackers as in [FMHN08] or, simpler, using our method. The augmentation remains accurate after the acquisition as long as the scene does not deforms too much. As we already demonstrated for our project, notably in Chapter 6, such relatively stationary phases often occur in various kinds of interventions. Nonetheless, as soon as surgical actions such as resections or mobilizations irreversibly change the scene, then the augmentation becomes obsolete and, if AR is still desired, a new scan has then to be performed. For CBCT imaging, each new acquisition requires to remove all instruments from the ROI in order to avoid artifacts and all the surgical staff has to exit the OR to limit unnecessary radiation exposure. Moreover, despite the low dose, a new acquisition also involves

a new exposition to radiation for the patient. For our project, we use the CBCT system Artis Zeego from Siemens, for which the radiation received during a 3D acquisition is only 30% of that for a standard CT. Therefore, in such scenario, ensuring an accurate augmentation at critical stages of the intervention comes at the cost of additional radiations for the patient. The trade-off between clinical benefits and radiation hazard is difficult to define and would typically require extensive clinical trials. For example, in [SDB+10], Shekhar et al. have demonstrated offline the possibility of continuously updating laparoscopic AR with 100 very low-dose acquisitions from a conventional CT scanner every 1.1 seconds. Given such a protocol, it is unclear whether the clinical benefits of having continuously updated AR outweigh such prolonged radiation exposure for both patient and staff.

In light of this, laparoscopic US has remained a popular means of intraoperative imaging for its low cost and its non-invasiveness, especially compared to CBCT. Nonetheless, this modality does not exactly produce 3D data, rather 2D images stitched together. To compensate the sparsity of images or the lack of depth, intraoperative US may be used to register preoperative data, as in[VRL+10], but then again such data are not accurate by the time of the intervention. Moreover, the quality of intraoperative US images highly depends on the operator's level of expertise.

As a result, intraoperative MRI appears to be the best solution to laparoscopic AR. Indeed, it offers fast, accurate and fully 3-dimensional images of the patient's anatomy. Besides, it is non-invasive and does not rely on iodine contrast agent, such acquisitions may be repeated several times as the intervention unfolds. Nevertheless, as explained in Chapter 2, intraoperative MR scanners presents a significant caveat regarding the presence of ferromagnetic materials which imposes the use of special instruments or a complex adaptation of the OR.

So, as of now there is no perfect solution of intraoperative imaging for laparoscopic AR. However, interesting perspectives emerge as significant progress is being made in taking down the last obstacles preventing a wide adoption of such systems.

Perception issues in surgical AR

So far, research on surgical AR has primarily focused on legitimate goals such as registration accuracy, real-time processing, robustness and especially how to durably fulfill those criteria for as long as necessary. In comparison, the quality of the surgeon's perception of the augmentation has commonly been overlooked. This may seem paradoxical as it basically constitutes the core goal of surgical AR. A few publications report progress about this topic, but most perceptual issues in surgical AR remain unsolved. As stated in [SML+08], we currently lack a robust quantitative evaluation framework for perceptual fidelity in laparoscopic AR, which would help identify and quantify the problems. However, two main issues have already been reported in the literature: the choice in augmentation style and content and especially the perception of depth.

What to display ?

As pointed out in [KOJC13], "what", "how" and "when" are paramount concerns about the digital data presented to the surgeon, but these features are often overlooked in an image guidance system. As already mentioned in the conclusion of Chapter 3, this may be a symptom of the lack of multidisciplinary collaboration between the engineers and medical staff [CP10][MIH11]. Such a lack of insight is regrettable, especially since AR may have a strong impact on the surgeon's attention [DDC+13], which underlines the importance of appropriate

rendering and content for the augmentation. For example, disorientation has been identified as the main cause of surgical errors in laparoscopic bile duct surgery, especially due to visual illusion [SYD12].

As far as "what" and "when" are concerned, a possible solution may lie in context-aware AR. For example, Katic et al. propose to automatically adapt the content of the augmentation to the various phases of an intervention, based on sensor feedback and logical rules [KWG+13]. This approach would then benefit from methods of automatic localization of different elements in the scene, such as the multi-structure segmentation of stereo endoscopic scenes using prior from preoperative model [NAP+15].

Estimation and display of uncertainty in the augmentation is also a promising feature for surgery guidance. Typically, this could help the surgeon during tumor resection to achieve negative surgical margins while sparing a maximum of healthy tissues. To this end, Ukimura et al. propose a color-coded road map with colored outlines around the target at various planned distances [UNS+10]. A similar approach can be used to display uncertainty around segmentations of preoperative data [AKNP+13]. If volume rendering is used, then the uncertainty can be displayed via probabilistic animations associated to adaptive TFs [LLPY07].

Concerning "how", an equilibrium between integration and highlight has to be reached. Indeed, the augmentation should ideally blend in the scene, so as to limit the distraction to the surgeon. However, it should also significantly stand out in order to be easily perceived. Such a balance is too difficult to strike. For example, transparency is commonly used in AR, but also decreases the contrast of either the augmentation or the scene behind. Various rendering approaches have been investigated in the literature, but, as demonstrated in [SBHN06], they are not equally efficient, especially when it comes to alleviating a prominent issue of surgical AR: depth perception.

The perception of depth

The most intuitive and thus most used approach in surgical AR is a simple superimposition of the augmentation onto the real endoscopic image. Such a projection of one 2D image onto another removes optical effects such as occlusions and motion parallax, which are the strongest clues for depth perception in human vision. Other secondary clues include binocular disparity, perspective, texture and shading/shadowing [CV95][WDC+05]. As listed hereafter, each of these cues may be enforced in an attempt to improve depth perception of an augmented scene.

- **Occlusion** is the strongest clue for the perception of depth and, as such, a strong candidate for its simulation. A common approach to suggesting occlusion in AR is virtual windowing [BSHN07] (see Fig. 8.2-a), but such this kind of augmentation also partially obstruct the endoscopic image. To partially alleviate this issue, several groups have reported the use of **inverse realism**. This technique is essentially the closest to approximate the concept of a superhero-like see-through vision by extracting and accentuating strong features from the scene and by superimposing them in turn on top of the augmentation (see Fig. 8.2-b). Such an occlusion provides the sensation that the tissue surface covers the augmentation, while keeping the latter visible and preserving the salient features of the former. Inverse realism may be obtained through a simple gradient of the endoscopic image. More advanced techniques also take into into account the topology of the scene, either through a pq-space representation [LCMY07] or an RGB-D image [KMS07]. Inverse realism can be rendered in real-time thanks to a GPU implementation [PMV+12].
  Occlusion of the augmentation from hovering items such as instruments may also be

simulated using image-based tracking [ATC+14]. Nonetheless, this technique primarily concerns depth over the surface which may be surgically less significant than depth within the tissues. At any case, despite being the strongest clue for depth perception, occlusion only informs about the order, not about the distance, neither relative nor absolute.

- **Motion parallax** is weaker than occlusion in terms of depth suggestion but it provides relative distance, which is very valuable in surgery. However, this cue is difficult to enforce as it requires motion. The most intuitive approach is to track the surgeons head or with a HMD as in [BWHN07](see Fig. 8.2-c), but such equipment is very constraining (see Section 3.3). An additional feature of this approach can be a virtual mirror [BHFN09], but requires instrument tracking.

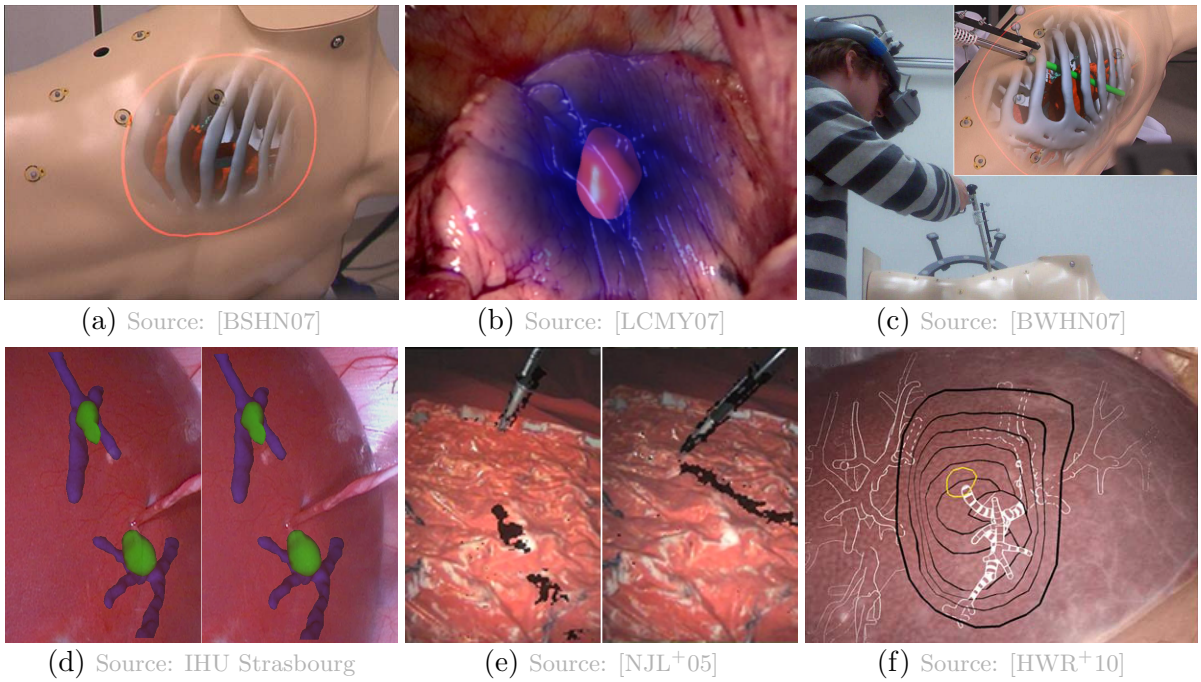
- **Stereo disparity** is based on the same depth perception principle as motion parallax but transcends the latter given its capacity of originating from static scenes. Moreover, thanks to the vergence of our eyes, stereo vision constitutes our best resource for estimating an absolute distance in a scene (see Fig. 8.2-d), which is a valuable resource for the surgeon. In fact, binocular laparoscopy has been shown to outperform monocular laparoscopy [TAP+12][AEBS+14][WBE+14]. However, in simulated stereoscopy, the instinctive relationship between vergence and accommodation in our eyes is not respected and needs to be overridden in order to maintain sharpness in our vision [RHFL10]. Depending on the person and thus on the surgeon, such a strain may provoke a discomfort and, over time, visual fatigue and even headaches that could seriously hinder the intervention.

- **Shadowing** compensates depth perception loss by virtually creating or enhancing shadows [NJL+05] (see Fig. 8.2-e). However, this approach requires a support for casting virtual shadows onto. Such a surface may originate from a model [BWHN07] or from a reconstruction of the scene, but the latter is challenging as explained in Section 3.5. On the other hand, despite providing some sense of depth, **shading** is rather a clue for shape [CV95]. Ideally, the shadowing/shading of virtual objects should match that of the scene. A simplified lighting model (as the one used in Chapter 7) might be enough.

- **Perspective** may be used to provide some relative depth perception, but is also strongly dependent on shape. Nonetheless, it can be used to enhance the visualization of certain structures. For example, Hansen et al. propose to relate the thickness of contours of tumors and vasculature to their depth (see Fig. 8.2-f) [HWR+10]. However, wireframe and contour rendering is unsuitable for larger, more complex structures.

Beside vision, another sense might compensate for the lack of depth perception, as in [BFS+15] where sound is used to inform about the distance to target.

As one can see, interesting approaches have been suggested in order to provide a sense of depth to the surgeon based on a single clue. However, enforcing depth perception for all the clues at once remains a complex challenge in visualization and rendering and is thus not yet available. As a result, conflicts might arise between enforced and unenforced clues and provoke more or less discomfort for the surgeon, depending on his/her own appreciation.



(a) Source: [BSHN07] (b) Source: [LCMY07] (c) Source: [BWHN07]

(d) Source: IHU Strasbourg (e) Source: [NJL+05] (f) Source: [HWR+10]

Figure 8.2: Simulating depth in surgical AR. (a) Occlusion simulated by a virtual window; (b) Inverse realism; (c) Motion parallax provided via a HMD; (d) Stereo disparity from binocular endoscope; (e) Casting virtual shadows; (f) Perspective simulated by varying vasculature contours.

So, while the pursuit of accuracy and robustness in surgical AR are founded, perception has also to be taken care of. Indeed, the augmentation ought to be displayed in a realistic way to enrich the surgeon's view on the scene, otherwise the alleged augmentation might turn into an actual degradation.

What's the ideal AR system ?

In this last part of the thesis, we propose to examine ideal concepts of laparoscopic AR to better grasp the remaining challenges in this field.

According to Lee et al. in [LLV+10], such an ideal system would first consist of a patient-specific biomechanical model built from preoperative data. This model would then be non-rigidly registered to the scene, whose surface is recovered in real-time and serves to simulate haptic feedback and to keep clear from forbidden zones.

Another ideal laparoscopic AR approach is described by Hughes et al. in [HHMM+14] as a "reliable, safe, accurate and user-friendly system". The latter would rely on the synergy between automated registration, tracking of both surgical scene and endoscopic camera and tissue deformation recovery. Multi-modal imaging, including for example US, would help in connecting those three elements.

In our opinion, those concepts are mostly valid, but we now propose to go a little deeper in the description. Our ideal laparoscopic AR system would ideally be constantly accurate, robust, safe for both patient and medical staff and would not disturb the surgical workflow. From these criteria, we imagine the following approach.

Our idea of an ideal laparoscopic AR system

A scan of the patient is performed prior to the intervention, using a non-invasive imaging modality such as MRI and if required, a harmless contrast agent. The surgeon and his team plan the whole procedure on these preoperative data using an intuitive yet complete set of tools allowing rapid segmentations, annotations and assisted or free-hand drawing. All these planning information is attached to the preoperative data, alongside biomechanical properties of the structures of interest. At the beginning of the intervention, the patient is placed on the operating table in a hybrid OR, which includes a semi-open low-field intraoperative MR scanner. To avoid any risk of interaction with the magnetic field, all equipment, including the instruments and the endoscopes, are mostly non-ferromagnetic since made of ceramic and polymers. Attached to the operating table are several robotic arms. At least one of them is holding the main endoscope, the others may hold other endoscopes for multiple simultaneous points of view or laparoscopic instruments to be maintained at a certain position. The robotic arms may work passively for human guidance or actively for simple trajectories such as a return to a certain position. These arms may also be blocked in a certain position and restrict the exploration of certain areas designated as hazardous. The endoscopes are stereoscopic and each displays 4K images at 60 fps on multiple screens next to the operating table. The visual sensors are also multispectral for a better distinction of tissues and structures in the scene, using fluorescence for example. The main monitor displays in 3D for the surgeon to guide his actions thanks to a realistic vision upon the surgical site with depth perception. Other 2D monitors are placed in the hybrid OR so that the remainder of the staff permanently follows the progress of the intervention. An illustration of such a hybrid OR is given in Fig. 8.3. As in [PHH13], a voice recognition system is also installed in the room and allows the surgical staff to send instructions to the various digital systems in place, limiting manual interaction and thus disturbance to the surgical workflow.

Source: IHU Strasbourg

Figure 8.3: Concept of hybrid OR in the near future.

During the intervention, when AR is required, an intraoperative scan of the patient is performed, with the tip of the endoscope(s) introduced in the field of acquisition. Dedicated markings along each endoscope appear in contrast in the intraoperative 3D data and allow to automatically determine the position and orientation of each camera, thanks to a method similar to the one described in this thesis. Virtual cameras are then created with the exact same point of view upon the scene than that of the real ones. Adaptive TFs allow for a volume rendering of the intraoperative 3D data that automatically shows structures of interest. The same structures were also localized in the preoperative data and a 3D non-rigid registration is performed taking into account the biomechanical properties of each tissue. Thus, planning information is also visible from the virtual cameras, whose images then constitutes the augmentation for the actual laparoscopic view. While AR is enabled, inverse realism is used to provide a more natural integration of the augmentation to the scene, thus limiting conflicts in depth perception. Moreover, the surgeon's gaze is tracked and given the vergence and accommodation of his/her eyes, the stereo disparity on the 3D display is adapted accordingly so as to reduce visual fatigue.

For the remainder of the intervention, the augmentation is also automatically updated. First, the scene is permanently automatically segmented according to the content of the laparoscopic image and a surface reconstruction based on stereo disparity. Based on dictionaries, elements of the scene such as instruments and organs are then identified based on logical rules and the stage of the intervention is deducted. This allows for an automatic adaptation of the augmentation according the advancement of the procedure, which limits unnecessary clutter in the scene. Moreover, motion feedback from the robotic arms holding the endoscopes allows to automatically update the position of the virtual cameras and thus to maintain the accuracy of the augmentation. Pseudo-periodical motions in the scene resulting from breathing and heartbeat are automatically compensated for the augmentation based on a real-time registration refinement in the image plane, similarly to our piecewise shading method. Finally, the augmentation is also updated with respect to scene deformations resulting from surgical acts such as resection and mobilization, based on the surface reconstruction and the biomechanical models.

We are aware that such an ideal system does not yet exist. Yet many of the features described above were inspired by current achievements in the field of laparoscopic AR. So, while many challenges remain ahead, we are convinced that such a concept may be concretized in a not-so-distant future and we hope that our work has brought us one step closer to this ideal.

Appendices

A Precedent reviews

AR in surgery has been an active topic of research for more than two decades. It is a multidisciplinary concept than spans various scientific fields and involves numerous tasks. As a result, the volume of literature related to surgical AR is significantly large. In fact, there are already many review papers concerning this subject. Here is a list of those published during the last 8 years, alongside a brief description of their content.

- [BFMR08] reviews computer assisted navigation systems for soft tissue surgery. While the focus is on navigation, a bit of a surgical AR background is still drawn as part of a larger scope.
- [SFN08] examines medical AR at large and does not focus on laparoscopy. Despite being slightly outdated, this constitutes a good review on the various techniques of rendering medical AR as well as the systems for displaying it. This paper also features an interesting consideration for AR perception issues. However, the focus is clearly on the equipment and not the methods: the review skims through registration, especially non-rigid.
- [CP10] is a review on image-guided surgery in general with technological considerations for tracking, registration, visualization and software. Numerous concrete applications for various kinds of interventions are mentioned, from neurosurgery to NOTES. Insightful points are raised that are applicable to AR, notably the transfer between engineers and practitioners as well as the need for rigorous validation. This article is a follow-up from their book of 2008 [PC08].
- [LLV+10] provides technical pointers and background on feature tracking, surface reconstruction and subject-specific modeling (including biomechanical models) geared towards robotic assisted MIS. These elements are very common in surgical AR, which is briefly considered.
- [LFC+10] consists in a short overview of AR and robotics in surgery, but the authors emphasize the importance of multidisciplinarity in development of AR solutions. They illustrate this point by reporting about their multidisciplinary team during a 4-year project.
- [NSMM11] provides a review of AR in laparoscopic surgery with details on the various displays and tracking methods. The paper revolves mostly around interactive AR methods, but some feedback about automatic methods is also reported.

- [MIH11] mostly aims at providing some technical background in computer vision and key elements to endoscopic navigation, in an effort to better understand how the former can improve the latter. A brief state-of-the-art is also provided, with only a few works mentioned about AR in laparoscopy and a certain focus on motion compensation.

- [NUFG12] briefly reviews some recent advancements in AR applied to laparoscopic surgery, especially in urology. Most of the article reports on the achievements from their group. A few other works are also mentioned.

- [KOJC12] proposes a taxonomy, coined DVV, for mixed reality visualization techniques applied for image-guided surgery (IGS) based on the type of data and their processing. The authors aim at unifying the syntax and framework for this field. Unfortunately, judging by the literature, it seems that their taxonomy has been little used in the 3 years following their publication. Nonetheless, their work has the merit of clarifying the critical components of a successful mixed reality visualization system. Moreover, a review of 17 state-of-the-art papers in IGS is provided and analyzed at the light of their taxonomy.

- [KOJC13] follows up on [KOJC12] by releasing a review of the state-of-the-art in mixed reality visualization using their DVV taxonomy. Their review spans across all surgical fields and focuses on visualization. Their thorough analysis helped to assert the points made in [KOJC12] about the trends and lacks in the field of mixed reality visualization in surgery. Some of their findings will be touched in the remainder of this chapter, but we encourage any interested reader to go through their article for a more complete and sound picture.

- [NMSA13] is a review about image guidance for interventions on liver and kidney. A solid surgical background is also provided about laparoscopy, laparotomy and percutaneous procedures.

- [Tro13] is a book delving into the realm of Medical Robotics as a whole. As a computer-based and care-oriented approach, surgical AR also falls into this category and is considered in depth in the fourth chapter. This chapter begins by pertinently dispensing the basics in terms of anatomy modeling and visualization, especially for the abdomen. Upon this base, a distinction is made between interactive and automatic AR methods. For the latter, the survey is restricted to external camera-based solutions which are listed with respect to the number of cameras used and their mobility. Nonetheless, numerous surgical cases are reported across different disciplines. Besides, the book also contains chapters dealing with related subjects such as image registration and acquisition.

- [HHMM+14] reports some major achievements in terms of AR for partial nephrectomy. The review is not exhaustive, but the different mentioned approaches are well categorized and compared against. In the discussion section, interesting points are made about the need to merge several modalities and technologies. Moreover, the authors conclude with an insightful description of an ideal AR system.

- [OOY+14] considers AR-based navigation applications dedicated to abdominal surgery. However, a significant part of the article consists in a description of accomplishments from their group, which concern open surgery for the abdomen (laparotomy), and not its

minimally invasive equivalent (laparoscopy). Nonetheless, an interesting consideration about the evaluation of prototypes concludes the paper.

- [KWN+15] examines some of the recent emerging technologies in computer-assisted abdominal surgery. This concerns a wide variety of techniques and applications, among which laparoscopic AR, briefly touched in their review. Nevertheless, this paper gives interesting perspectives for abdominal surgery.

B Tracking the camera

In order to guide his/her actions, the surgeon is often brought to change his point of view on the scene and move the camera, sometimes abruptly. Without a reliable tracking of the endoscope, its position becomes unknown and thus the augmentation inaccurate. Determining this position in real-time is inherent in surface-based methods. Therefore, the issue of camera motion concerns mostly volume-based methods. Various tracking technologies are available nowadays, but two are predominant: optical and magnetic tracking.

Optical Tracking
This technology consists of at least two dedicated infrared cameras which detect reflective beads fixed in a pattern and attached to the tracked element, as shown in Fig. B.1. By triangulation of the position of the markers in both views, their 3D location in the scene is calculated in real-time. The pattern is always attached to a part of the tracked object that minimizes the disturbance for the surgeon and also guarantees as much as possible its constant visibility by the cameras. A hand-eye calibration is also required in order to determine the relationship between the pattern and the optical center.

Figure B.1: Optical tracking. On the left, reflective beads are fixed in a pattern and attached to the tracked element, here a laparoscope. On the right, a device with two dedicated cameras is placed in the OR and detects the infrared reflections emitted by the beads for triangulation of the laparoscope's position.

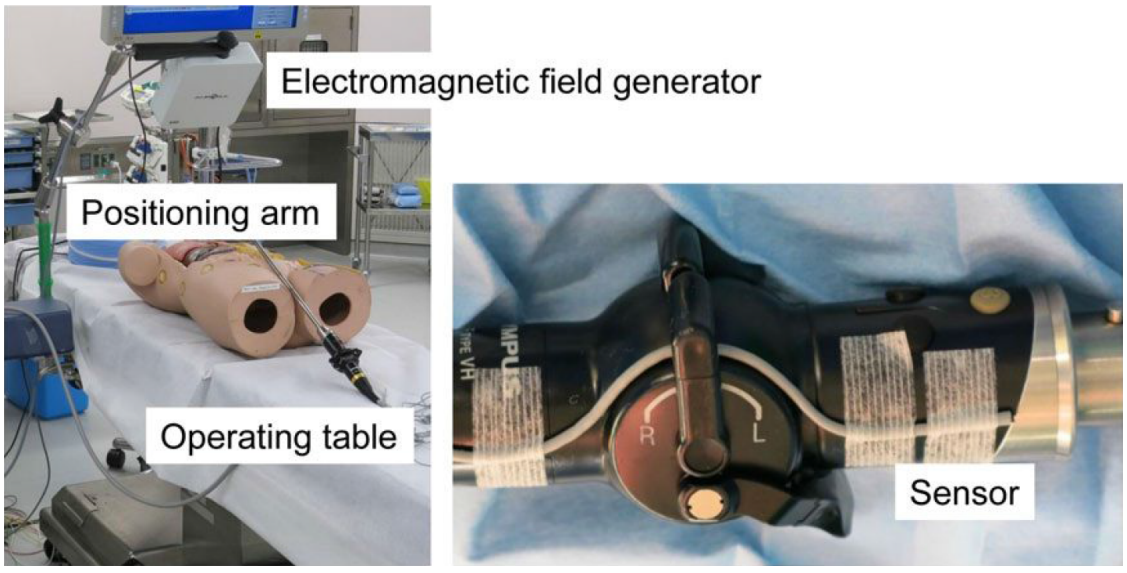
Optical tracking is a popular solution for endoscope and instrument tracking. Typically, several optical patterns are used to simultaneously track other objects in addition to the laparoscope, such as a C-arm [FMHN08] [ROU+12], surgical instruments [BVP+13] or a US probe [UNS+10] [KOW+14]. Clinical trials of laparoscopic AR using optical tracking have already been reported in [IUK+12] [SIU+13].

As shown in [WTF04], this tracking technology is able to provide a sub-millimetric accuracy at the location of the markers. However, as explained in the previous paragraph, the patterns are located at the proximal part in MIS and thus, by propagation, the error is much larger at the distal end, where accuracy actually counts for registration. According to the measurements on rigid bodies performed in [WTF04], a 0.25 mm RMS error at the marks would result in nearly 2 mm at the tip of a common 40 cm long laparoscope.

Besides, optical tracking equipment is costly and cumbersome. Many calibration steps are required [SBMJ+02] [GML+04] [NGS05] and can slow down the surgical workflow. Moreover, the line of sight between the infrared reflectors and the cameras have to remain clear at all times, which constrains the surgeon's movements. Finally, optical tracking is inappropriate for laparoscopic US transducers, which have a flexible tip. To alleviate this issue, an additional pattern can be drawn onto the tip of the probe and detected in the laparoscopic image, either as a colored tip [ONN+14] or as contrasted tags [JMPC13]. Nonetheless, for this kind of endoscopic tools, electromagnetic (EM) tracking is better suited.

Electromagnetic Tracking

This technology consists in tracking one or several small wired sensors inside a magnetic field generated by a device close to the surgical site, as shown in Fig. B.2. The magnetic field is of known geometry and rapidly changing, inducing a current in the sensors at various voltages which are measured. These measurements are then used to determine the position of the sensors with respect to the generator. Since the sensors are much smaller than the optical ones and given that the magnetic field traverses tissues, the markers can be attached at the exact location that requires tracking, even inside the patient. Therefore, EM tracking is well-suited for any insertable tools with a flexible tip, such laparoscopic US probes [LKW+14] or flexible endoscopes in NOTES [ARS+12]. This tracking technology has already been used for laparoscopic AR on the kidney [CWM+10] and the liver [KDE+12] [HIH+15].



Source: [HIH+15]

Figure B.2: Electromagnetic tracking. On the left, the generator is placed next to the operating table. On the right, the wired sensor is taped to the endoscope.

As pointed out by Franz et al. in their comprehensive review of the state-of-the-art of EM tracking in medicine [FHB+14], this technology is not as accurate as its optical counterpart concerning the accuracy at the marker level. Nonetheless, EM tracking may still compete at the level of tip of the tracked instrument, as it can reach millimetric accuracy, but only in a good environment. Indeed, EM tracking is very sensible to its surroundings as other nearby metallic objects in the OR distort the magnetic field [FHB+14]. These perturbations may make the tracking accuracy fall back into the centimetric range [LVR+12], due to the presence of a C-arm for instance. Konishi et al. propose to compensate these tracking errors by a preoperative calibration [KHN+05]. This may work for static metallic equipment such as the operating table, but is not applicable for mobile devices.

In light of the strengths and weaknesses of these two tracking technologies, attempts have been made to use both in an optical-EM hybrid approach [NNS+08] [FRV+09], but the reported accuracy still revolves between 2 and 3 mm. Robotic arms represent another emerging tracking solution [SDIH+10] [LJD+14] as they seem to obtain a significantly better accuracy, but this approach is still in its infancy.

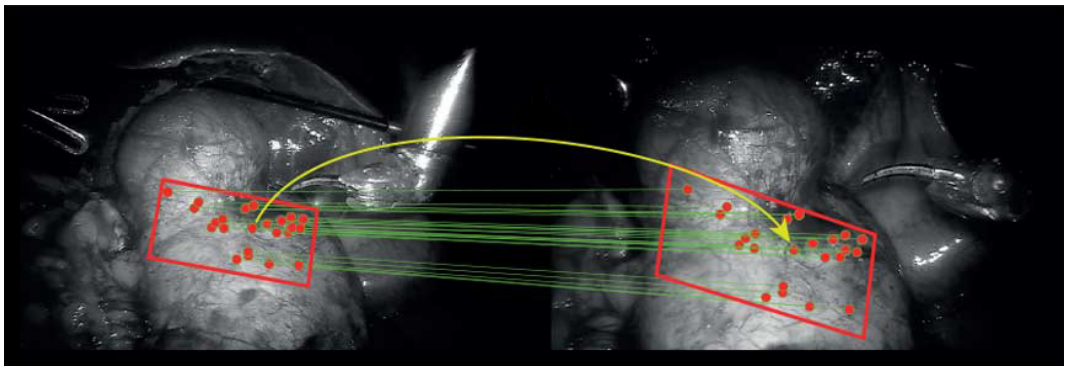
Overall, all currently available tracking technologies impractically introduce additional costly equipment in the OR and clutter the surrounding of the operating table. Moreover, these systems require extensive calibration procedures which are a major source of inaccuracy in the resulting registration [CKTB13]. Concerning endoscope tracking, a solution may lie in the exploitation of natural or artificial landmarks, using triangulation algorithms as in the surface reconstruction methods [SNP+07]. Such landmarks are already useful to recover the deformation of the scene.

C Tracking the scene

As seen in Section 3.4, a laparoscopic scene is prone to deformation during the intervention. Causes include anatomical motions such as breathing and heart beat, as well as surgeon's interactions with the tissues. While the former are approximately periodical, the latter is unpredictable. Either way, these deformations need to be evaluated and the augmentation updated accordingly. The most common approach is to track landmarks in the endoscopic image, which can be natural or artificial.

Tracking natural landmarks

A surgical scene contains naturally discriminative features such as edges, corners and rich textures. Tracking these in real-time in the endoscopic image provides information about the various dynamics in the scene [PMJ09], as shown in Fig. C.1.



Source: [PSM+13]

Figure C.1: Tracking of natural landmarks between two images of the same scene, but taken a different times. The chosen landmarks (red dots) are associated with the green lines.

For laparoscopic AR, these features can be selected automatically [SY09] or manually [SVA+09] [NDA+13]. Some tracking methods seek robustness [PSM+13] [YWLP14], while others aim at a long term accuracy [GVSY13] [YLS+12] [PSCM+14] [SNA+15]. The tracking may be designed to compensate a specific kind of natural motion like respiratory movements of the liver [MY10] or heart beat [SY09]. Feature tracking can also serve to update a biomechanical model of the liver to better determine its deformation [HDP+13]. In a more natural way, Stoyanov et al. propose to stabilize motion in the scene using gaze tracking [SML+08].

As seen in Section 3.5 with surface-based AR methods, the main difficulty of tracking natural landmarks is the paucity of distinctive features in certain areas of a laparoscopic scene, in addition to local deformation, illumination changes and specularities. The challenge of tracking in such conditions makes the need of validation data sets for feature matching in laparoscopy, but those require extensive manual interactions from experts. To circumvent this issue, Maier et al. propose to resort to crowdsourcing the ground truth of data sets [MHMK+14]. Also, a validation framework for feature tracking is proposed in [SNA+15].

Nonetheless, the lack of reliable features in parts of the laparoscopic scene imply an uncertainty about their motion and/deformation. In such cases, artificial landmarks, also known as fiducials, are a possible solution.

Tracking artificial landmarks

Contrary to natural landmarks, artificial ones are designed to facilitate their detection. As a result, tracking fiducials is robust and reliable in the long run. The most common approach consists in sticking needles with colored heads into the tissues of interest in the laparoscopic scene [SBM+11] [TGS+09]. Besides optical tracking, Nakamoto et al. propose to resort to wireless tracking of tiny EM transponders placed in the organ [NUG+08], but the accuracy of EM tracking is too sensible to its environment, as seen in the previous section.

As outlined in [LLV+10], the main disadvantage of this approach is its invasiveness. Inserting and removing fiducials in organs require extra surgical efforts and significantly increase the risk of complication. Therefore, only a limited number of markers are scattered across the area of interest, which makes the tracking uncertain between them. Sauvée et al. propose to address this issue by combining natural and artificial landmarks for a denser and more robust motion recovery [SNP+07]. Nevertheless, the presence of fiducials may also interfere with the intervention itself and thus are often placed away from the actual zone of surgery, which defeats its own purpose.

In conclusion, scene tracking and deformation recovery is a nuanced problem. Simple or periodical motions in the scene can be compensated for using prediction algorithms or dedicated equipment. However, non-rigid deformations compel for more complex methods. For this specific task, surface-based AR methods have an advantage, at the condition of providing an accurate reconstruction over the complete scene in real-time. However, such a method has not yet been developed for laparoscopic AR, as explained in Section 3.5.

D Popular software tools

Here is a list of the most popular software platforms and libraries for surgical AR development in research:

- **DICOM**. Standing for Digital Imaging and Communications in Medicine (DICOM)[1], it is widely considered to be the international standard format for medical images for both exchange and storage in every medical field from radiology to surgery.
- **MITK**. The Medical Imaging Interaction toolkit (MITK)[2] is a free open-source framework for developing interactive medical image processing software [WVW+05] [NZS+13]. MITK is based on two popular libraries: the Insight Toolkit (ITK)[3] for image processing tools and the Visualization Toolkit (VTK)[4] for display and user interaction capabilities.
- **MITK-IGT**. The Image Guided Therapy (IGT) framework[5] is an extension to MITK dedicated to the development of image-guided therapeutic applications using tracking. It offers native support of many commercial tracking systems and processing tools for the tracking data [BNM09] [FSS+12].
- **IGSTK**. The image-guided surgery toolkit (IGSTK)[6] is a framework dedicated to the development of image-guided surgery applications [ECG+07] [YCW+10]. It provides the components for loading and displaying medical images though a GUI and integrates a point-based registration tool. This framework also encapsulates the application programming interfaces (API) for common tracking systems.
- **Slicer**. Also known as 3D Slicer[7], it is a free open-source, modular and cross-platform software package for visualization and processing of images, especially medical ones [PHK04] [FBKC+12]. Also based on ITK/VTK. It offers several possibilities of rendering (slices, surface or volume) and tools for interactive segmentation and non-rigid registration.
- **CISST & SAW**. The former[8] is a collection of libraries to facilitate the development of computer assisted intervention systems. The latter[9] is an open-source cross-platform framework derived from CISST which eases the integration of various surgical systems such robots, haptic interfaces, tracking and imaging systems [VDD+08] [KDD+10].
- **OpenIGTLink**. It provides a standardized mechanism for communications among computers and devices in operating rooms (OR) for a wide variety of image-guided therapy (IGT) applications [TFP+09] [TTS+15a].
- **FW4SPL**. A component-oriented programming framework dedicated to medical image processing and visualization and developed at our institution (IRCAD France)[10].

[1] http://dicom.nema.org
[2] http://mitk.org/wiki/MITK
[3] http://www.itk.org
[4] http://www.vtk.org
[5] http://www.mitk.org/wiki/IGT
[6] http://www.igstk.org
[7] http://www.slicer.org
[8] https://github.com/jhu-cisst/cisst/wiki
[9] https://www.cisst.org/Saw
[10] https://code.google.com/p/fw4spl

- **Osirix**. A popular DICOM viewer[11] on Apple platforms (Mac, iPhone, iPad).
- **SOFA**. The Simulation Open Framework Architecture (SOFA)[12] is an free open-source framework dedicated to real-time simulation with a certain focus on medical applications and biomechanical models [ACF+07] [CJA+10].
- **MIDAS**. An open-source platform[13] to facilitate the creation and maintenance of customized web-enabled data storage, especially medical images.
- **MeVisLab**. A free modular cross-plateform software package[14] dedicated to image analysis with a focus on medical applications [RBH+11], including various segmentation, registration and visualization tools.
- **NifTK**. A compilation of various libraries[15] for medical image analysis [CZT+14]. It includes tools for registration, biomechanical simulation, segmentation and visualization.
- **OpenCV**. Probably the most popular computer vision library[16] with applications in various fields, notably medical image analysis [BK08] [PBKE12].

[11] http://www.osirix-viewer.com
[12] https://www.sofa-framework.org
[13] http://www.midasplatform.org
[14] http://www.mevislab.de
[15] http://cmic.cs.ucl.ac.uk/home/software/
[16] http://opencv.org

E Camera position and parallelism with respect to the endoscope

To determine the position of the camera from the endoscope tip, we typically perform an optimization of the position of the virtual camera along the endoscope axis for several experiments on the radio-opaque checkerboard (see Section 4.4). It has been observed that this distance d is approximately 3.4 ± 1 mm for standard FOV and focus settings. Also, this axial deviation of ± 1 mm introduces only between 0.5 and 3.5 pixels of reprojection error depending on the depth of the scene.

For the sake of consistency, we also calculated the position of the camera from its view of the checkerboard by solving the classic Perspective-n-Point (PnP) problem and measured its distance from the tip along the endoscope axis (see Fig. E.1). Across 34 experiments, we got $d = 4.1 \pm 0.45$ mm, which is close to what we previously found. Given the limited impact of the distance d on the registration accuracy, this parameter may be set anywhere between 3 and 4 mm.

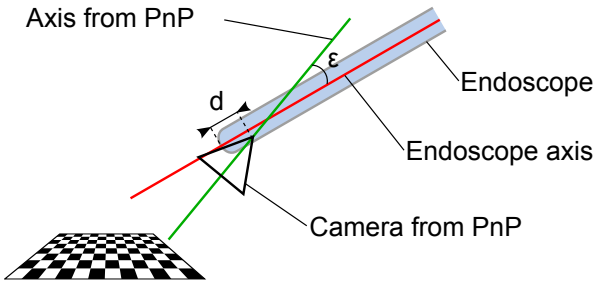


Figure E.1: Endoscope looking at checkerboard. By detecting the corners and solving the PnP problem, we can evaluate the position and orientation of the camera and compare them to what we find with our method.

PnP solving also provides the camera orientation, which is helpful to estimate how parallel the optical and endoscope axes are in the case of a 0° endoscope (see Fig. E.1). Across the same 34 experiments, we measured $\epsilon = 0.42 \pm 0.17°$, which introduces a registration error of approximately 0.51 ± 0.21 mm at nominal distance (70 mm). Therefore, it is safe to assume that the endoscope and the optical axes are parallel.

F Constant tilt sensitivity with a 3-axis accelerometer

Let us first consider a vertical 2-axis accelerometer tilting at an angle of θ in the gravity field g. As shown by the illustrations in Fig. F.1, the acceleration Ax along the X-axis is given by $Ax = \sin(\theta)$ and the one along the Y-axis by $Ay = \cos(\theta)$. Since it is placed vertically ($Y = -g$), we also have $\sqrt{Ax^2 + Ay^2} = 1g$.

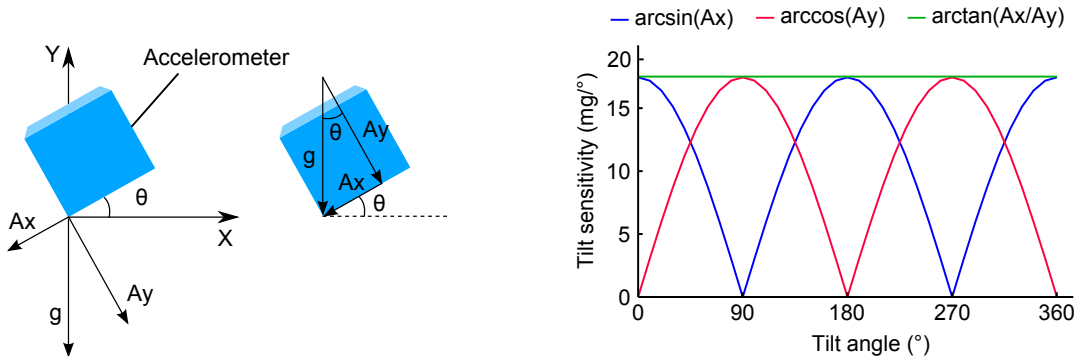


Figure F.1: Tilt sensitivity of a 2-axis accelerometer. Once placed in the gravity field g, the sensor measures the projection of the gravity acceleration $A = 1g$ onto the axes X and Y, resulting in Ax and Ay respectively (right). Considering that the accelerometer is tilted by an angle θ, Ax and Ay are geometrically related to θ (middle). However, the tilt sensitivity depends on the tilt angle θ if Ax and Ay are considered separately (right, blue and red). The solution lies in measuring the angle θ with a combination of both (right, green).

If we only consider the acceleration along the X-axis, there would be an angle orientation ambiguity as $sin(\theta) = sin(-\theta)$. Moreover, the non-linearity between Ax and θ introduces a variation of the tilt sensitivity with respect to the tilt angle (see graph in Fig. F.1). However, if we also consider the acceleration along the Y-axis, the angle ambiguity is lifted and a combination of both would result in $\frac{Ax}{Ay} = \tan(\theta)$, which removes the tilt sensitivity variations.

We can now extend these remarks to a 3-axis accelerometer. Similarly, the orientation ambiguity and the tilt sensitivity variations can be alleviated by combining the measurements of the gravity acceleration $A = 1g$ along the three axes X, Y and Z, resulting in Ax, Ay and Az respectively (see Fig. F.2).

For example, the determination of the pitch angle α can be reduced to a 2D tilting resolution in the plane $(X, Y+Z)$ and thus $\frac{Ax}{\sqrt{Ay^2+Az^2}} = \tan(\alpha)$, resulting in $\alpha = \arctan(\frac{Ax}{\sqrt{Ay^2+Az^2}})$. Similarly, we also get $\beta = \arctan(\frac{Ay}{\sqrt{Ax^2+Az^2}})$ and $\theta = \arctan(\frac{\sqrt{Ay^2+Ax^2}}{Az})$. Therefore, the tilt sensitivity of a 3-axis accelerometer can be made independent of the tilt angle itself.

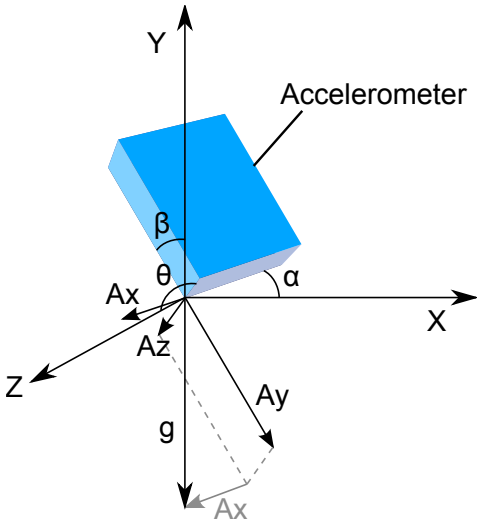


Figure F.2: Tilting of a 3-axis accelerometer. Placed in the gravity field g, the accelerometer tilts with the angles α, β and θ with respect to the axes X, Y and Z respectively.

Bibliography

[ACF+07] Jérémie Allard, Stéphane Cotin, François Faure, Pierre-Jean Bensoussan, François Poyer, Christian Duriez, Hervé Delingette, and Laurent Grisoni. Sofa-an open source framework for medical simulation. In *MMVR 15-Medicine Meets Virtual Reality*, volume 125, pages 13–18. IOP Press, 2007. [Cited on page 122]

[AEBS+14] B Alaraimi, W El Bakbak, S Sarker, S Makkiyah, A Al-Marzouq, R Goriparthi, A Bouhelal, V Quan, and B Patel. A randomized prospective study comparing acquisition of laparoscopic skills in three-dimensional (3D) vs. two-dimensional (2D) laparoscopy. *World journal of surgery*, 38(11):2746–2752, 2014. [Cited on page 109]

[AHM+92] SE Attwood, AD Hill, PG Murphy, J Thornton, and RB Stephens. A prospective randomized trial of laparoscopic versus open appendectomy. *Surgery*, 112(3):497–501, 1992. [Cited on page 19]

[AHR13] Anwar Abdalbari, Xishi Huang, and Jing Ren. Endoscopy-mr image fusion for image guided procedures. *Journal of Biomedical Imaging*, 2013:23, 2013. [Cited on page 43]

[AKHP+15] Alborz Amir-Khalili, Ghassan Hamarneh, Jean-Marc Peyrat, Julien Abinahed, Osama Al-Alao, Abdulla Al-Ansari, and Rafeef Abugharbieh. Automatic segmentation of occluded vasculature via pulsatile motion analysis in endoscopic robot-assisted partial nephrectomy video. *Medical image analysis*, 2015. [Cited on page 33]

[AKNP+13] Alborz Amir-Khalili, Masoud S Nosrati, Jean-Marc Peyrat, Ghassan Hamarneh, and Rafeef Abugharbieh. Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy: A phantom study. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 182–191. Springer, 2013. [Cited on pages 44 and 108]

[ALF+08] Mohamed R Ali, Jamie P Loggins, William D Fuller, Brian E Miller, Christopher J Hasser, Peter Yellowlees, Tamas J Vidovszky, Jason J Rasmussen, and Jonathan Pierce. 3-D telestration: a teaching tool for robotic surgery. *Journal of laparoendoscopic & advanced surgical techniques*, 18(1):107–112, 2008. [Cited on page 33]

[AOG+11] Hernan O Altamar, Rowena E Ong, Courtenay L Glisson, Davis P Viprakasit, Michael I Miga, Stanley Duke Herrell, and Robert L Galloway. Kidney deformation and intraprocedural registration: a study of elements of image-guided kidney surgery. *Journal of Endourology*, 25(3):511–517, 2011. [Cited on pages 43 and 44]

[AOT+13] Max Allan, Sébastien Ourselin, Susan Thompson, David J Hawkes, Jonathan Kelly, and Danail Stoyanov. Toward detection and localization of instruments in minimally invasive surgery. *Biomedical Engineering, IEEE Transactions on*, 60(4):1050–1058, 2013. [Cited on page 106]

[ARS+12] DE Azagury, M Ryou, SN Shaikh, R San José Estépar, BI Lengyel, J Jagadeesan, KG Vosburgh, and CC Thompson. Real-time computed tomography-based augmented reality for natural orifice transluminal endoscopic surgery navigation. *British Journal of Surgery*, 99(9):1246–1253, 2012. [Cited on pages 31 and 117]

[ATC+14] Max Allan, Steve Thompson, Matthew J Clarkson, Sébastien Ourselin, David J Hawkes, John Kelly, and Danail Stoyanov. 2D-3D pose tracking of rigid instruments in minimally invasive surgery. In *Information Processing in Computer-Assisted Interventions*, pages 1–10. Springer, 2014. [Cited on page 109]

[BB11] Jerrold T. Bushberg and John M. Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, third edition, 2011. [Cited on page 14]

[BBDK06] H Baumert, A Ballaro, F Dugardin, and AV Kaisary. Laparoscopic versus open simple prostatectomy: a comparative study. *The Journal of urology*, 175(5):1691–1694, 2006. [Cited on page 19]

[BCBC12] Adrien Bartoli, Toby Collins, Nicolas Bourdel, and Michel Canis. Computer assisted minimally invasive surgery: Is medical computer vision the answer to improving laparosurgery? *Medical hypotheses*, 79(6):858–863, 2012. [Cited on pages 19 and 20]

[BCG+09] Joseph F Buell, Daniel Cherqui, David A Geller, Nicholas ORourke, David Iannitti, Ibrahim Dagher, Alan J Koffron, Mark Thomas, Brice Gayet, Ho Seong Han, et al. The international position on laparoscopic liver surgery: The louisville statement, 2008. *Annals of surgery*, 250(5):825–830, 2009. [Cited on page 19]

[BCHG08] Anne B Benincasa, Logan W Clements, S Duke Herrell, and Robert L Galloway. Feasibility study for image-guided kidney surgery: Assessment of required intraoperative surface for accurate physical to image space registrations. *Medical physics*, 35(9):4251–4261, 2008. [Cited on page 44]

[BDN+96] L Michael Brunt, Gerard M Doherty, Jeffrey A Norton, Nathaniel J Soper, Mary A Quasebarth, and Jeffrey F Moley. Laparoscopic adrenalectomy compared to open adrenalectomy for benign adrenal neoplasms. *Journal of the American College of Surgeons*, 183(1):1–10, 1996. [Cited on page 19]

[BFD+07] G Belli, C Fantini, A Dagostino, L Cioffi, S Langella, N Russolillo, and A Belli. Laparoscopic versus open liver resection for hepatocellular carcinoma in patients with histologically proven cirrhosis: short-and middle-term results. *Surgical endoscopy*, 21(11):2004–2011, 2007. [Cited on page 19]

[BFM+03] Wolfgang Birkfellner, Michael Figl, Christian Matula, Johann Hummel, Rudolf Hanel, Herwig Imhof, Felix Wanschitz, Arne Wagner, Franz Watzinger, and Helmar Bergmann. Computer-enhanced stereoscopic vision in a head-mounted operating binocular. *Physics in medicine and biology*, 48(3):N49, 2003. [Cited on page 36]

[BFMR08] M. Baumhauer, M. Feuerstein, H.-P. Meinzer, and J. Rassweiler. Navigation in endoscopic soft tissue surgery: perspectives and limitations. *Journal of endourology / Endourological Society*, 22(4):751–766, 2008. [Cited on pages 37, 38, and 113]

[BFS+15] Felix Bork, Bernhard Fuerst, Anja-Katharina Schneider, Francisco Pinto, Christoph Graumann, and Nassir Navab. Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. In *Mixed and Augmented Reality (ISMAR), 2015 IEEE International Symposium on*, pages 7–12. IEEE, 2015. [Cited on page 109]

[BHE+08] Kristy K Brock, Maria Hawkins, Cynthia Eccles, Joanne L Moseley, Douglas J Moseley, David A Jaffray, and Laura A Dawson. Improving image-guided target localization through deformable registration. *Acta Oncologica*, 47(7):1279–1285, 2008. [Cited on page 46]

[BHFN09] Christoph Bichlmeier, Sandro Michael Heining, Marco Feuerstein, and Nassir Navab. The virtual mirror: a new interaction paradigm for augmented reality environments. *Medical Imaging, IEEE Transactions on*, 28(9):1498–1510, 2009. [Cited on page 109]

[BK08] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.", 2008. [Cited on page 122]

[BKM+02] SE Burpee, M Kurian, Y Murakame, S Benevides, and M Gagner. The metabolic and immune response to laparoscopic vs open liver resection. *Surgical Endoscopy and Other Interventional Techniques*, 16(6):899–904, 2002. [Cited on page 19]

[BLM+12] Mei Bai, Bin Liu, Hongqiu Mu, Xianghua Liu, and Yuan Jiang. The comparison of radiation dose between C-arm flat-detector CT (DynaCT) and multi-slice CT (MSCT): a phantom study. *European journal of radiology*, 81(11):3577–3580, 2012. [Cited on page 26]

[BNA+14] Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. Automatic detection of endoscope in intraoperative CT image: Application to ar guidance in laparoscopic surgery. In *Proceedings of ISBI*, pages 563–567. IEEE, 2014. [Cited on page 10]

[BNA+16] Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. Automatic localization of endoscope in intraoperative CT image: A simple approach to augmented reality guidance in laparoscopic surgery. *Medical Image Analysis*, 30:130–143, 2016. [Cited on page 10]

[BNB+15] Sylvain Bernhardt, Stéphane A. Nicolau, Adrien Bartoli, Vincent Agnus, Luc Soler, and Christophe Doignon. Using shading to register an intraoperative CT scan to a laparoscopic image. In *Computer-Assisted and Robotic Endoscopy*, pages 59–68. Springer, 2015. [Cited on page 10]

[BNH+13] J. Bano, S. A. Nicolau, A. Hostettler, C. Doignon, J. Marescaux, and L. Soler. Registration of preoperative liver model for laparoscopic surgery from intraoperative 3D acquisition. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 201–210, 2013. [Cited on page 46]

[BNM09] M Baumhauer, J Neuhaus, and H-P Meinzer. The mitk image guided therapy toolkit and its exemplary application for augmented reality guided prostate surgery. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, pages 224–227. Springer, 2009. [Cited on page 121]

[BOK+05] Roberto T Blanco, Risto Ojala, Juho Kariniemi, Jukka Perälä, Jaakko Niinimäki, and Osmo Tervonen. Interventional and intraoperative MRI at low field scanner - a review. *European journal of radiology*, 56(2):130–142, 2005. [Cited on page 26]

[BP11] Dirk Bartz and Bernhard Preim. Visualization and exploration of segmented anatomic structures. In *Biomedical Image Processing*, pages 379–401. Springer, 2011. [Cited on page 34]

[BSH+13] Sebastian Bauer, Alexander Seitel, Hannes Hofmann, Tobias Blum, Jakob Wasza, Michael Balda, Hans-Peter Meinzer, Nassir Navab, Joachim Hornegger, and Lena Maier-Hein. Real-time range imaging in health care: a survey. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, pages 228–254. Springer, 2013. [Cited on page 43]

[BSHN07] Christoph Bichlmeier, Tobias Sielhorst, Sandro M Heining, and Nassir Navab. Improving depth perception in medical ar. In *Bildverarbeitung für die Medizin 2007*, pages 217–221. Springer, 2007. [Cited on pages 108 and 110]

[BSMS+08] Matthias Baumhauer, Tobias Simpfendörfer, Beat P Müller-Stich, Dogu Teber, CN Gutt, Jens Rassweiler, H-P Meinzer, and Ivo Wolf. Soft tissue navigation for laparoscopic partial nephrectomy. *International Journal of Computer Assisted Radiology and Surgery*, 3(3-4):307–314, 2008. [Cited on pages 46 and 47]

[BVP+13] Nicolas C Buchs, Francesco Volonte, François Pugin, Christian Toso, Matteo Fusaglia, Kate Gavaghan, Pietro E Majno, Matthias Peterhans, Stefan Weber, and Philippe Morel. Augmented environments for the targeting of hepatic lesions during image-guided robotic liver surgery. *journal of surgical research*, 184(2):825–831, 2013. [Cited on page 116]

[BWHN07] Christoph Bichlmeier, Felix Wimmer, Sandro Michael Heining, and Nassir Navab. Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 129–138. IEEE, 2007. [Cited on pages 109 and 110]

[CB12a] Toby Collins and Adrien Bartoli. 3D reconstruction in laparoscopy with close-range photometric stereo. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 634–642. Springer, 2012. [Cited on pages 42 and 43]

[CB12b] Toby Collins and Adrien Bartoli. Towards live monocular 3D laparoscopy using shading and specularity information. In *Information Processing in Computer-Assisted Interventions*, pages 11–21. Springer, 2012. [Cited on page 91]

[CDT+05] Olivier Clatz, Hervé Delingette, I-F Talos, Alexandra J Golby, Ron Kikinis, Ferenc A Jolesz, Nicholas Ayache, and Simon K Warfield. Robust nonrigid registration to capture brain shift from intraoperative MRI. *Medical Imaging, IEEE Transactions on*, 24(11):1417–1427, 2005. [Cited on page 27]

[CGK+14] Alexis Cheng, Xiaoyu Guo, Hyun-Jae Kang, Behnoosh Tavakoli, Jin U Kang, Russell H Taylor, and Emad M Boctor. Concurrent photoacoustic markers for direct three-dimensional ultrasound to video registration. In *SPIE BiOS*, pages 89435J–89435J. International Society for Optics and Photonics, 2014. [Cited on page 47]

[CHD+14] Ping-Lin Chang, Ankur Handa, Andrew J Davison, Danail Stoyanov, et al. Robust real-time visual odometry for stereo endoscopy using dense quadrifocal tracking. In *Information Processing in Computer-Assisted Interventions*, pages 11–20. Springer, 2014. [Cited on page 42]

[CIN+03] Kevin Cleary, Luis Ibanez, Nassir Navab, Dan Stoianovici, Alexandru Patriciu, and Gabriela Corral. Segmentation of surgical needles for fluoroscopy servoing using the insight software toolkit (itk). In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 1, pages 698–701. IEEE, 2003. [Cited on page 55]

[CJA+10] Hadrien Courtecuisse, Hoeryong Jung, Jérémie Allard, Christian Duriez, Doo Yong Lee, and Stéphane Cotin. Gpu-based real-time soft tissue deformation with cutting and haptic feedback. *Progress in biophysics and molecular biology*, 103(2):159–168, 2010. [Cited on page 122]

[CJDL03] Rosie Conlon, Michael Jacobs, Dowmitra Dasgupta, and J Peter A Lodge. The value of intraoperative ultrasound during hepatic resection compared with improved preoperative magnetic resonance imaging. *European journal of ultrasound*, 16(3):211–216, 2003. [Cited on page 27]

[CKTB13] Alexis Cheng, Jin U Kang, Russell H Taylor, and Emad M Boctor. Direct three-dimensional ultrasound-to-video registration using photoacoustic markers. *Journal of biomedical optics*, 18(6):066013–066013, 2013. [Cited on pages 47 and 118]

[CM11] Carlos D Correa and Kwan-Liu Ma. Visibility histograms and visibility-driven transfer functions. *Visualization and Computer Graphics, IEEE Transactions on*, 17(2):192–204, 2011. [Cited on page 35]

[CMBP15] Elvis CS Chen, A Jonathan McLeod, John SH Baxter, and Terry M Peters. Registration of 3D shapes under anisotropic scaling. *International journal of computer assisted radiology and surgery*, pages 1–12, 2015. [Cited on page 44]

[CMC+10] Daniel Cohen, Erik Mayer, Dongbin Chen, Ann Anstee, Justin Vale, Guang-Zhong Yang, Ara Darzi, et al. Augmented reality image guidance in minimally invasive prostatectomy. In *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, pages 101–110. Springer, 2010. [Cited on pages 42 and 44]

[CMS+05] David M Cash, Michael Miga, Tuhin K Sinha, Robert L Galloway, William C Chapman, et al. Compensating for intraoperative soft-tissue deformations using incomplete surface data and finite elements. *Medical Imaging, IEEE Transactions on*, 24(11):1479–1491, 2005. [Cited on page 44]

[CP10] Kevin Cleary and Terry M Peters. Image-guided interventions: technology review and clinical applications. *Annual review of biomedical engineering*, 12:119–142, 2010. [Cited on pages 48, 107, and 113]

[CPB+13] Toby Collins, Daniel Pizarro, Adrien Bartoli, Michel Canis, and Nicolas Bourdel. Realtime wide-baseline registration of the uterus in laparoscopic videos using multiple texture maps. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 162–171. Springer, 2013. [Cited on page 42]

[CPB+14] Thomas Collins, Daniel Pizarro, Alberto Bartoli, Michel Canis, and Nicolas Bourdel. Computer-assisted laparoscopic myomectomy by augmenting the uterus with pre-operative MRI data. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 243–248. IEEE, 2014. [Cited on pages 44 and 46]

[CRHH01] M. J. Clarkson, D. Rueckert, D. L. G. Hill, and D. J. Hawkes. Using photo-consistency to register 2D optical images of the human face to a 3D surface model. *Trans. on Pattern Analysis and Machine Intelligence*, 23(11):1266–1280, 2001. [Cited on page 91]

[CRS+09] SS Chopra, J Rump, SC Schmidt, F Streitparth, C Seebauer, G Schumacher, I Van der Voort, and U Teichgräber. Imaging sequences for intraoperative mr-guided laparoscopic liver resection in 1.0-t high field open MRI. *European radiology*, 19(9):2191–2196, 2009. [Cited on page 27]

[CSB15] Ivan Cabrilo, Karl Schaller, and Philippe Bijlenga. Augmented reality-assisted bypass surgery: Embracing minimal invasiveness. *World neurosurgery*, 83(4):596–602, 2015. [Cited on page 32]

[CV95] James E. Cutting and Peter M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In William Epstein and Sheena Rogers, editors, *Perception of Space and Motion*, Handbook of Perception and Cognition, chapter 3, pages 69–117. Academic Press, 2nd edition edition, 1995. [Cited on pages 108 and 109]

[CVR+09] Denis Castaing, Eric Vibert, Luana Ricca, Daniel Azoulay, Rene Adam, and Brice Gayet. Oncologic results of laparoscopic versus open hepatectomy for colorectal liver metastases in two specialized centers. *Annals of surgery*, 250(5):849–855, 2009. [Cited on page 19]

[CWM+10] Carling L Cheung, Chris Wedlake, John Moore, Stephen E Pautler, and Terry M Peters. Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 408–415. Springer, 2010. [Cited on page 117]

[CZT+14] Matthew J Clarkson, Gergely Zombori, Steve Thompson, Johannes Totz, Yi Song, Miklos Espak, Stian Johnsen, David Hawkes, and Sébastien Ourselin. The niftk software platform for image-guided interventions: platform overview and niftylink messaging. *International journal of computer assisted radiology and surgery*, 10(3):301–316, 2014. [Cited on page 122]

[DBB08] Benoît Dagon, Charles Baur, and Vincent Bettschart. Real-time update of 3D deformable models for computer aided liver surgery. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008. [Cited on page 46]

[DBMDS07] Stijn De Buck, Frederik Maes, André DHoore, and Paul Suetens. Evaluation of a novel calibration technique for optically tracked oblique laparoscopes. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 467–474. Springer, 2007. [Cited on page 105]

[DBWC14] Patrick Dubach, Brett Bell, Stefan Weber, and Marco Caversaccio. Image-guided otorhinolaryngology. In *Intraoperative Imaging and Image-Guided Therapy*, pages 845–856. Springer, 2014. [Cited on page 32]

[DCF+99] Laurent Desbat, Guillaume Champleboux, Markus Fleute, P Komarek, Catherine Mennessier, B Monteil, Thomas Rodet, P Bessou, Max Coulomb, and Gilbert Ferretti. 3d interventional imaging with 2d x-ray detectors. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI'99*, pages 973–980. Springer, 1999. [Cited on page 24]

[DCY03] Fani Deligianni, Adrian Chung, and Guang-Zhong Yang. pq-space based 2D/3D registration for endoscope tracking. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pages 311–318. Springer, 2003. [Cited on page 91]

[DDC+13] Benjamin J Dixon, Michael J Daly, Harley Chan, Allan D Vescan, Ian J Witterick, and Jonathan C Irish. Surgeons blinded by enhanced navigation: the

effect of augmented reality on attention. *Surgical endoscopy*, 27(2):454–461, 2013. [Cited on page 107]

[DF01] Frederic Devernay and Olivier Faugeras. Straight lines have to be straight. *Machine vision and applications*, 13(1):14–24, 2001. [Cited on page 85]

[DFS08] J.-D. Durou, M. Falcone, and M. Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008. [Cited on page 91]

[dSSK+14] Thiago Ramos dos Santos, Alexander Seitel, Thomas Kilgus, Stefan Suwelack, Anna-Laura Wekerle, Hannes Kenngott, Stefanie Speidel, Heinz-Peter Schlemmer, Hans-Peter Meinzer, Tobias Heimann, et al. Pose-independent surface matching for intra-operative soft-tissue marker-less registration. *Medical image analysis*, 18(7):1101–1114, 2014. [Cited on page 44]

[EC08] E Christopher Ellison and Larry C Carey. Lessons learned from the evolution of the laparoscopic revolution. *Surgical Clinics of North America*, 88(5):927–941, 2008. [Cited on page 19]

[ECG+07] Andinet Enquobahrie, Patrick Cheng, Kevin Gary, Luis Ibanez, David Gobbi, Frank Lindseth, Ziv Yaniv, Stephen Aylward, Julien Jomier, and Kevin Cleary. The image-guided surgery toolkit igstk: an open source c++ software toolkit. *Journal of Digital Imaging*, 20(1):21–33, 2007. [Cited on page 121]

[EPY+15] Philip Edgcumbe, Philip Pratt, Guang-Zhong Yang, Christopher Nguan, and Robert Rohling. Pico lantern: Surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector. *Medical Image Analysis*, 2015. [Cited on page 43]

[FBKC+12] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging*, 30(9):1323–1341, 2012. [Cited on page 121]

[FCN+10] Norio Fukuda, Yen-Wei Chen, Masahiko Nakamoto, Toshiyuki Okada, and Yoshinobu Sato. A scope cylinder rotation tracking method for oblique-viewing endoscopes without attached sensing device. In *Software Engineering and Data Mining (SEDM), 2010 2nd International Conference on*, pages 684–687. IEEE, 2010. [Cited on page 105]

[FDM+05] Gabor Fichtinger, Anton Deguet, Ken Masamune, Emese Balogh, Gregory S Fischer, Herve Mathieu, Russell H Taylor, S James Zinreich, and Laura M Fayad. Image overlay guidance for needle insertion in CT scanner. *Biomedical Engineering, IEEE Transactions on*, 52(8):1415–1424, 2005. [Cited on page 36]

[FES+10] Julia Finkelstein, Elisabeth Eckersberger, Helen Sadri, Samir S Taneja, Herbert Lepor, and Bob Djavan. Open versus laparoscopic versus robot-assisted laparoscopic prostatectomy: the european and us experience. *Reviews in urology*, 12(1):35, 2010. [Cited on page 20]

[FHB+14] Alfred M Franz, Tamas Haidegger, Wolfgang Birkfellner, Kevin Cleary, Terry M Peters, and Lena Maier-Hein. Electromagnetic tracking in medicinea review of technology, validation, and applications. *Medical Imaging, IEEE Transactions on*, 33(8):1702–1725, 2014. [Cited on page 118]

[FKB+13] Barbara Flach, Jan Kuntz, Marcus Brehm, Rolf Kueres, Sönke Bartling, and Marc Kachelrieß. Low dose tomographic fluoroscopy: 4d intervention guidance with running prior. *Medical physics*, 40(10):101909, 2013. [Cited on page 15]

[FLD02] Markus Fleute, Stéphane Lavallée, and Laurent Desbat. Integrated approach for matching statistical shape models with intra-operative 2d and 3d data. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2002*, pages 364–372. Springer, 2002. [Cited on page 58]

[FLR+98] Henry Fuchs, Mark A Livingston, Ramesh Raskar, Kurtis Keller, Jessica R Crawford, Paul Rademacher, Samuel H Drake, Anthony A Meyer, et al. *Augmented reality visualization for laparoscopic surgery.* Springer, 1998. [Cited on page 36]

[FMHN08] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab. Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. *IEEE Transactions on Medical Imaging*, 27(3):355–369, 2008. [Cited on pages 39, 46, 47, 106, and 116]

[FMM+15] Ryo Furukawa, Ryunosuke Masutani, Daisuke Miyazaki, Masashi Baba, Shinsaku Hiura, Marco Visentini-Scarzanella, Hiroki Morinaga, Hiroshi Kawasaki, and Ryusuke Sagawa. 2-dof auto-calibration for a 3d endoscope system based on active stereo. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 7937–7941. IEEE, 2015. [Cited on page 43]

[FRH+10] M. Figl, D. Rueckert, D. Hawkes, R. Casula, M. Hu, O. Pedro, D. P. Zhang, G. Penney, F. Bello, and P. Edwards. Image guidance for robotic minimally invasive coronary artery bypass. *Computerized Medical Imaging and Graphics*, 34(1):61–68, 2010. [Cited on pages 45 and 91]

[FRV+09] Marco Feuerstein, Tobias Reichl, Jakob Vogel, Joerg Traub, and Nassir Navab. Magneto-optical tracking of flexible laparoscopic ultrasound: model-based online detection and correction of magnetic tracking errors. *Medical Imaging, IEEE Transactions on*, 28(6):951–967, 2009. [Cited on page 118]

[FSS+12] AM Franz, A Seitel, M Servatius, C Zöllner, I Gergel, I Wegner, J Neuhaus, S Zelzer, M Nolden, J Gaa, et al. Simplified development of image-guided therapy software with mitk-igt. In *SPIE medical imaging*, pages 83162J–83162J. International Society for Optics and Photonics, 2012. [Cited on page 121]

[FVW12] Moti Freiman, Stephan D Voss, and Simon Keith Warfield. Abdominal images non-rigid registration using local-affine diffeomorphic demons. In *Abdominal Imaging. Computational and Clinical Applications*, pages 116–124. Springer, 2012. [Cited on page 46]

[FZB09] Nicolas Foroglou, Amir Zamani, and Peter Black. Intra-operative MRI (iopmr) for brain tumour surgery. *British journal of neurosurgery*, 23(1):14–22, 2009. [Cited on page 27]

[GBC+14] Oscar G Grasa, Ernesto Bernal, Santiago Casado, Iñigo Gil, and JM Montiel. Visual slam for handheld monocular endoscope. *Medical Imaging, IEEE Transactions on*, 33(1):135–146, 2014. [Cited on page 42]

[GH11] Angela D Gupta and Misop Han. Imaging guidance in minimally invasive prostatectomy. In *Urologic Oncology: Seminars and Original Investigations*, volume 29, pages 343–346. Elsevier, 2011. [Cited on page 27]

[GKL+07] Inderbir S Gill, Louis R Kavoussi, Brian R Lane, Michael L Blute, Denise Babineau, J Roberto Colombo, Igor Frank, Sompol Permpongkosol, Christopher J Weight, Jihad H Kaouk, et al. Comparison of 1,800 laparoscopic and open partial nephrectomies for single renal tumors. *The Journal of urology*, 178(1):41–46, 2007. [Cited on page 19]

[GML+04] J. Garcia, P. Malik, J. Liu, J. Kowal, G. Marti, I. Pappas, and M. Caversaccio. Image-guided surgical microscope with mounted mini-tracker. In *International Congress Series*, volume 1268, page 1311, 2004. [Cited on page 117]

[GSH+11] A Groch, A Seitel, S Hempel, S Speidel, R Engelbrecht, J Penne, K Höller, S Röhl, K Yung, S Bodenstedt, et al. 3D surface reconstruction for laparoscopic computer-assisted interventions: comparison of state-of-the-art methods. In *SPIE Medical Imaging*, pages 796415–796415. International Society for Optics and Photonics, 2011. [Cited on page 43]

[GTB+11] Daniel J Greene, Christopher F Tenggadjaja, Ryan J Bowman, Gautum Agarwal, Kamyar Y Ebrahimi, and D Duane Baldwin. Comparison of a reduced radiation fluoroscopy protocol to conventional fluoroscopy during uncomplicated ureteroscopy. *Urology*, 78(2):286–290, 2011. [Cited on page 24]

[GVSY13] Stamatia Giannarou, Marco Visentini-Scarzanella, and Guang-Zhong Yang. Probabilistic tracking of affine-invariant anisotropic regions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(1):130–143, 2013. [Cited on page 119]

[GX14] Junping Geng and Junfeng Xie. Review of 3-D endoscopic surface imaging techniques. *Sensors Journal, IEEE*, 14(4):945–960, 2014. [Cited on page 41]

[HDH+02] Jae-Sung Hong, Takeyoshi Dohi, Makoto Hasizume, Kozo Konishi, and Nobuhiko Hata. A motion adaptable needle placement instrument based on tumor specific ultrasonic image segmentation. In *Medical Image Computing and Computer-Assisted InterventionMICCAI 2002*, pages 122–129. Springer, 2002. [Cited on page 55]

[HDP+13] Nazim Haouchine, Jérémie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stéphane Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *Mixed and*

Augmented Reality (ISMAR), 2013 IEEE International Symposium on, pages 199–208. IEEE, 2013. [Cited on pages 42, 44, 46, and 119]

[HH05] Andrew G. Harrell and B. Todd Heniford. Minimally invasive abdominal surgery: lux et veritas past, present, and future. *The American journal of surgery*, 190(2):239–243, 2005. [Cited on page 19]

[HHCS+03] Thomas Hartkens, Derek LG Hill, Andy D Castellano-Smith, David J Hawkes, Calvin R Maurer Jr, Alastair J Martin, Walter A Hall, Haiying Liu, and Charles L Truwit. Measurement and analysis of brain deformation during neurosurgery. *Medical Imaging, IEEE Transactions on*, 22(1):82–92, 2003. [Cited on page 4]

[HHMM+14] Archie Hughes-Hallett, Erik K Mayer, Hani J Marcus, Thomas P Cundy, Philip J Pratt, Ara W Darzi, and Justin A Vale. Augmented reality partial nephrectomy: examining the current status and future perspectives. *Urology*, 83(2):266–273, 2014. [Cited on pages 110 and 114]

[HIH+15] Yuichiro Hayashi, Tsuyoshi Igami, Tomoaki Hirose, Masato Nagino, and Kensaku Mori. Development and clinical application of surgical navigation system for laparoscopic hepatectomy. In *SPIE Medical Imaging*, pages 94151X–94151X. International Society for Optics and Photonics, 2015. [Cited on page 117]

[HKMG09] Stanley D Herrell, David Morgan Kwartowitz, Paul M Milhoua, and Robert L Galloway. Toward image guided robotic surgery: system validation. *The Journal of urology*, 181(2):783–790, 2009. [Cited on pages 31 and 43]

[HLF+07] Judy Hung, Roberto Lang, Frank Flachskampf, Stanton K Shernan, Marti L McCulloch, David B Adams, James Thomas, Mani Vannan, and Thomas Ryan. 3D echocardiography: a review of the current status and future directions. *Journal of the American Society of Echocardiography*, 20(3):213–233, 2007. [Cited on page 27]

[HM09] Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3D medical image segmentation: a review. *Medical image analysis*, 13(4):543–563, 2009. [Cited on pages 33 and 44]

[HMS+05] Cristiano GS Huscher, Andrea Mingoli, Giovanna Sgarzini, Andrea Sansonetti, Massimiliano Di Paola, Achille Recher, and Cecilia Ponzano. Laparoscopic versus open subtotal gastrectomy for distal gastric cancer: five-year results of a randomized prospective trial. *Annals of surgery*, 241(2):232, 2005. [Cited on page 19]

[HNY+06] Shozo Hirota, Norio Nakao, Satoshi Yamamoto, Kaoru Kobayashi, Hiroaki Maeda, Reiichi Ishikura, Koui Miura, Kiyoshi Sakamoto, Ken Ueda, and Rika Baba. Cone-beam CT with flat-panel-detector digital angiography system: early experience in abdominal interventional procedures. *Cardiovascular and interventional radiology*, 29(6):1034–1038, 2006. [Cited on page 26]

[HPF+12] Mingxing Hu, Graeme Penney, Michael Figl, Philip Edwards, Fernando Bello, Roberto Casula, Daniel Rueckert, and David Hawkes. Reconstruction of a 3D surface from video that is robust to missing data and outliers: Application to minimally invasive surgery using stereo and mono endoscopes. *Medical image analysis*, 16(3):597–611, 2012. [Cited on page 41]

[HPR+09] M. Hu, G. P. Penney, D. Rueckert, P. J. Edwards, F. Bello, R. Casula, M. Figl, and D. J. Hawkes. Non-rigid reconstruction of the beating heart surface for minimally invasive cardiac surgery. In *Proceedings of MICCAI*, pages 34–42, 2009. [Cited on page 42]

[HSN06] Mitsuhiro Hayashibe, Naoki Suzuki, and Yoshihiko Nakamura. Laser-scan endoscope system for intraoperative geometry acquisition and surgical robot safety management. *Medical Image Analysis*, 10(4):509–519, 2006. [Cited on page 43]

[HVGS+09] Tobias Heimann, Bram Van Ginneken, Martin Styner, Yulia Arzhaeva, Volker Aurich, Christian Bauer, Andreas Beck, Christoph Becker, Reinhard Beichel, György Bekes, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *Medical Imaging, IEEE Transactions on*, 28(8):1251–1265, 2009. [Cited on pages 33 and 46]

[HWR+10] Christian Hansen, Jan Wieferich, Felix Ritter, Christian Rieder, and Heinz-Otto Peitgen. Illustrative visualization of 3D planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2):133–141, 2010. [Cited on pages 109 and 110]

[HWT02] Ian Harris, Peter M Walker, and Lawrence Trieu. Radiation exposure using laser aiming guide in orthopedic procedures. *ANZ journal of surgery*, 72(5):349–351, 2002. [Cited on page 24]

[HZB+10] Oleg Heizmann, Stephan Zidowitz, Holger Bourquain, Silke Potthast, Heinz-Otto Peitgen, Daniel Oertli, and Christoph Kettelhack. Assessment of intraoperative liver deformation during hepatic resection: prospective clinical study. *World journal of surgery*, 34(8):1887–1893, 2010. [Cited on page 38]

[IUK+12] Satoshi Ieiri, Munenori Uemura, Kouzou Konishi, Ryota Souzaki, Yoshihiro Nagao, Norifumi Tsutsumi, Tomohiko Akahoshi, Kenoki Ohuchida, Takeshi Ohdaira, Morimasa Tomikawa, et al. Augmented reality navigation system for laparoscopic splenectomy in children based on preoperative CT image using optical tracking device. *Pediatric surgery international*, 28(4):341–346, 2012. [Cited on pages 105 and 116]

[JC04] Z. Jankó and D. Chetverikov. Photo-consistency based registration of an uncalibrated image pair to a 3D surface model using genetic algorithm. In *Proc. of 3D Data Processing, Visualization and Transmission*, pages 616–622, 2004. [Cited on page 91]

[JKB+04] Frank Willem Jansen, Wendela Kolkman, Erica A Bakkum, Cor D de Kroon, Trudy CM Trimbos-Kemper, and J Baptist Trimbos. Complications of laparoscopy: an inquiry about closed-versus open-entry technique. *American journal of obstetrics and gynecology*, 190(3):634–638, 2004. [Cited on page 19]

[JMPC13] Uditha L Jayarathne, A Jonathan McLeod, Terry M Peters, and Elvis CS Chen. Robust intraoperative us probe tracking using a monocular endoscopic camera. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 363–370. Springer, 2013. [Cited on page 117]

[JSWM02] David A Jaffray, Jeffrey H Siewerdsen, John W Wong, and Alvaro A Martinez. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 53(5):1337–1349, 2002. [Cited on page 26]

[KBCH12] Jae-Hak Kim, Adrien Bartoli, Toby Collins, and Richard Hartley. Tracking by detection for interactive image augmentation in laparoscopy. In *Biomedical Image Registration*, pages 246–255. Springer, 2012. [Cited on page 33]

[KDD+10] P Kazanzides, S DiMaio, A Deguet, B Vagvolgyi, M Balicki, C Schneider, R Kumar, A Jog, B Itkowitz, C Hasser, et al. The surgical assistant workstation (saw) in minimally-invasive surgery and microsurgery. In *MICCAI Workshop on Systems and Arch. for Computer Assisted Interventions*, 2010. [Cited on page 121]

[KDE+12] Markus Kleemann, Steffen Deichmann, Hamed Esnaashari, Armin Besirevic, Osama Shahin, Hans-Peter Bruch, and Tilman Laubert. Laparoscopic navigated liver resection: technical aspects and clinical practice in benign liver tumors. *Case reports in surgery*, 2012, 2012. [Cited on page 117]

[KHB+13] Thomas Köhler, Sven Haase, Sebastian Bauer, Jakob Wasza, Thomas Kilgus, Lena Maier-Hein, Hubertus Feußner, and Joachim Hornegger. Tof meets rgb: novel multi-sensor super-resolution for hybrid 3-d endoscopy. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 139–146. Springer, 2013. [Cited on page 43]

[KHN+05] Kozo Konishi, Makoto Hashizume, Masahiko Nakamoto, Yoshihiro Kakeji, Ichiro Yoshino, Akinobu Taketomi, Yoshinobu Sato, Shinichi Tamura, and Yoshihiko Maehara. Augmented reality navigation system for endoscopic surgery based on three-dimensional ultrasound and computed tomography: Application to 20 clinical cases. In *International Congress Series*, volume 1281, pages 537–542. Elsevier, 2005. [Cited on pages 32, 38, 47, and 118]

[KMS07] Denis Kalkofen, Erick Mendez, and Dieter Schmalstieg. Interactive focus and context visualization for augmented reality. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–10. IEEE Computer Society, 2007. [Cited on page 108]

[KNZI02] Ryo Kurazume, Ko Nishino, Zhengyou Zhang, and Katsushi Ikeuchi. Simultaneous 2D images and 3D geometric model registration for texture mapping

utilizing reflectance attribute. *Proc. Fifth ACCV*, pages 99–106, 2002. [Cited on page 91]

[KOJC12] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. Dvv: a taxonomy for mixed reality visualization in image guided surgery. *Visualization and Computer Graphics, IEEE Transactions on*, 18(2):332–352, 2012. [Cited on pages 49 and 114]

[KOJC13] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. The state of the art of visualization in mixed reality image guided surgery. *Computerized Medical Imaging and Graphics*, 37(2):98–112, 2013. [Cited on pages 34, 107, and 114]

[KOW+14] Xin Kang, Jihun Oh, Emmanuel Wilson, Ziv Yaniv, Timothy D Kane, Craig A Peters, and Raj Shekhar. Towards a clinical stereoscopic augmented reality system for laparoscopic surgery. In *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, pages 108–116. Springer, 2014. [Cited on pages 35 and 116]

[Kro00] Arnon Krongrad. Laparoscopic radical prostatectomy. *Current urology reports*, 1(1):36–40, 2000. [Cited on page 21]

[KRPH05] CH Kau, S Richmond, JM Palomo, and MG Hans. Current products and practice: Three-dimensional cone beam computerized tomography in orthodontics. *Journal of orthodontics*, 32(4):282–293, 2005. [Cited on page 26]

[KtMS+11] Pieter L Kubben, Karlien J ter Meulen, Olaf EMG Schijns, Mariël P ter Laak-Poort, Jacobus J van Overbeeke, and Henk van Santbrink. Intraoperative MRI-guided resection of glioblastoma multiforme: a systematic review. *The lancet oncology*, 12(11):1062–1070, 2011. [Cited on page 27]

[KVNL07] James N Kirkpatrick, Mani A Vannan, Jagat Narula, and Roberto M Lang. Echocardiography in heart failure: applications, utility, and new horizons. *Journal of the American College of Cardiology*, 50(5):381–396, 2007. [Cited on page 27]

[KWG+13] Darko Katić, Anna-Laura Wekerle, Jochen Görtler, Patrick Spengler, Sebastian Bodenstedt, Sebastian Röhl, Stefan Suwelack, Hannes Götz Kenngott, Martin Wagner, Beat Peter Müller-Stich, et al. Context-aware augmented reality in laparoscopic surgery. *Computerized Medical Imaging and Graphics*, 37(2):174–182, 2013. [Cited on page 108]

[KWG+14] Hannes G Kenngott, Martin Wagner, Matthias Gondan, Felix Nickel, Marco Nolden, Andreas Fetzer, Jürgen Weitz, Lars Fischer, Stefanie Speidel, Hans-Peter Meinzer, et al. Real-time image guidance in laparoscopic liver surgery: first clinical experience with a guidance system based on intraoperative CT imaging. *Surgical endoscopy*, 28(3):933–940, 2014. [Cited on page 26]

[KWN+15] HG Kenngott, M Wagner, F Nickel, AL Wekerle, A Preukschas, M Apitz, T Schulte, R Rempel, P Mietkowski, F Wagner, et al. Computer-assisted abdominal surgery: new technologies. *Langenbeck's Archives of Surgery*, 400(3):273–281, 2015. [Cited on pages 36 and 115]

[LBF+14] Miguel Lourenço, Joao P Barreto, Fernando Fonseca, Hélder Ferreira, Rui M Duarte, and Jorge Correia-Pinto. Continuous zoom calibration by tracking salient points in endoscopic video. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 456–463. Springer, 2014. [Cited on page 105]

[LCMY07] Mirna Lerotic, Adrian J Chung, George Mylonas, and Guang-Zhong Yang. pq-space based non-photorealistic rendering for augmented reality. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 102–109. Springer, 2007. [Cited on pages 108 and 110]

[LCS+08] B. Lo, A. J. Chung, D. Stoyanov, G. Mylonas, and G.-Z. Yang. Real-time intra-operative 3D tissue deformation recovery. In *Biomedical Imaging: From Nano to Macro, ISBI*, pages 1387–1390, 2008. [Cited on page 42]

[LFC+10] Pablo Lamata, Adinda Freudenthal, Alicia Cano, Denis Kalkofen, Dieter Schmalstieg, Edvard Naerum, Eigil Samset, Enrique J Gómez, Francisco M Sánchez-Margallo, Hugo Furtado, et al. Augmented reality for minimally invasive surgery: Overview and some recent advances. In Soha Maad, editor, *Augmented Reality*, chapter 5, pages 74–98. INTECH Open Access Publisher, 2010. [Cited on page 113]

[LHH+14] Ofir Livne, Ran Harel, Moshe Hadani, Roberto Spiegelmann, Zeev Feldman, and Zvi R Cohen. Intraoperative magnetic resonance imaging for resection of intra-axial brain lesions: A decade of experience using low-field magnetic resonance imaging, polestar n-10, 20, 30 systems. *World neurosurgery*, 82(5):770–776, 2014. [Cited on page 27]

[LISD10] Hongen Liao, Takashi Inomata, Ichiro Sakuma, and Takeyoshi Dohi. 3-d augmented reality for mri-guided surgery using integral videography autostereoscopic image overlay. *Biomedical Engineering, IEEE Transactions on*, 57(6):1476–1486, 2010. [Cited on page 36]

[LJD+14] Rudy J Lapeer, Samuel J Jeffrey, Josh T Dao, Gerardo González García, Minsi Chen, Steve M Shickell, Roger S Rowland, and Carl M Philpott. Using a passive coordinate measurement arm for motion tracking of a rigid endoscope for augmented-reality image-guided surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 10(1):65–77, 2014. [Cited on page 118]

[LKW+14] Xinyang Liu, Sukryool Kang, Emmanuel Wilson, Craig A Peters, Timothy D Kane, and Raj Shekhar. Evaluation of electromagnetic tracking for stereoscopic augmented reality laparoscopic visualization. In *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, pages 84–91. Springer, 2014. [Cited on page 117]

[LLPY07] Carl Lundstrom, Per Ljung, Anders Persson, and Anders Ynnerman. Uncertainty visualization in medical volume rendering using probabilistic animation. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1648–1655, 2007. [Cited on page 108]

[LLV+10] Su-Lin Lee, Mirna Lerotic, Valentina Vitiello, Stamatia Giannarou, Ka-Wai Kwok, Marco Visentini-Scarzanella, and Guang-Zhong Yang. From medical images to minimally invasive intervention: computer assistance for robotic surgery. *Computerized Medical Imaging and Graphics*, 34(1):33–45, 2010. [Cited on pages 110, 113, and 120]

[LLY06] Claes Lundstrom, Patric Ljung, and Anders Ynnerman. Local histograms for design of transfer functions in direct volume rendering. *Visualization and Computer Graphics, IEEE Transactions on*, 12(6):1570–1579, 2006. [Cited on page 35]

[LPH+09] Thomas Lange, Nils Papenberg, Stefan Heldmann, Jan Modersitzki, Bernd Fischer, Hans Lamecker, and Peter M Schlag. 3D ultrasound-ct registration of the liver using combined landmark-intensity information. *International journal of computer assisted radiology and surgery*, 4(1):79–88, 2009. [Cited on page 46]

[LRS+15] Wen P Liu, Jeremy D Richmon, Jonathan M Sorger, Mahdi Azizian, and Russell H Taylor. Augmented reality and cone beam CT guidance for transoral robotic surgery. *Journal of Robotic Surgery*, pages 1–11, 2015. [Cited on page 32]

[LSQ+15] Bingxiong Lin, Yu Sun, Xiaoning Qian, Dmitry Goldgof, Richard Gitlin, and Yuncheng You. Video-based 3D reconstruction, laparoscope localization and deformation recovery for abdominal minimally invasive surgery: a survey. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2015. [Cited on pages 41 and 91]

[LVR+12] Thomas Langø, Sinara Vijayan, Anna Rethy, Cecilie Våpenstad, Ole Vegard Solberg, Ronald Mårvik, Gjermund Johnsen, and Toril N Hernes. Navigated laparoscopic ultrasound in abdominal soft tissue surgery: technological overview and perspectives. *International journal of computer assisted radiology and surgery*, 7(4):585–599, 2012. [Cited on pages 27 and 118]

[MADdM12] Xavier Maurice, Chadi Albitar, Christophe Doignon, and Michel de Mathelin. A structured light-based laparoscope with real-time organs' surface reconstruction for minimally invasive surgery. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 5769–5772. IEEE, 2012. [Cited on page 43]

[MB14] A. Malti and A. Bartoli. Combining conformal deformation and the cook-torrance model for 3D reconstruction in laparoscopy. *IEEE Transactions on Biomedical Engineering*, 61(6):1684–1692, 2014. [Cited on page 42]

[MBA+94] AJ McMahon, JN Baxter, JR Anderson, G Ramsay, PJ O'Dwyer, IT Russell, S Ross, G Sutherland, D Galloway, and CG Morran. Laparoscopic versus mini-laparotomy cholecystectomy: a randomised trial. *The Lancet*, 343(8890):135–138, 1994. [Cited on page 19]

[MBC11] Abed Malti, Adrien Bartoli, and Toby Collins. Template-based conformal shape-from-motion from registered laparoscopic images. In *Conference on Medical Image Understanding and Analysis*, number 2 in 1, page 7, 2011. [Cited on page 41]

[MBF12] Rui Melo, Joao P Barreto, and Gabriel Falcao. A new solution for camera calibration and real-time image distortion correction in medical endoscopy–initial technical evaluation. *Biomedical Engineering, IEEE Transactions on*, 59(3):634–644, 2012. [Cited on page 105]

[MD97] Charles T Mehlman and Thomas G DiPasquale. Radiation exposure to the orthopaedic surgical team during fluoroscopy:" how far away is far enough?". *Journal of orthopaedic trauma*, 11(6):392–398, 1997. [Cited on page 24]

[MFB12] Rui Melo, Gabriel Falcao, and João P Barreto. Real-time hd image distortion correction in heterogeneous parallel computing systems using efficient memory access patterns. *Journal of Real-Time Image Processing*, pages 1–9, 2012. [Cited on page 105]

[MFN+14] Peter Mountney, Johannes Fallert, Stéphane A. Nicolau, Luc Soler, and Philip W. Mewes. An augmented reality framework for soft tissue surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 423–431. Springer, 2014. [Cited on pages 46, 47, and 52]

[MHFdS+12] Lena Maier-Hein, Alfred M Franz, Thiago R dos Santos, Mirko Schmidt, Markus Fangerau, Hans-Peter Meinzer, and J Michael Fitzpatrick. Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(8):1520–1532, 2012. [Cited on page 44]

[MHGB+14] L Maier-Hein, A Groch, Alberto Bartoli, S Bodenstedt, G Boissonnat, P-L Chang, NT Clancy, DS Elson, S Haase, E Heim, et al. Comparative validation of single-shot optical techniques for laparoscopic 3-d surface reconstruction. *Medical Imaging, IEEE Transactions on*, 33(10):1913–1930, 2014. [Cited on page 43]

[MHMB+13] L. Maier-Hein, P. Mountney, A. Bartoli, H. Elhawary, D. Elson, A. Groch, A. Kolb, M. Rodrigues, J. Sorger, S. Speidel, and D. Stoyanov. Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. *Medical Image Analysis Journal*, 17(8):974–996, 2013. [Cited on pages 41, 48, 91, and 106]

[MHMK+14] Lena Maier-Hein, Sven Mersmann, Daniel Kondermann, Christian Stock, Hannes Gotz Kenngott, Alexandro Sanchez, Martin Wagner, Anas Preukschas, Anna-Laura Wekerle, Stefanie Helfert, et al. Crowdsourcing for reference correspondence generation in endoscopic images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 349–356. Springer, 2014. [Cited on page 119]

[MHMMB13] Sergio E Martinez Herrera, Abed Malti, Oscar Morel, and Alberto Bartoli. Shape-from-polarization in laparoscopy. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 1412–1415. IEEE, 2013. [Cited on page 43]

[MHT+07] Dorothea Miller, Stefan Heinze, Wuttipong Tirakotai, Oliver Bozinov, Oguzkan Sürücü, Ludwig Benes, Helmut Bertalanffy, and Ulrich Sure. Is the image guidance of ultrasonography beneficial for neurosurgical routine? *Surgical neurology*, 67(6):579–587, 2007. [Cited on page 27]

[MIH11] Daniel J Mirota, Masaru Ishii, and Gregory D Hager. Vision-based navigation in image-guided interventions. *Annual review of biomedical engineering*, 13:297–319, 2011. [Cited on pages 107 and 114]

[MJB13] Uli Mezger, Claudia Jendrewski, and Michael Bartels. Navigation in surgery. *Langenbeck's Archives of Surgery*, 398(4):501–514, 2013. [Cited on page 32]

[MJHM+98] Calvin R Maurer Jr, Derek LG Hill, Alastair J Martin, Haiying Liu, Michael McCue, Daniel Rueckert, David Lloret, Walter A Hall, Robert E Maxwell, David J Hawkes, et al. Investigation of intraoperative brain deformation using a 1.5-t interventional mr system: preliminary results. *Medical Imaging, IEEE Transactions on*, 17(5):817–825, 1998. [Cited on page 27]

[MK94] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12):1321–1329, 1994. [Cited on page 30]

[MLG+01] Jacques Marescaux, Joel Leroy, Michel Gagner, Francesco Rubino, Didier Mutter, Michel Vix, Steven E Butner, and Michelle K Smith. Transatlantic robot-assisted telesurgery. *Nature*, 413(6854):379–380, 2001. [Cited on page 39]

[MLT+04] R. Mårvik, T. Langø, G. A. Tangen, J. O. Andersen, J. H. Kaspersen, B. Ystgaard, E. Sjølie, R. Fougner, H. E. Fjøsne, and T. A. Nagelhus Hernes. Laparoscopic navigation pointer for three-dimensional image-guided surgery. *Surgical endoscopy*, 18(8):1242–1248, 2004. [Cited on pages 38 and 47]

[MLZ+14] Oliver J Muensterer, Martin Lacher, Christoph Zoeller, Matthew Bronstein, and Joachim Kübler. Google glass in pediatric surgery: An exploratory study. *International Journal of Surgery*, 12(4):281–289, 2014. [Cited on page 36]

[MMR+03] M Morino, I Morra, E Rosso, C Miglietta, and C Garrone. Laparoscopic vs open hepatic resection: a comparative study. *Surgical Endoscopy and Other Interventional Techniques*, 17(12):1914–1918, 2003. [Cited on page 19]

[MMS+11] Sven Mersmann, Michael Müller, Alexander Seitel, Florian Arnegger, Ralf Tetzlaff, Julien Dinkel, Matthias Baumhauer, Bruno Schmied, Hans-Peter Meinzer, and Lena Maier-Hein. Time-of-flight camera technique for augmented reality in computer-assisted interventions. In *SPIE Medical Imaging*, pages 79642C–79642C. International Society for Optics and Photonics, 2011. [Cited on page 43]

[MNK+12] N. Mahmoud, S. Nicolau, A. Keshk, M. Ahmad, L. Soler, and J. Marescaux. Fast 3D structure from motion with missing points from registration of partial reconstructions. In *Proceedings of International Conference on Articulated Motion and Deformable Objects (AMDO'12) LNCS 7378*, pages 173–183, 2012. [Cited on page 42]

[MPHH+15] Hani J Marcus, Philip Pratt, Archie Hughes-Hallett, Thomas P Cundy, Adam P Marcus, Guang-Zhong Yang, Ara Darzi, and Dipankar Nandi. Comparative effectiveness and safety of image guidance systems in neurosurgery: a preclinical randomized study. *Journal of neurosurgery*, pages 1–7, 2015. [Cited on page 32]

[MPS+13] Ettore Marzano, Tullio Piardi, Luc Soler, Michele Diana, Didier Mutter, Jacques Marescaux, and Patrick Pessaux. Augmented reality-guided artery-first pancreatico-duodenectomy. *Journal of Gastrointestinal Surgery*, 17(11):1980–1983, 2013. [Cited on page 32]

[MRA+04] J. Marescaux, F. Rubino, M. Arenas, D. Mutter, and L. Soler. Augmented-reality–assisted laparoscopic adrenalectomy. *Journal of the American Medical Association*, 292(18):2211–2215, 2004. [Cited on pages 32, 38, and 47]

[MS10] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12):2262–2275, 2010. [Cited on page 44]

[MSLD08] Ken Masamune, Ikuma Sato, Hongen Liao, and Takeyoshi Dohi. Non-metal slice image overlay display system used inside the open type MRI. In *Medical Imaging and Augmented Reality*, pages 385–392. Springer, 2008. [Cited on page 36]

[MTLP12] Primoz Markelj, D Tomaževič, Bostjan Likar, and F Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661, 2012. [Cited on page 46]

[MUS+13] Daniel J Mirota, Ali Uneri, Stefan Schafer, Sajendra Nithiananthan, Douglas D Reh, Masaru Ishii, Gary L Gallia, Russell H Taylor, Gregory D Hager, and Jeffrey H Siewerdsen. Evaluation of a system for high-accuracy 3D image-based registration of endoscopic video to C-arm cone-beam CT for image-guided skull base surgery. *Medical Imaging, IEEE Transactions on*, 32(7):1215–1226, 2013. [Cited on page 32]

[MVFCM03] Fabien Mourgues, Thierry Vieville, Volkmar Falk, and Eve Coste-Maniere. Interactive guidance by image overlay in robot assisted coronary artery bypass. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pages 173–181. Springer, 2003. [Cited on page 39]

[MY10] P. Mountney and G.-Z. Yang. Motion compensated slam for image guided surgery. In *Proceedings of MICCAI*, pages 496–504, 2010. [Cited on pages 39, 41, 42, and 119]

[NAP+15] Masoud S Nosrati, Rafeef Abugharbieh, Jean-Marc Peyrat, Julien Abinahed, Osama Al-Alao, Abdulla Al-Ansari, and Ghassan Hamarneh. Simultaneous multi-structure segmentation and 3D non-rigid pose estimation in image-guided robotic surgery. 2015. Early access. [Cited on page 108]

[NBA+15] Stéphane A. Nicolau, Sylvain Bernhardt, Vincent Agnus, Christophe Doignon, Luc Soler, and Jacques Marescaux. Validation methodology for augmented

reality in laparoscopic surgery. *International Journal of Computer Assisted Radiology and Surgery*, 10(S1):S61–S62, 2015. [Cited on pages 9 and 48]

[NBG+08] SA Nicolau, J Brenot, L Goffin, P Graebling, L Soler, and J Marescaux. A structured light system to guide percutaneous punctures in interventional radiology. In *Photonics Europe*, pages 700016–700016. International Society for Optics and Photonics, 2008. [Cited on pages 36 and 43]

[NDA+13] S. A. Nicolau, M. Diana, V. Agnus, L. Soler, and J. Marescaux. Semi-automated augmented reality for laparoscopic surgery: First in-vivo evaluation. *International Journal of Computer Assisted Radiology and Surgery*, 8(1 supp):109–110, 2013. [Cited on pages 32, 38, 40, and 119]

[NFKF13] Tetsuo Nozaki, Yasuyoshi Fujiuchi, Akira Komiya, and Hideki Fuse. Efficacy of dynact for surgical navigation during complex laparoscopic surgery: an initial experience. *Surgical endoscopy*, 27(3):903–909, 2013. [Cited on page 26]

[NGS05] S. A. Nicolau, L. Goffin, and L. Soler. A low cost and accurate guidance system for laparoscopic surgery: Validation on an abdominal phantom. In *Proceedings of the ACM symposium on Virtual Reality Software and Technology*, pages 124–133, 2005. [Cited on page 117]

[NIM+12] Tetsuo Nozaki, Yasuhiro Iida, Akihiro Morii, Yasuyoshi Fujiuchi, and Hideki Fuse. Laparoscopic radical nephrectomy under near real-time three-dimensional surgical navigation with c-arm cone beam computed tomography. *Surgical innovation*, 19(3):263–267, 2012. [Cited on pages 26 and 28]

[NJL+05] Marios Nicolaou, Adam James, Benny PL Lo, Ara Darzi, and Guang-Zhong Yang. Invisible shadow for navigation and planning in minimal invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 25–32. Springer, 2005. [Cited on pages 109 and 110]

[NKL+12] Woo Hyun Nam, Dong-Goo Kang, Duhgoon Lee, Jae Young Lee, and Jong Beom Ra. Automatic registration between 3D intra-operative ultrasound and pre-operative CT images of the liver based on robust edge matching. *Physics in medicine and biology*, 57(1):69, 2012. [Cited on page 46]

[NMB+11] David Netuka, Václav Masopust, Tomáš Belšán, Filip Kramář, and Vladimír Beneš. *One year experience with 3.0 T intraoperative MRI in pituitary surgery.* Springer, 2011. [Cited on page 26]

[NMSA13] Nima Najmaei, Kamal Mostafavi, Sahar Shahbazi, and Mahdi Azizian. Image-guided techniques in renal and hepatic interventions. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 9(4):379–395, 2013. [Cited on pages 49 and 114]

[NMT+11] Kevin Tri Nguyen, J Wallis Marsh, Allan Tsung, J Jennifer L Steel, T Clark Gamblin, and David A Geller. Comparative benefits of laparoscopic vs open hepatic resection: a critical appraisal. *Archives of surgery*, 146(3):348–356, 2011. [Cited on page 19]

[NNS+08] Masahiko Nakamoto, Kazuhisa Nakada, Yoshinobu Sato, Kozo Konishi, Makoto Hashizume, and Shinichi Tamura. Intraoperative magnetic tracker calibration using a magneto-optic hybrid tracker for 3-d ultrasound-based navigation in laparoscopic surgery. *Medical Imaging, IEEE Transactions on*, 27(2):255–270, 2008. [Cited on page 118]

[NNZ+10] Kazuhiro Nakamura, Yukio Naya, Satoki Zenbutsu, Kazuhiro Araki, Shuko Cho, Sho Ohta, Naoki Nihei, Hiroyoshi Suzuki, Tomohiko Ichikawa, and Tatsuo Igarashi. Surgical navigation using three-dimensional computed tomography images fused intraoperatively with live video*. *Journal of Endourology*, 24(4):521–524, 2010. [Cited on page 32]

[NSMM11] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux. Augmented reality in laparoscopic surgical oncology. *Surgical oncology*, 20(3):189–201, 2011. [Cited on pages 37 and 113]

[NUFG12] Masahiko Nakamoto, Osamu Ukimura, Kenneth Faber, and Inderbir S Gill. Current progress on augmented reality visualization in endoscopic surgery. *Current opinion in urology*, 22(2):121–126, 2012. [Cited on page 114]

[NUG+08] Masahiko Nakamoto, Osamu Ukimura, Inderbir S Gill, Arul Mahadevan, Tsuneharu Miki, Makoto Hashizume, and Yoshinobu Sato. Realtime organ tracking for endoscopic augmented reality visualization using miniature wireless magnetic tracker. In *Medical Imaging and Augmented Reality*, pages 359–366. Springer, 2008. [Cited on page 120]

[NZS+13] Marco Nolden, Sascha Zelzer, Alexander Seitel, Diana Wald, Michael Müller, Alfred M Franz, Daniel Maleike, Markus Fangerau, Matthias Baumhauer, Lena Maier-Hein, et al. The medical imaging interaction toolkit: challenges and advances. *International journal of computer assisted radiology and surgery*, 8(4):607–620, 2013. [Cited on page 121]

[OGND10] Angel Osorio, Juan-Antonio Galan, Julien Nauroy, and Patricia Donars. Real time planning, guidance and validation of surgical acts using 3D segmentations, augmented reality projections and surgical tools video tracking. In *SPIE medical imaging*, pages 762529–762529. International Society for Optics and Photonics, 2010. [Cited on page 36]

[ONN+14] Ryo Oguma, Toshiya Nakaguchi, Ryoichi Nakamura, Tadashi Yamaguchi, Hiroshi Kawahira, and Hideaki Haneishi. Ultrasound image overlay onto endoscopic image by fusing 2D-3D tracking of laparoscopic ultrasound probe. In *Augmented Environments for Computer-Assisted Interventions*, pages 14–22. Springer, 2014. [Cited on pages 35 and 117]

[OOK+14] Shinji Onda, Tomoyoshi Okamoto, Masaru Kanehira, Fumitake Suzuki, Ryusuke Ito, Shuichi Fujioka, Naoki Suzuki, Asaki Hattori, and Katsuhiko Yanaga. Identification of inferior pancreaticoduodenal artery during pancreaticoduodenectomy using augmented reality-based navigation system. *Journal of hepato-biliary-pancreatic sciences*, 21(4):281–287, 2014. [Cited on page 49]

[OOY+14] Tomoyoshi Okamoto, Shinji Onda, Katsuhiko Yanaga, Naoki Suzuki, and Asaki Hattori. Clinical application of navigation surgery using augmented reality in the abdominal field. *Surgery today*, 45(4):397–406, 2014. [Cited on pages 49 and 114]

[oSTSG+04] Clinical Outcomes of Surgical Therapy Study Group et al. A comparison of laparoscopically assisted and open colectomy for colon cancer. *The New England journal of medicine*, 350(20):2050, 2004. [Cited on page 19]

[OWK+08] Robert C Orth, Michael J Wallace, Michael D Kuo, Technology Assessment Committee of the Society of Interventional Radiology, et al. C-arm cone-beam CT: general principles and technical considerations for use in interventional radiology. *Journal of Vascular and Interventional Radiology*, 19(6):814–820, 2008. [Cited on page 26]

[OZM+13] O. Oktay, L. Zhang, T. Mansi, P. Mountney, P. Mewes, S. Nicolau, L. Soler, and C. Chefdhotel. Biomechanically driven registration of pre- to intra- operative 3D images for laparoscopic surgery. In *Proceedings of MICCAI*, pages 1–9, 2013. [Cited on page 46]

[PB12] D. Pizarro and A. Bartoli. Feature-based deformable surface detection with self-occlusion reasoning. *International Journal of Computer Vision*, 97(1):54–70, 2012. [Cited on page 41]

[PBB+07] Philipp Pohlenz, Marco Blessmann, Felix Blake, Sven Heinrich, Rainer Schmelzle, and Max Heiland. Clinical indications and perspectives for intraoperative cone-beam computed tomography in oral and maxillofacial surgery. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 103(3):412–417, 2007. [Cited on page 26]

[PBDY14] Philip Pratt, Christos Bergeles, Ara Darzi, and Guang-Zhong Yang. Practical intraoperative stereo camera calibration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 667–675. Springer, 2014. [Cited on page 106]

[PBHC+05] Jean Pouliot, Ali Bani-Hashemi, Josephine Chen, Michelle Svatos, Farhad Ghelmansarai, Matthias Mitschke, Michele Aubin, Ping Xia, Olivier Morin, Kara Bucci, et al. Low-dose megavoltage cone-beam CT for radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 61(2):552–560, 2005. [Cited on page 26]

[PBKE12] Kari Pulli, Anatoly Baksheev, Kirill Kornyakov, and Victor Eruhimov. Real-time computer vision with opencv. *Communications of the ACM*, 55(6):61–69, 2012. [Cited on page 122]

[PC08] Terry Peters and Kevin Cleary. *Image-guided interventions: technology and applications*. Springer Science & Business Media, 2008. [Cited on page 113]

[PCF06] Emmanuel Prados, Fabio Camilli, and Olivier Faugeras. A unifying and rigorous shape from shading method adapted to realistic data and applications. *Journal of Mathematical Imaging and Vision*, 25(3):307–328, 2006. [Cited on page 91]

[PDS+15] Patrick Pessaux, Michele Diana, Luc Soler, Tullio Piardi, Didier Mutter, and Jacques Marescaux. Towards cybernetic surgery: robotic and augmented reality-assisted liver segmentectomy. *Langenbeck's Archives of Surgery*, 400(3):381–385, 2015. [Cited on page 32]

[Pet06] Terry M Peters. Image-guidance for surgical procedures. *Physics in medicine and biology*, 51(14):R505, 2006. [Cited on page 32]

[PHH13] Aristotelis Perrakis, Werner Hohenberger, and Thomas Horbach. Integrated operation systems and voice recognition in minimally invasive surgery: comparison of two systems. *Surgical endoscopy*, 27(2):575–579, 2013. [Cited on page 111]

[PHK04] Steve Pieper, Michael Halle, and Ron Kikinis. 3D slicer. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 632–635. IEEE, 2004. [Cited on page 121]

[PHS+09] Jochen Penne, Kurt Höller, Michael Stürmer, Thomas Schrauder, Armin Schneider, Rainer Engelbrecht, Hubertus Feußner, Bernhard Schmauss, and Joachim Hornegger. Time-of-flight 3-d endoscopy. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 467–474. Springer, 2009. [Cited on page 43]

[PJHH+15] Philip Pratt, Alexander Jaeger, Archie Hughes-Hallett, Erik Mayer, Justin Vale, Ara Darzi, Terry Peters, and Guang-Zhong Yang. Robust ultrasound probe tracking: initial clinical experiences during robot-assisted partial nephrectomy. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–9, 2015. [Cited on page 35]

[PMJ09] Perrine Paul, Xavier Morandi, and Pierre Jannin. A surface registration method for quantification of intraoperative brain deformations in image-guided neurosurgery. *Information Technology in Biomedicine, IEEE Transactions on*, 13(6):976–983, 2009. [Cited on page 119]

[PMV+12] Philip Pratt, Erik Mayer, Justin Vale, Daniel Cohen, Eddie Edwards, Ara Darzi, and Guang-Zhong Yang. An effective visualisation and registration system for image-guided robotic partial nephrectomy. *Journal of Robotic Surgery*, 6(1):23–31, 2012. [Cited on pages 4, 40, and 108]

[PSCM+14] Gustavo Puerto-Souza, Jeffrey Cadeddu, Gian-Luca Mariottini, et al. Toward long-term and accurate augmented-reality for monocular endoscopic videos. *Biomedical Engineering, IEEE Transactions on*, 61(10):2609–2620, 2014. [Cited on pages 40, 44, 46, and 119]

[PSM+13] Gustavo Puerto-Souza, Gian-Luca Mariottini, et al. A fast and accurate feature-matching algorithm for minimally-invasive endoscopic images. *Medical Imaging, IEEE Transactions on*, 32(7):1201–1214, 2013. [Cited on page 119]

[PVH09] Zachary Pezzementi, Sandrine Voros, and Gregory D Hager. Articulated object tracking by rendering consistent appearance parts. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 3940–3947. IEEE, 2009. [Cited on page 106]

[PXP00] Dzung L Pham, Chenyang Xu, and Jerry L Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2:315–337, 2000. [Cited on page 32]

[RBB+11] Marc Ruiz, Anton Bardera, Imma Boada, Ivan Viola, Miquel Feixas, and Mateu Sbert. Automatic transfer functions based on informational divergence. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):1932–1941, 2011. [Cited on page 35]

[RBH+11] Felix Ritter, Tobias Boskamp, André Homeyer, Hendrik Laue, Michael Schwier, Florian Link, and Heinz-Otto Peitgen. Medical image analysis: A visual approach. *Pulse, IEEE*, 2(6):60–70, 2011. [Cited on page 122]

[RBS+12] Sebastian Röhl, Sebastian Bodenstedt, Stefan Suwelack, Hannes Kenngott, Beat P Müller-Stich, Rüdiger Dillmann, and Stefanie Speidel. Dense gpu-enhanced surface reconstruction from stereo endoscopic images for intraoperative registration. *Medical physics*, 39(3):1632–1645, 2012. [Cited on pages 42 and 45]

[RFS+00] Y Raja Rampersaud, Kevin T Foley, Alfred C Shen, Scott Williams, and Milo Solomito. Radiation exposure to the spine surgeon during fluoroscopically assisted pedicle screw insertion. *Spine*, 25(20):2637–2645, 2000. [Cited on page 24]

[RHFL10] Stephan Reichelt, Ralf Häussler, Gerald Fütterer, and Norbert Leister. Depth cues in human visual perception and their realization in 3d displays. In *SPIE Defense, Security, and Sensing*, pages 76900B–76900B. International Society for Optics and Photonics, 2010. [Cited on page 109]

[ROU+12] S Reaungamornrat, Yoshito Otake, Ali Uneri, Sebastian Schafer, DJ Mirota, Sajendra Nithiananthan, J Webster Stayman, Gerhard Kleinszig, A Jay Khanna, Russell H Taylor, et al. An on-board surgical tracking and video augmentation system for c-arm image guidance. *International journal of computer assisted radiology and surgery*, 7(5):647–665, 2012. [Cited on page 116]

[RP09] S Rossitti and M Pfister. 3D road-mapping in the endovascular treatment of cerebral aneurysms and arteriovenous malformations. *Interventional Neuroradiology*, 15(3):283–290, 2009. [Cited on page 26]

[RS11] Daniel Rueckert and Julia A Schnabel. Medical image registration. In *Biomedical Image Processing*, pages 131–154. Springer, 2011. [Cited on page 46]

[RSH+86] David W Roberts, John W Strohbehn, John F Hatch, William Murray, and Hans Kettenberger. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *Journal of neurosurgery*, 65(4):545–549, 1986. [Cited on pages 4 and 32]

[SBHN06] Tobias Sielhorst, Christoph Bichlmeier, Sandro Michael Heining, and Nassir Navab. Depth perception–a major issue in medical ar: evaluation study by twenty surgeons. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 364–372. Springer, 2006. [Cited on page 108]

[SBK+09] Stefanie Speidel, Julia Benzko, Sebastian Krappe, Gunther Sudra, Pedram Azad, Beat Peter Müller-Stich, Carsten Gutt, and Rüdiger Dillmann. Automatic classification of minimally invasive instruments based on endoscopic image sequences. In *SPIE Medical Imaging*, pages 72610A–72610A. International Society for Optics and Photonics, 2009. [Cited on page 106]

[SBM+11] Tobias Simpfendörfer, Matthias Baumhauer, Michael Müller, Carsten N Gutt, Hans-Peter Meinzer, Jens J Rassweiler, Selcuk Guven, and Dogu Teber. Augmented reality visualization during laparoscopic radical prostatectomy. *Journal of endourology*, 25(12):1841–1845, 2011. [Cited on page 120]

[SBMJ+02] R. Shahidi, M. R. Bax, C. R. Maurer Jr, J. A. Johnson, E. P. Wilkinson, B. Wang, J. B. West, M. J. Citardi, K. H. Manwaring, and R. Khadem. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Transactions on Medical Imaging*, 21(12):1524–1535, 2002. [Cited on page 117]

[SDB+10] R. Shekhar, O. Dandekar, V. Bhat, M. Philip, P. Lei, C. Godinez, E. Sutton, I. George, S. Kavic, and R. Mezrich et al. Live augmented reality: a new visualization method for laparoscopic surgery using continuous volumetric computed tomography. *Surgical endoscopy*, 24(8):1976–1985, 2010. [Cited on pages 46 and 107]

[SDIH+10] Caitlin M Schneider, Gregory W Dachs II, Christopher J Hasser, Michael A Choti, Simon P DiMaio, and Russell H Taylor. Robot-assisted laparoscopic ultrasound. In *Information Processing in Computer-Assisted Interventions*, pages 67–80. Springer, 2010. [Cited on page 118]

[SDJM12] Amber L Simpson, Prashanth Dumpuri, William R Jarnagin, and Michael I Miga. Model-assisted image-guided liver surgery using sparse intraoperative data. In *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, pages 7–40. Springer, 2012. [Cited on pages 44 and 46]

[SDP13] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013. [Cited on page 46]

[SET+14] Antonio Siniscalchi, Giorgio Ercolani, Giulia Tarozzi, Lorenzo Gamberini, Lucia Cipolat, Antonio D Pinna, and Stefano Faenza. Laparoscopic versus open liver resection: Differences in intraoperative and early postoperative outcome among cirrhotic patients with hepatocellular carcinomaa retrospective observational study. *HPB Surgery*, 2014, 2014. [Cited on page 19]

[SF08] William C Scarfe and Allan G Farman. What is cone-beam CT and how does it work? *Dental Clinics of North America*, 52(4):707–730, 2008. [Cited on pages 24 and 26]

[SFN08] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *Journal of Display Technology*, 4(4):451–467, 2008. [Cited on pages 37 and 113]

[SFSA12] Christoph Schmalz, Frank Forster, Anton Schick, and Elli Angelopoulou. An endoscopic 3D scanner based on structured light. *Medical image analysis*, 16(5):1063–1072, 2012. [Cited on page 43]

[SFT+06] Tobias Sielhorst, Marco Feuerstein, Joerg Traub, Oliver Kutter, and Nassir Navab. Campar: A software framework guaranteeing quality for medical augmented reality. *International Journal of Computer Assisted Radiology and Surgery*, 1:29, 2006. [Cited on page 49]

[SG09] M Alper Selver and Cüneyt Güzeliş. Semiautomatic transfer function initialization for abdominal visualization using self-generating hierarchical radial basis function networks. *Visualization and Computer Graphics, IEEE Transactions on*, 15(3):395–409, 2009. [Cited on page 35]

[Sho96] Thomas B Shope. Radiation-induced skin injuries from fluoroscopy. *Radiographics*, 16(5):1195–1199, 1996. [Cited on page 24]

[SHR+13] Damien L Smith, Jonathan P Heldt, Gideon D Richards, Gautum Agarwal, Wayne G Brisbane, Catherine J Chen, Joshua D Chamberlin, and D Duane Baldwin. Radiation exposure during continuous and pulsed fluoroscopy. *Journal of Endourology*, 27(3):384–388, 2013. [Cited on page 24]

[SHT09] Aleksandr Segal, Dirk Haehnel, and Sebastian Thrun. Generalized-icp. In *Robotics: Science and Systems*, volume 2 of *4*, 2009. [Cited on page 44]

[Shu04] Jeffrey H Shuhaiber. Augmented reality in surgery. *Archives of surgery*, 139(2):170–174, 2004. [Cited on page 32]

[SIU+13] Ryota Souzaki, Satoshi Ieiri, Munenori Uemura, Kenoki Ohuchida, Morimasa Tomikawa, Yoshiaki Kinoshita, Yuhki Koga, Aiko Suminoe, Kenichi Kohashi, Yoshinao Oda, et al. An augmented reality navigation system for pediatric oncologic surgery based on preoperative CT and MRI images. *Journal of pediatric surgery*, 48(12):2479–2483, 2013. [Cited on pages 47 and 116]

[SLT+07] Ole Vegard Solberg, Frank Lindseth, Hans Torp, Richard E Blake, and Toril A Nagelhus Hernes. Freehand 3D ultrasound reconstruction algorithmsa review. *Ultrasound in medicine & biology*, 33(7):991–1009, 2007. [Cited on page 28]

[SMB+05] JH Siewerdsen, DJ Moseley, S Burch, SK Bisland, A Bogaards, BC Wilson, and DA Jaffray. Volume CT with a flat-panel detector on a mobile, isocentric C-arm: pre-clinical investigation in guidance of minimally invasive surgery. *Medical physics*, 32(1):241–254, 2005. [Cited on page 26]

[SML+08] Danail Stoyanov, George P Mylonas, Mirna Lerotic, Adrian J Chung, and Guang-Zhong Yang. Intra-operative visualizations: perceptual fidelity and human factors. *Display Technology, Journal of*, 4(4):491–501, 2008. [Cited on pages 107 and 119]

[SMMCL+11] F. M. Sánchez-Margallo, J. L. Moyano-Cuevas, R. Latorre, J. Maestre, L. Correa, J. B. Pagador, L. F. Sánchez-Peralta, J. A. Sánchez-Margallo, and J. Usón-Gargallo. Anatomical changes due to pneumoperitoneum analyzed by MRI: an experimental study in pigs. *Surgical and radiologic anatomy : SRA*, 33(5):389–396, 2011. [Cited on page 38]

[SNA+15] F Selka, S Nicolau, V Agnus, A Bessaid, J Marescaux, and L Soler. Context-specific selection of algorithms for recursive feature tracking in endoscopic image using a new methodology. *Computerized Medical Imaging and Graphics*, 40:49–61, 2015. [Cited on page 119]

[SNL+13] Caitlin Schneider, Christopher Nguan, Michelle Longpre, Robert Rohling, and Septimiu Salcudean. Motion of the kidney between preoperative and intraoperative positioning. *Biomedical Engineering, IEEE Transactions on*, 60(6):1619–1627, 2013. [Cited on page 38]

[SNP+07] Mickaël Sauvée, Aurélien Noce, Philippe Poignet, Jean Triboulet, and Etienne Dombre. Three-dimensional heart motion estimation using endoscopic monocular vision system: From artificial landmarks to texture analysis. *Biomedical Signal Processing and Control*, 2(3):199–207, 2007. [Cited on pages 118 and 120]

[SNP+14] Luc Soler, Stephane Nicolau, Patrick Pessaux, Didier Mutter, and Jacques Marescaux. Real-time 3D image reconstruction guidance in liver resection surgery. *Hepatobiliary surgery and nutrition*, 3(2):73, 2014. [Cited on pages 35 and 36]

[SPL+10] Sebastian F Schoppmann, Gerhard Prager, Felix B Langer, Franz M Riegler, Barbara Kabon, Edith Fleischmann, and Johannes Zacherl. Open versus minimally invasive esophagectomy: a single-center case controlled study. *Surgical endoscopy*, 24(12):3044–3053, 2010. [Cited on page 19]

[SRB+14] Stefan Suwelack, Sebastian Röhl, Sebastian Bodenstedt, Daniel Reichard, Rüdiger Dillmann, Thiago dos Santos, Lena Maier-Hein, Martin Wagner, Josephine Wünscher, Hannes Kenngott, et al. Physics-based shape matching for intraoperative image guidance. *Medical physics*, 41(11):111901, 2014. [Cited on page 44]

[SSD+13] Douglas P Slakey, Eric Simms, Barbara Drew, Farshid Yazdi, and Brett Roberts. Complications of liver resection: laparoscopic versus open procedures. *JSLS: Journal of the Society of Laparoendoscopic Surgeons*, 17(1):46, 2013. [Cited on page 19]

[SSPY10] Danail Stoyanov, Marco Visentini Scarzanella, Philip Pratt, and Guang-Zhong Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 275–282. Springer, 2010. [Cited on page 42]

[SSSW94] DL Stoker, DJ Spiegelhalter, R Singh, and JM Wellwood. Laparoscopic versus open inguinal hernia repair: randomised prospective trial. *The Lancet*, 343(8908):1243–1245, 1994. [Cited on page 19]

[STJ+11] Ruitian Song, Aaryani Tipirneni, Perry Johnson, Ralf B Loeffler, and Claudia M Hillenbrand. Evaluation of respiratory liver and kidney movements for MRI navigator gating. *Journal of Magnetic Resonance Imaging*, 33(1):143–148, 2011. [Cited on page 38]

[Sup39] Superman. *Action Comics*, (11), April 1939. [Cited on page 31]

[SVA+09] L.-M. Su, B. P. Vagvolgyi, R. Agarwal, C. E. Reiley, R. H. Taylor, and G. D. Hager. Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3D-CT to stereoscopic video registration. *Urology*, 73(4):896–900, 2009. [Cited on pages 38, 44, 45, and 119]

[SW97] Shelley Jane Spaner and Garth Loren Warnock. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *Journal of Laparoendoscopic & Advanced Surgical Techniques*, 7(6):369–373, 1997. [Cited on page 19]

[SWM97] Nelson H Stringer, Jean C Walker, and Peter M Meyer. Comparison of 49 laparoscopic myomectomies with 49 open myomectomies. *The Journal of the American Association of Gynecologic Laparoscopists*, 4(4):457–464, 1997. [Cited on page 19]

[SY09] Danail Stoyanov and Guang-Zhong Yang. Soft tissue deformation tracking for robotic assisted minimally invasive surgery. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 254–257. IEEE, 2009. [Cited on page 119]

[SYD12] Mikael Sodergren, Guang-Zhong Yang, and Lord Ara Darzi. Perception and orientation in minimally invasive surgery. *Archives of Surgery*, 147(3):210–211, 2012. [Cited on page 108]

[SYK+10] Maki Sugimoto, Hideki Yasuda, Keiji Koda, Masato Suzuki, Masato Yamazaki, Tohru Tezuka, Chihiro Kosugi, Ryota Higuchi, Yoshihisa Watayo, Yohsuke Yagawa, et al. Image overlay navigation by markerless surface registration in gastrointestinal, hepatobiliary and pancreatic surgery. *Journal of hepato-biliary-pancreatic sciences*, 17(5):629–636, 2010. [Cited on page 36]

[TAP+12] Youssef S Tanagho, Gerald L Andriole, Alethea G Paradis, Kerry M Madison, Gurdarshan S Sandhu, J Esteban Varela, and Brian M Benway. 2D versus 3D visualization: impact on laparoscopic proficiency using the fundamentals of laparoscopic surgery skill set. *Journal of Laparoendoscopic & Advanced Surgical Techniques*, 22(9):865–870, 2012. [Cited on page 109]

[TAS+98] Josef Tacke, Gerhard Adam, Ralf Speetzen, Kerstin Brucksch, Arno Sticker, Ingo Heschel, Andreas Prescher, Joop J van Vaals, David W Hunter, Günter Rau, et al. Mr-guided interstitial cryotherapy of the liver with a novel, nitrogen-cooled cryoprobe. *Magnetic resonance in medicine*, 39(3):354–360, 1998. [Cited on page 27]

[TCL+13] Gary KL Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C Langbein, Yonghuai Liu, David Marshall, Ralph R Martin, Xian-Fang Sun, and Paul L Rosin. Registration of 3D point clouds and meshes: a survey from rigid to nonrigid. *Visualization and Computer Graphics, IEEE Transactions on*, 19(7):1199–1217, 2013. [Cited on page 44]

[TFA+08] Baki Topal, Steffen Fieuws, Raymond Aerts, H Vandeweyer, and Freddy Penninckx. Laparoscopic versus open liver resection of hepatic neoplasms: comparative analysis of short-term results. *Surgical endoscopy*, 22(10):2208–2213, 2008. [Cited on page 19]

[TFP+09] Junichi Tokuda, Gregory S Fischer, Xenophon Papademetris, Ziv Yaniv, Luis Ibanez, Patrick Cheng, Haiying Liu, Jack Blevins, Jumpei Arata, Alexandra J Golby, et al. Openigtlink: an open network protocol for image-guided therapy environment. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 5(4):423–434, 2009. [Cited on page 121]

[TGS+09] Dogu Teber, Selcuk Guven, Tobias Simpfendörfer, Mathias Baumhauer, Esref Oguz Güven, Faruk Yencilek, Ali Serdar Gözen, and Jens Rassweiler. Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? preliminary in vitro and in vivo results. *European Urology*, 56(2):332–338, 2009. [Cited on pages 32, 38, and 120]

[TNK+09] Takaaki Takeshita, Yoshikazu Nakajima, MK Kim, Shinya Onogi, Mamoru Mitsuishi, and Yoichiro Matsumoto. 3D shape reconstruction endoscope using shape from focus. In *VISAPP (1)*, pages 411–416, 2009. [Cited on page 42]

[Tro13] Jocelyne Troccaz. *Medical robotics*. John Wiley & Sons, 2013. [Cited on page 114]

[TTS+14] Johannes Totz, Stephen Thompson, Danail Stoyanov, Kurinchi Gurusamy, Brian R Davidson, David J Hawkes, and Matthew J Clarkson. Fast semi-dense surface reconstruction from stereoscopic video in laparoscopic surgery. In *Information Processing in Computer-Assisted Interventions*, pages 206–215. Springer, 2014. [Cited on page 42]

[TTS+15a] Sebastian Tauscher, Junichi Tokuda, Günter Schreiber, Thomas Neff, Nobuhiko Hata, and Tobias Ortmaier. Openigtlink interface for state control and visualisation of a robot for image-guided therapy systems. *International journal of computer assisted radiology and surgery*, 10(3):285–292, 2015. [Cited on page 121]

[TTS+15b] Stephen Thompson, Johannes Totz, Yi Song, Stian Johnsen, Danail Stoyanov, Sébastien Ourselin, Kurinchi Gurusamy, Crispin Scheider, Brian Davidson, David Hawkes, et al. Accuracy validation of an image guided laparoscopy system for liver resection. In *SPIE Medical Imaging*, pages 941509–941509. International Society for Optics and Photonics, 2015. [Cited on page 44]

[TTU+13] Norifumi Tsutsumi, Morimasa Tomikawa, Munenori Uemura, Tomohiko Akahoshi, Yoshihiro Nagao, Kozo Konishi, Satoshi Ieiri, Jaesung Hong, Yoshihiko Maehara, and Makoto Hashizume. Image-guided laparoscopic surgery in an

open MRI operating theater. *Surgical endoscopy*, 27(6):2178–2184, 2013. [Cited on page 47]

[TWWW08] Andelle L Teng, Corey Wallach, Peter G Whang, and Jeffrey C Wang. Intraoperative and interventional magnetic resonance imaging: Applications in spine surgery. In *Seminars in Spine Surgery*, volume 20, pages 214–220. Elsevier, 2008. [Cited on page 27]

[UG08] Osamu Ukimura and Inderbir S Gill. Imaging-assisted endoscopic surgery: Cleveland clinic experience. *Journal of Endourology*, 22(4):803–810, 2008. [Cited on page 32]

[UGR04] Kazunori Umeda, Guy Godin, and Marc Rioux. Registration of range and color images using gradient constraints and range intensity images. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 3, pages 12–15. IEEE, 2004. [Cited on page 91]

[UNS+10] Osamu Ukimura, Masahiko Nakamoto, Yoshinobu Sato, Makoto Hashizume, Tsuneharu Miki, Mihir Desai, Monish Aron, and Inderbir S Gill. Augmented reality for image-guided surgery in urology. In *New Technologies in Urology*, pages 215–222. Springer, 2010. [Cited on pages 32, 46, 108, and 116]

[UOK+08] Osamu Ukimura, Koji Okihara, Kazumi Kamoi, Yoshio Naya, Atsushi Ochiai, and Tsuneharu Miki. Intraoperative ultrasonography in an era of minimally invasive urology. *International journal of urology*, 15(8):673–680, 2008. [Cited on page 27]

[URS+06] G Unsgaard, OM Rygh, T Selbekk, TB Müller, F Kolstad, F Lindseth, and TA Nagelhus Hernes. Intra-operative 3D ultrasound in neurosurgery. *Acta neurochirurgica*, 148(3):235–253, 2006. [Cited on page 27]

[VDD+08] Balazs Vagvolgyi, S DiMaio, Anton Deguet, Peter Kazanzides, Rajesh Kumar, Christopher Hasser, and R Taylor. The surgical assistant workstation. In *Proc MICCAI Workshop: Systems and Architectures for Computer Assisted Interventions*, pages 1–8, 2008. [Cited on page 121]

[VE06] Kirby G Vosburgh and R San Jose Estepar. Natural orifice transluminal endoscopic surgery (notes): an opportunity for augmented reality guidance. *Studies in health technology and informatics*, 125:485, 2006. [Cited on page 31]

[VKS06] Sebastian Vogt, Ali Khamene, and Frank Sauer. Reality augmentation for medical procedures: System architecture, single camera marker tracking, and system evaluation. *International Journal of Computer Vision*, 70(2):179–190, 2006. [Cited on page 36]

[VKZHCO11] Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. A survey on shape correspondence. In *Computer Graphics Forum*, volume 30 of *6*, pages 1681–1707. Wiley Online Library, 2011. [Cited on page 44]

[VPB+11] Francesco Volonté, François Pugin, Pascal Bucher, Maki Sugimoto, Osman Ratib, and Philippe Morel. Augmented reality and image overlay navigation with osirix in laparoscopic and robotic surgery: not only a matter of fashion. *Journal of hepato-biliary-pancreatic sciences*, 18(4):506–509, 2011. [Cited on page 36]

[VRH+14] Sinara Vijayan, Ingerid Reinertsen, Erlend Fagertun Hofstad, Anna Rethy, Toril A Nagelhus Hernes, and Thomas Langø. Liver deformation in an animal model due to pneumoperitoneum assessed by a vessel-based deformable registration. *Minimally Invasive Therapy & Allied Technologies*, 23(5):279–286, 2014. [Cited on page 38]

[VRL+10] Cecilie Våpenstad, Anna Rethy, Thomas Langø, Tormod Selbekk, Brynjulf Ystgaard, Toril A Nagelhus Hernes, and Ronald Mårvik. Laparoscopic ultrasound: a survey of its current and future use, requirements, and integration with navigation technology. *Surgical endoscopy*, 24(12):2944–2953, 2010. [Cited on pages 46 and 107]

[VSTH08] Balazs Vagvolgyi, Li-Ming Su, Russell Taylor, and Gregory D Hager. Video to CT registration for image overlay on solid organs. *Proc. Augmented Reality in Medical Imaging and Augmented Reality in Computer-Aided Surgery (AMIARCS)*, pages 78–86, 2008. [Cited on page 44]

[VSZ+04] Th J Vogl, R Straub, S Zangos, MG Mack, and K Eichler. Mr-guided laser-induced thermotherapy (litt) of liver tumours: experimental and clinical data. *International journal of hyperthermia*, 20(7):713–724, 2004. [Cited on page 27]

[WBE+14] Fokko P Wieringa, Henri Bouma, Pieter T Eendebak, Jean-Paul A van Basten, Harrie P Beerlage, Geert AHJ Smits, and Jelte E Bos. Improved depth perception with three-dimensional auxiliary display and computer generated three-dimensional panoramic overviews in robot-assisted laparoscopy. *Journal of Medical Imaging*, 1(1):015001–015001, 2014. [Cited on page 109]

[WBK+08] Wolfgang Wein, Shelby Brunke, Ali Khamene, Matthew R Callstrom, and Nassir Navab. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical image analysis*, 12(5):577–585, 2008. [Cited on page 46]

[WCD+06] Christian Wengert, Philippe C Cattin, John M Duff, Charles Baur, and Gábor Székely. Markerless endoscopic registration and referencing. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 816–823. Springer, 2006. [Cited on page 32]

[WDC+05] Andrew E Welchman, Arne Deubelius, Verena Conrad, Heinrich H Bülthoff, and Zoe Kourtzi. 3D shape perception from combined depth cues in human visual cortex. *Nature neuroscience*, 8(6):820–827, 2005. [Cited on page 108]

[WJN10] C Wu, B Jaramaz, and SG Narasimhan. A full geometric and photometric calibration method for oblique-viewing endoscopes. *Computer Aided Surgery*, 15(1-3):19–31, 2010. [Cited on page 105]

[WKG+08] Michael J Wallace, Michael D Kuo, Craig Glaiberman, Christoph A Binkert, Robert C Orth, Gilles Soulez, Technology Assessment Committee of the Society of Interventional Radiology, et al. Three-dimensional C-arm cone-beam CT: applications in the interventional suite. *Journal of vascular and interventional radiology*, 19(6):799–813, 2008. [Cited on page 26]

[WKS+11] Christian Winne, Martin Khan, Fabian Stopp, Emanuel Jank, and Erwin Keeve. Overlay visualization in endoscopic ent surgery. *International journal of computer assisted radiology and surgery*, 6(3):401–406, 2011. [Cited on page 32]

[WMF+11] Clifford R Weiss, David R Marker, Gregory S Fischer, Gabor Fichtinger, Antonio J Machado, and John A Carrino. Augmented reality visualization using image-overlay for mr-guided interventions: system description, feasibility, and initial evaluation in a spine phantom. *American Journal of Roentgenology*, 196(3):W305–W307, 2011. [Cited on page 36]

[WNL08] Clifford R Weiss, Sherif Gamal Nour, and Jonathan S Lewin. Mr-guided biopsy: A review of current techniques and applications. *Journal of Magnetic Resonance Imaging*, 27(2):311–325, 2008. [Cited on page 27]

[WT12] Dan Wang and Ahmed H Tewfik. Real time 3D visualization of intraoperative organ deformations using structured dictionary. *Medical Imaging, IEEE Transactions on*, 31(4):924–937, 2012. [Cited on page 45]

[WTF04] Andrew D Wiles, David G Thompson, and Donald D Frantz. Accuracy assessment and interpretation for optical tracking systems. In *Medical Imaging 2004*, pages 421–432. International Society for Optics and Photonics, 2004. [Cited on page 117]

[WVW+05] Ivo Wolf, Marcus Vetter, Ingmar Wegner, Thomas Böttger, Marco Nolden, Max Schöbinger, Mark Hastenteufel, Tobias Kunert, and Hans-Peter Meinzer. The medical imaging interaction toolkit. *Medical image analysis*, 9(6):594–604, 2005. [Cited on page 121]

[WWL+14] Jing-Ren Wu, Min-Liang Wang, Kai-Che Liu, Ming-Hsien Hu, and Pei-Yuan Lee. Real-time advanced spinal surgery via visible patient model and augmented reality system. *Computer methods and programs in biomedicine*, 113(3):869–881, 2014. [Cited on pages 4 and 36]

[WZH+10] Xianwang Wang, Qing Zhang, Qiong Han, Ruigang Yang, Melody Carswell, Brent Seales, and Erica Sutton. Endoscopic video texture mapping on prebuilt 3-D anatomical objects without camera tracking. *Medical Imaging, IEEE Transactions on*, 29(6):1213–1223, 2010. [Cited on page 91]

[YCW+10] Ziv Yaniv, Patrick Cheng, Emmanuel Wilson, Teo Popa, David Lindisch, Enrique Campos-Nanez, Hernan Abeledo, Vance Watson, Kevin Cleary, and Filip Banovac. Needle-based interventions with the image-guided surgery toolkit (igstk): from phantoms to clinical trials. *Biomedical Engineering, IEEE Transactions on*, 57(4):922–933, 2010. [Cited on page 121]

[YLS+12] Michael C Yip, David G Lowe, Septimiu E Salcudean, Robert N Rohling, and Christopher Y Nguan. Tissue tracking and registration for image-guided surgery. *Medical Imaging, IEEE Transactions on*, 31(11):2169–2182, 2012. [Cited on page 119]

[YNS+04] Tetsuzo Yamaguchi, Masahiko Nakamoto, Yoshinobu Sato, Kozo Konishi, Makoto Hashizume, Nobuhiko Sugano, Hideki Yoshikawa, and Shinichi Tamura. Development of a camera model and calibration procedure for oblique-viewing endoscopes. *Computer Aided Surgery*, 9(5):203–214, 2004. [Cited on page 105]

[YWLP14] Bo Yang, Wai-Keung Wong, Chao Liu, and Philippe Poignet. 3D soft-tissue tracking using spatial-color joint probability distribution and thin-plate spline model. *Pattern Recognition*, 47(9):2962–2973, 2014. [Cited on page 119]

[ZEP11] Qi Zhang, Roy Eagleson, and Terry M Peters. Volume visualization: a technical overview with a focus on medical applications. *Journal of digital imaging*, 24(4):640–664, 2011. [Cited on page 35]

[Zha00] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000. [Cited on pages 54 and 85]

[ZLH+12] Mariken Zijlmans, Thomas Langø, Erlend Fagertun Hofstad, Christiaan FP Van Swol, and Anna Rethy. Navigated laparoscopy–liver shift and deformation due to pneumoperitoneum in an animal model. *Minimally Invasive Therapy & Allied Technologies*, 21(3):241–248, 2012. [Cited on page 38]

[ZSR+13] Max J Zinser, Hermann F Sailer, Lutz Ritter, Bert Braumann, Marc Maegele, and Joachim E Zöller. A paradigm shift in orthognathic surgery? a comparison of navigation, computer-aided designed/computer-aided manufactured splints, and classic intermaxillary splints to surgical transfer of virtual orthognathic planning. *Journal of Oral and Maxillofacial Surgery*, 71(12):2151–e1, 2013. [Cited on page 32]

Résumé de la thèse de doctorat de **Sylvain Bernhardt**[*†‡]

Développement d'un système de guidage par réalité augmentée en chirurgie par laparoscopie basée sur une acquisition peropératoire 3D

Directeur de thèse : Christophe Doignon[†]
Co-directeur de thèse : Luc Soler[*‡]
Encadrant : Stéphane A. Nicolau[‡]

[*]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

1. Introduction

1.1 Résumé du contexte

La réalité augmentée en chirurgie mini-invasive a rapidement cru ces dernières années [1]. Communément, le site chirurgical est augmenté via la vue endoscopique au moyen d'un modèle 3D extrait d'une acquisition peropératoire [2, 3], par exemple des segmentations de lésions ou vaisseaux. Cependant, à cause du pneumopéritoine et du déplacement du patient, les organes d'intérêt souvent changent radicalement de forme et d'emplacement [4]. Des méthodes ont été développées afin d'estimer et compenser ce changement entre les états pré- et peropératoire.

Par exemple, des solutions de reconstruction de surface (par stéréovision par exemple) peuvent être utilisées afin d'estimer la surface de l'organe d'intérêt depuis l'image endoscopique, puis la recaler par rapport à la même surface extraite de l'image peropératoire [5-9]. Cependant, ce recalage n'est fiable que pour la partie visible de l'organe. D'autre part, l'intégration d'appareils d'imagerie tridimensionnelle au sein des blocs opératoires est de plus en plus répandue. Dans le cas où la qualité de cette image peropératoire est suffisante pour repérer automatiquement ou manuellement des points d'intérêt, l'augmentation peut se faire directement. Dans le cas contraire, cette donnée peut servir d'étape dans le recalage entre image préopératoire et vue endoscopique, par exemple en estimant la déformation entre les états pré- et peropératoire grâce à des modèles biomécaniques [10-12] ou par simple répétition d'acquisitions à faible dose [13, 14].

Cependant, quel que soit le cas, il reste à déterminer la relation entre l'image peropératoire et la vue endoscopique. La méthode la plus populaire consiste à utiliser un système de suivi optique grâce à des marqueurs réfléchissants, comme dans [14]. Néanmoins, cette méthode est couteuse, encombrante et restreint la liberté de mouvement du chirurgien car il ne peut obstruer les marqueurs pour les caméras de détection. De plus, ce type de système requiert de nombreuses étapes de calibration qui peuvent ralentir l'intervention.

Ainsi, dans le cadre de cette thèse, nous nous sommes proposés d'apporter une solution simple, efficace et sans système additionnel au problème de recalage entre image peropératoire et image endoscopique.

1.2 Approche proposée

L'endoscope étant composé principalement de métal (titane), celui-ci est très réfléchissant aux rayons X d'une acquisition CT interventionnelle. Par conséquent, en positionnant l'endoscope de manière à ce qu'il soit orienté vers la zone d'intérêt et introduit dans le champ d'acquisition, il est alors possible de distinguer dans l'image volumique la position de l'endoscope par rapport au reste de l'anatomie (voir Figure 1).

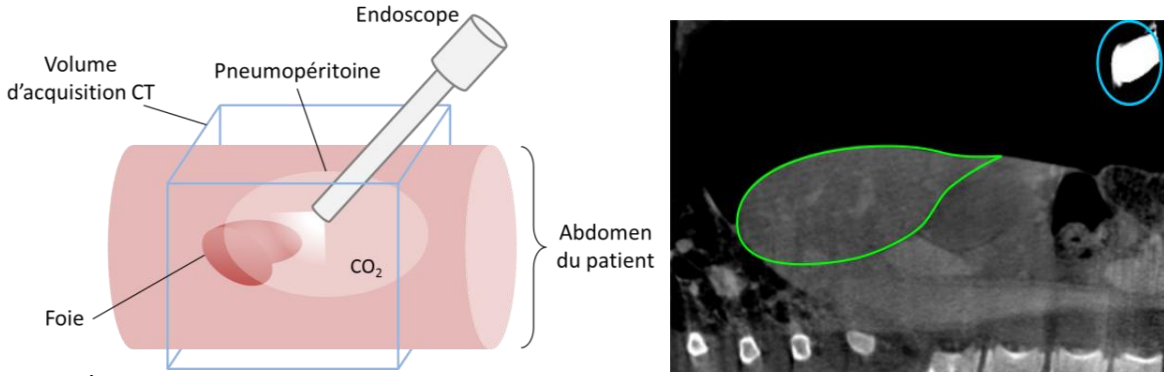


Figure 1: À gauche, illustration du placement de la caméra endoscopique dans l'abdomen et dans le champ d'acquisition du CT interventionnel. À droite, tranche de l'image volumique résultante, où l'on distingue bien le bout de l'endoscope cerclé en bleu, orienté vers le foie entouré de vert.

Dès lors, nous affirmons être en capacité d'extraire suffisamment d'informations de l'image volumique afin d'estimer la position et l'orientation de la caméra par rapport à celle-ci. Puis, en plaçant une caméra virtuelle avec ces mêmes paramètres rapport à un rendu 3D des données, sa vue sur la scène virtuelle devrait être la même que celle de l'endoscope sur la scène réelle. On est alors en mesure d'augmenter la vue endoscopique avec la vue virtuelle.

Cette relation statique peut déjà se montrer utile durant certaines étapes d'intervention chirurgicale telle que la planification d'ablation. De plus, si une navigation endoscopique est requise, nous pouvons toujours maintenir l'augmentation en recourant à un système classique de suivi (SLAM ou système optique) mais notre approche apporterait quand même une bonne initialisation au suivi et supprime donc tout besoin de calibration.

3. Extraction de l'endoscope dans l'image CT

Afin d'augmenter l'image endoscopique, on a besoin de déterminer quelle est l'exacte même vue de la caméra virtuelle sur le rendu volumique des données CT. Il s'agit d'obtenir tous les paramètres intrinsèques et extrinsèques de la caméra endoscopique. Pour les premiers, nous considérons la position du centre optique, la distorsion radiale de la lentille et le champ de vue (lié à la longueur focale). Les seconds sont essentiellement la position et l'orientation de la caméra dans le référentiel du rendu 3D.

3.1 Paramètres intrinsèques

Les paramètres intrinsèques sont obtenus via une calibration classique basée sur la méthode de Zhang [15] en filmant un damier une fois que le zoom de la caméra a été fixé par le chirurgien. Dès lors, nous connaissons la longueur focale, la position du centre optique dans le plan image, ainsi que la distorsion radiale de la lentille, permettant de détordre l'image endoscopique. Lors de nos tests de calibration, nous obtenons typiquement une erreur de reprojection en dessous de 0.7 pixel. Si jamais le zoom devait être changé en cours d'intervention, au lieu de refaire une calibration complète, nous pourrions recourir à une méthode similaire à [16, 17], qui déduit le champ de vue à partir de la taille du contour circulaire de l'image endoscopique.

3.2 Paramètres extrinsèques: méthode basée sur seuil

Pour les paramètres extrinsèques, nous extrayons nos informations de l'image volumique à l'endroit où l'endoscope est inséré. La composition métallique de celui-ci induit que l'intensité des voxels qui le représentent est très grande, bien plus grande que n'importe quelle autre partie du volume. Ainsi, il est trivial de seuiller l'image afin d'en extraire ces voxels seulement (voir Figure 2 à gauche).

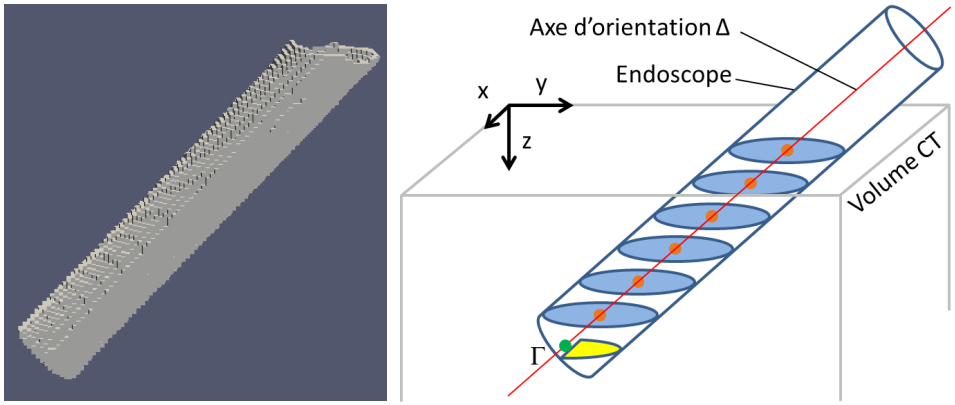


Figure 2: À gauche, exemple d'extraction de l'endoscope dans l'image volumique. À droite, illustration de l'extraction de l'axe de l'endoscope à partir des barycentres (orange) de tranches ellipsoïdales successives (bleues), jusqu'à atteindre le bout en jaune.

Après quoi, connaissant la forme cylindrique de l'endoscope, nous extrayons son axe de révolution en utilisant les barycentres de tranches successives dans les voxels déjà segmentés, le long de l'axe d'entrée dans le volume (l'axe z dans la Figure 2). Une régression linéaire en 3D sur ces points permet une estimation précise de l'orientation de l'endoscope. Le bout de ce dernier (point Γ dans la Figure 2) est également déterminé en considérant le profil d'intensité des voxels le long de l'axe de révolution calculé, en détectant une brusque chute des valeurs en partant de l'extérieur du volume. Il est cependant à noter que la position du bout de l'endoscope résultante ne représente pas la véritable position du centre optique car celle-ci dépend du zoom de la caméra. Néanmoins, de nombreux tests ont permis de montrer que cette position est quand même très proche du bout de l'endoscope (environ 4 mm) et que celle-ci variait relativement peu.

Ainsi, nous sommes capables de déterminer tous les paramètres extrinsèques sauf un: la forme tubulaire de l'endoscope ne permet pas de distinguer la rotation de celui-ci autour de son axe de révolution. Pour déterminer cette valeur, nous montons sur la caméra un accéléromètre qui mesure la position angulaire de celle-ci par rapport au champ de gravité terrestre.

3.3 Paramètres extrinsèques: méthode basée sur modèle

Nous nous sommes rendus compte que le choix de la valeur du seuillage avait une influence sur la détermination de l'orientation du tube et donc sur la précision du recalage final (voir Figure 3). Aussi, une nouvelle méthode, basée sur la mise en correspondance d'un modèle cylindrique tridimensionnel, a été développée.

La surface de l'endoscope produisant les voxels d'intensité la plus élevée (voir Figure 4 à gauche), il est alors possible de placer d'un modèle cylindrique sur cette surface. En échantillonnant le long du modèle, on obtient une valeur moyenne de l'intensité qui le parcourt. Cette valeur peut être considérée comme un score et peut donc être optimisée, grâce à un algorithme comme Powell. Les variables déterminant la direction sont un point C et deux angles ϕ et θ. Le rayon du cylindre est également un paramètre, car il y a un phénomène d'absorbance des rayons X.

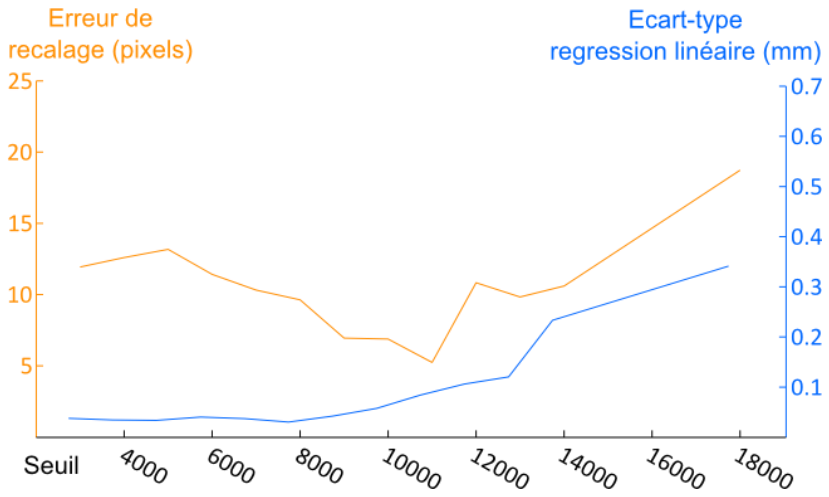


Figure 3: L'erreur de recalage a été mesurée sur un même jeu de données à différentes valeurs de seuil. La variation de l'erreur de recalage n'est pas corrélée à la valeur du seuil. D'autre part, la précision de la régression linéaire diminue lorsque le seuil augmente.

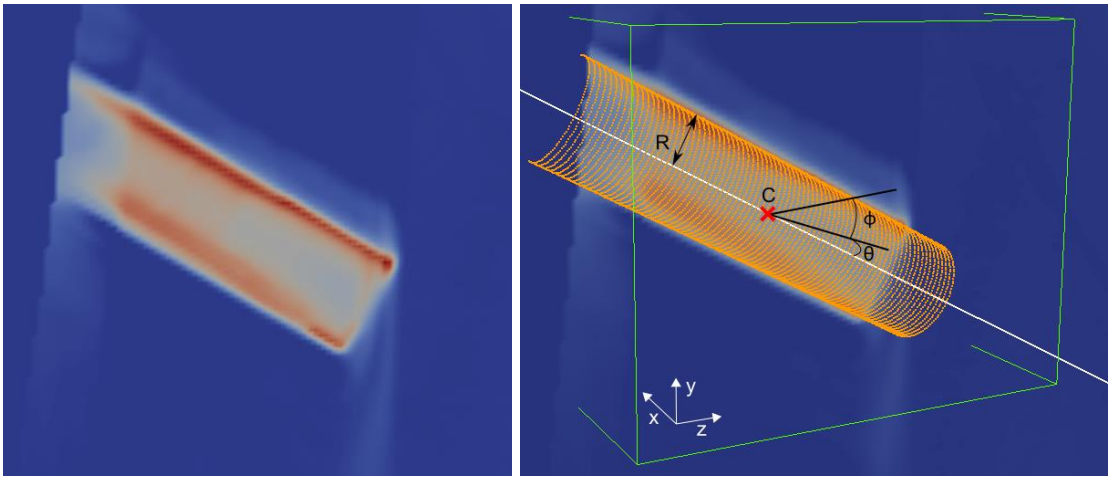


Figure 4: À gauche, image en coupe de l'endoscope issue du scan peropératoire. À droite, la direction de l'endoscope est déterminée par l'optimisation du placement d'un modèle cylindrique paramétré (R, C, ϕ, θ) dans l'image volumique.

4. Compensation du désalignement de l'axe optique

Malgré la précision de l'estimation de l'orientation du tube, nous nous sommes rendus compte qu'il pouvait y avoir potentiellement un désalignement entre ce dernier et l'axe optique. Or, nous positionnons la caméra par rapport au tube, donc un désalignement entre ces deux axes produit une erreur dans le recalage (voir Figure 5).

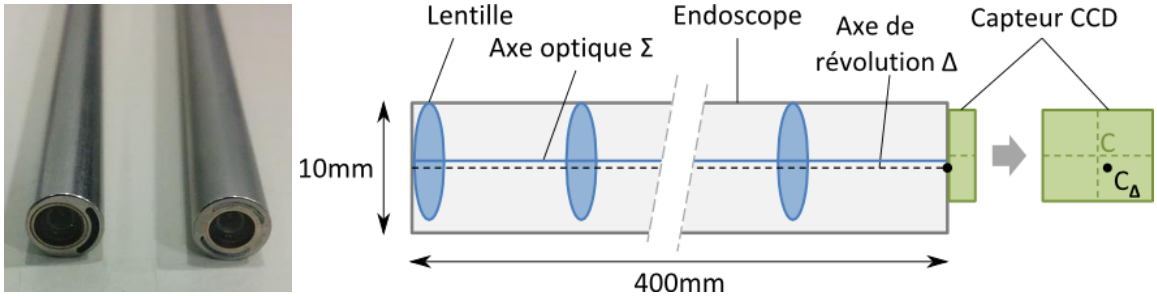


Figure 5: À gauche, deux laparoscopes dont l'un d'eux est décentré. À droite, illustration d'un désalignement axe du tube/axe optique qui résulte en une erreur de recalage dans le plan image.

Afin de compenser ce potentiel désalignement, nous proposons d'utiliser un tube carré stérile. En le faisant reposer sur le bout de l'endoscope et en le faisant tourner autour de celui-ci, on peut alors détecter les côtés du carré dans l'image et en déduire la projection de l'axe de révolution C_Δ dans celle-ci (voir Figure 6). En alignant l'image virtuelle tel que C_Δ soit sur C dans le plan image, on peut alors compenser le désalignement.

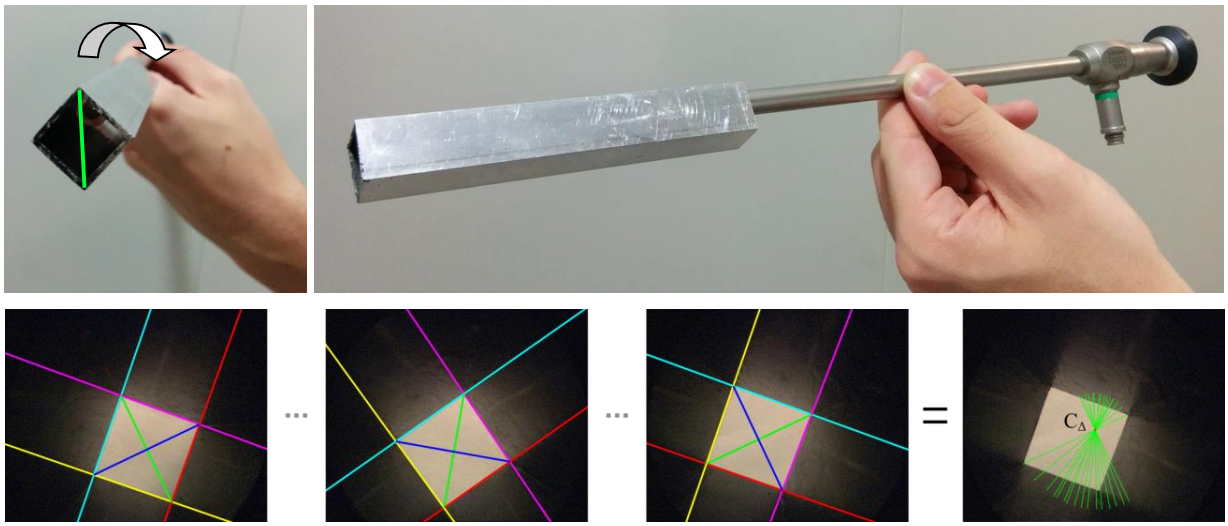


Figure 6: En haut, exemple de tube carré reposant sur un laparoscope. L'axe vert dénote la diagonale du carré qui croise l'axe de révolution du tube. En bas, détection, durant la rotation, des diagonales du carré dans l'image et déduction de la position de C_Δ.

5. Résultats quantitatifs sur mire radio-opaque

Afin de tester de manière quantitative notre méthode, nous avons imprimé en 3D une mire spéciale visible par les deux modalités d'imagerie (voir Figure 7). Une fois la caméra réglée et calibrée, le damier est acquis par l'endoscope et le scanneur en même temps suivant le protocole précédemment décrit. Ensuite, nous extrayons de façon automatique et indépendante les coins du damier à la fois dans l'image volumique (avec VTK) et dans l'image endoscopique (avec OpenCV) (voir Figure 8). Ainsi, en réalisant notre méthode de réalité augmentée, les positions des coins sont également recalées et on peut calculer l'erreur moyenne en pixels entre les deux nuages de points.

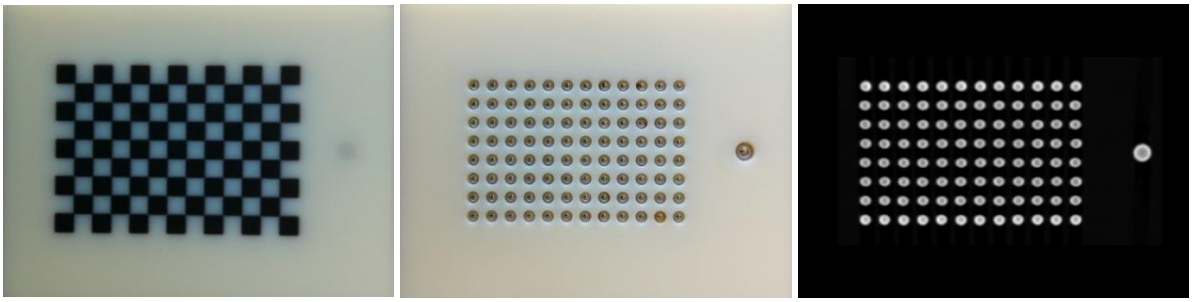


Figure 7: De gauche à droite : le dessus de la mire avec le damier visible par le laparoscope, le dessous de la mire avec des billes insérées exactement sous les coins du damier et l'image aux rayons X de ces billes.

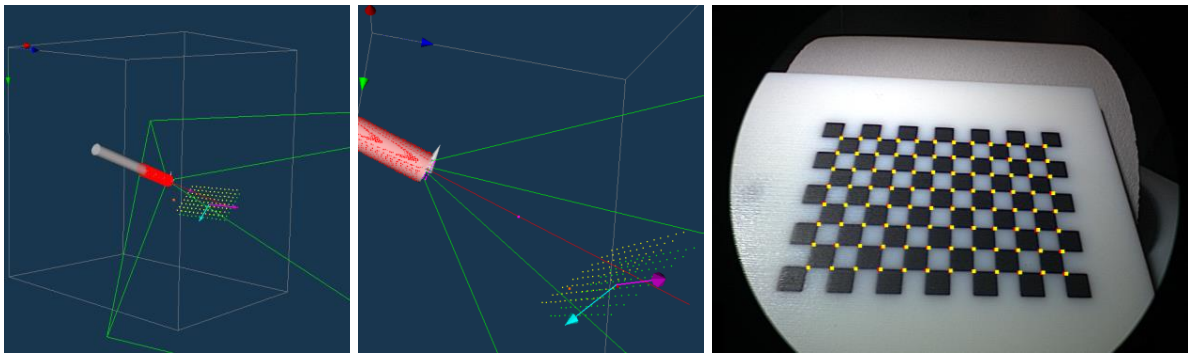


Figure 8: La caméra virtuelle placée suivant notre méthode (cylindre gris et champ en vert) et les le centre des billes extraites de l'image volumique (en vert). Les coins du damier en sont déduits (en jaune). À droite, les coins sont extraits dans l'image endoscopique (en rouge) et cette dernière est augmentée avec les coins du rendu volumique tels que vus par la caméra virtuelle.

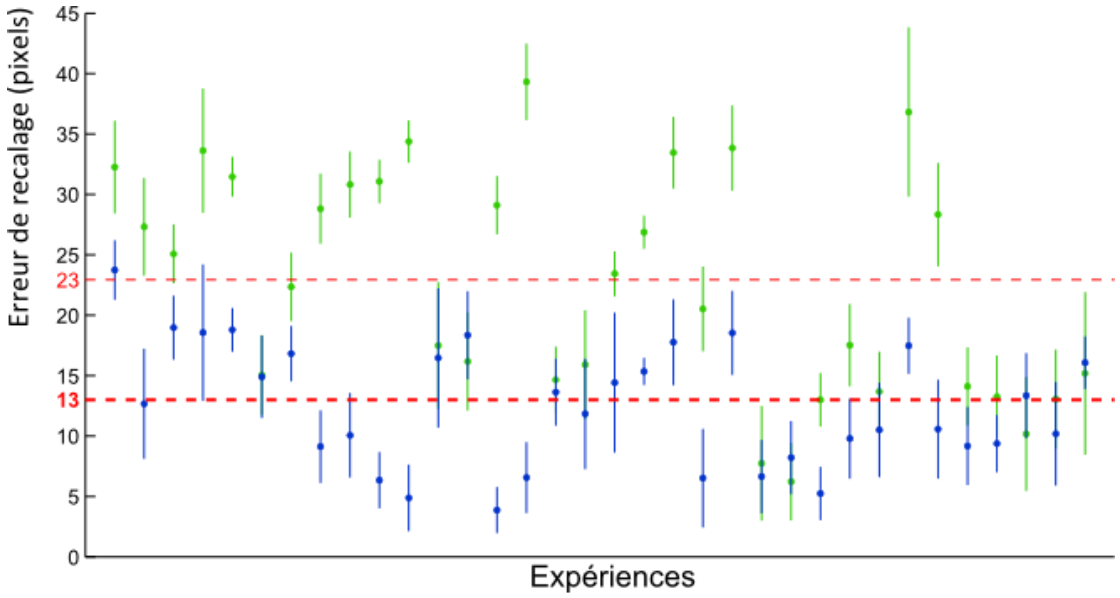


Figure 9: La moyenne et l'écart-type de l'erreur de recalage sur le damier radio-opaque pour 34 expériences, en vert sans compensation du désaxage et en bleu avec compensation.

Une série de 34 tests avec des positions et des réglages différents a été réalisée (Figure 9). En moyenne, notre méthode de recalage montre une erreur de 23 pixels sans compensation du désalignement axe optique/axe de révolution. Avec compensation, cette moyenne descend à 13 pixels, ce qui correspond à moins d'un millimètre dans la scène réelle à une distance nominale de 7 cm. Une telle précision est en deçà des marges chirurgicales de beaucoup d'interventions.

6. Tests précliniques et cliniques

Nous avons appliqué notre méthode de réalité augmentée dans des cas chirurgicaux réalistes et variés. Sur cochon, nous avons notamment réalisé une localisation d'une lésion hépatique (Figure 10). On peut constater que l'augmentation de la lésion est placée au bon endroit par rapport à la vue réelle.

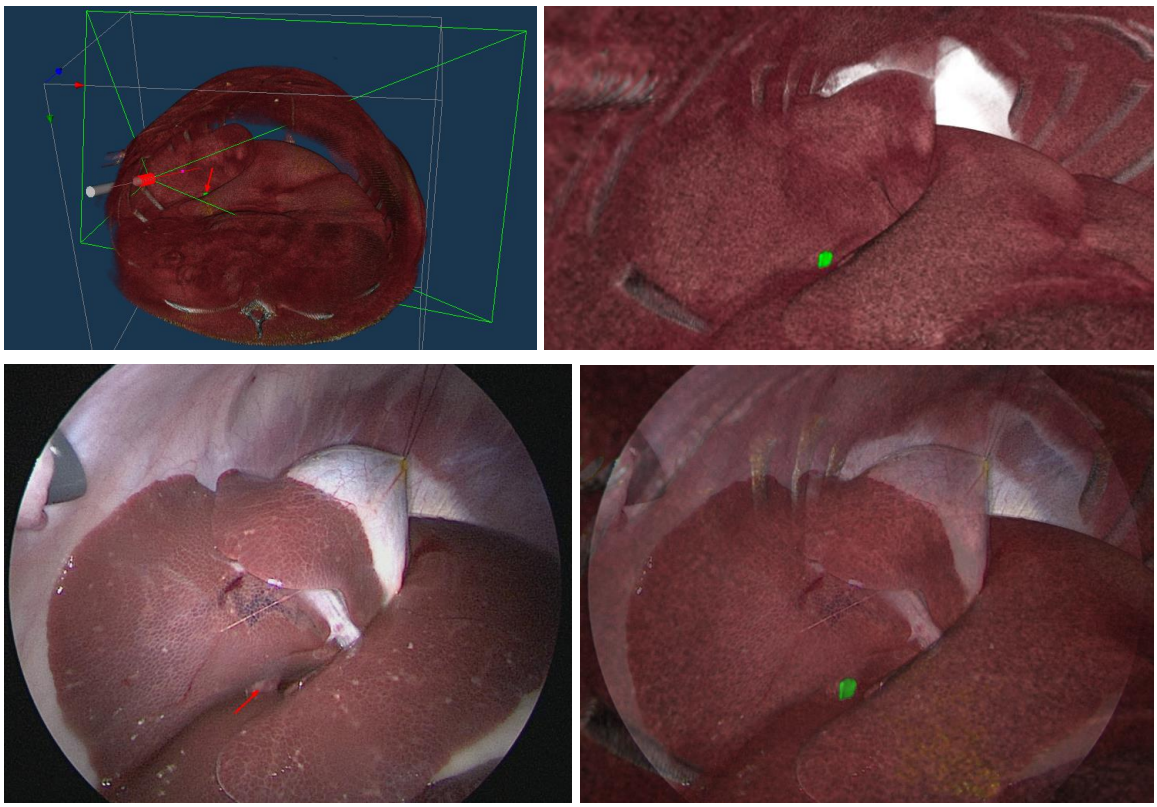


Figure 10: En haut à gauche, le rendu volumique des données issues du CT interventionnel est affiché à gauche avec la caméra virtuelle placée suivant notre méthode. Une lésion hépatique segmentée manuellement (vert) est pointée par la flèche rouge. En haut à droite, l'image virtuelle correspondante. En bas à gauche, l'image endoscopique montrant le foie. En bas à droite, la même image augmentée grâce aux informations issues de la scène virtuelle.

Toujours sur cochon, un produit de contraste a été injecté afin de rendre opaques la vessie et les uretères aux rayons X du scanneur peropératoire. Chez le porcin, les uretères sont également visibles à la surface dans la cavité via endoscopie mais chez l'humain ceux-ci sont couverts par des tissus adipeux. Nous nous sommes donc proposés de montrer la faisabilité d'une augmentation interventionnelle des uretères (Figure 11).

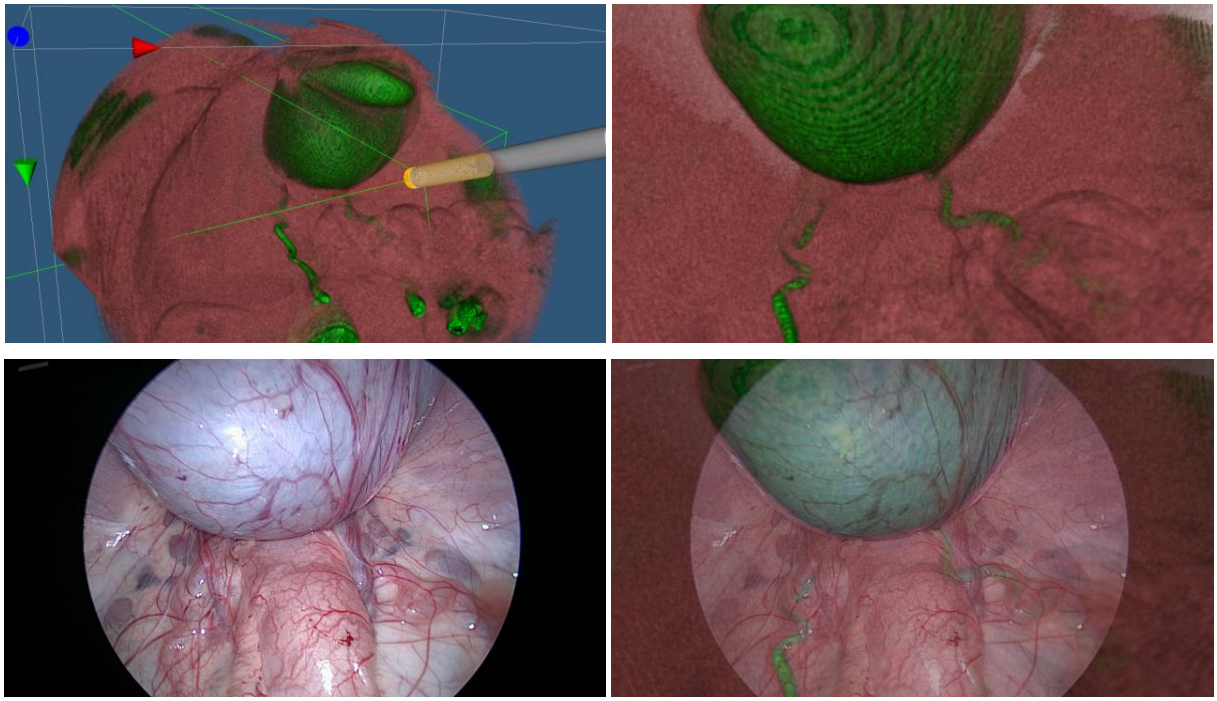
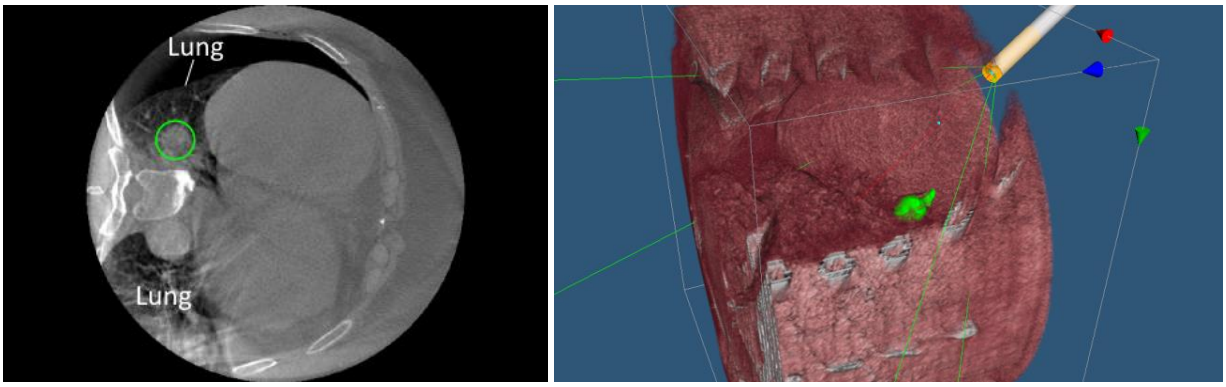


Figure 11: En haut à gauche, le rendu volumique des données issues du CT interventionnel est affiché à gauche avec la caméra virtuelle placée suivant notre méthode. Un produit de contraste a été injecté et, après réglage de la fonction de transfert de couleur, fait apparaitre la vessie et les uretères en vert. En haut à droite, l'image virtuelle correspondante. En bas à gauche, l'image endoscopique montrant la vessie et les uretères. En bas à droite, la même image augmentée grâce aux informations issues de la scène virtuelle.

En partenariat avec le CHU de Rennes, nous avons également pu tester notre méthode pour la localisation d'une tumeur pulmonaire chez un patient (Figure 12). Celle-ci a été repérée par le chirurgien avant l'intervention, mais l'affaissement nécessaire du poumon pour en faciliter l'accès a fait bouger de façon significative la cible. Nous avons alors demandé au chirurgien d'entourer la position de la tumeur dans la scène laparoscopique en se basant sur sa connaissance de l'image préopératoire. Préopérativement, la tumeur a été segmentée dans les données 3D et notre méthode a permis de révéler un désaccord entre notre position de la lésion et celle supposée par le praticien. Un examen de la vidéo de l'intervention a démontré que notre augmentation était correcte. Ainsi, notre approche a le potentiel d'empêcher le chirurgien de se tromper dans la localisation de cibles lors de fortes déformations entre états préopératoire et peropératoire.



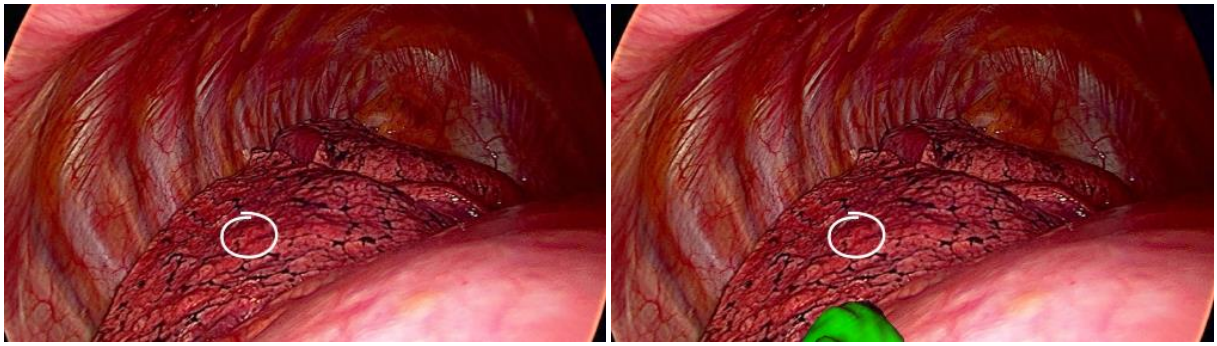


Figure 12: En haut à gauche, une coupe dans le scan CT interventionnel montre la tumeur (verte). En haut à droite, le rendu volumique de ces données 3D est complété par une segmentation de la tumeur et le positionnement de la caméra virtuelle selon notre méthode. En bas à gauche, l'image endoscopique et la position de la tumeur supposée par le chirurgien. En bas à droite, la même image augmentée par notre méthode montre l'erreur d'interprétation de la scène.

7. Raffinement du recalage basé sur l'ombre

7.1 Une nouvelle approche

En partenariat avec le Professeur Adrien Bartoli de l'Université d'Auvergne, nous avons développé une nouvelle méthode pour recaler la surface d'une scène intra-abdominale, extraite automatiquement de l'image volumique, et son image laparoscopique. Classiquement, les méthodes d'analyse de l'ombre supposent un albédo constant pour toute la scène, ce qui est souvent erroné, en particulier pour des scènes de laparoscopie. De plus, ce type de technique vise à reconstruire la topologie de la scène à partir des effets d'ombre. Nous nous proposons de faire l'inverse et d'estimer l'ombre à partir de la surface via une approche par morceau qui permet de maintenir un albédo constant. Par comparaison entre l'ombre estimée et l'ombre réelle, nous formulons une mesure de dissimilarité entre la surface et son image. Par minimisation de cette mesure, nous pouvons alors recaler l'image endoscopique rapport à la scène.

Pour l'instant, nous résolvons ce problème dans le plan image, donc pour trois degrés de liberté: la rotation (en X et Y) et la translation. Néanmoins, cela nous permet de nous affranchir de l'accéléromètre pour la détermination de la rotation axiale de l'endoscope. Egalement, cela nous permet de raffiner le recalage en translation issu de la calibration avec le tube carré.

7.2 Méthodologie

Afin de simuler l'éclairage de l'endoscope, nous adoptons un modèle de lumière ponctuelle où la source et le centre optique sont supposés identiques (Figure 13).

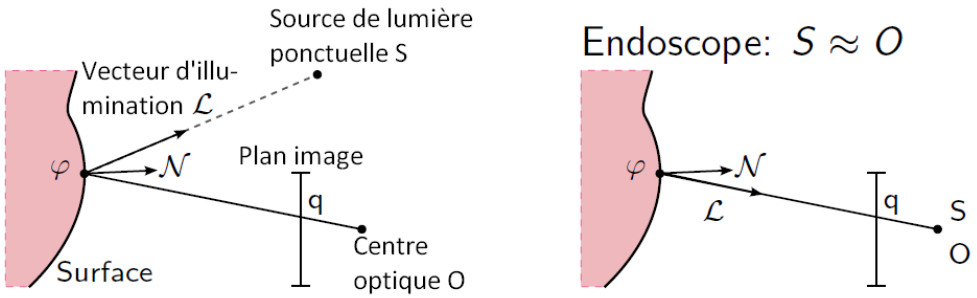


Figure 12: Modèle d'éclairage utilisé pour notre méthode.

Suivant ce modèle, on se propose d'estimer la luminance depuis la surface tel que :

$$I(q) = cJ \quad \text{avec} \quad c = ab\|\mathcal{L}\| \quad \text{et} \quad J = -\frac{\varphi(q)^T \mathcal{N}(q)}{\|\varphi(q)\|^2}$$

En assumant une réponse de caméra linéaire (I = a × lumière reçue) et un albédo constant localement. Afin de permettre cette approche par morceaux, l'image est divisée en zones de couleur homogène (Figure 13) grâce à l'algorithme dit *watershed*.

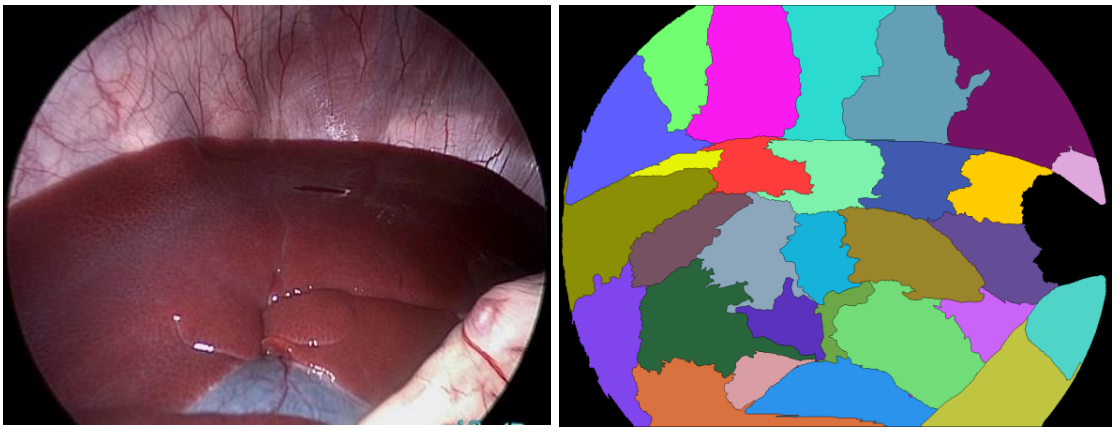


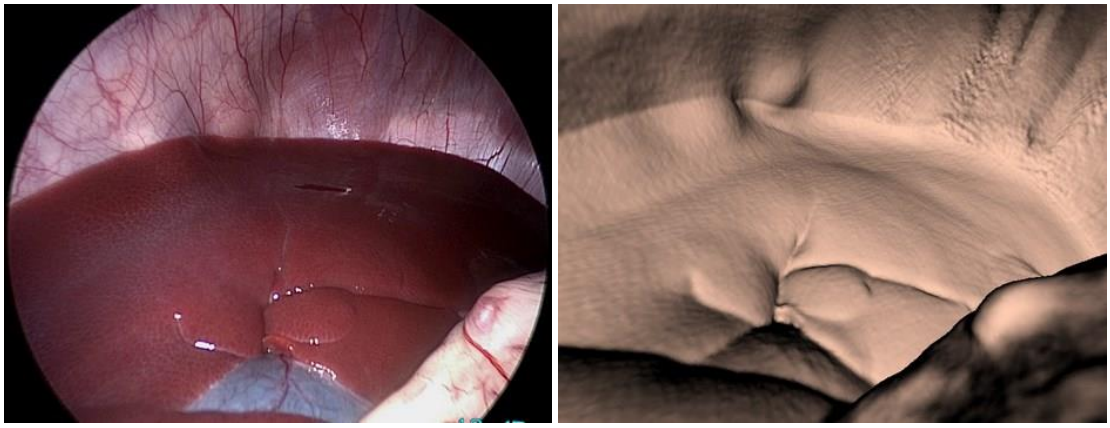
Figure 13: Division de l'image endoscopique en zones homogènes en se basant sur la couleur.

Pour une position ω dans le plan image donnée, on calcule c dans chaque morceau p par régression linéaire de I/J. Les résidus de cette régression f(ω) constituent une mesure de la dissimilarité locale entre l'image et la surface. En sommant ces dissimilarités pour tous les morceaux, on obtient une fonction de coût qui peut être optimisée.

$$f_p(\omega) = \min_{c\in\mathbb{R}} \sum_{q\in p} \left\| I(q) + c\frac{\varphi(q)^T \mathcal{N}(q)}{\|\varphi(q)\|^2} \right\|^2 \quad \text{et} \quad F(\omega) = \sum_p f_p(\omega)$$

7.3 Résultats

Des évaluations qualitatives sur des données *in vivo* donnent des résultats pertinents (Figure 14).



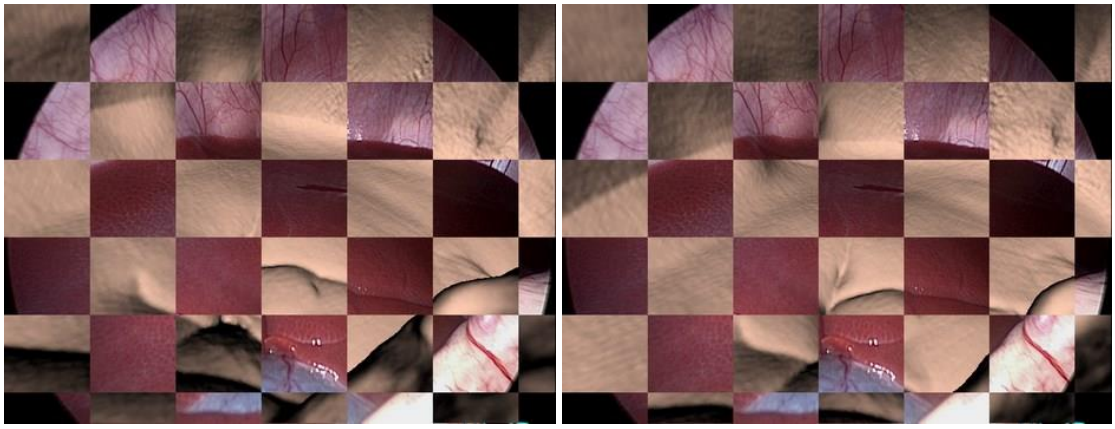


Figure 14: En haut à gauche, image endoscopique. En haut à droite, surface de la même scène, vue par la caméra virtuelle. En bas à gauche, mosaïque des deux avant recalage. En bas à droite, mosaïque après recalage. On peut observer la correspondance des différents contours.

8. Conclusion

Nous avons présenté un nouveau paradigme à la réalité augmentée en chirurgie laparoscopique. Se reposant uniquement sur l'équipement standard d'une salle opératoire hybride, notre approche peut fournir la relation statique entre l'endoscope et un scan peropératoire 3D. De nombreuses expériences sur un motif radio-opaque montrent quantitativement que nos augmentations sont exactes à moins d'un millimètre près. Des tests sur des données in vivo consolident la démonstration du potentiel clinique de notre approche dans plusieurs cas chirurgicaux réalistes.

Publications

Article de Journal

Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. "*Automatic Localization of Endoscope in Intraoperative CT Image: A Simple Approach to Augmented Reality Guidance in Laparoscopic Surgery.*" In Medical Image Analysis, pp. 130-143, 2016. – Disponible en ligne

Papiers de Conférence

Sylvain Bernhardt, Stéphane A. Nicolau, Adrien Bartoli, Vincent Agnus, Luc Soler, and Christophe Doignon. "*Using Shading to Register an Intraoperative CT Scan to a Laparoscopic Image.*" In Computer-Assisted and Robotic Endoscopy (CARE), 2015. MICCAI 2015 satellite event – Présentation orale, Prix du Meilleur Papier (2e place)

Stéphane A. Nicolau, Sylvain Bernhardt, Vincent Agnus, Christophe Doignon, Luc Soler, and Jacques Marescaux. "*Validation Methodology For Augmented Reality In Laparoscopic Surgery.*" In Computer Assisted Radiology and Surgery (CARS), 29th International Congress on, 2015. – Présentation orale

Sylvain Bernhardt, Stéphane A. Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. "*Automatic detection of endoscope in intraoperative CT image: Application to AR guidance in laparoscopic surgery.*" In Biomedical Imaging (ISBI), 11th International Symposium on, pp. 563-567. IEEE, 2014. – Présentation orale

Bibliographie

[1] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced Medical Displays : A Literature Review of Augmented Reality," *Journal of Display Technology*, vol. 4, no. 4, pp. 451–467, 2008.

[2] M. Baumhauer, M. Feuerstein, H.-P. Meinzer, and J. Rassweiler, "Navigation in endoscopic soft tissue surgery: perspectives and limitations," *Journal of endourology / Endourological Society*, vol. 22, no. 4, pp. 751–766, 2008.

[3] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, "Augmented reality in laparoscopic surgical oncology," *Surgical oncology*, vol. 20, no. 3, pp. 189–201, 2011.

[4] F. M. S´anchez-Margallo et al., "Anatomical changes due to pneumoperitoneum analyzed by MRI: an experimental study in pigs," *Surgical and radiologic anatomy : SRA*, vol. 33, no. 5, pp. 389–396, 2011.

[5] B. Lo, A. J. Chung, D. Stoyanov, G. Mylonas, and G.-Z. Yang, "Real-time intra-operative 3d tissue deformation recovery," in *Biomedical Imaging: From Nano to Macro, ISBI*, 2008, pp. 1387–1390.

[6] P. Mountney and G.-Z. Yang, "Motion compensated slam for image guided surgery," in *Proc. of MICCAI*, 2010, pp. 496–504.

[7] M. Hu et al., "Non-rigid reconstruction of the beating heart surface for minimally invasive cardiac surgery," in *Proc. of MICCAI*, 2009, pp. 34–42.

[8] M. Figl et al., "Image guidance for robotic minimally invasive coronary artery bypass," *Computerized Medical Imaging and Graphics*, vol. 34, no. 1, pp. 61–68, 2010.

[9] L.-M. Su, B. P. Vagvolgyi, R. Agarwal, C. E. Reiley, R. H. Taylor, and G. D. Hager, "Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3d-ct to stereoscopic video registration," *Urology*, vol. 73, no. 4, pp. 896–900, 2009.

[10] T. Kitasaka, K. Mori, Y. Hayashi, Y. Suenaga, M. Hashizume, and J. Toriwaki, "Virtual pneumoperitoneum for generating virtual laparoscopic views based on volumetric deformation," in *Proc. of MICCAI*, 2004, pp. 559–567.

[11] J. Bano et al., "Simulation of pneumoperitoneum for laparoscopic surgery planning," in *Proc. of MICCAI*, 2012, pp. 91–98.

[12] O. Oktay et al., "Biomechanically driven registration of pre- to intra- operative 3d images for laparoscopic surgery," in *Proc. of MICCAI*, 2013, pp. 1–9.

[13] R. Shekhar et al., "Live augmented reality: a new visualization method for laparoscopic surgery using continuous volumetric computed tomography," *Surgical endoscopy*, vol. 24, no. 8, pp. 1976–1985, 2010.

[14] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, "Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection," *IEEE Transactions on Medical Imaging*, vol. 27, no. 3, pp. 355–369, 2008.

[15] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Transac. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[16] T.-Y. Lee, T.-S. Chang, C.-H. Wei, S.-H. Lai, K.-C. Liu, and H.-S. Wu, "Automatic Distortion Correction of Endoscopic Images Captured With Wide-Angle Zoom Lens," *IEEE Transac. on Biomedical Engineering*, vol. 60, no. 9, pp. 2603–2613, 2013.

[17] C. Doignon et M. de Mathelin, "A Degenerate Conic-Based Method for a Direct Fitting and 3-D Pose of Cylinders with a Single Perspective View", *IEEE Int'l Conf. on Robotics and Automation*, ICRA'07, Roma, Italy.

Sylvain BERNHARDT

Automatic Localization of Endoscope in Intraoperative CT Image: A Simple Approach to Augmented Reality Guidance in Laparoscopic Surgery

Résumé:

Au cours des dernières décennies, la chirurgie mini invasive a progressivement gagné en popularité face à la chirurgie ouverte, grâce à de meilleurs bénéfices cliniques. Cependant, ce type d'intervention introduit une perte de vision directe sur la scène pour le chirurgien. L'introduction de la réalité augmentée en chirurgie mini invasive semble être une solution viable afin de remédier à ce problème et a donc été activement considérée par la recherche. Néanmoins, augmenter correctement une scène laparoscopique reste difficile à cause de la non-rigidité des tissus et organes abdominaux. En conséquence, la littérature ne fournit pas d'approche satisfaisante à la réalité augmentée en laparoscopie, car de telles méthodes manquent de précision ou requièrent un équipement supplémentaire, contraignant et onéreux.

Dans ce contexte, nous présentons un nouveau paradigme à la réalité augmentée en chirurgie laparoscopique. Se reposant uniquement sur l'équipement standard d'une salle opératoire hybride, notre approche peut fournir la relation statique entre l'endoscope et un scan intraopératoire 3D. De nombreuses expériences sur un motif radio-opaque montrent quantitativement que nos augmentations sont exactes à moins d'un millimètre près. Des tests sur des données in vivo consolident la démonstration du potentiel clinique de notre approche dans plusieurs cas chirurgicaux réalistes.

Mots-clés : intervention assistée par ordinateur, réalité augmentée, chirurgie laparoscopique, recalage

Abstract:

Over the past decades, minimally invasive surgery has progressively become more popular than open surgery thanks to greater clinical benefits. However, this kind of intervention introduced a loss of direct vision upon the scene for the surgeon. Introducing augmented reality to minimally invasive surgery appears to be a viable solution to alleviate this drawback and has thus been an attractive topic for the research community. Yet, correctly augmenting a laparoscopic scene remains challenging, due to the non-rigidity of abdominal tissues and organs. Therefore, the literature does not report a satisfactory approach to laparoscopic augmented reality, as such methods lack accuracy or require expensive and impractical additional equipment.

In light of this, we present a novel paradigm to augmented reality in abdominal minimally invasive surgery. Based only on standard hybrid operating room equipment, our approach can provide the static relationship between the endoscope and an intraoperative 3D scan. Extensive experiments on a radio-opaque pattern quantitatively show that the accuracy of our augmentations is less than one millimeter. Tests on in vivo data further demonstrates the clinical potential of our approach in several realistic surgical cases.

Keywords: computer-assisted intervention, augmented reality, laparoscopic surgery, registration