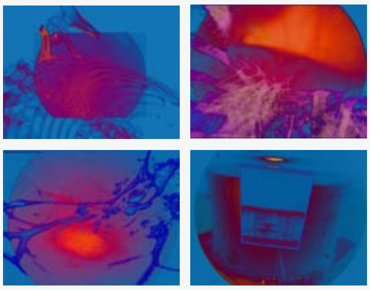
Computer Aided Medical Procedures
(CAMP)
Prof. Dr. Nassir Navab



Dissertation

Augmented Reality in Laparoscopic Surgery

New Concepts for Intraoperative Multimodal Imaging

Marco Feuerstein

Fakultät für Informatik
Technische Universität München

Technische Universität München, Fakultät für Informatik
Chair for Computer-Aided Medical Procedures & Augmented Reality

Augmented Reality in Laparoscopic Surgery
New Concepts for Intraoperative Multimodal Imaging

Marco Feuerstein

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. G. J. Klinker, PhD
Prüfer der Dissertation:		
	1.	Univ.-Prof. N. Navab, PhD
	2.	Assoc. Prof. K. Mori, PhD, Nagoya University, Japan

Die Dissertation wurde am 21.06.2007 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 01.10.2007 angenommen.

Abstract

A trend in abdominal surgery is the transition from open procedures to minimally invasive laparoscopic interventions, where visual feedback to surgical staff is only available through the laparoscope camera and direct palpation of organs is impossible. To successfully perform such sophisticated interventions, the provision of additional intraoperative feedback can be of great help to the surgical staff, especially in complex cases.

This work introduces several new concepts for the application of augmented reality techniques to laparoscopic surgery. One main idea is to utilize multiple intraoperative imaging devices for the acquisition of up-to-date patient data. Optical and electromagnetic tracking systems are applied to determine the position and orientation of both rigid (mobile C-arm, laparoscope) and flexible (laparoscopic ultrasound) imaging devices. The acquired patient data is intrinsically registered to the tracked laparoscope in one common coordinate system, so it can be directly superimposed on the images of the laparoscope camera in real time without intraoperative registration steps. This intuitive superimposition can visually assist and direct the surgeon, as hidden anatomy such as vessels or tumors below the surface of an organ are revealed.

The presented visualization aid can be used during critical phases in the surgical workflow such as port placement and intraoperative resection planning. Whereas superimposition for resection planning is based on intraoperative, implicitly registered imaging data, superimposition for port placement requires an interactive registration of preoperative imaging data to the patient. This interactive process is mostly automated by a newly introduced registration technique that results in a port placement procedure soundly integrated into the current surgical workflow. For resection planning and guidance, where navigated laparoscopic ultrasound can be used to acquire updated images of patient anatomy, a hybrid tracking approach including a method capable of estimating the reliability of electromagnetic tracking data is presented, which is able to automatically notify the surgical staff on possible tracking inaccuracies.

The dissertation bases its validation on many experiments, including animal experiments, performed in close partnership with several surgeons.

Keywords:
Medical Augmented Reality, Image-guided Surgery, Laparoscopic Surgery, Computer Aided Surgery

Zusammenfassung

Ein gegenwärtiger Trend in der abdominellen Chirurgie ist der Übergang von offenen zu minimalinvasiven laparoskopischen Eingriffen. Dabei erhält das chirurgische Team visuelles Feedback nur über die Laparoskop-Kamera und kann Organe nicht mehr direkt abtasten. Für die erfolgreiche Durchführung von laparoskopischen Eingriffen ist das Bereitstellen von zusätzlichem intraoperativen Feedback für das chirurgische Team von großer Hilfe, insbesondere bei komplizierten Befunden.

Diese Arbeit stellt diverse neue Konzepte für die Anwendung von Augmented Reality-Techniken in der laparoskopischen Chirurgie vor. Eine Hauptidee dabei ist die Verwendung von mehreren Geräten für die intraoperative Bildgebung, mit denen aktuelle Patientendaten gewonnen werden können. Die Position und Ausrichtung aller starren (C-Bogen, Laparoskop) sowie flexiblen (laparoskopischer Ultraschall) Bildgebungsgeräte wird von optischen und elektromagnetischen Tracking-Systemen verfolgt, was eine intrinsische Registrierung aller Geräte im selben Koordinatensystem ermöglicht. Dadurch können die Live-Bilder der Laparoskop-Kamera ohne zusätzliche intraoperative Registrierungsschritte sowie in Echtzeit mit den Patientendaten überlagert werden. Diese intuitive Überlagerung kann den Chirurgen visuell unterstützen und leiten, da unter der Organoberfläche verborgene anatomische Strukturen wie zum Beispiel Gefäße oder Tumore sichtbar gemacht werden.

Die vorgestellte Visualisierungshilfe kann während kritischer Phasen des chirurgischen Eingriffes verwendet werden, wie zum Beispiel zur Port-Platzierung und zur intraoperativen Resektionsplanung. Während die Überlagerung zur Resektionsplanung auf intraoperativen, implizit registrierten Bilddaten basiert, benötigt sie für die Port-Platzierung eine interaktive Registrierung der präoperativen Bilddaten zum Patienten. Diese interaktive Prozedur wird weitgehend automatisiert durch eine neu eingeführte Registrierungstechnik, aus der ein Port-Platzierungsverfahren hervorgeht, welches sich reibungslos in den gegenwärtigen chirurgischen Arbeitsablauf integrieren lässt. Für die Resektionsplanung und -führung, wo navigierter laparoskopischer Ultraschall zur Erfassung aktueller Bilder der Patientenanatomie verwendet werden kann, wird zudem ein hybrider Tracking-Ansatz einschließlich einer Methode zum Abschätzen der Zuverlässigkeit elektromagnetischer Tracking-Daten vorgestellt, welche das chirurgische Team automatisch über mögliche Tracking-Ungenauigkeiten informieren kann.

Alle vorgestellten Konzepte wurden in zahlreichen Experimenten sowie Tierversuchen validiert, welche in enger Zusammenarbeit mit mehreren Chirurgen durchgeführt wurden.

Schlagwörter:
Medizinische Augmented Reality, Bildgestützte Navigation, Laparoskopische Chirurgie, Computerunterstützte Chirurgie

ACKNOWLEDGMENTS

First of all, I would like to thank my PhD adviser Nassir Navab a lot, not only for making my dissertation possible, but also for giving me helpful advice, guidance, and support during the last years. I also owe many thanks to Gudrun Klinker, who arranged the contact to Nassir and agreed to chair my thesis commission, as well as Kensaku Mori, who was willing to be my second thesis reviewer and was not hesitating to accept me as a postdoc in Nagoya. I would also like to thank Martina Hilla for answering my questions on various applications, accounting, and business trips, and for organizing my defense.

I also owe a great deal to Jörg Traub for proofreading my thesis and providing me with a lot of helpful and corrective advise. I would like to thank him as well as my colleagues and students Tobias Sielhorst, Oliver Kutter, Stefan Wiesner, Jakob Vogel, Tobias Reichl, Christoph Bichlmeier, Thomas Wendler, Claudio Alcérreca, Konstantinos Filippatos, Mohammad Rustaee, and Julian Much, who all assisted me greatly throughout the whole thesis and contributed to many joint publications. Furthermore, many thanks go to the rest of the CAMP team, namely Martin Groher, Ben Glocker, Andreas Keil, Martin Bauer, Martin Horn, Moritz Blume, Pierre Georgel, Hauke Heibel, Wolfgang Wein, Darko Zikic, Tobias Lasser, Ruxandra Micu, Nicolas Padoy, and many more, for making it possible to work in the most collaborative and also fun environment.

I would also like to thank my medical partners at Klinikum Innenstadt and Klinikum rechts der Isar, Sandro M. Heining, Thomas Mussack, Armin Schneider, and Hubertus Feußner, and at Herzzentrum München, Eva U. Braun, Stephen M. Wildhirt, and Robert Bauernschmitt, for a very fruitful and active collaboration. Their medical advice and help in designing and organizing all the experiments was invaluable.

This work might not have been possible had it not been for the efforts of Inga Drosse, Roland Ladurner, Peter Scheuber, Philipe Khalil, Felix Hohenbleicher, Thomas Meindl, and Markus Körner, who greatly supported me in carrying out all experiments. I also appreciate the loan of miscellaneous equipment by Siemens Medical, KARL STORZ GmbH & Co. KG, and ART GmbH, and especially the support of Rainer Graumann, Christian Schmidgunst, Etienne Kneschke, Marc Schneberger, and Oliver Wenisch.

Finally, I owe a lot of thanks to my girlfriend Ching-Hsin Weng, my parents Marianne and Franz Feuerstein, my grandmother Centa Niedermayer, and my brother Ralph Feuerstein for their endless patience and aid during my studies.

CONTENTS

CHAPTER

ONE

MINIMALLY INVASIVE LAPAROSCOPIC SURGERY

"If we review the evolution of medicine from the ancient times of 'Metaphysical theory of disease' to the development of subspecialties in medicine called 'Modern medicine' (1900), and finally the 'Health for all in 2000' slogan (1981), we find that humankind has always tried to maintain the quality of health care, constantly improving the health of his society." [6] What matters most was to extend the quality and quantity of life, for instance by means of developing more and more advanced surgical treatment methods. During the last decades the advent of minimally invasive techniques such as laparoscopic surgery shifted this paradigm to "less can be more", i.e. less trauma as well as equivalent or better recurrence, cure, and complication rates.

This chapter gives a short overview on the history of laparoscopy and describes, how laparoscopic surgery is performed and which advantages and drawbacks in turn laparoscopy implicates.

1.1 History

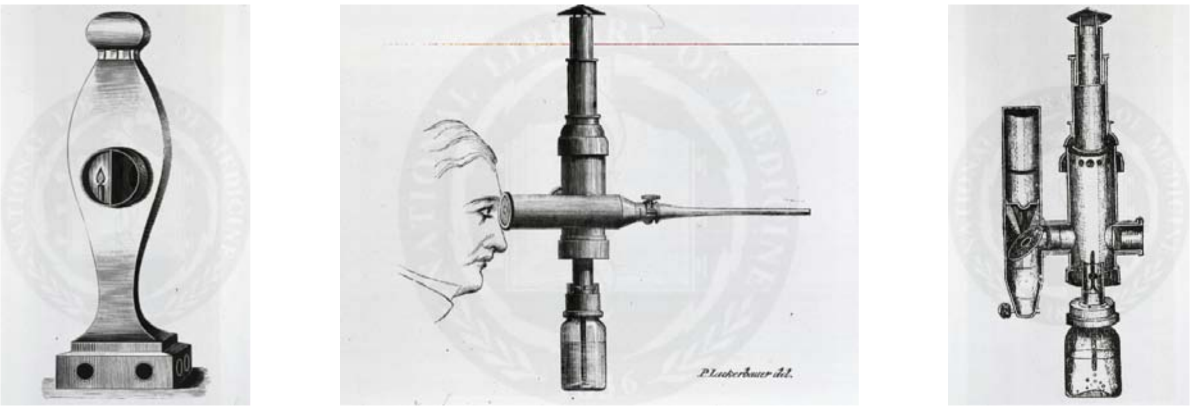
Already more than two millennia ago, first tools were developed to look inside the human body. The Greek Hippocrates (460 - ca. 370 BC), founder of a medical school on Cos, already describes a rectal speculum[1] for the treatise on fistula [91]. Ancient Greeks, Romans, and Egyptians already used similar instruments to examine all kinds of natural orifices of the human body such as rectum, vagina, ear, and nose. At that time, all examinations were dependent on natural light.

In 1805, the German Philipp Bozzini was the first utilizing artificial light coming from a candle placed inside a housing for his examinations [13]. To one side of the housing tubes of different sizes could be attached, which could be inserted into the orifices. Bozzini termed his invention "Lichtleiter" (cf. figure 1.1a). In 1853, the French Antonin Jean Desormeaux developed an open tube system incorporating mirrors and lenses to examine the urinary tract and the bladder. He was the first, who named his instrument "endoscope". Instead of a candle, he used a mixture of turpentine and alcohol as light

---

[1] A speculum is an instrument used to dilate an opening to look within a passage or a cavity.

source (cf. figures 1.1b and c). For many both Bozzini and Desormeaux are considered the “fathers of endoscopy”, Bozzini because of his early work and Desormeaux due to the great success of his endoscope, which was manufactured in rather large quantities.



(a) Bozzini’s Lichtleiter

(b) Desormeaux’s endoscope, illustrated by P. Lackerbauer del.

(c) Desormeaux’s endoscope incorporating lamp, chimney vent, and mirror

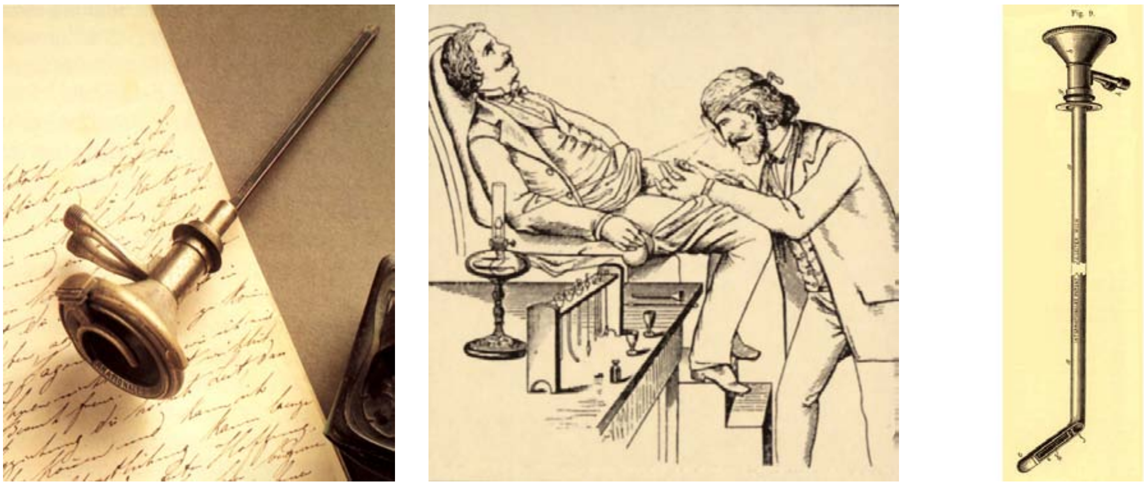
Figure 1.1: Historic endoscopes of Bozzini and Desormeaux (Images courtesy of the National Library of Medicine).

The first usage of endoscopes as telescopic instruments dates back to 1877, when the German Maximilian Nitze publicly presented an urethroscope and a cystoscope (cf. figure 1.2) with electrical lighting and lenses to examine the uthera and the bladder, respectively [142]. In 1879, together with the Austrian Josef Leiter he presented an improved version of the cystoscope, the so-called “Blasenspiegel”, which was appreciated by an international scientific community.

First diagnostic laparoscopic examinations were performed by the German Georg Kelling, who examined a dog’s peritoneal cavity and its contents using Nitze’s cystocope in 1901. 9 years later, the Swedish surgeon Hans Christian Jacobaeus was actually the first to coin the term “Laparothorakoskopie” for the examination of the human peritoneal, thoracic, and pericardial cavities. The word laparoscopy comes from the Greek words lapara (“the soft part of the body between ribs and the hip, flank, loin”, i.e. the abdomen) and skopein (“to look at or survey”) [91].

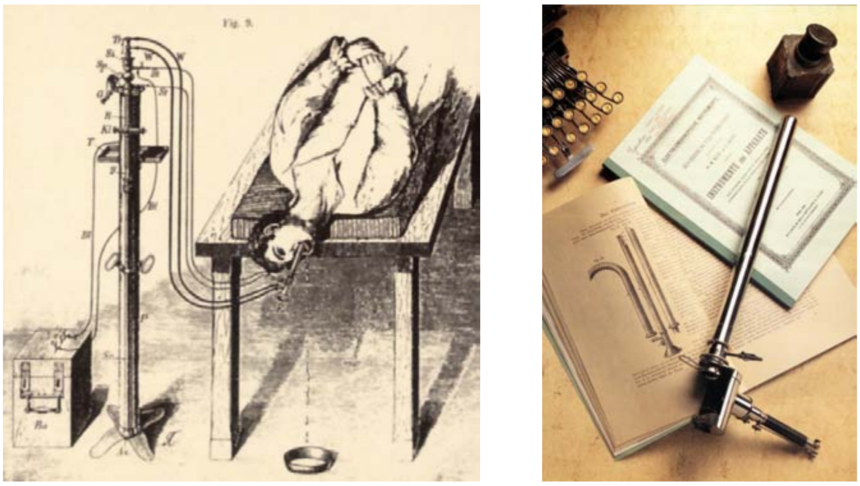
In the following decades, various rigid and flexible endoscopes were developed, for instance rectoscopes, esophagoscopes, gastroscopes, and bronchoscopes for the exploration of rectum, esophagus, stomach, and lungs (cf. figure 1.3). Endoscopy was mainly dedicated to diagnosis until the invention of video based systems in the 1980s, which are able to transfer the endoscope images to an external display. Thus, video endoscopy allows different team members to simultaneously see the endoscopic view. The operating surgeon can use both hands for the procedure while an assistant can position the endoscope. This feature was one of the major incentives for opening the field of endoscopic surgery.

Laparoscopic surgery started to evolve after the first successful laparoscopic cholecys-



(a) Urethroscope. (b) Examination using an urethroscope. (c) Cystoscope.

Figure 1.2: Nitze's telescopic instruments (Images courtesy of the Nitze-Leiter Museum of Endoscopy).



(a) Patient undergoing gastroscopy. (b) Panelectroscope for rectoscopy from 1907.

Figure 1.3: Gastroenterology (Images courtesy of the Nitze-Leiter Museum of Endoscopy).

tectomies[2] by O.D. Lukichev (1983), Erich Mühe (1985), and Phillipe Mouret (1987), respectively [182]. Since then, endoscopy was successfully introduced into other surgical disciplines as well.

More comprehensive reviews on the history of endoscopy and laparoscopy can be found for example in above mentioned references [13, 91, 97, 182].

1.2 Techniques

When Mouret performed his first laparoscopic cholecystectomy in 1987, he used four trocars[3] to insert the laparoscope along with a set of minimally invasive instruments into the patient's abdominal cavity [182]. Since then, laparoscopic cholecystectomy got a standard minimally invasive procedure to treat symptomatic gallstones. Nowadays up to about 98% of these interventions are performed laparoscopically with a very low conversion rate to open surgery [19].

In general, laparoscopic surgery is often applied for the (partial) resection of diseased organs. It is performed under general anesthetic. The procedure requires a few small incisions in the abdomen, which are used as trocar ports. Usually two to four plastic trocars of 11 and 12 mm diameter are placed to insert rigid surgical instruments. Another trocar is needed for the laparoscopic camera, which gives a magnified view onto the instruments and anatomy. The surgeon selects all ports by palpation of external anatomic landmarks, primarily based on his/her previous experience. An ideal selection of these ports can be one of the key issues in laparoscopic surgery, as the optimal choice of the instrument ports provides full access to the whole operation region as well as adequate surgeon dexterity.

The laparoscope usually has an oblique 30° optic to gain a wider perspective by rotating it about its own axis. This is especially useful when inserting the laparoscope camera relatively parallelly to an organ surface and for looking behind objects (cf. figure 1.5). To provide better visualization and exposure to the surgeon, pneumoperitoneum is applied, i.e. carbon dioxide (CO_2) is insufflated into the abdomen to enlarge the surgeon's working volume.

In contrast to cholecystectomy, where the whole gallbladder is removed, for liver resection only tumorous parts of the organ are resected, which are usually located in one of the eight liver segments (I–VIII) as defined by Couinaud [31, 169] (cf. figure 1.7). Liver resection may be used for metastasis from a colorectal cancer, hepatocellular carcinoma (HCC), and benign liver tumors or cysts.[4]

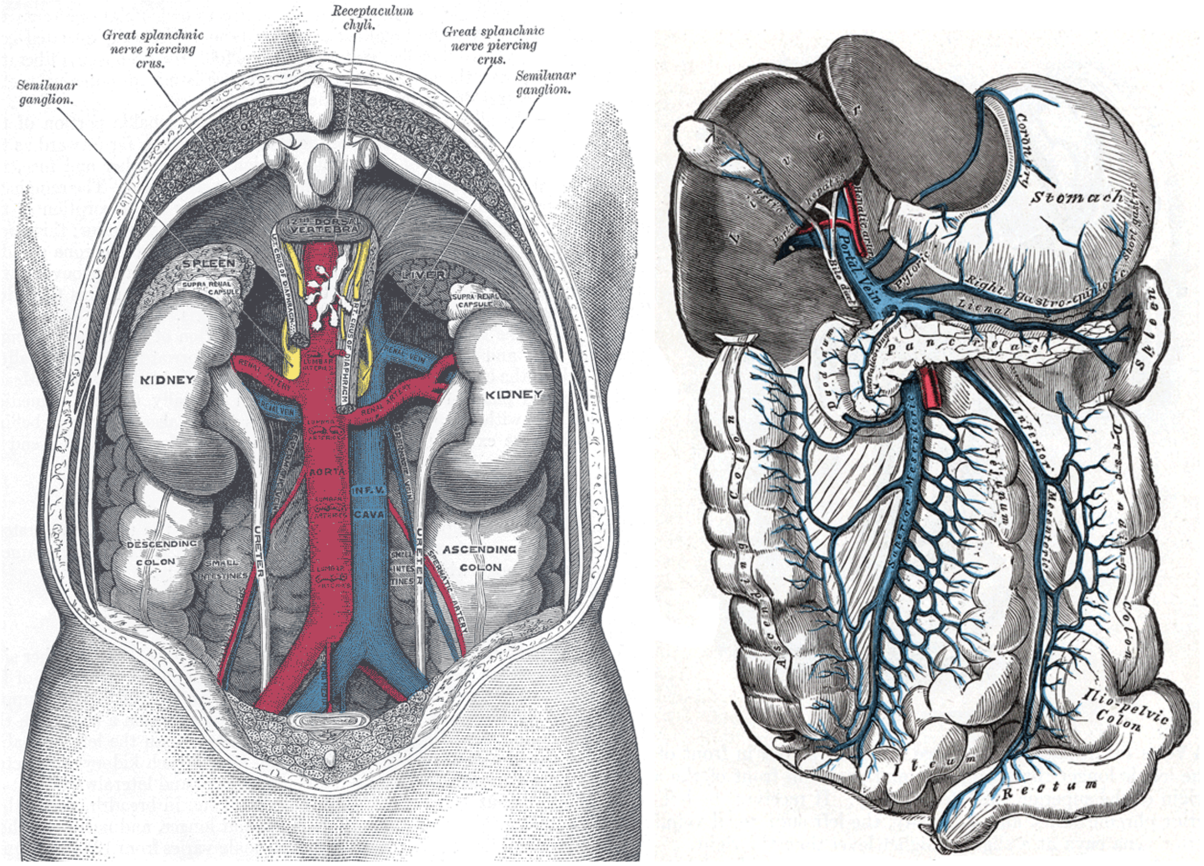
Mala and Edwin [102] provide a good insight into a typical totally laparoscopic liver

---

[2] Cholecystectomy is the surgical removal of the gallbladder.

[3] A trocar is a sharply pointed cylinder that can be used to insert instruments into the body cavity (various shape types of sharp trocar tips exist, e.g. pyramidal or conical). In figure 5.11, three plastic trocars are visible.

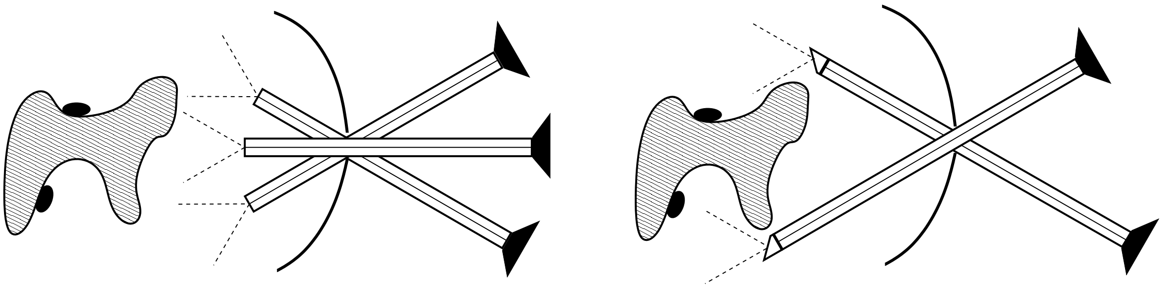
[4] Metastasis is the spread of cancer from its primary site to other places in the body. Hepatocellular carcinoma is a primary malignant tumor of the liver, which is capable of growth, invading into surrounding tissues, and spreading to distant tissues, contrary to benign tumors, which do not invade adjacent tissues and do not metastasize.



(a) Back view

(b) Frontal view

Figure 1.4: Illustration of the human abdominal anatomy (From Gray [58]).



(a) 0 ° optic.

(b) Forward oblique optic.

Figure 1.5: Advantage of a laparoscope with oblique optic (b) in comparison to a 0° optic (a): Using the same trocar, an oblique laparoscope allows to look behind objects (e.g. to see the black spots). Overall a wider field of view can be achieved by also rotating the oblique laparoscope. (Images courtesy of Vogt [184]).

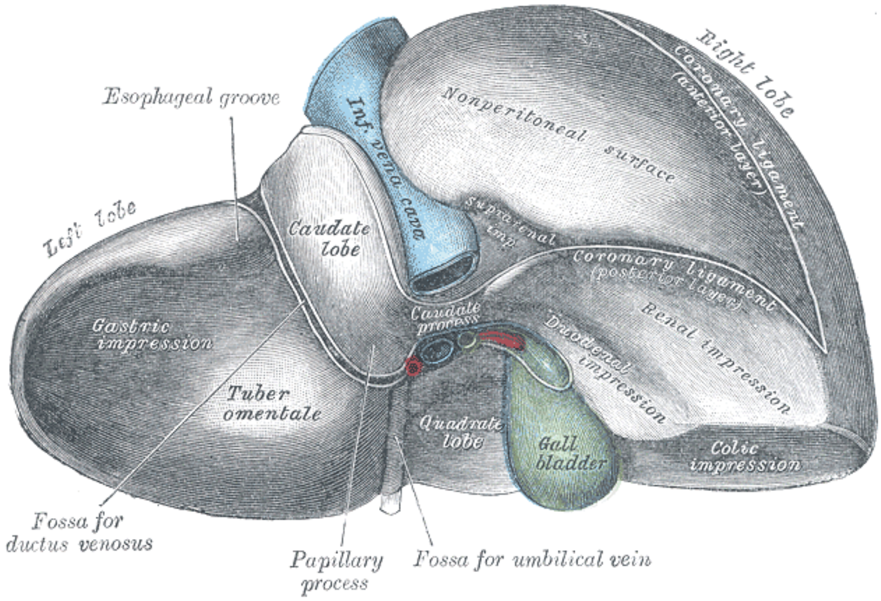
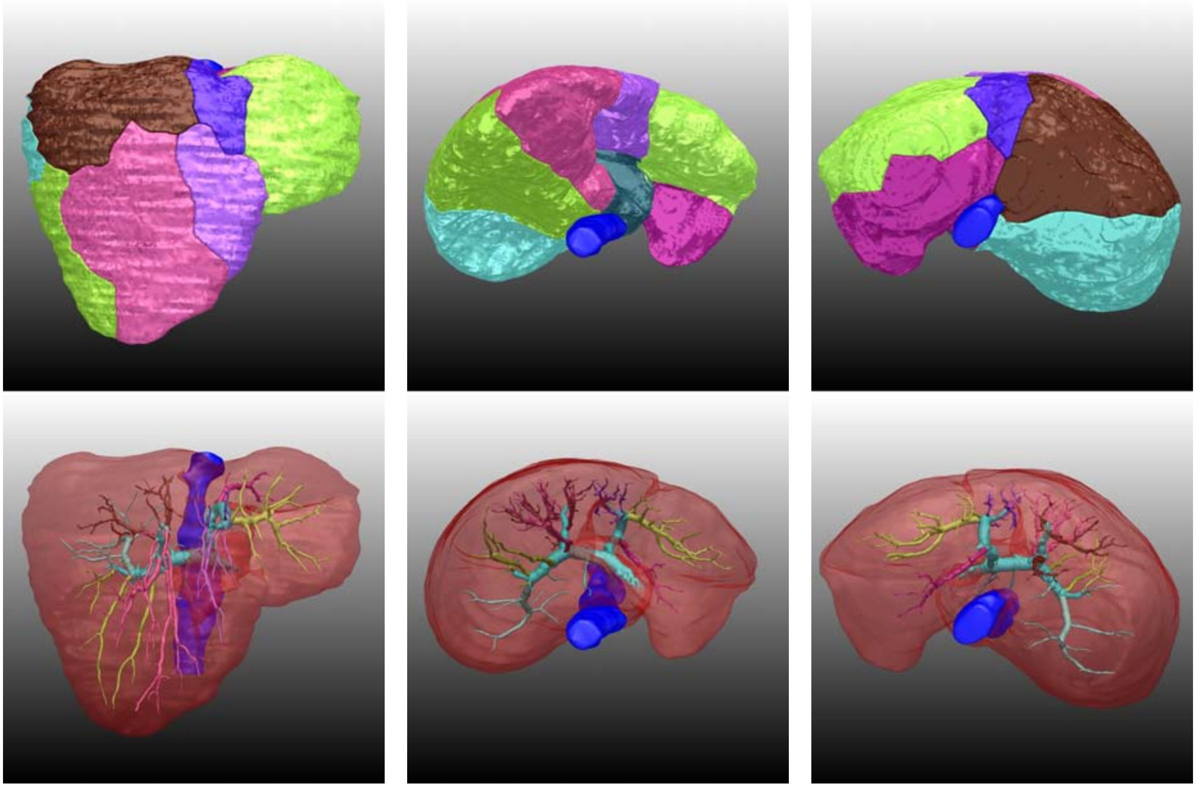


Figure 1.6: Anatomy of the liver (From Gray [58]).



(a) Frontal view. (b) Caudal view. (c) Cranial view.

Figure 1.7: Segmented liver segments (I–VIII) and their corresponding vessels as defined by Couinaud (Images courtesy of the German Cancer Research Center, Division of Medical and Biological Informatics).

resection procedure. After trocar port placement and CO_2 insufflation, electrocauterization[5] is utilized to mark the area to be resected on the liver surface. At the beginning of the resection, outer small vessels and bile ducts are sealed by an ultrasonic scalpel. An ultrasound (US) surgical aspirator can be used to fracture and evacuate liver tissue deeper inside the liver. An ultrasonic scalpel, diathermy[6], or clips can be applied to divide minor vessels and bile ducts, which remain preserved, while larger ones can be ligated by a stapling device or clips. To guide the division of vessels and eventually ensure adequate tumor clearance during resection, ultrasonography can be used for the assessment of vessel and tumor locations.

1.3 Advantages and Problems

Compared to traditional laparotomy (open surgery), several benefits were reported in the literature for laparoscopy (minimally invasive surgery). Besides leaving smaller scars, patient trauma and discomfort are reduced, which may result in shorter hospital stays, less postoperative complications, and faster rehabilitation. However, some laparoscopic procedures require a longer operating time and higher instrument costs. In the case of resection for malignancy, there is also the possibility of less tumor clearance [91].

If previously palpated external landmarks do not correspond to the individual internal anatomy of each patient, a misplacement of ports can occur, leading to time-consuming new port placement, which is considerable pain and takes extended recovery for every patient. Even experienced surgeons sometimes require port replacements during difficult interventions such as vessel dissection and lymph node dissection of the hepatoduodenal ligament[7] or along the vena cava inferior (see figures 1.4 and 1.6), so an exact port placement is of great importance.

To successfully perform laparoscopic interventions, highly trained and experienced specialists are required. When operating through small incisions, the surgeons are prevented from directly palpating organs, vessels, and tumors during the intervention. Besides this lack of tactile perception, they have to cope with a restricted 2D vision and a limited workspace compared to open surgery. At the same time, they need to maintain their dexterity and hand-eye coordination when handling minimally invasive instruments. Additionally, the target region often lies inside the organ, so it cannot be directly seen in the laparoscopic view onto the organ surface. In the case of liver resection, certain vessels need to be identified and ligated to avoid bleeding, which however mostly lie inside the liver. Therefore, advanced image guidance and visualization techniques are beneficial to support the surgical staff during the laparoscopic intervention [122].

---

[5] Electrocauterization is the process of burning or destroying tissue with electricity.

[6] Diathermy is a method of deep heating of tissues, accomplished by the use of high-frequency electrical current.

[7] The hepatoduodenal ligament is the portion of the lesser omentum that connects the porta hepatis of the liver and the duodenum.

CHAPTER

TWO

AUGMENTED REALITY IN IMAGE-GUIDED SURGERY

For many years, numerous imaging modalities and computer systems have been introduced and utilized to assist physicians in their everyday clinical life. Today, physicians are capable of performing more sophisticated as well as less invasive diagnosis and treatment of patients. Many procedures previously performed under open surgery can now be replaced by minimally invasive interventions, motivated by improved results and lower overall costs [206]. In abdominal surgery, an analogous trend from laparotomy to laparoscopy can be observed. It was not until the availability of image guidance that made this transition possible.

This chapter introduces the basic concepts of image-guided surgery and illustrates the detailed problems, which need to be tackled in laparoscopic surgery by means of image guidance. Furthermore, it describes, how augmented reality techniques can support image guided surgery and how they are utilized in this dissertation to facilitate new registration, visualization, and tracking concepts.

2.1 Image-guided Surgery

In minimally invasive surgery, direct visual feedback (as it was the case for open surgery) is replaced by indirect feedback. This indirect feedback relies solely on a combination of preoperative and intraoperative imaging data with additional information such as from tracking surgical instruments. In this way, a part of the information can be recovered, which got lost due to indirect feedback but which is needed to identify and to understand anatomical structures.

Many image-guided surgery systems have been developed, both on the research and commercial side. As summarized by Yaniv and Cleary [206], these systems typically rely on a surgical plan, which in turn is based on preoperative imaging data acquired during computer assisted diagnosis (CAD) [35, 40], where for instance suspicious lesions such as tumorous regions are detected and classified [110] or optimal trocar ports can be computed [2, 23, 25]. After registration, i.e. alignment of the coordinate system of the preoperative imaging data with the intraoperative coordinate system of patient and instruments, this plan is executed. During the intervention, image-guided surgery systems visually assist the surgeon by multiple interfaces and displays to perform a successful navigation of surgical instruments to a target region.

Recent reviews on (laparoscopic) image-guided surgery were published by Marvik et al., Peters, and Yaniv and Cleary [122, 137, 206]. In general, an image-guided surgery system can incorporate imaging (including low level image processing), segmentation, tracking, registration, interaction, as well as visualization techniques.

2.1.1 Imaging

Any surgical intervention is based on its underlying pathology, i.e. diagnosis of a disease through the examination of a patient. This examination is mainly achieved by preoperative anatomical and/or functional imaging such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission computed tomography (SPECT), or ultrasound (US) imaging (also referred to as (ultra)sonography), often combined with contrast agent administration to highlight e.g. vessels and tumors. Alternatively, more invasive methods such as diagnostic laparoscopy can be applied. The resulting images are two-dimensional (2D) slices or projections, three-dimensional (3D) volumes, or four-dimensional (4D) volumes over time, which support the physician in the diagnostic process. The same data sets or intraoperatively acquired updates can be used for instrument navigation during image guided interventions (to display the current position and orientation of surgical instruments within the data set).

More details on the physical principles, technology, equipment, and procedures related to image formation can be found for instance in the book of Hendee and Ritenour [65]. In a recent article [200], Wolbarst and Hendee also describe emerging trends and technologies in medical imaging such as optical and near-infrared imaging, terahertz imaging, microwave imaging, thermography, or intrinsic and applied electric and magnetic fields.

2.1.2 Segmentation

The data sets formed by afore said imaging technologies are generally discrete sets of pixels or voxels with certain intensities. As it can be difficult to distinguish between healthy or diseased tissues, organs, and bones, the data sets can be further processed to obtain labeled partitions of the patient anatomy such as liver segments, vessel trees, and tumors. This supports the physician during diagnosis to study anatomy, localize pathology, quantify tissue volumes, and plan the treatment, as well as during image-guided surgery [138], where a clear distinction of anatomy can be helpful. However, segmentation techniques are usually tailored to certain specialities and often require time-consuming interaction with the physician.

2.1.3 Tracking

In order to continuously determine the position and orientation ("pose") of surgical instruments in regard to the patient data set, tracking systems can be employed. Tracking bodies or sensors are integrated into or attached to the instruments and/or the patient (anatomy) and localized in the coordinate frame of the tracking system. Tracking systems usually provide pose information with six degrees of freedom, i.e. three degrees of freedom for translation and three degrees of freedom for orientation. Typical tracking devices used in medical applications are mechanical arms (which are mainly used in robotically assisted interventions), optical tracking systems, and electromagnetic tracking systems, all coming with their own advantages and drawbacks.

Martelli et al. compared an optical to a mechanical tracking system [103]. While both systems feature submillimeter accuracy, the authors slightly favor the optical system due to its easy use and capability of tracking multiple instruments, which can be important issues in the operating room. A comparison of two commercially available optical tracking systems was performed by Khadem et al. [76]. In detail, they evaluated the FlashPoint system, which is now offered by BIG (Boulder Innovation Group, Inc., Boulder, CO, USA)[1] and the Polaris system by NDI (Northern Digital Inc., Waterloo, OT, Canada)[2] in five different configurations. They conclude that both tracking systems are comparable in terms of jitter with the most jitter along the viewing direction of the tracking cameras getting worse with increasing distance. Therefore, they recommend to place the tracking cameras as close as possible to the operating field in such a way, that the least clinically significant direction is aligned with the viewing direction of the cameras. In their analysis of the stability of electromagnetic tracking systems, Schicho et al. conclude that electromagnetic trackers do not yet reach the stability and accuracy of their optical pendants [149]. They propose to perform a detailed risk analysis including the definition of accuracy security margins and also to test the stability of the electromagnetic tracking system with the use of surgical instruments, before it is utilized in the operating room.

The most important parts of instruments, which need to be located, are usually their tool tips, e.g. forceps or scissors. If the instruments are rigid, optical tracking systems or mechanical arms can be employed for tracking, while flexible instruments can almost only be tracked by electromagnetic systems because of the missing line of sight of the instrument tips, which are located inside the patient. An alternative to track flexible instruments may be the integration of a so-called ShapeTape (Measurand Inc, Fredericton, NB, Canada)[3], as done by Koizumi et al. for flexible endoscopic ultrasound [81].

More details on tracking concepts in general are provided for example by Bishop et al. [18], Rolland et al. [146], and Welch and Foxlin [192], details on tracking systems for surgical navigation e.g. by above references [76, 103, 149] and Birkfellner [15].

2.1.4 Registration

One of the key components of image-guided surgery systems is the registration of preoperative and/or intraoperative patient data to the intraoperative setup, i.e. patient and instruments. All entities need to be brought into one common world coordinate frame, which usually is the coordinate frame of the tracking system. Various noninvasive and invasive methods have been developed to register the patient and his/her imaging data,

---

[1] `http://www.boulderinnovators.com/`

[2] `http://www.ndigital.com/`

[3] `http://www.measurand.com/`

for instance the attachment of (stereotactic) frames, adapters, fiducial markers, or simply natural landmarks, which can be located in both the tracking coordinate system and the imaging data coordinate system. Alternatively, surface points of relevant patient anatomy can be intraoperatively collected in the coordinate frame of the tracking system, e.g. by a tracked pointing device or laser range scanner, and matched to surfaces generated from the patient data set.

Additionally, multiple imaging modalities may be used to guide the surgical staff, for instance PET and CT data for the fusion of functional and anatomical patient information. To align the coordinate systems of these modalities, image to image registration methods are commonly applied. A 2D to 3D image registration method is for instance presented by Grzeszczuk et al. [61] and Murphy [124], who use a fluoroscope to acquire intraoperative X-ray images and register them to digitally reconstructed radiographs (DRR) created from preoperative CT. This procedure has the advantage that no fiducials have to be added to the patient while keeping high accuracy. By also tracking the C-arm, its subsequent motions can be updated in the registered CT data set.

In abdominal and thoracic surgery, where the anatomy is deformed intraoperatively, the registration of patient and imaging data becomes very difficult, especially if preoperative data is to be aligned with the patient. The development of deformable registration algorithms is currently an issue of major interest to the scientific community.

Elaborate reviews on registration methods are provided by Maintz and Viergever [101], Hill et al. [67], as well as Zitová and Flusser [208].

2.1.5 Interaction

Interaction between the surgical staff and the image-guided surgery system is a topic, which was often underestimated especially for the first developed systems. A direct communication between the surgical staff and the system software was rarely possible, only via a system engineer [183]. Early systems often relied on input from standard non-sterilizable keyboards and mice. In the last years, several sterilizable interaction alternatives were proposed, for instance touch screens, binary input devices such as foot switches or tool embedded switches, tracked virtual keyboards, as well as speech and gesture recognition systems [206].

Another important interaction issue is the way how the data required for image-guided surgery is presented to the surgical staff. A standard technique utilizes a monitor that is placed next to the operating table and displays four major areas, where three of them are used for axial, sagittal, and coronal views of the patient data, and the fourth one is used for a 3D volume rendering (see below – section 2.1.6). To bring the displayed data closer to the physicians, several alternatives were developed for in-situ (in the place) visualization, for instance miniature LCD screens, stereo operating microscopes and binoculars, head mounted displays (HMDs), and semi-transparent mirrors [206].

2.1.6 Visualization

All imaging data needs to be presented to the surgical staff in an appropriate way. In the case of 3D volumetric data, three standard visualization methods are often used: Slice based, surface based, and direct volume rendering. Slice based techniques usually present orthogonal slices, i.e. axial, sagittal, and coronal views of the patient volume, or, more rarely used, oblique slices. Surface rendering methods require a previous data segmentation to obtain partitions of the anatomy, which are further processed in an intermediate step to generate 3D models of the anatomy, e.g. by using the marching cubes algorithm [100]. This intermediate step can be avoided by direct volume rendering techniques, which are able to directly present the 3D volume to the user as 2D projections by e.g. applying transfer functions, which map certain intensities and colors to the voxels of the volume. An overview of volume rendering techniques can be found in a tutorial of Meißner et al. [111].

The major advantage of image guidance is the simultaneous visualization of tracked surgical instruments with respect to the imaging data, for instance by showing (projected) 3D models of the surgical instruments or just lines or circles representing the axis or tool tip of an instrument within the volume.

Several groups also started to incorporate virtual reality (VR) and augmented reality (AR) techniques into image-guided surgery systems for an advanced and intuitive visualization, which is described in the following section.

2.2 Augmented Reality in Endoscopic Surgery

While virtual reality lets the user entirely immerge into a computer generated virtual world and interact with the computer, augmented reality takes the opposite approach: Virtual, computer generated objects are added to the real physical world [193]. Additional information is provided to the user and fused with the real world in an augmented reality environment or, synonymously, in a mixed reality environment. According to Milgram et al. [114], mixed reality also comprises the so-called augmented virtuality, i.e. a virtual environment in between augmented reality and virtual reality, which is enhanced by information from the outside reality, such as texture images or videos. However, nowadays mixed reality is commonly referred to as a synonym for augmented reality, along with the term “enhanced reality”, which some authors use [59, 105, 106, 112, 154]. Azuma et al. concisely define an augmented reality system as a system with the following properties [4, 5]:

1. Real and virtual objects are combined in a real environment, they appear to coexist in the same space.
2. The system runs interactively and in real time.
3. Real and virtual objects are registered, i.e. aligned with each other.

Applied to the surgical context, this means that real objects are for instance the patient and instruments, while their virtual counterparts are instrument models, imaging data, or additional information such as paths towards the target region, which are overlaid onto the surgeon's view.

The first medical system of this kind has been realized in the middle of the 1980s for neurosurgery [54, 145]. It integrates co-registered images of segmented CT volumes into the view of an operating microscope. Other medical augmented reality systems have been developed thereafter. Bajura et al. [7] report on a head mounted display for visualization of ultrasound images registered to the patient's anatomy. Lorensen et al. [99] developed an externally tracked camera for image overlay on a monitor for neurosurgery. Masutani et al. [106] report on an autostereoscopic display overlaying images via a semi transparent mirror on the operation site.

Even though the augmentation of additional imaging data on live endoscope images seems to be a straightforward idea, the first augmented reality systems in endoscopy did not appear before the end of the 1990s. Freysinger et al. developed an image guidance system for endoscopic ENT surgery, which is able to superimpose a bent 3D path towards a predefined target on the endoscopic images [53]. Shahidi et al. proposed a system for brain surgery in order to overlay preoperative volumes of MRI or CT, respectively, on live endoscope images [155]. Konen, Scholz et al. presented a navigation system for neurosurgery based on image processing [82, 152]. While Freysinger et al. employ electromagnetic tracking, the other two systems use optical infrared tracking technology. All systems are able to augment virtual objects on the images of a rigid endoscope and display them on a separate monitor.

A different augmentation approach was taken by Fuchs et al. [55]. They propose to superimpose the images of a laparoscope capable of depth extraction onto a stereo head mounted display for an intuitive 3D visualization, which may be able to restore the physician's natural point of view and head motion parallax.

2.2.1 Motivation

In general, the intraoperative augmentation of endoscope images is motivated by three major interests: Context finding, visualization of hidden structures, and enhancement of images.

2.2.1.1 Context Finding

The point of view and the horizon of an endoscopic image is constantly changing. Recovering each of them requires much concentration, since an operating surgeon generally does not move the endoscope him/herself and the endoscopic field of view is very limited.

Dey et al. [34] project endoscope images on segmented surfaces for providing context and creating endoscopic 3D panorama images. Similarily, Mountney et al. recover a 3D map of the scene from stereoscopic images [119]. Kawamata et al. [74] visualize the anatomical context by painting virtual objects in a larger area than endoscope images are available. Instead of augmenting the images of an endoscope camera, Ellsmere et al. [38, 39] and Estépar et al. [41] suggest to overlay endoscopic live ultrasound images onto CT slices and segmented CT data for improved context sensing. Similarly, Linte et al. visualize the relationship of ultrasound, instruments, and patient anatomy in a virtual environment for the guidance of mitral valve implantations [96].

2.2.1.2 Visualization of Hidden Structures

The visualization of hidden structures such as tissue that is covered or tissue that can only be distinguished by other imaging devices than by an endoscopic camera can be very helpful for both intraoperative surgery planning and navigation. Shahidi et al. [154] for example overlay structures that are not visible by an endoscope for the guidance of a surgical dissection during sinus surgery and ventriculostomy. Their system is used by Mayberg et al. for neurosurgery, where in axial, coronal, and sagittal MR images the location and trajectory of the endoscope tip is visualized and a virtual 3D endoscopic view containing lesions and adjacent healthy structures is shown [108]. Scheuering et al. propose a system that is able to overlay rigidly registered liver data on endoscope images for trocar placement and navigation [148].

2.2.1.3 Image Enhancement

Augmentation of endoscopic images does not necessarily mean fusion with other virtual objects such as imaging data. It can also refer to the enhancement of the endoscopic images, which however loosens the original definition of augmented reality. Scholz et al. [152] suggest several image based methods with a tracked endoscope to overcome typical limitations of endoscopy such as loss of sight or fixation of the endoscope by replay of former images, image mosaicing, and landmark tracking. They also propose to tackle brain tissue shift by a recalibration based on anatomical landmarks. Krüger et al. [86] evaluate endoscopic distortion correction, color normalization, and temporal filtering for clinical use.

Vogt et. al. [186] describe an image enhancement technique based on light fields. As the endoscope tip usually contains a strong point light source, specular highlights occur likely, which make the examination of anatomical structures difficult. Using a light field approach the highlights can be significantly reduced.

2.2.2 Specific Issues

In order to augment patient data directly on the endoscopic view, various issues need to be addressed. An adequate tracking method needs to be chosen to localize the endoscope. Offline, the endoscope needs to be calibrated in order to model its projection geometry. Additionally, all involved coordinate frames need to be registered with the patient data in a common world coordinate frame. Finally, all live data need to be synchronized to each other to ensure a smooth overlay containing data from the exactly same points of time.

2.2.2.1 Tracking

Tracking technology is one of the bottlenecks for augmented reality in general [5]. As an exception, for medical augmented reality this is quite different. As the working volume and hence the augmented space is indoors, predefined, and small, the environment, i.e. the operating room, can be prepared for the augmented reality system. Optical (infrared) tracking systems are already in use in modern operating rooms for intraoperative navigation. In orthopedics, trauma surgery, and neurosurgery, which only require a rigid body registration, available navigation systems proved to be sufficiently accurate. King et al. [78] proved in clinical studies to have overall errors in the submillimeter range for their microscope based augmented reality system for neurosurgery.

Tracking systems for endoscope localization are mainly optical, electromagnetic, or mechanical. Optical tracking systems are usually fiducial based, so they can guarantee a predictable quality of tracking.

Nicolau et al. [132] propose a registration with error prediction for endoscopic augmentation. An online error estimation is an important feature, since physicians have to rely on the visualized data. Bauer et al. presented a mathematical framework for the propagation of optical tracking errors [9], which can be used to visualize the covariance matrices of these errors [10, 159].

Rigid Endoscopes Most presented systems for endoscope augmentation use an optical tracking system to externally localize a body of fiducials [33, 34, 74, 104, 150, 152, 154, 170, 197, 204]. The body is attached close to the camera head of a rigid endoscope, so the required line of sight can be ensured, when the endoscope shaft is inside the patient.

Mourgues et al. [120] and Leven et al. [94] describe endoscope augmentation in a robotic surgery system. The tracking can be done implicitly by the robot. Therefore no additional tracking system is necessary.

Flexible Endoscopes Flexible endocopes cannot be tracked by optical tracking systems. Bricault et al. [21] describe the registration of bronchoscopy and virtual bronchoscopy images using only geometric knowledge and image processing. The algorithms in use did not have real time capability, however they proved to be stable in recorded videos. As opposed to Bricault's shape from shading approach, Mori et al. [118] use epipolar geometry for image processing. In order to improve the performance of their registration algorithm they suggest the addition of electromagnetic tracking of the bronchoscope [117]. For the fusion of the bronchoscopic video with a target-path, Wegner et al. restrict electromagnetic tracking data onto positions inside a previously segmented bronchial tree [191]. Some groups use electromagnetic tracking exclusively, as e.g. Klein et al. [80].

2.2.2.2 Calibration

Because of their wide angle optics, endoscopes suffer from a noticeable image distortion. If a perfect distortion-free pinhole camera model is assumed for superimposition, a particular source of error in the augmented image will be introduced [75], which can be neglected in other augmented reality systems with telephoto optics. Common types of distortion are radial distortion (also referred to as barrel distortion) and tangential distortion. Either the endoscope image has to be undistorted or the rendered overlay has to be distorted to achieve a perfect superimposition. While first approaches [163] took several minutes to undistort a single endoscope image, the undistortion can now be achieved in real time: De Buck et al. [33] undistort sample points in the image and map a texture of the endoscope image on the resulting tiles; Shahidi et al. [154] precompute a look-up table (LUT) for each pixel for real time undistortion.

In order to model the geometry of an endoscope camera, the intrinsic camera parameters focal length and principal point need to be determined. This can be achieved using well-established camera calibration techniques [64, 177, 207]. Most systems assume the focal length of an endoscope camera to be kept constant, although many endoscopes incorporate zoom lenses to change it intraoperatively, invalidating a certain calibration. Stoyanov et al. suggest a system to automatically adjust the calibration for intraoperative changes of the focal length of a stereoscopic camera [166]. Even though models for the calibration of monoscopic cameras with zoom lenses exist [199], they are not easily applicable to endoscopes. Preferably, these models require the (automatic) determination of the physical ranges for the lens settings e.g. in terms of motor units, but the zoom settings of endoscopes are usually manually adjusted and not by a precise motor.

To obtain the rigid Euclidean transformation from the camera coordinate frame to the coordinate frame of an attached tracking body or sensor, most authors avail themselves of hand-eye calibration techniques [14, 120, 131, 148, 150]. Alternatively, a tracked calibration pattern can be employed, whose physical coordinates are known with respect to the tracker [33, 104, 154].

For certain applications such as laparoscopy, oblique-viewing endoscopes are used, for which the viewing directions are changeable by rotating the scope cylinder. Yamaguchi et al. developed a calibration procedure for such endoscopes [204].

2.2.2.3 Registration

Registration algorithms are well discussed in the community, but their integration into the surgical workflow is always a trade-off between practicability and accuracy.

Registration of patient data can be performed with fiducials that are fixed on the skin or implanted [107]. These fiducials must be touched with a tracked pointer for the registration process. Alternatively, the fiducials can be segmented in the images of a tracked endoscope rather than touching them with a pointer for usability reasons. Stefansic et al. propose the direct linear transform (DLT) to map the 3D locations of fiducials into their corresponding 2D endoscope images [164]. Baumhauer et al. study different methods for endoscope pose estimation based on navigation aids stuck onto the prostate and propose to augment 3D transrectal ultrasound data on the camera images [11]. Using this method, no external tracking system is needed.

Especially for maxillofacial surgery, fiducials can be integrated in a reproducibly fixed geometry [78]. For spine surgery, Thoranaghatte et al. try to attach an optical fiducial to the vertebrae and use the endoscope to track it in situ [170].

The accuracy of a fiducial-based registration varies on the number of fiducials and quality of measurement of each fiducial, but also on the spatial arrangement of the fiducials [52].

Grimson et al. [59] follow a completely different approach by matching surface data of a laser range scanner to CT data of the head. For sinus surgery, Burschka et al. propose to reconstruct 3D structures using a non-tracked monocular endoscopic camera and register them to a preoperative CT data set [22]. For spine surgery, Wengert et al. describe a system that uses a tracked endoscope to achieve the photogrammetric reconstruction of the surgical scene and its registration to preoperative data [197].

When it comes to the registration of deformable anatomy such as liver or heart, very promising approaches for endoscope augmentation are based on the use of intraoperative imaging data. For instance, ultrasound images may be used, which are directly overlaid onto the endoscopic view to visualize their spatial relationship to endoscope images, as proposed by Nakamoto et al. [127] or Leven et al. [94].

2.2.2.4 Time Synchronization

Time synchronization of tracking data and video images is an important issue for an augmented endoscope system. In the unsynchronized case, data from different points of time would be visualized. Holloway et al. [68] investigated the source of errors for augmented reality systems. The errors of time mismatch can raise to be the highest error sources when the camera is moving. To overcome this problem, Jacobs et al. [72] suggest methods to visualize data from multiple input streams with different latencies from only the same point of time. Sauer et al. [147] describe an augmented reality system that synchronizes tracking and video data by hardware triggering. Their software waits for the slowest component before the visualization is updated. For endoscopic surgery, Vogt [184] also uses hardware triggering to synchronize tracking and video data by connecting the S-Video signal (PAL, 50 Hz) of the endoscope system to the synchronization card of the tracking system, which can also be run at 50 Hz.

2.3 Problem Statement

Today's available image-guided surgery systems are primarily used to assist surgeons during neurosurgery or orthopedics, where mainly rigid anatomy is involved. Abdominal and thoracic surgery, in contrast, involves a number of deformations in between preoperative diagnosis/planning and surgery: heartbeat, lung deflation, respiratory motion, patient relocation, carbon dioxide insufflation, and the intervention itself. Therefore, the accuracy requirements of the addressed surgical procedure have to be carefully analyzed in detail prior to the development of an image-guided surgery system, as they may vary between centimeters (e.g. for port placement) and (sub)millimeters (e.g. for intraoperative navigation for vessel clipping or thermal ablation). For the latter, preoperative imaging data can hardly be used. This is one of the main reasons why image-guided surgery systems for these disciplines are topic of current research and no commercially available solutions exist. Information on deformations of the patient anatomy needs to be incorporated into image-guided surgery systems, so abdominal and thoracic minimally invasive surgery dealing with soft tissue can be succesfully addressed.

Augmented reality has the potential of providing a smooth integration of visualization and guidance. As stated by Shuhaiber [158], it can support experienced surgeons to perform more complete and radical operative therapies as well as guide and advise novice surgeons of critical anatomic landmarks. “Further research is” however “needed to evaluate its long-term clinical impact of augmented reality on patients, surgeons, and hospital administrators. Its widespread use and the universal transfer of such technology remains limited until there is a better understanding of registration and ergonomics” [158].

The issue of patient registration is tightly coupled to the surgical workflow. An accurate patient registration method in abdominal surgery should not alter the workflow considerably in terms of time and costs. Additionally, a certain degree of confidence and the required accuracy for registration and navigation should be maintained. An image-guided surgery system will only be useful, if it does not change the conventional surgical workflow or only to a certain degree. A change however has to be justified by an improved patient outcome or at least equal patient outcome combined with less costs or time.

2.4 Main Contributions

This dissertation introduces new guidance, tracking, and visualization concepts for laparoscopic surgery based on intraoperative imaging and augmented reality, which improve currently available image-guided surgery solutions that are not able to deal with patient deformations and sometimes hamper the surgical workflow. Critical phases in the surgical workflow are supported by the presented system: Starting with the assistance for port placement by registered virtual flights of the laparoscope camera through the patient (see section 2.4.1), a complete medical augmented reality solution is presented in section 2.4.2, which incorporates novel intraoperative multimodal image guidance using a mobile C-arm capable of cone-beam CT and laparoscopic ultrasound. All components are embedded into the medical augmented framework CAMPAR (see also section 4.4).

Depending on the type and complexity of the intervention and the equipment available in the operating room, all proposed guidance components can be combined or used individually. All methods were validated in several phantom, ex vivo, and in vivo animal experiments[4] in close collaboration with surgeons (see chapter 5).

The dissertation work resulted in a series of mostly peer reviewed publications and patent applications, which are all listed in appendix A. Abstracts of major publications not addressed within the scope of this work can be found in appendix B.

2.4.1 Patient Registration for Port Placement

As already stated in chapter 1, optimal port placement is an important issue especially for complex interventions. A good port placement can improve the surgeon dexterity, and also additional pain to the patient caused by possible replacements can be avoided.

The accuracy requirements for a good port placement are around two centimeters. This is due to the fact that the patient’s skin and hence inserted trocars can be moved up

---

[4] Ex vivo means outside an organism, e.g. out of the living body of an animal. Analogously, in vivo means inside an organism, e.g. in the living body of an animal.

to a maximum of about two centimeters to compensate for possible port displacements. Therefore, image guidance based on preoperative data rigidly registered to the patient may be sufficient to support the surgical staff in choosing optimal port locations.

Related Work Several methods have been proposed to improve and automate the optimal placement of ports for minimally invasive surgery [2, 23, 25, 148, 171]. These methods all rely on the manual or semi-automatic segmentation of preoperative imaging data from CT or MRI, which is essential for reconstructing models of the anatomy, e.g. ribs, liver, and soft tissue. These 3D models can be used to automatically compute optimal port locations [2, 23, 25], which serve as important guidelines for surgeons. This can improve the learning curve especially of untrained surgeons.

A practical and accurate way to transfer the planned port locations to the operating room is however needed, meaning the patient has to be registered to his/her preoperative data. This patient registration process is usually based on matching anatomical or artificial landmarks, which are visible on both the patient and his/her preoperative data. Adhami and Coste-Manière use the end effectors of the da Vinci® surgical system to point to fiducials, which are attached to the patient [1]. Due to their shape and intensity, the fiducials can be segmented automatically in the CT data. Intraoperatively, the physician controlling da Vinci® moves the end effector of a robot arm to every single fiducial in order to get its position in the robot coordinate frame. As reported by Falk et al. [43], this task takes approximately two minutes. Similarly, Selha et al. use the sensor of an additional electromagnetic tracking system [153] as a pointing device, basing their registration on anatomical landmarks.

Contribution This dissertation proposes a practical alternative method to register the CT data to the patient and to visually assist the surgical staff during port placement [45, 48]. Spherical CT visible self-adhesive fiducials are affixed on the patient's skin, which can be done already for a diagnostic scan. The fiducials need to remain on the skin until the intervention. Alternatively, their locations can be marked, e.g. by a felt-tip pen, so they can be reattached before the the intervention. The fiducials can be segmented fully automatically in the patient's CT data.

Intraoperatively, instead of pointing to the fiducials, the tracked laparoscope is only moved around the fiducials and a set of images is acquired from differing, but arbitrary poses. To simplify the acquisition process, not all fiducials need to be seen by the camera in a single image. By automatically detecting the fiducials in these images, their 3D positions are reconstructed in the coordinate frame of the optical tracking system. Point based graph matching and registration methods enable their fully automated matching with the CT data. For port placement, a surgical staff member simply moves the tracked instruments or laparoscope to the positions where he/she wishes to place their corresponding ports. A virtual camera is placed on top of the instrument end effectors or the camera center of the laparoscope. It is able to simulate a flight through the patient's interior by rendering the CT volume as it would be seen by the laparoscope. In this natural way, optimal port placements can easily be identified without prior segmentation of patient's anatomy or the use of a pointing device. In addition, there is no need to identify anatomical landmarks or touch them, which for da Vinci® usually has to be performed by the physician controlling the system, as described above. The proposed method can be performed by any surgical staff member and is applicable to any tracked laparoscope, no matter whether it is tracked by an optical tracking system or a mechanical one such as da Vinci®. It could also be applied to other fields of minimally invasive surgery such as thoracoscopic surgery, where a good port placement is as important as in laparoscopic surgery.

2.4.2 Intraoperative Registration-free Multimodal Imaging and Visualization

Intraoperative accuracy requirements for laparoscopic surgery are different than for orthopedic surgery or neurosurgery. A discrimination of about half a centimeter is usually required. While lymph nodes are considered to be inflicted by a tumor in case the diameter is more than ten millimeters, canalicular structures such as vessels and bile ducts play a critical role in case they are equal to or thicker than five millimeters. To fulfill these requirements, it is hard to (deform and) register rigid preoperative data to match the intraoperative situation. Intraoperative imaging however can provide valuable up-to-date patient data.

A major novelty presented in this dissertation is the fusion of multiple intraoperative imaging modalities without need for tedious manual or interactive registration. Patient or patient imaging data is not used for registration, but is intrinsically registered to the tracking system. Therefore, there is no need for detection and matching of anatomical landmarks or fiducials on the patient, as used e.g. during port placement (cf. previous section 2.4.1). This makes the intraoperative visualization of the proposed system registration-free, i.e. it is solely based on imaging, navigation, and visualization, all in the same external tracking coordinate system. This dissertation introduces the use of a video-imaging system, i.e. laparoscope, within a multimodal registration-free navigation system providing imaging data from a mobile 3D C-arm [45, 46] and a laparoscopic ultrasound transducer [47]. This is the first time several different imaging systems are integrated into an augmented reality solution using the registration-free concept.

By means of optical and electromagnetic tracking systems, both rigid (C-arm) and flexible (laparoscopic ultrasound) imaging devices can be tracked. While the registration-free concept has great potential for further developments based on other rigid imaging modalities, e.g. interventional stationary C-arms such as DynaCT[5], it can surely also be extended to deal with procedures based on flexible endoscopy such as bronchoscopy or NOTES (natural orifice translumenal endoscopic surgery) in the future. The technology presented here can be the basis for such trends in minimally invasive surgery.

The following two sections exemplarily depict the advantages of the registration-free multimodal imaging concept in the context of laparoscopic surgery, where both a tracked mobile C-arm and a tracked flexible laparoscopic ultrasound transducer can be of great help to the surgical staff.

---

[5] `http://healthcare.siemens.com/dynact/`

2.4.2.1 Mobile C-arm Based Vessel Augmentation

As described by Mala and Edwin [102], laparoscopic liver resection is a technically demanding procedure, usually performed by well-trained surgeons. This especially applies to difficult cases, where the tumor is embedded into the vessels (in close proximity to vessels) or located between the hepatic veins (segment VIII or IVa). An intraoperative visualization of these blood vessels in regard to the laparoscope or other surgical instruments can assist the surgical team during such challenging procedures.

Related Work For orthopedics and neurosurgery, where mainly rigid structures are involved, navigation systems aligning imaging data in respect to the patient in order to guide the surgical team are commercially available[6]. Some of them, which are based on MRI or C-arm CT and used for neurosurgery, orthopedics, and trauma surgery, are even registration-free, as noted by Yaniv and Cleary [206]. For instance, one of these C-arm CT systems is used routinely on a daily basis at the Chirurgische Klinik und Poliklinik, Klinikum der LMU, Munich, for spine, pelvis, hip, knee, and ankle surgery[7]. For this system, both the C-arm and surgical instruments are optically tracked during surgery. The accuracy of such a registration-free C-arm-based navigation system was evaluated to be better than two millimeters for pedicle screw placement [42, 60], making it superior to conventional approaches or CT-based navigation procedures, where anatomical landmarks are required to register the patient to its preoperative CT volume set.

For laparoscopic surgery, the target region can be deformed due to the heartbeat and respiratory motion. As shown by Olbrich et al. [136], deformations in the abdominal area caused by the heartbeat are negligible. The rather large respiratory motion of about 1-2 cm [8] can be corrected for by gating [29, 66]. As expiration and inspiration plateaus are reproducible within about 1 mm under active breathing control [202], but also under normal breathing [8], they can be synchronized to e.g. an augmented visualization [136]. Nicolau et al. currently also investigate on respiratory motion correction, considering either gating or deformable registration [133]. Up to now, they use rigidly registered preoperative CT data and a tracked needle for the guidance of radio-frequency tumor ablation, where no pneumoperitoneum is applied. Their achieved average accuracy for tumor localization was 9.5 mm. They also presented initial experiments on a rigid abdominal phantom, where they applied their system to laparoscopic surgery [131].

Individual deformations of greater extent mainly occur between preoperative acquisi-

---

[6] e.g. by Aesculap, BrainLAB, Medtronic, ORTHOsoft, PI Systems, Praxim Medivision, and Stryker

[7] In detail, the SurgiGATE® system by Medivision is used for:

- Spine: Pedicle screw placement, decompression of the spinal canal, control of achieved reposition, spinal tumor resection
- Pelvis: Minimally invasive percutaneous placement of SI-screws (sacro-iliacal screws), minimally invasive acetabular reconstruction
- Hip: Screw osteosynthesis of femoral neck fractures
- Knee: Minimally invasive reconstruction of tibia plateau fractures, screw placement and control of reduction
- Ankle: Retrograde drilling (core decompression) in osteochondrosis dissecans tali (OD 2°-3°)

tion of the CT and the beginning of the resection, i.e. during patient and port placement, appliance of CO_2 pneumoperitoneum, and the intervention itself. Pneumoperitoneum alone can already cause large liver motions of e.g. 1.8 ± 12, 4.1 ± 6.4, and 0.1 ± 0.4 mm in x, y, and z directions, respectively, as shown for two pigs by Herline et al. [66]. In this case, using preoperative rigid imaging data to support the surgeon in updating the surgical resection planning is difficult to perform and hard to validate. For robot assisted coronary artery bypass, Mourgues et al. therefore proposed an intelligent way to intraoperatively update the model of a preoperative coronary tree [121]. Interactively, the surgeon identifies and marks visual clues in the endoscope images, so an algorithm can estimate a better intraoperative registration of the coronary tree model. In vivo experiments showed an accuracy of about 9.3 to 19.2 mm [43].

Several attempts were made to use intraoperative imaging to achieve a higher guidance accuracy. In general, standard MR scanners are too bulky to be used during laparoscopic surgery or require the patient to be moved for the acquisition, making a precise intraoperative registration almost impossible. Fichtinger et al. developed an inventive intraoperative CT image overlay system based on a semi-transparent mirror for the purpose of needle insertion, where no major deformations are involved [49]. Keeping it simple and inexpensive, only a single 2D CT slice is shown, which is sufficient for “in-plane” procedures such as needle placement. It is difficult to apply their system to laparoscopic vessel augmentation, where volumetric 3D data is essential. A promising alternative is however the use of supplementary laparoscopic ultrasound, as described in section 2.4.2.2.

Contribution To provide registered high-resolution 3D data supplementary to laparoscopic ultrasound, this dissertation proposes to use a tracked mobile isocentric C-arm providing cone-beam CT imaging capability to visualize contrasted liver vessels intraoperatively and co-align them with the images of the laparoscope camera. An optical tracking system determines the pose of both C-arm and laparoscope. Their acquired imaging data can be brought into the same tracking coordinate system by various offline calibration routines, as described in section 4.1. This makes the intraoperative soft tissue visualization of the proposed system registration-free.

Intraoperatively, after port and trocar[8] placement and application of CO_2 pneumoperitoneum, the vessel tree of the liver is contrasted, similarily to Beldi et al., who contrasted and reconstructed the biliary tree of the liver with a commercially available, image-intensifier based mobile C-arm [12]. At the same time as contrast agent administration, an image series is acquired during patient exhalation. Alternatively, C-arm projections could be gated and correlated to respiratory motion in order to acquire a high-quality scan, as Kriminski et al. suggest [83]. After reconstruction, the contrasted vessel tree can be precisely augmented directly on the laparoscopic view just before the beginning of the resection without any time-consuming patient registration process. The augmentation could be synchronized to the patient’s respiration and only be displayed during exhalation [136]. This provides the surgeon with valuable information on the location of veins, arteries, and bile ducts, which supply the liver segment to be resected and which therefore need to be divided. In general, the augmented visualization will only be shown to the

---

[8] All trocars are made of plastic, so they do not give artifacts in the cone beam reconstruction.

surgeon for the intraoperative in-situ planning of the resection to provide a detailed “road map” of the vessels, but not any more when the surgeon starts to cut, since this causes the liver to deform again and invalidates any prior intrinsic registration. Only if crucial problems appear, another image series may be acquired and an intrinsically registered volume may be reconstructed.

2.4.2.2 Laparoscopic Ultrasound Augmentation

Ultrasonography is an appealing technology to surgeons because of its noninvasiveness, wide availability, flexible handling, and low cost. Having been used primarily for diagnosis in the past, intraoperative use of ultrasound (IOUS) and laparoscopic ultrasound (LUS) nowadays plays an increasing role in abdominal surgery. Liver, biliary tract, and pancreas are main application areas of IOUS and LUS, for instance to detect liver lesions such as metastases. Unfortunately LUS is operator dependent; especially for novice surgeons it is often difficult or even impossible to perform laparoscopic ultrasonography [73]. Among others, the major reasons given for this are the missing tactile feedback, the difficulty to interpret LUS images, a limited degree of positioning through the trocar access, disorientations caused by the constantly changing imaging plane, and a lack of awareness of the transducer tip location (the tip needs to be constantly observed in the laparoscopic camera images in order to avoid inadvertent injury) [56, 73, 140].

Related Work Several groups tried to address some of these issues by providing navigated LUS: The pose of the ultrasound transducer is estimated, so its body and B-scan images can be visualized in relation to the patient, other surgical instruments, or preoperative and intraoperative imaging data. This may greatly support surgeons utilizing LUS in cancer staging, radio frequency ablation, and other procedures.

Ellsmere et al. propose an advanced system to intuitively display the laparoscopic US image plane relatively to a preoperative 3D model of the patient [38, 39]. This helps the physician to identify anatomical key structures and to learn the use of laparoscopic ultrasound. Another approach to improve the spatial relation of US images to the patient is taken by Leven et al. [94]. They propose a system to apprehensively overlay the laparoscopic ultrasound image plane or a reconstructed US volume, respectively, directly on the live images of a stereo endoscope. A point of criticism by the surgeons evaluating their system was the use of a rigid probe, which certain target regions could not be reached with.

To estimate the pose of a transducer with a rigid tip, a robot or optical tracking may be used [94]. In the latter case, a rigid body can be attached to the transducer handle to assure its continuous visibility. Several groups also try to localize rigid laparoscopic instruments in laparoscopic images by advanced image processing techniques, such as Voros et al. [187]. However, laparoscopic transducers most commonly used and preferred by surgeons feature a flexible tip providing rightward, leftward, forward, and backward steering. The tip also yields to external pressure from organ surfaces. Due to the missing line of sight to the flexible transducer tip, an optical tracking system cannot be used exclusively to localize this tip. A robot could only be utilized, if the ultrasound probe was fully integrated into the end-effector. To the author's knowledge, no such system currently exists. Promising alternatives are the use of an electromagnetic tracking sensor attached to the tip [39, 41, 79, 84] or fully incorporated into the tip [62], or magneto-optic tracking, i.e. the combination of optical tracking and electromagnetic tracking [81, 127].

When clinically using electromagnetic tracking, a considerable problem is the distortion of the electromagnetic field leading to erroneous tracking data. This distortion can be caused by metallic or electrically powered objects inside or in close vicinity to the working volume, for instance surgical instruments, an operating table, or imaging devices such as a C-arm or a computed tomography scanner. Depending on the operating room setup and instrumentation, tracking errors of several millimeters or even centimeters can occur [71, 125]. To compensate for erroneous measurements caused by stationary objects, various calibration techniques were proposed [77]. They usually require the user to acquire a set of well distributed measurements within the electromagnetic tracking volume. This set is compared to a set of reference measurements to compute a field distortion function that is based on look-up tables or polynomials. Unfortunately, this function can only compensate static errors of non-moving distortion fields, so that the calibration process has to be repeated for every new operating room setup before an intervention. Dynamic changes of the field distortion, for example caused by the intraoperative relocation of the electromagnetic transmitter or movement of instruments, cannot be compensated by the previously computed distortion functions. A first step towards the intraoperative detection of erroneous measurements caused by metallic objects distorting the field was presented by Birkfellner et al. [15, 16] and later on Mucha et al. [123]. They all incorporate two sensors into a pointer, so redundant measurements can be obtained. Deviations of the fixed distance between the two sensors are used as a plausibility value.

Contribution This dissertation introduces a new method to detect field distortions online [47], i.e. intraoperatively without a pre-computed distortion function. It is applied to a flexible laparoscopic ultrasound transducer, whose pose is determined by a magneto-optic tracking system, based on two electromagnetic sensors attached to the flexible and the rigid part of the transducer and another optical tracking body attached to the rigid part, so tracking redundancy between the sensor and the body on the rigid transducer part can be achieved. As optical tracking data is not affected by electromagnetic field distortions, distorted measurements of the rigid sensor can be detected. In this case, the surgical staff can be warned immediately.

Furthermore, the B-scan images of the transducer are overlaid on the live images of an optically tracked laparoscope in real time without intraoperative registration. This provides surgeons with a better understanding of the spatial relationship between the two imaging modalities. This overlay however may be inaccurate, as tracking and calibration errors are propagated, which can be even increased by the effect of electromagnetic field distortions. Based on a mathematical model of all possible transducer tip movements relatively to the optical tracking body, misalignments of the overlay can be partially corrected.

Tracked laparoscopic ultrasound, the mobile C-arm providing high-resolution 3D data, and the laparoscope providing in situ live images, together form a strong triplet for image-guided minimally invasive abdominal surgery.

CHAPTER

THREE

SYSTEM COMPONENTS

The laparoscope augmentation system proposed here depends on a set of essential software and hardware components, but also provides optional modules, which can be used depending on the necessity and availability of medical imaging equipment in the operating room and the kind of intervention. Currently not every clinic has a mobile C-arm capable of 3D cone-beam CT reconstructions, but intraoperative ultrasound is widely available.

This chapter shortly describes required tracking and imaging technology needed for the laparoscope augmentation system.

3.1 General Hardware

The augmented reality visualization software (see also section 4.4) runs on a standard workstation PC including two analog frame grabbers (FALCON, IDS Imaging Development Systems, Obersulm, Germany)[1] for capturing the videos of both ultrasound and laparoscope camera in real time. The workstation can be connected via Ethernet to the C-arm system and the PC required for the optical tracking system, and via USB to the electromagnetic tracking system.

Whenever the mobile C-arm is used intraoperatively, the operating table should be made of carbon in order to limit imaging artifacts during surgery. For electromagnetic tracking, a carbon or metal-free table is of great help, too, so static field distortions can be avoided.

3.2 Optical Tracking

An essential component of the laparoscope augmentation system is the optical tracking system. It is required to determine the pose of all intraoperatively utilized imaging devices, i.e. laparoscope, C-arm, laparoscopic ultrasound, and, optionally, other surgical instruments. The tracking system used throughout all experiments is made by A.R.T.

---

[1] `http://www.ids-imaging.com/`

GmbH, Weilheim, Germany. For both laboratory and operating room setups, four ARTtrack2 cameras are mounted on the ceiling, one in each corner of a rectangle, so the surgical staff occludes the line of sight of the cameras as little as possible and the tracking error distribution can be kept low [9].

Tracking bodies consisting of several retroreflective spherical markers are attached to all imaging devices. The markers can be segmented well in the 2D images of the tracking cameras, as the measurement volume is additionally illuminated by an infrared light flash for every measurement cycle. The segmented 2D markers can be triangulated in space to reconstruct their 3D positions [44, 63, 176] (see also section 4.2). If at least three non-collinear markers are combined to a tracking body, six degrees of freedom (6 DOF), i.e. the full pose of the body can be computed. The DTrack software of A.R.T. running on a book size PC does these computations and sends the tracking data via Ethernet to the visualization workstation.

The root mean squared (RMS) measurement errors of the optical tracking system are stated as 0.4 mm (position) and 0.12° (orientation) by the manufacturer[2].

3.3 Electromagnetic Tracking

To track the flexible tip of the laparoscopic ultrasound transducer, electromagnetic tracking is needed in addition due to the missing line of sight of the flexible tip to the optical tracking cameras. Therefore, the 3D Guidance unit of Ascension Technology Corporation, Burlington, VT, USA, was employed.

A mid-range transmitter with three orthogonal windings sequentially creates magnetic fields along its x, y, and z axis. Sensors with coils along three axes measure the transmitted field vectors at a certain point in space, relatively to the transmitter. The 3D Guidance unit processes the sensor signals and sends 6 DOF tracking data via USB to the visualization workstation. The static RMS errors of the electromagnetic tracking system are stated as 1.4 mm (position) and 0.5° (orientation).[3]

To co-register both optical and electromagnetic tracking, the transmitter is also equipped with an optical tracking body ("transmitter body"). In this way, the transmitter can be moved easily within the optical tracking volume, while all electromagnetic measurements can be transformed into the coordinate system of the optical tracking system.

3.4 Laparoscope

The proposed augmented reality system can be applied to almost any laparoscope commonly used in the operating room. Laparoscopes usually provide a magnified forward

---

[2] Specifications of the typical accuracy of the A.R.T. tracking system are available under `http://www.ar-tracking.de/`.

[3] Specifications of the typical accuracy of the 3D Guidance unit are available under `http://www.ascension-tech.com/`.

oblique view of the surgical site, which is achieved on the one hand by a negative element (a plano-concave field-widening lens) at the tip of the laparoscope, which reduces the inclination of incident rays, and on the other hand by a prism, which redirects the field of view to the side. Following are the objective lenses, an array of relay lenses, and the eyepiece, from where the rays are led to the image sensor of the camera, which often is a charge-coupled device (CCD), i.e. an array of linked light-sensitive capacitors. To illuminate the surgical site, white light is emitted from a high-intensity xenon, mercury, or halogen lamp and transmitted through fiber optic light guide bundles.

More technical details on the design, construction, and optics of laparoscopes can be found for instance in the papers of Bobbart et al. [20], Leiner [92], and Miller and Hollingsworth [115].

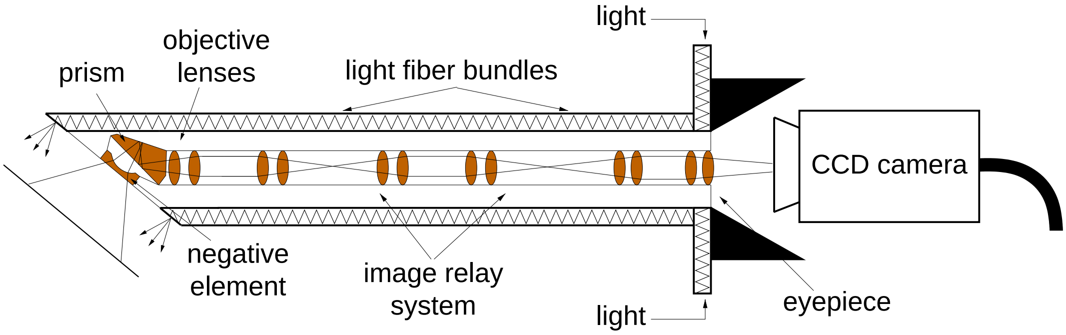


Figure 3.1: Technical details of the design and optics of a laparoscope (Image adapted from Boppart et al. [20] and Vogt [184]).

The laparoscopes used throughout all experiments are made by KARL STORZ GmbH & Co. KG, Tuttlingen, Germany[4]. They have a rigid forward-oblique 30° HOPKINS telescope, which is typically used for abdominal surgery. Depending on their availability for experiments, either an NTSC-based TELECAM® one-chip camera or a PAL-based TRICAM® three-chip camera was used, which both provide analog video images. To further process the images by the visualization software, one of the two frame grabbers incorporated into the visualization workstation is used for real-time capturing.

To track the pose of the laparoscope, the optical tracking system is employed. In a first and second iteration, wooden and aluminum blocks were manufactured, which rigidly connect the laparoscope shaft and the camera head (cf. figures 5.3 and 3.2a,b). This simplifies the calibration routine, as only one tracking body has to be attached to the laparoscope and hence only one rigid offset from this single body to the camera center needs to be computed (see section 4.1.1). However, this also prevents the laparoscope shaft from being freely rotated against the camera head, which is not very practical during surgery, as surgeons are used to keep the camera head and hence the image "horizon" fixed and only rotate the cylindric shaft. Therefore, in another iteration, a tracking body consisting of four retroreflective spherical markers was each attached to the camera head ("laparoscope body") and to the cylindric shaft of the laparoscope (cf. figure 3.2c).

---

[4] `http://www.karlstorz.com/`

(a) Initial sub-optimal marker configuration.

(b) Rearranged markers to ensure optimal tracking properties.

(c) Two optical tracking bodies to determine the laparoscope rotation.

Figure 3.2: Several developed laparoscope marker body configurations – according to the surgeons they are not disadvantageous for the laparoscope handling and comfort.

3.5 Mobile C-arm

To acquire high-resolution 3D data intraoperatively, an optically tracked mobile C-arm is utilized. The prototype mobile C-arm used during all experiments is based on a Siemens PowerMobil and incorporates a workstation deployed with acquisition and reconstruction software by Siemens Medical. The visualization workstation is connected to the C-arm workstation via Ethernet to directly access any reconstructed volumes. Four retroreflective spherical markers are attached to the flat-panel detector of the C-arm (cf. figure 3.3).

For 3D cone-beam CT reconstruction, the C-arm orbits around its isocenter. Synchronizing the X-ray tube emissions and the flat-panel detector readouts, a predefined set of X-ray projections is acquired under continuous rotation, where the X-ray tube is modeled as camera center of a pinhole camera and the flat-panel detector as image plane (cf. figure 4.1b). The acquisition of 200 C-arm projections takes 64 seconds, which stays within the time limits of holding breath during active breathing control. Alternatively 100 projections can be acquired, which only takes 32 seconds. The reconstruction of a 3D volume currently takes about six minutes for 200 projections or three minutes for 100 projections, respectively. The reconstruction software is uncoupled from the overall software package to allow better control and debugging of the reconstruction process. Once the prototype system is commercialized, the reconstruction algorithms will be further optimized and parallelized. Additionally, reconstruction will commence as soon as the first projection is acquired. This will lead to reconstruction times of approximately one minute after completion of a scan, as known from modern flat panel based angiographic computed tomography (ACT) scanners such as DynaCT, where a volume set is available for assessment in less than three minutes[5]. The reconstruction software generates a 16 bit gray level volume of 512x512x384 voxels.

Further details on the C-arm system are given e.g. by Siewerdsen et al. [162].

---

[5] `http://healthcare.siemens.com/dynact/`

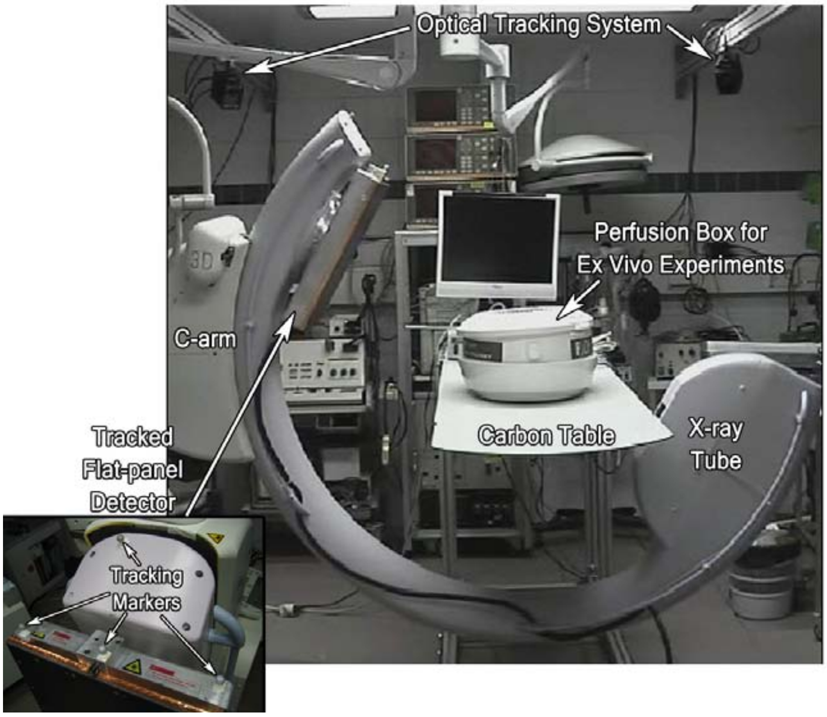


Figure 3.3: C-arm setup

3.6 Laparoscopic Ultrasound

To support the physician with ultrasound images, which are co-registered and aligned with the laparoscopic view, any standard ultrasound system outputting digital or analog 2D B-scan images may be used. Systems such as from Terason Ultrasound[6] or Ultrasonix Medical Corporation[7], which provide application programming or research interfaces and hence access to digital ultrasound images and the acquisition time stamps, are favored. However, currently they do not offer a laparoscopic transducer. Throughout all experiments made for this dissertation, a SONOLINE Omnia US system by Siemens Medical Solutions (Mountain View, CA, USA)[8] was used, which a flexible laparoscopic linear array transducer (LAP8-4, 5 MHz, 10 mm diameter) is connected to. The transducer features a flexible tip providing rightward, leftward, forward, and backward steering, which can be controlled by two steering levers. The tip also yields to external pressure, e.g. from organ surfaces.

Sound waves produced by the transducer are partially reflected inside the human body, whenever density changes occur, i.e. between different tissues. The reflected sound waves, which return to the transducer, usually have different directions, magnitudes, and durations. Based on the information on these differences, grayscale B-scan images can

---

[6] `http://www.terason.com/`

[7] `http://www.ultrasonix.com/`

[8] `http://www.medical.siemens.com/`

be generated by the ultrasound system. The analog output of the ultrasound system containing the B-scan images is connected to a frame grabber inside the visualization workstation.

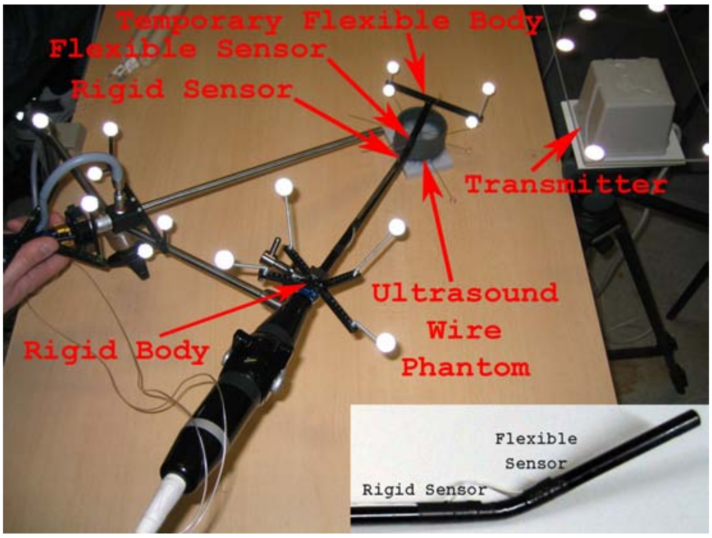


Figure 3.4: Ultrasound transducer setup.

In addition to an optical tracking body, which is attached to the transducer handle (below referred to as “rigid body”), two electromagnetic sensors are attached to the transducer shaft: One to the flexible tip (“flexible sensor”), the other one to the rigid part (“rigid sensor”), as close to each other as possible. As mentioned previously, another optical tracking body is mounted on the electromagnetic transmitter. This setup allows a co-calibration of electromagnetic tracking and optical tracking and a provision of redundant tracking information of the rigid part of the transducer shaft, since the rigid sensor theoretically has a constant offset to the optical tracking body. This is important to estimate electromagnetic tracking errors. For accuracy evaluation experiments, a temporary optical tracking body was also glued[9] to the flexible transducer tip (“flexible body”), which was removed after the experiments.

---

[9] Particularly suitable for (temporarily) adhering tracking bodies and sensors to surgical instruments are low-temperature hot glue guns of about 110°C such as UHU LT 110 (UHU GmbH & Co. KG, Bühl/Baden, Germany), as e.g. the specifications of the Ascension sensors state a maximum applicable temperature of 150°C.

CHAPTER

FOUR

METHODS

BESIDES the experimental evaluation of the introduced augmented reality solution, most work has been put into the design and development of methods, which enable and support the presented concepts. This chapter describes in detail, which offline calibration steps are required for the system to work, and introduces all registration, tracking, and visualization concepts that are combined to provide a powerful intraoperative augmented reality visualization based on multimodal imaging.

It must be noted that throughout all presented methods a few standard mathematical techniques are heavily used, for instance matrix and (dual) quaternion arithmetics, matrix decompositions such as singular value decomposition (SVD) and QR decomposition, and optimizers such as the Levenberg-Marquardt algorithm. These techniques are not described in detail here, but can be found in various textbooks.

4.1 System Calibration

In order to augment preoperative or intraoperative imaging data on the laparoscopic view, data represented in different coordinate systems (cf. figures 4.1a, 4.7, and 4.8) need to be brought into one common world coordinate frame. In the following, the coordinate frame of the optical tracking system is used synonymously to the world coordinate frame, as it is the coordinate frame common to all tracked imaging devices. While preoperative imaging data required for port placement needs to be manually registered to the patient (see section 4.2), all intraoperative imaging devices can be calibrated offline to automatically provide a registration-free intraoperative alignment of patient and data. Following main transformations have to be computed to bring the local coordinate frames of the intraoperative imaging devices into the world coordinate frame of the optical tracking system:

1. The transformation from the camera center to the tracking body attached to the camera head of the laparoscope along with the projection geometry of the laparoscope camera (cf. section 4.1.1).

2. The transformation from the isocenter of the C-arm to the tracking body attached to the flat-panel detector (cf. section 4.1.3).
3. The transformation from the ultrasound plane coordinates to the electromagnetic sensor attached to the flexible transducer tip (cf. section 4.1.4). Additionally, for the electromagnetic error estimation method the transformation from the sensor on the rigid shaft to the optical tracking body needs to be computed (cf. section 4.1.4.1).
4. The transformation from electromagnetic tracking coordinates into optical tracking coordinates (cf. section 4.1.4.1).

All calibration steps can be done offline, i.e. before the intervention or while manufacturing. As long as the tracking bodies and sensors are not repositioned, calibration results can be valid for a long period of time and only need to be repeated once in a while.

Repositioning Repositioning however is an important issue, especially during sterilization. Practical commercial solutions already exist for the intraoperative use of markers attached to a mobile C-arm, as for the system that is actively used in the operating room of our clinical partners (see section 2.4.2.1). For the sterilization of the laparoscope and its attached tracking bodies, autoclavable frames could be manufactured, which can be removed before sterilization and precisely reattached to the laparoscope afterwards. Currently frames manufactured from laser-range scanning data are used, which precisely fit the head and shaft of the laparoscope. Screws firmly fix the markers as well as the frames. All screw connections should however be replaced by snap-on connections in the future, as offered for instance by NDI for their passive markers. Sterilizable markers are also commercially available (e.g. IZI Spherz, IZI Medical Products, Baltimore, MD, USA[1], or NDI Passive Spheres™, Scanlan International, Saint Paul, MN, USA[2]). The marker body attached to the laparoscopic ultrasound transducer can be manufactured in a similar way. Additionally required electromagnetic sensors could also be integrated into the transducer tip [62] and shaft.

On-site Verification It is easy to setup on-site accuracy verification procedures to monitor the validity of all estimated parameters simply by imaging and superimposing volumetric data of an optically tracked phantom with known geometry on the laparoscopic view. This verification can be performed by a nurse or by the surgical staff in the operating room before patient anesthesia. In case of unsatisfactory validation results a system recalibration may be performed on site, similarly to the approach of Falk et al., where an interactive calibration of an endoscope was performed in three to eight minutes [43], or Vogt, who approximates one minute for his intraoperative endoscope calibration procedure [184].

4.1.1 Laparoscope Camera

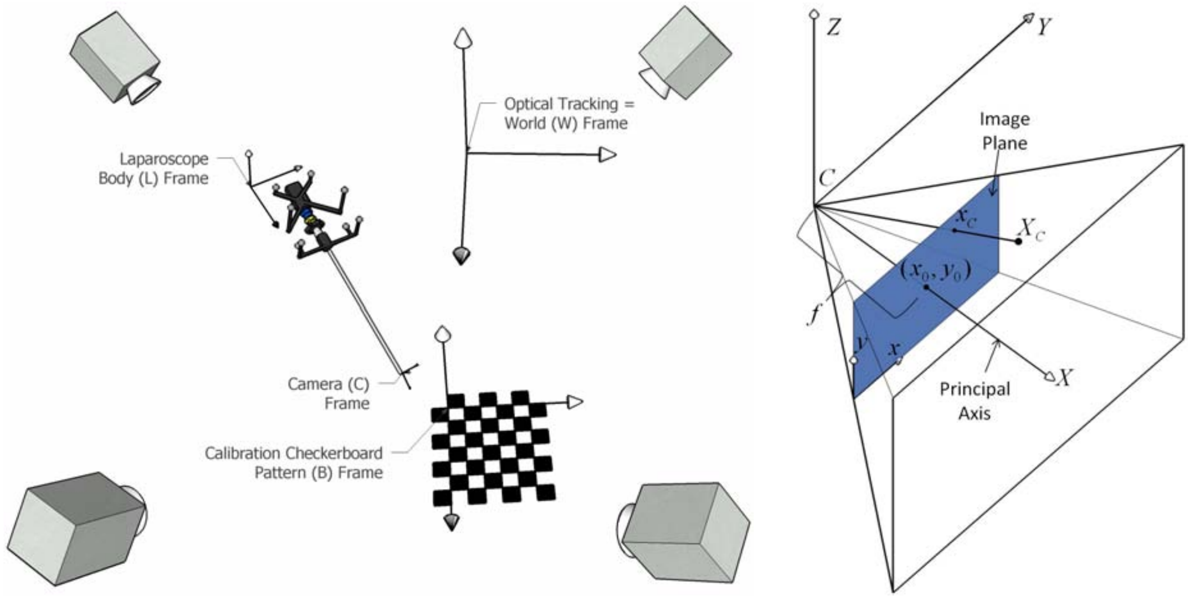
The main purpose of calibrating the laparoscope is to model the transformation of a 3D world point onto the 2D image plane of the camera, so the projection of any 3D world

---

[1] `http://www.izimed.com/`

[2] `http://www.scanlaninternational.com/`

scene onto the laparoscope image can be described mathematically, which generally is done by a pinhole camera model.



(a) Coordinate frames for camera calibration. (b) Pinhole camera model.

Figure 4.1: Coordinate frames and camera model

In detail, a 3D point $\boldsymbol{X}_W$ in the world frame is first transformed into the laparoscope body frame by ${}^{L}\boldsymbol{T}_W$, from where it is transformed into the camera frame by ${}^{C}\boldsymbol{T}_L$ (see also figure 4.1a):

$$\begin{bmatrix} \boldsymbol{X}_C \\ 1 \end{bmatrix} = \left({}^{C}\boldsymbol{T}_L\right)\left({}^{L}\boldsymbol{T}_W\right)\begin{bmatrix} \boldsymbol{X}_W \\ 1 \end{bmatrix} \tag{4.1}$$

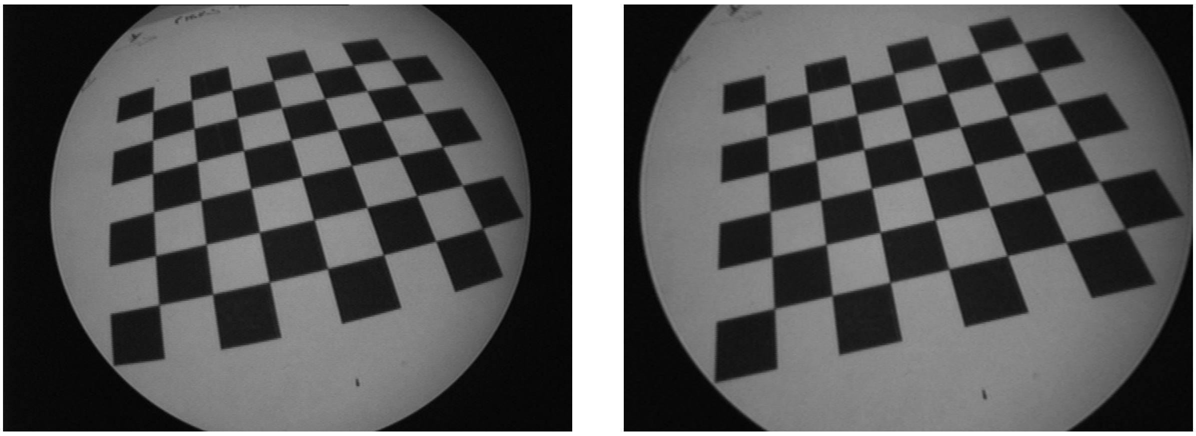
Finally, it is mapped onto the image plane by the camera calibration matrix $\boldsymbol{K}$ (up to scale λ). $\boldsymbol{x}_C = \begin{pmatrix} x_C \\ y_C \end{pmatrix}$ is the final 2D point projected onto the image plane:

$$\lambda \begin{bmatrix} \boldsymbol{x}_C \\ 1 \end{bmatrix} = \boldsymbol{K}\boldsymbol{X}_C \text{ , where } \boldsymbol{K} = \begin{bmatrix} \alpha_x & c & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.2}$$

As illustrated in figure 4.1b, the principal axis is orthogonal to the image plane. Their common intersection point defines the principal point. The coordinates (x_0, y_0) of the principal point, the parameter c describing the skewness of the two image axes x and y, and the scaled focal length (α_x, α_y) in image axes x and y are the intrinsic camera parameters, all represented in pixel coordinates. f in figure 4.1b also refers to the focal length, but represented in millimeters. To scale f to pixel space, it can be divided by the length of a single element (pixel) of the CCD of the camera. As such elements are not necessarily square, both their width s_{width} and height s_{height} in millimeters needs to be considered. The focal length in pixels therefore is given by $\alpha_x = f/s_{width}$ and

$\alpha_y = f/s_{height}$. The skew parameter c is always 0 for CCD and CMOS (complementary metal-oxide-semiconductor) sensors, as it is the case for laparoscope cameras, i.e. they have square pixels. Only when taking an image from an image, which can lead to non coinciding principal axes, c may be unequal to zero.

The transformation ${}^{L}\boldsymbol{T}_W$ can be directly received from the optical tracking system. (Note that tracking systems usually provide the inverse transformation of ${}^{L}\boldsymbol{T}_W$, i.e. the transformation ${}^{W}\boldsymbol{T}_L$ from the tracking body or sensor to the reference tracking coordinate system.) The transformation ${}^{C}\boldsymbol{T}_L$ from laparoscope body to camera coordinate frame and the intrinsic camera parameters stored in K need to be computed once [154, 164]. Additionally, the rather large radial and tangential lens distortion of laparoscopes needs to be corrected for. While being required for a perfect augmentation, undistorted images are also preferred by surgeons [185].



(a) Original distorted image.

(b) Undistorted image.

Figure 4.2: Calibration pattern as seen from the laparoscope.

To determine all these parameters, a planar checkerboard pattern with 8x7 squares of 10 mm unit length is used (see figure 4.2). The pattern is viewed from several arbitrary poses $i = 1 \dots m$, $m > 2$ satisfying distinct rotation axes for all possible motions between them (distinct rotation axes are required for hand-eye calibration; see next section). It must be ensured that the laparoscope shaft is not rotated against the camera head during these movements. For each pose, an image is recorded along with the 6D pose data of the laparoscope body provided by the optical tracking system. In each image, the inner corners of the checkerboard pattern are found automatically by utilizing methods of the Open Source Computer Vision Library (OpenCV)[3]. This gives a set of 42 sorted 2D points[4], which sub-pixel accuracy is guaranteed for. Knowing their corresponding 3D position on the pattern ($\boldsymbol{X}_B = (X_B, Y_B, 0)^T$, $X_B = 0, 10, \dots, 60$, $Y_B = 0, 10, \dots, 50$), the intrinsic camera parameters and distortion coefficients can be computed using well-established camera calibration techniques [64, 177, 207].

---

[3] `http://www.intel.com/technology/computing/opencv/`

[4] A checkerboard pattern with 8x7 squares has exactly $7 \cdot 6 = 42$ inner corners.

4.1.1.1 Camera Calibration

The camera calibration implementation, which is used for this thesis, can be found in OpenCV. As its concepts are largely borrowed from the method of Zhang [207], his ideas are presented here shortly to give an insight into camera calibration.

The objective of camera calibration is to estimate all parameters in ${}^C\boldsymbol{T}_B$ and K needed to transform and project a 3D point $\boldsymbol{X}_B$ on the planar pattern into the image plane:

$$\begin{aligned}\lambda\begin{bmatrix}\boldsymbol{x}_C\\1\end{bmatrix} &= \boldsymbol{K}\begin{bmatrix}{}^C\boldsymbol{R}_B & {}^C\boldsymbol{t}_B\end{bmatrix}\begin{bmatrix}\boldsymbol{X}_B\\1\end{bmatrix}\\ &= \boldsymbol{K}\begin{bmatrix}{}^C\boldsymbol{r}_{B1} & {}^C\boldsymbol{r}_{B2} & {}^C\boldsymbol{r}_{B3} & {}^C\boldsymbol{t}_B\end{bmatrix}\begin{bmatrix}X_B\\Y_B\\0\\1\end{bmatrix}\\ &= \boldsymbol{K}\begin{bmatrix}{}^C\boldsymbol{r}_{B1} & {}^C\boldsymbol{r}_{B2} & {}^C\boldsymbol{t}_B\end{bmatrix}\begin{bmatrix}X_B\\Y_B\\1\end{bmatrix}\end{aligned} \tag{4.3}$$

where ${}^C\boldsymbol{r}_{Bi}$ denotes the ith column of the rotation matrix ${}^C\boldsymbol{R}_B$. By defining $\tilde{\boldsymbol{m}} = [\boldsymbol{x}_C, 1]^T$, $\boldsymbol{H} = \boldsymbol{K}\begin{bmatrix}{}^C\boldsymbol{r}_{B1} & {}^C\boldsymbol{r}_{B2} & {}^C\boldsymbol{t}_B\end{bmatrix}$, and $\tilde{\boldsymbol{M}} = [X_B, Y_B, 1]^T$, above equation can be rewritten to:

$$\lambda\tilde{\boldsymbol{m}} = \boldsymbol{H}\tilde{\boldsymbol{M}} \tag{4.4}$$

which states that a pattern point $\tilde{\boldsymbol{M}}$ is related to its image point $\tilde{\boldsymbol{m}}$ by the homography $\boldsymbol{H}$ (up to scale λ).

Homography Estimation To estimate $\boldsymbol{H}$ for $j = 1 \ldots n$ point correspondences $\tilde{\boldsymbol{m}}_{(j)}$ and $\tilde{\boldsymbol{M}}_{(j)}$, which satisfy equation 4.4, Zhang uses a technique based on the maximum likelihood criterion. As the image points $\tilde{\boldsymbol{m}}_{(j)}$ are usually noisy, Gaussian noise with mean $\boldsymbol{0}$ and covariance matrix $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{m}}_j}$ can be assumed, so $\boldsymbol{H}$ can be determined by minimization of

$$\sum_{j=i}^{n}\left(\tilde{\boldsymbol{m}}_{(j)} - \hat{\boldsymbol{m}}_{(j)}\right)^T \boldsymbol{\Sigma}_{\tilde{\boldsymbol{m}}_{(j)}}{}^{-1}\left(\tilde{\boldsymbol{m}}_{(j)} - \hat{\boldsymbol{m}}_{(j)}\right) \tag{4.5}$$

where $\hat{\boldsymbol{m}}_{(j)} = \frac{1}{\bar{\boldsymbol{h}}_3^T}\begin{bmatrix}\bar{\boldsymbol{h}}_1^T\tilde{\boldsymbol{M}}_{(j)}\\ \bar{\boldsymbol{h}}_2^T\tilde{\boldsymbol{M}}_{(j)}\end{bmatrix}$ and $\bar{\boldsymbol{h}}_i$ is the ith row of $\boldsymbol{H}$.

Assuming $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{m}}_{(j)}} = \sigma^2\boldsymbol{I}$ for all j, equation 4.5 can be simplified to

$$\sum_{j=1}^{n}\left\|\tilde{\boldsymbol{m}}_{(j)} - \hat{\boldsymbol{m}}_{(j)}\right\|^2 \tag{4.6}$$

which is a non-linear least-squares problem and can be minimized by Levenberg-Marquardt. To provide an initialization for the optimization, equation 4.4 is rewritten as

$$\begin{bmatrix}\tilde{\boldsymbol{M}}^T\,\bar{\boldsymbol{h}}_1^T\\ \tilde{\boldsymbol{M}}^T\,\bar{\boldsymbol{h}}_2^T\\ \tilde{\boldsymbol{M}}^T\,\bar{\boldsymbol{h}}_3^T\end{bmatrix} - \lambda\begin{bmatrix}x_C\\y_C\\1\end{bmatrix} = \boldsymbol{0}$$

To satisfy this equation, λ needs to be equal to $\tilde{\boldsymbol{M}}^T \ \bar{\boldsymbol{h}}_3^T$. For all $j = 1 \dots n$ point correspondences, this gives

$$\underbrace{\begin{bmatrix} \tilde{\boldsymbol{M}}_{(1)}^T & \boldsymbol{0}^T & -x_{C(1)}\tilde{\boldsymbol{M}}_{(1)}^T \\ \boldsymbol{0}^T & \tilde{\boldsymbol{M}}_{(1)}^T & -y_{C(1)}\tilde{\boldsymbol{M}}_{(1)}^T \\ \tilde{\boldsymbol{M}}_{(2)}^T & \boldsymbol{0}^T & -x_{C(2)}\tilde{\boldsymbol{M}}_{(2)}^T \\ \boldsymbol{0}^T & \tilde{\boldsymbol{M}}_{(2)}^T & -y_{C(2)}\tilde{\boldsymbol{M}}_{(2)}^T \\ & \vdots & \\ \tilde{\boldsymbol{M}}_{(n)}^T & \boldsymbol{0}^T & -x_{C(n)}\tilde{\boldsymbol{M}}_{(n)}^T \\ \boldsymbol{0}^T & \tilde{\boldsymbol{M}}_{(n)}^T & -y_{C(n)}\tilde{\boldsymbol{M}}_{(n)}^T \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} \bar{\boldsymbol{h}}_1^T \\ \bar{\boldsymbol{h}}_2^T \\ \bar{\boldsymbol{h}}_3^T \end{bmatrix}}_{\boldsymbol{f}} = \boldsymbol{0} \tag{4.7}$$

As the 9D vector $\boldsymbol{f}$ is defined up to scale, at least four point pairs $\tilde{\boldsymbol{m}}_{(j)}$ and $\tilde{\boldsymbol{M}}_{(j)}$ are needed to solve the eight degrees of freedom of equation 4.7. The solution for $\boldsymbol{f}$ is provided e.g. by a singular value decomposition of $\boldsymbol{A}$ into $\boldsymbol{UDV}^T$, where $\boldsymbol{f}$ is set to the right singular vector (a row vector) in $\boldsymbol{V}^T$ associated with the smallest singular value in $\boldsymbol{D}$. Because $\boldsymbol{A}$ contains a mixture of constant elements, coordinates represented in millimeters, and products of millimeters and pixels, it is poorly conditioned. Therefore, an additional prior data normalization is recommended, as described by Hartley and Zisserman for their “gold standard algorithm” [63].

Closed-form Solution After the estimation of the homography $\boldsymbol{H}$, it still needs to be decomposed into its intrinsic camera matrix $\boldsymbol{K}$ and extrinsic parameters of the transformation ${}^C\boldsymbol{T}_B$. Therefore, equation 4.4 is rewritten as

$$\begin{bmatrix} \boldsymbol{h}_1 & \boldsymbol{h}_2 & \boldsymbol{h}_3 \end{bmatrix} = s\boldsymbol{K} \begin{bmatrix} {}^C\boldsymbol{r}_{B1} & {}^C\boldsymbol{r}_{B2} & {}^C\boldsymbol{t}_B \end{bmatrix}$$

where h_i is the ith column of $\boldsymbol{H}$ (remember that $\boldsymbol{H}$ is defined up to scale, here represented by s). Since ${}^C\boldsymbol{r}_{B1}$ and ${}^C\boldsymbol{r}_{B2}$ are orthonormal, their dot product ${}^C\boldsymbol{r}_{B1}^T \ {}^C\boldsymbol{r}_{B2}$ needs to be zero, i.e. for above equation

$$\boldsymbol{h}_1^T\boldsymbol{K}^{-T}\boldsymbol{K}^{-1}\boldsymbol{h}_2 = 0 \tag{4.8}$$

while ${}^C\boldsymbol{r}_{B1}^T \ {}^C\boldsymbol{r}_{B1} = 1$ and ${}^C\boldsymbol{r}_{B2}^T \ {}^C\boldsymbol{r}_{B2} = 1$, i.e.

$$\boldsymbol{h}_1^T\boldsymbol{K}^{-T}\boldsymbol{K}^{-1}\boldsymbol{h}_1 = \boldsymbol{h}_2^T\boldsymbol{K}^{-T}\boldsymbol{K}^{-1}\boldsymbol{h}_2 \tag{4.9}$$

Next, a symmetric matrix $\boldsymbol{B}$ is defined as

$$\begin{aligned} \boldsymbol{B} = \boldsymbol{K}^{-T}\boldsymbol{K}^{-1} &= \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\alpha_x^2} & -\frac{c}{\alpha_x^2\alpha_y} & \frac{y_0c-x_0\alpha_y}{\alpha_x^2\alpha_y} \\ -\frac{c}{\alpha_x^2\alpha_y} & \frac{c^2}{\alpha_x^2\alpha_y^2}+\frac{1}{\alpha_y^2} & -\frac{c(y_0c-x_0\alpha_y)}{\alpha_x^2\alpha_y^2}-\frac{y_0}{\alpha_y^2} \\ \frac{y_0c-x_0\alpha_y}{\alpha_x^2\alpha_y} & -\frac{c(y_0c-x_0\alpha_y)}{\alpha_x^2\alpha_y^2}-\frac{y_0}{\alpha_y^2} & -\frac{(y_0c-x_0\alpha_y)^2}{\alpha_x^2\alpha_y^2}+\frac{y_0^2}{\alpha_y^2}+1 \end{bmatrix} \end{aligned} \tag{4.10}$$

Defining the 6D vector $\boldsymbol{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T$ and the ith column of $\boldsymbol{H}$ as $\boldsymbol{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$, the relationship between $\boldsymbol{H}$ and $\boldsymbol{B}$ can be reformulated as

$$\boldsymbol{h}_i^T \boldsymbol{B} \boldsymbol{h}_j = \boldsymbol{e}_{ij}^T \boldsymbol{b} \tag{4.11}$$

where

$$\boldsymbol{e}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T$$

This relationship can be applied to equations 4.8 and 4.9:

$$\underbrace{\begin{bmatrix} \boldsymbol{e}_{12}^T \\ (\boldsymbol{e}_{11} - \boldsymbol{e}_{22})^T \end{bmatrix}}_{\boldsymbol{E}_{(i)}} \boldsymbol{b} = \boldsymbol{0} \tag{4.12}$$

where $\boldsymbol{E}_{(i)}$ is one of m matrices obtained from $i = 1 \dots m$ camera poses. $\boldsymbol{b}$ imposes five degrees of freedom (up to scale). As for a laparoscope camera the skew factor c can be assumed to be equal to zero, another equation, namely $[0, 1, 0, 0, 0, 0]\boldsymbol{b} = 0$, can be added to equation 4.12. Hence, even two camera poses would be already enough to solve for $\boldsymbol{b}$, as this additional equation reduces the degrees of freedom to four. However, results can greatly be improved by providing images from more than just two camera poses. Stacking up all matrices $\boldsymbol{E}_{(i)}$ and $[0, 1, 0, 0, 0, 0]$ to $\boldsymbol{E}$, equation 4.12 can be solved e.g. by a singular value decomposition, i.e. $\boldsymbol{E} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$, where $\boldsymbol{b}$ is set to the right singular vector (a row vector) in $\boldsymbol{V}^T$ associated with the smallest singular value in $\boldsymbol{D}$.

The intrinsic camera parameters can now be obtained from $\boldsymbol{b}$ (Again note that c can be set to zero for sensors generally found in laparoscope cameras):

$$\begin{aligned} y_0 &= (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2) \\ s &= B_{33} - [B_{13}^2 + y_0(B_{12}B_{13} - B_{11}B_{23})]/B_{11} \\ \alpha_x &= \sqrt{s/B_{11}} \\ \alpha_y &= \sqrt{sB_{11}/(B_{11}B_{22} - B_{12}^2)} \\ c &= -B_{12}\alpha_x^2\alpha_y/s \overset{!}{=} 0 \\ x_0 &= cy_0/\alpha_y - B_{13}\alpha_x^2/s \overset{c\overset{!}{=}0}{=} -B_{13}\alpha_x^2 \end{aligned} \tag{4.13}$$

Based on the intrinsics, the extrinsic camera parameters can be computed for every camera pose $i = 1 \dots m$ from equation 4.4:

$$\begin{aligned} {}^C\boldsymbol{r}_{B1} &= s\boldsymbol{K}^{-1}\boldsymbol{h}_1 \\ {}^C\boldsymbol{r}_{B2} &= s\boldsymbol{K}^{-1}\boldsymbol{h}_2 \\ {}^C\boldsymbol{r}_{B3} &= {}^C\boldsymbol{r}_{B1} \times {}^C\boldsymbol{r}_{B2} \\ {}^C\boldsymbol{t}_B &= s\boldsymbol{K}^{-1}\boldsymbol{h}_3 \end{aligned} \tag{4.14}$$

where $s = 1/\left\|\boldsymbol{K}^{-1}\boldsymbol{h}_1\right\| = 1/\left\|\boldsymbol{K}^{-1}\boldsymbol{h}_2\right\|$. Due to noise, ${}^C\boldsymbol{R}_B = \left[{}^C\boldsymbol{r}_{B1}{}^C\boldsymbol{r}_{B2}{}^C\boldsymbol{r}_{B3}\right]$ will not be an ideal rotation matrix consisting of three orthonormal vectors. To approximate an optimal rotation matrix ${}^{C}\breve{\boldsymbol{R}}_B$ for ${}^{C}\boldsymbol{R}_B$ provided that ${}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\breve{\boldsymbol{R}}_B = \boldsymbol{I}$, the Frobenius norm of the difference ${}^{C}\breve{\boldsymbol{R}}_B - {}^{C}\boldsymbol{R}_B$ is minimized by

$$\min_{{}^{C}\boldsymbol{R}_B} \left\| {}^{C}\breve{\boldsymbol{R}}_B - {}^{C}\boldsymbol{R}_B \right\|_F^2 \tag{4.15}$$

This is equivalent to maximizing trace$({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\boldsymbol{R}_B)$, as

$$\begin{aligned}\left\| {}^{C}\breve{\boldsymbol{R}}_B - {}^{C}\boldsymbol{R}_B \right\|_F^2 &= \mathrm{trace}\left(\left({}^{C}\breve{\boldsymbol{R}}_B - {}^{C}\boldsymbol{R}_B\right)^T \left({}^{C}\breve{\boldsymbol{R}}_B - {}^{C}\boldsymbol{R}_B\right)\right)\\ &= \mathrm{trace}({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\breve{\boldsymbol{R}}_B) - \mathrm{trace}({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\boldsymbol{R}_B)\\ &\quad - \mathrm{trace}({}^{C}\boldsymbol{R}_B^{\,T}\,{}^{C}\breve{\boldsymbol{R}}_B) + \mathrm{trace}({}^{C}\boldsymbol{R}_B^{\,T}\,{}^{C}\boldsymbol{R}_B)\\ &= 3 - 2\mathrm{trace}({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\boldsymbol{R}_B) + \mathrm{trace}({}^{C}\boldsymbol{R}_B^{\,T}\,{}^{C}\boldsymbol{R}_B)\end{aligned} \tag{4.16}$$

If ${}^{C}\boldsymbol{R}_B$ is decomposed into its singular values, so ${}^{C}\boldsymbol{R}_B = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$ with $\boldsymbol{D} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$, and an orthogonal matrix $\boldsymbol{Z}$ is defined as $\boldsymbol{Z} = \boldsymbol{V}^T\,{}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,\boldsymbol{U}$,

$$\begin{aligned}\mathrm{trace}({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,{}^{C}\boldsymbol{R}_B) &= \mathrm{trace}({}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T) = \mathrm{trace}(\boldsymbol{V}^T\,{}^{C}\breve{\boldsymbol{R}}_B^{\,T}\,\boldsymbol{U}\boldsymbol{D})\\ &= \mathrm{trace}(\boldsymbol{Z}\boldsymbol{D}) = \sum_{i=1}^{3} Z_{ii}\sigma_i \le \sum_{i=1}^{3} \sigma_i\end{aligned} \tag{4.17}$$

To maximize trace$(\boldsymbol{Z}\boldsymbol{D})$, Z_{ii} need to be as large as possible, optimally $\boldsymbol{Z} = \boldsymbol{I}$. This can be accomplished by setting ${}^{C}\breve{\boldsymbol{R}}_B = \boldsymbol{U}\boldsymbol{V}^T$.

Non-linear Optimization Including Distortion Estimation The distortion model of the camera calibration implementation provided by OpenCV is slightly different from Zhang's approach. Additionally to modeling two parameters κ_1, κ_2 for the radial distortion, it contains two parameters κ_3, κ_4 for the tangential distortion vector, inspired by the work of Heikkilä and Silvén [64]. The distortion center is assumed to coincide with the principal point.

To include the distortion model, equation 4.3 is split up into:

$$\boldsymbol{X}_C = \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \begin{bmatrix} {}^{C}\boldsymbol{R}_B & {}^{C}\boldsymbol{t}_B \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_B \\ 1 \end{bmatrix} \tag{4.18}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{Z_C} \begin{bmatrix} X_C \\ Y_C \end{bmatrix} \tag{4.19}$$

$$\breve{\boldsymbol{x}}_C = \underbrace{\left(1 + \kappa_1 r^2 + \kappa_2 r^4\right) \begin{bmatrix} a \\ b \end{bmatrix}}_{\text{radial distortion}} + \underbrace{\begin{bmatrix} 2\kappa_3 ab + \kappa_4(r^2 + 2a^2) \\ 2\kappa_4 ab + \kappa_3(r^2 + 2b^2) \end{bmatrix}}_{\text{tangential distortion}} \tag{4.20}$$

$$\begin{bmatrix} \boldsymbol{x}_C \\ 1 \end{bmatrix} = \boldsymbol{K} \begin{bmatrix} \breve{\boldsymbol{x}}_C \\ 1 \end{bmatrix} \tag{4.21}$$

where $r^2 = a^2 + b^2$. This means that after its transformation into camera coordinates, an ideal point in pinhole projection coordinates $(a, b)^T$ is first distorted into $\breve{\boldsymbol{x}}_C$, before it is converted into pixel coordinates by $\boldsymbol{K}$.

Finally, given $i = 1 \ldots m$ images and $j = 1 \ldots n$ points for each image, the complete set of camera parameters can be obtained by minimizing following functional using Levenberg-Marquardt:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \left\| \boldsymbol{x}_{C(ij)} - \hat{\boldsymbol{m}}(\boldsymbol{K}, \kappa_1, \kappa_2, \kappa_3, \kappa_4, {}^{C}\boldsymbol{R}_{B(i)}, {}^{C}\boldsymbol{t}_{B(i)}, \boldsymbol{X}_{B(ij)}) \right\|^2 \tag{4.22}$$

where $\hat{\boldsymbol{m}}$ is the projection of a point $\boldsymbol{X}_{B(ij)}$ of camera pose i according to equations 4.18 - 4.21. ${}^{C}\boldsymbol{R}_B$ can be parameterized by three parameters by e.g. utilizing the Rodrigues formula or Euler angles. The optimization can be initialized with the results of above closed-from solution and setting the distortion coefficients to zero.

To later undistort laparoscope images in real time (see figure 4.2b), the other direction than described by equation 4.20 is required. To map every distorted point to its corresponding undistorted point at real-time frame rates, a lookup table is created utilizing OpenCV. Linear interpolation is usually sufficient to fill any occurring gaps between the undistorted points.

4.1.1.2 Hand-eye Calibration

Hand-eye calibration is an important concept for augmented reality applications, which is used for several calibration steps throughout this thesis. Its naming originally comes from robotics, where a camera ("eye") is rigidly mounted on a robot arm ("hand") and the camera center coordinates need to be determined within the robot coordinate frame [157, 178, 179]. In this work, the robot is replaced by an optical tracking system. Speaking of the problem in a more generalized way, the fixed offset between two rigidly connected coordinate frames needs to be computed, where their poses are determined independently by two different tracking systems and hence in two different reference coordinate frames. The computation needs to be performed without previous co-calibration of the reference coordinate frames of the two independent tracking systems.

Hand-eye calibration solves this problem in the following elegant way: The two rigidly connected frames are moved to several stations (at least three) within a tracking volume covered by both tracking systems. For each pair of stations, the relative motion of each frame is recorded, i.e. the motion $\boldsymbol{A}$ between two "eye" stations and the motion $\boldsymbol{B}$ between two "hand" stations is stored. During all motions (at least two motions, all with distinct rotation axes), the transformation $\boldsymbol{X}$ between the two rigidly connected coordinate frames remains constant.

The camera calibration algorithms described above compute the transformations ${}^{C}\boldsymbol{T}_{B(i)}$ of pattern coordinate frame F_B to the camera frame for every station i (out of m stations). Each of them has a respective 6D pose ${}^{W}\boldsymbol{T}_{L(i)}$ that is provided by the tracking system. As visualized in figure 4.3, two of them can be each paired to a motion $\boldsymbol{T}_{C(l \leftarrow k)} = \left[{}^{C}\boldsymbol{T}_{B(l)}\right] \left[{}^{C}\boldsymbol{T}_{B(k)}\right]^{-1}$ and $\boldsymbol{T}_{L(l \leftarrow k)} = \left[{}^{W}\boldsymbol{T}_{L(l)}\right]^{-1} \left[{}^{W}\boldsymbol{T}_{L(k)}\right]$ respectively, giving $M = m(m-1)/2$ possible motions (excluding inverse motions) for the camera center as

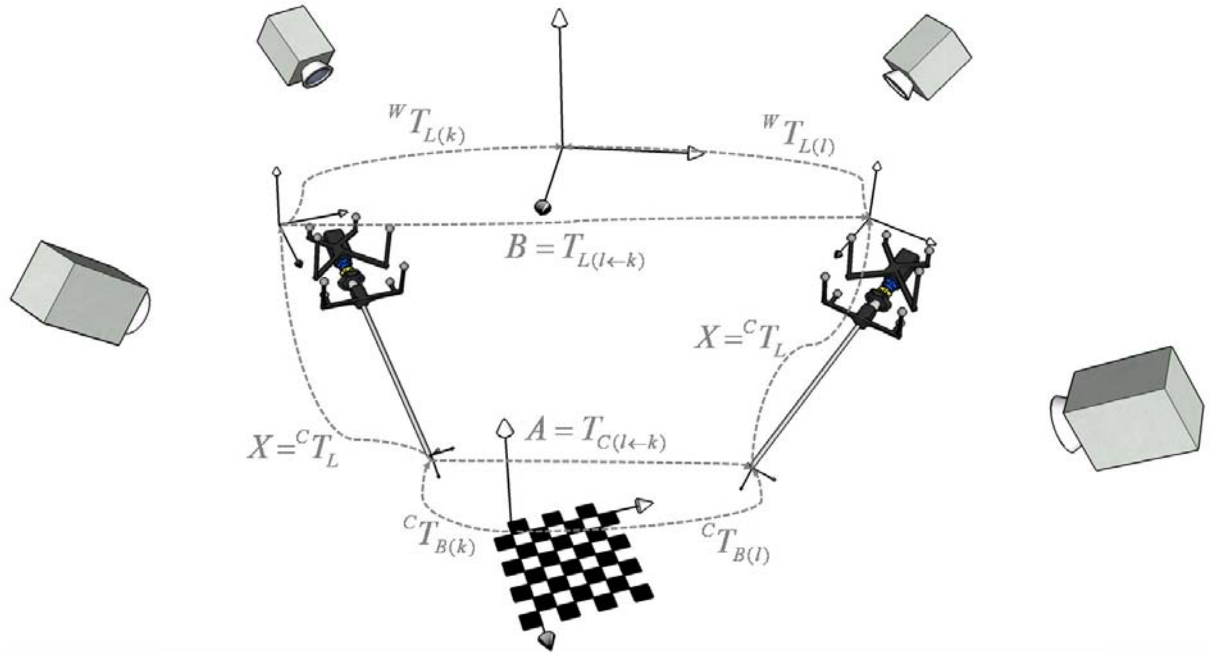


Figure 4.3: Coordinate transformations during hand-eye calibration.

well as the laparoscope body. One motion pair gives following equation:

$$\left[\boldsymbol{T}_{C(l \leftarrow k)}\right]\left[{}^{C}\boldsymbol{T}_{L}\right] = \left[{}^{C}\boldsymbol{T}_{L}\right]\left[\boldsymbol{T}_{L(l \leftarrow k)}\right] \tag{4.23}$$

Herewith, the so-called hand-eye formulation can be formalized, where $\boldsymbol{X} = \left[{}^{C}\boldsymbol{T}_{L}\right]$, $\boldsymbol{A} = \boldsymbol{T}_{C(l \leftarrow k)}$, and $\boldsymbol{B} = \boldsymbol{T}_{L(l \leftarrow k)}$:

$$\boldsymbol{AX} = \boldsymbol{XB} \tag{4.24}$$

or, equivalently:

$$\begin{pmatrix} \boldsymbol{R}_A & \boldsymbol{t}_A \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{R}_X & \boldsymbol{t}_X \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}_X & \boldsymbol{t}_X \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{R}_B & \boldsymbol{t}_B \\ 0 & 1 \end{pmatrix} \tag{4.25}$$

After stacking up the equations $\boldsymbol{AX} = \boldsymbol{XB}$ for all M motions, $\boldsymbol{X}$ can be solved for instance by first computing the rotational and then the translational part of $\boldsymbol{X}$ by QR decomposition, as proposed by Tsai and Lenz [178, 179], or by a dual quaternions approach, first described by Daniilidis [32]. Various other methods have been proposed, for instance by Chou and Kamel [26], who use a quaternion representation of the rotation, by Li and Betsis [95], who use a canonical matrix representation of the rotation, or by Horaud and Dornaika [69], who propose a non-linear equation system to solve the hand-eye equation. All methods however give similar results in terms of accuracy [116]. In this work, the methods of Tsai and Lenz as well as Daniilidis were compared (see below). A more complete review on hand-eye calibration methods is provided by Strobl and Hirzinger [167].

Several authors also propose the selection of optimal motion pairs. Shi et al. describe an algorithm to avoid degenerate cases such as pure translations and small rotations to decrease the calibration error [156]. Schmidt et al. propose a modified hand-eye calibration routine providing improved numerical stability by selecting an optimal set of relative motions and also eliminating erroneous robot data. A stable hand-eye calibration can be provided, although the error of the pose data provided by the robot can be large depending on the motion actually executed [150].

The transformation from the checkerboard to the world coordinate frame ${}^{W}\boldsymbol{T}_{B}$ can be computed in two ways: The obvious method is to just close the transformation loop after computation of $\boldsymbol{X}$, for instance by ${}^{W}\boldsymbol{T}_{B} = {}^{W}\boldsymbol{T}_{L(k)}\ \boldsymbol{X}\ {}^{C}\boldsymbol{T}_{B(k)}$ for an arbitrary k or by averaging/optimizing over all M. Alternatively, another hand-eye calibration can be performed using the same $i = 1 \dots m$ poses ${}^{C}\boldsymbol{T}_{B(i)}$ and ${}^{W}\boldsymbol{T}_{L(i)}$, respectively. However, the difference for this hand-eye calibration is that all computations are done in the camera center frame and laparoscope body frame as reference coordinate frames, i.e. theoretically the pattern and optical tracking system frames are now moved.

Once ${}^{W}\boldsymbol{T}_{B}$ is known, the 3D corner points of the checkerboard pattern can be projected into the (undistorted or original) laparoscope video in real time, enabling an easy evaluation of the calibration accuracy.

In the following, the hand-eye calibration methods of Tsai and Lenz as well as Daniilidis are shortly depicted, as they were both evaluated in this dissertation.

Tsai and Lenz Tsai and Lenz decouple the determination of the transformation matrix $\boldsymbol{X}$ into its rotational part $\boldsymbol{R}_X$ and translational part $\boldsymbol{t}_X$:

$$\boldsymbol{R}_A \boldsymbol{R}_X = \boldsymbol{R}_X \boldsymbol{R}_B \tag{4.26}$$

$$\boldsymbol{R}_A \boldsymbol{t}_X + \boldsymbol{t}_A = \boldsymbol{R}_X \boldsymbol{t}_B + \boldsymbol{t}_X \tag{4.27}$$

They define $\boldsymbol{p}$ as the rotation axis of a rotation $\boldsymbol{R}$, which is actually the vector part of the equivalent quaternion. First, a scaled version of $\boldsymbol{p}_X$, defined as $\boldsymbol{p}_X^* = \frac{1}{2\cos\left(\frac{\theta_X}{2}\right)}\boldsymbol{p}_X = \frac{1}{\sqrt{4-\|\boldsymbol{p}_X\|^2}}\boldsymbol{p}_X$, is computed, which can be achieved by a QR decomposition:

$$[\boldsymbol{p}_A + \boldsymbol{p}_B]_\times \boldsymbol{p}_X^* = \boldsymbol{p}_A - \boldsymbol{p}_B \tag{4.28}$$

where $\boldsymbol{p}_A$ and $\boldsymbol{p}_B$ are one of M pairs of rotation axes and a skew symmetric matrix $[\boldsymbol{s}]_\times$ is defined as

$$[\boldsymbol{s}]_\times = \begin{bmatrix} 0 & -s_z & s_y \\ s_z & 0 & -s_x \\ -s_y & s_x & 0 \end{bmatrix}$$

Next, $\boldsymbol{p}_X$ is computed as:

$$\boldsymbol{p}_X = \frac{2\boldsymbol{p}_X^*}{\sqrt{1 + \|\boldsymbol{p}_X^*\|^2}} \tag{4.29}$$

$\boldsymbol{R}_X$ can be computed from $\boldsymbol{p}_X$ with following equation:

$$\boldsymbol{R}_X = \left(1 - \frac{\|\boldsymbol{p}_X\|^2}{2}\right) I + \frac{1}{2}\left(\boldsymbol{p}_X \boldsymbol{p}_X^T + \sqrt{4 - \|\boldsymbol{p}_X\|^2}\,[\boldsymbol{p}_X]_\times\right) \tag{4.30}$$

$\boldsymbol{t}_X$ can now be determined using another QR decomposition:

$$(\boldsymbol{R}_A - \boldsymbol{I})\boldsymbol{t}_X = \boldsymbol{R}_X\boldsymbol{t}_B - \boldsymbol{t}_A \tag{4.31}$$

where, again, $\boldsymbol{R}_A$, $\boldsymbol{t}_B$, and $\boldsymbol{t}_A$ correspond to one of M motion pairs.

Reasoning Parts of the proofs for above equations are stated in the following. More complete proofs than the ones presented here can be found in the original works of Tsai and Lenz [178, 179].

The illustration provided by figure 4.4, which shows the rotation axes corresponding to one motion pair, helps to understand the reasoning.

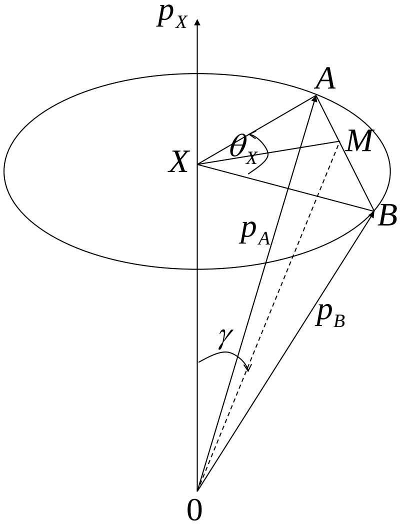


Figure 4.4: Rotation axes involved in hand-eye calibration.

1. $\boldsymbol{R}_X$ rotates the rotation axis of $\boldsymbol{R}_B$ into that of $\boldsymbol{R}_A$:

$$\boldsymbol{p}_A = \boldsymbol{R}_X\boldsymbol{p}_B \tag{4.32}$$

2. The rotation axis of $\boldsymbol{R}_X$ is perpendicular to the vector joining the ends of the rotation axes for $\boldsymbol{R}_B$ and $\boldsymbol{R}_A$:

$$\boldsymbol{p}_X \perp (\boldsymbol{p}_A - \boldsymbol{p}_B) \tag{4.33}$$

because

$$\begin{aligned}(\boldsymbol{p}_A - \boldsymbol{p}_B)^T\boldsymbol{p}_X &= (\boldsymbol{p}_A - \boldsymbol{p}_B)^T\boldsymbol{R}_X^T\boldsymbol{R}_X\boldsymbol{p}_X = (\boldsymbol{R}_X(\boldsymbol{p}_A - \boldsymbol{p}_B))^T\boldsymbol{R}_X\boldsymbol{p}_X \\ &= (\boldsymbol{R}_X\boldsymbol{p}_A - \boldsymbol{R}_X\boldsymbol{p}_B)^T\boldsymbol{p}_X = ((\boldsymbol{R}_X - \boldsymbol{I})\boldsymbol{p}_A)^T\boldsymbol{p}_X \\ &= \boldsymbol{p}_A^T(\boldsymbol{R}_X - \boldsymbol{I})^T\boldsymbol{p}_X = 0\end{aligned} \tag{4.34}$$

3. As it can be observed in figure 4.4, $\boldsymbol{p}_A - \boldsymbol{p}_B$ is colinear with $(\boldsymbol{p}_A + \boldsymbol{p}_B) \times \boldsymbol{p}_X$. Furthermore, $\boldsymbol{p}_A - \boldsymbol{p}_B$ and $(\boldsymbol{p}_A + \boldsymbol{p}_B) \times \boldsymbol{p}_X^*$ have the same length, because

$$\begin{aligned} |(\boldsymbol{p}_A + \boldsymbol{p}_B) \times \boldsymbol{p}_X^*| &= |\boldsymbol{p}_A + \boldsymbol{p}_B| 2 \sin\frac{\theta_X}{2}(4 - |2\sin\frac{\theta_X}{2}|^2)^{-\frac{1}{2}} \sin\gamma \\ &= |\boldsymbol{p}_A + \boldsymbol{p}_B| \sin\frac{\theta_X}{2}(1 - \sin^2\frac{\theta_X}{2})^{-\frac{1}{2}} \sin\gamma \\ &= |\boldsymbol{p}_A + \boldsymbol{p}_B| \tan\frac{\theta_X}{2} \sin\gamma = 2|\overline{\boldsymbol{OM}}| \sin\gamma \tan\frac{\theta_X}{2} \\ &= 2|\overline{\boldsymbol{XM}}| \tan\frac{\theta_X}{2} = 2|\overline{\boldsymbol{MB}}| = |\overline{\boldsymbol{AB}}| = |\boldsymbol{p}_A - \boldsymbol{p}_B| \end{aligned} \tag{4.35}$$

Because of this collinearity and their same lengths, $\boldsymbol{p}_A - \boldsymbol{p}_B = (\boldsymbol{p}_A + \boldsymbol{p}_B) \times \boldsymbol{p}_X^*$, or, written as a skew symmetric matrix, $[\boldsymbol{p}_A + \boldsymbol{p}_B]_\times \boldsymbol{p}_X^* = \boldsymbol{p}_A - \boldsymbol{p}_B$. As $[\boldsymbol{p}_A + \boldsymbol{p}_B]_\times$ is singular and has rank two, at least two motions around distinct rotation axes are necessary to compute $\boldsymbol{X}$.

Daniilidis Different from the work of Tsai and Lenz, where two QR decompositions are required to solve the hand-eye equation for $\boldsymbol{X}$, Daniilidis proposes a dual quaternion approach, which can be done in a single step.

A unit dual quaternion $\check{\boldsymbol{q}}$ consists of a pure rotational quaternion $\boldsymbol{q}$ with dual part equal to zero (the nondual part) and a pure translational unit dual quaternion $\boldsymbol{q}'$ (the dual part):

$$\check{\boldsymbol{q}} = \boldsymbol{q} + \epsilon\boldsymbol{q}' = \boldsymbol{q} + \epsilon\frac{1}{2}(0, \boldsymbol{t})\boldsymbol{q}$$

where $(0, \boldsymbol{t})$ is a translational quaternion, $\epsilon^2 = 0$, and $\boldsymbol{q}$ is a quaternion corresponding to the rotation matrix $\boldsymbol{R}$. More details on quaternions, dual quaternions, and their arithmetic are given in the original paper of Daniilidis [32].

The main idea of Daniilidis' approach is to stack up a $6M \times 8$ matrix $\boldsymbol{M}$ concatenated from $i = 1 \dots M$ matrices $\boldsymbol{M}_i$, which satisfy following equation (where M is the number of motions):

$$\boldsymbol{M}\begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{q}' \end{pmatrix} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \\ \vdots \\ \boldsymbol{M}_M \end{pmatrix} \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{q}' \end{pmatrix} = \boldsymbol{0} \tag{4.36}$$

where a single motion described by a 6×8 matrix $\boldsymbol{M}_i$ is given by

$$\boldsymbol{M}_i = \begin{pmatrix} \boldsymbol{a} - \boldsymbol{b} & [\boldsymbol{a} + \boldsymbol{b}]_\times & \boldsymbol{0}_{3\times1} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{a}' - \boldsymbol{b}' & [\boldsymbol{a}' + \boldsymbol{b}']_\times & \boldsymbol{a} - \boldsymbol{b} & [\boldsymbol{a} + \boldsymbol{b}]_\times \end{pmatrix} \tag{4.37}$$

and

- $\boldsymbol{a}$ is the vector part of the quaternion $\boldsymbol{q}_A$ corresponding to the rotation $\boldsymbol{R}_A$ of $\boldsymbol{A}$,
- $\boldsymbol{a}'$ is the vector part of the dual part $\boldsymbol{q}_A' = \frac{1}{2}(0, \boldsymbol{t}_A)\boldsymbol{q}_A$ of the dual quaternion corresponding to $\boldsymbol{A}$, and

- $\boldsymbol{b}$ and $\boldsymbol{b}'$ are defined in an analog way.

To solve above equation, a singular value decomposition is applied to $\boldsymbol{M}$ to obtain $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$. The two last right-singular 8×1 vectors (columns) of $\boldsymbol{V}$, $\boldsymbol{v}_7$ and $\boldsymbol{v}_8$, are written as $\boldsymbol{v}_7 = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{v}_1 \end{pmatrix}$ and $\boldsymbol{v}_8 = \begin{pmatrix} \boldsymbol{u}_2 \\ \boldsymbol{v}_2 \end{pmatrix}$, where $\boldsymbol{u}_1$, $\boldsymbol{v}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{v}_2$ are 4×1 vectors.

Next, s is computed:

$$s_{1,2} = \frac{-(\boldsymbol{u}_1^T\boldsymbol{v}_2 + \boldsymbol{u}_2^T\boldsymbol{v}_1) \pm \sqrt{(\boldsymbol{u}_1^T\boldsymbol{v}_2 + \boldsymbol{u}_2^T\boldsymbol{v}_1)^2 - 4(\boldsymbol{u}_1^T\boldsymbol{v}_1)(\boldsymbol{u}_2^T\boldsymbol{v}_2)}}{2(\boldsymbol{u}_1^T\boldsymbol{v}_1)} \tag{4.38}$$

For each of the two solutions for s, the solution r for following trinomial is computed:

$$r_{1,2} = s^2(\boldsymbol{u}_1^T\boldsymbol{u}_1) + 2s(\boldsymbol{u}_1^T\boldsymbol{u}_2) + (\boldsymbol{u}_2^T\boldsymbol{u}_2) \tag{4.39}$$

For the larger r, λ_1 and λ_2 are computed:

$$\lambda_2 = \sqrt{1/r} \tag{4.40}$$

$$\lambda_1 = s\lambda_2 \tag{4.41}$$

The resulting nondual part $\boldsymbol{q}$ and dual part $\boldsymbol{q}'$ of the dual quaternion $\check{\boldsymbol{q}}$ can now be determined as

$$\begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{q}' \end{pmatrix} = \lambda_1 \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{v}_1 \end{pmatrix} + \lambda_2 \begin{pmatrix} \boldsymbol{u}_2 \\ \boldsymbol{v}_2 \end{pmatrix} \tag{4.42}$$

Compared to Tsai and Lenz, the results provided by above dual quaternion $\check{\boldsymbol{q}}$ actually describe the inverse transformation $\boldsymbol{X}^{-1}$. Therefore, the rotational and translational parts need to be inverted to comply with the definition of Tsai and Lenz. Hence, the final quaternion $\boldsymbol{q}_X$ corresponding to $\boldsymbol{R}_X$ is $\boldsymbol{q}_X = \bar{\boldsymbol{q}}$, i.e. the conjugate quaternion of $\boldsymbol{q}$, while the final translation vector $\boldsymbol{t}_X$ is the vector part of the result of the quaternion multiplication $\bar{\boldsymbol{q}}(-2\boldsymbol{q}'\bar{\boldsymbol{q}})\boldsymbol{q}$.[5]

4.1.1.3 Oblique Scope Calibration

The transformation between laparoscope body and camera center ${}^C\boldsymbol{T}_L$ will only remain static, if either a 0° laparoscope is used or a 30° one with the scope shaft not being rotated around its longitudinal axis against the camera head. However, especially for abdominal surgery the shaft of the frequently used 30° laparoscope is often rotated, which invalidates the rigid transformation ${}^C\boldsymbol{T}_L$. To correct for this rotational offset, a second marker body is attached to the scope shaft to measure its rotation against the camera head. Mathematically, this single physical rotation can be modeled by two successive mathematical rotations, as proposed by Yamaguchi et al. [204, 205].

Therefore, first the camera and hand-eye calibration is performed keeping the scope shaft fixed at an initial constant rotation angle 0°. This results in an initial transformation ${}^C\boldsymbol{T}_{L(0)}$. To incorporate the laparoscope rotation into the final transformation ${}^C\boldsymbol{T}_L$, two mathematical rotations that are multiplied by the initial transformation are modeled:

---

[5] In general, a translation vector $\boldsymbol{t}$ can be recovered from a dual quaternion by $2\boldsymbol{q}'\bar{\boldsymbol{q}}$.

1. The inverse rotation around the longitudinal axis of the shaft (shaft rotation $\boldsymbol{T}_{Cs}(-\phi)$) by the rotation angle $-\phi$.
2. The rotation around the principal axis described by the viewing direction of the laparoscope (principal rotation $\boldsymbol{T}_{Cp}(\phi)$) by ϕ.

Following equation gives the overall transformation from the camera center to the laparoscope markers:

$$^{C}\boldsymbol{T}_L = \boldsymbol{T}_{Cp}(\phi)\boldsymbol{T}_{Cs}(-\phi)\,^{C}\boldsymbol{T}_{L(0)} \tag{4.43}$$

Since in reality both rotation axes are not perfectly aligned with the ideal longitudinal and principal axis, respectively, their actual positions need to be estimated (compare also figure 4.5).

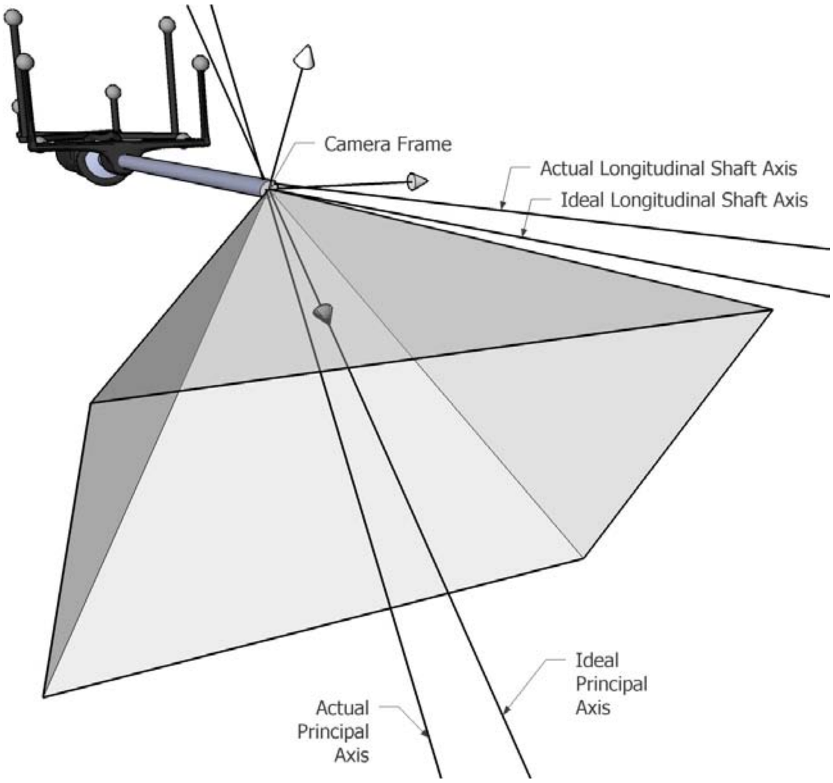


Figure 4.5: Ideal and actual rotation axes of the forward oblique laparoscope.

Shaft Rotation The shaft rotation axis is determined by performing a full 360° rotation while measuring the 3D positions of the tracking body attached to the laparoscope shaft. These $k = 1 \ldots p$ measurements are transformed by $^{C}\boldsymbol{T}_{L(0)}$ into the camera coordinate frame. After estimating their gravity center $\boldsymbol{X}_{Grav} = \left(\sum_{k=1}^{p} \boldsymbol{X}_{C(k)}\right)/p$, the eigenvector $\boldsymbol{v}_{\lambda min}$ corresponding to the minimum absolute eigenvalue of $\boldsymbol{A}$ is determined:

$$\boldsymbol{A} = \sum_{k=1}^{p}(\boldsymbol{X}_{C(k)} - \boldsymbol{X}_{Grav})(\boldsymbol{X}_{C(k)} - \boldsymbol{X}_{Grav})^T \tag{4.44}$$

This eigenvector, which is normal to the plane described by all measured circle points, corresponds to the directional vector of the shaft axis. To also estimate the center $\boldsymbol{X}_{Cen}$ of the circle, following functional needs to be minimized e.g. by Levenberg-Marquardt, as described by Yamaguchi et al. [204]:

$$\sum_{k=1}^{p} \rho^2 - \left\| \boldsymbol{X}_{C(k)} - \boldsymbol{X}_{Cen} \right\|^2 + (\boldsymbol{v}_{\lambda min}{}^T (\boldsymbol{X}_{C(k)} - \boldsymbol{X}_{Cen}))^2 \tag{4.45}$$

where ρ is the radius of the circle, which is only needed for the optimization, but not any more for modeling the shaft rotation. The center of the circle corresponds to the origin of the shaft rotation axis.

Principal Rotation The principal rotation axis can be computed acquiring images of the checkerboard pattern at several poses with different axis rotations $\phi_{(i)}$, $i = 1 \ldots m$, where the checkerboard pattern must not be moved between camera, hand-eye, and oblique scope calibration. Again, for each image at pose i, the 3D position of the shaft body is measured to estimate its corresponding rotation angle. Additionally, at the same time the pose of the camera head body is measured and $j = 1 \ldots n$ sorted 2D points $\boldsymbol{x}_{C(ij)}$ are automatically detected for each rotation angle. Using the initial laparoscope to camera transformation ${}^{C}\boldsymbol{T}_{L(0)}$, the origin and directional vector of the principal rotation axis can be computed by minimizing following functional using Levenberg-Marquardt:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \left\| \boldsymbol{K} \left[\boldsymbol{R}_{Cp}(\phi_{(i)}) | \boldsymbol{t}_{Cp}(\phi_{(i)}) \right] \boldsymbol{T}_{Cs}(-\phi_{(i)}) {}^{C}\boldsymbol{T}_{L(0)} {}^{L}\boldsymbol{T}_{W(i)} {}^{W}\boldsymbol{T}_{B} \begin{bmatrix} \boldsymbol{X}_{B(j)} \\ 1 \end{bmatrix} - \begin{bmatrix} \boldsymbol{x}_{C(ij)} \\ 1 \end{bmatrix} \right\|^2 \tag{4.46}$$

Above functional is only valid for point coordinates $\boldsymbol{x}_{C(ij)}$ detected in undistorted images. If they are segmented in original (distorted) images, the final mapping of a 3D point in camera coordinates onto the image plane needs to be extended by equations 4.19 - 4.21 instead of only applying $\boldsymbol{K}$.

4.1.2 Pointer

To acquire the 3D coordinates of points of a C-arm calibration phantom in the coordinate frame of the optical tracking system (cf. section 4.1.3), a tracked pointer is used, which is illustrated in figure 4.6a. To calibrate the pointer, i.e. to determine its tip position in its local tracking body coordinate frame, the tip is fixed inside an indentation and tracking data is recorded, while the pointer is pivoted around the tip describing a hemisphere. Hence, a series of $i = 1...m$ translational and rotational measurements are acquired, which represent the transformation from pointer body coordinates into world coordinates. For each transformation ${}^{W}\boldsymbol{T}_{P(i)}$, the tip of the pointer remains static in local tracking body coordinates ($\boldsymbol{X}_P$) as well as world coordinates ($\boldsymbol{X}_W$), as described by following formula:

$$\boldsymbol{X}_W = \left[{}^{W}\boldsymbol{R}_{P(i)} \right] \boldsymbol{X}_P + {}^{W}\boldsymbol{t}_{P(i)} \tag{4.47}$$

As depicted by Tuceryan et al. [180], following equation needs to be solved (e.g. by QR decomposition) to determine the pointer tip coordinates:

$$\begin{pmatrix} \boldsymbol{I} & -\left[{}^{W}\boldsymbol{R}_{P(1)}\right] \\ \boldsymbol{I} & -\left[{}^{W}\boldsymbol{R}_{P(2)}\right] \\ \vdots & \vdots \\ \boldsymbol{I} & -\left[{}^{W}\boldsymbol{R}_{P(m)}\right] \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_W \\ \boldsymbol{X}_P \end{pmatrix} = \begin{pmatrix} {}^{W}\boldsymbol{t}_{P(1)} \\ {}^{W}\boldsymbol{t}_{P(2)} \\ \vdots \\ {}^{W}\boldsymbol{t}_{P(m)} \end{pmatrix} \tag{4.48}$$

(a) Pointer.

(b) Hollow, cylindrical plastic phantom containing a helical pattern of 54 BBs of two different diameters.

Figure 4.6: Calibration utensils used for C-arm calibration.

4.1.3 C-arm

The computation of the transformation ${}^{F}\boldsymbol{T}_I$ from the isocenter to the flat panel coordinate frame is performed in two steps. First, the transformation ${}^{I}\boldsymbol{T}_G$ from the calibration phantom coordinate frame to the isocenter frame is determined during geometric calibration of the cone-beam CT capable C-arm, second the transformation ${}^{W}\boldsymbol{T}_G$ from the phantom to the world coordinate frame is computed applying a point based registration.

4.1.3.1 Geometric C-arm Calibration

For geometric C-arm calibration [128], the phantom and software of the C-arm manufacturer is used. A helical pattern of 54 BBs[6] of two different diameters inside a hollow, cylindrical plastic phantom is utilized, as visualized in figure 4.6b. The center of each BB is known very precisely in 3D in its local coordinate frame F_G. The geometric calibration is used to determine the reproducible non-ideal orbit of the C-arm around the scan

---

[6] A BB is a metal sphere.

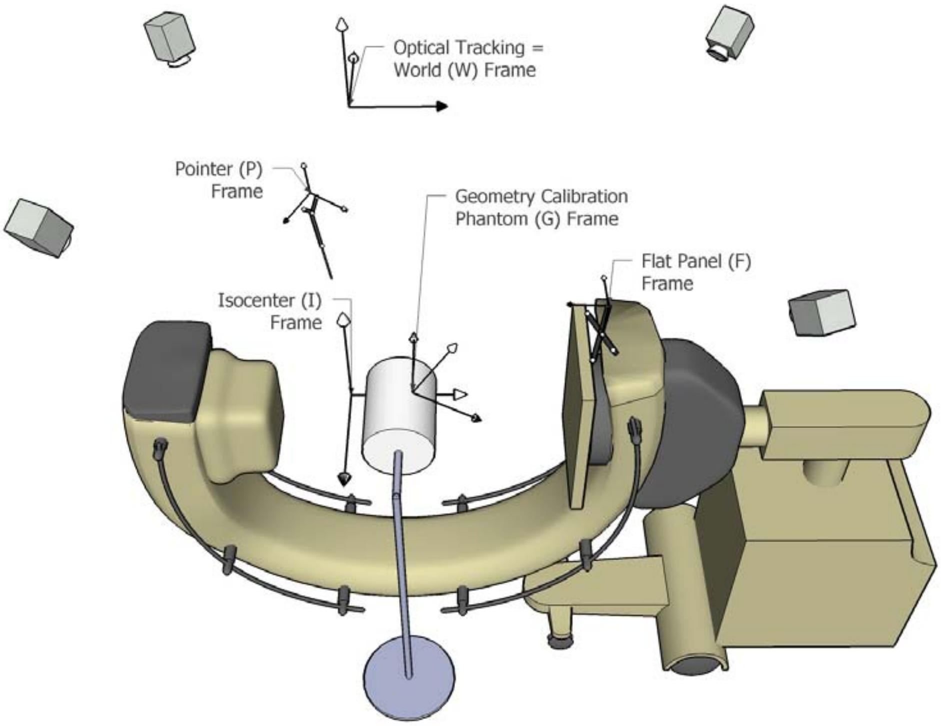


Figure 4.7: Coordinate frames for C-arm calibration.

volume. For every pose of the C-arm source (X-ray tube) and its corresponding X-ray image across the orbit (detected on the flat panel), it determines the associated projection matrix. Knowing all projection matrices, a rigid transformation ${}^{I}\boldsymbol{T}_{G}$ from the geometric calibration phantom coordinate frame to the isocenter can be estimated, which is done internally by the C-arm manufacturer's calibration software.

4.1.3.2 Point Based 3D Transformation

The last missing transformation ${}^{W}\boldsymbol{T}_{G}$ from the phantom coordinate frame to the world coordinate frame is then estimated as in Ritter et al. [144]. The tracked pointer is used to acquire the 3D coordinates of the outward surface of the BBs on the phantom in the world coordinate frame. To compensate for the fact that the pointer only touches the surface of a BB and not its center, an offset needs to be added to the coordinate of each BB. This offset equals the radius of a BB and is along the orthogonal to the tangent touching the surface of the phantom.

By matching the coordinates of the point set acquired with the pointer and the corresponding adjusted coordinates of the BBs in the calibration phantom, the rigid transformation ${}^{W}\boldsymbol{T}_{G}$ can be computed. Several ways of computation have been proposed in the literature, e.g. by applying singular value decomposition [3, 181] or dual quaternions [188]. However, there are no noticeable differences in terms of accuracy and stability of these algorithms for any practical application [37]. The method of Arun et al. [3], later refined by Umeyama [181], which is based on singular value decomposition, is used in this work to estimate the Euclidean transformation between two ordered corresponded 3D point sets $\boldsymbol{m}_i$ and $\boldsymbol{d}_i$, $i = 1 \ldots n$:

$$\boldsymbol{d}_i = \boldsymbol{R}\boldsymbol{m}_i + \boldsymbol{t} \tag{4.49}$$

First, the means of both point sets are calculated to determine their centroids, along with their variances:

$$\begin{aligned} \boldsymbol{\mu}_d &= \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{d}_i \\ \boldsymbol{\mu}_m &= \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{m}_i \end{aligned} \tag{4.50}$$

$$\begin{aligned} \sigma_d^2 &= \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{d_i} - \boldsymbol{\mu}_d\| \\ \sigma_m^2 &= \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{m_i} - \boldsymbol{\mu}_m\| \end{aligned} \tag{4.51}$$

Now the covariance matrix $\boldsymbol{\Sigma}_{dm}$ can be generated:

$$\boldsymbol{\Sigma}_{dm} = \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{m}_i - \boldsymbol{\mu}_m)(\boldsymbol{d}_i - \boldsymbol{\mu}_d)^T \tag{4.52}$$

which is decomposed into $\boldsymbol{\Sigma}_{dm} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$ by a singular value decomposition.

The rotation $\boldsymbol{R}$ rotating points $\boldsymbol{m}_i$ into points $\boldsymbol{d}_i$ can now be calculated from:

$$\boldsymbol{R} = \boldsymbol{V}\boldsymbol{S}\boldsymbol{U}^T, \text{ where } \boldsymbol{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(\boldsymbol{V}\boldsymbol{U}^T) \end{pmatrix} \tag{4.53}$$

The additional matrix $\boldsymbol{S}$ is introduced to take care of cases, where the determinant of $\boldsymbol{R}$ is not 1, but -1 (which represents a reflection rather than a rotation), as it can happen for planar or very noisy point sets.

The translation $\boldsymbol{t}$ can finally be derived from the centroids of the point sets and the rotation matrix:

$$\boldsymbol{t} = \boldsymbol{\mu}_d - s\boldsymbol{R}\boldsymbol{\mu}_m \tag{4.54}$$

where $s = 1$, if a Euclidean transformation is sought for, as it is the case here. More generally, if a similarity transformation with one scaling factor is wanted, which can be especially useful for data sets given in different units, s can be set to:

$$s = \frac{\operatorname{trace}(\boldsymbol{D}\boldsymbol{S})}{\sigma_d^2} \tag{4.55}$$

4.1.4 Ultrasound

Ultrasound calibration refers to the determination of the pixel scaling of the ultrasound B-scan plane and the transformation of the ultrasound plane to a tracking sensor or body attached to the transducer. This is useful to describe the transformation chain, which transforms a point in ultrasound plane coordinates $\boldsymbol{X}_{US}$ to tracking system coordinates. In the scope of this work, the tracking system coordinates are defined by the electromagnetic tracking (EMT) system:

$$\boldsymbol{X}_{EMT} = {}^{EMT}\boldsymbol{T}_{FlexS}{}^{FlexS}\boldsymbol{T}_{US}\boldsymbol{X}_{US} \tag{4.56}$$

where $\boldsymbol{X}_{US} = \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}$. s_x and s_y scale pixel units to millimeters, for the x and y direction of the image plane, respectively.

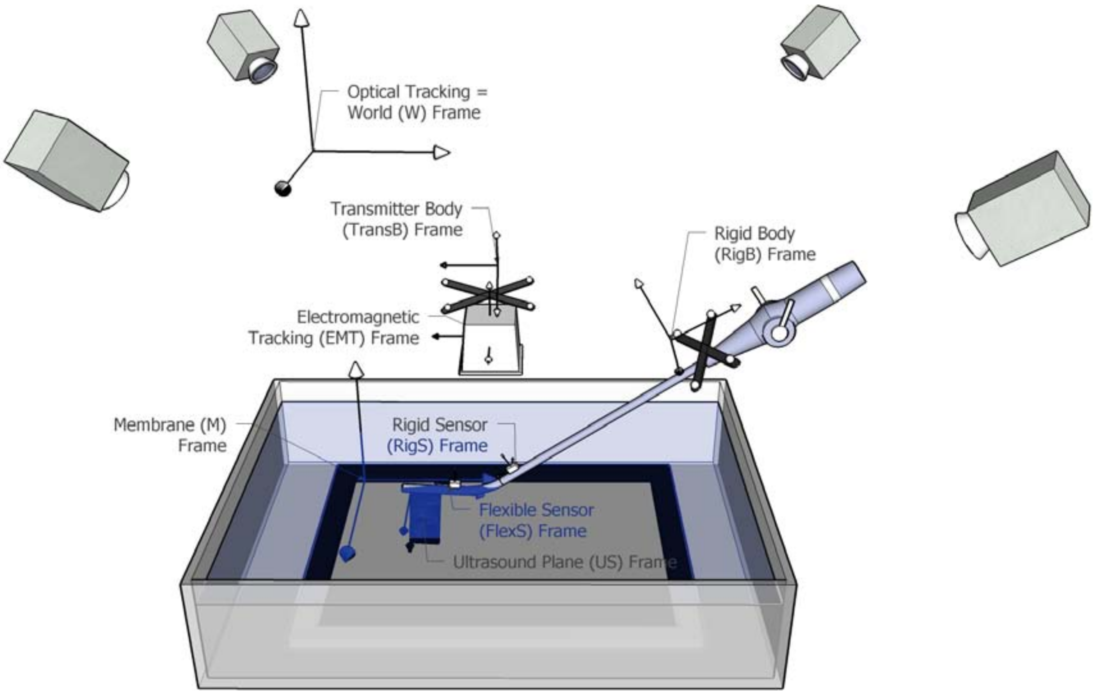


Figure 4.8: Coordinate frames for laparoscopic ultrasound calibration.

Ultrasound calibration usually aims at the computation of the parameters s_x, s_y, and ${}^{FlexS}\boldsymbol{T}_{US}$, wherefore various methods have been proposed in the literature [113]. Because it does not require the manufacturing of a specific elaborate calibration phantom, in this work a single-wall calibration [70, 139, 175] is performed to estimate the ultrasound parameters. In the original works on single-wall calibration, a flat-bottomed water bath, preferably uniformly roughened by very fine emery cloth or sand paper, is scanned, or alternatively the so-called Cambridge phantom is used for more accurate results. In this work, a nylon membrane (which can be from simple panty hose) stretched over a planar frame is used as calibration phantom, as introduced by Langø [90].

Single-wall calibration is based on the idea that when a point on the ultrasound plane $\boldsymbol{X}_{US}$ is transformed into the local coordinate frame F_M of the planar calibration phantom, its z coordinate will be always equal to 0, as the x-y plane of the local coordinate frame can be defined in such a way that it is exactly aligned with the plane of the calibration phantom. This relation is reflected by following equation:

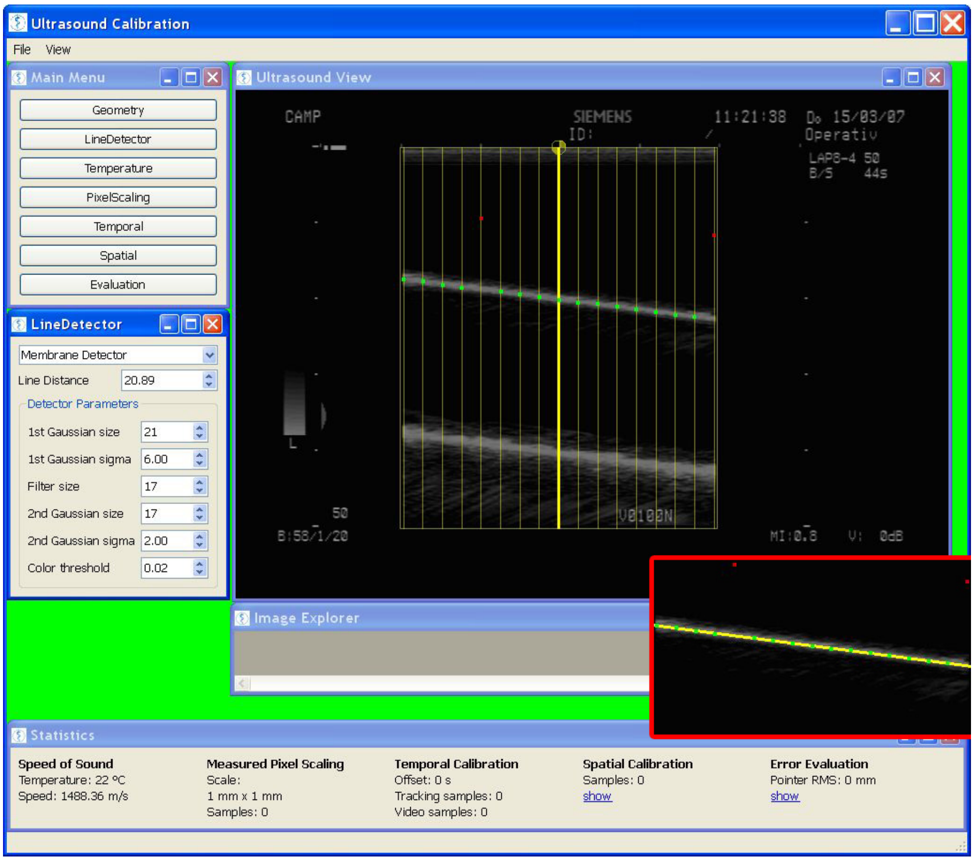
$$\boldsymbol{X}_M = \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = {}^{M}\boldsymbol{T}_{EMT}\,{}^{EMT}\boldsymbol{T}_{FlexS}\,{}^{FlexS}\boldsymbol{T}_{US} \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{4.57}$$

To collect samples for solving above equation, several poses of the flexible sensor of the transducer as well as corresponding B-scan images showing the membrane inside a water bath are acquired. The lines corresponding to the planar nylon membrane are automatically segmented and used for the computation of all calibration parameters, i.e. pixel scaling and the transformation ${}^{FlexS}\boldsymbol{T}_{US}$. To ensure numeric stability for all degrees of freedom of the transformation, the calibration protocol of the Cambridge group is adapted to acquire a set of about 40 lines and their corresponding tracking data [70, 139, 175]:

1. The transducer is moved vertically up and down without changing its orientation.
2. The transducer is rotated from side to side (both clockwise and counterclockwise) while keeping the scan plane and the phantom plane perpendicular.
3. The transducer is rotated towards the user and away from him/her, changing the angle between the scan plane and the phantom plane. The angle of the image of the phantom plane should not change significantly during these moves.
4. The transducer is rotated through 45 degrees in each direction about its vertical axis.
5. The transducer is translated across the phantom plane and motions 1, 2, 3, and 4 are performed in a total of three non-collinear locations.

For the automatic line segmentation, a two-step approach inspired by Prager et al. [139] is applied. First, features are detected on predefined vertical scan lines of the image, smoothed using a Gaussian filter kernel, median-filtered, then differentiated and thresholded. In the second step, the random sample consensus algorithm (RANSAC) [51] is used to match several candidate lines through these feature points and to choose the definite line from these candidates, taking into account the number of features supporting it and the proximity to the line detected in the last frame (cf. figure 4.9).

Finally, all collected sensor poses and segmented lines are used to solve equation 4.57. At least 11 independent samples are needed to determine the 11 identifiable parameters of all 14 unknown parameters in equation 4.57: Two for s_x, s_y, six for ${}^{FlexS}\boldsymbol{T}_{US}$, but only three of six for ${}^{M}\boldsymbol{T}_{EMT}$. The other three parameters of ${}^{M}\boldsymbol{T}_{EMT}$ are not identifiable, as the rotation of the planar membrane about its z-axis and the translation of the membrane in its own x-y plane do not affect the z-component of $\boldsymbol{X}_M$. A Levenberg-Marquardt optimization is applied to best fit the acquired samples.



Membrane Detector
Line Distance 20.89

Detector Parameters

1st Gaussian size 21
1st Gaussian sigma 6.00
Filter size 17
2nd Gaussian size 17
2nd Gaussian sigma 2.00
Color threshold 0.02

Speed of Sound
Temperature: 22 °C
Speed: 1488.36 m/s

Measured Pixel Scaling
Scale:
1 mm x 1 mm
Samples: 0

Temporal Calibration
Offset: 0 s
Tracking samples: 0
Video samples: 0

Spatial Calibration
Samples: 0
show

Error Evaluation
Pointer RMS: 0 mm
show

Figure 4.9: User interface developed for ultrasound calibration based on CAMPAR. It shows a B-scan image of the nylon membrane generated by the laparoscopic transducer. On the yellow vertical scan lines feature points are detected, which RANSAC either accepts (green points) or rejects (red points). The red framed inlet shows the line, which is finally fitted to all feature points.

4.1.4.1 Magneto-optic Hand-eye Calibration

As ultrasound calibration only determines the transformation into the electromagnetic tracking frame, the transformation into the optical world coordinate frame still needs to be determined. Furthermore, the technique for online electromagnetic error estimation proposed in section 4.3.1 requires the determination of the offset between the rigid electromagnetic sensor and the rigid optical tracking body attached to the rigid shaft of the ultrasound transducer.

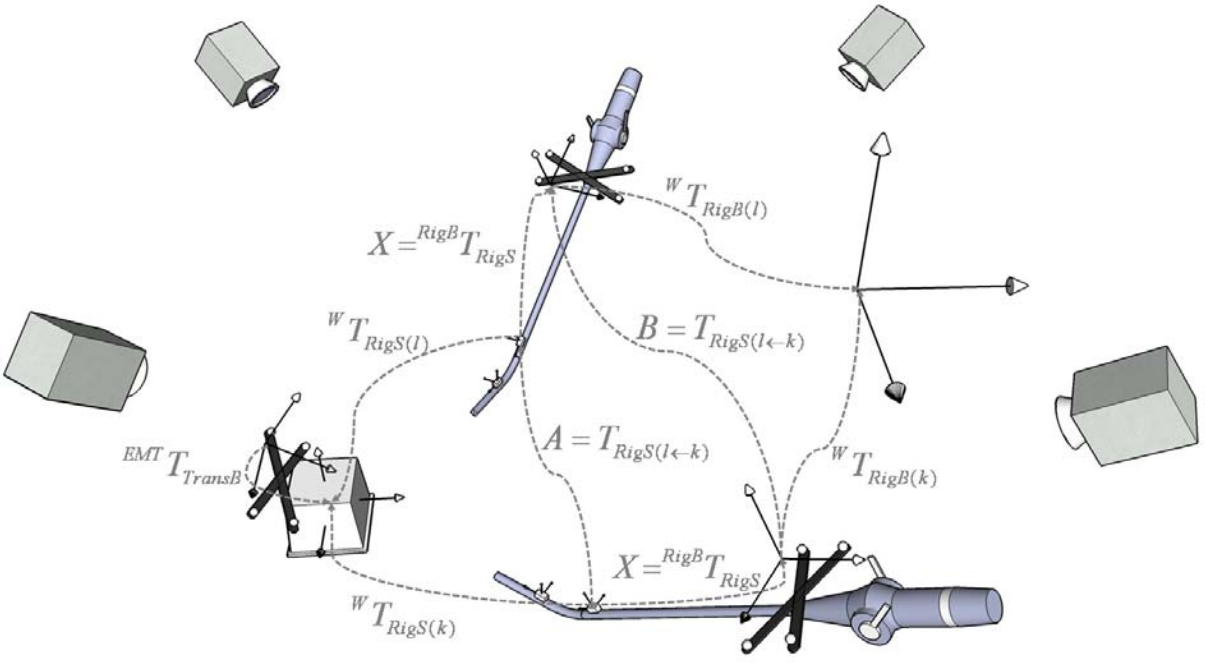


Figure 4.10: Coordinate transformations during magneto-optic hand-eye calibration.

To compute the Euclidean transformation ${}^{RigB}\boldsymbol{T}_{RigS}$ between the rigid sensor and the rigid body frames, the hand-eye calibration theory introduced in section 4.1.1.2 can be applied anew. This time, several poses of the rigid body and the rigid sensor with distinct rotation axes are recorded in both the optical tracking and electromagnetic tracking coordinate frames (cf. figure 4.10). Stacked matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are generated again from all movements between these poses and related to each other by $\boldsymbol{AX} = \boldsymbol{XB}$. Inverting the same poses, they can be used to estimate the rigid hand-eye transformation ${}^{EMT}\boldsymbol{T}_{TransB}$ between the electromagnetic transmitter coordinate frame and its optical tracking body.

In a final optimization step, the two hand-eye calibration matrices ${}^{RigB}\boldsymbol{T}_{RigS}$ and ${}^{EMT}\boldsymbol{T}_{TransB}$ are optimized for all recorded poses by the Levenberg-Marquardt optimizer. The matrix $\boldsymbol{T}_\delta$ resulting from the transformation chain "rigid sensor to rigid body to optical tracking (world) to transmitter body to electromagnetic tracking to rigid sensor frame", which theoretically is an identity matrix, represents the accumulated transformation errors:

$$\boldsymbol{T}_\delta = \begin{bmatrix} \boldsymbol{R}_\delta & \boldsymbol{t}_\delta \\ \boldsymbol{0} & 1 \end{bmatrix} = {}^{RigS}\boldsymbol{T}_{EMT} \; {}^{EMT}\boldsymbol{T}_{TransB} \; {}^{TransB}\boldsymbol{T}_{W} \; {}^{W}\boldsymbol{T}_{RigB} \; {}^{RigB}\boldsymbol{T}_{RigS} \tag{4.58}$$

For optimization, a cost function δ was chosen that weights translational to rotational errors 1:3, reflecting the RMS error ratio provided independently by the two tracking system manufacturers (0.4 mm and 0.12° for the optical tracking system, 1.4 mm and 0.5° for the electromagnetic tracking system).

$$\delta = \|\boldsymbol{t}_\delta\| \; + \; 3 \cdot \frac{180}{\pi} \cdot \arccos\left(\frac{\mathrm{trace}(\boldsymbol{R}_\delta) - 1}{2}\right) \tag{4.59}$$

The maximum error δ_{max} determined after optimization is chosen as a measure of distrust for the overall performance of the hand-eye calibration (cf. also section 4.3.1).

4.1.4.2 Temporal Calibration

In order to later provide a smooth visualization without lag, time stamps are used to bring all data into the same time frame. While the optical tracking PC and the visualization workstation are synchronized via the network time protocol (NTP) to the same reference time, the ultrasound and electromagnetic tracking systems require a more advanced synchronization. As these systems do not automatically provide reliable time stamps corresponding to the actual data acquisition time, a time stamp is generated when their data arrives at the visualization workstation. Therefore, a fixed offset is subtracted from this time stamp to compensate for any lag introduced while traveling to the workstation. To determine this offset, the magneto-optically tracked transducer is moved up and down (as suggested by Treece et al. [175]), while the line is automatically segmented in the B-scan image, as described above. All three motion sequences are stored, brought into 2D (time & translation) by a principal component analysis (PCA), and normalized. By translating the time axis to match their curves, the ultrasound offset and the electromagnetic tracking offset in regard to the reference time is computed.

4.1.4.3 Transducer Tip Model

Modeling the transducer tip can serve two purposes. First, the model can provide all necessary 3D information to overlay a virtual model of the ultrasound transducer onto the measured location. Second, if built relatively to the rigid optical tracking body, it can be used to correct tracking inaccuracies (see section 4.3.1).

The ultrasound transducer does not contain a single joint, but a lengthy bending region extending over approximately three centimeters, similar to the one visualized in figure 4.12. The horizontal and vertical bending of the tip is controlled by two steering levers. Each lever offers seven positions, giving 49 manually selectable tip poses. As the tip yields to external pressure as well, continuous motions on an ellipsoid-like shape are modeled to approximate all possible tip poses, i.e. poses of the flexible sensor in relation to the rigid sensor. Therefore, a chain of transformations is modeled to come from the coordinate frame of the flexible sensor into that of the rigid sensor, as visualized on figure 4.11.

The proposed model is built relatively to the rigid sensor, but could also be built relatively to the rigid body by applying the additional static hand-eye transformation ${}^{RigB}\boldsymbol{T}_{RigS}$, which is more useful for error correction.

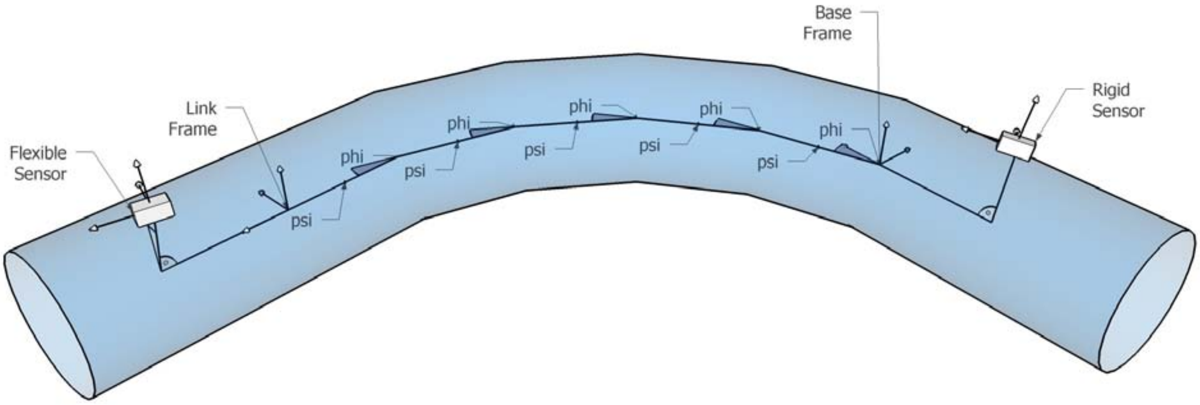


Figure 4.11: Mathematical model of the tip of the flexible ultrasound transducer – only the rotation ϕ about the x axis is visualized here, the rotation ψ about the y axis is zero.

First, the flexible sensor coordinate frame is rotated and translated into the "link frame" in such a way that the z axis points along the transducer axis. This transformation ${}^{Link}\boldsymbol{T}_{FlexS}$ has six degrees of freedom (DOF), i.e. three rotations and one translation to align the sensor with the transducer axis, followed by a rotation about and a translation along the transducer axis.

Starting at the link frame, n short links follow, which approximate the possible tip movements. In an alternating manner, each link rotates the transducer axis by ψ about the y axis and by ϕ about the x axis, respectively. No rotation about the z axis is assumed, but each link introduces a translational offset $\boldsymbol{t}_z$ along the z axis, so in total ${}^{Base}\boldsymbol{T}_{Link}$ has four DOF.

$$ {}^{Base}\boldsymbol{T}_{Link} = \underbrace{\boldsymbol{T}_{Link}(\phi) \cdot \boldsymbol{T}_{Link}(\psi) \cdot \boldsymbol{T}_{Link}(\phi) \cdot \ldots}_{n\ times} \tag{4.60} $$

where

$$ \boldsymbol{T}_{Link}(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) & 0 \\ 0 & \sin(\phi) & \cos(\phi) & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} $$

$$ \boldsymbol{T}_{Link}(\psi) = \begin{pmatrix} \cos(\psi) & 0 & \sin(\psi) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\psi) & 0 & \cos(\psi) & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} $$

This behavior of alternating links is inspired by the real nature of the bending region, where four cables controlled by the steering levers are pairwise and alternately connected to every second link.

After the bending region, i.e. at the base frame, a final translational and rotational offset ${}^{RigS}\boldsymbol{T}_{Base}$ is applied to align the final tip coordinate frame with the rigid sensor, which has five DOF. Actually, ${}^{Link}\boldsymbol{T}_{FlexS}$ and ${}^{RigS}\boldsymbol{T}_{Base}$ would both have six DOF. However, one DOF of the rotation about the transducer (z) axis is not needed for ${}^{RigS}\boldsymbol{T}_{Base}$, as

Figure 4.12: Typical bending region usually found in endoscopic instruments. Four cables are pairwise and alternately connected to every second link of the bending region. They are often controlled by steering levers, where two each are connected to the same lever.

it can be entirely compensated by just altering the combination and magnitudes of ϕ and ψ. Contrary, one more DOF is needed for the rotation about the transducer axis included in ${}^{Link}\boldsymbol{T}_{FlexS}$ to adjust for the fact that both sensors can be mounted at different angles about the transducer axis, i.e. assuming ϕ and ψ to be 0, the line through the origins of the two sensor coordinate frames will not be parallel to the transducer axis.

The final model transformation from flexible sensor coordinates into rigid sensor coordinates can be described by

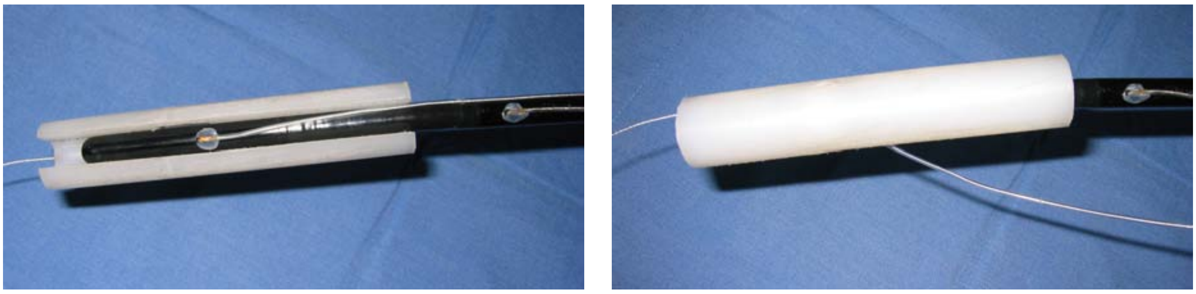
$$\boldsymbol{T}_{Model} = {}^{RigS}\boldsymbol{T}_{Base}\ {}^{Base}\boldsymbol{T}_{Link}\ {}^{Link}\boldsymbol{T}_{FlexS} \tag{4.61}$$

For the offline computation of the model, first the transducer axis is estimated, both in relation to the rigid sensor and to the flexible sensor. This can be done by putting a plastic cylinder over the transducer shaft, which contains an additional electromagnetic sensor on one end (cf. figure 4.13). For each the rigid part and the flexible part of the transducer shaft, the plastic cylinder is rotated at least 360 degrees, flipped, and rotated another 360 degrees or more. During each rotation, the measurements of the additional electromagnetic sensor are transformed into the local coordinate frames of the rigid sensor and the flexible sensor, respectively. Each acquired point cloud $\boldsymbol{X}_i$, $i = 1 \dots m$ is then fitted to a cylinder with radius r around the transducer axis, which is defined by the 3D point $\boldsymbol{b}$ on the axis closest to the rigid sensor (or flexible sensor, for the other point cloud) and a 3D unit vector $\boldsymbol{d}$ pointing along the axis to the tip of the transducer. Following functional is minimized by the Levenberg-Marquardt optimizer:

$$\sum_{i=1}^{m} |r - \|\boldsymbol{d} \times (\boldsymbol{b} - \boldsymbol{X}_i)\|| \tag{4.62}$$

Finally, two transducer axes are defined, one in flexible sensor coordinates by the base point $\boldsymbol{b}_{FlexS}$ and the unit directional vector $\boldsymbol{d}_{FlexS}$ and the other one in rigid sensor coordinates by $\boldsymbol{b}_{RigS}$ and $\boldsymbol{d}_{RigS}$.

All parameters except angles ϕ and ψ remain constant for a given configuration and can be computed offline. Given that the transducer axes in relation to the rigid and the flexible sensor have been calibrated, only five model parameters have to be estimated offline. These are the translation along the rigid shaft axis (contained in ${}^{RigS}\boldsymbol{T}_{Base}$), the length of the bending region ($n \cdot t_z$), the number of links in the bending region (which



(a) Plastic cylinder put over the transducer shaft with an additionally attached sensor on the left end.

(b) The plastic cylinder is rotated at least 360 degrees, flipped, and rotated another 360 degrees or more.

Figure 4.13: Transducer shaft calibration.

can be initialized or fixed to $n = 12$)[7], the angle of rotation about the tip axis, and the length of translation along the tip axis (both contained in ${}^{Link}\boldsymbol{T}_{FlexS}$).

For every selectable position of the two control levers, the position of the flexible sensor (attached to the transducer tip) in relation to the rigid sensor (attached to the transducer shaft) is recorded in a distortion-free environment, i.e. without ferrous-magnetic or electrically powered materials in close vicinity. Then the remaining model parameters are optimized numerically using the Levenberg-Marquardt algorithm. The model has to be computed only once, when the electromagnetic sensors are attached to the transducer. At run-time the remaining two DOF of the rotations about the x and y axis can be numerically optimized by the Levenberg-Marquardt optimizer.

4.2 Registration for Port Placement

Because tracked and intrinsically registered intraoperative imaging devices are primarily used for the augmented reality solution presented here, no manual or interactive registration is required for the alignment and fusion of all intraoperatively acquired images. However, the proposed port placement method is based on preoperative CT data, which is in many cases available anyway, but requires an additional intraoperative registration step. To avoid this registration step, one may argue to also use cone-beam C-arm data (instead of preoperative data) for port placement planning. Unfortunately, this can hardly be justified due to following reasons:

- The size of a reconstructed C-arm volume is $20 \times 20 \times 15$ cm^3, which only partially covers the patient anatomy needed for port placement planning.
- Patients usually get one or more preoperative diagnostic CT scans. If fiducials are already attached to a patient for these scans, their data can be used for port placement planning.
- To compensate for organ movements, complex laparoscopic interventions require an updated data set *after* port placement and CO_2 insufflation, so a second data set

---

[7] Many surgical instruments as the one in figure 4.12 have 12 links.

would have to be acquired just for port placement planning. Radiation and contrast agent exposure caused by an additional scan can hardly be justified.

To smoothly integrate the registration procedure into the current clinical workflow, most registration steps for the proposed method are automated. Manual interaction is mainly needed, when the laparoscope is moved around the fiducial markers attached to the patient, which takes approximately two minutes. In general, the patient registration procedure can be split into three essential steps:

1. All fiducials must be segmented in the CT volume to determine the positions of their centroids.
2. Their positions in the world coordinate frame need to be reconstructed using the images, which are acquired by moving the calibrated laparoscope around them and show the fiducials.
3. The resulting point sets need to be matched in order to register the patient to the CT data set.

The automatic segmentation of the fiducials in the CT volume can be achieved by using standard image processing techniques based on thresholding, filling, morphology, and subtraction [130, 189]. The centroids of all segmented fiducials can be computed very precisely by weighing their associated voxel intensities and incorporating partial volume effects.

4.2.1 3D Reconstruction

For finding the 3D positions of the fiducials in the world coordinate frame, two iterations are performed for each image i containing an arbitrary number m of fiducials. First, the 2D positions $\boldsymbol{x}_{C(i1)} \dots \boldsymbol{x}_{C(im)}$ of all visible fiducials are extracted automatically after undistortion of the image utilizing OpenCV. In detail, the fiducial segmentation techniques in use are based on thresholding, contour finding, polygonal curve approximation, moments calculation, roundness factor, and color information of the fiducials. Second, epipolar geometry is used to match and reconstruct their 3D positions as follows.

For each image i, the extrinsic parameters of the laparoscope camera, meaning the transformation ${}^{C}\boldsymbol{T}_{W(i)}$ from world to camera coordinate frame can be computed:

$$ {}^{C}\boldsymbol{T}_{W(i)} = {}^{C}\boldsymbol{T}_{L}\,{}^{L}\boldsymbol{T}_{W(i)} \tag{4.63}$$

Thus, for each image pair i and j, the transformation between its associated camera poses, $\boldsymbol{T}_{C(j\leftarrow i)}$, and its corresponding essential matrix $\boldsymbol{E}_{(j\leftarrow i)}$ can be computed [63].

$$\boldsymbol{T}_{C(j\leftarrow i)} = \begin{bmatrix} \boldsymbol{R}_{C(j\leftarrow i)} & \boldsymbol{t}_{C(j\leftarrow i)} \\ \boldsymbol{0} & 1 \end{bmatrix} = {}^{C}\boldsymbol{T}_{W(j)}\left({}^{C}\boldsymbol{T}_{W(i)}\right)^{-1} \tag{4.64}$$

$$\boldsymbol{E}_{(j\leftarrow i)} = [\boldsymbol{t}_{C(j\leftarrow i)}]_{\times}\boldsymbol{R}_{C(j\leftarrow i)} \tag{4.65}$$

$$\text{where } [\boldsymbol{t}]_{\times} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

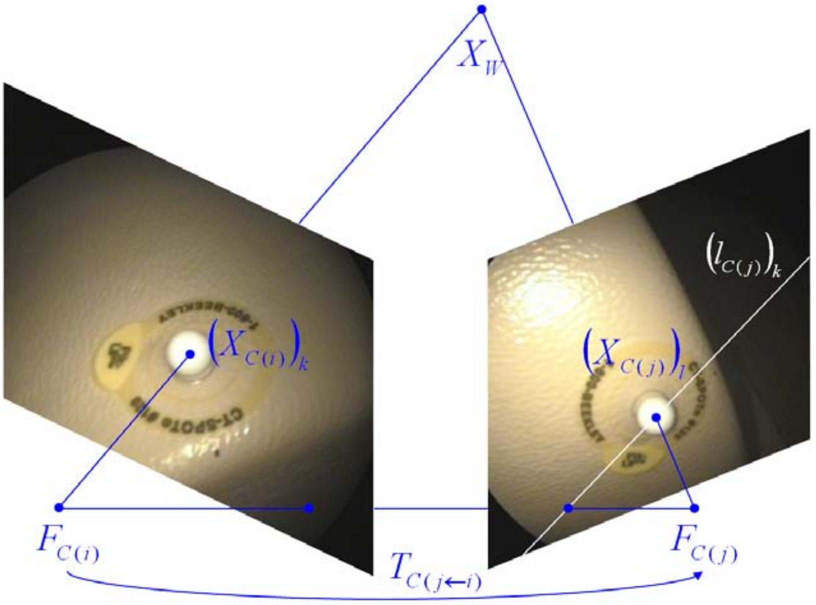


Figure 4.14: Epipolar geometry used for reconstruction.

Since the camera is calibrated, all epipolar lines $\boldsymbol{l}_{C(j1)} \ldots \boldsymbol{l}_{C(jm)}$ in image j corresponding to the points in image i can be determined:

$$\boldsymbol{l}_{C(jk)} = \boldsymbol{K}^{-T}\boldsymbol{E}_{(j\leftarrow i)}\boldsymbol{X}_{C(ik)} \tag{4.66}$$

$$\text{where } \boldsymbol{X}_{C(ik)} = \boldsymbol{K}^{-1}\begin{bmatrix} \boldsymbol{x}_{C(ik)} \\ 1 \end{bmatrix}$$

If any of the points in image j lies on an epipolar line or very close to it (e.g. within one pixel), it is very likely that this point and its corresponding point in image i are projections of the same 3D point. So this point is reconstructed by computing the intersection of the two rays back-projected from $\boldsymbol{X}_{C(ik)}$ and $\boldsymbol{X}_{C(jl)}$, respectively. However, since the intrinsic and extrinsic parameters of the camera were only estimated during calibration and tracking, the two rays will not exactly intersect in space. Hence, the midpoint of the two rays is computed by determining their intersection points with the segment that is orthogonal to both rays, which can be achieved by solving a linear system of equations [176].

$$\boldsymbol{t}_{C(i\leftarrow j)} = a\boldsymbol{X}_{C(ik)} - b\boldsymbol{R}_{C(i\leftarrow j)}\boldsymbol{X}_{C(jl)} + c\left(\boldsymbol{X}_{C(ik)} \times \left(\boldsymbol{R}_{C(i\leftarrow j)}\right)\boldsymbol{X}_{C(jl)}\right),$$

where $\boldsymbol{R}_{C(i\leftarrow j)} = \left(\boldsymbol{R}_{C(j\leftarrow i)}\right)^T$ and $\boldsymbol{t}_{C(i\leftarrow j)} = -\left(\boldsymbol{R}_{C(j\leftarrow i)}\right)^T \boldsymbol{t}_{C(j\leftarrow i)}$. The intersection points in the camera coordinate frame now are $a\boldsymbol{X}_{C(ik)}$ and $\boldsymbol{t}_{C(i\leftarrow j)} + b\left(\boldsymbol{R}_{C(i\leftarrow j)}\right)\boldsymbol{X}_{C(jl)}$, so their midpoint can be easily computed and transformed into the world coordinate frame, as described by following equation:

$$\begin{bmatrix} \boldsymbol{X}_W \\ 1 \end{bmatrix} = \left({}^{C}\boldsymbol{T}_{W(i)}\right)^{-1}\begin{bmatrix} \left(a\boldsymbol{X}_{C(ik)} + \boldsymbol{t}_{C(i\leftarrow j)} + b\left(\boldsymbol{R}_{C(i\leftarrow j)}\right)\boldsymbol{X}_{C(jl)}\right)/2 \\ 1 \end{bmatrix} \tag{4.67}$$

Wrong reconstructions can always happen, if a 2D point close to or on the epipolar line is found, which actually corresponds to another 3D point than the one the epipolar line belongs to. Following approach is taken to identify incorrect reconstructions. The fiducials are positioned on the patient at a minimum distance of 30 mm. Therefore, all reconstructed points having a certain maximal distance, e.g. are closer to each other than 10 mm, are assumed to represent the same point in 3D. For each such 3D point, its corresponding 2D point pairs are combined into a list. All lists are compared to each other. If a 2D point happens to be in more than one list, its corresponding point pair is kept in the list with most point pairs and deleted from all other lists. By validating lists with more than two point pairs only, 3D points can be reconstructed with the utmost probability.

In the next reconstruction step, the lists with point pairs corresponding to the same 3D point are used to optimally reconstruct the 3D point. For all 2D points, their associated projection rays $\boldsymbol{r}_1 \ldots \boldsymbol{r}_s$ are constructed, which intersect the camera center $\boldsymbol{C}_r = {}^W\boldsymbol{t}_{C(i)}$ and the projection of the point onto the image plane $\boldsymbol{P}_r = {}^W\boldsymbol{R}_{C(i)}(\boldsymbol{X}_{C(i)})_k + {}^W\boldsymbol{t}_{C(i)}$, where ${}^W\boldsymbol{R}_{C(i)} = ({}^C\boldsymbol{R}_{W(i)})^T$ and ${}^W\boldsymbol{t}_{C(i)} = -({}^C\boldsymbol{R}_{W(i)})^{T\,C}\boldsymbol{t}_{W(i)}$. They can be represented using the camera center $\boldsymbol{C}_r$ as starting point and a directional unit vector $\boldsymbol{d}_r = \frac{\boldsymbol{P}_r - \boldsymbol{C}_r}{\|\boldsymbol{P}_r - \boldsymbol{C}_r\|}$:

$$\boldsymbol{r}_r = \boldsymbol{C}_r + \lambda_r \boldsymbol{d}_r = \boldsymbol{C}_r + \lambda_r \frac{\boldsymbol{P}_r - \boldsymbol{C}_r}{\|\boldsymbol{P}_r - \boldsymbol{C}_r\|} \tag{4.68}$$

Again, the associated midpoint $\boldsymbol{X}_w$ can be computed, which is closest in average to all s rays. Therefore, following expression has to be minimized:

$$\sum_{r=1}^{s} \|\boldsymbol{C}_r + \lambda_r \boldsymbol{d}_r - \boldsymbol{X}_W\|^2 \tag{4.69}$$

As stated by Sturm et al., this linear least squares problem may be solved using the Pseudo-inverse [168]:

$$\begin{bmatrix} \boldsymbol{X}_W \\ \lambda_1 \\ \vdots \\ \lambda_s \end{bmatrix} = \begin{bmatrix} n\boldsymbol{I} & -\boldsymbol{d}_1 & \cdots & -\boldsymbol{d}_s \\ -(\boldsymbol{d}_1)^T & 1 & & \\ \vdots & & \ddots & \\ -(\boldsymbol{d}_s)^T & & & 1 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} & \cdots & \boldsymbol{I} \\ -(\boldsymbol{d}_1)^T & & \\ & \ddots & \\ & & -(\boldsymbol{d}_s)^T \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1 \\ \vdots \\ \boldsymbol{C}_s \end{bmatrix} \tag{4.70}$$

Finally, these results can be further improved by using the Levenberg-Marquardt iteration to minimize following functional:

$$\sum_{r=1}^{s} \left\| \boldsymbol{K} \left[{}^C\boldsymbol{R}_L | {}^C\boldsymbol{t}_L \right] {}^L\boldsymbol{T}_{W(r)} \begin{bmatrix} \boldsymbol{X}_W \\ 1 \end{bmatrix} - \begin{bmatrix} \boldsymbol{x}_{C(r)} \\ 1 \end{bmatrix} \right\|^2 \tag{4.71}$$

It must be noted that the described reconstruction approach solves the general case that several fiducials are visible in a single laparoscope image but some may be missing, similar to the work of Nicolau et al. [133], where two stationary cameras are used to provide a stereo view of fiducials. This general approach was implemented in order to remain flexible with respect to the utilized camera or endoscope. Future advances of digital and high-definition endoscopes may provide a laparoscope camera, which is able to see several fiducials in a single view with a very high resolution. However, currently a laparoscope with an analog camera and therefore a more restricted reconstruction protocol is utilized: The laparoscope is hold close to a fiducial, so that no other fiducial can be seen at the same time, and is pivoted around this fiducial. The fiducial is segmented in each image and its position is added to a list of 2D points, which is in turn used to reconstruct the 3D position of the fiducial, as described above. This process is repeated for all other fiducials.

4.2.2 Point Matching and Registration

After the reconstruction of all 3D points from the 2D points of their associated lists, they need to be matched with the points segmented in the CT data set. Therefore, the correct point correspondences need to be identified and the transformation from the CT coordinate frame into the world coordinate frame, where the patient is registered in, needs to be computed. This is done by a distance-weighted graph matching approach along with a point based registration algorithm.

The graph matching algorithm in use is adapted from the work of Gold and Rangarajan, who combine the concepts of softassign, graduated nonconvexity, and sparsity [57]. Here, the reconstructed 3D points are matched to the segmented 3D points from CT, each represented as nodes in one of two graphs. The number of nodes is A and I, respectively. In each graph, all nodes are connected to each other by edges, which are weighted by the normalized Euclidean distances between the 3D points corresponding to the nodes. In this way, two symmetric adjacency matrices $\boldsymbol{G}$ and $\boldsymbol{g}$ (with elements G_{ab} and g_{ij}, respectively, and dimensions A and I, respectively) can be generated with the normalized Euclidean distances as their elements. The graduated assignment algorithm uses these matrices to find the best matching nodes. The matching nodes are stored in the match matrix $\boldsymbol{M}$, a square permutation matrix holding zeros and ones, whose rows and columns add up to one.

A major advantage of the algorithm is that a node in one graph can match to at most one node in the other graph, meaning even if wrong or additional 3D points were reconstructed or segmented before, the algorithm would still be capable of filtering these out by just not assigning a corresponding node. This is achieved by extending $\boldsymbol{M}$ with an extra row and column containing so-called slack variables [57]. This extended matrix $\hat{\boldsymbol{M}}$ is of dimension $(A+1, I+1)$, where one of its elements $\hat{M}_{ai}$ describes the probability that a node a from one graph corresponds to node i from the other graph. An overview of the algorithm is provided below (see algorithm 1).

Finally, using the corresponding point lists, a point based registration algorithm (see section 4.1.3.2) is applied to register the patient's CT volume in the same coordinate frame as the patient. An alternative approach for the computation of this final registration, which considers noise in 2D as well as in 3D data, was introduced by Nicolau et al. [132]. Based on radio-opaque fiducials, they propose an extended 3D/2D criterion to register a 3D model generated from preoperative CT data to the patient for augmented reality guided radio frequency ablation.

Algorithm 1: The graduated assignment algorithm of Gold and Rangarajan [57].

```
Initialization: $\beta = 0.5$, $\beta_f = 10.0$, $\beta_r = 1.075$, $I_{Bmax} = 4$, $I_{Cmax} = 30$, $\epsilon_B = 0.5$,
$\epsilon_C = 0.05$, $\hat{M}_{ai} = 1 + \epsilon_C$
repeat (Loop A)
    $I_B \leftarrow 0$
    repeat (Loop B)
        $I_C \leftarrow 0$
        for $a \leftarrow 1$ to $A$ do
            for $i \leftarrow 1$ to $I$ do
                $M^B_{ai} \leftarrow M_{ai}$
                $Q_{ai} \leftarrow \sum_{b=1}^{A} \sum_{j=1}^{I} M_{bj} C_{aibj}$, where
                $C_{aibj} = \begin{cases} 0 & \text{if } G_{ab} = 0 \text{ or } g_{ij} = 0 \\ 1 - 3\left\|G_{ab} - g_{ij}\right\| & \text{otherwise} \end{cases}$
                $M_{ai} \leftarrow \exp(\beta Q_{ai})$
        repeat (Loop C)
            for $a \leftarrow 1$ to $A + 1$ do
                for $i \leftarrow 1$ to $I + 1$ do
                    $\hat{M}^C_{ai} \leftarrow \hat{M}_{ai}$
                    $\hat{M}_{ai} \leftarrow \frac{\hat{M}_{ai}}{\sum_{i=1}^{I+1} \hat{M}_{ai}}$ (Row Normalization)
                    $\hat{M}_{ai} \leftarrow \frac{\hat{M}_{ai}}{\sum_{a=1}^{A+1} \hat{M}_{ai}}$ (Column Normalization)
            $I_C \leftarrow I_C + 1$
        until $I_C > I_{Cmax}$ or $\sum_{a=1}^{A} \sum_{i=1}^{I} \left\|\hat{M}^C_{ai} - \hat{M}_{ai}\right\| < \epsilon_C$
        $I_B \leftarrow I_B + 1$
    until $I_B > I_{Bmax}$ or $\sum_{a=1}^{A} \sum_{i=1}^{I} \left\|M^B_{ai} - M_{ai}\right\| < \epsilon_B$
    $\beta \leftarrow \beta_r \beta$
until $\beta \geq \beta_f$
Cleanup Heuristic: Set one element in each column of $\hat{\boldsymbol{M}}$ to 1, all other elements
to 0
```

4.3 Hybrid Tracking

Several tracking approaches are fused to achieve a more reliable and stable determination of the position and orientation of the flexible tip of the ultrasound transducer. The two-sensor-one-body-concept described in section 4.3.1 allows to determine electromagnetic field distortions to warn the surgical staff. To also correct erroneous measurements of the flexible transducer tip, the measurements can be constrained to lie on one of the poses described by the mathematical model of the transducer tip introduced in section 4.1.4.3. Furthermore, it can be supposed that the projection of the electromagnetically tracked axis of the ultrasound tip has to exactly match the axis as seen in the laparoscope image (4.3.2). Fusing all this information, an improved superimposition accuracy can be obtained.

4.3.1 Electromagnetic Distortion Estimation

Intraoperatively, every measured pose of the rigid sensor of the ultrasound transducer is transformed applying equation 4.58 and its corresponding error δ is determined using equation 4.59. If the error is bigger than the previously computed distrust level δ_{max}, the surgical staff is automatically warned. Such errors are often caused by dynamic or static field distortions. Whenever the occurrence of an error is determined, it is visualized by drawing a red frame around the ultrasound plane that is augmented on the camera image. Otherwise the frame is drawn in green. Additionally, as the flexible sensor is in close proximity to the rigid one, its measurements will be most likely affected by field distortions as well.

Simple Error Correction In order to approximate a correction of erroneous measurements of the flexible sensor, a simple approach is to multiply/add the rotational/translational part of the deviation between the previously hand-eye calibrated (“calib”) and the measured (“meas”) transformation of the rigid sensor to the measured flexible sensor transformation, all relatively to the fixed optical tracking (OT, world) reference frame:

$$\begin{aligned} {}^{OT}\boldsymbol{R}_{FlexS(corr)} &= {}^{OT}\boldsymbol{R}_{RigidS(meas)}{}^{T} \cdot {}^{OT}\boldsymbol{R}_{RigidS(calib)} \cdot {}^{OT}\boldsymbol{R}_{FlexS(meas)} & (4.72) \\ {}^{OT}\boldsymbol{t}_{FlexS(corr)} &= -{}^{OT}\boldsymbol{t}_{RigidS(meas)} + {}^{OT}\boldsymbol{t}_{RigidS(calib)} + {}^{OT}\boldsymbol{t}_{FlexS(meas)} & (4.73) \end{aligned}$$

An attempt to correct the error is optional and can be performed only e.g. when the distortion detection was successful. Activated error correction can be visualized by drawing a yellow frame around the ultrasound superimposition.

Model Based Error Correction Especially dynamic field distortions, e.g. caused by moving instruments close to the sensors, are very difficult to reproduce and often distort the electromagnetic field at the rigid sensor in a completely different direction and magnitude than at the flexible sensor, even though they are placed close to each other.

Therefore, also the mathematical model describing the flexible tip of the ultrasound transducer can be used (cf. 4.1.4.3) to improve the correction method. This can be done by mapping erroneous measurements of the flexible sensor onto the “hemisphere” of possible sensor movements relatively to the rigid optical tracking body, which is attached to the transducer shaft and not influenced by electromagnetic field distortions. Therefore, only the hand-eye calibrated transformation ${}^{RigB}\boldsymbol{T}_{RigS}$ and the tip model parameters (compare section 4.1.4.3) are needed, when optimizing ψ and ϕ by the Levenberg-Marquardt algorithm.

4.3.2 Image Based Transducer Tip Tracking

As the intrinsic and extrinsic camera parameters of the laparoscope and hence the spatial location of the image plane are known, another approach to improve the tracking accuracy of the flexible transducer tip is to automatically localize the tip in the images of the laparoscope camera and adjust the superimposition of the B-scan plan accordingly. A correction transformation can be calculated, which can further improve the results of the model based error correction method, especially to correct errors introduced by calibration inaccuracies. As for above methods to correct tracking errors, this approach also provides additional information, which is entirely optional and can be made available upon request of the surgeon.

It must be noted that the utilized segmentation techniques based on simple edge detection are by far not as sophisticated as other methods described in the literature [30, 36, 98, 109] and will very likely fail in the operating room. However, the presented method is only a proof of concept and will be further addressed in future work to make it more robust and to also include additional features such as the determination of the end of the transducer tip.

Line Detection To find the 2D image coordinates of the transducer tip axis, OpenCV is utilized to automatically segment the transducer tip axis in the undistorted laparoscope images. First, the Canny algorithm [24] is applied to provide a binary image of edges, which is fed into a Hough transform to give a set of lines in the camera image. For each line, its end points are obtained.

To find the two lines corresponding to the two edges of the transducer tip, the whole set of segmented lines is first back-projected into 3D space, i.e. each end point $\boldsymbol{x}_C$ given in image coordinates (pixels) gets back-projected to $\boldsymbol{X}_C$ given in camera coordinates (millimeters).

$$\boldsymbol{X}_C = \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \begin{bmatrix} X_C \\ Y_C \\ 1 \end{bmatrix} = \boldsymbol{K}^{-1} \begin{bmatrix} \boldsymbol{x}_C \\ 1 \end{bmatrix} \tag{4.74}$$

Together with the camera center, each line represented by its two end points $\boldsymbol{X}_{C1}$ and $\boldsymbol{X}_{C2}$ forms a plane, which can be defined by its normal $\boldsymbol{n} = \boldsymbol{X}_{C1} \times \boldsymbol{X}_{C2}$, scaled to unit length.

All planes are now compared to the measurements of the transducer tip axis (which is defined by $\boldsymbol{b}_{FlexS}$ and $\boldsymbol{d}_{FlexS}$), acquired by electromagnetic tracking and transformed

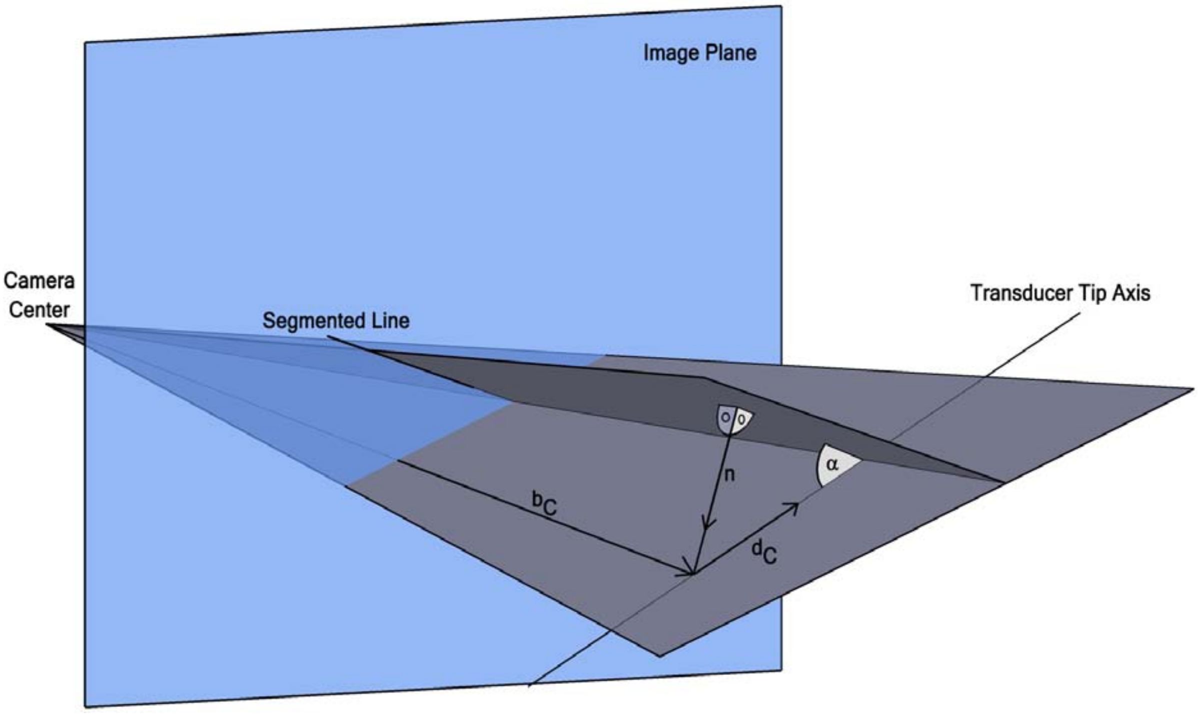


Figure 4.15: Back-projection of a segmented line and its comparison to the transducer tip axis.

into camera coordinates (after applying the model based distortion correction method proposed in the previous section):

$$\begin{aligned} \boldsymbol{b}_C &= {}^{C}\boldsymbol{T}_{FlexS}\boldsymbol{b}_{FlexS} \\ \boldsymbol{d}_C &= {}^{C}\boldsymbol{T}_{FlexS}\boldsymbol{d}_{FlexS} \end{aligned} \tag{4.75}$$

where ${}^{C}\boldsymbol{T}_{FlexS} = {}^{C}\boldsymbol{T}_{L}{}^{L}\boldsymbol{T}_{W}{}^{W}\boldsymbol{T}_{TransB}{}^{TransB}\boldsymbol{T}_{EMT}{}^{EMT}\boldsymbol{T}_{FlexS}$.

An angle α between the measured transducer tip axis and each plane can be determined by

$$\alpha = \arcsin(\boldsymbol{n} \cdot \boldsymbol{d}_C) \tag{4.76}$$

The distance d between the base point of the measured transducer tip axis and the plane is described by

$$d = \boldsymbol{n} \cdot \boldsymbol{b}_C \tag{4.77}$$

Depending on whether d is positive, negative, or zero, the base point $\boldsymbol{b}_C$ of the measured transducer tip axis will be above (on the half-space, the normal is pointing to), beneath, or on the plane.

For each line, $|\alpha|$ and $|d|$ are compared to a certain threshold, e.g. $\alpha_{thresh} = 5$ (degrees) and $d_{thresh} = 30$ (millimeters). If both parameters are below the corresponding threshold, it can be assumed that the current line corresponds to an edge on the transducer tip.

Correction Transformation Iterating over all lines belonging to a transducer tip edge, the minimum and maximum distances d_{min} and d_{max} between the base point of the measured transducer tip axis and the plane described by the back-projection of a line are stored. Ideally, the difference between previously computed maximum and minimum distances $|d_{max} - d_{min}|$ needs to be equal to the diameter of the ultrasound transducer tip, which is 10 millimeters. If this difference stays within certain limits, say 10 ± 2 mm, it can be assumed with high probability, that lines were extracted, which belong to both edges of the transducer. These $i = 1 \dots n$ lines are included into any further computations, along with the mean plane normal $\bar{\boldsymbol{n}} = \frac{\sum_{i=1}^{n} \boldsymbol{n}_i}{\|\sum_{i=1}^{n} \boldsymbol{n}_i\|}$ and the mean angle $\bar{\alpha} = \frac{\sum_{i=1}^{n} \alpha_i}{n}$ between transducer tip axis and plane.

The distance d_{est} between segmented transducer axis and electromagnetically measured transducer axis can be estimated as the average of the minimum and maximum distances $d_{est} = 0.5(d_{max} + d_{min})$.

When translating the measured transducer tip axis along the mean plane normal $\bar{\boldsymbol{n}}$ by the estimated distance d_{est}, the axis origin will be in the middle of the segmented transducer tip. Next, the tip axis needs to be rotated into the plane. Since the rotation axis $\boldsymbol{r}$ has to be orthogonal to the plane normal as well as to the measured tip axis direction, it can be computed as $\boldsymbol{r} = \bar{\boldsymbol{n}} \times \boldsymbol{d}_C$. Together with the mean angle $\bar{\alpha}$ between measured tip axis and plane, a homogeneous correction transformation can be estimated. This transformation maps the electromagnetically measured tip axis to a pose, from where it can be projected onto the image plane in such a way that it is exactly aligned with the segmented axis of the transducer tip.

4.4 Augmented Reality Visualization

In order to provide a medical augmented reality solution for various applications including laparoscope augmentation, the CAMPAR framework was jointly developed within our department over the past three years [161]. While guaranteeing intraoperability between different hardware components, it at the same time provides reliable data synchronization and visualization usable and suitable to the operating room.

4.4.1 Reliable Synchronization

An augmented reality visualization is not possible without certain hardware components such as tracking systems or imaging devices, which usually have unsynchronized clocks running at different speeds. As already introduced in section 2.2.2.4, a lack of synchronization of tracking systems and imaging devices can be a major cause of superimposition errors [68] – the point of time of data acquisition is at least as important as the corresponding value of the data. The CAMPAR framework is able to provide a synchronization mechanism between all components and can also check in real time, whether the visualization system currently provides the desired synchronization accuracy. Its synchronization technique is based on the network time protocol (NTP) and is integrated into the system core.

The software collects all input data and marks them with a time stamp of the point in time, when the data was measured. This should not be mistaken with the point in time when the data arrived at the visualization system. Even if certain hardware components neither provide time information nor hardware triggering, an accurate synchronization is still possible by generating own time stamps and utilizing temporal calibration algorithms (cf. section 4.1.4.2) to estimate the travel time from data generation to data arrival at CAMPAR.

In this way, time lags between different components are eliminated as error sources and only data from the same point of time is visualized, which can be also very helpful for hybrid or redundant tracking methods such as magneto-optic tracking.

It is important to note that this synchronization is not a compromise that slows down the system. Since the software is aware of all components and thus of the component with the slowest update rate, it can use this information to update the visualization at the earliest possible point of time, when all essential data are available.

Monitoring and control of the data flow of all components is possible at run time providing a temporal accuracy of below a millisecond, as a high precision timer has been integrated [134]. Even if some software or hardware components are distributed to several remote PCs, the augmented reality architecture allows to keep track of time offsets with an accuracy in the magnitude of one to just a few milliseconds (This lower precision is mainly caused by the variable network lag and can be kept low in a dedicated local area network connecting only hardware components required for the augmented reality system). Therefore, every remote PC serves as NTP server[8], so CAMPAR can regularly update its local time offsets to all remote PCs.

4.4.2 Usability

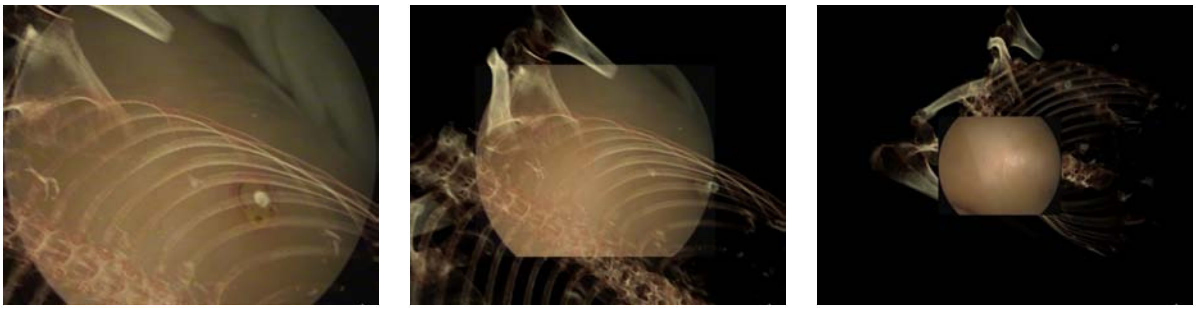
As for every image-guided surgery system, its general aim should be to support the surgical team to concentrate on their actual work, while still taking advantage of the latest imaging and navigation components. This design goal always has to be kept in mind during development of a software framework for augmented reality visualization. The visualization of complex intraoperative imaging data has to be easy to use and must be adapted to the needs of the physicians, for instance to be switched on and off easily during surgery, as it will typically not be needed longer than just a few minutes during complicated situations. Visualization modes can be limited to only comprise a low set of selectable states, so sterilizable interaction devices such as touch screens or switchable input devices can be utilized for the augmented reality visualization.

Visualization Only visualization algorithms have been selected that do not need much manual interaction by physicians. Common visualization paradigms such as orthogonal slices, arbitrary 3D slices, volume rendering, and instrument based navigation provide an intuitive start to any new system, since they are well known among surgeons and do not

---

[8] It is advisable to use an open source NTP implementation, as e.g. available for Windows from MEINBERG Funkuhren GmbH & Co. KG, Bad Pyrmont, Germany (`http://www.meinberg.de/english/sw/ntp.htm`), to properly configure the NTP server.

require prior segmentation by a physician. For example, using the Cg shader language[9], 3D volumes acquired by CT or a mobile C-arm can be directly rendered in real time as view-aligned 3D textures and augmented directly on the automatically undistorted live laparoscope images. This direct visualization technique does not require much interaction time, since time-consuming segmentation, as it would be needed for surface-based rendering, can be completely avoided. It is also possible to virtually zoom in or out the laparoscope image, with its superimposition enabled or disabled (see figure 4.16). To visualize certain tissues or anatomy only, default transfer functions can be loaded, for instance tailored to contrasted vessels. These can still be fine-tuned by interactively assigning specific colors and opacities to the rendered textures, provided by a graphical transfer function editor.



(a) Volume rendered overlay.

(b) Virtually zoomed overlay.

(c) Virtual zoom with an overlay only around the laparoscope image.

Figure 4.16: Superimposition of CT data onto the laparoscope view of a human cadaver.

System Lag and Frame Rate As stated by Ware and Balakrishnan [190], system lag and frame rate can greatly influence the task performance. The system lag of CAMPAR is below 100 milliseconds with a frame rate restricted to the slowest system component, which is in the case of analog laparoscope cameras 25 Hz for PAL and 30 Hz for NTSC images, respectively. Psychological research suggests that this speed offers an efficient interaction with a visualization system.

4.4.3 Interoperability

The environment of an advanced visualization system must be borne in mind, i.e. it must be usable in the operating room and not be yet another system with some functionality. For a smooth integration into the operating room and intuitive usability for the surgical team, the system must interact with other (imaging and computer aided) systems in the operating room and access their data and functionality. Current operating room hardware does not have a common standardized user interface. A unified user interface with a single visualization system per operating room but not per imaging or navigation device

---

[9] `http://developer.nvidia.com/object/cg_toolkit.html`

should be desired. This ideal visualization system should have access to all computerized procedures in the operating room and provide a single user interface to the surgeon instead of a display and a user interface for each piece of hardware. Therefore, the CAMPAR framework is designed to operate with differing hardware. This means, that imaging standards such as DICOM or networking standards such as TCP/UDP, but also special data handling routines are integrated. This is also one of the goals of the DICOM Working Group 24, which aims to develop DICOM objects and services related to image-guided surgery [93].

As important as abiding by standards is an abstract programming interface for similar hardware, e.g. for the tracking systems provided by A.R.T., NDI, and Ascension Technology Corporation. CAMPAR allows for interaction with similar hardware of different vendors in exactly the same way by introducing abstractions of the provided data such as pose information from tracking systems. This idea appears to be obvious; however the existing solutions in the operating room do not usually support this kind of collaboration. Only research frameworks such as the open source toolkits IGSTK [28] and MITK [201] mostly incorporate such interfaces.

The data synchronization routines for data from different computers offer a smooth collaboration without changing or opening devices in the operating room. The open design of the software framework uses standard programming interfaces like OpenGL, OpenInventor, Cg, DICOM, and NTP. This offers compatibility to common and stable libraries.

4.4.4 Implementation

The CAMPAR framework is implemented in C++ to fulfill its high real-time requirements and possibility to address hardware as directly as possible while keeping a clear object-oriented design. To guarantee a high visualization quality and performance, the core of the graphics routines is based on OpenGL, a reliable and stable industry standard for computer graphics. Also OpenGL based scene graph engines such as OpenInventor can be used. To further speed up the visualization of huge medical imaging data, for instance during direct volume rendering, the processing unit of the graphic card (GPU) is utilized by the integration of shader languages such as Cg.

The most common format for all kinds of medical imaging data is DICOM, which can be easily processed within the framework via the gdcm library. OpenGLUT or Qt can be chosen as the graphical user interface (and even be exchanged by just modifying a single line in the configuration file). All external libraries can be downloaded and used for free for non-commercial purposes.

4.4.5 Flexibility

The concept of the software framework proved to be a strong base for projects aiming at the operating room. Various applications, demonstrations, and evaluations based on the framework have been presented to the scientific community, both for monoscopic visualization in laparoscopy [47, 48, 129] and molecular imaging [195, 196] and for stereoscopic visualization in trauma surgery, based on head mounted displays [160, 172, 174]. Its maturity could be shown in two cadaver studies [173], ex vivo studies [46, 194], as well as in vivo animal experiments [45].

CHAPTER

FIVE

EXPERIMENTS AND EVALUATION

In order to evaluate the augmented reality visualization for laparoscopic surgery, various offline, phantom, ex vivo, and in vivo experiments have been performed in close collaboration with surgeons. The main focus of all studies was on usability and especially accuracy.

Many factors play a role regarding the usability of an image-guided surgery system, for example its integration into the surgical workflow (including sterilization issues, differences to conventional procedures in terms of time and complexity), patient outcome, costs, or accuracy. One must be aware that accuracy requirements very much depend on the type of application. While, for instance, port placement only requires an accuracy of about two centimeters, liver resection requires a more stringent accuracy between five and ten millimeters for a fine preparation or achievement of secure resection margins or an identification of vessels.

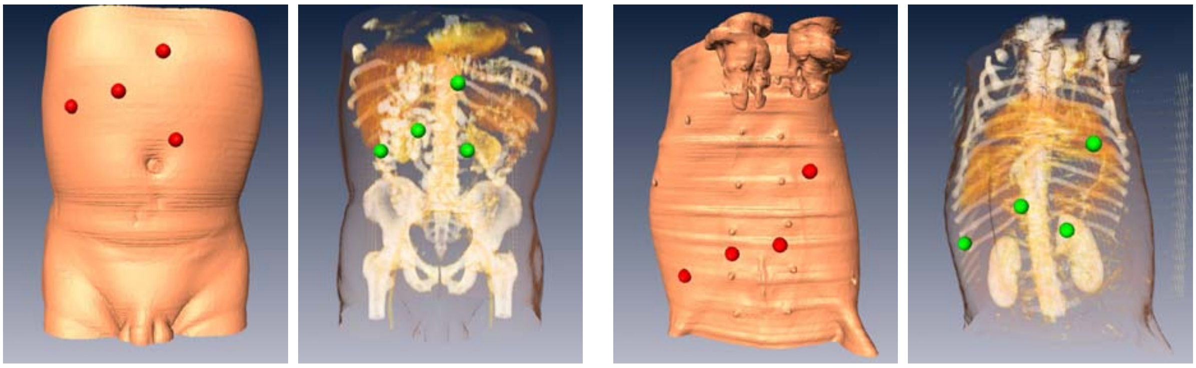
5.1 Port Placement

The first iteration of experiments was solely dedicated to port placement. As shown by offline studies conducted with four surgeons on preoperative data, port placement may be performed more confidently and accurately by a visualization aid. To evaluate the accuracy of this visualization aid, preoperative imaging data was overlaid on the laparoscope video on both a rigid phantom and later on in in vivo studies to analyze the influence of deformations of the anatomy between CT acquisition and surgery.

5.1.1 Offline Studies

To evaluate the necessity of precise port placement planning, offline experiments on human and porcine CT data were conducted. For the human data set a typical venous CT data set was chosen, for the porcine data set one of two data sets acquired for in vivo experiments was chosen (cf. sections 5.1.3 and 5.2.3). Four surgeons, who are all experienced in laparoscopic surgery, were asked to place four typical ports for liver resection (one laparoscope and three instrument ports), based on two different visualization modes, which were shown consecutively.

In the first mode (cf. left images of figures 5.1a and 5.1b), only orthogonal 2D slices and a 3D view of the skin were shown to the surgeons to simulate their conventional way of placing ports (except for palpation). In the second mode (cf. right images of figures 5.1a and 5.1b), the skin was made transparent, so the surgeons were able to see the abdominal interior, based on previously described volume rendering techniques (cf. section 4.4). To independently place a set of four ports in each data set (represented by colored spheres), the surgeons simply had to click inside the 3D view.



(a) Human CT data set.

(b) Porcine CT data set.

Figure 5.1: Offline port placement experiments. The four ports were chosen using two different visualization modes (opaque, transparent). All views are from the same perspective.

The Euclidean 3D distances between corresponding ports chosen with the first and the second visualization method were calculated. For the patient data set, distances between 6.57 mm and 66.20 mm were computed with an RMS of 40.42 mm and a standard deviation (SD) of 17.59 mm. For the pig, the distances were between 6.38 mm and 48.74 mm with an RMS of 22.28 mm and an SD of 11.07 mm. Although the distances of corresponding ports in the porcine data set seem to be much smaller, for a direct comparison they would need to be scaled, as the pig was relatively young (3-6 months, 16-20 kg), while the patient was adult.

Interpreting these numbers, it can be confirmed that, in comparison to the standard port placement technique, a visualization aid considering individual patient anatomy significantly altered port placement both in human and animal data sets. All surgeons agreed that the presented visualization method supports the surgical procedure and port placement accuracy.

5.1.2 Accuracy Evaluation on Rigid Phantoms

To evaluate the accuracy of hand-eye calibration and augmented reality visualization, very first experiments were performed on rigid phantom data.

It must be noted that the accuracy evaluation performed here was done in a very early stage of the dissertation work. At this point of time, both synchronization as well as tracking methods were very rudimentary. For instance, only two tracking cameras were used and only a wooden block was used as optical tracking body with rather old markers and a suboptimal marker configuration (cf. figure 5.3). This introduced additional errors, which could be avoided in later iterations on accuracy evaluation.

5.1.2.1 Hand-eye Calibration

Initially, both implemented hand-eye calibration algorithms (i.e. based on the works of Tsai/Lenz and Daniilidis, respectively) were evaluated. Therefore, the intrinsic and extrinsic parameters of the laparoscope camera were computed from 32 frames, each acquired at a different station within the optical tracking volume. To estimate the performance of hand-eye calibration relative to the number of stations, $n = 3 \dots 32$ stations and the corresponding $M = 3 \dots 496$ (remember that $M = n(n-1)/2$) unidirectional motions between them were used to compute the hand-eye formulation $\boldsymbol{AX} = \boldsymbol{XB}$, i.e. in this specific case the transformation from the laparoscope body to the camera frame ${}^{C}\boldsymbol{T}_{L}$. This means that 29 different matrices for ${}^{C}\boldsymbol{T}_{L}$ were computed.

To validate these matrices, the positions of nine retroreflective spherical markers were reconstructed from six laparoscopic images using the algorithm described in section 4.2. These reconstructions were compared to the “ground truth” measurements of the 3D positions of the nine markers provided by the optical tracking system. The average distance of the reconstructed points to the measurements of the tracking system was computed for each of the 29 transformation matrices. As visualized in figure 5.2, the best reconstruction results with errors between 1.4 and 2 millimeters were achieved, when 10 to 25 stations were used for hand-eye calibration, which also means for camera calibration that images and tracking data from at least 10 different laparoscope poses should be acquired. No significant difference in the performance of the two evaluated hand-eye calibration methods was observed.

5.1.2.2 Laparoscope Augmentation

To determine the augmentation error during port placement in a rigid environment, 13 CT visible spherical fiducials with a diameter of four millimeters (CT-SPOTS®, Beekley Corporation, Bristol, CT, USA)[1] were attached to a plastic thorax phantom containing a heart model (The Chamberlain Group, Great Barrington, MA, USA)[2]. After a CT scan and segmentation of all fiducials, the phantom was placed arbitrarily. The 3D positions of four fiducials were reconstructed automatically by moving the tracked laparoscope around them, using three to four images from differing poses for each fiducial. The other nine fiducials were just used later for validating the augmentation from many different viewing directions, so in practice they are not needed. Next, the CT-to-tracking transformation of the four fiducials was computed.

---

[1] `http://www.beekley.com/`

[2] `http://www.thecgroup.com/`

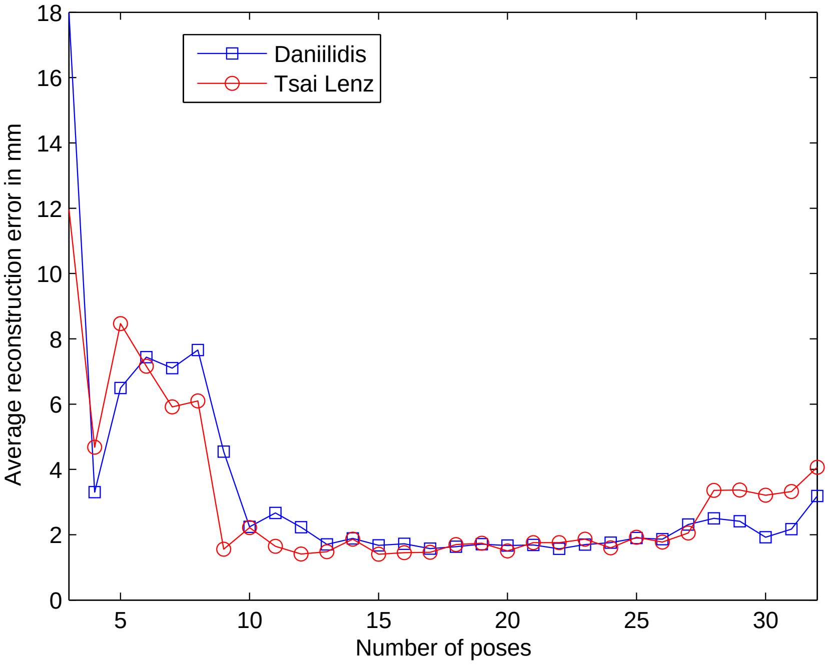


Figure 5.2: Average reconstruction error for retroreflective markers depending on the hand-eye calibration method and number of poses

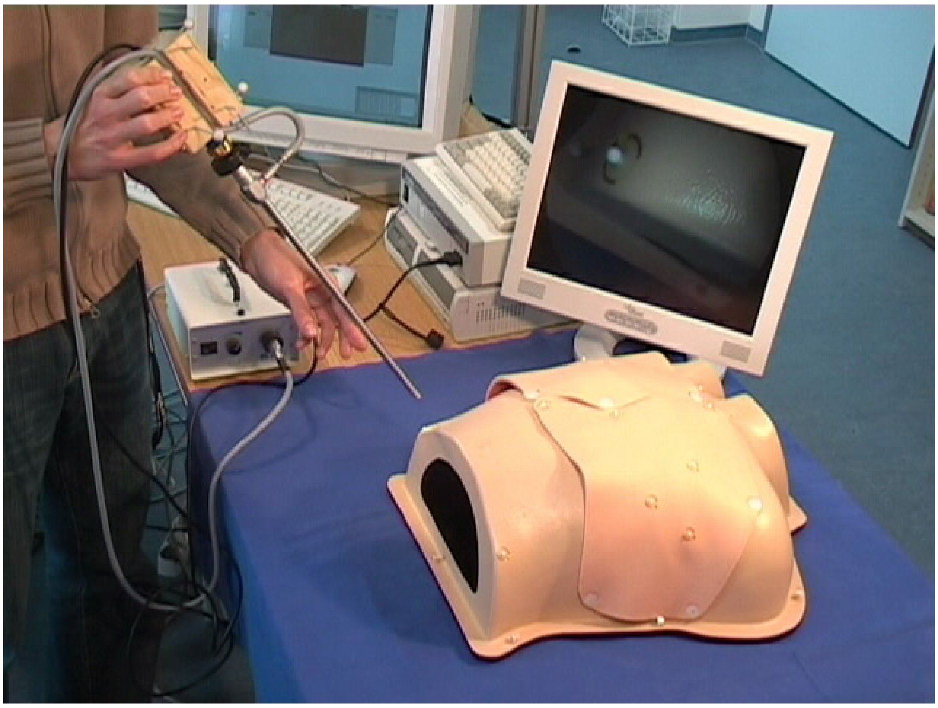


Figure 5.3: Very first evaluation of the port placement accuracy on a static phantom.

Having the sensor-to-camera and CT-to-tracking transformations as well as intrinsic camera parameters and distortion coefficients, the laparoscopic images can be undistorted and the CT volume can be augmented on the images. To verify the augmentation, the distances of all 13 fiducials from the real images to a semi-transparent augmentation in an orthogonal view were measured in pixels and scaled to millimeters. A scaling is possible, since the real diameter of a fiducial is known in millimeters. An RMS error of 2.97 mm could be assessed (with an SD of 1.43 mm). This is fully sufficient for a precise port placement, where an accuracy in the range of two centimeters is required. The proposed 3D reconstruction method was also compared to a pointer based approach. Therefore, a calibrated pointing device tracked by the optical tracking system was used to record the positions of the four fiducials. Again, the CT-to-tracking transformation was computed and used for the augmentation. With this method, an RMS error of 3.36 mm was achieved (with an SD of 1.11 mm), i.e. the reconstruction based method performs at least as good as the pointer based one. The comparison of augmentation errors is visualized in figure 5.4.

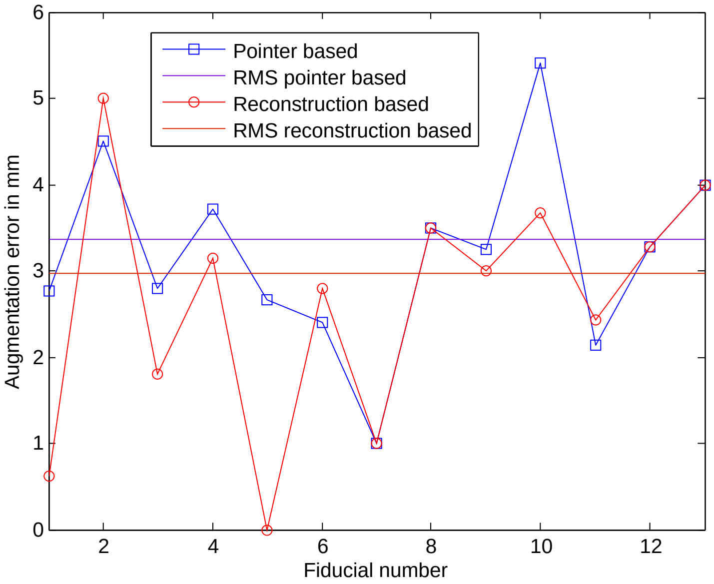
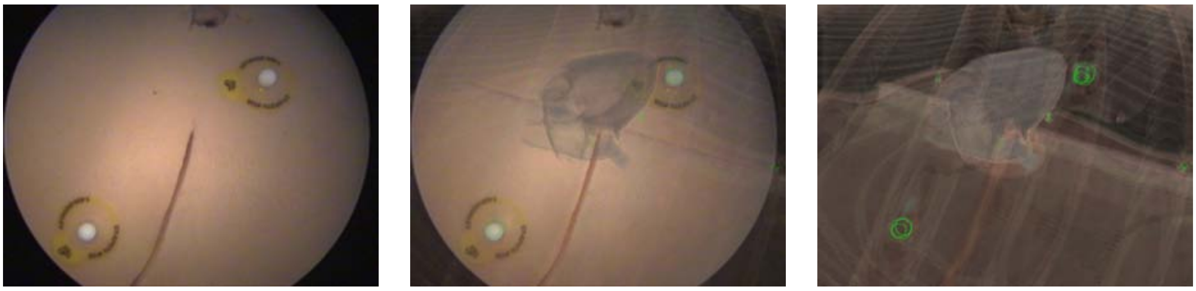


Figure 5.4: The augmentation error for all 13 fiducials measured directly in the video.

A port placement application was integrated into the CAMPAR framework offering three visualization modes, as displayed on figure 5.5. In the first mode, the undistorted real laparoscopic image is displayed. The second mode additionally augments the volume on the phantom in a half-transparent mode, so the accuracy of the overlay can be visually and qualitatively verified by the surgeon. In a third purely virtual mode the surgeon can switch the laparoscope optics from 30 degrees to 0 degrees and move the camera in and out the volume along the laparoscope shaft axis to validate a possible port. The superimposition of a 512×512×444 CT volume, undistortion of the 640×480 camera frames, and a final visualization at a resolution of 800×600 pixels was achieved in real time.



(a) Real (b) Augmented (c) Virtual

Figure 5.5: Visualization modes for the same laparoscope pose: (a): Real camera image, (b): Transparent augmented view outlining fiducials, ribs, and heart (the virtual green contours correctly match the white fiducials in the video image), (c): Purely virtual view, which can be used for port placement to move the camera in and out

5.1.3 In Vivo Porcine Studies

To evaluate the port placement method in a more realistic setup, a series of in vivo experiments was performed in a special animal operating room of our clinical partners. First, the system was evaluated for port placement planning based on preoperative CT data. Later on (see section 5.2.3), liver resection planning based on cone-beam CT data was evaluated.

Two live pigs of 16-20 kg were anesthetized by injections given in the auricular veins on the dorsal surface of the ears. Spherical fiducials were adhered to each pig's skin using DERMABOND adhesive (Ethicon GmbH, Norderstedt, Germany)[3] to ensure their fixation during transportation of the pig between animal operating room and CT scanner. The fiducials were positioned on the abdominal wall in areas such that they are not moving too excessively during breathing. The pig was carried to the radiology department, where 50 ml contrast agent (SOLUTRAST® 300) was administered at a flow of 3 ml/s. CT data was acquired in both arterial and venous phases. Four fiducials were segmented in the venous CT data set. Back in the animal operating room, the pig was connected to an artificial respiration system to keep its breathing constant. The pig's vena jugularis was cannulated, so that it could be used for future contrast injections during cone-beam CT acquisition.

The tracked laparoscope was moved around the fiducials to reconstruct their 3D positions, which took approximately two minutes. Matching them automatically to the four fiducials previously segmented in CT, the pig could be registered to its venous CT data set. The CT data set was augmented directly on the laparoscope by means of direct volume rendering. Since the pig was placed in a different position than during the acquisition of the CT, the fiducials were displaced to a certain extent. Using four fiducials for registration, an RMS error of 11.05 mm, a maximum error of 17.11 mm, and an SD of 4.03 mm were calculated, which complies with previous reports on port placement accuracy of Falk et al. [43]. Even though, due to the large displacement of the subject between

---

[3] `http://www.ethicon.de/`

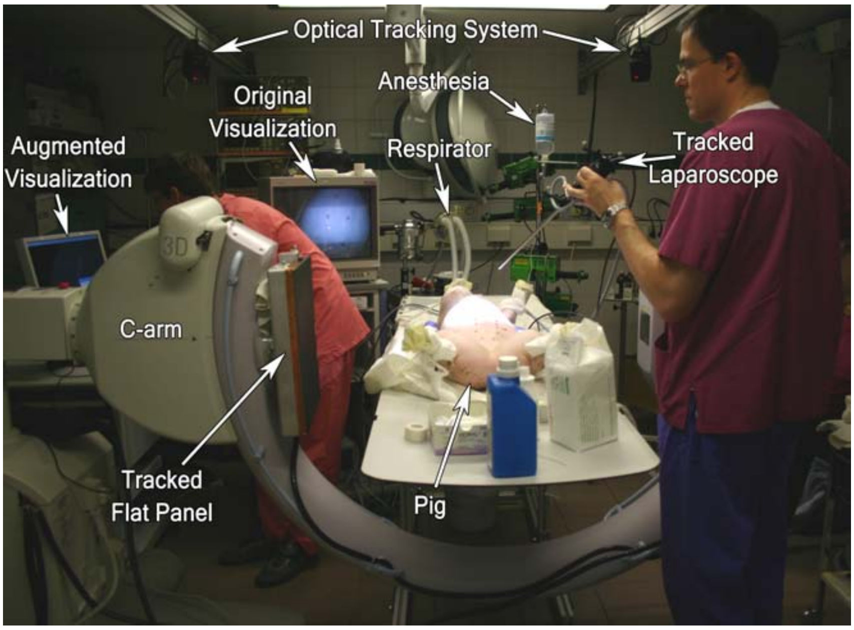


Figure 5.6: In vivo studies. The tracked laparoscope defines the viewing direction of the virtual visualization along the instrument axis for port placement planning.

preoperative and intraoperative imaging, the CT data set can not be used directly for a precise intraoperative augmentation during resection planning, it still fulfills the rather low accuracy requirements of port placement (about two centimeters). Switching to a co-registered virtual visualization with 0° optics (cf. figure 5.7), the bones and all contrasted sections such as aorta, kidneys, and hepatic veins were visualized in respect to the instrument axis.

All ports were planned successfully and the trocars could be placed appropriately, as confirmed by the two participating surgeons, who were interviewed on the advantages and disadvantages of the augmented reality visualization system. Both surgeons confirmed that the augmentation for port planning is integrated smoothly into the surgical workflow and gives valuable as well as sufficiently accurate information on trocar port placement.

5.2 C-Arm Based Vessel Visualization

Analogously to the port placement studies, the evaluation of superimposing intraoperative C-arm data on the laparoscopic view was also first performed on rigid phantoms and later on complemented by ex vivo and in vivo studies.

5.2.1 Accuracy Evaluation on Rigid Phantoms

To evaluate the overall accuracy of the presented registration-free system in a rigid environment, two types of experiments were performed on non-deformable phantoms. First, the navigation accuracy with a tracked pointer was assessed, second, the augmentation

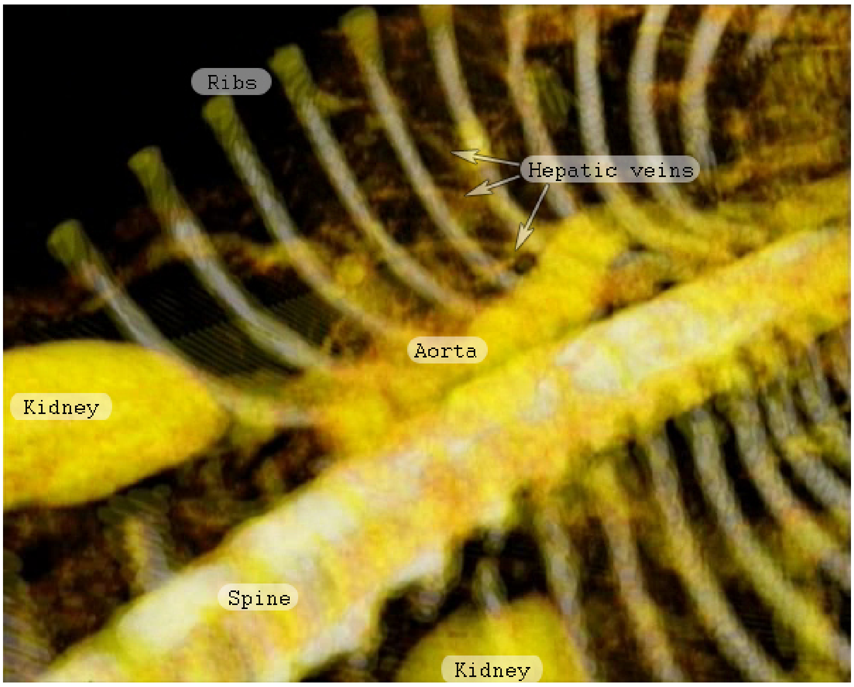


Figure 5.7: Virtual visualization for port placement during in vivo experiments, which enables the surgeon to see important anatomy – annotations added manually (Note that these are static frames of dynamic video sequences, which provide better perception of anatomical structure to the surgeons).

accuracy when overlaying C-arm reconstructed volumes on the images of the tracked laparoscope was estimated.

5.2.1.1 Navigation Error

To determine the obtainable navigation accuracy with the registration-free system, a cubic evaluation phantom developed by Ritter et al. was used, which contains cylindrical bore holes of varying depth [144], as shown in figure 5.8. The phantom was adhered to the operating table with double-sided adhesive tape. Then the varying depth of 20 bore holes was measured in terms of the 3D position of the pointer tip in the world coordinate frame of the optical tracking system. The bottom of all bore holes could be entirely reached, as the pointer has a sharp tip. The measured tip coordinates were transformed into the volume coordinate system by ${}^{I}\boldsymbol{T}_{F}{}^{F}\boldsymbol{T}_{W}{}^{W}\boldsymbol{T}_{P}$.

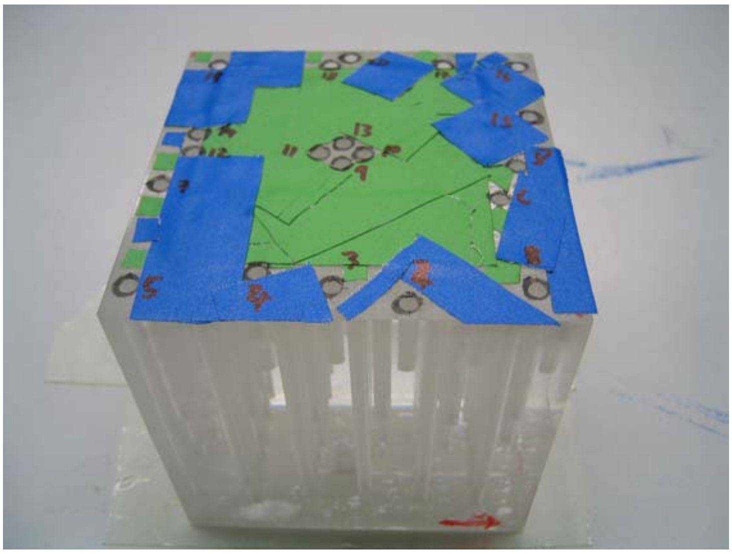


Figure 5.8: Cubic evaluation phantom containing cylindrical bore holes of varying depth.

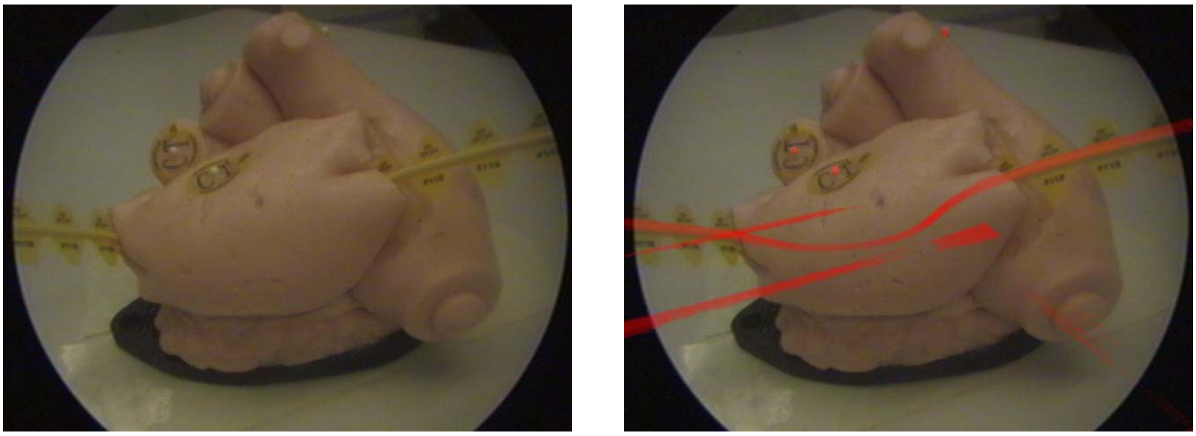
Now each of the 20 bore holes was filled with a single BB of radius $r = 1.5$ mm. Afterwards, a volume was reconstructed, wherein the centroids of the BBs were extracted automatically by a hybrid segmentation algorithm based on thresholding, region growing, the shape of the BBs, and weighing of voxel intensities. The segmented coordinates of all BBs were compared to the measured coordinates.

Since the bore holes have an inclination α of 56°, the distance from a BB centroid to the end of the bore hole equals $r/\cos\alpha$. This distance offset needs to be applied to all segmented centroids to be theoretically aligned with the tip of the pointer.

The overall Euclidean RMS error between the measured and segmented coordinates of BB centroids was 1.10 mm, which confirms previous results of Ritter et al. [144]. A second experiment after a complete system recalibration using BBs of 1.0 mm radius resulted in a reproducible Euclidean RMS error of 1.05 mm.

5.2.1.2 Augmentation Error

For the determination of the laparoscopic augmentation error, a volume of a plastic model of a heart with three adhered spherical fiducials of 2.3 mm diameter and a curved line marker (CT-SPOTS®, Beekley Corporation, Bristol, CT, USA) representing a vessel was reconstructed. The laparoscope was moved around this plastic model arbitrarily. Augmenting the volume on the live laparoscope video, images were taken from a large number of views covering the interventional working space. On all images the fiducials and their augmented counterparts, visualized in a different color, are visible (cf. figure 5.9). The absolute distance in millimeters was measured from the midpoint of a spherical fiducial and its corresponding virtual projection. The obtained RMS error was 1.78 mm, the maximum error 4.14 mm, and the SD 1.12 mm. This rather high error was caused by an inappropriate configuration of the markers on the laparoscope (cf. figure 3.2a), leading to unstable tracking, e.g. by partial or complete occlusions of the markers. This behavior could be anticipated, since the calibrated camera center is located about 350 mm away from the markers, leading to a large extrapolation displacement for small angular or translational tracking errors of the markers.



(a) Real, undistorted laparoscope image showing the 3 spherical fiducials and a curved line marker on the plastic heart.

(b) Augmented laparoscope image. Volume rendered are only the fiducials; the additional lines are artifacts in the reconstructed volume.

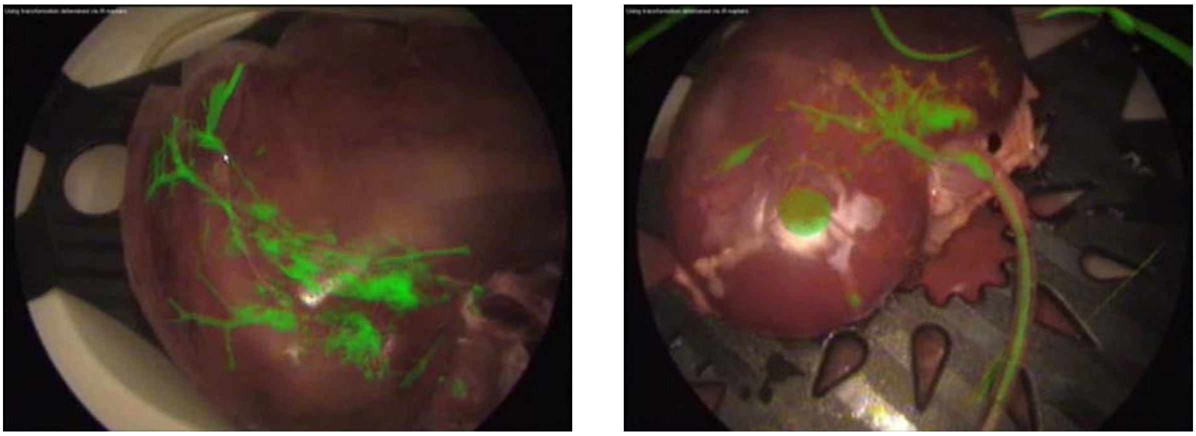
Figure 5.9: Plastic heart used for the determination of the augmentation error.

Therefore, in a second experiment the four markers on the laparoscope body were rearranged, so they could be seen optimally by the tracking system in almost every possible position (cf. figure 3.2b), following the heuristics described by West and Maurer [198] (For a detailed description of the equations and errors involved in tracking distal tips of long instruments and optimal marker design refer for example to recent results presented by Fischer and Taylor [50] as well as Bauer et al. [10].). After a recalibration of the laparoscope, the augmentation accuracy from all positions of the laparoscope was significantly improved. The RMS error was decreased to only 0.81 mm, the maximum error to 1.38 mm, and the SD to 0.44 mm.

During surgery however the full flexibility of the laparoscope is required to change viewing directions, i.e. the laparoscope shaft shall be able to be freely rotated against the camera head. Hence two new marker targets were designed and attached to the laparoscope (cf. figure 3.2c). After re-calibration (including an additional calibration of the oblique scope – see section 4.1.1.3), the augmentation accuracy when rotating the laparoscope head was determined. The RMS error was 1.58 mm, the maximum error 2.81 mm, and the SD 0.74 mm.

5.2.2 Ex Vivo Perfusion Studies

The second series of experiments for C-arm based image superimposition was conducted on a freshly harvested whole porcine liver and a defrosted whole ovine kidney. In both cases, the organ was placed in a perfusion box (Pulsating Organ Perfusion Trainer, OPTIMIST Handelsges.m.b.H., Bregenz, Austria)[4]. Our collaborating surgeons catheterized the portal vein of the liver and the renal artery of the kidney, respectively. Then the iodinated nonionic contrast agent SOLUTRAST® 300 was administered into the organ. It was diluted in normal saline and conveyed into the organ by the pump of the perfusion system. A C-arm acquisition was started immediately and a 3D volume of the organ was reconstructed. In another experiment, the pure contrast agent was directly injected.



(a) Augmented porcine liver. The speckled areas are reconstruction artifacts caused by a high concentration of contrast agent.

(b) Augmented ovine kidney. The big spot is the plug of the perfusion system, which apparently has a similar Hounsfield unit like the contrast agent.

Figure 5.10: Ex vivo vessel augmentation.

The contrasted part was clearly visible in the reconstructed volume for both cases, directly as well as via the perfusion system. The vessels were directly augmented on the laparoscope view by volume rendering utilizing pre-defined transfer functions in such a way that only voxel intensities of the contrast agent are rendered. The vessels could be accurately overlaid for most laparoscope positions, as one can see in figure 5.10b in the case of the kidney, where the real and virtual catheters appear to perfectly overlap. In

---

[4] `http://www.optimist.at/`

the case of the porcine liver, only the vessels of one lobe could be contrasted partially, as some liver vessels were accidentally cut at the abattoir while harvesting.

Although in general more vessels could be visualized for the direct administration of contrast agent, artifacts appeared during the 3D volume reconstruction due to the high concentration. Further studies have to be conducted to find an optimal contrast agent concentration for an artifact-free reconstruction and a clear visualization of all relevant vessels.

5.2.3 In Vivo Porcine Studies

Motivated by the positive results of the ex vivo studies, the augmented reality visualization system was tested in vivo. This is the first time that a system capable of augmenting intraoperative cone-beam CT data on the laparoscope view was tested on animals. As described in section 5.1.3, two pigs were anesthetized and their vena jugularis was cannulated for contrast injections during cone-beam CT acquisition.

Figure 5.11: In vivo porcine studies for cone-beam CT superimposition.

As the previous accuracy evaluation on the rigid phantom already gave good results, one focus of these studies was to analyze any deviations from the correct augmentation caused by deformations of the liver due to breathing. Additionally, the feasibility of contrast enhanced C-arm reconstructions for laparoscope augmentation was evaluated in terms of timing and protocol of contrast agent administration as well as volume rendering capabilities of intraoperatively contrasted and reconstructed liver vessels.

After insufflation of CO_2, two spherical fiducials of 4 mm diameter were glued to the surface of the liver using a fibrin pad, as shown on figure 5.12a. The artificial respiration system was temporarily turned off in exhalation. While injecting about 130 ml of contrast agent into the vena jugularis (and the auricular vein for the second pig, respectively), 100 orbiting C-arm projections were acquired consecutively within 32 seconds. A volume was reconstructed and transferred to the navigation workstation. As plastic trocars were used and all instruments were removed for the scan, reconstruction artifacts could be largely avoided. Since laparoscope and C-arm are both co-registered in the tracking coordinate system, the transferred volume could be immediately augmented on the laparoscopic view.

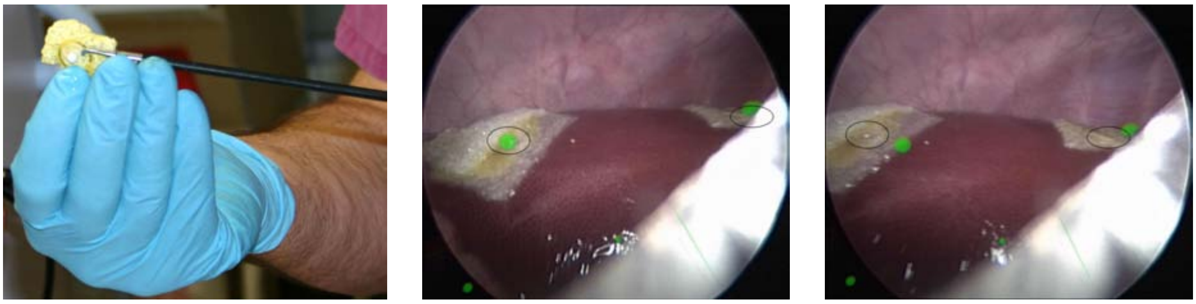

(a) Fibrin pad wrapped around a fiducial, which is glued to the liver surface.
(b) Intrinsically registered augmented fiducials as seen during exhalation.
(c) Displaced fiducials as seen during inhalation.

Figure 5.12: Fiducials adhered to the liver surface – black ellipses added manually to outline the current (actual) fiducial positions.

First, the movement of the liver caused by breathing was analyzed in the permanently augmented laparoscope images. As visualized on figures 5.12b and 5.12c, a deformation of about 1 cm could be confirmed between exhalation and inhalation, as found previously by Herline et al. [66]. Also the expiration and inspiration plateaus were reproducible within approximately 1 mm, as stated by Balter et al. and Wong et al. [8, 202]. Consequently, the final augmentation error can be approximated to be between 2 mm (during expiration) and 12 mm (during inspiration), depending on the current respiratory phase.

Next, the contrasted hepatic veins were augmented on the surface of the liver to plan the resection, using an appropriate transfer function for direct volume rendering (cf. figure 5.13). After successful planning, the visualization was switched off and the surgeon started the resection.

Finally, the two participating surgeons were interviewed on the assets and drawbacks of the augmented reality visualization system. Both surgeons agreed on the successful validation of the visualization in an in vivo model. Since the contrast in the peripheral hepatic veins was too low, for the augmentation during liver resection planning only the major vessels could be visualized. Therefore, the timing and protocol of contrast agent administration needs to be further improved. Furthermore the amount of contrast agent needs to be reduced to a maximum of 50 ml, which is the default dose for a human patient. As this was one of the first experiments with intraoperative cone beam reconstruction of contrasted liver, these problems were expected.

The surgeons did not ask for the synchronization of the pig's respiration and pulse to the laparoscopic augmentation. Knowing that the augmentation is correct only at expiration, both surgeons claimed not to be distracted by the continuous augmentation during breathing, even without additional fiducials showing the augmentation offset. They were

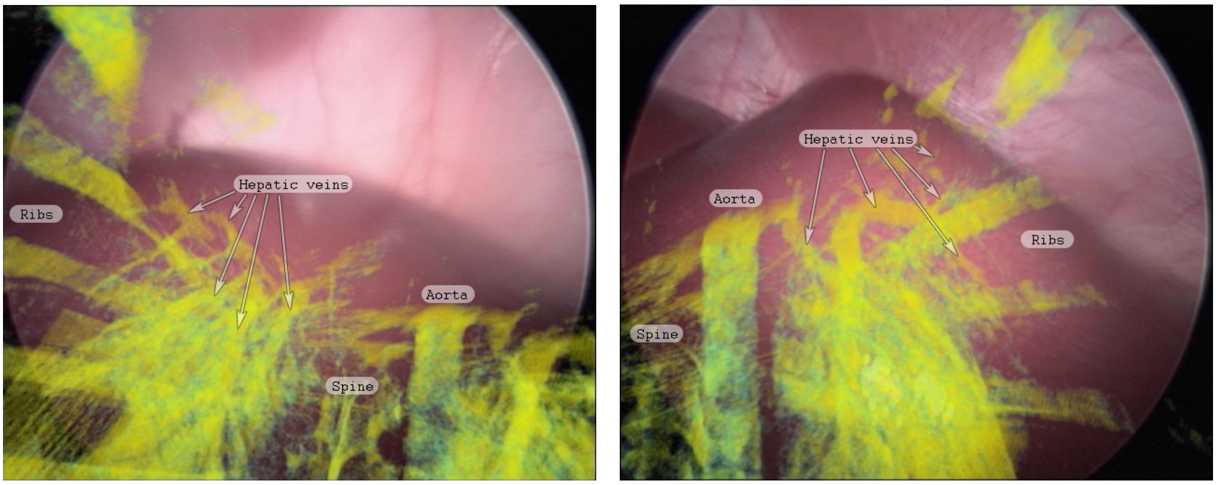


Figure 5.13: In vivo augmentation of contrasted hepatic veins and aorta as well as ribs and spine on the laparoscope – annotations added manually (Again note that these are static frames of dynamic video sequences, which provide better perception of anatomical structure to the surgeons, who are used to look at 2D laparoscopic projections during the operation).

satisfied with the current augmented visualization, which could allow them to perform appropriate intraoperative resection planning. The surgeons however asked for intraoperative segmentation of the contrasted vessels so that the system can provide them with the exact (metric) positioning of tracked instruments relative to the vascular structure at expiration.

5.3 Ultrasound Visualization

While the augmented reality visualization approaches based on preoperative and intraoperative (cone-beam) CT were already successfully validated in vivo, the ultrasound visualization is ongoing research, which is currently evaluated in a laboratory setup, where a precisely defined working environment can be guaranteed, which is important to ensure no or repeatable electromagnetic field distortions and a specified working volume. For all experiments, the electromagnetic tracking measurements were acquired in a restricted volume of 20–36 cm for x, and ±15 cm for y and z, to avoid too many outliers due to lower field excitation.

Two sensors of 1.3 mm diameter were used for tracking the ultrasound transducer, and an additional one for calibrating the transducer axis.

5.3.1 Ultrasound Calibration Error

After acquiring 40 flexible sensor poses and their corresponding lines that were automatically detected in the B-scan images, the ultrasound calibration matrix was computed using the Levenberg-Marquardt optimizer. To determine the ultrasound calibration accuracy, a single electromagnetic sensor with tip coordinates given in the electromagnetic tracking frame was submerged into the water bath. Its tip was segmented manually in 5 regions of the B-scan plane, which was repeated for 4 poses of the transducer differing from the ones used during calibration. The coordinates of the pointer tip were transformed into the B-scan plane coordinates and compared to the segmented tip coordinates (scaled to millimeters). An RMS error of 1.69 mm with standard deviation of 0.51 mm and maximum error of 2.39 mm was obtained.

5.3.2 Augmentation Error

In order to estimate the laparoscope augmentation errors automatically, an additional optical tracking body ("flexible body") was temporarily attached to the transducer tip and co-calibrated to the flexible sensor by another hand-eye calibration (cf. section 4.1 and figure 3.4). One marker of the flexible body was chosen as a reference and automatically segmented whenever visible in the laparoscopic video. Its center coordinates were compared to the projection of its respective optical tracking coordinates onto the image plane. Additionally, the corresponding electromagnetic tracking measurements as well as their approximated corrections (calculated by the simple correction method) were projected using the previously determined hand-eye calibration transformations.

Evaluation data was recorded using a laparoscope-to-marker distance of five to ten centimeters, which is a typical intraoperative working distance. The current distance can be recovered from optical tracking data and the camera calibration parameters. This information was also used to scale pixel units to millimeters.

For each of six evaluation series, the transducer was fixed at a different pose and the laparoscope was used to measure the projected distances from five differing poses, each in an undistorted and a distorted environment. To distort the electromagnetic field, two alternatives were evaluated. A metal plate was placed on the table to simulate primarily static distortions caused for instance by an operating table. For dynamic distortions, a steel rod of 10 mm diameter was brought close to the transducer to simulate a surgical instrument, changing its proximity and angle to the transducer in five measurements.

The RMS errors are given in figure 5.14. For each of the six series, the errors of the three distortion cases (no distortion, static, and dynamic distortion) were plotted, each scenario with the simple correction function enabled and disabled. While static interferences could be predicted and corrected with high reliability, dynamic distortions yielded even worse results when attempting a simple correction.

In order to evaluate the distrust function statistically, the distrust level (cf. equation 4.59) was computed for each of the poses. An offset between the segmented marker and the electromagnetic tracking projections of more than 2 mm was regarded as erroneous measurement. In this case, a distrust level δ of more than δ_{max} is expected (during hand-eye calibration, δ_{max} was empirically determined to be 20). The following cases were defined for the evaluation:

- A *true positive* is a measurement, in which the electromagnetic tracking error was above 2 mm with a distrust level of above 20 – the detector rejected an erroneous reading correctly.

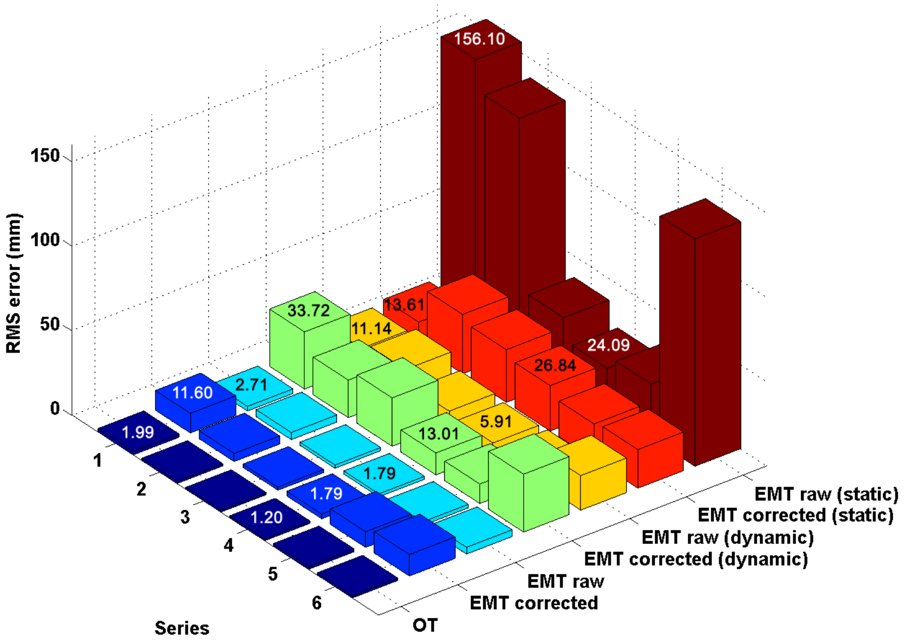


Figure 5.14: RMS projection errors.

- A *true negative* is a measurement, in which the electromagnetic tracking error was below 2 mm with a distrust level below 20 – the original electromagnetic tracking data was correctly accepted.
- A *false positive* (type 1 error) is a measurement, in which the electromagnetic tracking error was below 2 mm, but the distrust level above 20 – a correct value was not detected, so it was rejected without necessity.
- A *false negative* (type 2 error) is a measurement, in which the electromagnetic tracking error was above 2 mm, but the distrust level below 20 – the record was accepted although the real error was large.

The results are listed in table 5.15. In about 85 % of all cases, the true situation (*true positives* and *true negatives*) was correctly detected.

distortion		true	false
w/o:	positive	40.0%	10.0%
	negative	30.0%	20.0%
static:	positive	100.0%	0.0%
	negative	0.0%	0.0%
dynamic:	positive	73.8%	13.8%
	negative	12.4%	0.0%
avg:	positive	71.3%	7.9%
	negative	14.1%	6.7%

Figure 5.15: Distortion detection rate by the proposed distrust level.

To visually inspect the augmentation when overlaying the B-scan plane on the laparoscopic live video, a cylindric phantom containing straight wires was constructed, which extend through the walls of the phantom. It was filled with water of known temperature. Adjusting the pixel scaling factors to an adequate speed of sound, the B-scan plane was augmented, allowing the camera to view a wire on the augmented B-scan plane and its extension outside the phantom walls. A typical augmented laparoscope image can be seen on figure 5.16.

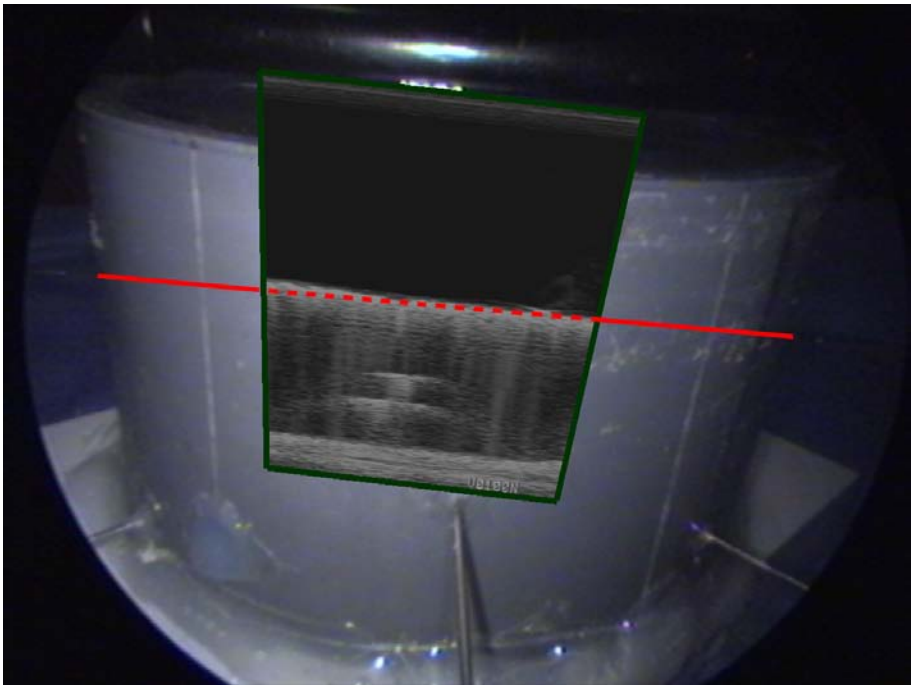
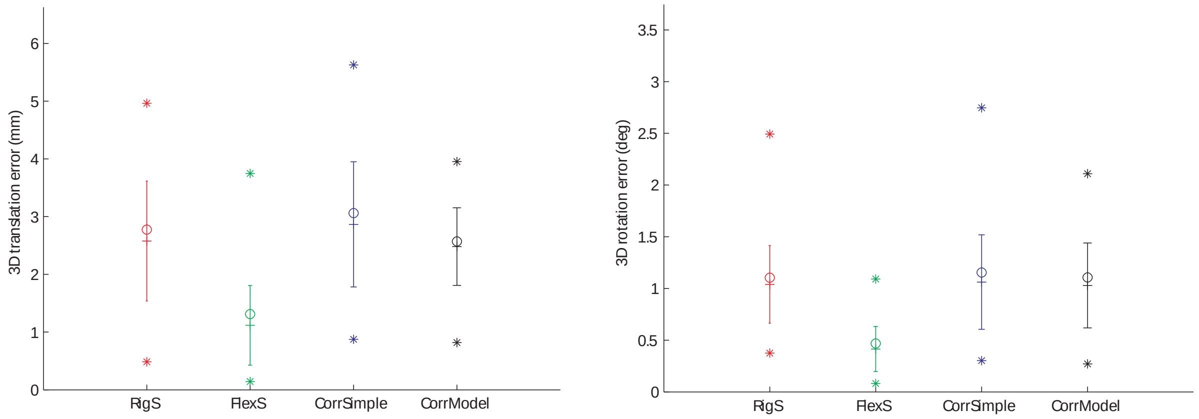


Figure 5.16: Ultrasound plane augmented on the laparoscope video – Red line added manually to visualize the extension of the straight wire.

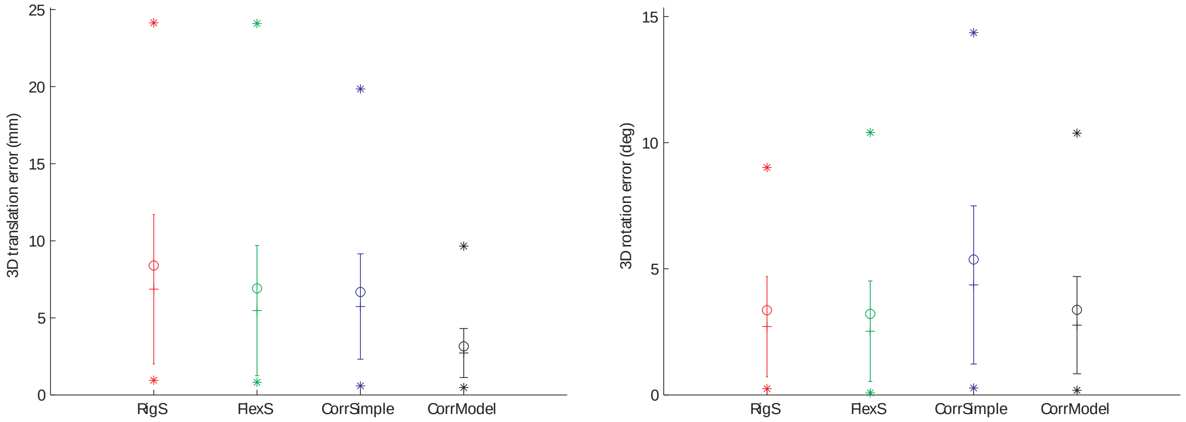
5.3.3 Model Based and Image Based Correction

To compare the performance of the two more advanced model based and image based error correction methods with the previously evaluated simple correction approach, which turned out not to be satisfying especially for no or dynamic field distortions, another series of experiments was performed. These experiments mainly focused on the evaluation of the real 3D errors obtained for the measurements of the two sensors attached to the transducer shaft (rigid sensor) and to its tip (flexible sensor) and for the corrected measurements of the flexible sensor, all in a distortion-free and a dynamically distorted electromagnetic tracking environment. An error computation entirely in 3D is possible, as for both the rigid and the flexible sensors also reference measurements via the rigid and the flexible optical bodies and their corresponding static (hand-eye calibrated) transformations into the sensor coordinate frames are available. Furthermore, the evaluation of the 2D augmentation errors, as described in section 5.3.2, was repeated, but this time additionally the performance of the model based and image based correction methods was tested.

3D Errors First, in a distortion-free environment 70 measurements of the flexible sensor and body as well as the rigid sensor and body were acquired for varying transducer poses. The Euclidean distance in millimeters between the sensor measurements and their corresponding reference measurements via the optical body was considered as translation error, while the rotation angle in degrees corresponding to the quaternion rotating the sensor pose into the reference pose was considered as rotation error. Furthermore, the simple and model based error correction methods were applied to the tracking data of the flexible sensor and compared to the reference measurements via the flexible body.



(a) Translation error (distortion-free environment). (b) Rotation error (distortion-free environment).

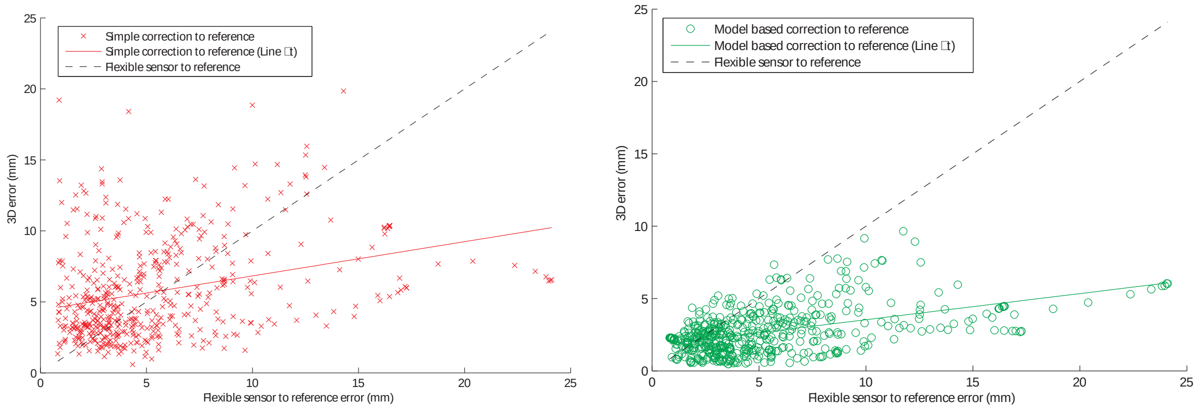


(c) Translation error (distorted environment). (d) Rotation error (distorted environment).

Figure 5.17: 3D errors measured at the rigid sensor (RigS), flexible sensor (FlexS), and corrected for the flexible sensor using the simple approach (CorrSimple) and the model based approach (CorrModel). The minimum and maximum error is each represented by a star, the RMS error by a circle, and mean and SD by a bar.

As can be seen in figures 5.17a and b, rigid sensor measurements compared to their reference measurements via the rigid body diverge more than the flexible sensor measurements. This is because the distance between the rigid body and the rigid sensor is much longer than between the flexible body and the flexible sensor, wherefore more errors propagate into the computation of the hand-eye calibration of the rigid sensor-body-transformation. It can also be seen that in the distortion-free case the model based correction method can be at most as good as the measurements of the rigid sensor, as the transducer model is built relatively to the rigid body. As already observed in the previous experiment, the simple correction method can even worsen results.

Thereafter, 500 measurements were acquired, again for varying transducer poses, but this time also dynamically distorted by miscellaneous ferromagnetic objects (steel rod, knife, and power supply unit). As visualized in figure 5.17c, the simple correction method is not able to improve the results, while the model based approach greatly reduces distortion errors in terms of translation (from 6.91 mm RMS and 24.10 mm maximum to 3.15 mm RMS and 9.65 mm maximum). Rotation errors, however, got worse with the simple approach and remained similar for the model based approach (cf. figure 5.17d). This means for the model based approach that either the weighting of the rotational part in equation 4.58 needs to be increased or the model is not yet accurate enough to represent rotations.



(a) Simple correction.

(b) Model based correction.

Figure 5.18: Residual 3D errors after a simple correction (left) in comparison to those after a model based correction (right), each in relation to the magnitude of initial errors between the flexible sensor measurements and the reference measurements via the flexible optical body.

Another observation can be made in figure 5.18, where the residual 3D errors after a simple correction are compared to those after a model based correction, each in relation to the magnitude of errors between the flexible sensor measurements and the reference measurements via the flexible optical body. Whereas the simple correction approach has a large error variance, the model based approach always stays within an error of 10 mm and mostly even within 5 mm, even for large initial errors of up to 25 mm. A similar statement is supported by figure 5.19. It shows the distances of the initial flexible sensor measurements to the reference measurements, compared to the distances between the initial flexible sensor measurements and their corrected counterparts that the model based correction method computed. In most cases, the differences between the distances stay within about 5 mm.

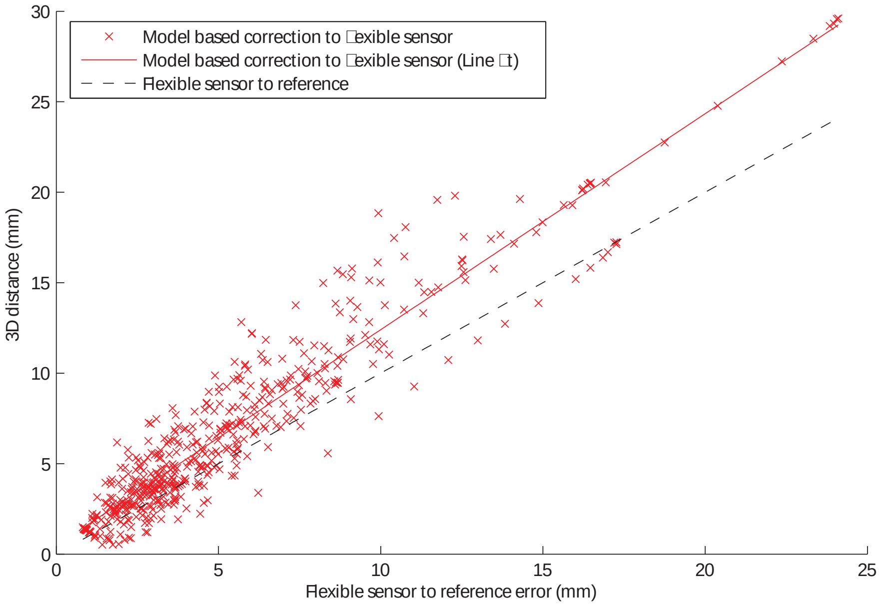


Figure 5.19: Distances between initial flexible sensor measurements and their (model based) corrected counterparts in relation to the magnitude of errors between the flexible sensor measurements and the reference measurements via the flexible optical body.

More plots confirming the superiority of the model based correction approach compared to the simple approach can be found in the work of Reichl [141].

2D Augmentation Errors In a final experiment, the 2D augmentation errors where evaluated when applying (1) the simple error correction method, (2) the model based correction approach, and (3) a hybrid approach initialized by the results of the model based correction and further refined by the image based approach (cf. section 4.3.2). A similar experimental setup was followed as for the error evaluation in section 5.3.2 except for the difference that a laparoscope-to-marker distance of five to 20 centimeters was maintained to make both the marker and the transducer tip visible (and hence segmentable) in the laparoscope image for a large number of laparoscope and ultrasound tip poses, altogether for 300 varying poses. Only dynamic distortions using the steel rod were enforced, which is capable of enforcing large field distortions of up to 60 mm. Knowing the laparoscope-to-marker distance and the camera geometry, the 2D errors were again scaled from pixel units to millimeters.

As visualized in figure 5.20, the simple correction method performs worst, while the model based approach highly improves the error correction. The best results were obtained by the combination of model based and image based error correction methods.

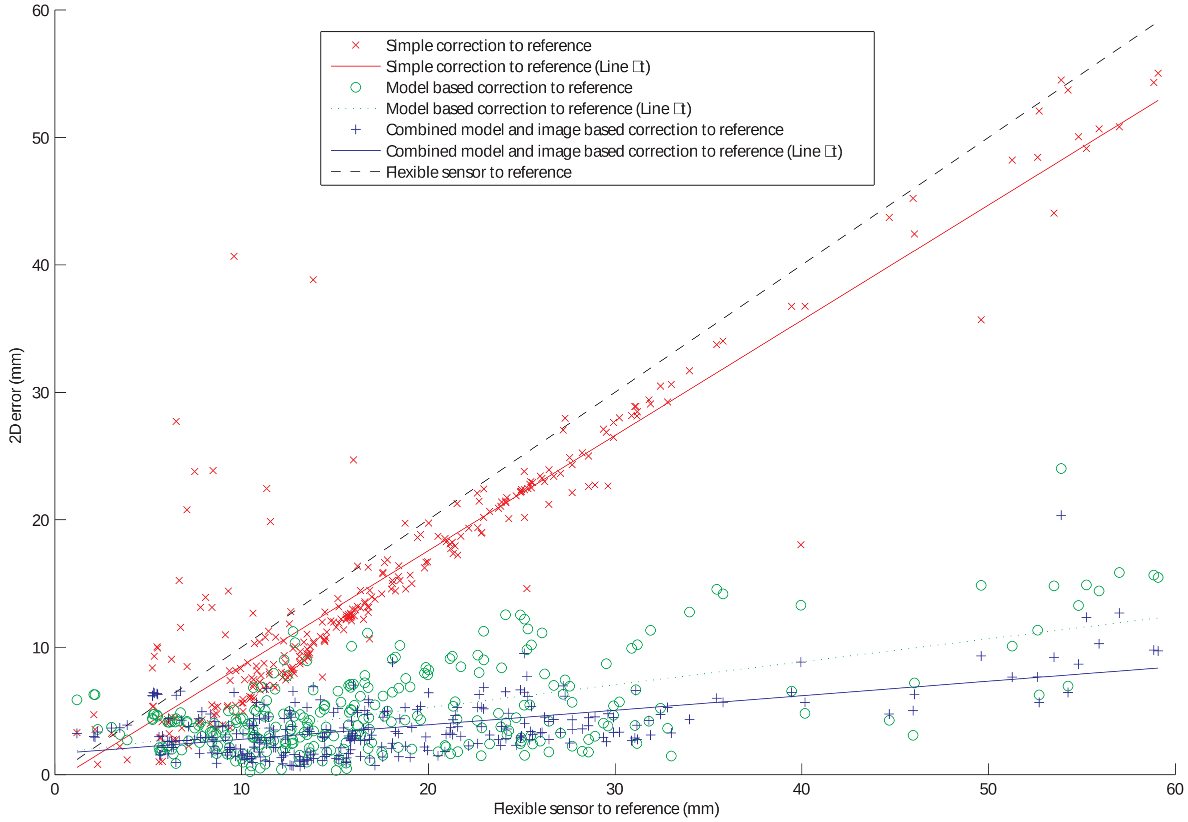


Figure 5.20: Residual 2D augmentation errors after a simple correction (red), a model based correction (green), and a combination of model based and image based correction (blue), each in relation to the magnitude of initial errors between the flexible sensor measurements and the reference measurements via the flexible optical body.

CHAPTER

SIX

CONCLUSION

This chapter summarizes the achievements made in the dissertation, followed by a discussion of current problems and a presentation of ideas for future work to further advance the laparoscopic augmented reality solution.

6.1 Summary

This work introduces the use of three independent intraoperative imaging systems, i.e. X-ray C-arm, laparoscopic ultrasound, and laparoscopic camera within a registration-free navigation framework enriched by augmented reality visualization. The combination of multiple intraoperative imaging modalities, tracking, and visualization allows to implicitly take into account organ movements between preoperative and intraoperative imaging, making this method especially useful for minimally invasive abdominal surgery.

For visualization during port placement planning, which relies on preoperative data and therefore requires manual registration, a largely automated method has been developed that can be soundly integrated into the current surgical workflow, because it is kept simple and still leaves decisions and control to the surgeon during port placement. Apart from the preoperative attachment of a few fiducials to the patient and a short and intuitive intraoperative patient registration procedure, the conventional clinical workflow is not altered. The method can be applied to any minimally invasive endoscopic procedure provided that preoperative patient data is available.

During the intervention, the intraoperative use of C-arm and ultrasound enables an advanced visualization for laparoscopic surgery, especially liver resection planning and navigation, without tedious manual registration.

This is the first time that intraoperative data of a co-registered mobile C-arm is augmented on the view of a laparoscope. Its application is not restricted to liver surgery, where it has been extensively evaluated in vivo. It may also be utilized in various other endoscopic interventions, for instance in tumor resection procedures such as partial adrenalectomy, partial pancreatectomy (especially endocrine neoplasms in the pancreatic tail), and partial nephrectomy, or in stone removal procedures such as nephrolithotomy or percutaneous nephrostolithotomy, where C-arm fluoroscopic guidance is used anyway during stent placement, tract creation, calyceal puncture, stone localization, and placement of the nephrostomy tube. The system is especially useful for difficult cases, where a tumor is located very close to or in between vessels or the hepatic veins. The result provided here could allow different methods proposed in the literature [39, 41, 94, 133] to utilize co-registered intra-operative cone beam reconstruction in order to improve their advanced visualization solutions.

Furthermore, the state of art [94, 127] for superimposing laparoscopic ultrasound images directly on the laparoscopic live images to give surgeons a better understanding of the spatial relationship between ultrasound and camera images is improved. In order to provide a trustable superimposition, a new method was presented to estimate electromagnetic field distortions online based on a hybrid magneto-optic tracking setup. By using two electromagnetic sensors and one optical tracking body, a distrust level of the current electromagnetic measurements is provided. Therefore, the system is able to automatically update and warn the surgical staff of possible inaccuracies. In laboratory studies, first online approaches for model based error correction were also successfully tested. In addition, as the laparoscopic ultrasound transducer tip is flexible, the proposed method could be applied to a large set of applications in abdominal surgery, e.g. liver, biliary tract, and pancreas.

All experiments and studies conducted in the course of this work show satisfying qualitative results for the improvement of intraoperative visualization in laparoscopic surgery.

6.2 Discussion and Future Work

Recent patient studies showed that the CT-SPOTS® pellets, which were used for patient registration during port placement, remain on the patient's skin for a quite limited time only, e.g. not overnight [89]. Therefore, either their positions need to be marked e.g. by a felt-tip pen, so they could be reattached before the intervention in case they were teared off or displaced, or they could be entirely replaced by flat fiducials such as CT-SPOTS® Crosshair or D-SPOTS® arrow (Beekley Corporation, Bristol, CT, USA), as used by Kutter et al. [89], or multi-modality markers (IZI Medical Products, Baltimore, MD, USA)[1], as used by Nicolau et al. [133]. They may guarantee a better adherence, even during moderate exercise. However, the image processing algorithms would have to be adapted to the new shape and color of these fiducials.

Addtional radiation and contrast agent exposure due to the intraoperative use of a mobile C-arm should not be disregarded. The trade-off of risks, benefits, and costs of such a system need to be considered for each case. For easy cases, already available data from preoperative CT or MRI may be sufficient to successfully perform an intervention, more difficult cases may justify additional radiation and contrast agent exposure.

While the presented augmented reality system has been successfully evaluated in vivo for port placement and C-arm based resection planning, the laparoscopic ultrasound subsystem is ongoing research and has only been evaluated in a laboratory setup. To analyze

---

[1] `http://www.izimed.com/`

and improve the stability of the methods for the error estimation and segmentation of the transducer shaft axis in the laparoscopic images, all algorithms need to be tested with tracking data and videos of real interventions in order to be adapted and extended to the new conditions. After this, the next logical step is to also set up in vivo experiments for the evaluation of the laparoscopic ultrasound augmentation and the error estimation methods.

The flat tablet transmitter recently presented by Ascension [151] may be an alternative to overcome electromagnetic field distortions, e.g. caused by the operating table. It is, however, far from providing as accurate measurements as the mid-range transmitter because of its lower excitation. In the same setup as described in sections 3.6 and 5.3, the flat tablet transmitter was used for ultrasound calibration, giving errors of about 4-8 mm for the 1.3 mm sensors (compared to errors of about 2 mm for the mid-range transmitter). Bigger sensors could be used to improve the accuracy, but this would probably require bigger trocars. Using 1.3 mm sensors, the total diameter of the laparoscopic transducer can be kept below 12 mm (including sterile cover), so it still fits a regular trocar.

An interesting idea is also to use the co-registered rigid electromagnetic sensor and optical tracking target attached to the shaft of the ultrasound transducer to generate a distortion function of the electromagnetic field on the fly, contrary to conventional electromagnetic tracker calibration techniques, which usually need to be performed before every new intervention [17, 27, 126, 203]. Additionally, an integration of methods to simulate electromagnetic errors would come in handy [50].

Not only the simulation of errors will be an important issue in future image-guided surgery systems. Also the in situ integration and visualization of uncertainty information, motivated e.g. by electromagnetic field distortions or (partial) occlusions during optical tracking, could be of great help to the surgical staff. All errors involved in an augmented reality visualization need to be modeled all their way of propagation, considering internal tracking parameters (electromagnetic field, camera parameters), receiver poses (transmitter, tracking cameras), as well as poses, geometry, and visibility of a tracking target (electromagnetic sensor, optical tracking body), as proposed by Bauer et al. and Sielhorst et al. for an optical tracking system [9, 10, 159].

Besides the visualization of navigation errors, more intuitive visualization methods need to be developed, so the surgeon better perceives overlaid data, while not even noticing the differences between real world and augmented reality any more. This may be achieved by applying advanced visualization techniques such as the ClearView system [85] to medical augmented reality for an improved fusion of virtual and real data, or by utilizing techniques to improve the depth perception during laparoscopy, e.g. by a virtual mirror [129]. Furthermore, an intraoperative segmentation of C-arm data sets would be of help to the surgeon for a more distinct differentiation of contrasted vessels, bones, and surrounding soft tissue, best performed automatically right after acquisition. Combining this segmentation with tracking of additional surgical instruments, important quantitative information for navigation could be provided, such as distances between vessels and instrument tip. As for the ultrasound transducer, tracking of these instruments may also be performed in a hybrid setup, probably extended by tracking methods based on 3D ultrasound image processing [135, 165].

To take full advantage of all visualization techniques, the surgical team requires adequate user interfaces. While the current system still depends on the use of keyboard and mouse, they need to be replaced by sterilizable and user-friendly interaction devices in future versions. This also requires an intensive collaboration with manufacturers of minimally invasive instruments, as interaction capabilities may be even integrated into their instruments. Similarly, manufacturers need to assist in the design of tracking sensors and bodies that can be smoothly integrated into laparoscopes and other surgical instruments.

Finally, a logical continuation of this work would be the extension of the proposed methods to other surgical imaging modalities and tools. To enhance future approaches for the tracking of flexible endoscopes, e.g. in bronchoscopy or NOTES, the fusion of electromagnetic tracking, image based tracking, and instrument specific models [87, 88] has a great potential. Furthermore, one should focus on the superimposition of data acquired by other emerging intraoperative imaging modalities. Fixed imaging systems such as stationary C-arms, open CT, or MRI scanners (for easy intraoperative access), and also movable systems such as mobile C-arms or the O-arm™ (Breakaway Imaging, LLC, Littleton, MA, USA)[2] will have a great impact on future surgical interventions. Tracking the X-ray source, gantry, or the patient table [143] of these systems, the multimodal imaging concepts presented in this dissertation can be applied and extended to further advance image-guided surgery in the operating room of the future.

---

[2] `http://www.breakawayimaging.com/`

APPENDIX
A

AUTHORED AND CO-AUTHORED PUBLICATIONS

i. M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, *Intra-operative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection*, To appear in IEEE Transactions on Medical Imaging (2007).

ii. M. Feuerstein, T. Reichl, J. Vogel, A. Schneider, H. Feussner, and N. Navab, *Magneto-Optic Tracking of a Flexible Laparoscopic Ultrasound Transducer for Laparoscope Augmentation*, To appear in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), N. Ayache, S. Ourselin, and A. Maeder, eds., vol. 4791 of Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007, Springer-Verlag, pp. 458–466.

iii. T. Wendler, M. Feuerstein, J. Traub, T. Lasser, J. Vogel, S. Ziegler, and N. Navab, *Real-Time Fusion of Ultrasound and Gamma Probe for Navigated Localization of Liver Metastases*, To appear in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), N. Ayache, S. Ourselin, and A. Maeder, eds., vol. 4792 of Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007, Springer-Verlag, pp. 252–260.

iv. N. Navab, J. Traub, T. Sielhorst, M. Feuerstein, and C. Bichlmeier, *Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures*, IEEE Computer Graphics and Applications, 27 (2007), pp. 10–14.

v. M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, *Registration-free laparoscope superimposition for intra-operative planning of liver resection*, in 3rd Russian-Bavarian Conference on Biomedical Engineering, Erlangen, Germany, July 2007, pp. 88–92.

vi. M. Feuerstein and N. Navab, *Registration-free augmentation device and method.* European Patent Application EP 06 009 222.8, filed April 11, 2007.

vii. N. Navab, M. Feuerstein, and C. Bichlmeier, *Virtual penetrating mirror device and method for visualizing virtual objects in endoscopic applications.* European Patent Application EP 06 009 221.0, filed April 11, 2007.

viii. R. Bauernschmitt, M. Feuerstein, J. Traub, E. Schirmbeck, G. Klinker, and R. Lange, *Optimal port placement and enhanced guidance in robotically assisted cardiac surgery*, Surgical Endoscopy, 21 (2007), pp. 684–687.

ix. C. Alcérreca, J. Vogel, M. Feuerstein, and N. Navab, *A new approach to ultrasound guided radio-frequency needle placement*, in Bildverarbeitung für die Medizin 2007, A. Horsch, T. M. Deserno, H. Handels, H.-P. Meinzer, and T. Tolxdoff, eds., Informatik aktuell, Springer Verlag, March 2007, pp. 26–30.

x. N. Navab, M. Feuerstein, and C. Bichlmeier, *Laparoscopic virtual mirror - new interaction paradigm for monitor based augmented reality*, in Virtual Reality, Charlotte, North Carolina, USA, March 2007, pp. 43–50.

xi. M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, *Registration-free laparoscope augmentation for intra-operative liver resection planning*, in Medical Imaging 2007: Visualization and Image-Guided Procedures, K. R. Cleary and M. I. Miga, eds., Proceedings of SPIE, San Diego, California, USA, February 2007.

xii. M. Feuerstein, K. Filippatos, O. Kutter, E. U. Schirmbeck, R. Bauernschmitt, and N. Navab, *A novel segmentation and navigation tool for endovascular stenting of aortic aneurysms*, International Journal of Computer Assisted Radiology and Surgery, 1 (2006), pp. 280–282.

xiii. T. Sielhorst, M. Feuerstein, J. Traub, O. Kutter, and N. Navab, *Campar: A software framework guaranteeing quality for medical augmented reality*, International Journal of Computer Assisted Radiology and Surgery, 1 (2006), pp. 29–30.

xiv. M. Feuerstein, S. M. Wildhirt, R. Bauernschmitt, and N. Navab, *Automatic patient registration for port placement in minimally invasive endoscopic surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), J. S. Duncan and G. Gerig, eds., vol. 3750 of Lecture Notes in Computer Science, Palm Springs, CA, USA, September 2005, Springer-Verlag, pp. 287–294.

xv. J. Traub, S. Wiesner, M. Feuerstein, H. Feussner, and N. Navab, *Evaluation of calibration methods for laparoscope augmentation*, in 4. Jahrestagung der Deutschen Gesellschaft für Computer-und Roboter-Assistierte Chirurgie (CURAC 2005), Berlin, Germany, September 2005.

xvi. M. Feuerstein, N. Navab, E. U. Schirmbeck, S. M. Wildhirt, R. Lange, and R. Bauernschmitt, *Endocsope augmentation for port placement and navigation for rigid anatomy*, in 4. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboter-Assistierte Chirurgie (CURAC 2005), Berlin, Germany, September 2005.

xvii. R. Bauernschmitt, M. Feuerstein, E. Schirmbeck, N. Augustin, S. Wildhirt, and R. Lange, *Improvement of endovascular stent grafting by* *augmented reality*, in Biomedizinische Technik (BMT 2005), Supplementary vol. 1, Part 2, vol. 50, Nuremberg, Germany, Septemper 2005, pp. 1262 – 1263.

xviii. R. Bauernschmitt, M. Feuerstein, E. Schirmbeck, J. Traub, G. Klinker, S. Wildhirt, and R. Lange, *Improved preoperative planning in robotic heart surgery*, in IEEE Proceedings of Computers in Cardiology (CinC 2004), Chicago, USA, Septemper 2004, pp. 773–776.

xix. J. Traub, M. Feuerstein, M. Bauer, E. U. Schirmbeck, H. Najafi, R. Bauernschmitt, and G. Klinker, *Augmented reality for port placement and navigation in robotically assisted minimally invasive cardiovascular surgery*, in Computer Assisted Radiology and Surgery, Chicago, USA, June 2004, pp. 735–740.

APPENDIX

B

ABSTRACTS OF MAJOR PUBLICATIONS NOT DISCUSSED IN THE DISSERTATION

Real-time Fusion of Ultrasound and Gamma Probe for Navigated Localization of Liver Metastases

Thomas Wendler, Marco Feuerstein, Joerg Traub, Tobias Lasser, Jakob Vogel, Sibylle Ziegler, and Nassir Navab

Liver metastases are an advanced stage of several types of cancer, usually treated with surgery. Intra-operative localization of these lesions is currently facilitated by intra-operative ultrasound (IOUS) and palpation, yielding a high rate of false positives due to benign abnormal regions. In this paper we present the integration of functional nuclear information from a gamma probe with IOUS, to provide a synchronized, real-time visualization that facilitates the detection of active metastases intra-operatively. We evaluate the system in an ex-vivo setup employing a group of physicians and medical technicians and show that the addition of functional imaging improves the accuracy of localizing and identifying malignant and benign lesions significantly. Furthermore we are able to demonstrate that the inclusion of an advanced, augmented visualization provides more reliability and confidence on classifying these lesions.

Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures

Nassir Navab, Joerg Traub, Tobias Sielhorst, Marco Feuerstein, and Christoph Bichlmeier

One key to the success of a user interface that includes AR visualization is its ability to automatically recognize different phases of a workflow, which each require various levels of augmentation. It is also important for the AR system to be transparent to the user during the rest of the procedure. These issues have greater importance when dealing with computer-aided surgery applications. In most of these applications, a surgeon needs augmentation for only quite brief periods, such as choosing the ports for a laparoscopic intervention or localizing the major arteries before starting a liver resection. These augmentations, however, can play an important role in the overall procedure's success. During the past three years, the authors have tried to develop such integrated AR solutions in the context of minimally invasive surgery. In this article, they discuss their activities and recent results.

A New Approach to Ultrasound Guided Radio-Frequency Needle Placement

Claudio Alcérreca, Jakob Vogel, Marco Feuerstein, and Nassir Navab

The transformation from an ultrasound (US) image plane to the coordinate system of a position sensor attached to the US transducer can be computed with US calibration algorithms. This knowledge can be used in many applications, including freehand 3D US and US guided surgical navigation. We present a software system assisting the surgeon to position a radio-frequency (RF) tumor ablation needle using augmented ultrasound, thus simplifying the treatment by (1) dividing it into two simple consecutive tasks, lesion finding and needle placement, and (2) relating the needle to the US plane at any time.

Laparoscopic Virtual Mirror – New Interaction Paradigm for Monitor Based Augmented Reality

Nassir Navab, Marco Feuerstein, and Christoph Bichlmeier

A major roadblock for using augmented reality in many medical and industrial applications is the fact that the user cannot take full advantage of the 3D virtual data. This usually requires the user to move the virtual object, which disturbs the real/virtual alignment, or to move his head around the real objects, which is not always possible and/or practical. This problem becomes more dramatic when a single camera is used for monitor based augmentation, such as in augmented laparoscopic surgery. In this paper we introduce an interaction and 3D visualization paradigm, which presents a new solution to this old problem. The interaction paradigm uses an interactive virtual mirror positioned into the augmented scene, which allows easy and complete interactive visualization of 3D virtual data.

This paper focuses on the exemplary application of such visualization techniques to laparoscopic interventions. A large number of such interventions aims at regions inside a specific organ, e.g. blood vessels to be clipped for tumor resection. We use high-resolution intra-operative imaging data generated by a mobile C-arm with cone-beam CT imaging capability. Both the C-arm and the laparoscope are optically tracked and registered in a common world coordinate frame. After patient positioning, port placement, and carbon dioxide insufflation, a C-arm volume is reconstructed during patient exhalation and superimposed in real time on the laparoscopic live video without any need for an additional patient registration procedure. To overcome the missing perception of 3D depth and shape when rendering virtual volume data directly on top of the organ's surface view, we introduce the concept of a laparoscopic virtual mirror: A virtual reflection plane within the live laparoscopic video, which is able to visualize a reflected side view of the organ and its interior. This enables the surgeon to observe the 3D structure of, for example, blood vessels by moving the virtual mirror within the augmented monocular view of the laparoscope.

A Novel Segmentation and Navigation Tool for Endovascular Stenting of Aortic Aneurysms

Marco Feuerstein, Konstantinos Filippatos, Oliver Kutter, Eva U. Schirmbeck, Robert Bauernschmitt, and Nassir Navab

Endovascular stenting is a minimally invasive technique to exclude an aortic aneurysm or dissection from the circulatory system. Currently, there is no technical aid to guide the surgical staff during the intervention, except the default visualization interface provided with the CT scanner and mobile C-arm. The purpose of our proposed system is two-fold: (1) In the planning phase, a modified graph cuts algorithm automatically segments the aorta and aneurysm, so the surgical staff can choose an appropriate type of stent to match the segmented location, length, and diameter of the aneurysm and aorta. (2) During implantation of the stent graft, after a landmark based registration of CT and angiography data, the current position of the stent can be visualized in the 3D CT data set at any time. This will enhance the accuracy of the actions of the surgeon, along with a minimum use of angiography, leading to less radiation exposure and less contrast agent injection.

LIST OF FIGURES

REFERENCES

[1] L. Adhami and È. Coste-Manière, *A versatile system for computer integrated mini-invasive robotic surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), Tokyo, Japan, 2002, pp. 272–281.

[2] L. Adhami and È. Coste-Manière, *Optimal planning for minimally invasive surgical robots*, IEEE Transactions on Robotics and Automation, 19 (2003), pp. 854–863.

[3] K. S. Arun, T. S. Huang, and S. D. Blostein, *Least-squares fitting of two 3-d point sets*, IEEE Trans. Pattern Anal. Mach. Intell., 9 (1987), pp. 698–700.

[4] R. T. Azuma, *A survey of augmented reality*, Presence: Teleoperators and Virtual Environments, 6 (1997), pp. 355–385.

[5] R. T. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, *Recent advances in augmented reality*, IEEE Computer Graphics and Applications, 21 (2001), pp. 34–47.

[6] A. Bagheri, *Spiritual care and a practical suggestion*, in Bioethics in and from Asia: The Fifth International Tsukuba Bioethics Roundtable, and Intensive Cross-Cultural Bioethics Course (TRT5), November 1999.

[7] M. Bajura, H. Fuchs, and R. Ohbuchi, *Merging virtual objects with the real world: seeing ultrasound imagery within the patient*, in Proceedings of the 19th annual conference on Computer graphics and interactive techniques, ACM Press, 1992, pp. 203–210.

[8] J. M. Balter, L. A. Dawson, S. Kazanjian, C. McGinn, K. K. Brock, T. Lawrence, and R. T. Haken, *Determination of ventilatory liver movement via radiographic evaluation of diaphragm position*, International Journal of Radiation Oncology*Biology*Physics, 51 (2001), pp. 267–270.

[9] M. Bauer, M. Schlegel, D. Pustka, N. Navab, and G. Klinker, *Predicting and estimating the accuracy of vision-based optical tracking systems*, in Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR), Santa Barbara (CA), USA, October 2006, pp. 43–51.

[10] M. A. Bauer, *Tracking Errors in Augmented Reality*, PhD thesis, Technische Universität München, 2007.

[11] M. Baumhauer, T. Simpfendörfer, R. Schwarz, M. Seitel, B. Müller-Stich, C. Gutt, J. Rassweiler, H.-P. Meinzer, and I. Wolf, *Soft tissue navigation for laparoscopic prostatectomy: evaluation of camera pose estimation for enhanced visualization*, in Medical Imaging 2007: Visualization and Image-Guided Procedures, K. R. Cleary and M. I. Miga, eds., Proceedings of SPIE, San Diego, California, USA, February 2007.

[12] G. Beldi, M. Styner, S. Schindera, D. Inderbitzin, and D. Candinas, *Intraoperative three-dimensional fluoroscopic cholangiography*, Hepatogastroenterology, 53 (2006), pp. 157–159.

[13] G. Berci and K. A. Forde, *History of endoscopy - what lessons have we learned from the past?*, Surgical Endoscopy, 14 (2002), pp. 5–15.

[14] G. Bianchi, C. Wengert, M. Harders, P. Cattin, and G. Székely, *Camera-marker alignment framework and comparison with hand-eye calibration for augmented reality applications*, in ISMAR 2005, October 2005, pp. 188–189.

[15] W. Birkfellner, *Tracking Systems in Surgical Navigation*, PhD thesis, Department of Biomedical Engineering and Physics, General Hospital, University of Vienna, 2000.

[16] W. Birkfellner, F. Watzinger, F. Wanschitz, G. Enislidis, M. Truppe, R. Ewers, and H. Bergmann, *Concepts and results in the development of a hybrid tracking system for cas*, in Proceedings of the First International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI), I. W. M. Wells, A. C. F. Colchester, and S. L. Delp, eds., vol. 1496 of Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998, pp. 343–351.

[17] W. Birkfellner, F. Watzinger, F. Wanschitz, R. Ewers, and H. Bergmann, *Calibration of tracking systems in a surgical environment*, IEEE Trans. Med. Imag., 17 (1998).

[18] G. Bishop, G. Welch, and B. D. Allen, *Course 11 – tracking: Beyond 15 minutes of thought*, 2001.

[19] R. Bittner, *The standard of laparoscopic cholecystectomy*, Langenbeck's Archives of Surgery, 389 (2004), pp. 157–163.

[20] S. A. Boppart, T. F. Deutsch, and D. W. Rattner, *Optical imaging technology in minimally invasive surgery – current status and future directions*, Surgical Endoscopy, 13 (1999), pp. 718–722.

[21] I. Bricault, G. Ferretti, and P. Cinquin, *Registration of real and ct-derived virtual bronchoscopic images to assist transbronchial biopsy*, IEEE Trans. Med. Imag., 17 (1998), pp. 703–714.

[22] D. Burschka, M. Li, R. Taylor, and G. D. Hager, *Scale-invariant registration of monocular endoscopic images to ct-scans for sinus surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), C. Barillot, D. Haynor, and P. Hellier, eds., vol. 3217 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 413–421.

[23] J. Cannon, J. Stoll, S. Selha, P. Dupont, R. Howe, and D. Torchiana, *Port placement planning in robot-assisted coronary artery bypass*, IEEE Transactions on Robotics and Automation, 19 (2003), pp. 912–917.

[24] J. Canny, *A computational approach to edge detection*, IEEE Trans. Pattern Anal. Machine Intell., 8 (1986), pp. 679–698.

[25] A. M. Chiu, D. Dey, M. Drangova, W. D. Boyd, and T. M. Peters, *3-d image guidance for minimally invasive robotic coronary artery bypass*, The Heart Surgery Forum, 3 (2000), pp. 224–231.

[26] J. C. K. Chou and M. Kamel, *Finding the Position and Orientation of a Sensor on a Robot Manipulator Using Quaternions*, The International Journal of Robotics Research, 10 (1991), pp. 240–254.

[27] A. Chung, P. Edwards, F. Deligianni, and G. Yang, *Freehand cocalibration of optical and electromagnetic trackers for navigated bronchoscopy*, in Proceesings of Medical Imaging and Augmented Reality: Second International Workshop, MIAR 2004, Beijing, China, August 19-20, 2004., 2004.

[28] K. Cleary and the Insight Software Consortium, *IGSTK: The Book*, February 2007.

[29] M. Clifford, F. Banovac, E. Levy, and K. Cleary, *Assessment of hepatic motion secondary to respiration for computer assisted interventions*, Computer Aided Surgery, 7 (2002), pp. 291–299.

[30] J. Climent and P. Marés, *Automatic instrument localization in laparoscopic surgery*, Electronic Letters on Computer Vision and Image Analysis, 4 (2004), pp. 21–31.

[31] C. Couinaud, *Le foie: Etudes anatomiques et chirurgicales.* New York, NY: Masson Publishing USA Inc, 1957.

[32] K. Daniilidis, *Hand-eye calibration using dual quaternions*, International Journal of Robotics Research, 18 (1999), pp. 286–298.

[33] S. De Buck, J. Van Cleynenbreuge, I. Geys, T. Koninckx, P. R. Koninck, and P. Suetens, *A system to support laparoscopic surgery by augmented reality visualization*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), 2001, pp. 691–698.

[34] D. Dey, D. Gobbi, P. Slomka, K. Surry, and T. Peters, *Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: creating stereoscopic panoramas*, Medical Imaging, IEEE Transactions on, 21 (2002), pp. 23–30.

[35] K. Doi, *Current status and future potential of computer-aided diagnosis in medical imaging*, The British Journal of Radiology, 78 (2005), pp. 3–19.

[36] C. Doignon, F. Nageotte, and M. de Mathelin, *Segmentation and guidance of multiple rigid objects for intra-operative endoscopic vision*, in Workshop on Dynamic Vision, European Conference on Computer Vision, 2006.

[37] D. Eggert, A. Lorusso, and R. Fisher, *Estimating 3-d rigid body transformations: a comparison of four major algorithms*, Machine Vision and Applications, 9 (1997), pp. 272–290.

[38] J. Ellsmere, J. Stoll, D. W. Rattner, D. Brooks, R. Kane, W. M. Wells, III, R. Kikinis, and K. Vosburgh, *A navigation system for augmenting laparoscopic ultrasound*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 184–191.

[39] J. Ellsmere, J. Stoll, W. Wells, R. Kikinis, K. Vosburgh, R. Kane, D. Brooks, and D. Rattner, *A new visualization technique for laparoscopic ultrasonography*, Surgery, 136 (2004), pp. 84–92.

[40] B. J. Erickson and B. Bartholmai, *Computer-aided detection and diagnosis at the start of the third millennium*, Journal of Digital Imaging, 15 (2001), pp. 59–68.

[41] R. S. J. Estépar, N. Stylopoulos, R. E. Ellis, E. Samset, C.-F. Westin, C. Thompson, and K. Vosburgh, *Towards scarless surgery: An endoscopic-ultrasound navigation system for transgastric access procedures*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., vol. 4190 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 445–453.

[42] E. Euler, S. Heining, C. Riquarts, and W. Mutschler, *C-arm-based three-dimensional navigation: a preliminary feasibility study*, Computer Aided Surgery, 8 (2003), pp. 35–41.

[43] V. Falk, F. Mourgues, L. Adhami, S. Jacobs, H. Thiele, S. Nitzsche, F. W. Mohr, and È. Coste-Manière, *Cardio navigation: Planning, simulation, and augmented reality in robotic assisted endoscopic bypass grafting*, The Annals of Thoracic Surgery, 79 (2005), pp. 2040–2047.

[44] O. Faugeras, *Three-Dimensional Computer Vision. A Geometric Viewpoint*, Artificial Intelligence, MIT Press, 2001.

[45] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, *Intra-operative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection*, To appear in IEEE Trans. Med. Imag., (2007).

[46] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, *Registration-free laparoscope augmentation for intra-operative liver resection planning*, in Medical Imaging 2007: Visualization and Image-Guided Procedures, K. R. Cleary and M. I. Miga, eds., Proceedings of SPIE, San Diego, California, USA, February 2007.

[47] M. Feuerstein, T. Reichl, J. Vogel, A. Schneider, H. Feussner, and N. Navab, *Magneto-optic tracking of a flexible laparoscopic ultrasound transducer for laparoscope augmentation*, in To appear in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), N. Ayache, S. Ourselin, and A. Maeder, eds., vol. 4791 of Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007, Springer-Verlag, pp. 458–466.

[48] M. Feuerstein, S. M. Wildhirt, R. Bauernschmitt, and N. Navab, *Automatic patient registration for port placement in minimally invasive endoscopic surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), J. S. Duncan and G. Gerig, eds., vol. 3750 of Lecture Notes in Computer Science, Palm Springs, CA, USA, September 2005, Springer-Verlag, pp. 287–294.

[49] G. Fichtinger, A. Deguet, K. Masamune, E. Balogh, G. S. Fischer, H. Mathieu, R. H. Taylor, S. J. Zinreich, and L. M. Fayad, *Image overlay guidance for needle insertion in ct scanner*, IEEE Transactions on Biomedical Engineering, 52 (2005), pp. 1415–1424.

[50] G. S. Fischer and R. H. Taylor, *Electromagnetic tracker measurement error simulation and tool design*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), J. S. Duncan and G. Gerig, eds., vol. 3750 of Lecture Notes in Computer Science, Palm Springs, CA, USA, September 2005, Springer-Verlag, pp. 73–80.

[51] M. A. Fischler and R. C. Bolles, *Random sample consensus: A paradigm for model fitting with applications to image*, Analysis and Automated Cartography, Commun. ACM, 24 (1981).

[52] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, Jr., *Predicting error in rigid-body point-based registration*, IEEE Trans. Med. Imag., 14 (1998), pp. 694–702.

[53] W. Freysinger, A. Gunkel, and W. Thumfart, *Image-guided endoscopic ent surgery*, European archives of Otorhinolaryngology, 254 (1997), pp. 343–346.

[54] E. M. Friets, J. W. Strohbehn, J. F. Hatch, and D. W. Roberts, *A frameless stereotaxic operating microscope for neurosurgery*, IEEE Transactions on Biomedical Engineering, 36 (1989).

[55] H. Fuchs, M. A. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer, *Augmented reality visualization for laparoscopic surgery*, in Proceedings of the First International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI), I. W. M. Wells, A. C. F. Colchester, and S. L. Delp, eds., vol. 1496 of Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998, Springer-Verlag, pp. 934–943.

[56] S. Ganguli, J. B. Kruskal, D. D. Brennan, and R. A. Kane, *Intraoperative laparoscopic ultrasound*, Radiologic Clinics of North America, 44 (2006), pp. 925–935.

[57] S. Gold and A. Rangarajan, *A graduated assignment algorithm for graph matching*, IEEE Trans. Pattern Anal. Mach. Intell., 18 (1996), pp. 377–388.

[58] H. Gray, *Anatomy of the human body*, Lea & Febiger, 20th ed., 1918.

[59] W. E. L. Grimson, T. Lozano-Perez, W. M. Wells, III, G. J. Ettinger, S. J. White, and R. Kikinis, *An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization*, IEEE Trans. Med. Imag., 15 (1996), pp. 129–140.

[60] P. A. Grützner, A. Hebecker, H. Waelti, B. Vock, L.-P. Nolte, and A. Wentzensen, *Clinical study for registration-free 3d-navigation with the siremobil iso-c^{3D} mobile c-arm*, electromedica, 71 (2003), pp. 7–16.

[61] R. Grzeszczuk, S. Chin, R. Fahrig, H. Abassi, H. Holz, J. A. Daniel Kim1, and R. Shahidi, *A fluoroscopic x-ray registration process for three-dimensional surgical navigation*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), Lecture Notes in Computer Science, Springer-Verlag, 2000.

[62] J. Harms, H. Feussner, M. Baumgartner, A. Schneider, M. Donhauser, and G. Wessels, *Three-dimensional navigated laparoscopic ultrasonography*, Surgical Endoscopy, 15 (2001), pp. 1459–1462.

[63] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd ed., 2003.

[64] J. Heikkilä and O. Silvén, *A four-step camera calibration procedure with implicit image correction*, in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 1997, pp. 1106–1112.

[65] W. R. Hendee and E. R. Ritenour, *Medical Imaging Physics*, John Wiley & Sons Inc, 4th ed., 2002.

[66] A. Herline, J. Stefansic, J. Debelak, S. Hartmann, C. Pinson, R. Galloway, and W. Chapman, *Image-guided surgery: preliminary feasibility studies of frameless stereotactic liver surgery*, Archives of Surgery, 134 (1999), pp. 644–650.

[67] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, *Medical image registration*, Phys. Med. Biol., 46 (2001), pp. R1–R45.

[68] R. Holloway, *Registration error analysis for augmented reality*, Presence: Teleoperators and Virtual Environments, 6 (1997), pp. 413–432.

[69] R. Horaud and F. Dornaika, *Hand-eye calibration*, IJRR, 14 (1995), pp. 195–210.

[70] P.-W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece, *Rapid, easy and reliable calibration for freehand 3d ultrasound*, Ultrasound in Medicine & Biology, 32 (2006), pp. 823–835.

[71] J. B. Hummel, M. R. Bax, M. L. Figl, Y. Kang, C. Maurer, Jr., W. W. Birkfellner, H. Bergmann, and R. Shahidi, *Design and application of an assessment protocol for electromagnetic tracking systems*, Medical Physics, 32 (2005), pp. 2371–2379.

[72] M. C. Jacobs, M. A. Livingston, and A. State, *Managing latency in complex augmented reality systems*, in Proceedings of the 1997 symposium on Interactive 3D graphics, ACM Press, 1997, pp. 49–54.

[73] J. J. Jakimowicz, *Intraoperative ultrasonography in open and laparoscopic abdominal surgery: an overview*, Surgical Endoscopy, 20 (2006), pp. 425–435.

[74] T. Kawamata, H. Iseki, T. Shibasaki, and T. Hori, *Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: technical note*, Neurosurgery, 50 (2002), pp. 1393–1397.

[75] R. Khadem, M. R. Bax, J. A. Johnson, E. P. Wilkinson, and R. Shahidi, *Endoscope calibration and accuracy testing for 3d/2d image registration*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), 2001, pp. 1361–1362.

[76] R. Khadem, C. C. Yeh, M. Sadeghi-Tehrani, M. R. Bax, J. A. Johnson, J. N. Welch, E. P. Wilkinson, and R. Shahidi, *Comparative tracking error analysis of five different optical tracking systems*, Computer Aided Surgery, 5 (2000), pp. 98–107.

[77] V. V. Kindratenko, *A survey of electromagnetic position tracker calibration techniques*, Virtual Reality: Research, Development, and Applications, 5 (2000), pp. 169–182.

[78] A. P. King, P. J. Edwards, C. R. Maurer, Jr., D. A. de Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. J. Strong, C. L. Chandler, A. Richards, and M. J. Gleeson, *Design and evaluation of a system for microscope-assisted guided interventions*, IEEE Trans. Med. Imag., 19 (2000), pp. 1082–1093.

[79] M. Kleemann, P. Hildebrand, M. Birth, and H. P. Bruch, *Laparoscopic ultrasound navigation in liver surgery: technical aspects and accuracy*, Surgical Endoscopy, 20 (2006), pp. 726–729.

[80] T. Klein, J. Traub, H. Hautmann, A. Ahmadian, and N. Navab, *Fiducial free registration procedure for navigated bronchoscopy*, 2007. To appear at Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI).

[81] N. Koizumi, K. Sumiyama, N. Suzuki, A. Hattori, H. Tajiri, and A. Uchiyama, *Development of three-dimensional endoscopic ultrasound system with optical tracking*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), T. Dohi and R. Kikinis, eds., vol. 2488 of Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 60–65.

[82] W. Konen, M. Scholz, and S. Tombrock, *The vn project: endoscopic image processing for neurosurgery.*, Comput Aided Surg, 3 (1998), pp. 144–8.

[83] S. Kriminski, M. Mitschke, S. Sorensen, N. M. Wink, P. E. Chow, S. Tenn, and T. D. Solberg, *Respiratory correlated cone-beam computed tomography on an isocentric c-arm*, Physics in Medicine and Biology, 50 (2005), pp. 5263–5280.

[84] J. Krücker, A. Viswanathan, J. Borgert, N. Glossop, Y. Yanga, and B. J. Wood, *An electro-magnetically tracked laparoscopic ultrasound for multimodality minimally invasive surgery*, in Computer Assisted Radiology and Surgery, Berlin, Germany, June 2005, pp. 746–751.

[85] J. Krüger, J. Schneider, and R. Westermann, *ClearView: An interactive context preserving hotspot visualization technique*, IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2006), 12 (2006), pp. 941–948.

[86] S. Krüger, F. Vogt, W. Hohenberger, D. Paulus, H. Niemann, and C. H. Schick, *Evaluation of computer-assisted image enhancement in minimal invasive endoscopic surgery*, Methods Inf Med, 43 (2004), pp. 362–366.

[87] M. Kukuk, *A model-based approach to intraoperative guidance of flexible endoscopy*, PhD thesis, Universität Dortmund, May 2003.

[88] M. Kukuk and B. Geiger, *A real-time deformable model for flexible instruments inserted into tubular structures*, in MICCAI '02: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II, London, UK, 2002, Springer-Verlag, pp. 331–338.

[89] O. Kutter, S. Kettner, E. Braun, N. Navab, R. Lange, and R. Bauernschmitt, *Towards an integrated planning and navigation system for aortic stent-graft placement*, in Computer Assisted Radiology and Surgery, 2007.

[90] T. Langø, *Ultrasound Guided Surgery: Image Processing and Navigation*, PhD thesis, Norwegian University of Science and Technology, October 2000.

[91] W. Lau, *History of Endoscopic and Laparoscopic Surgery*, World Journal of Surgery, 21 (1997), pp. 444–453.

[92] D. C. Leiner, *Miniature optics in the hospital operating room*, in Gradient-Index Optics and Miniature Optics, J. D. Rees and D. C. Leiner, eds., April 1988, pp. 52–+.

[93] H. U. Lemke, *Summary of the white paper of dicom wg24 'dicom in surgery'*, in Proceedings of the SPIE Medical Imaging 2007: PACS and Imaging Informatics., S. C. Horii and K. P. Andriole, eds., vol. 6516 of Proceedings of SPIE, March 2007.

[94] J. Leven, D. Burschka, R. Kumar, G. Zhang, S. Blumenkranz, X. D. Dai, M. Awad, G. D. Hager, M. Marohn, M. Choti, C. Hasser, and R. H. Taylor, *Davinci canvas: A telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 3749 of Lecture Notes in Computer Science, Springer-Verlag, September 2005, pp. 811–818.

[95] M. Li and D. Betsis, *Head-eye calibration*, in ICCV '95: Proceedings of the Fifth International Conference on Computer Vision, Washington, DC, USA, 1995, IEEE Computer Society, p. 40.

[96] C. A. Linte, A. D. Wiles, N. Hill, J. Moore, C. Wedlake, G. Guiraudon, D. Jones, D. Bainbridge, and T. M. Peters, *An augmented reality environment for image-guidance of off-pump mitral valve implantation*, in Medical Imaging 2007: Visualization and Image-Guided Procedures, K. R. Cleary and M. I. Miga, eds., Proceedings of SPIE, San Diego, California, USA, February 2007.

[97] G. Litynski, *Endoscopic Surgery: The History, the Pioneers*, World Journal of Surgery, 23 (1999), pp. 745–753.

[98] B. P. Lo, A. Darzi, and G.-Z. Yang, *Episode classification for the analysis of tissue/instrument interaction with multiple visual cues*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., no. 2878 in Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 230–237.

[99] W. Lorensen, H. Cline, C. Nafis, R. Kikinis, D. Altobelli, L. Gleason, G. Co, and N. Schenectady, *Enhancing reality in the operating room*, in IEEE Conference on Visualization, 1993, pp. 410–415.

[100] W. E. Lorensen and H. E. Cline, *Marching cubes: A high resolution 3d surface construction algorithm*, in SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques, New York, NY, USA, 1987, ACM Press, pp. 163–169.

[101] J. B. A. Maintz and M. A. Viergever, *A survey of medical image registration*, Medical Image Analysis, 2 (1998), pp. 1–36.

[102] T. Mala and B. Edwin, *Role and limitations of laparoscopic liver resection of colorectal metastases*, Digestive Diseases, 23 (2005), pp. 142–150.

[103] S. Martelli, S. Bignozzi, M. Bontempi, S. Zaffagnini, and L. Garcia, *Comparison of an optical and a mechanical navigation system*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., no. 2879 in Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 303–310.

[104] G. Marti, V. Bettschart, J.-S. Billiard, and C. Baur, *Hybrid method for both calibration and registration of an endoscope with an active optical tracker*, in Computer Assisted Radiology and Surgery, 2004.

[105] K. Masamune, G. Fichtinger, A. Deguet, D. Matsuka, and R. Taylor, *An image overlay system with enhanced reality for percutaneous therapy performed inside ct scanner*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), 2002.

[106] Y. Masutani, M. Iwahara, O. Samuta, Y. Nishi, N. Suzuki, M. Suzuki, T. Dohi, H. Iseki, and K. Takakura, *Development of integral photography-based enhanced reality visualization system for surgical support*, Proc. of ISCAS, 95 (1995), pp. 16–17.

[107] C. R. Maurer, Jr., J. M. Fitzpatrick, M. Y. Wang, J. Robert L. Galloway, R. J. Maciunas, and G. S. Allen, *Registration of head volume images using implantable fiducial markers*, IEEE Trans. Med. Imag., 16 (1997), pp. 447–462.

[108] M. R. Mayberg, E. Lapresto, and E. J. Cunningham, *Image-guided endoscopy: description of technique and potential applications*, Neurosurg Focus, 19 (2005), pp. 1–5.

[109] S. J. McKenna, H. N. Charif, and T. Frank, *Towards video understanding of laparoscopic surgery: Instrument tracking*, in Image and Vision Computing New Zealand, 2005.

[110] H.-P. Meinzer, M. Thorn, and C. E. Cárdenas, *Computerized planning of liver surgery – an overview*, Computers & Graphics, 26 (2002), pp. 569–576.

[111] M. Meissner, H. Pfister, R. Westermann, and C. Wittenbrink, *Volume visualization and volume rendering techniques.* Tutorial presented at Eurographics, 2000.

[112] J. Mellor, *Enhanced reality visualization in a surgical environment*, Tech. Report 1544, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, January 1995.

[113] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins, *A review of calibration techniques for freehand 3-d ultrasound systems*, Ultrasound in Medicine and Biology, 31 (2005), pp. 449–471.

[114] P. Milgram and F. Kishino, *A taxonomy of mixed reality visual displays*, in IEICE Trans. Information Systems, 1994, pp. 1321–1329.

[115] R. Miller and T. Hollingsworth, *Rigid endoscopes: Optical and design considerations*, British Medical Bulletin, 42 (1986), pp. 226–229.

[116] M. Mitschke and N. Navab, *Recovering projection geometry: How a cheap camera can outperform an expensive stereo system*, in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), vol. 1, 2000, pp. 193–200.

[117] K. Mori, D. Deguchi, K. Akiyama, T. Kitasaka, C. R. Maurer, Jr., Y. Suenaga, H. Takabatake, M. Mori, , and H. Natori, *Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), Springer-Verlag, 2005, pp. 543–550.

[118] K. Mori, D. Deguchi, J. ichi Hasegawa, Y. Suenaga, J. ichiro Toriwaki, H. Takabatake, and H. Natori, *A method for tracking the camera motion of real endoscope by epipolar geometry analysis and virtual endoscopy system*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 2208 of Lecture Notes in Computer Science, Springer-Verlag, 2001, pp. 1–8.

[119] P. Mountney, D. Stoyanov, A. Davison, and G.-Z. Yang, *Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., vol. 4190 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 347–354.

[120] F. Mourgues and È. Coste-Manière, *Flexible calibration of actuated stereoscopic endoscope for overlay in robot assisted surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), 2002, pp. 25–34.

[121] F. Mourgues, T. Vieville, V. Falk, and È. Coste-Manière, *Interactive guidance by image overlay in robot assisted coronary artery bypass*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 173–181.

[122] R. Mårvik, T. Langø, G. Tangen, F. Lindseth, Y. Yavuz, and T. Nagelhus Hernes, *Image-guided laparoscopic surgery – review and current status*, Minerva Chirurgica, 60 (2005), pp. 305–325.

[123] D. Mucha, B. Kosmecki, and J. Bier, *Plausibility check for error compensation in electromagnetic navigation in endoscopic sinus surgery*, International Journal of Computer Assisted Radiology and Surgery, 1 (2006), pp. 316–318.

[124] M. J. Murphy, *An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery*, Medical Physics, 24 (1997), pp. 857–866.

[125] C. Nafis, V. Jensen, L. Beauregard, and P. Anderson, *Method for estimating dynamic em tracking accuracy of surgical navigation tools*, in Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display, K. R. Cleary and R. L. Galloway, Jr., eds., vol. 6141 of Proceedings of SPIE, March 2006.

[126] K. Nakada, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, and S. Tamura, *A rapid method for magnetic tracker calibration using a magneto-optic hybrid tracker*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., vol. 2879 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 285–293.

[127] M. Nakamoto, Y. Sato, M. Miyamoto, Y. Nakamjima, K. Konishi, M. Shimada, M. Hashizume, and S. Tamura, *3d ultrasound system using a magneto-optic hybrid tracker for augmented reality visualization in laparoscopic liver surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), T. Dohi and R. Kikinis, eds., vol. 2489 of Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 148–155.

[128] N. Navab, A. R. Bani-Hashemi, M. M. Mitschke, D. W. Holdsworth, R. Fahrig, A. J. Fox, and R. Graumann, *Dynamic geometrical calibration for 3d cerebral angiography*, in Medical Imaging 1996: Physics of Medical Imaging, vol. 2708 of Proceedings of SPIE, 1996, pp. 361–370.

[129] N. Navab, M. Feuerstein, and C. Bichlmeier, *Laparoscopic virtual mirror - new interaction paradigm for monitor based augmented reality*, in Virtual Reality, Charlotte, North Carolina, USA, March 2007, pp. 43–50.

[130] S. Nicolau, A. Garcia, X. Pennec, L. Soler, and N. Ayache, *Augmented reality guided radio-frequency tumor ablation*, Computer Animation and Virtual World (previously the Journal of Visualization & Computer Animation), 16 (2005).

[131] S. Nicolau, L. Goffin, and L. Soler, *A low cost and accurate guidance system for laparoscopic surgery: Validation on an abdominal phantom*, in ACM Symposium on Virtual Reality Software and Technology, November 2005, pp. 124–133.

[132] S. Nicolau, X. Pennec, L. Soler, and N. Ayache, *An accuracy certified augmented reality system for therapy guidance*, in Proc. of the 8th European Conference on Computer Vision (ECCV 04), vol. 3023 of Lecture Notes in Computer Science, Prague, May 2004, Springer-Verlag, pp. 79–91.

[133] S. Nicolau, X. Pennec, L. Soler, and N. Ayache, *A complete augmented reality guidance system for liver punctures: First clinical evaluation*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), J. Duncan and G. Gerig, eds., vol. 3749 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 539–547.

[134] J. Nilsson, *Timers: Implement a continuously updating, high-resolution time provider for windows*, MSDN Magazine, 19 (2004).

[135] P. M. Novotny, J. A. Stoll, N. V. Vasilyev, P. J. del Nido, P. E. Dupont, and R. D. Howe, *Gpu based real-time instrument tracking with three dimensional ultrasound*, in Proceedings of Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science, 2006, pp. 58–65.

[136] B. Olbrich, J. Traub, S. Wiesner, A. Wiechert, H. Feussner, and N. Navab, *Respiratory motion analysis: Towards gated augmentation of the liver*, in Computer Assisted Radiology and Surgery, Berlin, Germany, June 2005, pp. 248–253.

[137] T. Peters, *Image-guidance for surgical procedures*, Physics in Medicine and Biology, 51 (2006), pp. R505–R540.

[138] D. L. Pham, C. Xu, and J. L. Prince, *Current methods in medical image segmentation*, Annual Review of Biomedical Engineering, 2 (2000), pp. 315–337.

[139] R. Prager, R. Rohling, A. Gee, and L. Berman, *Rapid calibration for 3-d freehand ultrasound*, Ultrasound in Medicine and Biology, 24 (1998), pp. 855–869.

[140] B. Rau, M. Hünerbein, and P. M. Schlag, *Is there additional information from laparoscopic ultrasound in tumor staging?*, Digestive Surgery, 19 (2002), pp. 479–483.

[141] T. Reichl, *Online error correction for the tracking of laparoscopic ultrasound*, master's thesis, Technische Universität München, July 2007.

[142] M. Reuter, *Maximilian Nitze (1848–1906)*, Der Urologe, 45 (2006), pp. 1076–1083.

[143] K. Rhode, M. Sermesant, D. Brogan, S. Hegde, J. Hipwell, P. Lambiase, E. Rosenthal, C. Bucknall, S. Qureshi, J. Gill, R. Razavi, and D. Hill, *A system for real-time xmr guided cardiovascular intervention*, IEEE Trans. Med. Imag., 24 (2005), pp. 1428–1440.

[144] D. Ritter, M. Mitschke, and R. Graumann, *Markerless navigation with the intra-operative imaging modality siremobil iso-c*^{3D}, electromedica, 70 (2002), pp. 31–36.

[145] D. Roberts, J. Strohbehn, J. Hatch, W. Murray, and H. Kettenberger, *A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope.*, J Neurosurg, 65 (1986), pp. 545–9.

[146] J. Rolland, L. Davis, and Y. Baillot, *Fundamentals of Wearable Computers and Augmented Reality*, Mahwah, NJ: Lawrence Erlbaum, 2001, ch. A survey of tracking technology for virtual environments, pp. 67–112.

[147] F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi, *Augmented workspace: designing an ar testbed*, in Proc. IEEE and ACM Int'l Symp. on Augmented Reality, 2000, pp. 47–53.

[148] M. Scheuering, A. Schenk, A. Schneider, B. Preim, and G. Greiner, *Intraoperative augmented reality for minimally invasive liver interventions*, in Medical Imaging 2003: Visualization, Image-Guided Procedures, and Display, Proceedings of SPIE, 2003.

[149] K. Schicho, M. Figl, M. Donat, W. Birkfellner, R. Seemann, A. Wagner, H. Bergmann, and R. Ewers, *Stability of miniature electromagnetic tracking systems*, Physics in Medicine and Biology, 50 (2005), pp. 2089–2098.

[150] J. Schmidt, F. Vogt, and H. Niemann, *Robust Hand-Eye Calibration of an Endoscopic Surgery Robot Using Dual Quaternions* , in Pattern Recognition, 25th DAGM Symposium , B. Michaelis and G. Krell, eds., vol. 2781 of Lecture Notes in Computer Science , Berlin, Heidelberg, New York, 2003, Springer-Verlag, pp. 548–556.

[151] M. Schneider and C. Stevens, *Development and testing of a new magnetic-tracking device for image guidance*, in Medical Imaging 2007: Visualization and Image-Guided Procedures, K. R. Cleary and M. I. Miga, eds., Proceedings of SPIE, San Diego, California, USA, February 2007.

[152] M. Scholz, W. Konen, S. Tombrock, B. Fricke, and L. Adams, *Development of an endoscopic navigation system based on digital image processing*, Computer Aided Surgery, 3 (1998), pp. 134–143.

[153] S. D. Selha, P. E. Dupont, R. D. Howe, and D. F. Torchiana, *Dexterity optimization by port placement in robot-assisted minimally invasive surgery*, in Telemanipulator and Telepresence Technologies VIII, vol. 4570 of Proceedings of SPIE, Februar 2002, pp. 97–104.

[154] 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 Trans. Med. Imag., 21 (2002), pp. 1524–1535.

[155] R. Shahidi, B. Wang, M. Epitaux, R. Grzeszczuk, and J. Adler, *Volumetric image guidance via a stereotactic endoscope*, in Proceedings of the First

International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI), I. W. M. Wells, A. C. F. Colchester, and S. L. Delp, eds., vol. 1496 of Lecture Notes in Computer Science, Cambridge, MA, USA, October 1998, Springer-Verlag, pp. 241–252.

[156] F. Shi, J. Wang, and Y. Liu, *An approach to improve online hand-eye calibration*, in Pattern Recognition and Image Analysis, vol. 3522, Springer-Verlag Berlin / Heidelberg, 2005, pp. 647–655.

[157] Y. Shiu and S. Ahmad, *Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form ax=xb*, IEEE Transactions on Robotics and Automation, 5 (1989), pp. 16–29.

[158] J. H. Shuhaiber, *Augmented reality in surgery*, Arch Surg, 139 (2004), pp. 170–174.

[159] T. Sielhorst, M. A. Bauer, O. Wenisch, G. Klinker, and N. Navab, *Online estimation of the target registration error for n-ocular optical tracking systems*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), 2007. To appear at Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI).

[160] T. Sielhorst, C. Bichlmeier, S. M. Heining, and N. Navab, *Depth perception a major issue in medical ar: Evaluation study by twenty surgeons*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., Lecture Notes in Computer Science, 2006.

[161] T. Sielhorst, M. Feuerstein, J. Traub, O. Kutter, and N. Navab, *Campar: A software framework guaranteeing quality for medical augmented reality*, International Journal of Computer Assisted Radiology and Surgery, 1 (2006), pp. 29–30.

[162] J. H. Siewerdsen, D. J. Moseley, S. Burch, S. K. Bisland, A. Bogaards, B. C. Wilson, and D. A. 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 (2005), pp. 241–254.

[163] W. E. Smith, N. Vakil, and S. A. Maislin, *Correction of distortion in endoscope images*, IEEE Trans. Med. Imag., 11 (1992), pp. 117–122.

[164] J. D. Stefansic, A. J. Herline, Y. Shyr, W. C. Chapman, J. M. Fitzpatrick, and R. L. Galloway, *Registration of physical space to laparoscopic image space for use in minimally invasive hepatic surgery.*, IEEE Transactions on Medical Imaging, 19 (2000), pp. 1012–1023.

[165] J. Stoll, P. Novotny, R. Howe, and P. Dupont, *Real-time 3d ultrasound-based servoing of a surgical instrument*, in IEEE International Conference on Robotics and Automation (ICRA), 2006, pp. 613–618.

[166] D. Stoyanov, A. Darzi, and G. Z. Yang, *Laparoscope self-calibration for robotic assisted minimally invasive surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), J. Duncan and G. Gerig, eds., vol. 3750 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 114–121.

[167] K. H. Strobl and G. Hirzinger, *Optimal Hand-Eye Calibration*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, October 2006, pp. 4647–4653.

[168] P. Sturm, S. Ramalingam, and S. K. Lodha, *On calibration, structure-from-motion and multi-view geometry for panoramic camera models*, in Panoramic Photogrammetry Workshop, Berlin, Germany, 2005.

[169] F. Sutherland and J. Harris, *Claude Couinaud – a passion for the liver*, Arch Surg, 137 (2002), pp. 1305–1310.

[170] R. U. Thoranaghatte, G. Zheng, F. Langlotz, and L.-P. Nolte, *Endoscope-based hybrid navigation system for minimally invasive ventral spine surgeries*, Computer Aided Surgery, 10 (2005), pp. 351–356.

[171] J. Traub, M. Feuerstein, M. Bauer, E. U. Schirmbeck, H. Najafi, R. Bauernschmitt, and G. Klinker, *Augmented reality for port placement and navigation in robotically assisted minimally invasive cardiovascular surgery*, in Computer Assisted Radiology and Surgery, Chicago, USA, June 2004, pp. 735–740.

[172] J. Traub, P. Stefan, S.-M. Heining, T. Sielhorst, C. Riquarts, E. Euler, and N. Navab, *Hybrid navigation interface for orthopedic and trauma surgery*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., vol. 4190 of Lecture Notes in Computer Science, Copenhagen, Denmark, Oct. 2006, MICCAI Society, Springer, pp. 373–380.

[173] J. Traub, P. Stefan, S.-M. Heining, T. Sielhorst, C. Riquarts, E. Euler, and N. Navab, *Stereoscopic augmented reality navigation for trauma surgery: cadaver experiment and usability study*, International Journal of Computer Assisted Radiology and Surgery, 1 (2006), pp. 30–31.

[174] J. Traub, P. Stefan, S.-M. Heining, T. Sielhorst, C. Riquarts, E. Euler, and N. Navab, *Towards a hybrid navigation interface: Comparison of a slice based navigation system with in-situ visualization*, in Medical Imaging and Augmented Reality, G.-Z. Yang, T. Jiang, D. Shen, L. Gu, and J. Yang, eds., vol. 4091 of Lecture Notes in Computer Science, Shanghai, China, Aug. 2006, Springer, pp. 179–186.

[175] G. M. Treece, A. H. Gee, R. W. Prager, C. J. C. Cash, and L. H. Berman, *High-definition freehand 3-d ultrasound*, Ultrasound in Medicine and Biology, 29 (2003), pp. 529–546.

[176] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, 1998.

[177] R. Tsai, *A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses*, IEEE Journal of Robotics and Automation, RA-3 (1987), pp. 323–344.

[178] R. Tsai and R. Lenz, *Real time versatile robotics hand/eye calibration using 3d machine vision*, in Proceedings of the IEEE International Conference on Robotics and Automation, vol. 1, 1988, pp. 554–561.

[179] R. Y. Tsai and R. K. Lenz, *A new technique for fully autonomous and efficient 3d robotics hand/eye calibration*, IEEE Transactions on Robotics and Automation, 5 (1989), pp. 345–358.

[180] M. Tuceryan, D. S. Greer, R. T. Whitaker, D. E. Breen, C. Crampton, E. Rose, and K. H. Ahlers, *Calibration requirements and procedures for a monitor-based augmented reality system*, IEEE Transactions on Visualization and Computer Graphics, 1 (1995), pp. 255–273.

[181] S. Umeyama, *Least-squares estimation of transformation parameters between two point patterns*, IEEE Trans. Pattern Anal. Machine Intell., 13 (1991), pp. 376–380.

[182] R. Vecchio, B. MacFayden, and F. Palazzo, *History of laparoscopic surgery*, Panminerva Medica, 42 (2000), pp. 87–90.

[183] H. Visarius, J. Gong, C. Scheer, S. Haralamb, L. Nolte, and M. E. Müller, *Man-machine interfaces in computer assisted surgery*, Computer Aided Surgery, 2 (1997), pp. 102–107.

[184] F. Vogt, *Augmented Light Field Visualization and Real-Time Image Enhancement for Computer Assisted Endoscopic Surgery*, PhD thesis, Universität Erlangen-Nürnberg, 2005.

[185] F. Vogt, S. Krüger, H. Niemann, and C. Schick, *A system for real-time endoscopic image enhancement*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 356–363.

[186] F. Vogt, D. Paulus, and H. Niemann, *Highlight substitution in light fields*, in Proceedings of the IEEE International Conference on Image Processing (ICIP), IEEE Computer Society Press, 2002, pp. 637–640.

[187] S. Voros, J.-A. Long, and P. Cinquin, *Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., vol. 4190 of Lecture Notes in Computer Science, Springer, 2006, pp. 535–542.

[188] M. W. Walker, L. Shao, and R. A. Volz, *Estimating 3-d location parameters using dual number quaternions*, CVGIP: Image Underst., 54 (1991), pp. 358–367.

[189] M. Y. Wang, C. R. Maurer, Jr., J. M. Fitzpatrick, and R. J. Maciunas, *An automatic technique for finding and localizing externally attached markers in ct and mr volume images of the head*, IEEE Trans. Biomed. Eng., 43 (1996), pp. 627–637.

[190] C. Ware and R. Balakrishnan, *Reaching for Objects in VR Displays: Lag and Frame Rate*, ACM Transactions on Computer-Human Interaction, 1 (1994), pp. 331–356.

[191] I. Wegner, M. Vetter, M. Schoebinger, I. Wolf, and H. Meinzer, *Development of a navigation system for endoluminal brachytherapy in human lungs*, in Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display, K. R. Cleary and R. L. Galloway, Jr., eds., vol. 6141 of Proceedings of SPIE, March 2006, pp. 23–30.

[192] G. Welch and E. Foxlin, *Motion tracking: no silver bullet, but a respectable arsenal*, IEEE Computer Graphics and Applications, 22 (2002), pp. 24–38.

[193] P. Wellner, W. Mackay, and R. Gold, *Computer augmented environments: Back to the real world*, Commun. ACM, 36 (1993), pp. 24–26.

[194] T. Wendler, M. Feuerstein, J. Traub, T. Lasser, J. Vogel, F. Daghighian, S. Ziegler, and N. Navab, *Real-time fusion of ultrasound and gamma probe for navigated localization of liver metastases*, in To appear in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), N. Ayache, S. Ourselin, and A. Maeder, eds., vol. 4792 of Lecture Notes in Computer Science, Brisbane, Australia, October/November 2007, Springer-Verlag, pp. 252–260.

[195] T. Wendler, A. Hartl, T. Lasser, J. Traub, F. Daghighian, S. Ziegler, and N. Navab, *Towards intra-operative 3d nuclear imaging: reconstruction of 3d radioactive distributions using tracked gamma probes*, 2007. To appear at Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI).

[196] T. Wendler, J. Traub, S. Ziegler, and N. Navab, *Navigated three dimensional beta probe for optimal cancer resection*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., vol. 4190 of Lecture Notes in Computer Science, Copenhagen, Denmark, Oct. 2006, MICCAI Society, Springer, pp. 561–569.

[197] C. Wengert, P. Cattin, J. M. Duff, C. Baur, and G. Székely, *Markerless endoscopic registration and referencing*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. Larsen, M. Nielsen, and J. Sporring, eds., Lecture Notes in Computer Science, 2006.

[198] J. West and C. Maurer, *Designing optically tracked instruments for image-guided surgery*, IEEE Transactions on Medical Imaging, 23 (2004), pp. 533–545.

[199] R. Willson, *Modeling and Calibration of Automated Zoom Lenses*, PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, January 1994.

[200] A. Wolbarst and W. Hendee, *Evolving and experimental technologies in medical imaging*, Radiology, 238 (2006), pp. 16–39.

[201] I. Wolf, M. Vetter, I. Wegner, T. Böttger, M. Nolden, M. Schöbinger, M. Hastenteufel, T. Kunert, and H.-P. Meinzer, *The medical imaging interaction toolkit*, Medical Image Analysis, 9 (2005), pp. 594–604.

[202] J. W. Wong, M. B. Sharpe, D. A. Jaffray, V. R. Kini, J. M. Robertson, J. S. Stromberg, and A. A. Martinez, *The use of active breathing control (abc) to reduce margin for breathing motion*, International Journal of Radiation Oncology*Biology*Physics, 44 (1999), pp. 911–919.

[203] X. Wu and R. Taylor, *A direction space interpolation technique for calibration of electromagnetic surgical navigation systems*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., vol. 2879 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 215–222.

[204] T. Yamaguchi, M. Nakamoto, Y. Sato, K. Konishi, M. Hashizume, N. Sugano, H. Yoshikawa, and S. Tamura, *Development of a camera model and calibration procedure for oblique-viewing endoscopes*, Computer Aided Surgery, 9 (2004), pp. 203–214.

[205] T. Yamaguchi, M. Nakamoto, Y. Sato, Y. Nakajima, K. Konishi, M. Hashizume, T. Nishii, N. Sugano, H. Yoshikawa, K. Yonenobu, and S. Tamura, *Camera model and calibration procedure for oblique-viewing endoscope*, in Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI), R. E. Ellis and T. M. Peters, eds., Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 373–381.

[206] Z. Yaniv and K. Cleary, *Image-guided procedures: A review*, tech. report, Computer Aided Interventions and Medical Robotics, Imaging Science and Information Systems Center, Department of Radiology, Georgetown University Medical Center, Washington, DC, 20007, USA, April 2006.

[207] Z. Zhang, *Flexible camera calibration by viewing a plane from unknown orientations*, in Seventh International Conference on Computer Vision (ICCV'99), vol. 1, Los Alamitos, CA, USA, 1999, IEEE Computer Society, pp. 666–673.

[208] B. Zitová and J. Flusser, *Image registration methods: a survey*, Image and Vision Computing, 21 (2003), pp. 977–1000.