



**HAL**  
open science

# Automatic Registration of CT/MRI Data on Laparoscopic Liver Images for Surgical Gesture Assistance by Augmented Reality

Yamid Espinel Lopez

► **To cite this version:**

Yamid Espinel Lopez. Automatic Registration of CT/MRI Data on Laparoscopic Liver Images for Surgical Gesture Assistance by Augmented Reality. Medical Imaging. Université Clermont Auvergne, 2022. English. NNT : 2022UCFAC088 . tel-04112080

**HAL Id: tel-04112080**

**<https://theses.hal.science/tel-04112080v1>**

Submitted on 31 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UFR DE MÉDECINE  
ET DES PROFESSIONS PARAMÉDICALES**  
Université Clermont Auvergne

ÉCOLE DOCTORALE DES SCIENCES POUR L'INGÉNIEUR

## THÈSE

pour obtenir le grade de Docteur délivré par

**Université Clermont-Auvergne**

(Décret du 5 Juillet 1984)

**Spécialité “Vision par Ordinateur”**

*présentée et soutenue publiquement par*

**Yamid ESPINEL LÓPEZ**

le 13 Décembre, 2022

### **Recalage automatique de données scannographiques sur des images de cœlioscopie hépatique pour l'assistance au geste chirurgical par réalité augmentée**

Directeur de thèse : **Adrien BARTOLI**

Co-encadrant de thèse : **Christophe TILMANT**

#### **Jury**

<b>Guillaume MOREL,</b>	Professeur, Sorbonne Université, Président du Jury, Examineur
<b>Stamatia GIANNAROU,</b>	Professeur, Imperial College London, Rapporteur
<b>David FOFI,</b>	Professeur, Université de Bourgogne, Rapporteur
<b>Emmanuel BUC,</b>	Chirurgien, CHU Clermont-Ferrand, Invité
<b>Lilian CALVET,</b>	Chercheur Postdoctoral, Toulouse INP, Invité



**EnCoV, Institut Pascal, UMR 6602 CNRS,**  
**Université Clermont Auvergne, CHU Clermont-Ferrand**  
Faculté de Médecine  
28 place Henri Dunant, 63000 Clermont-Ferrand, France  
Tel: +33 4 73 17 81 23

---

---

# Abstract

Laparoscopic liver resection (LLR) is a less invasive procedure compared to open surgery, as it does not require making big incisions in the abdominal wall. Instead, small incisions are made through which thin instruments are inserted, including the endoscope, scissors, cauterisers, among others. Due to the lack of tactile feedback and the liver's big dimensions, it is difficult for surgeons to localise inner structures such as tumours and vessels. This increases the risk of causing damage to the patient because of accidental haemorrhages and cancerous cell remnants from inaccurate resections. Augmented Reality (AR) can help to overcome these issues by showing the internal structures, hence improving resection planning. AR works by overlaying a preoperative 3D model reconstructed from CT/MRI data onto the laparoscopic images. The preoperative 3D model is usually composed of the liver's outer surface and internal structures such as vessels and tumours. A typical AR-based guidance system for laparoscopy is composed of two stages: (i) an initial registration stage in which the preoperative 3D model is registered to the intraoperative laparoscopic images; and (ii) an update stage in which the model is automatically registered to every new image by tracking the organs. In this thesis we focus on the initial registration stage (i), which is a challenging and highly researched problem.

In the case of LLR, the liver may deform significantly between the moment when the CT/MRI scans are taken (preoperative stage) and the actual surgery (intraoperative stage). Thus, a deformable 3D-2D registration should be done for AR to work properly. This registration is done by using visual information present on the liver's surface. Visual information may consist of texture information, if available, and anatomical landmarks. However, due to the liver's big dimensions and the proximity of the laparoscope, the liver is only partially visible in the laparoscopic images. This makes the existing single-view registration methods to leave big portions of the preoperative 3D model unconstrained, reducing the accuracy of the augmentations. Existing multi-view methods do not work well on patient data due to the slight deformations undergone by the liver between views, even if at rest. Therefore, there is a need for registration methods that can cope with the partial views of the liver and the deformations across views, in a way that the final accuracy is improved.

Another issue in AR for LLR is the way the augmentations are presented to the surgeon. Most of the existing methods just project the registered 3D models to the optical axis of the laparoscope. This may create confusions to the surgeon regarding the true location of the inner structures. Some methods initially project a shape of the tumour towards a manually chosen location in the liver surface, which does not provide an accurate indication on how to reach the tumour, for example, using a particular tool. All of this can make surgeons to perform bad resection plannings, increasing the probability of missing the tumours.

In this thesis we aim to solve the aforementioned problems in three ways: In the first part, we propose a method combining visual cues with tactile interactions to solve registration on a single image with a partial view of the liver, taking advantage of the surgeon's understanding of the scene. Here, the surgeon refines registration by manually deforming the hidden parts of the liver using an enclosing cage. This method is useful to do a quick registration without having to compute inter-image correspondences. In the second part, we propose several methods to solve registration using multiple intraoperative views. This is done in an attempt to merge the information coming

---

from the different views, as well as to overcome possible human-induced errors in the method proposed at the beginning. These methods aim to work in two cases: when the liver does not show noticeable deformations between views (rigidly linked views), and when it does (non-rigidly linked views). These two cases can be related to when the liver is at rest, and when the liver is being manipulated with surgical tools, respectively. In the third part, we develop a keyhole-aware AR pipeline that provides the surgeon with an indication on how to reach a tumour from a particular keyhole. It does so by estimating the keyhole location from the visible tools in the image, and projecting the tumour's shape towards the estimated keyhole. In this way, the surgeon will have a clear idea of the tumour's location and perform resection planning accordingly. Although the proposed methods are designed to work with monocular endoscopic systems, they can also be applied to stereoscopic systems, including surgical robots.

---

# Résumé

La résection hépatique par coelioscopie (LLR) est une procédure moins invasive que la chirurgie traditionnelle, car elle ne nécessite pas de grandes incisions de la paroi abdominale. Au lieu de cela, de petites incisions sont ouvertes à travers lesquelles des instruments fins sont insérés, notamment un endoscope, des ciseaux et un cautérisateur. En raison du manque d'une palpation directe et des grandes dimensions du foie, il est difficile pour les chirurgiens de localiser les structures internes telles que les tumeurs et les vaisseaux. Cela augmente le risque de dommages au patient en raison d'hémorragies accidentelles et de restes de cellules cancéreuses en raison de résections imprécises. La réalité augmentée (RA) aide à surmonter ces problèmes en montrant les structures internes du foie pendant la chirurgie, améliorant ainsi la planification des résections. La RA fonctionne en superposant un modèle 3D préopératoire reconstruit à partir d'images TDM/IRM sur les images coelioscopiques. Le modèle 3D préopératoire est généralement composé de la surface externe du foie et des structures internes telles que les vaisseaux et les tumeurs. Un système de guidage basé sur la RA est essentiellement composé de deux étapes : (i) une étape de recalage initiale dans laquelle le modèle 3D préopératoire est recalé sur les images coelioscopiques préopératoires ; et (ii) une étape de mise à jour dans laquelle le modèle est recalé automatiquement dans chaque nouvelle image en suivant des repères visuels. Dans cette thèse, nous nous concentrons sur l'étape de recalage initial (i), qui est un défi et un problème très étudié.

Pour la LLR, le foie peut se déformer significativement entre le moment où les images TDM/IRM sont prises (étape préopératoire) et l'intervention proprement dite (étape peropératoire). Pour résoudre ce problème, un recalage 3D-2D déformable doit être effectué pour que la RA fonctionne correctement. Ce recalage est fait en utilisant des indices visuelles présents à la surface du foie. Ces indices peuvent consister en des informations de texture, s'il y en a, et des repères anatomiques. Cependant, en raison des grandes dimensions du foie et de la proximité de l'endoscope, le foie n'est que partiellement visible dans les images coelioscopiques. Cela fait que les méthodes existantes de recalage à vue unique laissent de grandes parties du modèle 3D préopératoire sans contraintes de recalage, ce qui réduit la précision des augmentations. De la même façon, les méthodes multi-vues existantes ne fonctionnent pas bien sur les données patients en raison des légères déformations subies par le foie entre les vues, même s'il est au repos. Par conséquent, il y a un besoin de méthodes de recalage pouvant travailler avec des vues partielles et avec des déformations entre les vues, de telle manière que la précision de l'augmentation soit améliorée.

Un autre problème en RA pour la LLR est la façon dont les augmentations sont présentées au chirurgien. La plupart des méthodes existantes projettent simplement les modèles 3D recalés vers l'axe optique de l'endoscope. Cela peut induire le chirurgien en erreur par rapport au vrai emplacement des structures internes. Certaines méthodes projettent initialement une silhouette de la tumeur vers un emplacement choisi manuellement dans la surface du foie, ce qui ne fournit pas une indication précise sur la manière d'atteindre la tumeur, par exemple à l'aide d'un outil. Tout cela peut amener les chirurgiens à faire de mauvaises planifications de résection, augmentant ainsi la probabilité de rater les tumeurs.

Dans cette thèse, nous visons à résoudre ces problèmes de trois manières. Dans la première partie, nous proposons une méthode qui combine les repères visuels avec des interactions tac-

---

tiles pour recalculer le modèle préopératoire sur une image unique avec une vue partielle du foie, en tirant parti de la compréhension du chirurgien de la scène. Ici, le chirurgien affine le recalage en déformant manuellement les parties cachées du foie à l'aide de points de contrôle positionnés sur une cage. Cette méthode est utile pour faire un recalage rapide sans avoir à calculer des correspondances entre les vues. Dans la deuxième partie, nous proposons plusieurs méthodes pour faire du recalage à l'aide de plusieurs vues peropératoires. Le but est de fusionner les informations provenant de toutes les vues, ainsi que de surmonter des éventuelles erreurs induites par l'utilisateur dans la méthode proposée initialement. Ces méthodes visent à fonctionner dans deux cas : lorsque le foie ne présente pas de déformations notables entre les vues (vues liées de manière rigide) et lorsqu'il en présente (vues liées de manière non rigide). Ces deux cas correspondent au moment où le foie est au repos, et au moment où le foie est manipulé avec des outils chirurgicaux, respectivement. Dans la troisième partie, nous développons une méthode de RA qui tient compte de l'emplacement du trocart, et qui fournit au chirurgien une indication sur la façon d'atteindre une tumeur depuis ce trocart. Pour ce faire, la méthode estime l'emplacement du trocart à partir des outils visibles dans l'image, pour ensuite projeter la silhouette de la tumeur vers le trocart estimé. De cette façon, le chirurgien aura une idée claire de l'emplacement de la tumeur et planifiera la résection en conséquence. Bien que les méthodes proposées soient conçues pour fonctionner avec des systèmes endoscopiques monoculaires, elles peuvent aussi être utilisées dans des systèmes stéréoscopiques, y compris des robots chirurgicaux.

---

# List of publications

## In international conferences:

- Espinel, Y., Özgür, E., Calvet, L., Le Roy, B., Buc, E., and Bartoli, A. (2019). Combining visual cues with interactions for 3D–2D registration in liver laparoscopy. In *International Conference on Information Processing in Computer-Assisted Interventions - IPCAI 2019*.
- Espinel, Y., Calvet, L., Botros, K., Buc, E., Tilmant, C., and Bartoli, A. (2021). Using multiple images and contours for deformable 3D-2D registration of a preoperative CT in laparoscopic liver surgery. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2021*.

## In academic journals:

- Espinel, Y., Özgür, E., Calvet, L., Le Roy, B., Buc, E., and Bartoli, A. (2020). Combining visual cues with interactions for 3D–2D registration in liver laparoscopy. *Annals of Biomedical Engineering*, 48(6):1712–1727.
- Espinel, Y., Calvet, L., Botros, K., Buc, E., Tilmant, C., and Bartoli, A. (2022). Using multiple images and contours for deformable 3D-2D registration of a preoperative CT in laparoscopic liver surgery. In *International Journal of Computer Assisted Radiology and Surgery* (accepted for publication in October, 2022).

## In collaboration:

- Teluob, G., Calvet, L., Espinel, Y., Savareux, L., Guandalino, M., Ravel, A., Bourdel, N., Chauvet, P., Guy, L., Chabrot, P., and Bartoli, A. (2019). Preliminary Trial of Augmented Reality Performed on a Regular and a Robot-Assisted Laparoscopic Partial Nephrectomies. *Videourology*, 33(3).
- Bertrand, L. R., Abdallah, M., Espinel, Y., Calvet, L., Pereira, B., Ozgur, E., Pezet, D., Buc, E., and Bartoli, A. (2020). A case series study of augmented reality in laparoscopic liver resection with a deformable preoperative model. *Surgical Endoscopy*, 34(12), 5642–5648.
- Rabbani, N., Calvet, L., Espinel, Y., Le Roy, B., Ribeiro, M., Buc, E., and Bartoli, A. (2022). A methodology and clinical dataset with ground-truth to evaluate registration accuracy quantitatively in computer-assisted Laparoscopic Liver Resection. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging Visualization*, 10(4), 441-450.
- Adballah, M., Espinel, Y., Calvet, L., Pereira, B., Le Roy, B., Bartoli, A., and Buc, E. (2022). Augmented reality in laparoscopic liver resection evaluated on an ex-vivo animal model with pseudo-tumours. *Surgical Endoscopy*, 36(1), 833–843.





---

# Acknowledgements

I'm deeply grateful to my supervisors, Adrien Bartoli, Christophe Tilmant, Emmanuel Buc, and Lilian Calvet for their guidance, support, carefulness, responsiveness and patience during all this process. They trusted me to work on a difficult problem such as Augmented Reality for Laparoscopic Liver Resection. They listened to me whenever I needed it and gave me good advice to tackle the different challenges, including the conceptual and technical ones. Adrien, thanks for believing in me, you have always been willing to help me with your expertise and everything you could. Thanks to your guidance I've been able to publish several high quality articles during my PhD and conclude it successfully. Christophe, thanks a lot for the tips you gave me to improve the experimental protocols, the writing of the articles, and on how to plan my work in order to finish it on time. This has improved the skills I will require for my future career in research. It was a pleasure to work with someone as kind as you. Emmanuel, thank you for teaching me a lot about digestive surgery. Your feedback gave me a clever idea of the problems surgeons face during laparoscopy, which was a key component to develop the methods that we propose in this thesis. Also, the opportunity you gave me to go to the hospital and do the experiments that we required was of great help to validate our methods. Lilian, your kindness and willingness to help others are out of this world. I'm glad to have earned such a good friend: you have been there to give me advice in the research, technical, and general aspects of my PhD. Your emotional support also helped me to go through these difficult pandemic times, and I will always carry that in my heart.

I want to express my gratitude as well to the jury members Stamatia Giannarou, David Fofi, and Guillaume Morel, for accepting to review my work and their constructive comments. They have been important to clarify some aspects of the work, and to better define the problems that should be solved in future work. Their feedback will also help me in my future career in research. I really feel honoured to have discussed with them about my work on the defence day.

My deepest gratitude goes for Esperanza Espinel, my mother, who has always been a pillar in everything I do in my life. Thanks to her love and support I have been able to achieve all my goals. I hope I can have you by my side for many more years. I also would like to thank Mario Espinel, my uncle, for his love and carefulness, for being like a father to me. To all my friends in Colombia and France, thanks for the funny moments we have spent together and for always being with me. They are so many that it would take several pages to name them all. I am lucky to have met such wonderful people, especially during these difficult pandemic times.

Finally, I would like to thank the Université Clermont-Auvergne, the Institut Pascal, the École Doctorale des Sciences pour l'Ingénieur and the Centre Hospitalier Universitaire de Clermont-Ferrand for giving me the opportunity to pursue this PhD, along with their logistical, financial and academic support. I enjoyed working in their facilities, which offered a comfortable environment and all the technical means to do my research. I also want to thank the Centre International de Chirurgie Endoscopique in Clermont-Ferrand for kindly letting us use their endoscopic systems to perform some of the experiments presented in this work.



# Contents

<b>Contents</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Minimally Invasive Surgery	2
1.2 Laparoscopic Liver Resection	12
1.3 Augmented Reality for LLR	16
1.4 Contributions and Thesis Structure	19
<b>2 Previous Work in AR for LLR</b>	<b>21</b>
2.1 Introduction	22
2.2 3D-3D Registration Methods	22
2.3 3D-2D Registration Methods	29
2.4 State-of-the-Art Review	33
<b>3 Technical Background on 3D-3D and 3D-2D Registration, 3D Reconstruction, and AR for LLR</b>	<b>37</b>
3.1 Introduction	38
3.2 3D-3D Registration	38
3.3 3D-2D Registration	43
3.4 3D Reconstruction	47
3.5 Augmented Reality for Laparoscopic Liver Resection	50
<b>4 Single-View Registration Combining Visual Cues with Interactions</b>	<b>57</b>
4.1 Introduction	58
4.2 Background	59
4.3 Hybrid Registration Combining a Biomechanical Model, Visual Cues and Manual Interactions	59
4.4 Quantitative Evaluation on Phantom Data	63
4.5 Quantitative Evaluation on Ex-Vivo Animal Data	67
4.6 Qualitative Evaluation on Patient Data	68
4.7 Conclusions	71
<b>5 Multi-View Registration Techniques</b>	<b>73</b>
5.1 Introduction	74
5.2 Multi-View Registration of Rigidly Linked Images	76
5.3 Multi-View Registration on Non-Rigidly Linked Images	82

---

5.4	Quantitative Evaluation on Phantom Data . . . . .	83
5.5	Qualitative Evaluation on Patient Data . . . . .	86
5.6	Conclusion . . . . .	89
<b>6</b>	<b>Keyhole-Aware AR with Tumour-Projection Visualisation</b>	<b>91</b>
6.1	Introduction . . . . .	92
6.2	Keyhole-Aware AR for LLR . . . . .	94
6.3	Quantitative Evaluation on a Geometric Model . . . . .	97
6.4	Quantitative Evaluation on Phantom Data . . . . .	103
6.5	Qualitative Evaluation on Patient Data . . . . .	107
6.6	Conclusion . . . . .	108
<b>7</b>	<b>Conclusions and Future Work</b>	<b>109</b>
<b>A</b>	<b>Acronyms</b>	<b>III</b>
<b>B</b>	<b>Notation</b>	<b>V</b>

# Chapter 1

## Introduction

---

We introduce some medical and technological aspects necessary to understand the context and goal of this thesis. First, we present Minimally Invasive Surgery (MIS), the existing types of MIS, and the advantages and drawbacks compared traditional open surgery. Then, we discuss about the Laparoscopic Liver Resection (LLR) problem, including the anatomy of the liver, the most common liver diseases, how LLR is used to treat some of these pathologies, and the difficulties the surgeons face while performing LLR. We continue by showing how Augmented Reality (AR) can help to overcome such difficulties, with a general explanation of how AR works. Finally, we present a summary of the contributions we make to improve AR in LLR.

---

## 1.1 Minimally Invasive Surgery

### 1.1.1 General Points

Minimally Invasive Surgery (MIS) is a set of surgical techniques that make use of an endoscope and thin instruments to perform an operation. Instead of a big incision in the abdominal wall, small incisions are made through which the surgical tools are inserted. The surgeon navigates the intraoperative scene with the help of the camera, observing it from an external screen. The commonly used instruments in MIS include a small camera, a pair of scissors, a cauteriser, and a vacuum/irrigation probe, as shown in figure 1.1, although this can vary depending on the MIS technique (Robinson and Stiegmann, 2004). MIS is usually prescribed for procedures of low to middle complexity, where direct access to an organ is not required (Himal, 2002).

In the case of hepatobiliary, gastric, urologic, and gynaecologic MIS, the usual procedure is as follows: 1. The peritoneal cavity is cut down and a trocar inserted; 2. The cavity is insufflated with carbon dioxide to create a bigger working area; 3. The camera is inserted and a first exploration of the scene is made; 4. The trocars for the other instruments are placed; 5. The surgery is performed by inserting and driving the instruments through the trocars; 6. Once the surgery is finished, the trocars are removed and the carbon dioxide is released; 7. The trocar incisions are closed (Hageman et al., 2008). The size of the incisions can be as little as 3 mm, causing less injury to the patient and reducing both the amount of bleeding and the risk of infection.



Figure 1.1: Basic setup for Laparoscopy, a Minimally Invasive Surgery technique. The surgeon operates using narrow instruments that are inserted through small incisions, observing the scene from an external screen (picture taken from (B. Braun, 2021)).

There exist several techniques for MIS, each of them targeting different organs or parts of the human body. They can use different sets of instruments and surgical techniques, and may operate by opening small incisions or through natural orifices. In all cases, surgeons do not have direct access to the organ or area under operation, and must follow the procedure from an external screen. This is why they are required to follow a rigorous training process, which can take several years (Bouré, 2005).

Since the first laparoscopic cholecystectomy performed by Philippe Mouret in 1987 (Litynski, 1999), MIS has been increasingly adopted in the medical community thanks to its demonstrated

benefits over open surgery for both the patient and the hospital. For the patients, benefits include reduced injury to the organs and faster recovery times. For the hospital, reduced costs thanks to shorter hospitalisation times. Methods and protocols for MIS have also evolved in the last decades, improving the safety and precision of the dissections. This has led to oncologic outcomes for the patient similar to open surgery. However, MIS still remains a challenging procedure due to the lack of palpation and the limited range of movements (Lee et al., 2013).

### 1.1.2 Types of Minimally Invasive Surgery

MIS techniques can be classified depending on the area of the human body where it is applied to. According to (Sun and Fong, 2017), there exist techniques for head and neck surgery, esophageal surgery, lung surgery, gastric surgery, colorectal surgery, hepatobiliary surgery, gynaecologic surgery, and urologic surgery. Other specialities include maxillofacial surgery, orthopedic surgery like arthroscopy of the knee, cardiovascular surgery, and plastic surgery. Most of these techniques target the removal of cancerous tissues, and may use classical or robotic endoscopic systems. Each of them has its own preparation and operation protocol, as well as its own set of instruments. Some of the most commonly used techniques are:

- **Head and neck:** One of the most popular procedures is Transoral Robotic Surgery. It uses a surgical robot to remove tumours from hard-to-reach areas in the mouth and throat (Hockstein et al., 2006). The instruments are inserted through the patient's mouth, without the need of extra incisions. Therefore, it avoids cutting the jaw to access the base of the tongue or the throat, which is a delicate procedure. This technique is used to treat pharyngeal, laryngeal, and tongue cancers, as well as obstructive sleep apnea. An example of TORS applied in a pediatric patient is shown in figure 1.2. Another popular technique is Minimally Invasive Brain Surgery, which aims to precisely and safely remove tumours located under the skull, in a way that minimises the damage to the surrounding tissues (Rosenfeld, 1996). It opens small incisions of less than 2 cm in the skull, to insert the instruments that will be used to resect the tumours. Some approaches also insert the instruments through the nasal passage, without opening holes in the skull. Common pathologies treated using these techniques include pituitary tumours, craniopharyngiomas, and sinus carcinomas.



Figure 1.2: Transoral Robotic Surgery on a pediatric patient (Leonardis et al., 2013).



- **Esophagus:** Thoracoscopic Esophagectomy is a minimally invasive procedure to remove cancerous tissues from the esophagus. It does so by opening small incisions in the abdomen and the chest, through which the camera and instruments are inserted (Alvelo-Rivera et al., 2004). This avoids the opening of large incisions and the mobilisation of the ribs, which increases the risk of morbidity and the recovery times for the patient. As most minimally invasive approaches, it is technically more complex than a traditional open esophagectomy, with a very steep learning curve. An example of how the ports are placed in thoracoscopic esophagectomy is shown in ???. A similar technique is Robotic-assisted Esophagectomy, which follows the same principle as the thoracoscopic case but uses a surgical robot to guide the resection. It is especially useful when the tumour is located in places of difficult access or require a fine dissection (Jha et al., 2020).

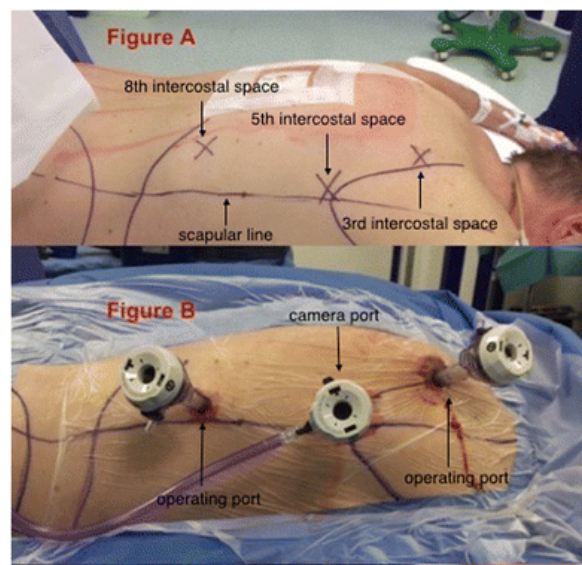


Figure 1.3: Patient and keyhole positioning for Thoracoscopic Esophagectomy (Charalabopoulos et al., 2017).

- **Lung:** Video-assisted Thoracoscopic Surgery is a popular procedure to diagnose and treat problems in the lungs and chest. Problems may include the biopsy and removal of lung cancer, esophageal cancer, removal of the thymus gland, among others. It opens several small incisions between the ribs to insert the thoracoscope and the surgical instruments. Compared to conventional laparoscopy, this technique does not require insufflation with carbon dioxide, thanks to the rigidity of the thoracic cage. It also avoids the opening of large incisions, and the separation or cutting of ribs and muscles, reducing recovery times and long term complications. A similar approach that uses a surgical robot instead of a conventional thoracoscope is Robotic-assisted Thoracic Surgery. It offers increased accuracy in the surgical gestures and reduced damage to the organs compared to the conventional approach (Sun and Fong, 2017).
- **Gastric:** Laparoscopic Gastrectomy is an increasingly popular procedure for the treatment of stomach cancer, thanks to the reduced blood loss, faster recovery times, and increased survival rates compared to traditional gastrectomy. It commonly uses four 12 mm trocars placed in the abdomen: one for the laparoscope and three for the instruments. Three types



Figure 1.4: Keyhole positioning for video-assisted thoracoscopic lobectomy (Agasthian, 2010).

of gastrectomy can be done: Partial gastrectomy, where the lower part of the stomach is removed; sleeve gastrectomy, where up to three quarters are removed; and total gastrectomy, where the whole stomach is removed. An example of total laparoscopic gastrectomy with the corresponding keyhole positioning is shown in figure 1.5. However, Laparoscopic Gastrectomy is usually recommended for early stage cancers rather than advanced ones, due to the difficulty of resecting big tumours and other tissues using the laparoscopic approach (Caruso and Scatizzi, 2022). A variant of this technique is the Robotic-assisted Laparoscopic Gastrectomy, which is gaining attention for its ability to perform delicate procedures like intracorporeal reconstruction after gastrectomy with high precision. It also has a reduced learning curve compared to traditional laparoscopy, which translates into a faster diffusion of minimally invasive surgery for gastric cancer (Caruso et al., 2016).

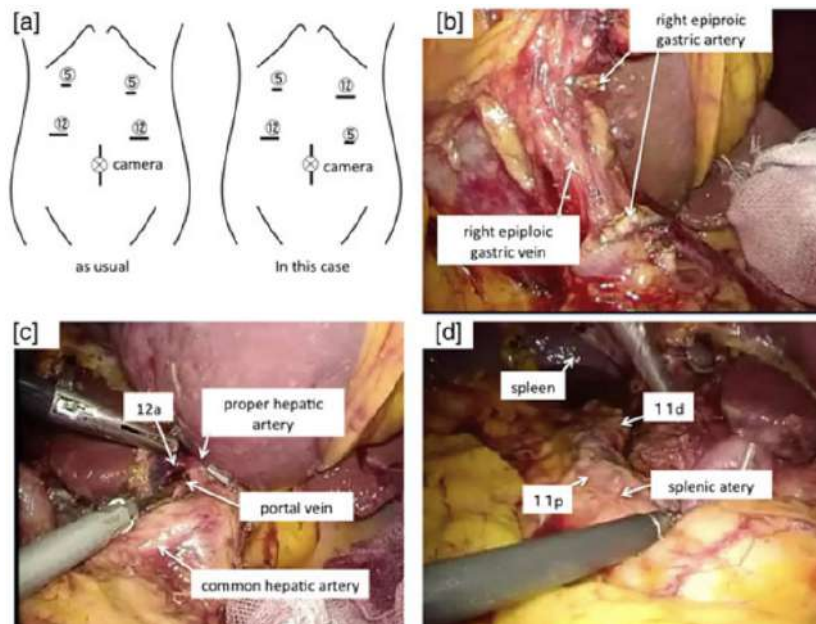


Figure 1.5: Example of keyhole positionings and intraoperative images for total laparoscopic gastrectomy (Shibata et al., 2018).

- **Colorectal:** A common approach in laparoscopic colorectal surgery is Laparoscopic Colec-

tomy. It is used to treat several affections including colon cancer, bowel obstruction, uncontrollable bleeding, among others. Depending on the severity of the affection, a partial or total colon resection can be performed. The partial colon resection may consist of a right or left colectomy. Three to five trocars are placed in the abdomen, and their locations depend on the section of the colon to be removed (Ahad and Figueredo, 2007). Example views of a laparoscopic colectomy are shown in 1.6. Similarly to other organs, the robotic version of this technique is called Robotic-assisted Colectomy, for which several reports have shown less complications for the patient during surgery, a reduced hospital stay, and a reduced probability of switching to an open colectomy (Benlice et al., 2017).

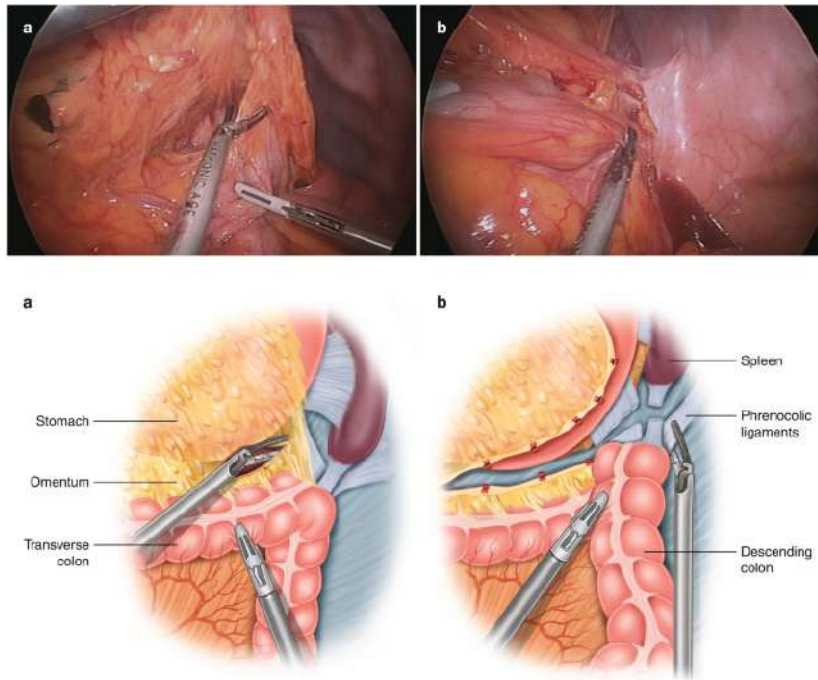


Figure 1.6: Example views of a laparoscopic colectomy (Eglinton and Frizelle, 2019).

- **Gynecologic:** Laparoscopic Surgery of Uterus is a popular choice for the removal of benign and malignant uterus cancers, removal of uterine fibroids (myomectomy), complete removal of the uterus (hysterectomy), among others. Three to five incisions are opened in the abdomen, depending on the size and the location of the tumour, or the type of surgery to be done. Compared to open techniques, laparoscopy of the uterus has shown reduced blood loss, low conversion rates, less postoperative pain, and shorter hospital stays (Bhave Chittawar et al., 2014). A basic setup for laparoscopic hysterectomy is shown in 1.7 Robotic-assisted Uterus Surgery has been related with a reduced difficulty in dissection and suturing, although no clear evidence of lower intraoperative and postoperative complications over laparoscopic surgery has been demonstrated in the literature (Gingold et al., 2018).
- **Urologic:** Laparoscopic Prostatectomy is a well known technique to remove a patient's cancerous prostate. Compared to traditional open prostatectomy, laparoscopic approaches show similar resection accuracies, while offering a magnified view of the anatomical structures, reduced blood loss, and increased postoperative quality of life to the patient. It operates in a similar way as open prostatectomy, but using minimally invasive instruments

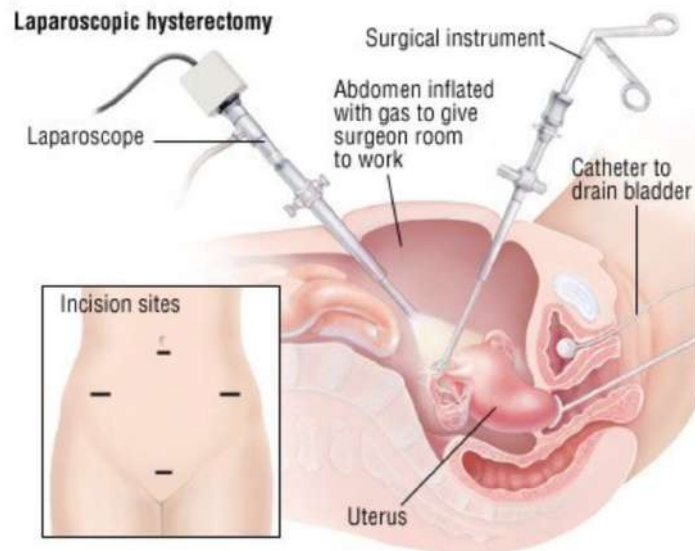


Figure 1.7: Basic keyhole placement and setup for laparoscopic hysterectomy (Harvard Health Publishing, 2022).

that are inserted through five or six trocars placed in the patient's lower abdomen (Touijer and Guillonnet, 2004; Bove et al., 2009). For the robotic-assisted case, several studies show a global increase in the resection accuracy thanks to the more precise surgical gestures, with similar intraoperative and postoperative complications compared to laparoscopic approaches (Berryhill et al., 2008). Laparoscopic Nephrectomy is the minimally invasive technique for the resection of tumours in the kidney. It is divided into partial nephrectomy (partial removal of the kidney) and radical nephrectomy (complete removal of the kidney). Partial nephrectomy is advised for solitary and small tumours with less than 4 cm in size; otherwise, a radical nephrectomy is advised (Rashid et al., 2008). Four to five trocars are placed in one of the sides of the abdomen, with their arrangement depending on the size of the kidney, as well as the location of the tumour (Johnston and Wolf, 2005). An example of trocar positioning for laparoscopic nephrectomy is shown in figure 1.8 Regarding Robotic-assisted Nephrectomy, several studies have found that it does not bring clear advantages of efficiency and safety over traditional laparoscopy, which does not compensate for the high costs of a robotic surgical system (Caruso et al., 2006).

- **Hepatobiliary:** Laparoscopic Hepatectomy and Laparoscopic Pancreatectomy are the minimally invasive approaches for the treatment of cancer of liver and pancreas, respectively. Laparoscopic Hepatectomy, also called Laparoscopic Liver Resection (LLR), is the preferred procedure to remove early-stage Hepatocellular Carcinoma (HCC). HCC is the most frequent liver cancer and one of the main causes of cancer deaths worldwide (Njei et al., 2015). In a typical LLR setup, four to six trocars may be placed in the patient's abdomen, with their distribution depending on the location of the tumour and the type of resection to be done (Thiruchelvam et al., 2021). The difficulty of LLR may vary depending on the size and position of the tumours, being less complex for tumours smaller than 5 cm and located at the anterior (frontal) part of the liver (Jia et al., 2018). In the case of the pancreas, Laparoscopic Pancreatectomy is a challenging procedure due to the location of the organ behind the stomach and its proximity to major vasculature. It is usually recommended for early-stage,

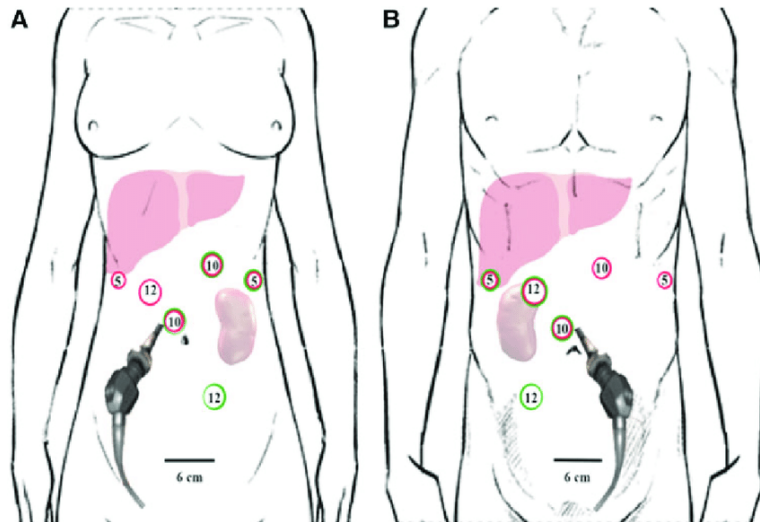


Figure 1.8: Keyhole positioning for lateral sectionectomy and nephrectomy. (A) Keyhole position for laparoscopic left lateral sectionectomy (red) and left nephrectomy (green). (B) Trocar position for laparoscopic left lateral sectionectomy (red) and right-hand-assisted nephrectomy (green) (Angelico et al., 2019).

isolated tumours located in the distal pancreas (Liang et al., 2014). During the operation, four to five trocars are placed in the upper abdomen, the stomach is lifted and the spleen may be removed to have better access to the pancreas (de Rooij et al., 2015). Both laparoscopic approaches can be performed using surgical robots, which are referred to as robotic-assisted hepatectomy and robotic-assisted pancreatectomy, respectively. In the first case, several studies have found no significant differences between the robotic and laparoscopic approaches with respect to their effectiveness in resection, intraoperative blood loss, and postoperative complications (Liu et al., 2017). In the second case, the robotic approach has shown advantages when performing resection while preserving the spleen, which requires fine and precise gestures. Also, robotic assistance has shown decreased rates of conversion to open surgery, and shorter recovery times compared to the laparoscopic approach (Kauffmann et al., 2020). An example of keyhole placement and the corresponding laparoscopic image for robotic hepatectomy is shown in 1.9. A more detailed description of LLR is given in section 1.2.

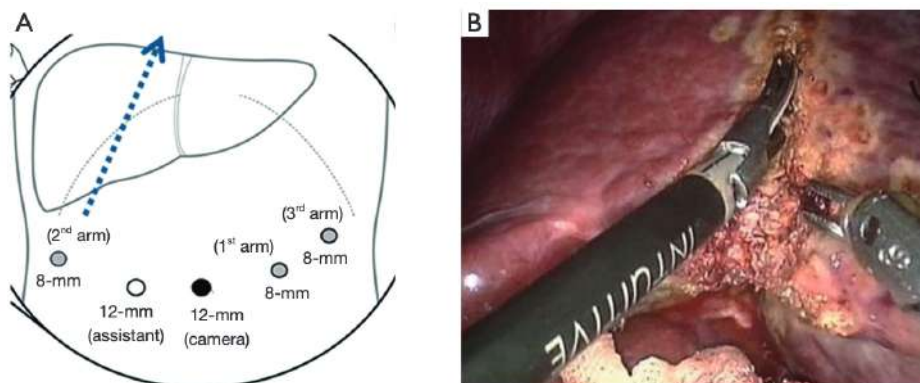


Figure 1.9: Example of keyhole placement and laparoscopic view for robotic hepatectomy (Choi et al., 2017).

### 1.1.3 Advantages and Drawbacks of Minimally Invasive Surgery over Open Surgery

- **Advantages:** Compared to traditional open surgery, in which big incisions are made, patients can recover quickly with MIS thanks to the faster wound healing times. The small incisions translate into a reduced trauma to the patient, less wound infections and a lower risk of deadly intraoperative and postoperative complications. While MIS techniques may be more complicated and operating times can be longer, hospitalisation times are significantly shorter, reducing the total costs for the patient and the hospital. Some MIS procedures can even be ambulatory if the level of complexity allows it. Reports have shown that there is a significant reduction in blood loss, along with less postoperative morbidity in patients undergoing MIS (Chen et al., 2017). Thanks to these benefits, demand for MIS has increased in the last decade, although its adoption remains low compared to open surgery. This is mainly due to the fact that MIS requires a special training and a solid experience of the surgeon, which can take several years. However, advances in training strategies and endoscopic systems allow surgeons to be more prepared for solving most of the complications inherent to MIS. This has made MIS the first option to consider when choosing a treatment for a particular pathology, especially in the gastrointestinal, colorectal and gynecologic fields (Okunrintemi et al., 2016; Siddaiah-Subramanya et al., 2017; St John et al., 2020).
- **Drawbacks:** One of the main drawbacks of MIS is that it requires specialised training and different skills than open surgery, with usually long learning curves. As the surgeon does not have direct access to the area under intervention, they must watch the scene through an external screen. The lack of tactile feedback, depth perception and the reduced space can increase the difficulty of the operating conditions. This translates into a higher risk of damage to the organ, hemorrhages due to accidental cut of vessels, and inaccurate tumour resections, to name a few issues. Because of such operational constraints, procedures usually take a longer time as surgeons have to be more cautious with the surgical gestures (Driessen et al., 2017; Kaplan et al., 2016; Li et al., 2021). From the patient side, some of the common complications from MIS include bleeding of the surgical incisions, hemorrhages due to accidental vascular cuts, clip slippage, bile leakage, bowel injury and postoperative adhesions (Doddamani et al., 2018).

Nevertheless, MIS continues to grow in popularity within the medical community as the benefits largely outweigh the drawbacks. With proper training and the right equipment, surgeons can solve most of the aforementioned complications without having to switch to an open intervention. In the same way, several assistive techniques have been developed in order to better guide surgeons throughout the surgical procedure. Among these techniques are fluorescence imaging, ultrasound-based navigation, preoperative and intraoperative CT imaging, and augmented reality.

### 1.1.4 Assistive Interventional Techniques for MIS

- **Fluorescence-guided Surgery (FGS):** Near-infrared Fluorescence is a common application of FGS. The usage of this technique requires the injection of Indocyanine Green (ICG), a cyanine dye used in medical diagnostics. It helps in determining cardiac output, hepatic function, liver and gastric blood flow, among other functionalities (Nagaya et al., 2017; Zheng

et al., 2019). ICG-enhanced fluorescence has proved to be useful at improving intraoperative anatomical view during minimally invasive interventions and thus increase patients' safety during difficult surgical procedures (Watanabe et al., 2015; Aoki et al., 2018). This is thanks to the fact that ICG is absorbed and retained by cancerous cells for several hours. Regarding the working principle, fluorescence is produced when the water-soluble ICG is excited using a laser beam or a light of specific wavelength in the near-infrared spectrum, of about 800 nm, and is visualised using specific optics and camera systems. In the case of MIS, a system of filters, lenses and cameras is attached to the end of the probe (Gray et al., 2012). ICG is almost entirely metabolised by the liver and excreted into the bile. It is thus a cheap way of highlighting the internal structures for real-time image-guided interventions. (Alander et al., 2012; Urade et al., 2020). An example of this visualisation technique for laparoscopic liver resection is shown in figure 1.10.



Figure 1.10: Example of hepatic tumours highlighted using ICG on two different patients. Tumours are visualised by fusing data from the ICG sensor with the camera's RGB image (Hamabe et al., 2019).

- Laparoscopic Ultrasound (LUS):** This is another common tool for finding inner structures intraoperatively in MIS. It consists mainly of a mobile computing machine with a screen, along with the Ultrasound (US) endoscopic probe. LUS can provide high-resolution images of the abdominal cavity and help in examining the anatomical structures of abdominal viscera and their relative positions. Therefore, acquisition of laparoscopic images via ultrasound is useful for guiding the surgery intraoperatively, reducing potential errors in the procedures. LUS has been widely applied to many parts of the human body for the diagnosis and treatment of multiple diseases. For instance, it has been used for treating gallbladder diseases, such as Intraoperative Cholecystectomy, Choledochal Cysts Excision and Biliary Drainage (Dili and Bertrand, 2017). It also plays an important role in Hepatectomy and Hepatic Cysts Excision (Majbar et al., 2014), as well as in Splenectomy, Spleen Fixation, and Pancreatectomy (Sun et al., 2010). It can be also applied in the assessment of a cancer stage, diagnosis of intraperitoneal tuberculosis, treatment of peritoneal adhesion, and resection of various tumours on abdominal organs (Berber and Siperstein, 2004). The usual procedure for tumour resection consists in finding the tumour and its surrounding vessels with LUS, marking the resection margin in a way that the whole tumour is contained inside, performing the planned resection, and doing a final check with LUS for potential remains of the tumour. The LUS images are usually seen from an external machine, as shown in figure 1.11.
- Augmented Reality (AR):** In the last two decades, AR has become a source of interest in

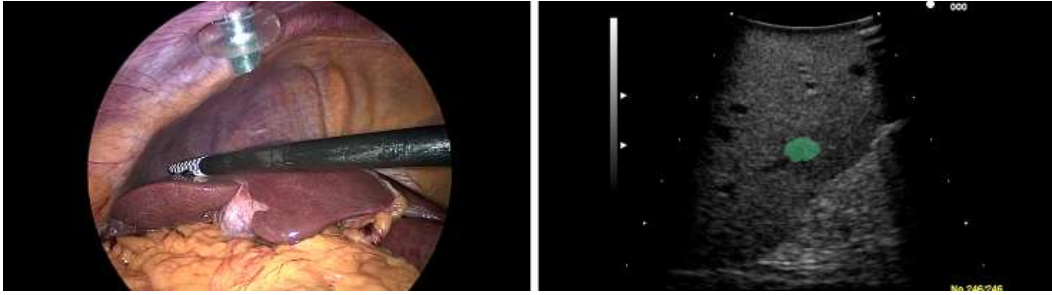


Figure 1.11: Localization of inner hepatic tumour using LUS: Laparoscopic image showing the LUS probe (left) and corresponding US image with segmented tumour in green (right) (Rabbani et al., 2022).

MIS thanks to the possibility of overlaying different kinds of information into the intraoperative 2D images. The information shown can include preoperative 3D models, resection paths and LUS images. The preoperative 3D models are reconstructed from Computed Tomography (CT) or Magnetic Resonance (MR) data and include tumours, vessels and other structures. AR avoids the need of mentally mapping different sources of information into the laparoscopic scene. AR has been introduced in almost all the specialties concerned by MIS. One of the most studied ones neurosurgery, in which structures are generally rigid, facilitating the registration process and making AR the standard in this field (Guha et al., 2017). Other fields where AR is widely used include Maxillofacial Surgery (Zhu et al., 2017), Otorhinolaryngology (Wong et al., 2018), and Orthopedics (Jud et al., 2020). Regarding abdominal surgery, AR for MIS is still in early stages due to the deformable and mobile nature of the inner organs and structures. Preliminary tests have been done for Gynaecologic Surgery (Collins et al., 2021), Urologic Surgery (Gadjiev et al., 2020), Hepatobiliary Surgery (Luo et al., 2020), and Colorectal Surgery (Shen et al., 2020). Another reason for the slow adoption of AR in laparoscopy is that most of the developed strategies make use of stereo endoscopes, tracking devices and other accessories that are not commonly available in surgery rooms. Two examples of AR for MIS using LUS images and preoperative 3D models are shown in figure 1.12.

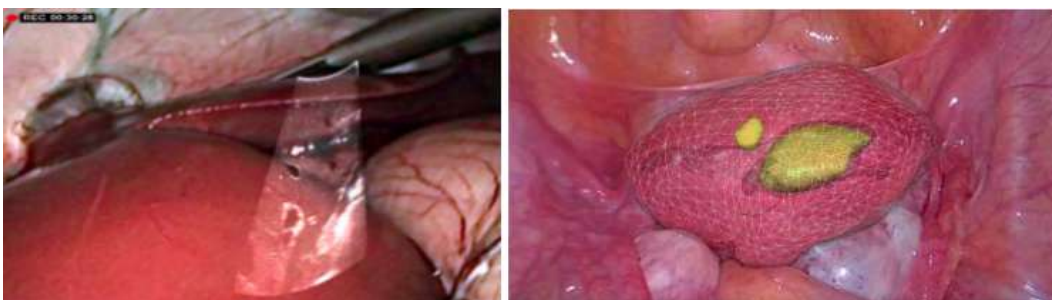


Figure 1.12: Examples of AR for MIS: Laparoscopic image of a porcine liver augmented with LUS (left) (Kang et al., 2014); Laparoscopic image of a human uterus augmented with preoperative 3D models (right) (Collins et al., 2021).



## 1.2 Laparoscopic Liver Resection

### 1.2.1 General Points

LLR is a minimally-invasive procedure whose goal is to remove a liver mass (benign or cancerous) safely and efficiently. The procedure is done with the patient under general anaesthesia, while three to seven small incisions are made to remove the liver mass. The number of incisions and the duration of the surgery will depend on the number and locations of the tumours, and how much of the liver needs to be removed. For a patient to be suitable for LLR, surgeons must take into account the size and location of the tumours, as well as the presence of cirrhosis. Tumours no larger than 5 cm in diameter and located in peripheral areas (segments II to VI) are more suitable for LLR. Depending on the expertise of the surgeon, difficult segments VII, VIII and IVa could also be treated, even though operation times might be longer (Jia et al., 2018). These segments are illustrated in section 1.2.2. In an effort to improve efficiency, a security resection margin of 1 cm has been advised in the literature for resection of HCC (Lazzara et al., 2017). This will help to ensure the whole tumour mass is extracted and no remnants of cancerous cells are left in the liver.

### 1.2.2 Anatomy of the Liver

The liver is a peritoneal organ positioned in the right upper quadrant of the abdomen. It is the largest visceral structure in the abdominal cavity, and the largest gland in the human body. It is an accessory digestive gland whose main goal is to clean the blood coming from the digestive tract, before passing it to the rest of the body. The liver also detoxifies chemicals and metabolises drugs. While doing so, the liver secretes bile that ends up back in the intestines. The liver also makes proteins important for blood clotting and other functions. A human liver usually weighs approximately 1.5 kg and has a width of about 15 cm, although its size varies a lot across subjects. The liver has two large sections, called the right and the left lobes. The falciform ligament makes a superficial division between the two lobes, as shown in figure 1.13. The gallbladder sits under the liver, along with parts of the pancreas and intestines.

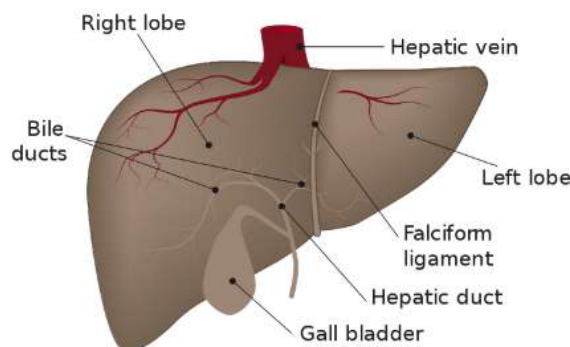


Figure 1.13: Anatomy of the human liver (Sheasby, 2011).

The liver can be also divided in eight functional segments using the Couinaud system (Germain et al., 2014). This is done by following a transverse plane through the bifurcation of the main portal vein. Each segment has its own vascular inflow, outflow and biliary drainage, having branches of the portal vein, hepatic artery, and bile duct. The classification uses the vascular supply in the liver to separate the segments (going from I to VIII) with segment I, the caudate

lobe, receiving its supply from the portal vein's right and left branches. The rest of the segments (II to VIII) are numbered in a clockwise fashion, as shown in figure 1.14. The division of the liver into segments allows the surgical resection of individual sections without damaging the remaining ones. Therefore, for the liver to remain functional, resections are done along the hepatic and portal veins in the planes that define the boundaries of the segments (Majno et al., 2014).

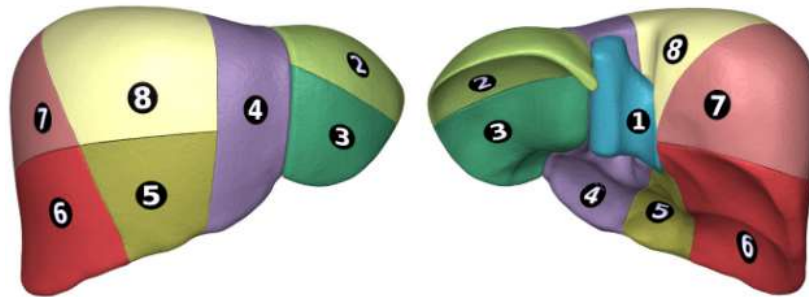


Figure 1.14: Anterior and posterior views of the liver with its functional segments according to the Couinaud system (Database Center for Life Science (DBCLS), 2015).

### 1.2.3 Common Liver Diseases

Liver diseases are one of the main causes of mortality worldwide, with cirrhosis and liver cancer being the 11th and 16th most common causes of death respectively. Alcoholic habits, obesity and diabetes are frequent risk factors for chronic liver diseases (Asrani et al., 2019). Some of the most common diseases are:

- **Alcohol-associated Liver Disease (AALD):** It is a spectrum of diseases which includes alcoholic fatty liver with or without hepatitis, alcohol hepatitis (reversible) to cirrhosis (irreversible). Over 50% of mortality with cirrhosis is attributable to alcohol globally. Evolution to cirrhosis may be higher in patients with AALD compared to Non-alcoholic Fatty Liver Disease.
- **Non-alcoholic Fatty Liver Disease (NAFLD):** It has a strong relationship with obesity, hyperlipidemia, and diabetes mellitus. NAFLD encompasses two different conditions: Non-alcoholic Fatty Liver and Non-alcoholic Steatohepatitis. The former is characterised for developing steatosis and lobular inflammation, while the latter shows varying degrees of fibrosis, cirrhosis and HCC.
- **Hepatitis:** It is an inflammation of the liver, usually caused by a viral infection or a liver damage for drinking alcohol. There are several types of hepatitis, some of which will pass without serious problems, while others can be chronic and cause scarring of the liver (cirrhosis), loss of liver function and, in some cases, liver cancer. While Hepatitis A, B, C, D and E are caused by viruses, there is also the Alcoholic Hepatitis caused by drinking excessive amounts of alcohol. Most types of hepatitis can degenerate in cirrhosis or liver cancer if they are not treated properly.

### 1.2.4 Pathologies Treated in LLR

LLR can be used in the treatment of several liver pathologies, whether they are benign or malignant and regardless of the liver condition (Schmelzle et al., 2020). These include:

- **Benign Liver Tumours (BLT):** They are masses of cells originating in the liver and without the ability of metastasising or invading neighbouring organs. They can consist of focal nodular hyperplasia, hemangioma and hepatocellular adenoma. Using LLR on BLT has shown a significant improvement compared to open surgery with respect to lower overall morbidity, reduced blood loss and shorter hospital stay (Troisi et al., 2008). One of the challenging parts of BLT is that they are mostly asymptomatic, only showing symptoms when they reach several centimeters in diameter (Wabitsch et al., 2019).
- **Colorectal Liver Metastases (CRLM):** They are malignant tumours that originally develop in the colon or rectum and have spread to the liver through the portal vein. Up to 70% of patients with colorectal cancer eventually develop liver metastases. Similarly to BLT, usage of LLR on CRLM has shown reduced blood loss, shorter hospital stay and reduced morbidity compared to open surgery (Zhang et al., 2017).
- **HCC:** It is the most common type of liver cancer in adults and the most common cause of death in people with cirrhosis. It is closely related to viral hepatitis B and C infections, or exposure to alcoholic toxins. Metabolic syndrome and non-alcoholic steatohepatitis are also strong risk factors for developing HCC. It is also the main pathology treated with LLR, with advantages on reduced trauma for the patient and faster recovery times compared to open surgery (Andreou et al., 2018).

### 1.2.5 Determining Conditions for Surgery

Feasibility of LLR depends on the size of the tumour, its location, presence of cirrhosis and whether the remaining liver tissue will ensure a correct hepatic function. Optimal candidates are solitary tumours with no evidence of vascular invasion, and located far from major vessels (Ocuin and Tsung, 2016). It has been usually difficult to perform LLR on the posterosuperior parts of the liver (segments VII and VIII) due to their poor exposure and difficulty to control bleeding. The Louisville consensus advises that solitary lesions not bigger than 5 cm and located in the peripheral regions of the liver (segments II to VI) are appropriate for LLR (Belli et al., 2013). However, thanks to advancements in surgical techniques and increased experience of the surgeon, segments VII and VIII can now be safely treated with LLR, as well as tumours bigger than 5 cm (Xiang et al., 2015). For patients suffering from cirrhosis, LLR should be done more cautiously as it is more difficult to identify the anatomical structures as well as to dissect the cirrhotic liver parenchyma. Compared to healthy liver, LLR for cirrhosis liver thus requires longer operation times (Yoon et al., 2017). Depending on the size and location of the tumour, different LLR techniques can be applied to ensure a successful resection.

### 1.2.6 Common Procedures for LLR

As mentioned before, there are multiple resection techniques that can be applied depending on the location and size of a particular tumour. A resection can be anatomical if it extracts one or

several anatomical segments of the liver, or non-anatomical if only a small portion of parenchyma is extracted. Some of the most common techniques include:

- **Wedge Resection:** It is a non-anatomical type of resection that involves cutting the parenchyma around the tumour, with a margin of healthy tissue surrounding it. It is advised only for superficial HCC with less than 3 cm in diameter, or for benign tumours where no security margin is required (Bilal and Samah, 2013).
- **Segmentectomy:** It is the anatomical resection of a complete section supplied by a segmental branch of the portal vein, and is the most common procedure for HCC. For this procedure to be feasible, the tumour should be confined to a single segment and all vascularity under the tumour's influence should be resected (Yoshida et al., 2019). It is advised over non-anatomical resection because of the high probability of the HCC cancerous cells of spreading through the portal venous system (Yasui and Shimizu, 2005).
- **Left and Right Hepatectomy:** It consists on removing several segments of the liver, generally divided in segments V to VIII for right hepatectomy, and II to IV for left hepatectomy. This kind of resection is suitable for tumours that are big enough to cross several segments, or for those located in a position where it is not possible to perform wedge resection or segmentectomy (Aragon and Solomon, 2012).
- **Left and Right Lobectomy:** It is an extended version of hepatectomy and involves the resection of all the segments on the sides of the falciform ligament, namely segments IV to VIII (and sometimes I) for right lobectomy, and segments II to III for left lobectomy (Aragon and Solomon, 2012).

These techniques can be performed using traditional laparoscopy, where the surgeon stands in front of the patient and directly manipulates the surgical tools, or using robot-assisted surgery, where the tools are attached to a surgical robot and the surgeon controls them from a separate console. Usage of robot-assisted surgery has increased in the last decade, mainly due to the increased comfort and precision offered to the surgeon. However, clinical outcomes in terms of resection accuracy and postoperative recovery times remain similar to those of traditional laparoscopy (Sheetz et al., 2020; Wang et al., 2021).

### 1.2.7 Difficulties Specific to the Liver

The liver is the heaviest internal organ and the largest gland in the human body, measuring up to 15 cm in its longest dimension (Patzak et al., 2014). It is also very elastic, easily deforming with heartbeats, breathing, pneumoperitoneum and instrument manipulations. Surgeons must then take into account such deformations while mentally mapping preoperative 3D data into the laparoscopic images, which does not guarantee a precise location of the inner structures.

A standard 10 mm laparoscope usually has a Field-of-View (FoV) between 50° and 60°. When inserted in the abdominal cavity, its maximum distance from the organ is around 15 cm (Tsai et al., 2010). Even with the zoom at the minimum settings, such conditions will make surgeons to mostly get partial views of the organ, as shown in figure 1.15. Therefore, it can be difficult for surgeons to navigate the scene and to keep track of the inner structures. They should constantly move the laparoscope in an effort to improve the spatial perception and better understand the scene. These

partial views can also be problematic when performing major resections like hepatectomies or lobectomies.

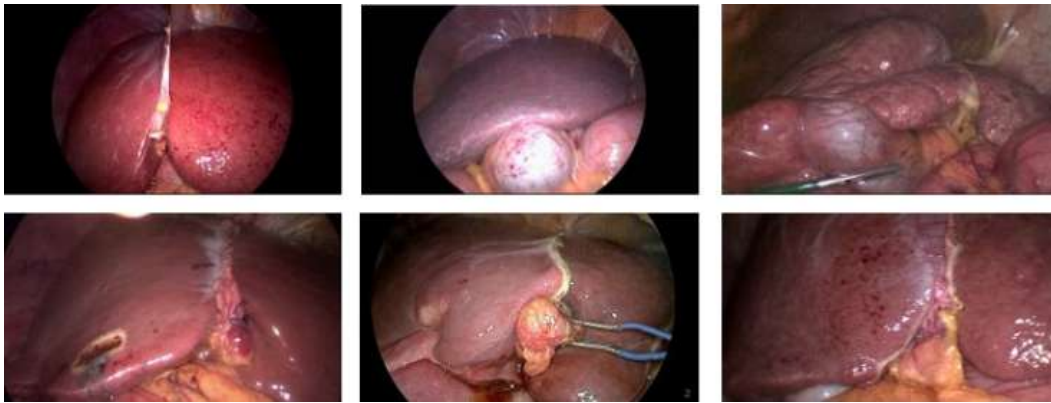


Figure 1.15: Examples of partial visibility of the liver in laparoscopy. This is caused by a combination of the liver's big dimensions, the short distance between the camera and the liver, and the camera's reduced FoV.

### 1.3 Augmented Reality for LLR

#### 1.3.1 Principle

AR aims to overcome the aforementioned issues by showing to the surgeon the liver's inner structures directly on top of the laparoscopic images. In this work we focus on augmenting the images with preoperative 3D models. These models represent the inner structures the surgeon wants to see, which usually include tumours, arteries and veins. They are reconstructed with the help of the surgeon from CT/MR data before the surgery (preoperative stage). The day of the surgery (intraoperative stage), the preoperative 3D models are deformed and aligned into the laparoscopic images. This process is called 3D-2D deformable registration, as illustrated in figure 1.16. In this way, the surgeon can be confident that the 3D models shown represent the real locations of the inner structures.

#### 1.3.2 Benefits of AR in Laparoscopy

The main advantage of AR is that it helps the surgeon to see the inner structures from a single video source, without having to look at separate screens and to mentally map other information into the laparoscopic images. Another advantage is that, if properly used, there would not be a need for third party devices, including LUS and tracking systems, which can liberate the surgeon from using extra probes or instruments. This means that a surgeon would not have to double check if they are looking at the right liver zone, reducing the total operation time and improving the quality of the resections, as the surgeon could operate using the augmented images directly. For these reasons, an AR system should guarantee an accurate registration at all times.

#### 1.3.3 General Pipeline

Although several AR solutions have been proposed for liver laparoscopy (Bernhardt et al., 2017; Vávra et al., 2017), they all share some similarities. We thus provide a general pipeline for AR in LLR in figure 1.17.

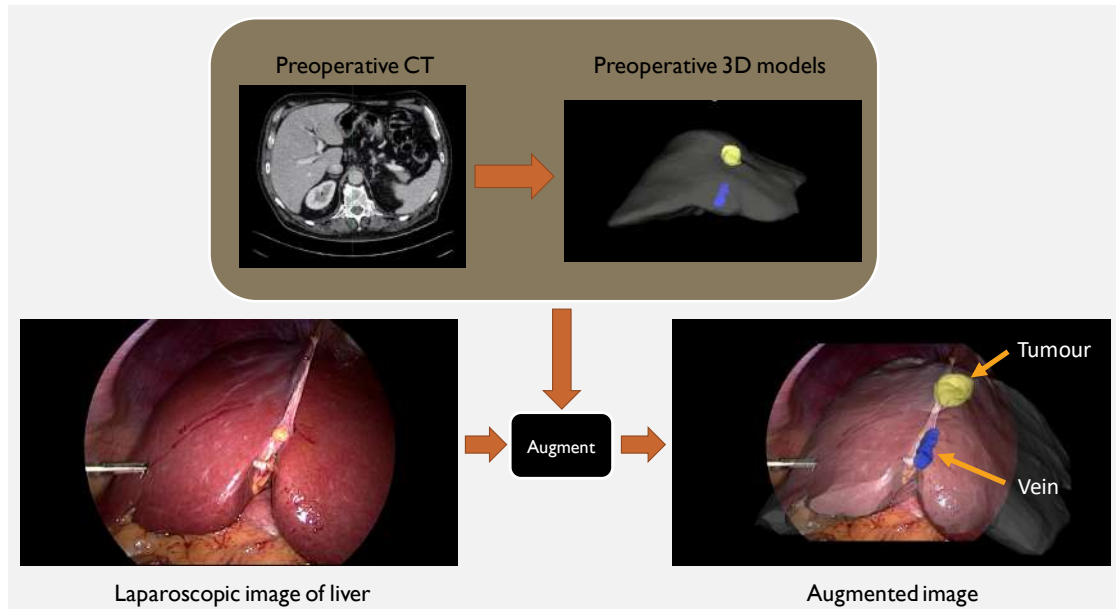


Figure 1.16: Augmented reality from the registration of preoperative 3D models into a laparoscopic image of the liver.

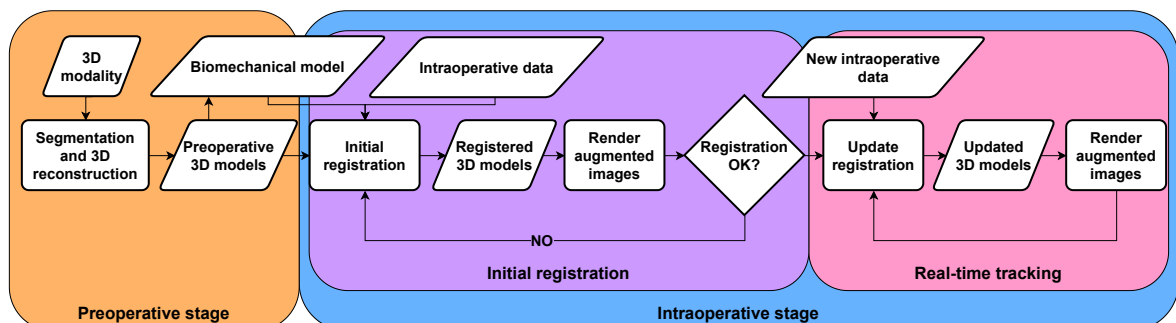


Figure 1.17: General pipeline for AR in laparoscopy.

As illustrated in figure 1.17, the AR pipeline is divided in two main stages: preoperative and intraoperative. In the preoperative stage, the CT/MR data is segmented by the surgeon/radiologist and the preoperative 3D models of the desired structures are generated. In the intraoperative stage, the process is divided in two steps: initial registration and real-time tracking. For the initial registration step, once the patient has been insufflated and the instruments inserted, a video of the scene is taken by covering all the visible parts of the liver. From the recorded video, a set of images is chosen and visual cues are marked or detected in each of them. These visual cues are the key to deform the preoperative 3D models in a correct way, and may include the organ's boundaries, anatomical landmarks and surface texture. Then, an initial registration is done of the preoperative 3D models to the intraoperative 2D images by combining the preoperative 3D models with a biomechanical model and the visual cues. The biomechanical model describes the elastic behaviour of the organ and it is necessary to achieve realistic deformations (Haouchine et al., 2013; Özgür et al., 2018). Because the objective of such registration is to locate the preoperative 3D models in real-world positions, the result should be as precise as possible. The registration is then rendered on all the images to be validated by the surgeon. An example of initial registration is illustrated in figure 1.18. If the registration is good, we proceed to the real-time tracking step. For every new incoming image, the 3D models are registered accordingly in a real-time fashion, taking the initial registrations as reference. The surgeon can then use such augmentations to guide the resection.

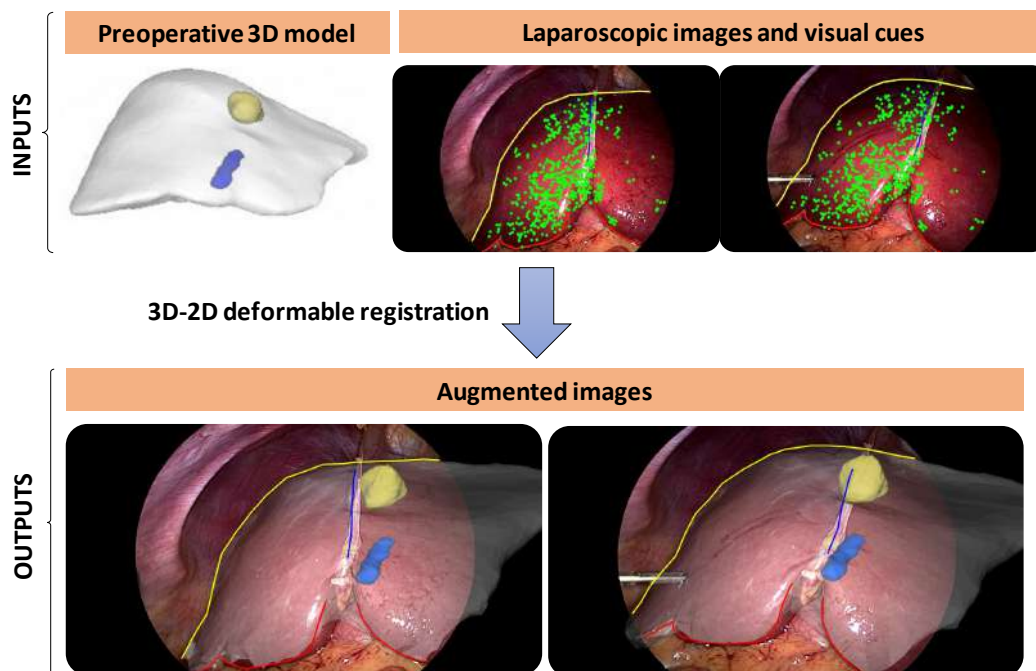


Figure 1.18: Example of initial registration in AR for liver laparoscopy. The preoperative 3D model is combined with the visual cues in the laparoscopic images to perform a 3D-2D deformable registration that fits the 3D model in the images. The yellow model corresponds to the tumour, while the blue model is the vein. The red, yellow, and blue curves correspond to anatomical landmarks of the liver used in registration. The green points are keypoint features extracted from the liver's surface, also used in registration.

In this thesis we focus on the initial registration step, which is a critical part of the AR pipeline and an open research problem. Because the subsequent tracking step usually depends on the initial registration, our target is to have an error lower than 1 cm, which is the advised security resec-

tion margin for HCC interventions (Zhong et al., 2017). Our work takes as basis the method developed in (Koo et al., 2017), which performs 3D-2D deformable registration on a single laparoscopic image, using the liver’s contours as deformation constraints. We will refer to it as the *Single-View Biomechanical (SVB)* method throughout the manuscript. We also address the problem of camera-based visualisations in AR, which can lead to inaccurate resections due to a wrong understanding of the tumour locations.

## 1.4 Contributions and Thesis Structure

The goal of this work is twofold. First, we aim to alleviate the problems coming from the liver’s partial visibility for the initial registration step by integrating information from the whole scene. We do this by either combining visual cues with manual interactions for registering the hidden parts of the liver in an image, or by performing registration using multiple images simultaneously. This way, we improve the registration accuracy and set a basis for future real-time tracking methods. Second, we enhance AR visualisation by projecting a tumour’s shape towards a particular keyhole, and showing the corresponding resection path. With this, the surgeon can be more confident of the location of the tumour and the direction to follow for resection. Concretely, we make three contributions in this work:

- **Contribution #1:** We propose a single-view 3D-2D deformable registration method that combines visual cues with manual interactions to take advantage of the information available in the whole scene. Here, the surgeon can register the hidden parts of a preoperative 3D model in a laparoscopic image, according to the information they see from the rest of the scene. The manual interactions are done with the use of a ‘cage’ that surrounds the preoperative 3D model and serves as a handle to deform the model. This cage is a simplified version of the preoperative 3D model. We have validated this method quantitatively on phantom and ex-vivo animal data, and qualitatively on patient data. We found this method is able to improve global registration accuracy compared to a traditional single-view 3D-2D registration. This method has been published as a journal article in (Espinel et al., 2020) and is presented in chapter 4 of this thesis.
- **Contribution #2:** We propose several multi-view 3D-2D deformable registration methods that combine visual information from multiple images for registration. These methods work for cases where the liver does not deform between images (rigidly linked images), and when it does (non-rigidly linked images). Thanks to the integration of information from multiple images, a bigger area of the liver is constrained, improving the global registration accuracy. The methods were validated quantitatively on synthetic and phantom data, and qualitatively on patient data. They show a noticeable improvement in registration accuracy compared to a normal single-view 3D-2D registration. Part of this work has been published as a conference article in (Espinel et al., 2021), while an extended version of it has been accepted for publication in the IJCARS journal as an special issue of MICCAI 2021. These multi-view methods are presented in chapter 5 of this thesis.
- **Contribution #3:** We propose to improve the surgical gesture by enhancing AR visualisation through the automatic estimation of the keyhole location from where an instrument is inserted. This location is used to project a tumour resection path, which provides an optimal



way to resect the tumour using that instrument. This has been validated quantitatively using a synthetic geometric model and 3D-printed phantom data, as well as qualitatively on patient data. It has been compared to traditional resections without AR and with the assistance of LUS. This work is pending to be published, and it is presented in chapter 6 of this thesis.

## Chapter 2

# Previous Work in AR for LLR

---

We present here the most relevant state-of-the-art methods that deal with AR for LLR. They are divided in 3D-3D and 3D-2D registration approaches. For the 3D-3D case, we classify the methods as SfM-based, SLAM-based, and non-monocular approaches. For the 3D-2D case we do not make a particular classification instead. For each of the methods, we present a short description of the registration pipeline and a summary of the experimental results. Finally, we discuss about the general status of the available registration methods, where we highlight some of the problems that have not been solved so far. We also make a comparison table of the reviewed methods in this chapter.

---

## 2.1 Introduction

AR for LLR has been an active field of research during the last decade, with different teams proposing several reconstruction and registration methods aimed at either monocular or non-monocular endoscopic systems. In this chapter we navigate through the state-of-the-art AR methods for LLR, summarising both their contributions and limitations. We discuss the general pipeline they follow, as well as the strategy used to validate the quality of the augmentations. They are classified according to the type of data they register and the registration method they use. Regarding the type of data, the works are divided in 3D-3D and 3D-2D registration methods. 3D-3D registration aligns preoperative 3D data to intraoperative 3D data, while 3D-2D registration aligns preoperative 3D to intraoperative 2D data. The 3D-3D approaches are divided according to the method used to reconstruct the intraoperative data in SfM-based, SLAM-based, and non-monocular-based approaches. The existing 3D-2D approaches only use one image to perform registration. Thus, we do not classify them in a particular way.

## 2.2 3D-3D Registration Methods

### 2.2.1 SfM-Based Approaches

We present the methods proposed in (Modrzejewski et al., 2019) and (Cheema et al., 2019) as examples of SfM-based registration approaches for AR in LLR. The method presented in (Modrzejewski et al., 2019) uses a reconstructed intraoperative shape from Structure from Motion (SfM) as a target to align and deform the preoperative liver model. It uses an iterative energy minimisation strategy to solve the constraints used for registration. The constraints are composed of the 3D-3D correspondences, the biomechanical liver model, the silhouette contours, and a self-collision correction term. The registration pipeline can be summarised as follows. First, a 3D surface model is generated from the preoperative CT images. Second, a tetrahedral model is generated from the surface model using TetGen (Weierstrass Institute for Applied Analysis and Stochastics, 2015). This tetrahedral model serves as the biomechanical liver model. Third, a dense pointcloud is reconstructed from the intraoperative laparoscopic images using SfM. Fourth, 3D contours are manually marked in the lobes of both the preoperative and the intraoperative models. Fifth, a manual rigid registration is done between the preoperative and the intraoperative models. Sixth, the preoperative model is deformably registered to the intraoperative model following a stiff-to-flexible energy minimisation approach (Amberg et al., 2007). The registration continues until convergence is detected or a maximum number of iterations is reached. This work also provides a new porcine dataset that can be used to evaluate deformable registration approaches. It contains the 3D models, the locations of 3D markers, and a 3D reconstruction from laparoscopic videos for several liver deformations, as well as for the original undeformed liver. An example of the information contained in the dataset is shown in figure 2.1. The proposed registration is validated using this dataset, where the 3D models generated from both the deformed CT scans and the laparoscopic videos are used as registration targets. Target Registration Error (TRE) was measured on the markers, for which the average TRE was of 27 mm for all deformations. Results also show that registrations using the laparoscopy-based models as target provides less accurate results than using the CT-based models, revealing that the method is sensitive to partial and noisy reconstructions. Because the anatomy of a porcine liver is substantially different from that of a human liver, it remains

difficult to extrapolate these results to the human case.

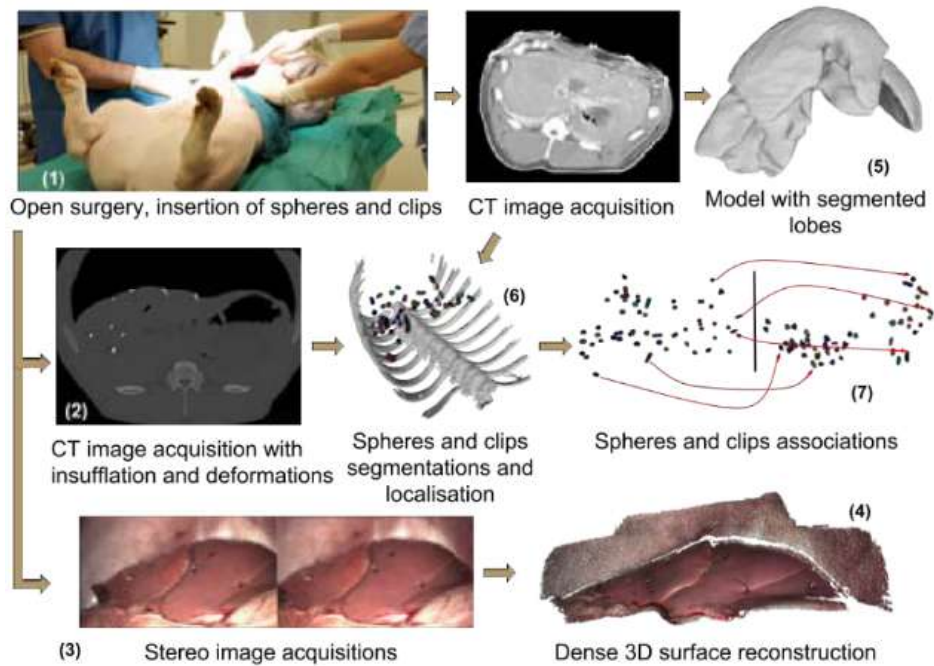


Figure 2.1: Porcine dataset for assessment of registration methods (Modrzejewski et al., 2019).

Another method that performs deformable reconstruction and registration is proposed in (Cheema et al., 2019). It uses SfM to reconstruct the intraoperative model and the camera poses from several partial views of the liver. Feature correspondences and shading cues are then used to perform registration between the preoperative and intraoperative models. However, in this case it is the intraoperative model that is deformed to fit the preoperative one. Generally speaking, the method operates as follows. First, Scale-invariant Feature Transform (SIFT) features are extracted and matched across all the images. Second, a dense 3D pointcloud is generated from the feature matches by using triangulation, Clustering View for Multi-View Stereo and Patch-based Multi-View Stereo. Third, the pointcloud points are connected to generate a triangular mesh. Fourth, feature correspondences are found between the intraoperative and preoperative models, using a correlation-based matching. Fifth, a rigid transformation is found to map the feature correspondences. Sixth, the intraoperative model is deformed to fit the preoperative model by projecting the pointcloud points to the corresponding preoperative points. Seventh, the deformed model is smoothed by applying shading cues generated from an albedo function of the liver surface. This method has been evaluated quantitatively on synthetic data, for which the quality of the intraoperative reconstruction was compared against state-of-the-art methods like classical SfM, Deformable SfM (DSfM), DSfM with Shading (DSfMS), and Shape-from-Shading (SfS). Compared to DSfMS, which was the best performing method from the literature with a TRE of 0.32 mm, the proposed method reduced this error to 0.18 mm. This method was also better in other metrics like rotational errors and reprojection errors. Qualitative experiments were made on a patient case, for which the method deformed the reconstruction to compensate for effects of pneumoperitoneum and surgical manipulations. This patient case is shown in figure 2.2, where the reconstructed intraoperative shape is compared with the preoperative model.

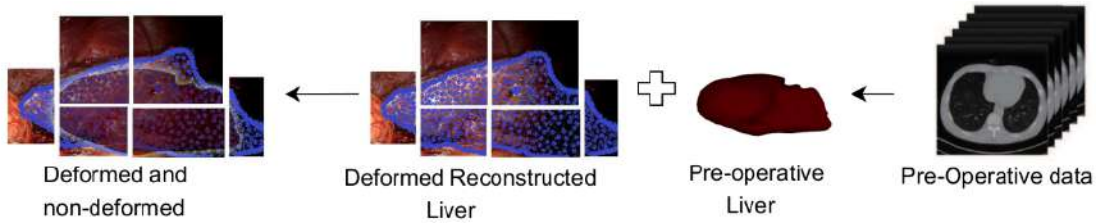


Figure 2.2: Reconstruction of intraoperative data from multiple images and comparison to the preoperative liver model using the approach from (Cheema et al., 2019).

### 2.2.2 SLAM-Based Approaches

In this section, we introduce the method proposed in (Chen et al., 2018) as a relevant example of SLAM-based registration for AR in LLR. It proposes a system that overlays the laparoscopic videos with custom 3D annotations. This method reconstructs an intraoperative shape from the laparoscopic video stream using Simultaneous Localisation and Mapping (SLAM), which is then used to annotate or attach any virtual information that the surgeon wants to see augmented in the video stream. This work is preliminary and only works on rigid MIS scenes. Its pipeline goes as follows. First, ORB-SLAM reconstructs a sparse pointcloud and the camera poses from the video stream. Such pointcloud is noisy and contains errors due to calibration issues and specularities inherent to MIS scenes. Second, the outliers are filtered out by removing points with very few neighbours. A voxel-grid filter is also applied to resample the pointcloud and make it more evenly distributed. Third, a smooth, continuous surface pointcloud is generated through a Moving Least Squares algorithm (Levin, 2004). Fourth, the pointcloud is triangulated using a Poisson surface reconstruction algorithm, generating the final mesh that will be used to perform AR. Fifth, the virtual annotations, which can include tissue measurements and markers, are attached to the intraoperative reconstruction. Sixth, the laparoscopic video is augmented by fusing the camera poses obtained from SLAM and the intraoperative reconstruction. This method is evaluated quantitatively on a synthetic 3D model representing a human digestive system, and qualitatively on real patient images. For the synthetic case, the reconstructed SLAM model is rigidly aligned to the synthetic 3D model using Iterative Closest Point (ICP), and the depth values of the SLAM model are compared to the ones of the synthetic model. Quantitative results show root-mean-square distances going from 2.54 mm to 3.66 mm for several levels of noise added to the camera images. For the real patient case, the method is tested on several laparoscopic video sequences, for which virtual objects and annotations are attached to the reconstructed intraoperative shapes, as shown in figure 2.3. However, this method was only tested on the same videos used to reconstruct the intraoperative shape, rather than on new ones. This method also fails in the presence of large tissue deformations and moving surgical tools.

### 2.2.3 Non-Monocular Approaches

We present in this section the non-monocular methods proposed in (Haouchine et al., 2013), (Luo et al., 2020), (Robu et al., 2018), (Thompson et al., 2015), (Bernhardt et al., 2016), (Soler et al., 2014), which deal with both rigid and deformable registration. For example, the method in (Haouchine et al., 2013) performs deformable registration and tracking for AR in stereoscopic LLR. The registration is achieved by reconstructing an intraoperative shape of the liver using stereoscopic tech-

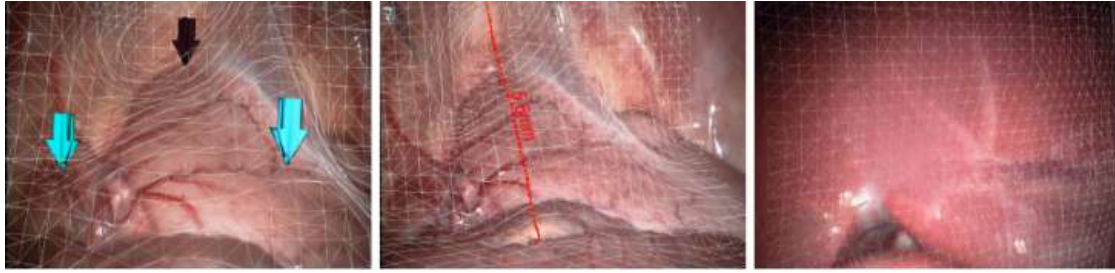


Figure 2.3: Examples of AR in MIS using an intraoperative reconstructed model with SLAM, augmented with virtual markers and annotations (Chen et al., 2018).

niques and using it as the registration target. Liver deformation is modelled by using a biomechanical model based on a finite-element method with a co-rotational element formulation (Felippa and Haugen, 2005). Vessel deformation is modelled by using serially linked beam elements (Petterlik et al., 2012), also employed to simulate catheters (Duriez et al., 2006). The general pipeline is divided into an initial registration and a real-time tracking steps: During the initial registration step, the liver's preoperative 3D model is first manually aligned to the first pair of laparoscopic images. These images are chosen so that a large part of the liver is visible. Then, the intraoperative sparse pointcloud is reconstructed from the first pair of images by matching and triangulating Speeded Up Robust Features (SURF) keypoint features in both images. The preoperative model is then deformed to fit the intraoperative pointcloud. This is achieved by projecting the visible preoperative model's points to the closest intraoperative pointcloud points and keeping them fixed during the deformation process. After deformation, all the intraoperative pointcloud points are projected towards the deformed liver surface using ray-casting, and the links between the projected pointcloud points and the intersected barycentric points in the liver surface are kept. During the tracking step, the SURF features are tracked over time using the Lucas-Kanade algorithm (Lucas and Kanade, 1981). The motion of these features over time is translated into stretching forces that will make the liver model deform. Due to the requirement of a stereo laparoscope to work, this method cannot be applied on the standard monocular laparoscopic systems available in most surgery rooms. This method has been validated quantitatively on synthetic and phantom data, and qualitative on real patient data. For the quantitative case, experiments are done using an homogeneous liver model without the vascular system, and an heterogeneous model including the vascular system modelled through the beam-elements. On synthetic data, experiments show that vascularity has a big influence on the tumour's location after registration, especially when the tumour is located in strongly deforming areas. On phantom data, experiments show that the accuracy of registration is stable in time for the heterogeneous model, while varying greatly for the homogeneous model. The final registration accuracy varies between 3.41 mm and 4.40 mm. For the qualitative case, experiments are done on a laparoscopic sequence of a real patient, where the liver is deformed by heart beating, respiratory motion and instrument manipulations. Results show a sufficiently good tracking of the liver deformation with a coherent registration of the 3D models, as shown in figure 2.4.

In the AR system presented in (Luo et al., 2020), both the preoperative 3D model and the intraoperative pointcloud are reconstructed using deep learning techniques. This method performs an initial rigid registration of the preoperative model to the intraoperative model using ICP, and uses an optical tracker to retrieve the camera pose and update the registration during the real-

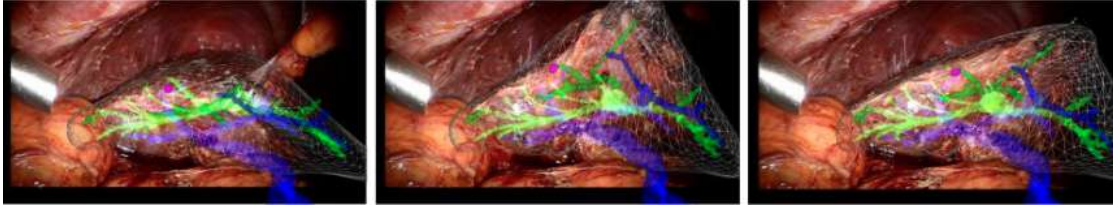


Figure 2.4: AR for LLR using the deformable registration method from (Haouchine et al., 2013).

time step. Generally speaking, this method starts by automatically segmenting the liver and its inner structures from CT data using a Convolutional Neural Network (CNN) called V-Net (Milletari et al., 2016). The preoperative 3D model is then reconstructed from the segmentation masks with the help of MITK (German Cancer Research Center (DKFZ), 2008). Then, a dense intraoperative pointcloud is reconstructed from the stereo image pair by using an unsupervised learning method based in CNN proposed in (Godard et al., 2017). This CNN predicts disparity maps from the input stereo images, which are then used to reconstruct the dense pointcloud. The preoperative model is then registered to the intraoperative pointcloud by using the globally-optimal ICP algorithm (Go-ICP) (Yang et al., 2016). Go-ICP is a pointcloud alignment method that is robust to initial solutions and is able to find a global optimal registration. This method has been validated on ex-vivo and in-vivo porcine livers. In both cases, artificial tumours were injected and fiducial markers were attached to the livers for evaluation purposes. The registration accuracy was measured by triangulating and measuring the 3D distances between the real and the augmented markers in the stereo image pairs. Tumour resections were also made to evaluate the functionality of the system. For the ex-vivo case, the mean TRE is of 6 mm, while for the in-vivo case the mean TRE is of 8.7 mm. A qualitative example of AR on an in-vivo porcine liver is shown in figure 2.5. However, due to the rigid nature of the approach, it is prone to suffer from high registration errors if the liver deforms significantly.

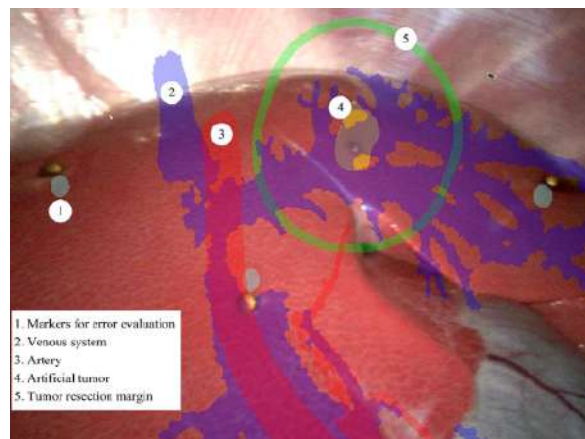


Figure 2.5: AR on porcine liver using the rigid registration approach from (Luo et al., 2020).

An initial rigid registration method is proposed in (Robu et al., 2018), where the preoperative 3D model is rigidly aligned to an intraoperative pointcloud obtained from stereo reconstruction. The registration process uses a shape matching technique that measures the geometric consistency between two sets of 3D features automatically extracted from the preoperative and intraoperative shapes, as well as between two 3D contour landmarks that are manually marked in both shapes. The process starts by extracting 3D TOLDI features (Yang et al., 2017) in the preoperative

and intraoperative models. The 3D contours are then marked in both models. Then, an initial set of features correspondences are found through a kd-trees matching strategy. A spectral analysis (Leordeanu and Hebert, 2005) and a filtering based on Random Sample Consensus (RANSAC) are done on this initial set to keep the most similar matches. Finally, the optimal rigid transformation is estimated using a least squares strategy that minimises the dissimilarities between the feature descriptors and maintains the geodesic distances between pairs of features in the two models. After this global alignment is made, a refinement can be done using ICP-based approaches. Validation was made through experiments using synthetic, phantom and patient data. For the synthetic case, the robustness of the method was evaluated against varying sizes of the intraoperative reconstructions, as well as varying levels of intraoperative deformations. Results show that the proposed method is more stable in both scenarios, compared to RANSAC-based methods. For the phantom case, the method is tested using partial views of the original deformed phantoms, as well as stereo-reconstructed intraoperative shapes as the registration targets. In the first case, the final TRE reduces to 7.77 mm, while in the second case the TRE is 12.10 mm. This demonstrates the sensitivity of the method to noise in the reconstructed preoperative shapes. For the patient case, the method is used on a real patient sequence, for which a comparison is made against a manual alignment. These results are illustrated in figure 2.6.

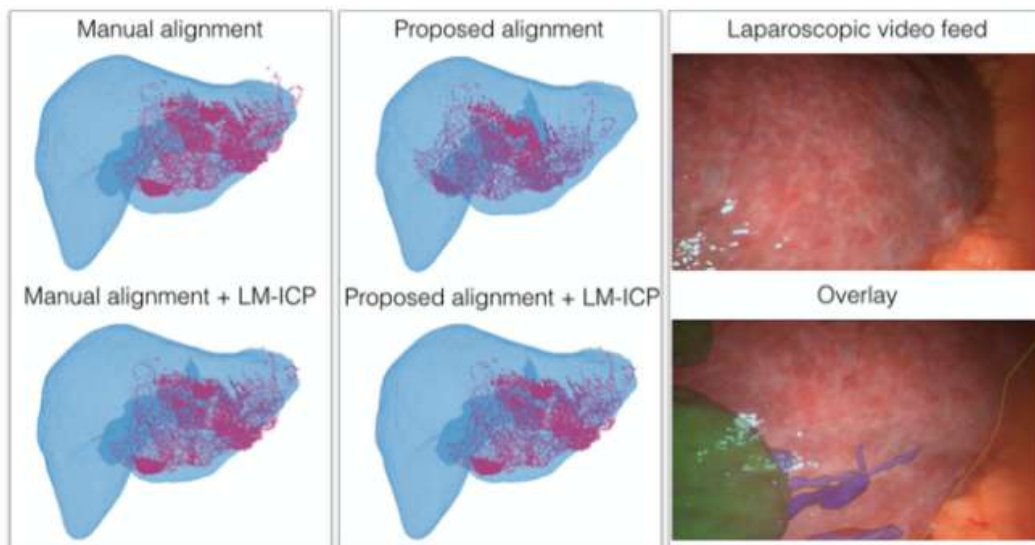


Figure 2.6: Qualitative evaluation of (Robu et al., 2018) on a patient case, compared to a manual initial alignment. The preoperative model is in blue, the intraoperative model in pink.

Another AR system that performs initial rigid registration is shown in (Thompson et al., 2015), where several intraoperative pointcloud patches are stereoscopically reconstructed and tracked using an optical tracking system. The reconstructed patches are then used as the target to rigidly register the preoperative 3D models. This method starts by reconstructing dense pointclouds from multiple pairs of stereo images in order to have a wider view of the intraoperative scene. The pointclouds are then tracked and merged together with the help of the optical tracker. Then, an initial manual registration aligns the preoperative 3D models with the visible liver in the laparoscopic image. Finally, a point-to-surface ICP algorithm finely aligns the preoperative 3D models with the intraoperative 3D pointclouds. This approach was validated using phantom data and in-vivo porcine data. For the phantom case, experiments were made using a single rigid silicon model and



TRE was measured on subsurface landmarks. The measured TRE for the proposed method is 2.9 mm. For the in-vivo porcine case, experiments were made on 3 different specimens. Registration errors were measured on artificial surface landmarks, for which the mean TRE is 9.4 mm. An example of an augmented laparoscopic image using this method is shown in figure 2.7. Similarly to the approach from (Robu et al., 2018), due to the rigid nature of the registration, strong deformations between the preoperative and intraoperative data will have a negative impact on the registration accuracy.

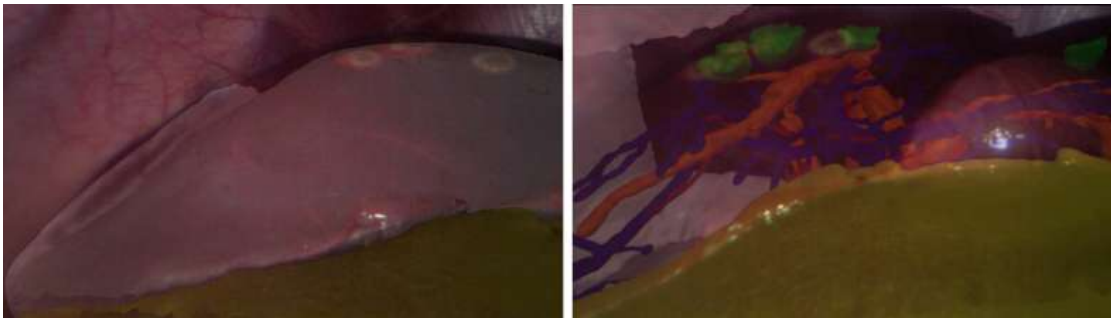


Figure 2.7: AR on an in-vivo porcine liver using the method from (Thompson et al., 2015). Left: Image augmented with the registered liver model. Right: Image augmented with vessels and tumours.

A rigid registration approach that aligns an intraoperative shape acquired from a rotational C-arm to a laparoscopic image is presented in (Bernhardt et al., 2016). It solves registration by gathering the pose of the laparoscope in the intraoperative volume, and complementing it with shading information coming from the laparoscopic image. Thus, the process is fully automatic and takes little time to complete. It starts by performing a 3D scan of the patient using the rotational C-arm, in a way that both the liver and the laparoscope are inside the scanner's acquisition field. Then, the pose of the laparoscope in the volume is used to place a virtual camera and create a virtual render of the intraoperative scene. The extracted pose is up to a rotation angle around the camera's optical axis and a 2D translation along the image frame. Such missing information is finally estimated through an intensity-based algorithm that reduces a shading dissimilarity between the virtual and the laparoscopic images. This method was evaluated on three in-vivo porcine livers, for which the registration error was measured on multiple visual cues as a 2D Target Visualisation Error (TVE). The measured average TVE over the three livers is 11.3 px, with a total computation time varying between 25 to 55s. The proposed shading model used for registration was also compared to other state-of-the-art models, showing that the dissimilarity metric was minimal at the optimal pose when using the new model, which was not the case for the others. An example of registration between the laparoscopic image and the intraoperative volume for one porcine case is shown in figure 2.8.

An AR system for LLR that assists surgeons with the generation of preoperative 3D models, the surgery planning, and the intraoperative AR stages is shown in (Soler et al., 2014). For the generation of preoperative 3D models, a tool was developed to segment multiple tissues and organs from CT images in a semi-automatic way. For the surgical planning stage, another interactive tool was developed to visualise the reconstructed preoperative 3D models and simulate tumour resections using virtual tools. For the AR-based surgery assistance tool, a manual registration approach rigidly aligns the liver's anatomical landmarks visible in both the preoperative and intraoperative data. The main benefit of this system is the removal of third party software in the AR pipeline,



Figure 2.8: Example of a laparoscopic image of porcine liver (left), an intraoperative volume of the scene generated using a C-arm (middle), and registration between the two modalities using (Bernhardt et al., 2016) (right).

which makes the process easier and reduces the total time required to perform AR. This system has been tested qualitatively on 40 resection procedures for organs including liver, adrenal gland, and pancreas. Although no statistics are given for the utility of the AR on those interventions, a general remark is made regarding the requirement of a deformable registration in order to improve the reliability of the augmentations. An example of this system applied to a patient case is shown in figure 2.9.

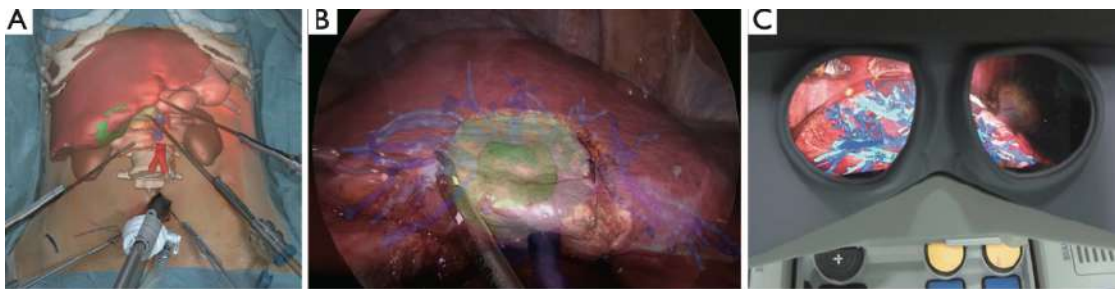


Figure 2.9: AR realised on a patient using the system from (Soler et al., 2014) with Da Vinci robot, providing an augmentation from outside of the abdominal cavity (A), an augmentation on a laparoscopic image (B) and the augmented image on the robot's stereoscopic vision system (C).

### 2.3 3D-2D Registration Methods

We introduce here the methods proposed in (Adagolodjo et al., 2017), (Koo et al., 2017), (Koo et al., 2022), and (Labrunie et al., 2022), being the most relevant 3D-2D registration approaches for LLR in the literature. The method described in (Adagolodjo et al., 2017) is a semi-automatic approach for registering a preoperative 3D model to an intraoperative laparoscopic image. It uses the visible occluding contours of the liver in the image to deform the preoperative 3D model, combined with a biomechanical model based on a co-rotational finite element method to simulate the elasticity parameters of the human liver. The method starts by marking the 2D contours in the image using a graphical user interface. Then, an initial rigid registration is done to roughly align the preoperative 3D model to the liver in the image. A Gauss-Seidel iterative scheme is then used to solve the contour and the biomechanical constraints in an alternate way. The contour constraints are solved as follows: First, the outer 3D model contour is extracted from the visible area of the preoperative 3D model by selecting the edges that connect the visible and the non-visible triangles. Second, the 3D model contour is projected in 2D using a camera projection matrix. Third, the

projected 2D model contour points are matched to the closest 2D image contour points. Fourth, the matched 2D contour points are back-projected in 3D as line-of-sights, which will serve as constraints to deform the preoperative 3D model. The biomechanical constraints are thus solved by projecting the 3D model contour points towards the line-of-sights, and then by displacing the rest of the model points to find an equilibrium between the forces applied by the image contours and the elasticity constraints. This process continues until a certain number of iterations has passed. This registration method has been validated quantitatively on a phantom gel model, for which the ground-truth data has been obtained by attaching several markers and locating them using optical tracking cameras. This phantom was deformed in four different ways and registrations were done on each of them by using a rigid manual registration and the proposed method. Results show a mean TRE of 7 mm for the rigid registration method, which is reduced to 1.2 mm for the proposed deformable method. A qualitative test was also made on human liver and kidney cases, for which the visual results were satisfactory to surgeons. An example of such registrations is shown in figure 2.10.

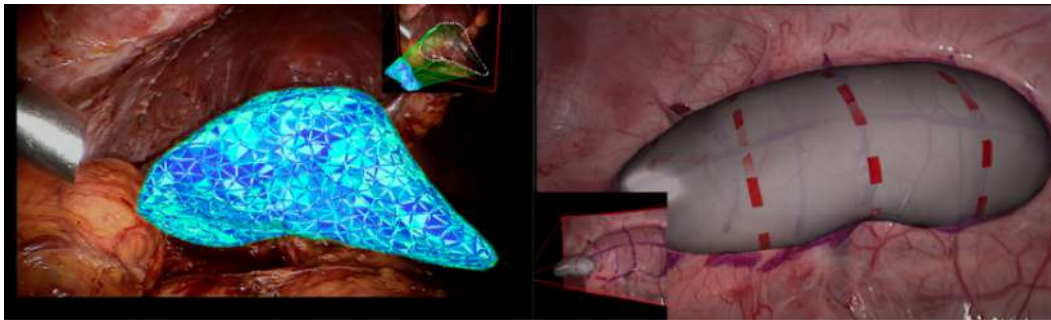


Figure 2.10: Example of 3D-2D deformable registration using the method from (Adagolodjo et al., 2017) on a human liver case (left) and a kidney case (right).

The method proposed in (Koo et al., 2017) is a semi-automatic approach where different types of liver anatomical landmarks are used to guide the registration. On top of the occluding contours, other landmarks are used such as the anterior ridge and the falciform ligament. Biomechanical and shading constraints are also used in the registration process. The biomechanical constraints are simulated through a Position-Based Dynamics (PBD) approach with a Neo-Hookean mechanical model. Similarly to (Adagolodjo et al., 2017), this method uses a Gauss-Seidel iterative algorithm to solve the constraints in an alternate way. This approach serves as a base for the registration methods developed in this thesis. The proposed registration algorithm is explained in detail in section 3.5. This method was evaluated quantitatively on synthetic liver data and qualitatively on human liver data. For the quantitative evaluation, preoperative 3D liver models from two patients were deformed in several ways, and the deformed shapes used as ground-truth data. Several synthetic images were generated from the deformed shapes, and registration was made using the proposed method with different initialisations. Results show that the method is robust against varying initial registrations, with TREs below 1cm. The method was also compared against a manual rigid registration, showing an improvement in the TRE of approximately 20 mm with respect to the rigid method. An example of AR using the rigid method and the proposed deformable method is shown in figure 2.11.

An automatic rigid registration method is developed in (Koo et al., 2022), where the liver landmarks are detected using a CNN. These landmarks correspond to the anterior ridge and the oc-

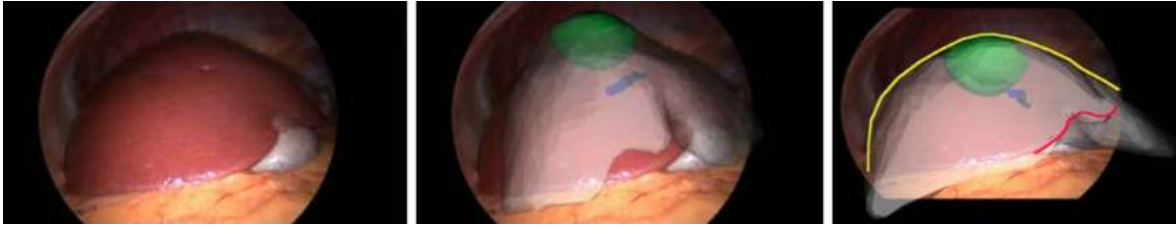


Figure 2.11: AR performed on a patient case, with the original laparoscopic image (left), the image augmented with the 3D models after rigid registration (middle), and after deformable registration using the method from (Koo et al., 2017) (right).

cluding contours. A CASENet network (Yu et al., 2017) is trained using a small dataset of 133 images from real laparoscopic interventions, combined with a synthetic dataset of 100,000 laparoscopic images. Augmentations are done for both the real and synthetic datasets. Results show that the prediction accuracy of the network is sufficiently good for the registration task, being able to generalise well to changes in the liver’s appearance. Regarding its registration pipeline, it can be globally described as follows: During the preoperative stage, the preoperative 3D models are segmented from the CT scans, and a top surface model is extracted from the preoperative liver model. The anterior ridge is also manually annotated on the top surface model. For registration initialisation purposes, a canonical transformation is computed between the liver model and a virtual camera, assuming the camera is inserted through a trocar placed around the patient’s belly button. During the intraoperative stage, a number of camera poses are generated in the canonical space by randomly rotating the camera around the  $x, y, z$  axes. These transformations are transferred to the preoperative 3D model using the canonical liver transformation, and the visible ridge and occluding 3D contours are extracted for each of the transformed 3D models. Then, the 3D contour points are projected in 2D and matched to the closest image contour points. An iterative algorithm that combines Perspective-n-Point (PnP) and RANSAC is then used to estimate the optimal camera pose. The optimal pose is chosen such that it reduces the Hausdorff distance between the corresponding model and image contours. This process is repeated from the generation of the random camera poses by taking as reference the newly estimated camera pose, until a certain number of iterations has passed. This approach was quantitatively validated on synthetic data and qualitatively on patient data. For the quantitative validation, multiple camera poses are randomly generated around a preoperative human liver model. For every camera pose, an image is rendered by extracting the visible ridge and silhouette 3D landmarks and projecting them in 2D. The proposed registration algorithm is run several times on every image, and the mean TRE is measured between the ground-truth and the registered 3D models. The ratio of liver visibility was also measured for every image. The results show that the method’s accuracy is generally stable regardless of the liver visibility ratio, with a mean TRE of 10 mm for visibility ratios going from 30% to 50%. For the qualitative validation, experiments were made on images from 5 different patients and errors were measured in 2D and 3D. The 2D errors were measured as the Hausdorff distance between the ground truth image contours and the projected 3D model contours. The 3D errors were measured as the Root Mean Square (RMS) error between the vertices of a manually registered 3D model and those of a registered model using the proposed method. The results show 2D reprojection errors going from 62px to 136px, and 3D registration errors going from 15.8 mm to 91 mm. An example of such registrations is shown in figure 2.12, where it can be seen that the overlaid registered models

do not exactly fit the liver in the images. This is mainly due to the rigid nature of the registration, which does not take into account the deformations exerted by the liver between the preoperative and intraoperative stages.

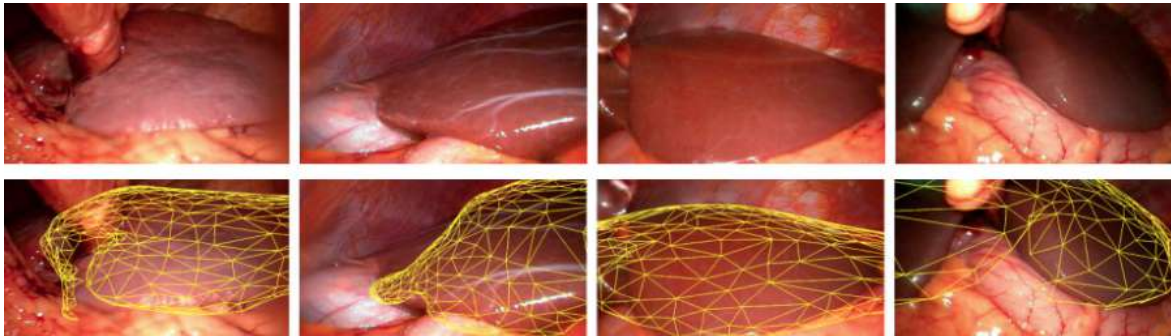


Figure 2.12: Transferred model after registration using the registration method from (Koo et al., 2022), showing the original laparoscopic images (top) and the images overlaid with the registered models (bottom).

Another automatic rigid registration method for initialisation purposes is developed in (Labrunie et al., 2022). This method detects the 2D contour landmarks automatically using U-Net, while the 3D landmarks are marked manually before surgery. It uses a coarse-to-fine registration scheme that varies the type of landmarks used in every step. The process starts with a coarse registration using RANSAC-based PnP that takes all the available landmarks into account. Because RANSAC depends heavily on the reprojection error threshold to define the set of inliers, the method estimates the pose using several thresholds and keeps the pose with the lowest reprojection error. Then, the process does a first refinement using only the visible ridge and ligament landmarks. It finishes by doing a second refinement using the visible ridge, ligament and silhouette landmarks. The method is validated on phantom and patient data. For the phantom validation case, the Mean Absolute Distance Error (MAE) is measured using both the groundtruth and the estimated poses from the proposed method, getting a final MAE of 7.7 mm for the first case, compared to a 15.9 mm for the second case. An evaluation using the deformable registration method from (Collins et al., 2021) and taking the optimal and the estimated poses as initial solutions is done. Using the optimal pose, the final MAE was 8.1 mm, while for the estimated pose the final MAE was 13.1 mm. For the patient validation case, the proposed method is compared to a manual rigid registration, and the TRE is measured on intraoperative tumours. The mean TRE using the manual registration method on 3 patients was 20.6 mm, while the mean TRE using the proposed method was 14.8 mm. An evaluation is also done using deformable registration, with the manual initialisation getting a mean TRE of 19.43 mm, and the automatic initialisation getting 13.4 mm. An example of registration using the manual and the automatic methods for the rigid and deformable cases is shown in figure 2.13.

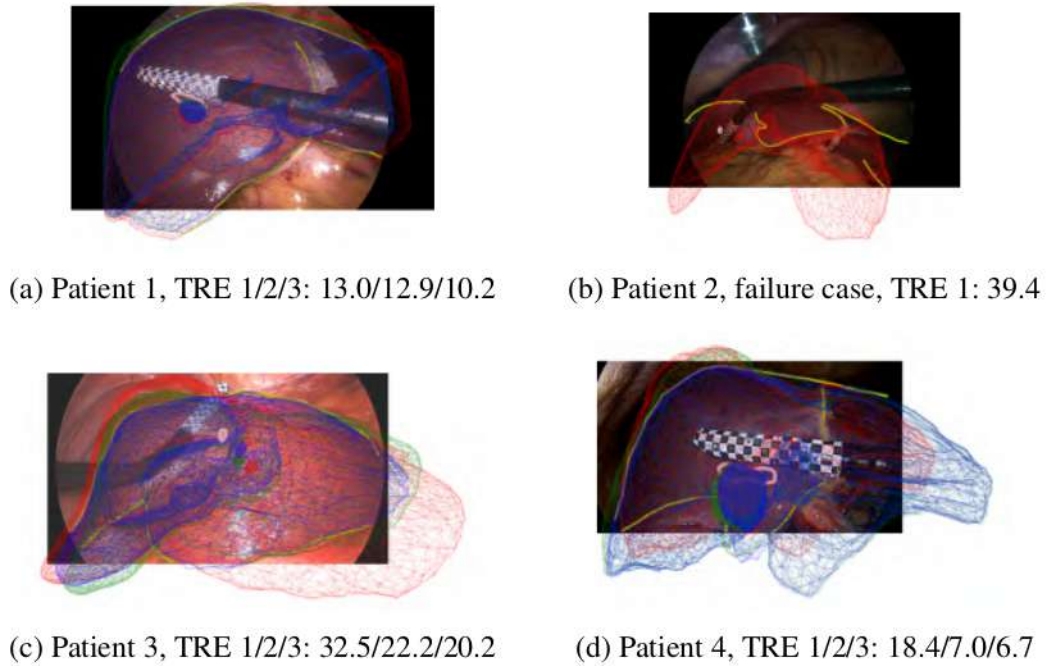


Figure 2.13: Comparison of 3D-2D registrations using a manual rigid initialisation (1-red), the automatic rigid initialisation method from (Labrunie et al., 2022) (2-green) and deformable registration using (Collins et al., 2021) after automatic initialisation (3-blue). TRE is given in mm.

## 2.4 State-of-the-Art Review

According to the previously reviewed methods, advances were made in registration for monocular (Modrzejewski et al., 2019; Cheema et al., 2019; Chen et al., 2018; Bernhardt et al., 2016; Adagolodjo et al., 2017; Koo et al., 2017, 2022; Labrunie et al., 2022) and non-monocular LLR (Haouchine et al., 2013; Luo et al., 2020; Robu et al., 2018; Thompson et al., 2015; Soler et al., 2014). They may use different types of data in the registration process such as anatomical liver landmarks (Modrzejewski et al., 2019; Soler et al., 2014; Adagolodjo et al., 2017; Koo et al., 2017, 2022; Labrunie et al., 2022), texture or shading information available in the liver surface (Modrzejewski et al., 2019; Cheema et al., 2019; Chen et al., 2018; Haouchine et al., 2013; Luo et al., 2020; Robu et al., 2018; Thompson et al., 2015; Bernhardt et al., 2016; Koo et al., 2017), or external devices for recovering intraoperative data (Bernhardt et al., 2016) or tracking the camera poses (Luo et al., 2020; Thompson et al., 2015). Most of the works in the literature have dealt with AR for non-monocular LLR, as well as monocular LLR with external tracking devices. This is because the data coming from stereo endoscopes and external tracking devices are reliable and fast to obtain, and can be directly used in registration. One of the presented methods has dealt with both deformable registration and tracking (Haouchine et al., 2013), while the rest only perform rigid registration. Due to the liver's high elasticity, deformable registration is important in LLR for increasing the registration accuracy. The problem with non-monocular methods is that most surgery rooms are not equipped with stereo endoscopes or tracking devices, but rather have traditional monocular endoscopes. Research on monocular AR is comparatively more challenging as it is difficult to obtain reliable 3D data from one or several unordered sets of images. However, monocular methods have dealt more with deformable registration (Modrzejewski et al., 2019; Cheema et al., 2019; Adagolodjo et al., 2017; Koo et al., 2017), which compensates for the lack of accurate 3D data from specialised devices. Sev-

eral issues arise from methods using a single image for deformable registration (Adagolodjo et al., 2017; Koo et al., 2017): First, a lack of anatomical landmarks due to a partial visibility of the liver. Second, a lack of depth constraint, making multiple registered models located at different depths to satisfy the same set of 2D landmark constraints. Regarding the multi-view methods in (Modrzejewski et al., 2019; Cheema et al., 2019), they perform a 3D reconstruction from the intraoperative images using SfM, which can be a challenging task in LLR due to changes in the liver's shape and appearance between images. A summary of the reviewed methods, stating the registration approach they use, the required devices and the type of validation done is provided in table 2.1. It is important to note that, due to the different strategies used to evaluate the accuracy of registration and the nature of the measured errors, the provided error values are not comparable between each other.

Method	Type of registration	Required devices	Registration constraints	Reconstruction and registration strategies	Type of validation	Error evaluation
(Modrzejewski et al., 2019)	3D-3D deformable	Monocular endoscope	Biomechanical model, intraoperative shape, anatomical landmarks	SfM reconstruction, Stiff-to-flexible iterative registration	Quantitative average TRE on in-vivo porcine data	27 mm
(Cheema et al., 2019)	3D-3D deformable	Monocular endoscope	Intraoperative shape, shading cues	SfM reconstruction, rigid registration + projection of points for deformation	Quantitative RMSE on synthetic data, qualitative on patient data	0.18 mm
(Chen et al., 2018)	3D-3D rigid	Monocular endoscope	Camera poses obtained from SLAM	SLAM reconstruction, rigid registration	Quantitative RMSD on synthetic data, qualitative on patient data	2.54 mm
(Haouchine et al., 2013)	3D-3D deformable	Stereo endoscope	Biomechanical model, intraoperative shape	Stereo reconstruction, manual initialization, force-based deformable registration	Quantitative tumour registration error on synthetic and phantom data, qualitatively on patient data	3.41 mm
(Luo et al., 2020)	3D-3D rigid	Stereo endoscope, optical tracker	Intraoperative shape, camera poses from optical tracker	CNN-based stereo reconstruction, ICP registration	Quantitative 3D reprojection error on ex-vivo and in-vivo porcine data, qualitative on in-vivo porcine data	8.7 mm
(Robu et al., 2018)	3D-3D rigid	Stereo endoscope	Intraoperative shape	Stereo reconstruction, least-squares and ICP registration	Quantitative mean TRE on synthetic and phantom data, qualitative on patient data	12.1 mm
(Thompson et al., 2015)	3D-3D rigid	Stereo endoscope, optical tracker	Intraoperative shape	Stereo reconstruction, manual and ICP registration	Quantitative RMS error on phantom and in-vivo porcine data, qualitative on in-vivo porcine data	9.4 mm
(Bernhardt et al., 2016)	3D-2D rigid	Monocular endoscope, rotational C-arm	Camera pose obtained from C-arm, shading cues	Similarity-based optimisation	Quantitative TVE on in-vivo porcine data	11.3 px
(Soler et al., 2014)	3D-2D rigid	Stereo endoscope	Anatomical landmarks	Manual registration	Qualitative on patient data	None



(Adagolodjo et al., 2017)	3D-2D deformable	Monocular endoscope	Biomechanical model, anatomical landmarks	Manual rigid initialisation, Iterative force-based deformable registration	Quantitative mean 3D error on phantom data, qualitative on patient data	1.2 mm
(Koo et al., 2017)	3D-2D deformable	Monocular endoscope	Biomechanical model, anatomical landmarks, shading cues	ICP rigid initialisation, iterative PBD optimization	Quantitative tumour registration error on synthetic data, qualitative on patient data	<10 mm
(Koo et al., 2022)	3D-2D rigid	Monocular endoscope	Anatomical landmarks	Iterative ICP, PnP and RANSAC rigid registration	Quantitative RMSE on synthetic data, qualitative on patient data	10 mm
(Labrunie et al., 2022)	3D-2D rigid	Monocular endoscope	Anatomical landmarks	Coarse-to-fine RANSAC-PnP rigid registration	Quantitative TRE on phantom and patient data	14.8 mm

Table 2.1: Comparison between the state-of-the-art methods in AR for LLR

## Chapter 3

# Technical Background on 3D-3D and 3D-2D Registration, 3D Reconstruction, and AR for LLR

---

We introduce here the concepts of 3D-3D and 3D-2D registration, which are necessary to perform for AR in LLR. We also describe some of the most relevant strategies used, like PCA, ICP, and RANSAC for the 3D-3D case, and feature-based and intensity-based registration for the 3D-2D case. For each strategy, we present some state-of-the-art examples in order to better illustrate its possibilities and limitations. Similarly, we introduce the concept of 3D reconstruction, being an important step in 3D-3D registration. We describe the pipelines of SfM and SLAM, as they are the two most popular methods used in 3D reconstruction. Finally, we explain the general pipeline of AR in LLR, going from the preoperative to the intraoperative stages.

---

### 3.1 Introduction

Augmented Reality requires the usage of a number of techniques to register two different sets of 2D or 3D data. In this chapter we discuss the principle of these registration techniques, showing examples in the literature for each of them, as well as their advantages and limitations. These examples are not limited to AR for LLR, in an effort to have a wider view of their capabilities. The presented techniques are classified in 3D-3D and 3D-2D approaches, being the most popular strategies in AR. We also introduce two of the most popular 3D reconstruction techniques, SfM and SLAM, which are widely used in AR. Then, we explain in detail the different stages required to perform AR, along with some examples of existing AR systems for LLR.

### 3.2 3D-3D Registration

The 3D-3D registration problem deals with the alignment of two objects in 3D space, by finding the best transformation between them. This is important for many applications such as object recognition, tracking, sensor data fusion, medical image alignment, robotics, and augmented reality (Li and Hartley, 2007; Bellekens et al., 2014). Most of the time, the objects to be registered are in the form of 3D pointclouds which may come from a variety of sources like an RGBD camera, a LiDAR sensor or a CT scan-based reconstruction. This registration can be rigid or deformable, depending if one of the models has suffered a noticeable change in its shape (Castellani and Bartoli, 2012). Registration is done by applying a transformation function  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  that aligns a moving pointcloud  $M \in \mathbb{R}^{3 \times \Phi}$  with a static reference pointcloud  $S \in \mathbb{R}^{3 \times \Psi}$ , where  $\Phi$  and  $\Psi$  are the number of points in  $M$  and  $S$  respectively, as shown in figure 3.1.

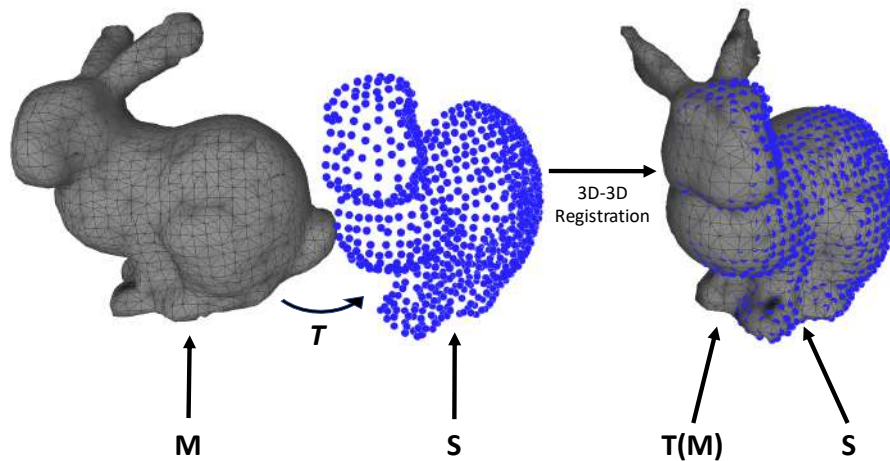


Figure 3.1: 3D-3D registration between a moving model  $M$  and a fixed pointcloud  $S$ . Left: A transformation  $T$  is computed between  $M$  and  $S$ . Right: The model  $M$  is aligned to  $S$  using  $T$ .

The registration problem can be seen as the estimation of an optimal transformation  $T^*$  so that it minimises the distance between the transformed moving pointcloud  $T(M)$  and the static reference pointcloud  $S$  (Castellani and Bartoli, 2012; Jian and Vemuri, 2011):

$$T^* = \underset{T}{\operatorname{argmin}} E(T(M), S) \quad (3.1)$$

where  $E$  is an error function that measures the registration error.  $E$  can be a measure of point-to-point or point-to-plane distances between  $T(M)$  and  $S$  (Besl and McKay, 1992; Chen and Medioni, 1992).

The transformation function  $T$  may contain rigid transformation parameters like a rotation matrix  $R \in SO(3)$ , and a translation vector  $t \in \mathbb{R}^3$ , as well as deformable parameters like Thin-Plate Spline models (Chui and Rangarajan, 2003) or Linear Bend Skinning models (Chang and Zwicker, 2009). This estimation is usually done between two sets of correspondences previously extracted from the pointclouds  $M$  and  $S$  using feature matching. This approach is called *correspondence-based registration*. If no correspondences are known a priori, they have to be estimated in parallel with the transformation  $T$ , which is known as *simultaneous-pose-and-correspondence registration*. In this section we explore some of the most popular correspondence-based and simultaneous-pose-and-correspondence registration methods and their application in several fields.

### 3.2.1 Random Sample Consensus

#### Principle

RANSAC is a *correspondence-based* algorithm capable of estimating  $T$  from a set of correspondences containing errors. It was first proposed by (Fischler and Bolles, 1981) as a method that robustly estimates the parameters of a mathematical model from a set of observations containing errors. In 3D-3D registration tasks, it starts by randomly drawing three pairs of correspondences and computing a hypothesis of transformation using Absolute Orientation (ABSOR). For each hypothesis, the method checks how many correspondences agree with it by evaluating a transformation constraint. The correspondences that agree with the hypothesis are the set of inliers. If the number of inliers is greater than a threshold, the transformation is re-estimated using these inliers and the process finishes. Otherwise, the process continues until a number of trials has been reached, moment at which the transformation is re-estimated using the largest set of inliers (Chin and Suter, 2017; Hartley and Zisserman, 2003). The general pipeline is illustrated in figure 3.2.

The set of inliers that explains a particular hypothesis is called the *consensus set*. For every iteration, the new hypothesis is kept if its consensus set is larger than the one from the previous iteration; otherwise, it is discarded. The main advantage of RANSAC is the ability to robustly estimate  $T$  with a good degree of accuracy, even with a significant number of errors. However, it might not be able to estimate an optimal transformation if the ratio of errors is larger than 50% of all the correspondences.

#### Examples

Most of the existing works try to improve RANSAC by increasing its accuracy and efficiency in different ways. For example, in (Torr and Zisserman, 2000) a modified RANSAC algorithm is proposed where, instead of choosing the best hypothesis by the total number of inliers, it selects it by estimating the maximum likelihood from the sets of correspondences, which makes it more robust against a strong presence of outliers. In (Han et al., 2015) region covariance descriptors are used to extract and match the strongest features in both pointclouds. This reduces the number of iterations required to find a correct set of matches, and increases accuracy thanks to a fewer number of outliers. In (Koguciuk, 2017), a parallel RANSAC strategy is proposed, reducing the total processing time and the global accuracy compared to serial RANSAC. This is because all possible hypotheses

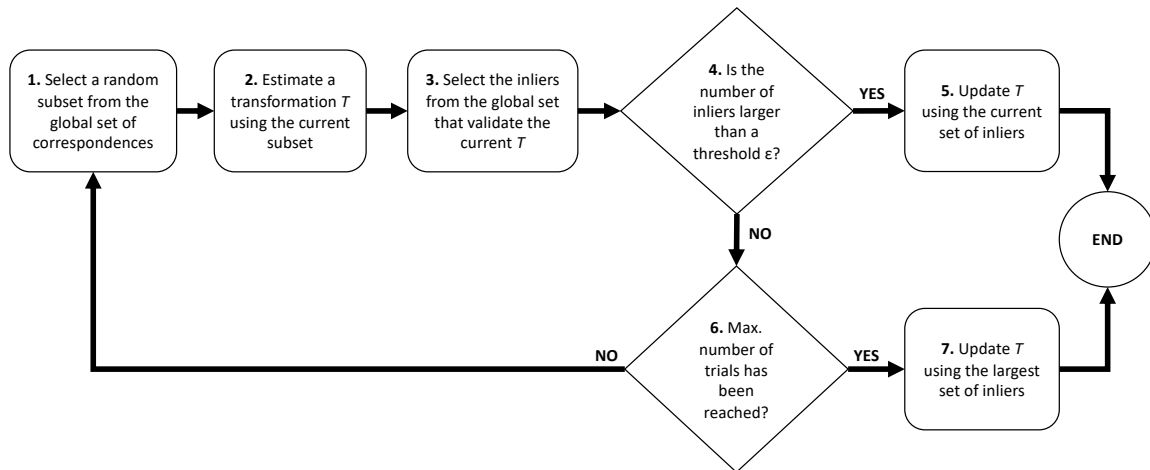


Figure 3.2: Pipeline for 3D-3D registration using RANSAC.

are computed in parallel and the best one is selected from this complete set, compared to a serial RANSAC in which a partial set of hypotheses is computed. In (Li et al., 2020a) an improved RANSAC algorithm is proposed for robust coarse registration of pointclouds. The method uses graphs to represent the feature points, and finds reliable matches through graph matching. This allow the usage of larger subsets of correspondences, which reduces the required number of iterations to find an accurate  $T$ . It also makes the method more robust to a high presence of errors, working for rates larger than 90%.

RANSAC-based methods are mainly used to find an initial alignment between the pointclouds, being complemented with other methods to refine such alignment. This approach is called coarse-to-fine registration (Xian et al., 2016), in which RANSAC is combined with other methods like ICP (Besl and McKay, 1992; Rusinkiewicz and Levoy, 2001; Han et al., 2016), Normal Distribution Transform (Biber and Strasser, 2003; Takeuchi and Tsubouchi, 2006; Magnusson, 2009), or Genetic Algorithms (Chow et al., 2004; Silva et al., 2003).

### 3.2.2 Iterative Closest Point

#### Principle

ICP is a *simultaneous-pose-and-correspondence* algorithm that works in an iterative way to find an optimal  $T^*$ , starting from an initial transformation  $T_i$ . The concept of ICP is intuitive and easy to implement, being used to solve numerous real-world applications. The algorithm starts by matching the points in  $M$  to the closest points in  $S$  given the initial  $T_i$ . Then, given this set of correspondences, an updated transformation  $T$  is found through a least squares optimisation. The moving pointcloud  $M$  is then transformed using  $T$  and the process continues. The algorithm will stop once a minimal distance criterion is met (Castellani and Bartoli, 2012). This process is illustrated in figure 3.3.

Because the transformations are computed from the closest point neighbours at every iteration, it will find a local minima of the cost function most of the time. As such, it strongly depends on a good initial transformation  $T_i$  to find the optimal transformation. This is why it is mainly used in fine registration stages. Several modifications to the original ICP algorithm have been proposed over the years, in an effort to improve its efficiency and accuracy, as well as to deal with the

local minima issues (Li et al., 2020b). Regarding the efficiency, the main improvements are made on the point-matching stage either by reducing the number of points through downsampling of the pointclouds (Gelfand et al., 2003), or by increasing the search speed of the nearest neighbours (Benjemaa and Schmitt, 1999; Greenspan and Godin, 2001). Regarding the accuracy, some methods try to improve the precision of the point matching, which translates into more accurate transformations (Censi, 2008; Grant et al., 2012). Other methods try to solve the transformation by using Total Least Squares (Estépar et al., 2004) or Weighted Total Least Squares (Tao et al., 2018). Regarding the local minima problem, several approaches combine an initial global registration (or coarse registration) with a local registration using ICP (or fine registration) (Gelfand et al., 2005; Rusu et al., 2008; Li et al., 2016).

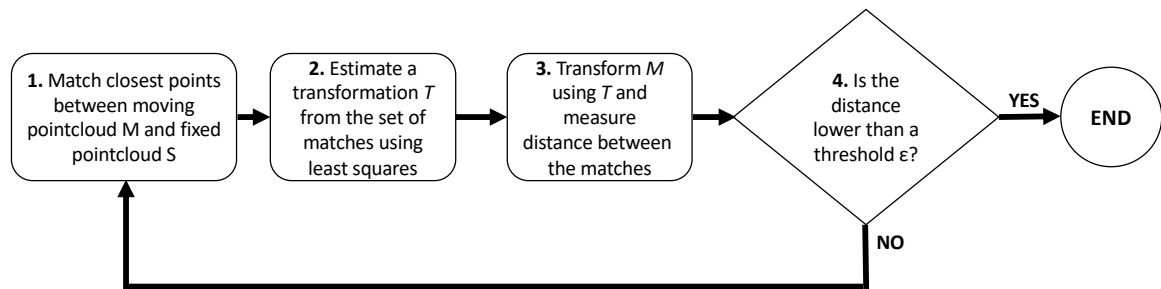


Figure 3.3: Pipeline for 3D-3D registration using ICP.

### Examples

As an example of a method that improves both the efficiency of the point matching and the accuracy of the transformation computation, the method in (Billings et al., 2015) computes an optimal shape alignment under anisotropic noise conditions. It does so by incorporating a probabilistic framework within both the matching and registration phases of the algorithm. To compute the matches, the pointclouds are modelled as gaussian models and the point pairs are selected in a way that they maximise a likelihood function. To compute registration, a generalised total-least-squares optimisation strategy is used, bringing improvements in terms of speed and accuracy. In general, this method is able to cope with highly misaligned pointclouds, as well as with a high number of errors. An automatic method for aligning two sets of unordered scattered pointclouds is presented in (Li et al., 2017), in which a first coarse registration is made using Principal Component Analysis with ICP, followed by a fine registration using ICP. This method is also robust to noisy data, and does not require any manual operation. However, it requires the pointclouds to have an overlapping of at least 50%. In an effort to improve the speed and efficiency of ICP, the method in (Liu et al., 2016) performs an initial rough alignment using centre superposition, followed by a fine alignment using ICP with a kd-tree correspondence search approach. This method proves to be particularly good for registering large pointclouds with varying densities. In order to improve the efficiency of ICP on large LiDAR pointclouds, the method in (Gressin et al., 2013) makes use of geometrical features to perform the matching. These features describe the shape of the neighborhood for each 3D point (linear, planar or volumetric), as well as its optimal size. Tests on indoor and outdoor datasets using this method have shown improvements in both speed and accuracy. An example of ICP used to find the transformation parameters between an US probe and a X-ray mammography device is presented in (Vitrani et al., 2015). The transformation is found by

registering a set of fiducial markers between the US and the X-ray coordinate frames. The transformation is then used to assist the surgeon in positioning the robotic-controlled US probe around a lesion that was previously identified in the mammography images. In this way, the US probe can be quickly positioned in the lesion with respect to a conventional non-assisted approach.

### 3.2.3 Principal Component Analysis

#### Principle

Principal Component Analysis (PCA) is a statistical method that reduces the dimensionality of a dataset so that it fits a smaller one. The key idea is to reduce the number of variables of the dataset in a way that it is easier to explore and visualise, while preserving as much information as possible. This method is often used in classification and compression techniques, as it projects the data onto a new orthonormal basis in the direction of the largest variance (Bellekens et al., 2014). In 3D-3D registration, PCA can be used to estimate the principal axes of every pointcloud. The two pointclouds can then be registered by aligning their principal axes. However, PCA may fail if the pointclouds are not symmetric or have a lot of missing parts. Concretely, the registration pipeline starts by centering the pointclouds such that their origins coincide. This is done by subtracting the centroid coordinates of the two pointclouds. Then, the covariance matrices are computed for each of the pointclouds. The diagonal values of the covariance matrix correspond to the variances while the off-diagonal values are the covariances. The eigenvectors and eigenvalues are then computed for the two covariance matrices. The largest eigenvector represents the direction of the pointcloud's largest variance and, therefore, its rotation. Then, the transformation matrix  $T$  that aligns the moving pointcloud  $M$  with the reference pointcloud  $S$  is computed. Finally, the centroid of  $S$  is added to the transformed  $M$  so that its centre corresponds to the centre of  $S$ . This pipeline is illustrated in figure 3.4. It is worth mentioning that the transformation computed with PCA is an approximate one, making this method suitable for performing initial rough registrations.

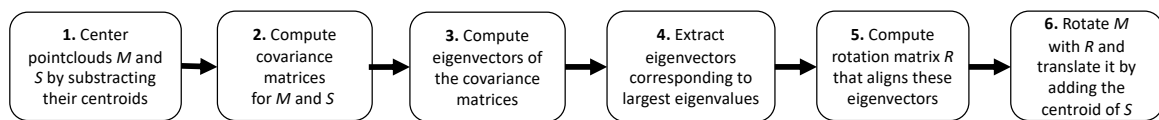


Figure 3.4: Pipeline for 3D-3D registration using PCA.

#### Examples

In an attempt to register large pointclouds coming from laser scanning devices, the method in (Xue et al., 2019) proposes a hybrid approach using PCA to perform an initial rough registration, followed by ICP to refine the registration. During the first PCA stage, two sets of feature points are extracted and matched using Feature Point Histogram. During the ICP stage, kd-tree search is used to find the point correspondences between  $M$  and  $S$ . This approach manages to increase the speed and accuracy of registration compared to a classical ICP algorithm. Another example of a method using PCA+ICP for pointcloud registration can be found in (Li et al., 2017), which tries to align two sets of unordered points without requiring any manual operation. Here, the pointclouds are first simplified to reduce the computation time and then roughly aligned with PCA. From these

initial poses, the pointclouds are finely aligned with ICP, by rotating the moving pointcloud  $M$  by a factor of  $\pi/4$  each time and computing the RMS residuals. Thus, the best alignment will be the one with the lowest RMS. This method is able to improve efficiency and accuracy on pointclouds with an overlap of at least 50%, as well as with raw data without any kind of preprocessing. A robust method for registering complex pointclouds in a fast and efficient way is presented in (Xin et al., 2018). There, a coarse-to-fine registration method using PCA, neural networks, and ICP is proposed. The method first performs a rough alignment of  $M$  and  $S$  with PCA; second, it simplifies the pointclouds using a back-propagation neural network that removes all the redundant points; third, the contour points are retrieved and inserted into the simplified pointclouds; and fourth, a fine registration is done using ICP based on M-estimation. This method has proven to be more efficient than classical PCA+ICP approaches based on least squares.

### 3.3 3D-2D Registration

Compared to 3D-3D registration in which the alignment is done between 3D structures, the 3D-2D registration problem deals with the alignment of 3D data (pointcloud, mesh and related structures) to a 2D image target, as shown in figure 3.5. It is widely used for medical imaging and robotic applications (Markelj et al., 2012; Otake et al., 2015; Crombez et al., 2018). In medical imaging, the 3D data is usually pre-operative and comes from CT and MR images, while the 2D data is usually intraoperative and comes from optical, US, X-ray or CT images. In robotics, the 3D data is usually obtained as a first step using reconstruction methods like SLAM or SfM, while the 2D data is obtained later and comes usually from optical data. In this case, the goal is to compute a transformation  $T$  from a set of 3D-2D feature matches between the two types of data. Similarly to the 3D-3D case, the registration can be rigid or deformable, depending if the 2D landmarks have a noticeably different shape than the 3D ones. The registration methods can also be classified depending on the nature of the data and, thus, the type of information to be compared.

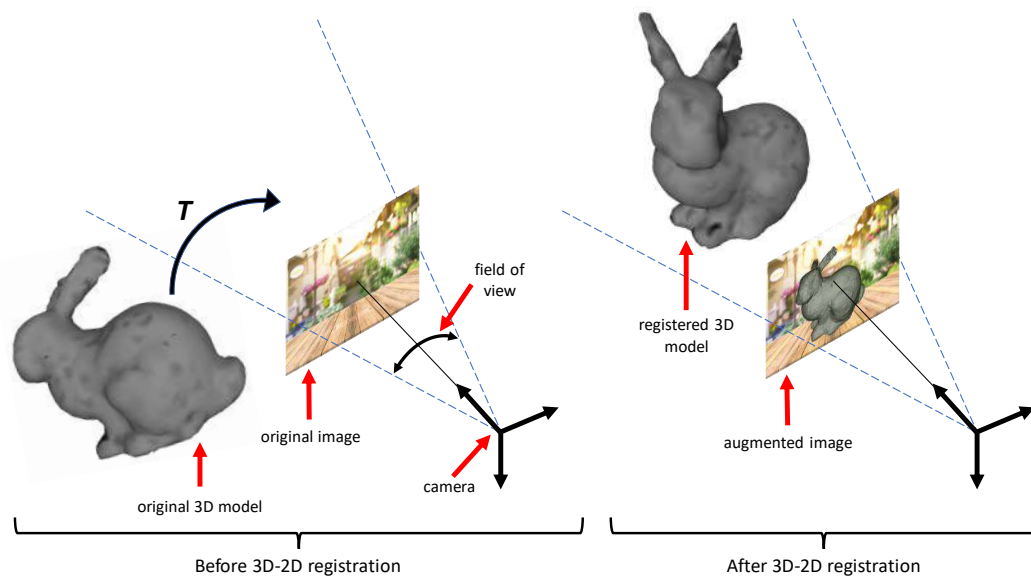


Figure 3.5: Registration of a 3D model into a 2D image.



### 3.3.1 Feature-Based Methods

#### Principle

Feature-based registration is mainly used when distinctive and reliable structures are available in both the 3D and 2D data. Such features can be points, curves or surface regions. Therefore, feature-based methods can be classified as point-to-point, curve-to-curve and surface-to-curve (Markelj et al., 2012). Point-to-point methods require the user to identify the corresponding points in both data, which is a subjective operation that depends heavily on the user's skills and his perception of the scene. In the end, the accuracy of these methods will depend on the precision of the selected correspondences. Thus, they are usually used to perform an initial rough registration that will be later refined through more precise methods. Curve-to-curve and surface-to-curve registration both reduce the need of finding precise point correspondences, as they usually correspond to occluding contours or landmarks with different lengths and/or different numbers of 3D-2D point correspondences. In this case, the point correspondences can be determined iteratively from the contour segments using algorithms like ICP, in a two-step fashion. In the first step, the point correspondences are found, while in the second step the pose of the 3D data is updated with the transformation computed from the current set of point correspondences. The transformation matrix  $T$  is obtained in a way that the distances between the corresponding 3D and 2D points are minimised. These distances can be measured in 3D and 2D, either by back-projecting the 2D points in 3D or by projecting the 3D points in 2D respectively.

For the 3D back-projection method, the estimation of the optimal  $T^*$  is an optimisation problem that minimises the distance between the transformed 3D points  $T(P)$  and the back-projected image points  $\pi^\dagger(p)$ , where  $\pi^\dagger()$  is the back-projection function:

$$T^* = \underset{T}{\operatorname{argmin}} E(T(P), \pi^\dagger(p)) \quad (3.2)$$

For the 2D projection method, the estimation of the optimal  $T^*$  is an optimisation problem that minimises the distance between the 2D projections of the transformed 3D points  $\pi(T(P))$  and the image points  $p$ , where  $\pi()$  is the projection function:

$$T^* = \underset{T}{\operatorname{argmin}} E(\pi(T(P)), p) \quad (3.3)$$

#### Examples

An example of 3D-2D registration using direct matching is presented in (Sattler et al., 2011), where the authors demonstrate that a direct 3D-2D matching yields better registration results compared to indirect methods. The matching process is based on the association of the 3D points with visual words obtained from clustering the feature descriptors. The visual words are then used to match the feature points of the query image with the 3D pointcloud, which proves to be efficient in large datasets. A 3D-2D non-rigid registration approach for medical imaging is presented in (Zhong et al., 2016), where a 3D non-uniform tetrahedral mesh reconstructed from CT data is registered to Cone-Beam Computed Tomography (CBCT) scans. The registration process uses displacement vector fields and 2D boundaries extracted from the CBCT images to drive the deformation. In this case, using a tetrahedral mesh shows better registration results compared to voxel-based models, uniform-grid meshes and uniform tetrahedral meshes. A registration method that combines

multimodal data is presented in (Crombez et al., 2018), where sparse and dense features are both used to accurately register 2D images to a 3D model. It takes advantage of the high domain of convergence offered by sparse features, and the high alignment precision given by dense features. A 3D-2D registration method using the back-projection approach is shown in (Chen et al., 2006), where a 3D mesh reconstructed from CT data is rigidly aligned to two orthogonal X-ray images. It combines the ICP and Z-buffer algorithms to find the correspondences and perform registration. In this case, the 3D and 2D contours of the bones are first extracted for the 3D mesh and the X-ray images respectively. Then, a transformation matrix is computed iteratively such that it minimises the distance between the 3D points and the back-projected 2D point correspondences.

### 3.3.2 Intensity-Based Methods

#### Principle

Intensity-based registration, compared to the feature-based case, is used when no distinctive landmarks are available in the data. It makes use of all the voxel and pixel information present in the 3D and 2D images, respectively. It is widely used in the medical imaging field. The 3D-2D correspondences are obtained using pixel-wise similarity measures, like mutual information (Dey and Napel, 2006; Zheng, 2007), normalised cross correlation (Hipwell et al., 2003; Munbodh et al., 2006), sum of square differences (Jans et al., 2006; Kamen et al., 2006), among others. The choice of the similarity measure will depend on the modality and nature of the data. In medical imaging, the most common registration method consists of generating simulated X-ray projection images called Digitally-reconstructed Radiographs (DRR). They are generated by ray-casting a CT image and applying a filter that creates the X-ray-like image. The registration is then done by iteratively optimising the similarity measure between the X-ray image and the transformed DRR image. Unlike the feature-based case, intensity-based methods do not suffer from segmentation errors and, because they use all the information in the images, they are generally more reliable and accurate. However, due to numerous local maxima of the similarity measures, they require an initialisation close to the optimal one. Regarding the registration of preoperative CT data to intraoperative X-ray images, the estimation of this optimal pose  $T^*$  can be thus seen as an optimisation problem that maximises the similarity measure  $s$  between the DRR projection of the transformed CT data  $\pi(T(M_{CT}))$  and the X-ray image  $I_{Xray}$ :

$$T^* = \underset{T}{\operatorname{argmax}} s(\pi(T(M_{CT})), I_{Xray}) \quad (3.4)$$

Gradient-based methods are a special type of intensity-based registration that use gradient information generated from CT/MR and X-ray data. They combine characteristics from both feature- and intensity-based methods. Like the feature-based case, registration can either be done in 2D by comparing the CT/MR 2D gradient projections with the X-ray gradient images, or in 3D by comparing the CT/MR 3D gradient vectors with the backprojected X-ray gradients. Like the intensity-based case, gradient-based registration uses similarity measures such as gradient correlation as the optimisation criterion (Penney et al., 1998; Livyatan et al., 2003). The key idea is to find an optimal transformation  $T^*$  that matches the magnitudes and directions of the gradients. These gradients are extracted at every iteration depending if the gradient comparison is done in 2D or 3D. Although being an accurate registration approach, it also suffers from local minima is-

sues, which requires the computation of a coarse initial alignment (Livyatan et al., 2003; Tomažević et al., 2001).

### Examples

A model-to-image registration approach is shown in (Toth et al., 2018), where a 3D model reconstructed from CT or MR data is registered to an X-ray single image for image-guided cardiac interventions. The registration algorithm uses a pre-trained neural network that aligns the preoperative 3D model to the intraoperative X-ray by comparing the 2D projections of the preoperative shape with the organ in the X-ray image. The system is trained with synthetic X-ray images using a reinforcement learning approach, and it is robust to segmentation errors and image modalities. A registration method that modifies the mutual information similarity measure to incorporate spatial information is presented in (Zheng, 2007), in an effort to improve robustness and accuracy. Instead of maximising the mutual information, it computes a variational approximation that includes the spatial information. This similarity measure has a least squares form that is minimised using a Levenberg-Marquardt optimiser. A method for rigidly registering cardiac CT data to Dual-energy Digital Radiography (DEDR) was developed in (Chen et al., 2007), where the authors evaluate the convenience of using DEDR to detect coronary artery calcification instead of the more expensive CT data. They used normalised mutual information as a similarity measure to compare the DDR projections and the DEDR images. The method is able to accurately register whole cardio-thoracic structures if provided with a good initial alignment. A hybrid approach for spine image registration is proposed in (Russakoff et al., 2005), where the intensity-based similarity measure is combined with a feature-based measure using a single fiducial marker. Thus, it combines the advantages of intensity-based registration like avoiding segmentation errors, and of feature-based registration like speed and robustness. A deformable registration method for microscopic imaging data is presented in (Striewski and Wirth, 2022). It is applied in the study of leukocyte extravasation in inflamed biological tissues. This method aligns 3D confocal microscopic volumes into 2D Intravital Microscopic Images (IVM). It deforms the 3D microscopic volume in a way that its 2D projection matches the 2D IVM image, which is achieved by maximising a custom-designed similarity measure.

For the gradient-based case, one of the first 3D-2D registration methods is proposed in (Livyatan et al., 2003), where a 3D CT model is registered to fluoroscopic X-ray images. It first performs a global initial registration using manually selected landmarks, followed by a coarse registration using a contour-based approach, to finalise by the gradient-based registration refinement. For the gradient-based refinement, it uses a 2D projection approach to obtain the optimal transformation. Another gradient-based method is presented in (Tomažević et al., 2001), where the gradient vectors of a CT/MR model are aligned with the back-projected 2D gradients of an X-ray image. To achieve this, a similarity measure that compares the gradients' magnitudes and directions is proposed. This rigid registration method can be applied to bony structures such as vertebrae, pelvis and head. A method that performs registration using gradient-based DDR images from CT data is shown in (Wein et al., 2005). Only gradients above a certain threshold are used, which lets the algorithm run up to 10 times faster than a conventional intensity-based approach. It also proposes a method for directly computing the gradient similarity measure without having to generate DDR images, which reduces the computation time even more.

## 3.4 3D Reconstruction

Reconstructing 3D data from a set of 2D images is a common task in 3D-3D registration and AR systems. Such data is usually in the form of 3D pointclouds, which correspond to the detected keypoint features in the images, and may serve as the target data to perform registration. While it is possible to obtain such 3D data from laser scanners, RGBD cameras or ultrasound sensors, they are not widely available or can be expensive to buy. Therefore, using conventional RGB cameras to recover 3D data is the option by default in most cases. In this section we present two reconstruction methods, SfM and SLAM, and show how they are used to recover 3D data from a scene in AR applications.

### 3.4.1 Structure-from-Motion

#### Principle

SfM is a photogrammetric method that recovers 3D information from a set of 2D images (Hartley and Zisserman, 2003). It is widely used in augmented reality and robotics for the reconstruction of 3D rigid scenes. The images should be related by a local set of motion vectors, which can be obtained if there is a partial overlap between them. The recovered information usually consists of a 3D pointcloud that represents the scene from where the images were taken, along with the 3D camera poses. This reconstruction is up to a scale factor, which means that the scene dimensions do not correspond to the real ones. SfM can be classified as incremental SfM, where reconstruction is done by adding image after image, and global SfM, where reconstruction is done using all the images simultaneously.

In the case of global SfM, a common pipeline is illustrated in figure 3.6. First, keypoints are detected in every image using descriptors like SIFT or SURF and then matched across the images, from which the incorrect matches are filtered out using RANSAC. Second, the fundamental matrices are estimated for every pair of images from the keypoint matches using algorithms like Least Median Squares or the Normalised 8-point Algorithm. This matrix describes the epipolar geometry between a pair of cameras. Third, relative camera poses between pairs of views are computed from the fundamental matrices. Fourth, the pairwise relative camera poses are transformed to the coordinate system of the first camera. Fifth, keypoints are tracked across all the views from the pairwise correspondences. Sixth, the 3D points are computed from the tracked 2D points and the camera projection matrices. This problem is called “triangulation” and can be solved using Linear Least Squares. Seventh, the 3D point locations and the camera poses are refined using Bundle Adjustment, which aims to find the best 3D points and camera poses such that they minimise the 2D reprojection errors of the 3D points. This nonlinear optimisation problem is usually solved using the Levenberg-Marquardt algorithm.

#### Examples

SfM has been of great utility in the field of AR. For example, the work in (Fukuda et al., 2017) performs real-time AR by visualising 3D virtual structures in construction sites. In this case, a 3D pointcloud of the construction site is built with SfM and an initial alignment is done between the virtual structure and the pointcloud. Then, a real-time step compares the new images with the ones used for the reconstruction and automatically registers the virtual model to this image. Sim-

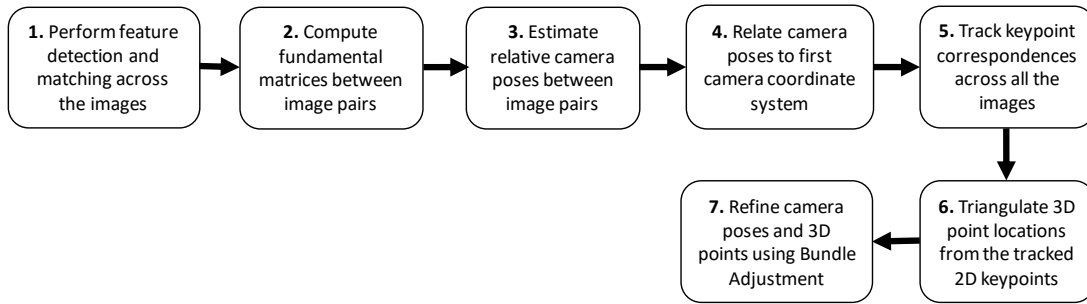


Figure 3.6: General pipeline for 3D reconstruction using SfM.

ilarly, the approach in (Bae et al., 2016) first reconstructs 3D pointclouds of urban structures and performs annotations on them. The pointcloud and the annotations are stored in a remote server, while the client is run in a mobile device. Then, during the tracking step, the client sends the new images to the server which will be compared with the existing pointcloud. Through a 2D-3D matching technique, the client’s relative pose is estimated and the images are augmented with the annotations in real-time. Another example is shown in (Mooser et al., 2009), where a markerless AR system is developed for augmenting urban images with artificial annotations. During the offline step, the user can attach the artificial annotations to the reconstructed pointcloud. In the online step, the method performs 2D-3D matching between every new frame and the original pointcloud to estimate the camera pose, but it is also able to include new features that are not available in the pointcloud. An AR method for navigation in large urban environments is proposed in (Gay-Bellile et al., 2010), where street images are augmented with virtual information in real-time. In this case, the reconstructed SfM pointcloud is corrected for drifts and deformations by registering it to coarse 3D city models from satellite data. In the case of MIS, the method in (Giannarou et al., 2012) proposes a deformable SfM method to reconstruct monocular endoscopic scenes. It combines visual data with inertial measurement data to recover the deforming areas of a previously made SfM reconstruction. This approach is targeted to model the deformations caused by tool-tissue interactions, which can be used for motion stabilisation of robotic surgical tools.

### 3.4.2 Simultaneous Localisation and Mapping

#### Principle

SLAM is a method for reconstructing and updating a map of an unknown environment, while simultaneously localising the agent which does the reconstruction over time. It is mainly used in robotics to automatically map unknown environments, but it has been increasingly used in other fields, including autonomous driving and AR. It can be divided between visual SLAM, which uses optical sensors like RGB or infrared cameras, and non-visual SLAM, which uses Light Detection and Ranging (LiDAR) or ultrasound sensors. Depending on the application and the type of sensors, different SLAM algorithms can be used. Concretely, the goal of SLAM is to estimate the agent’s location and the environment map through which the agent is moving, given a set of odometry measurements and a set of scene observations. As the sensors used are generally inaccurate, the SLAM problem has to deal with uncertainties in the measurements when estimating both the agent’s location and the environment map (Stachniss et al., 2016).

Common approaches used to solve SLAM include statistical methods like the Extended Kalman Filter (EKF), particle filters and graph-based optimisation techniques. They all compute accurate maps and agent locations from multiple inaccurate sensor data. More recently, methods like Parallel-Tracking-and-Mapping (PTAM) and ORB-SLAM have been used in Visual SLAM (VSLAM) to reconstruct a 3D pointcloud of the scene and recover the camera poses. This is especially useful in MIS, as the endoscopic camera is the only source of information. However, SLAM is designed to work in rigid scenes with non-deforming or moving objects, which is not common in MIS scenarios.

In the case of VSLAM (Mur-Artal et al., 2015), the reconstruction pipeline is shown in figure 3.7. First, the map is initialized by triangulating the feature correspondences from the first two image frames. Second, for every new frame, the camera pose is estimated from the feature tracks between the current frame and the last keyframe. Third, the algorithm checks if the current frame is a potential keyframe. A frame is considered to be a keyframe if it fulfills certain conditions regarding the number of elapsed frames since the last keyframe, and the number of tracked points. Fourth, if the current frame is a keyframe, the new feature points are triangulated and added to the global map. The map and the camera poses are refined using bundle adjustment. Fifth, Detect if the loop has been closed by comparing the current keyframe with all the others to decide if the place has already been visited. Sixth, if the loop is closed, optimise the global map and the camera poses to correct for possible drifts.

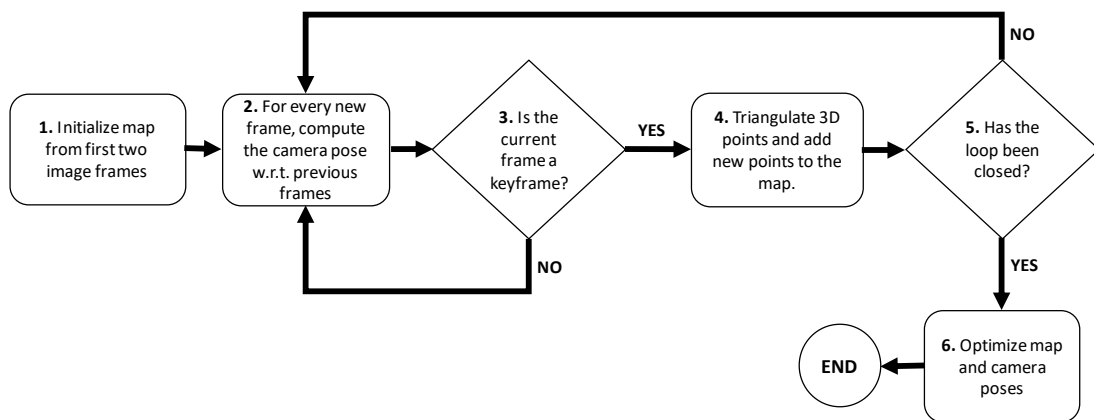


Figure 3.7: General pipeline for 3D reconstruction VSLAM.

### Examples

Several works use SLAM as a means to perform AR, like the approach in (Gao et al., 2017) where a 3D pointcloud of the scene is reconstructed using VSLAM and a 3D virtual model is registered into it. This method uses FAST and ORB features to robustly compute the camera poses and reconstruct the 3D pointcloud. Here, the registration accuracy is improved by detecting the dominant 2D plane in the scene and aligning the 3D model onto it. The method in (Liu et al., 2020) proposes a markerless mobile AR system that integrates data from an inertial measurement unit (IMU) into the SLAM process to reduce the drift over time. It starts by collecting several images to construct the initial map. Then, for every new image the feature points are extracted and matched with those from the reconstructed map. Using these matches and the IMU data, the camera pose is estimated

and the map is updated. The virtual model is simultaneously registered to every new image automatically. In (Abergel et al., 2019), a SLAM-based mobile platform for visualising 3D annotations of archaeological data. It extends an already existing AR platform that lets the user annotate on the images and reconstructs a dense pointcloud of the scene. This preexisting pointcloud is then registered to the SLAM pointcloud and the new images are augmented with the preexisting annotations. Another example of a SLAM-based mobile AR system is presented in (Munoz-Montoya et al., 2021), where it is used for conducting a study of short-term spatial memory. The system reconstructs a pointcloud of the room by using two RGB cameras and one depth camera. Once the pointcloud is obtained, the user can place virtual objects on any planar surface detected by the system.

### 3.5 Augmented Reality for Laparoscopic Liver Resection

As shown in section 1.3.3 of chapter 1, AR for LLR is usually divided into preoperative and intraoperative stages. The preoperative stage refers to the preparatory operations made before the surgery. The intraoperative stage refers to the operations made during the surgery that will ultimately lead to achieving AR. In this section we look in more detail on how these stages operate, the common software used and the data generated by each of them.

#### 3.5.1 The Preoperative Stage

The preoperative stage mostly deals with the preparation of the data required to perform AR. In our case, such data includes information like the type of surgery to be carried out, the location of the liver, tumours and vessels in the CT/MR data, and the reconstruction of the corresponding preoperative 3D models. These tasks are done as soon as the CT/MR data is available, with the support of the main surgeon. The reconstruction of the preoperative 3D models can be done using medical annotation software like MITK Workbench (German Cancer Research Center (DKFZ), 2008) or 3D Slicer (Kikinis, 2012). As shown in figure 3.8, the process starts by segmenting the organs and tissues of interest in all the slices in which they appear. Then, the 3D models are generated for each annotated structure by interpolating the 2D masks across the slices and reconstructing a triangular 3D mesh using methods like the Marching-Cubes Algorithm (Lorensen and Cline, 1987). These meshes contain hundreds of thousands of vertices and faces, and might have holes or non-manifold vertices and edges. When dealing with non-rigid registrations, it is advisable to simplify the models to a few thousands of vertices and remove the non-manifold components. In this way, the registration algorithm will run more efficiently and with less risk of numerical issues. This can be done using mesh processing software like Blender (The Blender Foundation, 1995) or Meshlab (Visual Computing Lab - ISTI-CNR, 2005).

#### 3.5.2 The Intraoperative Stage

The main goal of the intraoperative stage is to augment the laparoscopic images with the 3D models reconstructed during the preoperative stage. This process is done in two phases: The Initial Registration and the Real-time Tracking phases. In the Initial Registration phase, the intrinsic parameters of the camera required to perform AR are estimated, and the preoperative 3D data is registered to the intraoperative patient data. This registration brings the preoperative data into

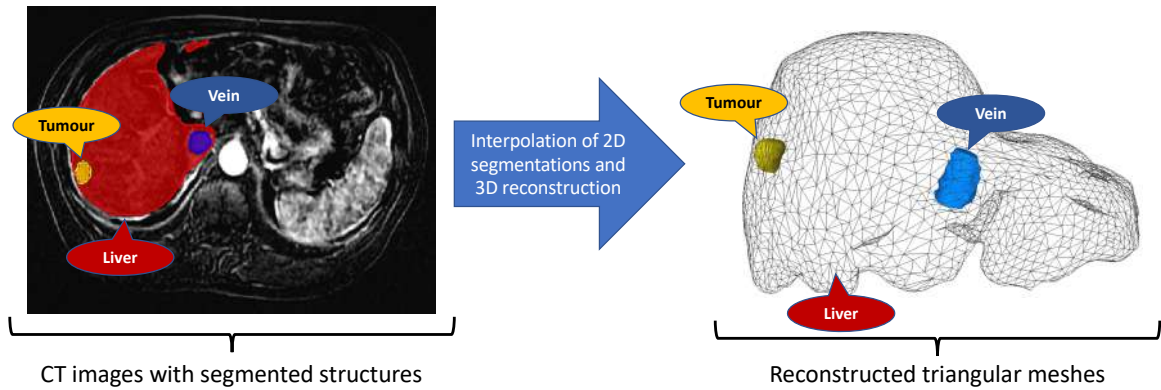


Figure 3.8: Segmentation of CT images and reconstruction of preoperative 3D models.

the reference frame of the intraoperative data, and can compensate for any deformation that may occur in between. In the Real-time Tracking phase, the new incoming images are augmented with the 3D models by updating the registration made during the preoperative stage. Depending on the nature of the organ being augmented, the registration update can be a rigid or a non-rigid one.

### Initial Registration

This step starts by estimating the intrinsic parameters of the camera, also referred to as camera calibration, which usually consist of the focal length  $\{f_x, f_y\}$  and the principal point  $\{u_0, v_0\}$ . The lens radial and tangential distortions are also computed. Then, the registration between the preoperative 3D data and the intraoperative data is done. It requires the matching of features between the two kinds of data, which can be done manually or automatically. The intraoperative data may come from optical devices, electromagnetic devices, or any kind of sensing device. It may consist of 2D images, or 3D data such as pointclouds, meshes, or 3D locations. Methods like SfM or SLAM can be used to reconstruct the intraoperative 3D data. If the intraoperative data is three-dimensional, 3D to 3D registration methods may be used like the ones shown in section 3.2. If it is two-dimensional, 3D to 2D feature-based methods are the advised way to go. The registration can be rigid or non-rigid, depending if the shape of the intraoperative data has changed with respect to the preoperative data. When the object to be registered is an organ that deforms in time, it is common to use a biomechanical model that simulates its elasticity parameters. This biomechanical model, along with the feature correspondences are used as constraints in the registration process. In this way, the preoperative 3D models will deform in a more realistic manner and the registration will be more accurate. The obtained transformation matrix  $T$  is then used to translate and rotate the preoperative 3D model so that it is inside the camera's FoV. Then, the image can be rendered by projecting the registered model in 2D and using any commercially available rendering software like OpenGL (Khronos Group, 2022) or Unity (Unity, 2022).

In order to render a scene of a registered 3D model for AR purposes, several transformations should be done to project the 3D model into the 2D image, as shown in figure 3.9. For instance, in OpenGL the 3D model is first expressed in a local coordinate frame. This frame is usually located at the centre of the model, although it can also originate at the coordinate frame of the device that generated the model. Second, the 3D model is transformed from local to world coordinates using a model matrix  $T_{model}$ . The world coordinate frame is usually located at the origin (no rotation or translation). Third, the model is transformed into view coordinates using the view matrix  $T_{view}$ .



This transformation aligns the model in a way that it is located inside the camera's FoV. In OpenGL, the camera is also located at the origin (no translation), but with its positive x axis pointing towards the negative x axis of the world coordinate frame. Fourth, the model is projected into the camera frame, also called clip coordinates, using the projection matrix  $K$ . This matrix can define a perspective or an orthographic projection, and uses the intrinsic camera parameters obtained in the calibration process. The clip coordinates are normalised between -1 and 1 in every axis, and the 3D data that falls outside this range after projection does not appear in the final image. Fifth, the clip coordinates are transformed into screen coordinates through a viewport transform function. Here, the coordinates are transformed from -1 to 1 to a range defined by the size of the viewport. The viewport is the image canvas where the final rendering is done. Once the 3D model  $M$  is transformed from local coordinates to clip coordinates  $m_{clip}$  as shown in equation 3.5, OpenGL transforms the model from clip coordinates to screen coordinates automatically.

$$m_{clip} = KT_{view}T_{model}M \quad (3.5)$$

In our case, the estimated transformation  $T$  is the combination of the model and view matrices  $\{T_{model}, T_{view}\}$ . Therefore, equation 3.5 can be simplified to:

$$m_{clip} = KTM \quad (3.6)$$

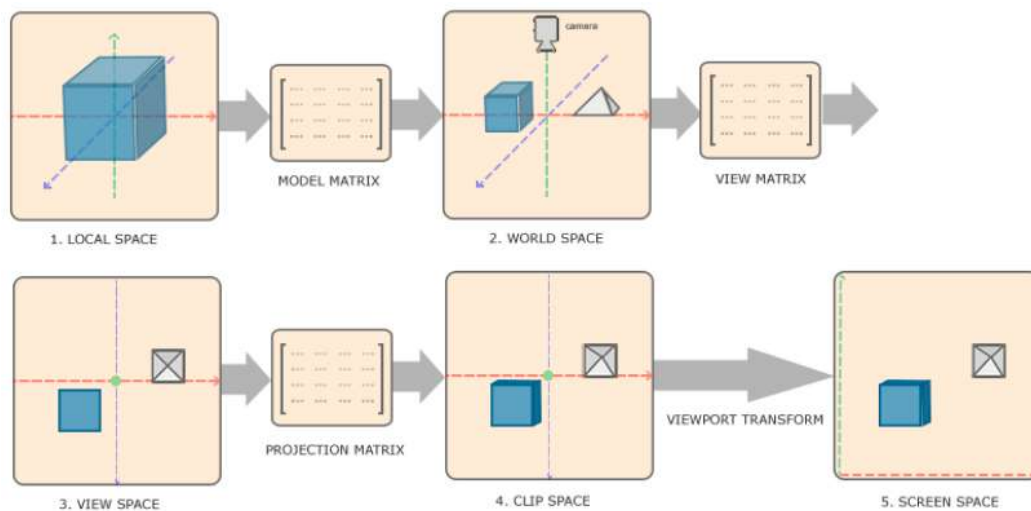


Figure 3.9: Transformation pipeline for rendering in OpenGL (de Vries, 2014).

As mentioned in section 1.3.3 of chapter 1, our work is based on the SVB method proposed in (Koo et al., 2017) and (Özgür et al., 2018). It is a monocular 3D-2D approach that uses several anatomical landmarks of the liver to guide the registration. It has been implemented in C++ and OpenGL, in a stand-alone AR software called *Hepataug*. SVB uses three types of landmarks, as shown in figure 3.10: the *ridge* contour, the *falciform ligament* contour and the *silhouette* contour. The ridge contour corresponds to the pronounced curve located at the bottom-anterior part of the liver. It is the most distinguishable and stable landmark among the three, as finding a direct correspondence between the laparoscopic image and the 3D model is a relatively easy task. The falciform ligament is the thin, fibrous tissue that connects the anterior part of the liver to the abdominal wall. It is usually cut during a laparoscopic procedure to facilitate the manipulation of

the liver. The remnant of this tissue on the liver's surface is what we use as a contour. The silhouette contour corresponds to the boundary of the liver at a given image and, thus, does not have a direct correspondence in the 3D model. The 3D correspondences are found by the registration algorithm using ICP. Along with the landmark constraints, SVB also uses a biomechanical model based on PBD (Müller et al., 2007) with a Neo-Hookean material model to simulate the elasticity properties of the liver.

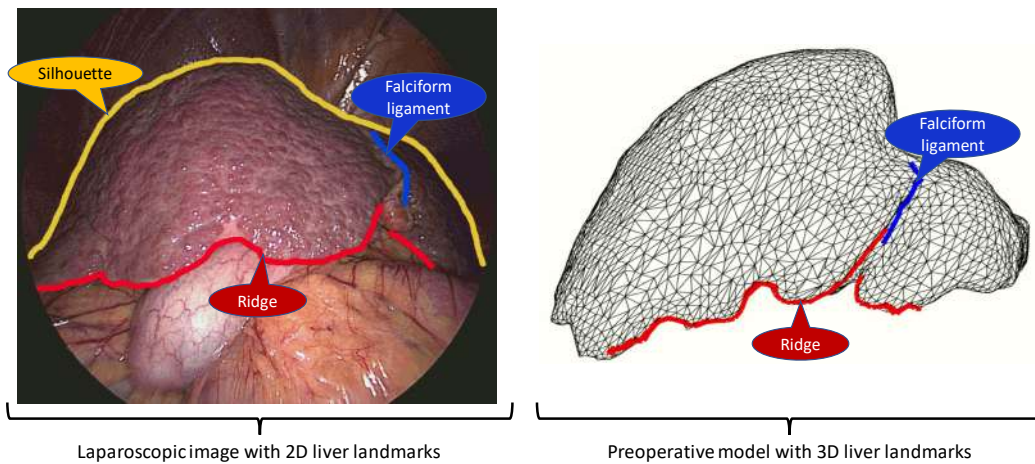


Figure 3.10: Liver anatomical landmarks used as constraints for SVB registration.

As shown in algorithm 3.1, SVB starts by marking the 2D image contours and the corresponding 3D model contours. Then, the preoperative model containing the liver's surface, tumours and vessels is tetrahedrised using a *Delaunay triangulation* technique. The resulting shape is a volumetric convex model, whose tetrahedrons are used to simulate the biomechanical constraints of the liver. Two main problems arise from this strategy: First, the global shape changes drastically as the original non-convex liver model becomes a convex one. Second, as no inner points are generated in the preoperative 3D model, the generated tetrahedra are small in quantity, and their shapes are big and elongated. This may drastically reduce the accuracy of the deformation. After Delaunay triangulation, the 3D model is rigidly registered to the laparoscopic image. This can be done manually or automatically using ICP with the annotated landmarks. From this initial rigid registration, the 3D model is deformed by solving the biomechanical and landmark constraints in an alternate way, using a *Gauss-Seidel* iterative approach. The biomechanical constraints are optimised following the PBD approach from (Bender et al., 2014) and using the elasticity parameters of a real human liver. The landmark constraints are solved in a 2D-projection way, for which the 3D-2D correspondences are recomputed at every iteration. The correspondences are found by projecting the 3D landmarks and, as there is a higher number of 2D landmark points than 3D points, each projected 3D landmark point is matched to the closest 2D landmark point. Then, the 2D point correspondences are back-projected in 3D as line-of-sights. The 3D landmark points are then projected to these line-of-sights and an iteration of biomechanical optimisation is done with the 3D landmarks fixed in place. The alternation between biomechanical and landmark constraints continues until the difference between the models of the previous and current iterations is lower than a threshold, or a number of iterations has passed. A common problem of PBD-based approaches is the slow decrease of the residual and the lack of convergence (both theoretically and practically) as the constraints are solved locally, compared to global solvers like *Gauss-Newton*

(Bender et al., 2014; Müller et al., 2007). Instead, the goal of PBD is to achieve fast simulations with visually plausible results, at the cost of less accurate deformations. An example of AR using SVB on different patients is presented in figure 3.11.

**Algorithm 3.1:** SVB registration method proposed in (Koo et al., 2017)

```

1 Function SVBRegistration
   Input:  $M_0, I$  // Preoperative model, laparoscopic image
   Output:  $M$  // Registered model
2   max_iter ← 1000 // Maximum number of iterations
3    $\epsilon \leftarrow 10^{-3}$  // Model distance threshold
4    $\{c_{image}\} \leftarrow \text{markImageContours}(I)$   $\{C_{model}\} \leftarrow \text{markModelContours}(M_0)$ 
5    $M_0 \leftarrow \text{rigidRegistration}(M_0)$  // Perform initial rigid registration
6    $M_0 \leftarrow \text{tetrahedrizeModel}(M_0)$  // Generate tetrahedral model using Delaunay triangulation
7    $i \leftarrow 0$  // Initialise iteration counter
8   while  $\text{dist}(M_{i+1}, M_i) > \epsilon$  or  $i < \text{max\_iter}$  // Main registration loop
9   do
10     $M_{i+1} \leftarrow M_i$  // Update model from previous iteration
11     $M_{i+1} \leftarrow \text{solveBiomechanicalConstraints}(M_{i+1})$ 
         $\{c_{model}\} \leftarrow \text{projectModelContoursIn2D}(C_{model}, M_{i+1})$ 
         $\{p_{image}\} \leftarrow \text{findImageCorrespondences}(c_{model}, c_{image})$  // Find correspondences
        between projected model contours and image contours
12     $\{P_{image}\} \leftarrow \text{backprojectImageCorrespondences}(p_{image})$  // Backproject 2D contour
        correspondences in 3D as line-of-sights
13     $M_{i+1} \leftarrow \text{projectModelContoursToImageCorrespondences}(M_{i+1}, P_{image}, C_{model})$  // Project
        model contour points to the backprojected line-of-sights
14     $M_{i+1} \leftarrow \text{solveBiomechanicalConstraintsModelContoursFixed}(M_{i+1}, C_{model})$  // Perform
        biomechanical optimisation with the projected model contour points fixed in place
15     $i++$  // Increment iteration counter

```

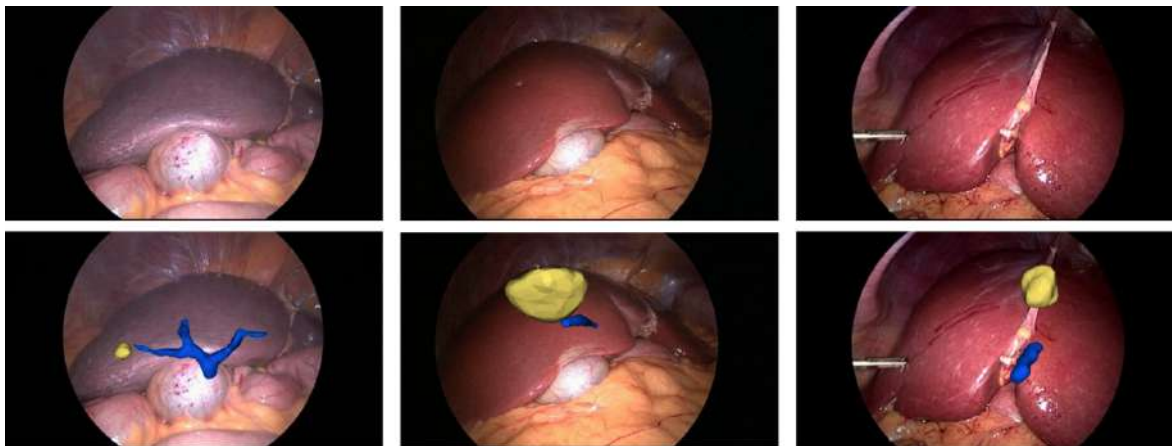


Figure 3.11: Examples of AR using SVB in LLR for 3 patients. Top: Original laparoscopic images. Bottom: Augmented images with 3D models of tumours (yellow) and veins (blue).

### Real-Time Tracking

This step augments each of the incoming images from the camera's video stream with the preoperative 3D models, taking as basis the registration made in the initial step. The tracking is made at a sufficiently high speed, usually at the camera's original frame rate, in a way that the organ's movements remain smooth and coherent, hence the real-time adjective. Most of the state-of-the-art works compare the new image with the ones from the initial step, and choose the one that resembles the most. The relative camera pose is computed between these two images and it is used

to update the 3D model's pose from the chosen initial image. As for the initial registration, the real-time tracking can be rigid or non-rigid depending on the organ's elasticity properties. A good example of this is shown in (Collins et al., 2021), where an AR system for monocular laparoscopy of the uterus is developed. The system is composed of a multiple-view initial non-rigid registration step and a real-time rigid tracking step. The registration pipeline uses an intraoperative model reconstructed through SfM as the registration target. Another example is presented in (Haouchine et al., 2013), where an AR system for stereo laparoscopy of liver is proposed. Because the liver is highly deformable, they proposed a non-rigid tracking pipeline that updates the model's shape from the previous image frame. Similarly to the method in (Collins et al., 2021), this method uses an intraoperative pointcloud reconstructed from stereo techniques as a target for registration. An AR system for monocular laparoscopy of the kidney is proposed in (Puerto-Souza et al., 2014), where a single image is used to perform initial 3D-2D registration. After registration, the visible 3D model points are associated with the 2D image features for tracking purposes. During the real-time step, the feature tracking is done in a sequential way but the model's pose is updated from the initial image according to the associated features. Therefore, the accuracy of the augmentation will depend on the quality of the initial registration and the amount of available features in the initial image.



## Chapter 4

# Single-View Registration Combining Visual Cues with Interactions

---

We present in here the Single-View Hybrid (SVH) registration method, which combines visual landmarks with tactile interactions to perform registration. The goal of this method is to overcome the lack of visual constraints in registration due to partial views of the liver. It uses a cage that encloses the 3D liver model to register the parts that are not constrained by any visual landmark. In this way, the registration accuracy can be improved, as demonstrated by experiments made on phantom, ex-vivo animal data, and in-vivo patient data. We compare our method with a single-view rigid method and a single-view deformable method.

---

## 4.1 Introduction

One of the main limitations of laparoscopy is the difficulty to accurately localise the target organ's internal anatomy, owing to the absence of tactile feedback. This is a particularly important issue for the liver, which may contain malignant tumours to be precisely resected with an oncologic margin. AR overcomes this limitation by overlaying the information extracted from a preoperative CT/MR volume onto the laparoscopy images, allowing the surgeon to perform resection planning. To achieve this, the preoperative 3D model is deformably registered into the laparoscopic image. Compared to classical mental mapping approaches used in laparoscopy such as (Lamata et al., 2010), AR systems like the proposed one cope with the deformation undergone by the liver from the preoperative to the intraoperative stages. Also, by directly overlaying the laparoscopic images with the registered preoperative model of the intraparenchymal structures, rather than mentally mapping them from a separate screen to the laparoscopic image, surgeons can be more confident regarding the real location of these structures.

However, due to the fact that the liver is only partially visible in the laparoscopy image due to its large size and proximity to the laparoscope, only a small part of the 3D model is constrained during registration. This is especially true for single-view registration methods like the ones presented in (Nicolau et al., 2011; Özgür et al., 2018; Koo et al., 2017; Adagolodjo et al., 2017). In (Nicolau et al., 2011), the rigidity assumption is far too restrictive to accurately model the liver deformation. In (Özgür et al., 2018; Koo et al., 2017; Adagolodjo et al., 2017), the visual cues are sparse and do not convey enough information to unambiguously constrain registration. This may lead to inaccuracies in the augmentations, making AR not reliable for resection planning.

We propose a *hybrid* registration approach. The key idea is that the manual and automatic approaches are highly complementary. Our hybrid approach extends and draws on the strengths of both by combining user interaction with visual cues and a biomechanical model. In other words, the rationale is that both the machine and the user perception are valuable and should be taken into account via the visual cues and interaction respectively. In the presence of both user interaction and visual cues, our hybrid approach bundles all constraints in a single registration. In the absence of user interaction, it behaves similarly to the existing automatic approaches. In the absence of visual cues, it allows the user to edit the registration under guidance of the biomechanical model. This is a significant improvement compared to the existing manual approach as it allows the user to fully express their expertise in anatomy, prior experience and spatial understanding of the case at hand to the system. We have implemented this idea following the cage-based paradigm from the field of shape editing. The cage is a set of control points enclosing the organ. Dragging these control points interactively deforms the model. Shape editing is a widely studied problem. The cage-based paradigm is well-adapted to registration owing to its flexibility. Concretely, we implemented our hybrid method with a Qt Graphical User Interface (GUI). Our system is entirely controllable by tactile interaction and may be used in a fast and intuitive manner. We compared our method named *Single-View Hybrid (SVH)* quantitatively and qualitatively against the *Single-View Rigid (SVR)* method proposed in (Nicolau et al., 2011), and the SVB method (Koo et al., 2017). The quantitative evaluation was done on phantom and ex-vivo animal data, while the qualitative evaluation was done on in-vivo patient data.

## 4.2 Background

### 4.2.1 Interactive Shape Editing

Shape editing refers to the change of a model's surface through a set of control points either part of or connected to it. Existing approaches can be divided in four main categories, depending on how the control points are distributed: point-based (Bechmann and Gerber, 2003), curve-based (Barr, 1984), surface-based (Feng et al., 2006) and cage-based deformations (Sederberg and Parry, 1986). In a point-based approach, the user provides a set of point displacements, each comprising a point along with its intended motion and region of influence. The way points are distributed does not depend on the shape of the model, but more on the user's preference. When these points are moved, the object is then warped to match the displacement constraints (Bechmann and Gerber, 2003). In curve-based approaches, the deformations are controlled by one or more curves. The control points are distributed to form a line that the user deforms. The deforming object is distorted to map from the source to the destination curves (Barr, 1984). The surface-based approaches consist in deforming the object when a surface patch is modified by translating a set of control points. The control points are directly located on the surface of the model. One of the main difficulties is to find a way to attach sample points on the object to the deforming patch (Feng et al., 2006). The cage-based approaches use a cage that encloses the object. This cage can have a fixed shape such as a cuboid (Sederberg and Parry, 1986), or can be adapted to the shape of the object to be deformed (Sacht et al., 2015). The shape of this cage is altered by repositioning control points. The resulting cage distortion is then transferred to the object.

## 4.3 Hybrid Registration Combining a Biomechanical Model, Visual Cues and Manual Interactions

### 4.3.1 Registration Pipeline

Our hybrid registration method takes as input a preoperative 3D model and a single laparoscopic image. Its principle is to combine a biomechanical model and the manual and automatic registration approaches. These respectively use user interaction and visual cues extracted from the image to solve for registration. Our method thus rests on three sets of constraints. The first two are borrowed from (Koo et al., 2017). These are a biomechanical model based on the Neo-Hookean elastic model and the use of the falciform ligament and inferior ridge as curve correspondences, and the silhouette. The third set of constraints are the cage-based constraints to model user interactions. As explained in section 3.5.2 of chapter 3, the preoperative 3D model is represented by a tetrahedral mesh and optimisation follows the principle of PBD (Bender et al., 2014).

The pipeline of our method is illustrated in figure 4.1. It has 7 main steps, with the first two being similar to SVB. In step (1), the user marks the falciform ligament, inferior ridge and silhouette on the laparoscopy image. In step (2), the user marks the corresponding contours on the preoperative 3D model. In step (3), the system generates a cage enclosing the preoperative 3D model, to be used for user interaction at step (6). In step (4), the cage's control points and the preoperative 3D model are co-tetrahedrised in order to obtain a single tetrahedral model. In step (5), an initial registration is computed using only the visual cues, following SVB. This initial registration is required to initiate interactive registration. In step (6), the user interactively edits the registration



by moving the cage's control points. The registration is updated in real-time to provide the user with live feedback. Importantly at this step, both the cage's control points and the visual cues are used to update the registration. Finally, once the user is satisfied with the registration, step (7) augments the laparoscopic image with hidden anatomical elements transferred from the preoperative 3D model.





### 4.3.2 Cage-Based User Interaction

An intuitive and easy-to-use interface allowing the user to edit the liver's shape in a way that respects its properties and the visual constraints must be proposed. This is achieved through the use of a cage. This has a good trade-off between editing flexibility, namely the possibility to edit at an appropriate spatial frequency, and user friendliness. The cage is represented by a mesh composed of a very limited number of control points. These are obtained following the cage initialisation procedure defined in (Sacht et al., 2015) so that the cage encloses the input preoperative 3D model. Once the cage is generated, it is linked to the preoperative 3D model through a Delaunay triangulation applied on all the vertices, namely the cage, the liver and its inner structures vertices. During the registration, they are all optimised using the same material model. An example of a generated cage is shown in figure 4.2.

Making the constraints derived from the cage movable during the optimisation procedure is not trivial. We propose to embed once and for all the cage's control points into the volumetric model of tetrahedral topology built from the preoperative 3D liver model. Another possible solution would consist in creating a cage according to the model deformation at every iteration, which would however significantly harm the software usability. When a cage's control point is moved during the optimisation, all the vertices belonging to the liver model adjacent to it are moved accordingly. This allows the user to handle a set of vertices simultaneously over a model region. These deformations are compensated by the contour constraints simultaneously. During optimisation, a single iteration of the contour-based optimisation is run for every change in position of the cage vertices in order to increase the responsiveness of the deformation.

### 4.3.3 Tactile Graphical User Interface

The proposed GUI is shown in figure 4.2. It is divided in four sections. First, the visualisation area in which the laparoscopic image and the preoperative 3D model are shown. The user can position the preoperative 3D model and mark the contours using tactile gestures or the keyboard and mouse. Second, the left toolbar, which is used to either import or export the laparoscopic image and the preoperative 3D model. Third, the right toolbar, which is used to modify the appearance of the preoperative 3D model, to mark the visual cues and to launch registration. Fourth, the bottom toolbar, which controls the size of the visualisation area, lets the user activate the cage-based editing mode, and implements miscellaneous other functionalities.

Registration begins with a user click on  to set the laparoscope parameters obtained from a prior calibration procedure. The button  is then used to load the laparoscopic image and the button  to load the preoperative 3D model. The visual cues are marked both in the laparoscopic image and the preoperative 3D model with the help of the controls located in the right toolbar. A rigid registration is done automatically so that the preoperative model fits the liver approximately in the laparoscopic image. The nested cage is generated by clicking on . The user then proceeds

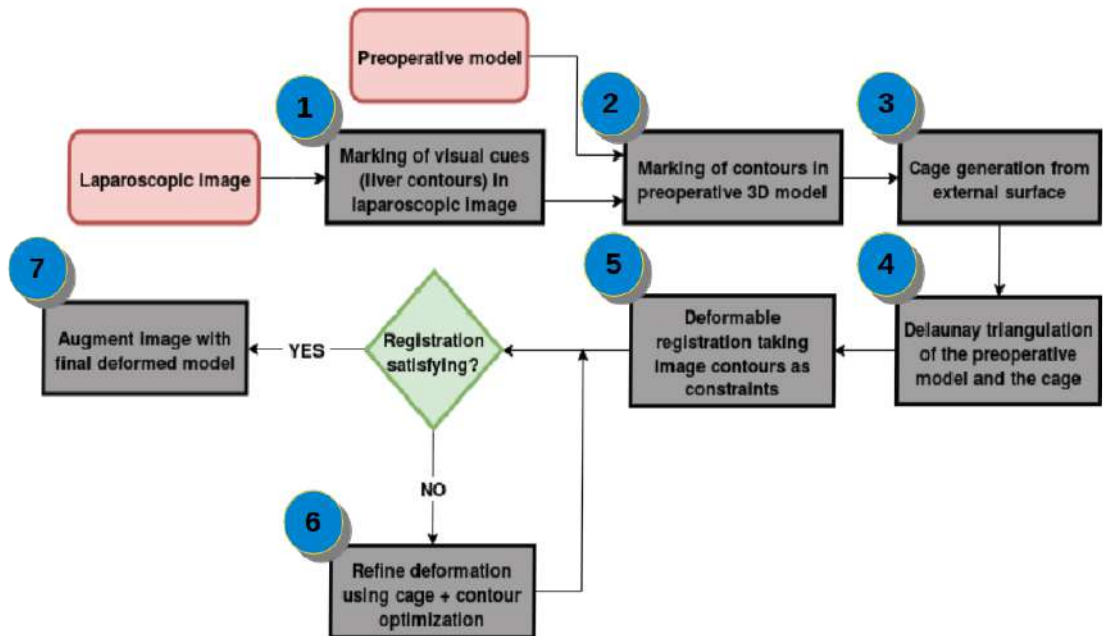
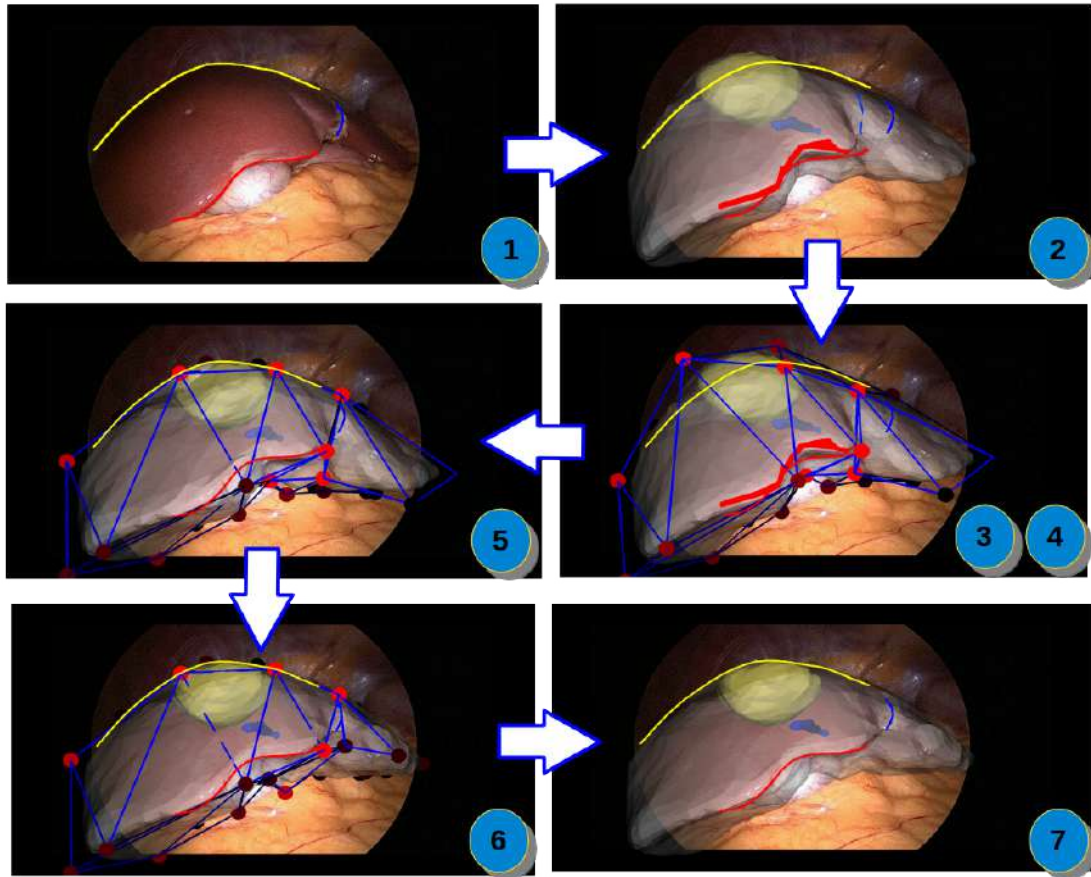


Figure 4.1: Pipeline of the proposed hybrid 3D to 2D deformable liver registration method. The liver surface mesh is overlaid in gray, its subsurface tumour in yellow and vein in blue. The contours associated to the silhouette, the falciform ligament and the ridge are marked in yellow, blue and red, respectively. The cage is rendered in blue wireframe and its associated control points with red dots.

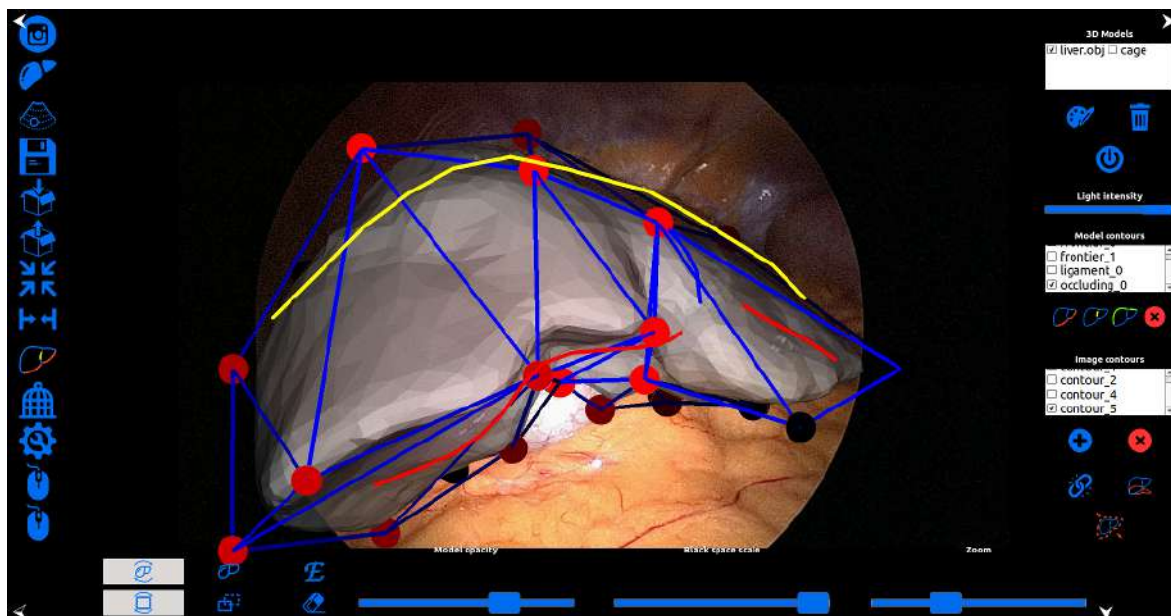





Figure 4.2: Proposed tactile Graphical User Interface for our hybrid method.

to match the contours by clicking on . The preoperative 3D model can then be translated and rotated so that they approximately fit the image. Automatic contour-based deformation is then launched by clicking on . Once it completes, the user can proceed to edit the registration using the cage by clicking on the button . The vertices of the cage may be dragged while the system displays the registration combining the visual cues and the cage in real-time.

Our AR software is setup on a PC computer running Linux. In the Operating Room (OR), this computer is connected via a capture card to the laparoscopy column in order to capture the laparoscopic video stream. The computer is located close to the other screens so that the surgeon has a direct view of the augmentation (see figure 4.3). It is equipped with a tactile screen, which can be directly used by the surgeon. For the surgeon to be able to interact with the computer, the device is sterilised before entering the surgery room and then covered with a sterile plastic film.



Figure 4.3: Usage of our AR guidance system (left computer) in the operating room to perform a laparoscopic tumour resection.

## 4.4 Quantitative Evaluation on Phantom Data

### 4.4.1 Validation Pipeline

The accuracy of the proposed registration method is evaluated using a 3D printed liver phantom made of silicon, aiming to simulate the bio-mechanical properties of a real liver. The liver phantom was built as follows. The preoperative 3D model was segmented from CT data of a real patient. Then, a mold of this 3D liver was generated and 3D-printed. Finally, the mold was filled with silicon (figure 4.4). We used an Ecoflex 00-20 silicon material made by Smooth-On Inc. which has a Young elastic modulus of 60kPa (Adams et al., 2017), very close to the 50 - 60kPa of a human liver.

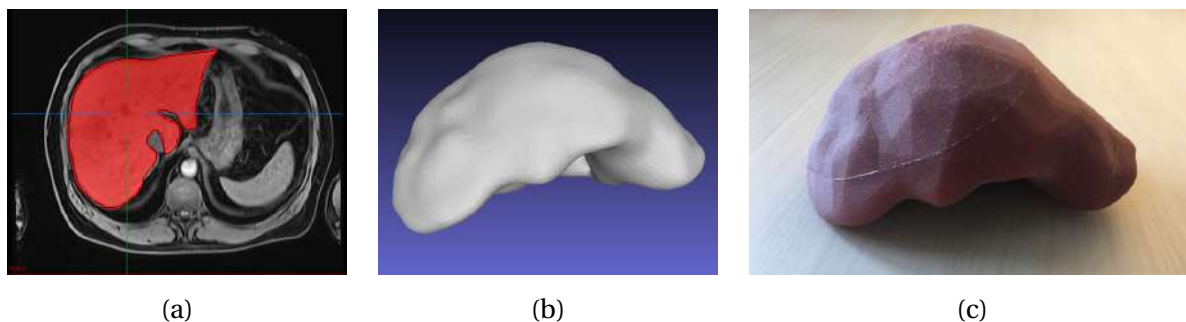


Figure 4.4: (a) Segmentation of the liver in a patient CT image. (b) Preoperative 3D liver model obtained from the segmentation and used to 3D print a model for (c) the liver phantom used for the proposed accuracy evaluation of the registration methods.

The principle of this experiment is as follows. The liver phantom is deformed and its shape reconstructed using the Structure-from-Motion software Agisoft Metashape (Agisoft LLC, 2022), as shown in figure 4.6. Then, we take  $N$  views out of those used to reconstruct the phantom's

shape as input images for the registration procedure. The CAD model from which the phantom was printed is used as the input preoperative 3D model and is registered following the proposed registration method (figure 4.5).

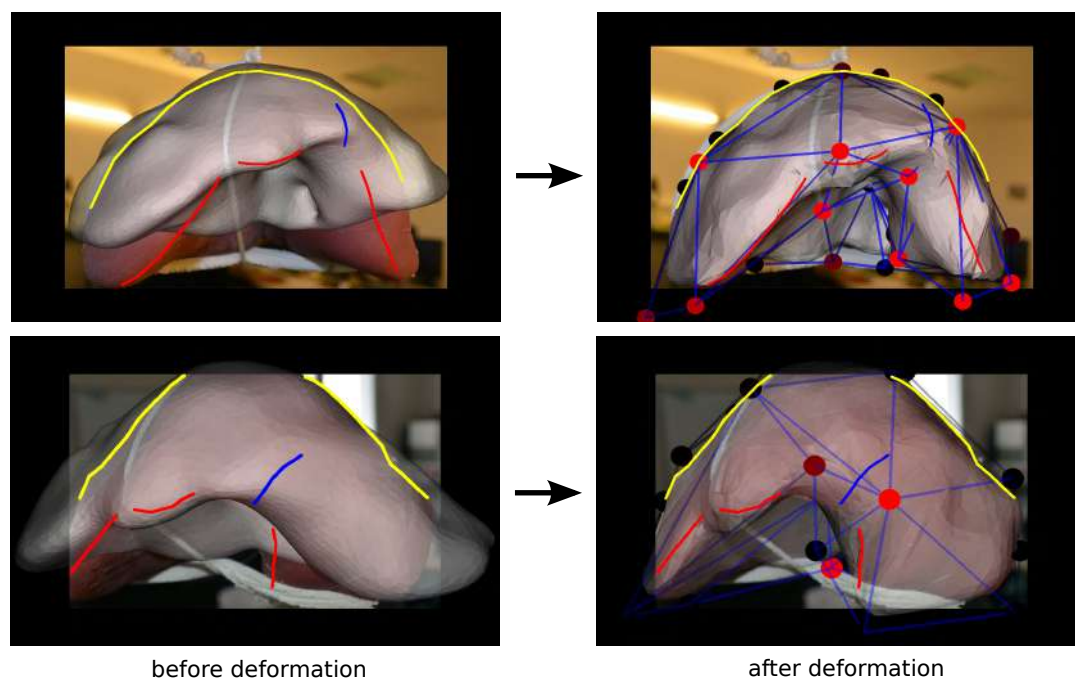


Figure 4.5: Registration of the preoperative 3D model on phantom data with the SVH method.

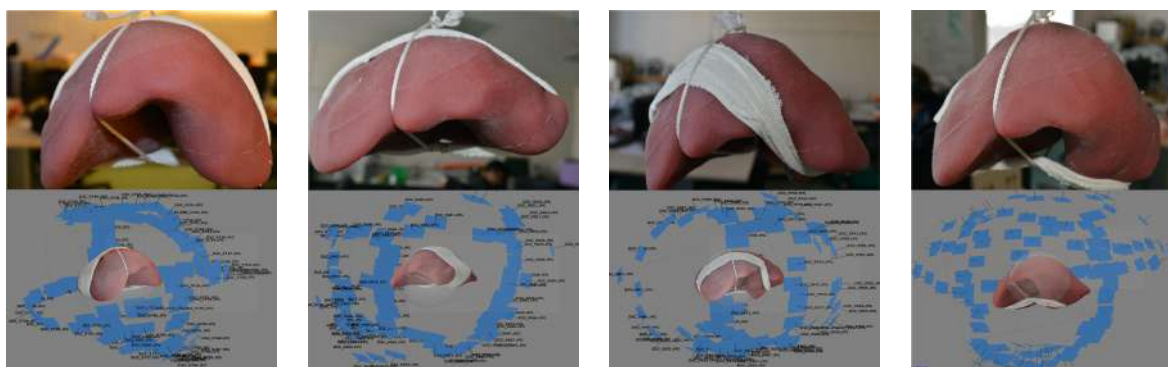


Figure 4.6: The deformations applied to the 3D printed phantom used in our first set of experiments. Top: deformed phantom. Bottom: 3D model, reconstructed with the Agisoft Metashape software, used as ground-truth in our experiments.

This experiment is performed for 4 phantom deformations, shown in figure 4.6, and 5 different views per deformation. The registration error, defined as the average distance between the vertices of the registered preoperative and ground-truth models, is reported in Table 4.1. As we compared the distances between all the vertices and not only the ones involved in the registration, it can be considered a measurement of TRE (Moghari et al., 2008).

#### 4.4.2 Experimental Results

We report the errors for two sets of vertices: those associated to the entire registered model and those restricted to its visible part, namely the anterior part, corresponding to a usual laparoscopy

image. The visible part of the model is defined manually by creating masks on the rendered views of the reconstructed shape in order to create realistic liver occlusions. In order to replicate the typical occlusions the liver presents during the surgery (due to, e.g., other organs or fat tissues), some visible parts of the reconstructed shape were manually masked in the rendered views (red region in figure 4.4(c)).

Registration error for whole liver (mm)			
Dataset ↓	SVR	SVB	SVH
1	09.00 ± 2.82	<u>05.35 ± 1.26</u>	<b>04.10 ± 0.39</b>
2	<u>06.19 ± 2.26</u>	08.65 ± 5.04	<b>05.05 ± 0.9</b>
3	12.23 ± 1.84	<u>10.32 ± 2.17</u>	<b>08.46 ± 1.26</b>
4	08.60 ± 1.9	<u>06.78 ± 0.8</u>	<b>05.70 ± 0.42</b>
Average	09.01 ± 2.2	<u>07.75 ± 2.31</u>	<b>05.82 ± 0.74</b>

Registration error for visible part (mm)			
Dataset ↓	SVR	SVB	SVH
1	11.09 ± 4.66	<u>07.96 ± 5.3</u>	<b>04.62 ± 1.02</b>
2	<u>06.77 ± 4.43</u>	07.78 ± 4.9	<b>04.11 ± 1.36</b>
3	12.67 ± 5.26	<u>09.43 ± 3.32</u>	<b>05.60 ± 1.66</b>
4	10.46 ± 7.47	<b>06.67 ± 1.03</b>	<u>07.80 ± 1.63</u>
Average	10.24 ± 5.45	<u>07.96 ± 3.63</u>	<b>05.53 ± 1.41</b>

Table 4.1: Registration error with respect to ground truth for the phantom experiment. The errors are expressed in millimetres and correspond to the mean and standard deviation of the distances between the registered model's vertices. Best results are in bold, second best are underlined.

SVH shows the lowest registration error. The error of SVR is noticeably higher as the method does not deal with the phantom deformations. The registration error of SVB is overall lower than SVR's. It shows that the visual cues in SVB well constrain the biomechanical model. SVH shows lower errors than SVB as misaligned parts can be corrected while preserving the visual cues and biomechanical constraints. The standard deviations are the lowest for SVH which shows that the method also provides the most stable results. In some cases, such as for example the registrations SVB performed on dataset 1 or SVH on dataset 4, the average error over the entire liver is lower than the error over its visible part, which reveals lower registration errors on hidden parts. To better illustrate this, error distributions over the entire liver's vertices are shown in figure 4.7.

The registration accuracy is evaluated with varying visibility of the liver phantom. A decreasing FoV was simulated by adding a circular black border to the images. The registration were performed for a FoV of 100%, 70% and 50%. One image per dataset were used to perform SVR, SVB and SVH registrations. The results for the whole liver are reported in Table 4.2. Some registration results are illustrated in figure 4.8.

The accuracy is also assessed for SVB with a varying number of tetrahedrons that compose the biomechanical model. Three preoperative 3D models were created, comprising tetrahedrons obtained by the triangulation applied on 8000, 4000 and 2000 vertices respectively. SVB was run on one image per dataset. The registration error for the whole liver is reported in Table 4.3.

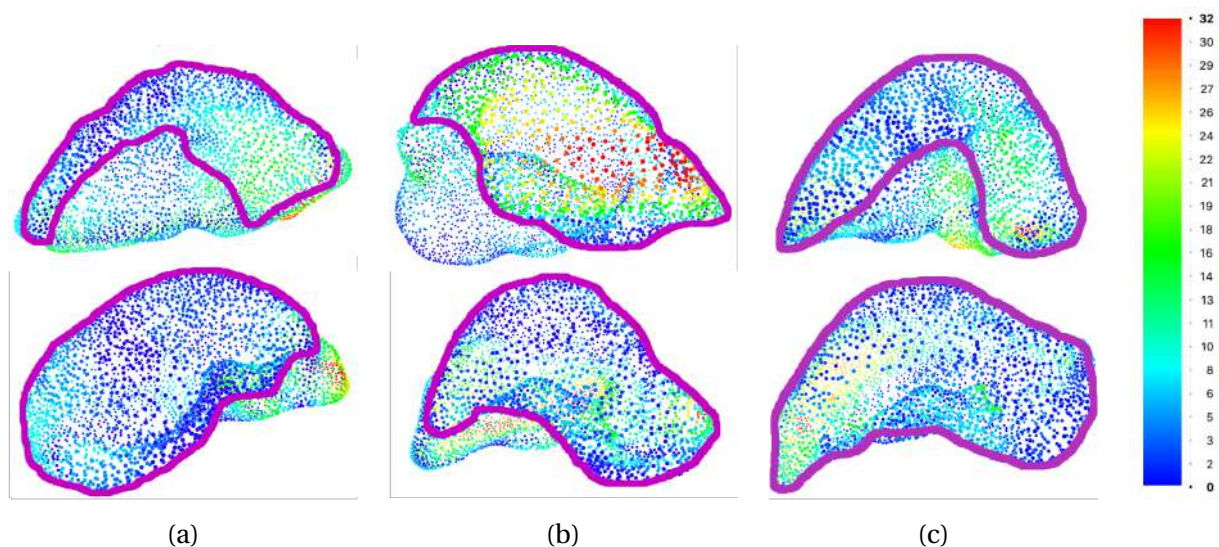


Figure 4.7: Error distribution over the registered phantom models using (a) SVR, (b) SVB and (c) SVH. The colors range from blue which corresponds to the lowest registration errors to red which corresponds to the highest ones. Distances are in millimeters. The visible parts correspond to the areas limited by the purple curve. Top: cases associated to a registration error higher on the visible parts than on the hidden parts. Bottom: cases associated to a registration error higher on the hidden parts than on the visible parts.

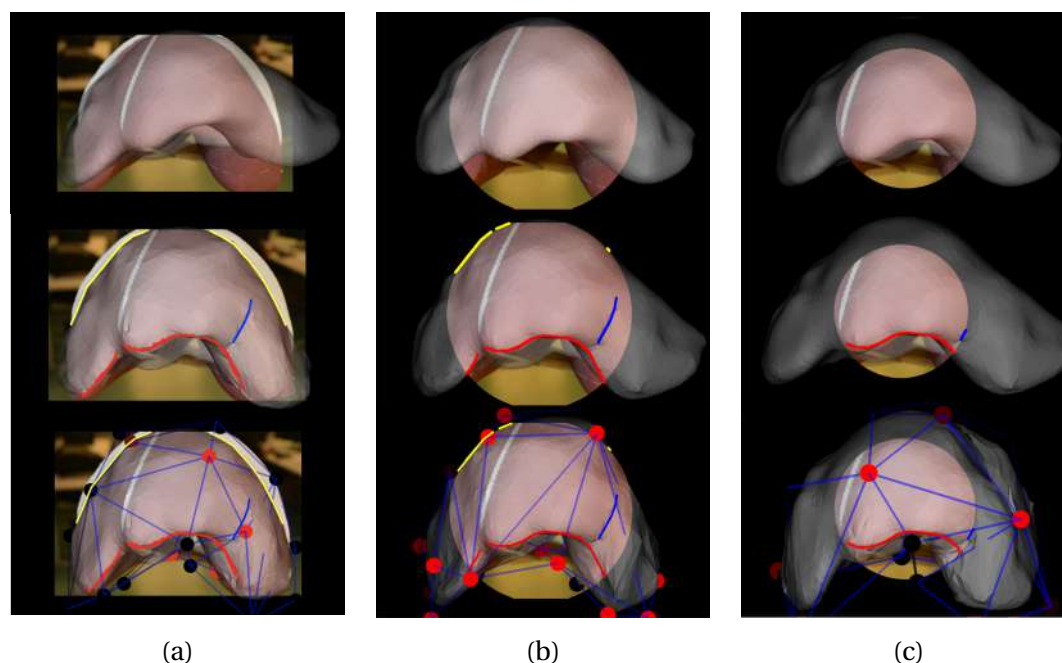


Figure 4.8: Registration results for 3 different FoV on the first dataset: a FoV of (a) 100%, (b) 70% and (c) 50%. The top images correspond to registrations using SVR, the middle images to registrations using SVB, and the bottom images to registrations using SVH. The circles represent the FoV applied in each case.

FoV →	100%	70%	50%	Avg
SVR	07.68	12.47	15.51	11.89
SVB	<u>06.75</u>	<u>07.15</u>	<u>07.59</u>	<u>07.16</u>
SVH	<b>05.88</b>	<b>06.47</b>	<b>06.60</b>	<b>06.32</b>

Table 4.2: Registration error for the whole liver (in mm) for decreasing FoV using SVR, SVB and SVH. The best results are in bold and the second best underlined.

Dataset →	1	2	3	4	Avg
2000 vertices	<u>06.22</u>	05.46	<b>07.43</b>	07.62	<b>06.68</b>
4000 vertices	<b>04.28</b>	<b>03.88</b>	<u>11.46</u>	<u>07.39</u>	<u>06.75</u>
8000 vertices	07.06	<u>04.59</u>	11.55	<b>06.71</b>	07.47

Table 4.3: Registration errors for the whole liver using the SVB method with varying number of vertices. The errors are expressed in millimetres and correspond to the average distance between the registered model’s vertices.

## 4.5 Quantitative Evaluation on Ex-Vivo Animal Data

### 4.5.1 Validation Pipeline

We assess the accuracy of our method with respect to inner tumours, by means of an ex-vivo sheep liver which was injected with alginate to create three artificial tumours. Two CT scans of the liver were made. The first one was performed to build the preoperative 3D model to register. The liver was then deformed. The second CT scan was performed on the deformed liver together with a SfM-based 3D reconstruction to obtain a registration ground truth (see figure 4.9). The registrations were made on two laparoscopic views of the deformed liver using SVR, SVB and SVH for two degrees of visual cues visibility: low and regular. This emulates the possible occlusions from fat and the surrounding organs.

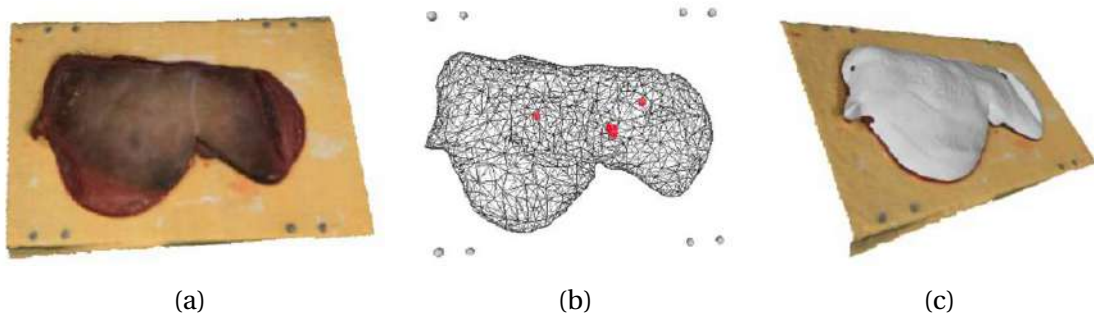


Figure 4.9: (a) SfM-reconstructed model of the deformed liver. (b) CT-reconstructed models of the deformed liver and tumours (in red). (c) SfM and CT models aligned using ABSOR.

### 4.5.2 Experimental Results

The distances between the three registered tumours and their respective ground-truths were measured for the SVR, SVB, and SVH methods. The per-tumour and average results are reported in Table 4.4.

We observe that method SVH outperforms for all three tumours and both visibility levels. Methods SVR and SVB compete for the second best performance, depending on the tumour, though SVR is overall slightly better.



SVR method			
Tumour ↓   View →	1	2	Avg
1	02.46	07.02	<u>04.74</u>
2	02.07	03.38	02.72
3	01.57	02.02	<u>01.79</u>

SVB method			
Tumour ↓   View →	1	2	Avg
1	02.93	06.65	04.79
2	01.39	02.88	<u>02.13</u>
3	03.79	01.49	02.64

SVH method			
Tumour ↓   View →	1	2	Avg
1	01.08	03.65	<b>02.36</b>
2	01.25	02.58	<b>01.91</b>
3	01.19	01.51	<b>01.35</b>

SVR method			
Tumour ↓   View →	1	2	Avg
1	00.82	06.56	<u>03.69</u>
2	02.82	02.91	02.86
3	01.50	01.14	<u>01.32</u>

SVB method			
Tumour ↓   View →	1	2	Avg
1	01.67	05.78	03.72
2	02.32	02.28	<u>02.30</u>
3	01.36	01.86	01.61

SVH method			
Tumour ↓   View →	1	2	Avg
1	01.02	05.58	<b>03.30</b>
2	01.87	02.35	<b>02.11</b>
3	01.31	01.23	<b>01.27</b>

(a)
(b)

Table 4.4: Registration errors for three synthetic inner tumours on an ex-vivo sheep liver using (a) low visibility and (b) regular visibility of the visual cues. The errors are in millimeters and correspond to the average of the deviations of the registered model’s vertices. For each tumour, the best result is in bold and the second best is underlined.

## 4.6 Qualitative Evaluation on Patient Data

### 4.6.1 Validation Pipeline

A high variability of the registration results obtained from different operators is a sign of unreliability of the registration solutions. High reprojection errors in control views reveal a bad registration. We propose to assess both registration variability and reprojection errors in control views on real patient data (figure 4.10).

### 4.6.2 Experimental Results

#### Variability for SVR, SVB and SVH registrations

To compare the variability between all the registration methods, we asked 5 surgeons to perform SVR, SVB and SVH registrations on 7 different patient datasets. Two of the surgeons had more than 20 years of experience in laparoscopy, while the others had less than 5 years. Before performing registration, the surgeons were also provided with short videos acquired during the surgery’s exploratory stage. The laparoscope was inserted in different keyholes to let the surgeons have a wider scene perception. Table 4.5 reports the average of the vertex-to-vertex root-mean-square deviation (RMSD) over the surgeons for SVR, SVB, and SVH registrations. For a patient, the RMSD measures how different the registered shapes are between surgeons. It differs from the standard deviations reported in Table 1 which correspond to the deviations of the registration errors computed from ground truths.

The average variability for all the patients is of 9.1 mm for SVR, 13.2 mm for SVB, and 11.1 mm for SVH. One of the key results is that, while SVH offers a higher flexibility on the model de-

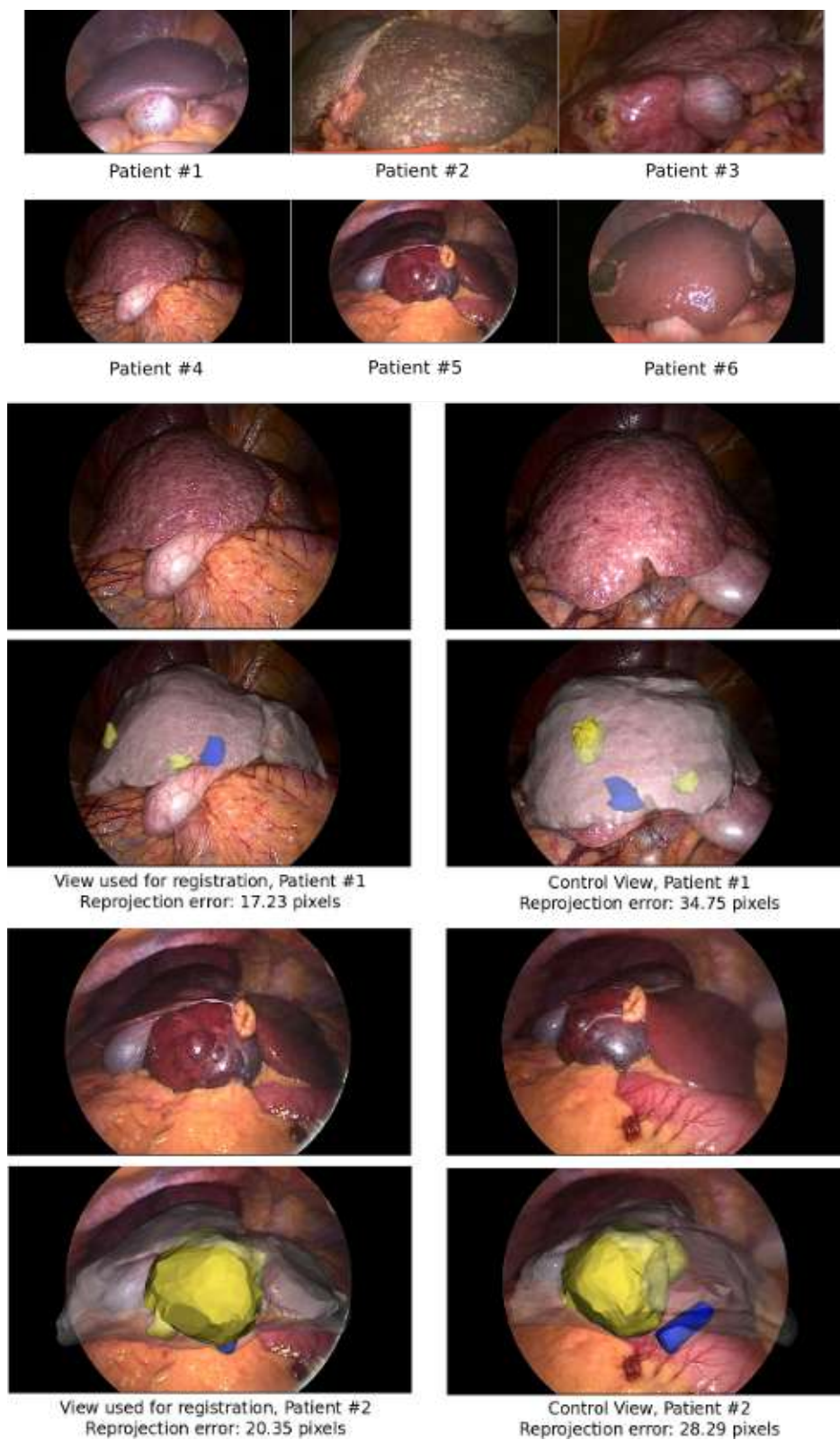


Figure 4.10: Patient images used for control view evaluation, along with two examples of augmented images and their reprojection errors after SVH registration. Laparoscopic images have Full HD resolution (1920 x 1080 pixels). The liver surface mesh is rendered in gray, their subsurface tumours in yellow and veins in blue.

Model	Method ↓   Patient →	1	2	3	4	5	6	7	Avg
Whole liver	SVR	<b>06.50</b>	<b>14.42</b>	<u>17.33</u>	<b>10.38</b>	<b>09.89</b>	<u>11.93</u>	<b>11.53</b>	<b>11.71</b>
	SVB	07.51	20.93	33.78	14.89	<u>12.99</u>	<b>10.51</b>	13.97	16.37
	SVH	<u>06.82</u>	<u>19.39</u>	<b>16.45</b>	<u>13.29</u>	13.23	17.37	<u>12.46</u>	<u>14.14</u>
Visible	SVR	<u>04.88</u>	<b>12.16</b>	<u>09.61</u>	<u>09.57</u>	<b>07.98</b>	<u>09.99</u>	<b>05.28</b>	<b>08.49</b>
	SVB	06.56	18.83	30.57	09.63	<u>08.79</u>	<b>08.80</b>	<u>10.96</u>	13.45
	SVH	<b>04.80</b>	<u>17.76</u>	<b>09.05</b>	<b>06.96</b>	09.58	14.32	11.59	<u>10.58</u>
Tumour	SVR	<b>06.07</b>	<b>12.09</b>	<u>08.71</u>	<b>09.52</b>	<b>07.05</b>	<u>14.86</u>	<b>05.26</b>	<b>09.08</b>
	SVB	07.41	18.33	29.87	11.60	<u>07.52</u>	<b>10.29</b>	<u>07.08</u>	13.16
	SVH	<u>06.59</u>	<u>17.60</u>	<b>07.69</b>	<u>10.79</u>	11.05	16.83	07.21	<u>11.11</u>

Table 4.5: Registration variability (in mm) over the surgeons on 7 in vivo datasets. Best results are in bold and second best in underline.

formation than for SVB, the overall registration variability remains lower. SVR shows the lowest variability because it has very little flexibility.

### Control View

From the registrations made by the surgeons on the in-vivo data, we selected 6 patients for which we had views of the liver acquired from a different optical keyhole. We measured the 2D reprojection errors as the distance from occluding contours manually extracted from the laparoscopic images to the boundaries of the registered model’s reprojections. It was performed for both the reference and additional views. Tables 4.6 and 4.7 report the 2D reprojection errors in pixels for the reference and the additional views respectively. Table 4.8 reports the average of the reprojection errors for both views.

Patient →	1	2	3	4	5	6	Avg
SVR	32.22	34.12	34.68	30.04	18.63	37.99	31.28
SVB	<b>15.16</b>	<u>25.67</u>	<u>22.99</u>	<u>15.19</u>	<b>09.20</b>	<u>09.47</u>	<u>16.28</u>
SVH	<u>17.23</u>	<b>20.35</b>	<b>13.20</b>	<b>14.97</b>	<u>09.83</u>	<b>08.40</b>	<b>14.00</b>

Table 4.6: Reprojection error (in pixels) of the in-vivo patient data in the reference view.

Patient →	1	2	3	4	5	6	Avg
SVR	<u>33.09</u>	31.69	<u>53.01</u>	<b>37.35</b>	17.40	44.87	36.24
SVB	<b>24.07</b>	<b>26.07</b>	58.40	42.78	<u>17.13</u>	<u>26.71</u>	<u>32.53</u>
SVH	34.75	<u>28.29</u>	<b>42.19</b>	<u>37.57</u>	<b>15.58</b>	<b>21.95</b>	<b>30.05</b>

Table 4.7: Reprojection error (in pixels) of the in-vivo patient data in the control view.

The average reprojection errors are much higher for SVR than for SVB and SVH, while SVH has the lowest values. The rigid model in SVR cannot be correctly aligned to fit the imaged liver.

Patient →	1	2	3	4	5	6	Avg
SVR	32.66	32.90	43.85	33.69	18.02	41.43	33.76
SVB	<b>19.62</b>	<u>25.87</u>	<u>40.69</u>	<u>28.98</u>	<u>13.17</u>	<u>18.09</u>	<u>24.40</u>
SVH	<u>25.99</u>	<b>24.32</b>	<b>27.69</b>	<b>26.27</b>	<b>12.70</b>	<b>15.17</b>	<b>22.02</b>

Table 4.8: Reprojection error average (in pixels) of the in-vivo patient data from both reference and control views.

### Registration Time

The total setup time was 5'56" on average in our experiments. This time can be split between time requiring the surgeon's attention (understanding the scene, marking the landmarks and performing SVH registration), which is 4'05", with standard deviation 0'38", and time not requiring the surgeon attention (for the software to initialise the system and compute SVB registration), which is 1'05" on average. It is worth noting that, once the surgeon has understood the scene and made the first registration, subsequent registrations on the same patient will take less time.

## 4.7 Conclusions

The obtained registration errors confirm that the proposed method improves the registration accuracy, thanks to the refinement of the non-visible areas of the liver in the laparoscopic image. The user can expect a similar or lower range of error nearby the tumour area, an error which is below the 1 cm oncologic margin advised in the literature for tumour resection in laparoscopic hepatectomy. The low variability obtained from our method suggests that surgeons have a similar interpretation of the scene and were provided with an appropriate tool to edit the model shape accordingly. The lowest variability shown by SVR can be explained by the limited control on the model compared to SVB and SVH, namely restricted to the model's rigid pose.

The time spent by surgeons to perform registration represents a very small portion of the total surgery time. Nonetheless, automating the detection of the landmarks could drastically decrease the manual interaction required from the surgeon, reducing the total registration time and thus improving usability. The problem of landmark detection in the laparoscopic image could be tackled within the framework of deep neural networks. However, it is a difficult problem which to date remains open. Contrarily to organ detection and segmentation, for which recent techniques show compelling results, landmark detection would require the machine to detect curves (which are more difficult to represent than regions in a deep neural network) and to classify them in a type related to their semantics (lying on or off the liver) and geometric properties (being part of the silhouette, for instance). This problem is still open in the computer vision and medical image processing literature.

The amount of visible liver also plays an important role in the registration, as shown in Table 4.2. The lack of visibility affects greatly SVR, while SVB and SVH have better and consistent errors regardless the FoV size. This indicates that in such cases both SVB and SVH are able to recover the shape of the hidden parts successfully. In general, we see an increase in the registration error for a higher number of vertices/tetrahedrons in the preoperative 3D model, as seen in Table 4.3. Nevertheless this does not always hold for each individual dataset, which means that factors

such as the viewpoint and the liver shape play a more important role in the registration than the number of tetrahedrons in the preoperative 3D model. Preliminary results on the ex-vivo experiments show that our method is able to accurately recover the location of the inner tumours for a varying visibility degree of visual cues, even if they are far from any visual constraint and regardless the viewpoint used for registration, as shown in Table 4.4. The two levels of visibility bring an interesting observation: the stronger the visibility, the smaller the differences between the methods. Specifically, we observe that SVH brings a substantial improvement when visibility decreases. This is a sensible result, because when visibility decreases, the added value of the surgeon expertise expressed by their interactions increases, maintaining the performance, while SVR and SVB may only worsen.

The registration performance remains correlated to the technical difficulties inherent to laparoscopic surgery, such as a reduced FoV and limited viewpoint range, which may substantially vary with patient anatomy. For example, the registration of a laparoscopic image where the liver is entirely visible and whose anatomical landmarks can be accurately localised (such as patient #5 in Table 4.6) is more accurate than one performed on an image where the liver is a partly visible and whose landmarks localisation is ambiguous (such as patient #3 in Table 4.6).

Our approach works on static laparoscopic images, which represent weak inputs, but nonetheless captures the effects of respiration, diaphragm interactions and pneumoperitoneum via the extracted visual cues. In other words, the visual cues inherently represent these complex constraints, which are not capturable otherwise in the routine surgical context of the problem at hand. The strength of our approach is to complement these visual cues which are also weak constraints, by surgeon interactions. This allows our system to take advantage of the observable landmarks from the input laparoscopic image (via the visual cues) and of the surgeon's expertise and understanding of the intraoperative scene (via their interactions). Our results confirm that combining a biomechanical model constrained by visual cues and manual interactions is very fruitful.

As future work, the registration method will be extended to let the surgeon choose the number of control points in the generated cage, according to the complexity of the liver's shape. The influence of using a more advanced biomechanical model on the performance should also be evaluated. Further clinical tests have to be made in order to validate our method, notably regarding the location of inner structures after registration on human cases. If such tests confirm an overall registration error lower than 1 cm, then the proposed method will give surgeons a reliable basis to guide resection.

## Chapter 5

# Multi-View Registration Techniques

---

We propose in this chapter several multi-view registration methods that also aim to solve the partial visibility issue in LLR. However, compared to the SVH method from chapter 4, they combine the visual constraints from the multiple views and perform registration automatically. The proposed methods aim to work in cases where the liver does not deform noticeably between views, and when it does. We call them the rigidly linked and the non-rigidly linked cases. For the rigidly-linked case, we propose the Multi-View Rigid Base (MV-B), the Variant of Multi-View Rigid Base (MV-B\*), and the Multi-View Rigid with Correspondences (MV-C) methods. They differ on the type of constraints and the registration strategy they use. For the non-rigidly linked case, we propose the Multi-View Deformable with Correspondences (MV-D) method. We evaluate these methods on phantom and patient data, comparing the registration accuracy between them and against a single-view method.

---

## 5.1 Introduction

The big dimensions of the liver and the proximity of the camera produces partial views of the liver in the laparoscopic images. This reduces landmark visibility and negatively impacts the registration accuracy in single-view registration methods (Koo et al., 2017; Adagolodjo et al., 2017; Espinel et al., 2020). While the SVH method presented in chapter 4 is able to refine the non-visible parts and improve registration accuracy, it heavily depends on the surgeon’s understanding of the intraoperative scene. Regarding the existing multi-view approaches, they usually reconstruct an intraoperative 3D shape and use it as a target for registration. This reconstruction is usually done through SfM or SLAM, as shown in chapter 2. However, getting a reliable intraoperative shape is difficult and requires very good imaging conditions. Consequently, there is a need for monocular registration methods that are compatible with the clinical constraints and the desired clinical outcomes, such as improved positive resection margins and reduced adverse effects to the patient (Thompson et al., 2011).

We propose several deformable registration methods that use multiple views simultaneously from standard monocular endoscopic systems. They bring the preoperative 3D models of the liver and its internal structures to camera coordinates, with the right shape deformation state. After registration is done, the 3D models of the inner structures are projected on and composed with the laparoscopic views. Our methods may assume that the liver does not deform between the laparoscopic views, which is the *rigidly linked view case*, or that it may deform, which is the *non-rigidly linked view case*, as illustrated by figure 5.1. We assume the views show the liver from multiple viewpoints. Similarly to SVH, our multi-view methods are also based on the SVB method (Koo et al., 2017). For the rigidly linked view case we propose three methods. The first one, named *Multi-View Rigid Base (MV-B)*, uses the liver landmarks and the inter-image rigidity as constraints for registration. The second one, *Variant of Multi-View Rigid Base (MV-B\*)*, uses the same constraints as MV-B but, instead of computing the camera poses through ICP, it does so with SfM. The third one, *Multi-View Rigid with Correspondences (MV-C)*, uses available texture information via inter-image keypoint correspondences to constrain registration. These three methods produce a single deformed shape as output. For the non-rigidly linked view case, we propose one method. It differs from the rigidly linked ones in that it does not use the inter-image rigidity to solve registration. This method is named *Multi-View Deformable with Correspondences (MV-D)* and uses the landmarks and keypoint correspondences as constraints, producing several deformed shapes according to the number of views. We illustrate and classify our methods, along with the constraints they use in figure 5.2.

We evaluated our methods quantitatively on phantom data and qualitatively on patient data. For the quantitative evaluation, TRE was measured on uniformly-distributed control points inside the liver models. For the qualitative evaluation, the 2D reprojection errors of the registered livers were measured on control views. Our goal is to have a TRE lower than 1 cm, which is the safety resection margin advised for resection of HCC (Zhong et al., 2017). We do not measure TRE on patient data due to the difficulty of having a reliable groundtruth with the available devices in the surgery room. We found that our multi-view methods are able to improve registration accuracy compared to the single-view case, being a step forward in AR for monocular laparoscopy of the liver.

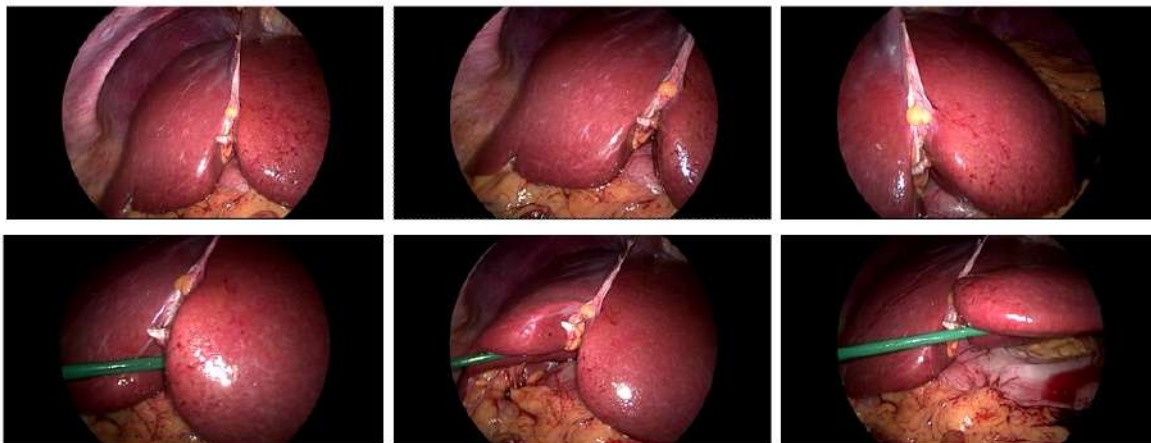


Figure 5.1: Rigidly linked (top) and non-rigidly linked (bottom) laparoscopic views of the liver. The liver does not undergo noticeable deformations between the rigidly linked views and may do between the non-rigidly linked views.

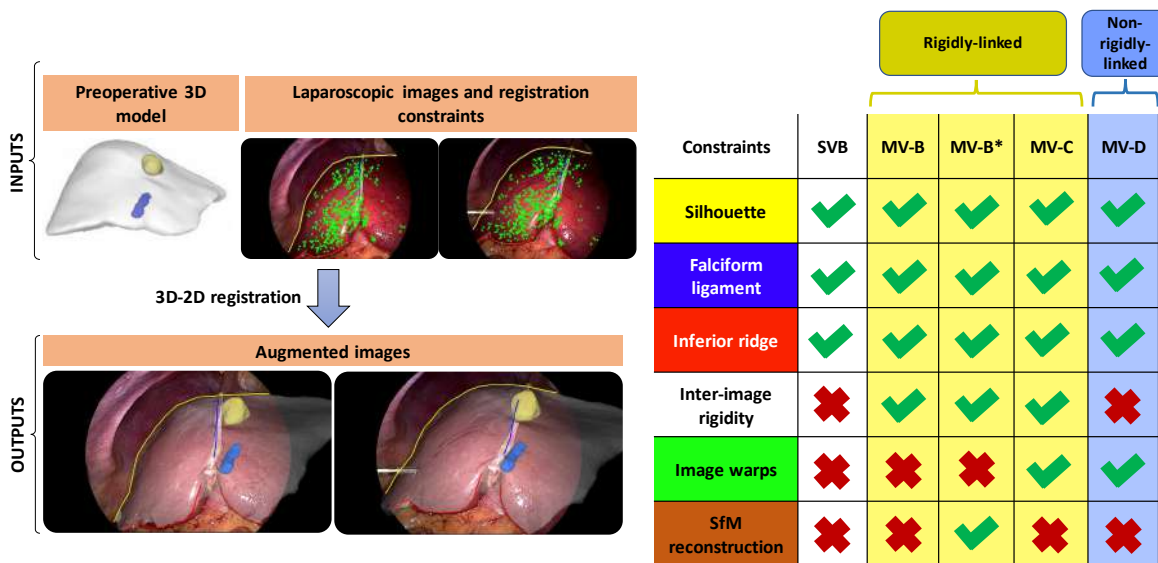


Figure 5.2: Characteristics of the state-of-the-art and proposed registration methods.



## 5.2 Multi-View Registration of Rigidly Linked Images

### 5.2.1 Multi-View Rigid Base (MV-B)

MV-B extends SVB to solve registration on  $N$  rigidly linked laparoscopic views. At each iteration, an SVB refinement is done on every view and then, an average model is computed from all the views. To obtain this average model, all the individual models from views  $N > 1$  are transformed into the first view using transformation matrices computed through ABSOR. Then, the average model is transformed back into all the views and a rigid ICP refinement is done in every view. The rigidity constraint comes thus from the averaging step, which produces a single deformed model.

The pipeline for MV-B is shown in algorithm 5.1. The main function **MVBRegistration** takes as inputs the preoperative 3D model  $M_0$  and the sets of 2D and 3D contours  $\mathcal{C} = \{C_1, \dots, C_N\}$  for every view  $I$ . A main loop performs SVB registration on every view using the **SVBRegistration** algorithm, computes the average model  $M$  from all the views using the **ComputeAverageModel** function, and rigidly registers  $M$  in every view using ICP through the **RigidICPRegistration** function. **ComputeAverageModel** first transforms all the models  $\{L_2, \dots, L_N\}$  into the first view using ABSOR and then proceeds to compute the average model  $M$ . The main loop iterates until the difference between the models of two consecutive iterations becomes lower than a threshold  $\epsilon$ , or until a maximum number of iterations `max_iter` has passed. We have set  $\epsilon = 10^{-3}$  and `max_iter` = 100 after running our methods for a large number of iterations on multiple phantom and patient cases.

---

**Algorithm 5.1:** MV-B registration.

---

```

1 Function MVBRegistration
   Input:  $\mathcal{I} = \{I_1, \dots, I_N\}, M_0, \mathcal{C} = \{C_1, \dots, C_N\}$  // Laparoscopic images, Preoperative 3D model, Set
   of 2D-3D contours
   Output:  $M$  // Registered model
2   max_iter  $\leftarrow$  100 // Maximum number of iterations
3    $\epsilon \leftarrow 10^{-3}$  // Precision threshold
4    $M_i \leftarrow M_0$  // Initialise model
5    $i \leftarrow 1$  // Current iteration
6   repeat
7      $n \leftarrow 1$  // View counter
8     repeat
9        $L_n \leftarrow$  SVBRegistration( $M_i, C_n$ ) // Perform single-view registration
10      until  $n++ \geq \text{size}(\mathcal{I})$ 
11       $M_{i+1} \leftarrow$  ComputeAverageModel( $\{L_1, \dots, L_N\}$ ) // Average model from all views
12       $n \leftarrow 1$  // View counter
13      repeat
14         $L_n \leftarrow$  RigidICPRegistration( $M_{i+1}, C_n$ ) // Perform rigid ICP registration
15        until  $n++ \geq \text{size}(\mathcal{I})$ 
16      until  $\text{dist}(M_{i+1}, M_i) \leq \epsilon$  or  $i++ \geq \text{max\_iter}$ 

```

---

### 5.2.2 Variant of Multi-View Rigid Base (MV-B\*)

MV-B\* is a variant of MV-B that, instead of computing inter-image rigid transformations using ABSOR, which may be inaccurate, directly uses the camera poses recovered by SfM. As such, SfM is firstly run on the keypoint correspondences extracted from SIFT, from which we obtain an intraoperative pointcloud  $Q$  along with the camera poses. However, these SfM reconstructions are up to a scale factor  $l$ , which should be recovered in a precise way during the registration process. Obtaining a precise  $l$  can be a difficult task especially if we want to keep the process fully automatic. To achieve this, we rely on anisometry to find a metric that can indicate when a registration

is done using the right  $l$ . Anisometry is a measure of shape variation that reflects to which extent a shape has stretched or shrunk locally, by integrating the local measures to form a global statistic. An anisometry measure vanishes if and only if the pair of shapes it is measured on are identical. Therefore, it is sensitive to scale changes, reaching its minimum value at the right scale. There are several ways of measuring anisometry on a discrete mesh model. The simplest way is to measure the Euclidean or geodesic distance between pairs of points chosen as the mesh vertices. We use three types of vertices to measure anisometry: the vertices defining the ridge landmark, the vertices forming the liver surface, and the vertices within the liver inner volume. We search for a suitable metric using both types of Euclidean and geodesic distances on the liver ridge, the liver surface and the liver inner volume, as shown in figure 5.3. The near anisometry measures are computed from connected pairs of vertex points and the far anisometry measures are computed from non-connected pairs of vertex points. The vertex pairs are selected randomly and the mean and median anisometry from all pairs are then used. Concretely, we have compared a total of 24 anisometry measures on the phantom data (euclidean vs geodesic, ridge vs surface vs volume, near vs far, mean vs median). To measure anisometry, we first look for an interval around a correct  $l$  for which MV-B\* has a better registration accuracy than SVB. Then, we run MV-B\* by varying  $l$  around the previously defined interval and measure the aforementioned anisometry measures between the preoperative and the registered models. We did these experiments on  $B = 2$  different phantoms, using  $G = 10$  combinations of views per phantom. Finally, we compute a global mean anisometry  $A$  from all the phantoms and all the view combinations, as:

$$A = \frac{\sum_{b=1}^B \delta_b}{B}. \quad (5.1)$$

We use  $\delta_b$ , the mean of the anisometry measures for phantom  $b$ , computed as:

$$\delta_b = \frac{\sum_{g=1}^G \alpha_g}{G}. \quad (5.2)$$

We use  $\alpha_g$  as either the mean or median of the anisometry measures for view combination  $g$ , computed as:

$$\alpha_{\text{mean}} = \frac{\sum_{h=1}^H |D_{F_h} - D_{I_h}|}{H} \quad (5.3)$$

$$\alpha_{\text{median}} = \text{med}_{h \in [1;H]} (|D_{F_h} - D_{I_h}|), \quad (5.4)$$

where  $H$  is the number of vertex pairs and  $\{D_{I_h}, D_{F_h}\}$  are the initial preoperative and final registered distances for the point pair  $h$ . These distances can be either euclidean or geodesic.

After running MV-B\* on the 2 phantoms by varying  $l$  within  $\pm 12\%$  around the ground-truth, we found that the Euclidean near volume anisometry shows the desired convex behaviour with its minimum value at the ground-truth. The behaviour of all the 24 anisometry measures is shown in figure 5.4. Using this metric, we run MV-B\* starting from an initial  $s$  and look for the minimum anisometry through a bisection search strategy. In the end, the resulting deformed model  $M$  is the product of a registration using the right  $l$ .

The pipeline for MV-B\* is shown in algorithm 5.2. The main function **MV-B\*Registration** uses bisection search to search for a local minimum in anisometry and perform registration with the correct scale  $l$ . It starts by doing single-view registration using the **SVBRegistration** function to obtain an initial middle scale  $l_M$ . This scale is computed as the depth ratio of the closest points between the registered 3D model  $M$  and the SfM pointcloud  $Q$ . From the initial  $l_M$ , a lower and

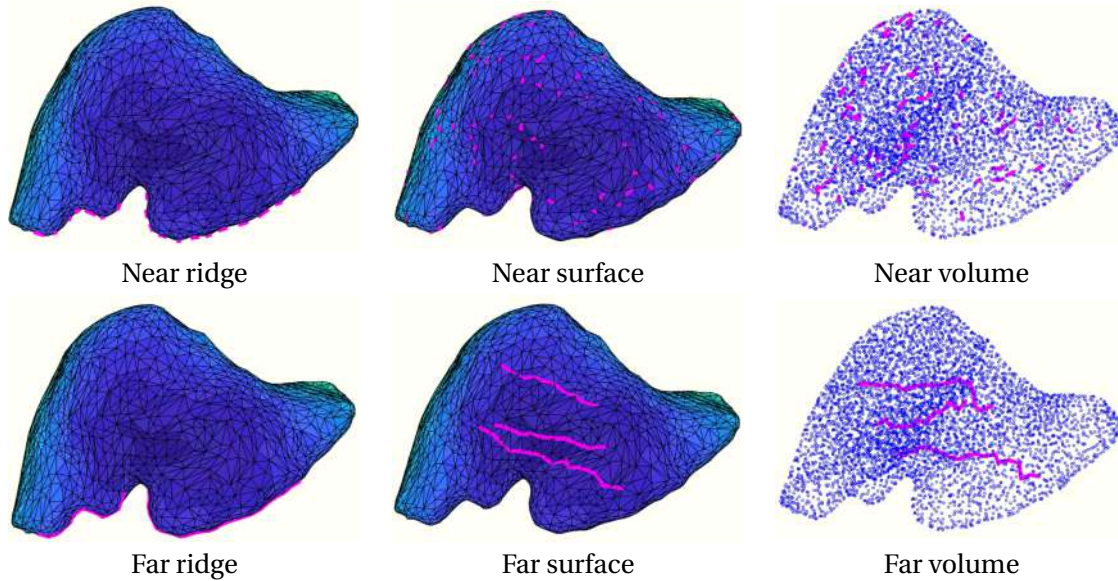


Figure 5.3: Illustration of the anisotropy measures on the liver 3D model. The near anisotropy is measured in contiguous pairs of points. The far anisotropy is measured in non-contiguous pairs of points.

upper pair of interval bounds  $\{l_L, l_U\}$  is set as  $\pm 30\%$  of  $l_M$ . Then, multi-view registration is done on  $\{l_L, l_U\}$  by scaling the SfM camera poses  $\mathcal{T}$  and calling the **DeformableRegistration** function. The respective anisotropies  $D_L, D_U$  are then computed. A main loop then compares  $D_L$  against  $D_U$  and selects the sub-interval  $l_L, l_M$  or  $l_M, l_U$  that has the lowest anisotropy. The new middle scale  $l_M$  is thus computed, multi-view registration is run on  $l_M$  and the new interval bound  $D_L$  or  $D_U$  will correspond to the anisotropy of  $l_M$  depending on the chosen sub-interval. The loop iterates until the difference between the lower and upper bounds becomes lower than  $\epsilon$  and deems the scale to be found. The resulting model  $M$  is thus the one obtained using the last middle scale  $l_M$ .

### 5.2.3 Multi-View Rigid with Correspondences (MV-C)

MV-C is an extension of MV-B that uses the inter-image keypoint correspondences to constrain registration. This strengthens the contour correspondences with the more reliable keypoint correspondences in the ICP refinement stage. The keypoint correspondences are obtained using SIFT and the mismatches are filtered out using Feature-Based Deformable Surface Detection (FBDS) (Pizarro and Bartoli, 2010). They are used to compute inter-image warp transformation functions  $\eta_{mn}$ , which are based on the Rigid-Perspective Thin-Plate Spline (Bartoli et al., 2010). The image warp  $\eta_{mn}$  allows one to transfer any point from view  $m$  to view  $n$ , as shown in figure 5.5. It is used to iteratively minimise the Euclidean distance between the projected particles, transferred to a common reference view  $i$ . This allows one to estimate an average, optimal camera pose, by ensuring that the  $U$  model particles  $x_u$ ,  $u \in [1, U]$ , are projected to the same location of the liver surface in all the views. First, each visible model particle  $x_u$  in the reference view  $n$  is projected in every view  $m$  wherever it is visible. Second, these 2D particle projections are transferred from all views  $m$  into the reference view  $n$  using  $\eta_{mn}$ . Third, a barycentre point  $\beta$  is computed in view  $n$  as the average of the transferred points. Fourth,  $\beta$  is used as the 2D correspondence of the model point  $x_u$  to refine camera pose. Fifth, the four steps are repeated, choosing each view as the reference view in turn, so as to refine its camera pose. In the end, MV-C generates a single deformed

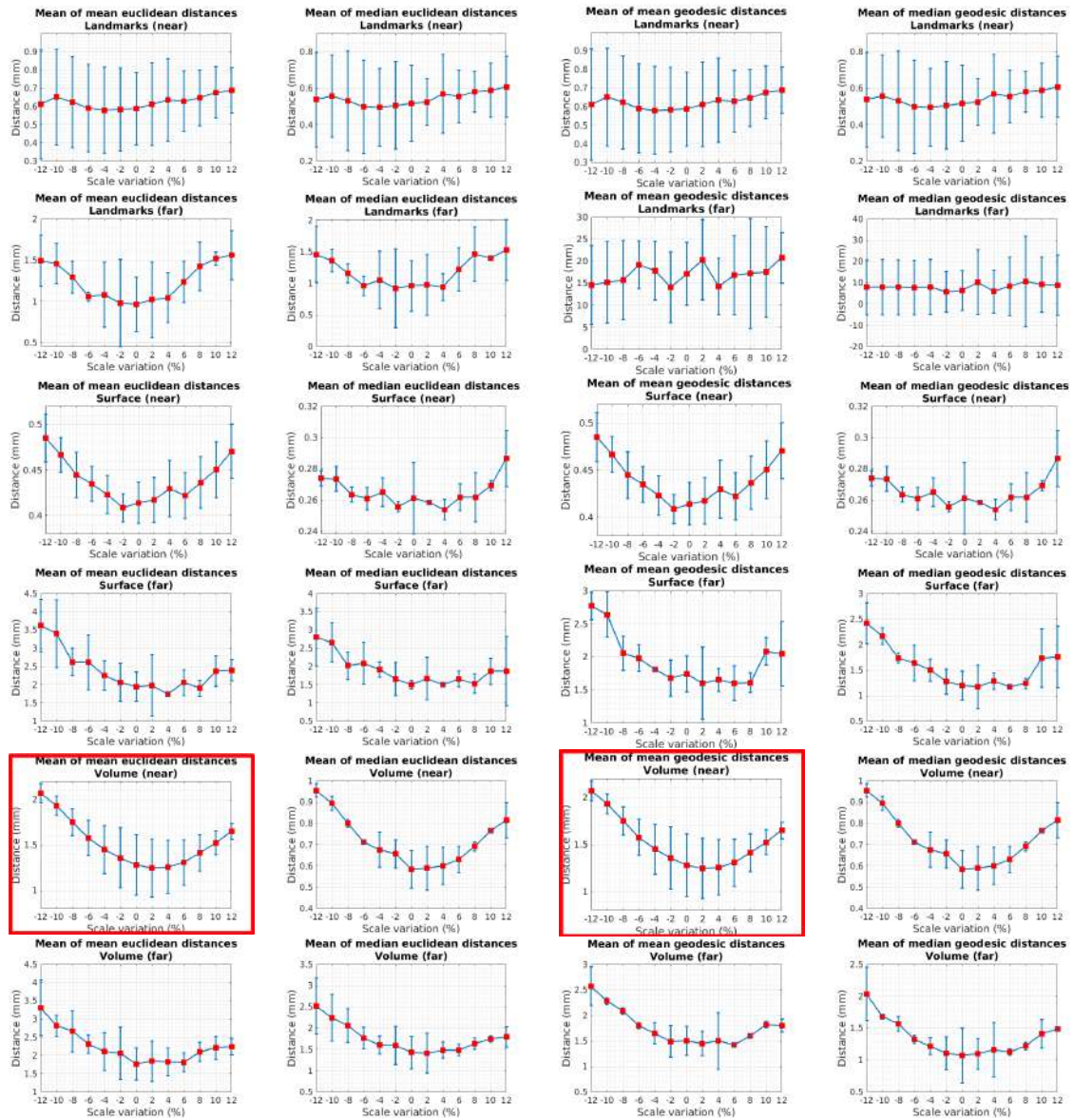


Figure 5.4: Behaviours of the 24 anisometry measures for 2 phantoms, using 10 combinations of 8 views per phantom. Two anisometry measures have a convex behaviour: They are highlighted in red and correspond to the near volume mean euclidean anisometry, and the near volume mean geodesic anisometry.

**Algorithm 5.2:** MV-B\* registration.

---

```

1 Function MVRegistration
   Input:  $\mathcal{I}, M, \mathcal{C}, \mathcal{T}, \text{max\_iter}$  // Laparoscopic images, Preoperative 3D model, Set of 2D-3D
   contours, SfM camera poses, Maximum number of iterations
   Output:  $M$  // Registered model
2    $i \leftarrow 1$  // Current iteration
3   repeat
4      $M \leftarrow \text{SolveBiomechanicalConstraints}(M)$  // Biomechanical optimisation
5      $n \leftarrow 1$  // View counter
6     repeat
7        $L_n \leftarrow \text{SolveContourConstraints}(M, C_n, T_n)$  // Contour landmarks optimisation
8     until  $n++ \geq \text{size}(\mathcal{I})$ 
9      $M_{i+1} \leftarrow \text{ComputeAverageModel}(\{L_1, \dots, L_N\}, \mathcal{T})$ 
10  until  $i++ \geq \text{max\_iter}$ 
11 Function MVB* Registration
   Input:  $\mathcal{I} = \{I_1, \dots, I_N\}, M_0, \mathcal{C} = \{C_1, \dots, C_N\}, Q, \mathcal{T} = \{T_1, \dots, T_N\}$  // Laparoscopic images,
   Preoperative 3D model, Set of 2D-3D contours, SfM pointcloud, SfM camera poses
   Output:  $M$  // Registered model
12   $\text{max\_iter} \leftarrow 20$  // Maximum number of iterations
13   $\epsilon \leftarrow 0.01$  // Minimum interval size parameter
14   $\text{converged} \leftarrow \text{false}$  // Binary convergence indicator
15   $M \leftarrow \text{SVBRegistration}(M_0, C_1)$  // Perform single-view registration on the first view
16   $l_M \leftarrow \text{computeInitialScale}(M, Q)$  // Middle SfM scale bound
17   $l_L \leftarrow l_M - (0.3l_M)$  // Lower scale bound
18   $l_U \leftarrow l_M + (0.3l_M)$  // Upper scale bound
19   $\mathcal{T}_L \leftarrow \mathcal{T} l_L$  // Scale SfM poses for lower scale
20   $\mathcal{T}_U \leftarrow \mathcal{T} l_U$  // Scale SfM poses for upper scale
21   $M_L \leftarrow \text{MVRegistration}(\mathcal{I}, M_0, \mathcal{C}, \mathcal{T}_L, \text{max\_iter})$  // Multi-view registration on lower bound
22   $M_U \leftarrow \text{MVRegistration}(\mathcal{I}, M_0, \mathcal{C}, \mathcal{T}_U, \text{max\_iter})$  // Multi-view registration on upper bound
23   $D_L \leftarrow \text{ComputeAnisometry}(M_0, M_L)$  // Compute anisometry for lower bound
24   $D_U \leftarrow \text{ComputeAnisometry}(M_0, M_U)$  // Compute anisometry for upper bound
25  repeat
   // If upper bound has a lower anisometry, choose the sub-interval  $\{D_M, D_U\}$ :
26  if  $D_L \geq D_U$  and  $(l_L - l_M) > \epsilon$  then
27     $D_L \leftarrow D_M$  // Update lower bound anisometry
28     $l_L \leftarrow l_M$  // Update lower bound scale
29     $l_M \leftarrow (l_M + l_U)/2$  // Compute new middle bound scale
   // If upper bound has a lower anisometry, choose the sub-interval  $\{D_L, D_M\}$ :
30  else if  $D_L < D_U$  and  $(l_L - l_M) > \epsilon$  then
31     $D_U \leftarrow D_M$  // Update upper bound anisometry
32     $l_U \leftarrow l_M$  // Update upper bound scale
33     $l_M \leftarrow (l_L + l_M)/2$  // Compute new middle bound scale
   // If the difference between bounds is less than a threshold, we have converged:
34  else
35     $M \leftarrow M_M$  // The resulting model comes from the last middle bound scale
36     $\text{converged} \leftarrow \text{true}$ 
37     $\mathcal{T}_M \leftarrow \mathcal{T} l_M$  // Scale SfM poses
38     $M_M \leftarrow \text{MVRegistration}(\mathcal{I}, M_0, \mathcal{C}, \mathcal{T}_M, \text{max\_iter})$  // Multi-view registration on medium
   bound
39     $D_M \leftarrow \text{ComputeAnisometry}(M_0, M_M)$  // Compute anisometry for medium bound
40  until  $\text{converged}$ 

```

---

model  $M$  from the contour, the inter-image correspondence and the rigidity constraints given by all the views.

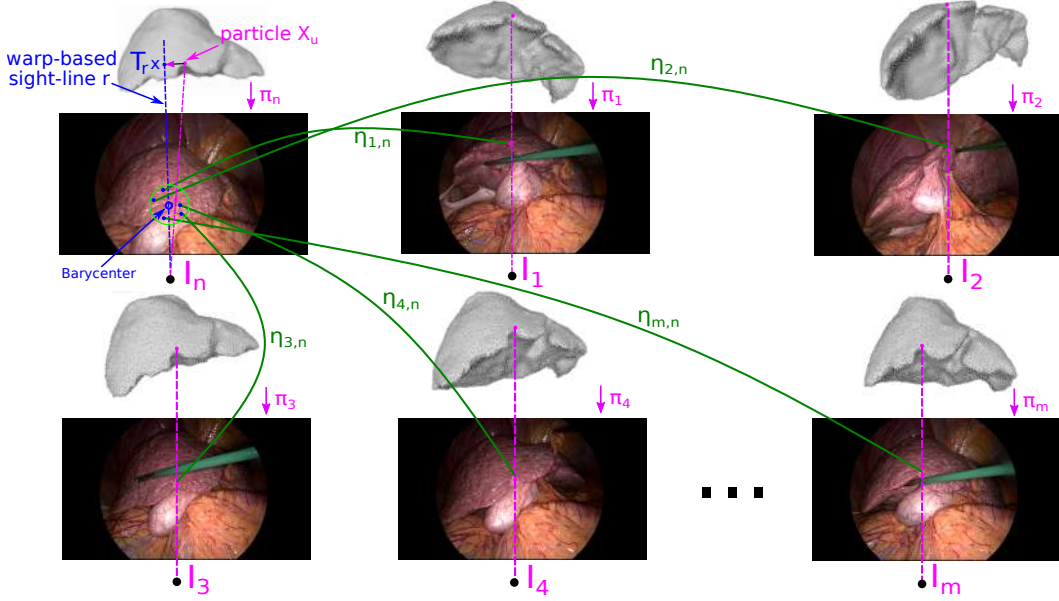


Figure 5.5: Warp-based particle location update. The particle locations are updated as follows. The inter-image warps are precomputed for all view-pairs  $m, n$ . Then, during an iteration of the optimisation algorithm, for every view  $n$ , a particle  $x_u$  is updated such that its new location corresponds to its orthogonal projection onto its warp-based sight line, namely the backprojections of the barycentre of the imaged particles  $\eta_{mn}(\pi_j(x_u))$  through the warps.

The pipeline for MV-C is shown in algorithm 5.3. As a first step, the keypoint correspondences  $\Omega$  are extracted from the views using SIFT and the warp functions  $\eta$  are computed. Then, a main loop progressively deforms  $M$  in four stages. First, SVB registration is done on every view using the **SVBRegistration** algorithm. Second, the average model  $M$  is computed from all the views using the **ComputeAverageModel** function. Third, the inter-image projection distances of the model particles  $x_u$  are minimised by transferring the 2D projections of the particles from view  $m$  to  $n$  using the warp function  $\eta_{mn}$ , and then by computing a set barycentric points  $b_n$  from the transferred particles  $x_{umn}$  using the **BarycentresFromTransferredParticles** function. Fourth, the barycentric points  $\beta_n$  and the 2D-3D contours  $C_n$  are used to rigidly register  $M$  in every view through the **RigidICPRegistration** function. The loop iterates until the difference between the models of two consecutive iterations becomes lower than a threshold  $\epsilon = 10^{-3}$ , or until a maximum number of iterations  $\text{max\_iter} = 100$  has passed.

**Algorithm 5.3:** MV-C registration.

---

```

1 Function MVCRegistration
   Input:  $\mathcal{I} = \{I_1, \dots, I_N\}, M_0, \mathcal{C} = \{C_1, \dots, C_N\}$  // Laparoscopic images, Preoperative 3D model, Set
     of 2D-3D contours
   Output:  $M$  // Registered model
2   max_iter  $\leftarrow 100$  // Maximum number of iterations
3    $\epsilon \leftarrow 10^{-3}$  // Precision threshold
4    $M_i \leftarrow M_0$  // Initialise model
5    $i \leftarrow 1$  // Current iteration
6    $\{\Omega\} \leftarrow \text{DetectKeypointCorrespondences}(\mathcal{I})$  // Detect inter-image correspondences
7    $\{\eta\} \leftarrow \text{ComputeWarps}(\Omega)$  // Compute warp functions
8   repeat
9      $n \leftarrow 1$  // View counter
10    repeat
11       $L_n \leftarrow \text{SVBRegistration}(M_i, C_n)$  // Perform single-view registration
12    until  $n++ \geq \text{size}(\mathcal{I})$ 
13     $M_{i+1} \leftarrow \text{ComputeAverageModel}(\{L_1, \dots, L_N\})$  // Average model from all views
14     $n \leftarrow 1$  // View counter
15    repeat
16       $m \leftarrow 1$  // View counter
17      repeat
18         $x_{um} \leftarrow \text{ProjectVisibleParticlesIn2D}(L_m)$  // Project model particles in 2D
19         $x_{umn} \leftarrow \text{TransferParticlesUsingWarps}(x_{um}, \eta_{mn})$  // Transfer particles to reference
          view  $n$ 
20      until  $m++ \geq \text{size}(\mathcal{I})$ 
21       $x_{un} \leftarrow \text{ProjectVisibleParticlesIn2D}(L_n)$  // Project model particles in 2D
22       $\beta_n \leftarrow \text{BarycentresFromTransferredParticles}(x_{un}, \{x_{u2n}, \dots, x_{uNn}\})$  // Compute
          barycentres from transferred particles
23       $L_n \leftarrow \text{RigidICPRegistration}(M, C_n, x_{un}, \beta_n)$  // Perform rigid ICP registration
24    until  $n++ \geq \text{size}(\mathcal{I})$ 
25  until  $\text{dist}(M_{i+1}, M_i) \leq \epsilon$  or  $i++ \geq \text{max\_iter}$ 

```

---

## 5.3 Multi-View Registration on Non-Rigidly Linked Images

### 5.3.1 Multi-View Deformable with Correspondences (MV-D)

MV-D is a variant of MV-C that does not use the inter-image rigidity to solve registration, but keeps the inter-image correspondences and the contours  $\mathcal{C}$  as constraints for registration. Compared to MV-C, MV-D produces multiple shapes according to the number of views used and the liver deformations in each of them. This is particularly useful in cases where the liver is being manipulated by tools or other external forces.

The pipeline for MV-D is shown in algorithm 5.4. The procedure is similar to MV-C with the difference that no average model is computed from all the views. Instead, the individually deformed models  $\{L_1, \dots, L_N\}$  are rigidly register with ICP in every view using the correspondences between the 2D projections of the visible particles and their corresponding barycentres  $\beta$ , along with the set of contour landmarks  $\mathcal{C}$ . Barycentres  $\beta$  are also obtained after transferring the particles from other views using the warp functions  $\eta$ .

**Algorithm 5.4:** MV-D registration.

---

```

1 Function MVDRegistration
   Input:  $\mathcal{I} = \{I_1, \dots, I_N\}, M_0, \mathcal{C} = \{C_1, \dots, C_N\}$  // Laparoscopic images, Preoperative 3D model, Set
   of 2D-3D contours
   Output:  $\{L_1, \dots, L_N\}$  // Registered models
2   max_iter  $\leftarrow$  100 // Maximum number of iterations
3    $\epsilon \leftarrow 10^{-3}$  // Precision threshold
4    $M_i \leftarrow M_0$  // Initialise model
5    $i \leftarrow 1$  // Current iteration
6    $\{\Omega\} \leftarrow$  DetectKeypointCorrespondences( $\mathcal{I}$ ) // Detect inter-image correspondences
7    $\{\eta\} \leftarrow$  ComputeWarps( $\Omega$ ) // Compute warp functions
8   repeat
9      $n \leftarrow 1$  // View counter
10    repeat
11       $L_n \leftarrow$  SVBRegistration( $M_i, C_n$ ) // Perform single-view registration
12    until  $n++ \geq \text{size}(\mathcal{I})$ 
13     $n \leftarrow 1$  // View counter
14    repeat
15       $m \leftarrow 1$  // View counter
16      repeat
17         $x_{um} \leftarrow$  ProjectVisibleParticlesIn2D( $L_m$ ) // Project model particles in 2D
18         $x_{umn} \leftarrow$  TransferParticlesUsingWarps( $x_{um}, \eta_{mn}$ ) // Transfer particles to reference
           view  $n$ 
19      until  $m++ \geq \text{size}(\mathcal{I})$ 
20       $x_{un} \leftarrow$  ProjectVisibleParticlesIn2D( $L_n$ ) // Project model particles in 2D
21       $\beta_n \leftarrow$  BarycentresFromTransferredParticles( $x_{un}, \{x_{u2n}, \dots, x_{uNn}\}$ ) // Compute
           barycentres from transferred particles
22       $L_n \leftarrow$  RigidICPRegistration( $L_n, C_n, x_{un}, \beta_n$ ) // Perform rigid ICP registration
23    until  $n++ \geq \text{size}(\mathcal{I})$ 
24     $M_{i+1} \leftarrow L_1$ 
25  until  $\text{dist}(M_{i+1}, M_i) \leq \epsilon$  or  $i++ \geq \text{max\_iter}$ 

```

---

## 5.4 Quantitative Evaluation on Phantom Data

### 5.4.1 Rigidly Linked Views

We reconstructed a 3D liver mesh from a patient's CT and generated 10 virtual deformations using Abaqus (Dassault Systèmes, 2022), by simulating gravity, pneumoperitoneum and manipulations with surgical instruments. We added randomly-spaced carved markers in the model's surface to simulate keypoints. We 3D printed these deformations using Polylactic Acid (PLA). Then, we used a surgical laparoscope and a pelvitrainer box to take 10 views of every printed model. Laparoscopic views of the 10 phantoms are shown in figure 5.6. We implemented our methods using Qt and unoptimized C++ code, and they were run on a PC with a processor AMD Ryzen 7 3700x, a graphics card NVIDIA RTX 2080 Ti, 32 GB of RAM, and with Ubuntu 20.04. We estimated registration for every phantom by varying the number of views, going from 1 to 8, and measured TRE as the average prediction error for uniformly sampled points within the virtual 3D liver. The contour marking is done manually in less than 5 minutes for all the 8 views. We use 8 views as a maximum to keep the computation time reasonable. The results are shown in figure 5.7(a). We observe that the TRE for SVB is steady around  $11.89 \pm 7.77$  mm, that for MV-B it varies with the number of views from  $10.21 \pm 5.87$  mm to  $12.08 \pm 12.46$  mm, that for MV-B\* it remains steady around  $10.77 \pm 6.26$  mm, that for MV-C it consistently decreases until  $8.32 \pm 4.58$  mm, and that for MV-D it has a maximum at  $14.14 \pm 7.24$  mm for 2 views and a final value of  $12.58 \pm 6.69$  mm. MV-B\* estimated the SfM scales close to the groundtruth ones, with differences comprising between -13% and +7%. In order to better observe the performance of MV-B\* using a sufficiently good scale, we computed TRE for phantoms with lower than 5% scale variation. This corresponds to 7 out of 10 phantoms, with



performance shown in figure 5.7(b). In this case, MV-B\* has better performance than SVB, with a TRE of  $9.86 \pm 6.36$  mm vs  $10.91 \pm 6.41$  mm. The total running time using 8 views, after marking the contours, was of 2'48" for MV-B, 5'37" for MV-B\*, 4'52" for MV-C, and 4'28" for MV-D.

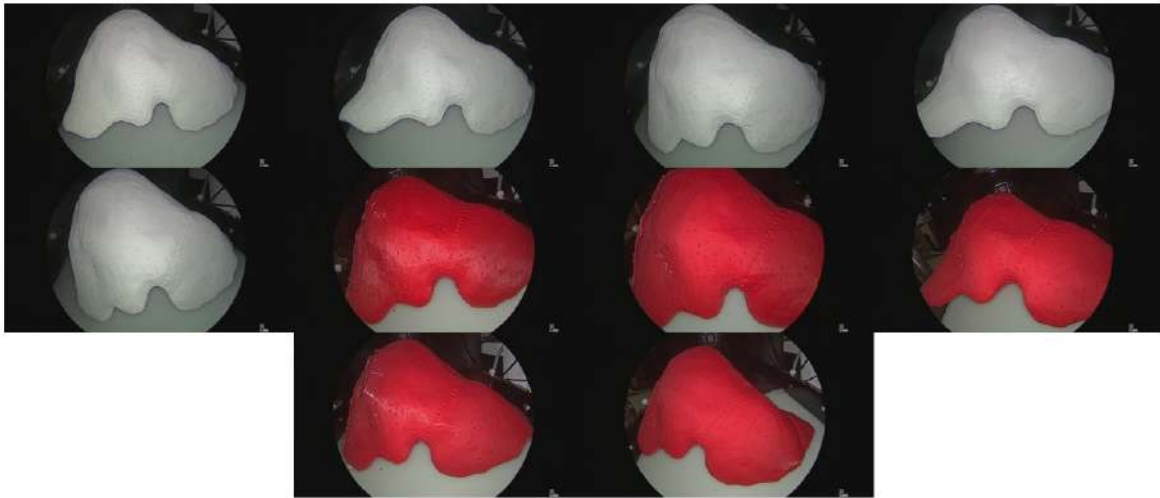
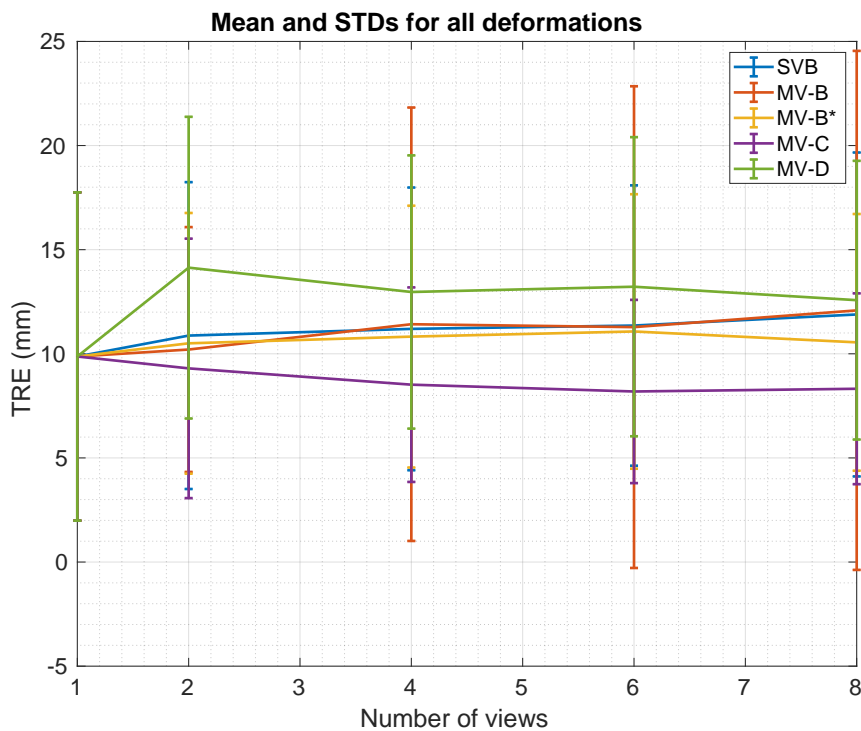


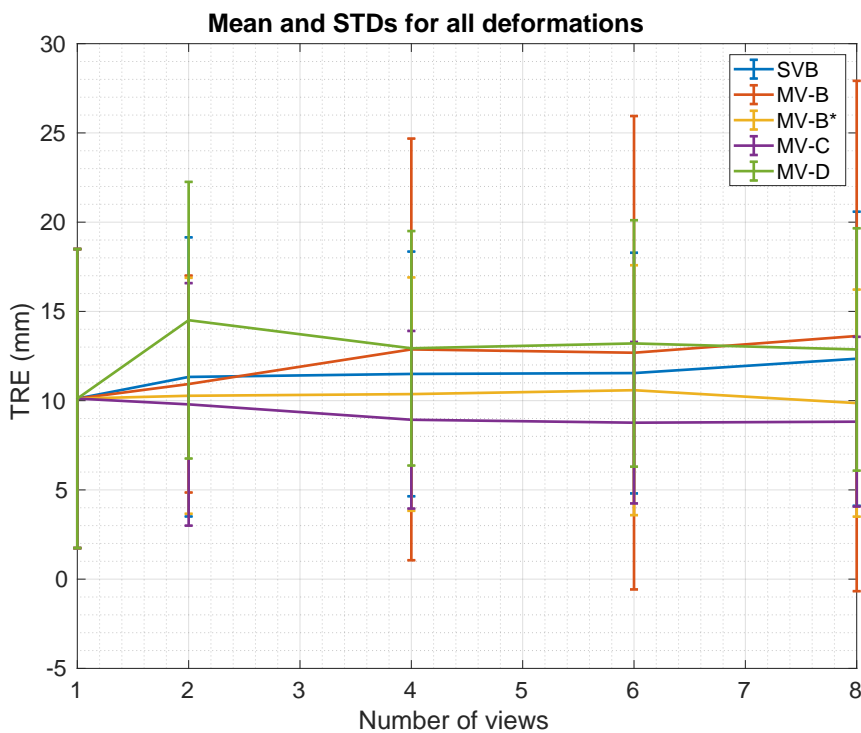
Figure 5.6: Laparoscopic views of the phantoms used in registration. Each view corresponds to a different deformation.

#### 5.4.2 Non-Rigidly Linked Views

We ran all methods on the generated phantom data, using 10 combinations of 8 views, with every view corresponding to a different deformation. In this way, we simulate a sequence of a liver being deformed with the surgical tools. Similarly to the rigidly linked case, experiments are done using an increasing number of views, going from 1 to 8 views for every combination. MV-B\* was excluded from the experiments, as SfM could not be ran in the absence of rigidity. TRE results for a varying number of views are shown in figure 5.8. For the experiments using 8 views, SVB has a TRE of  $11.59 \pm 6.95$  mm on the control points, MV-B has a TRE of  $13.34 \pm 9.51$  mm, MV-C has a TRE of  $11.03 \pm 9.45$  mm, and MV-D has a TRE of  $10.71 \pm 6.35$  mm.



(a)



(b)

Figure 5.7: Mean TRE and standard deviations using all the methods on the rigidly linked views from all the 10 phantoms (a) and on phantoms with estimated scale with less than 5% of deviation (b).

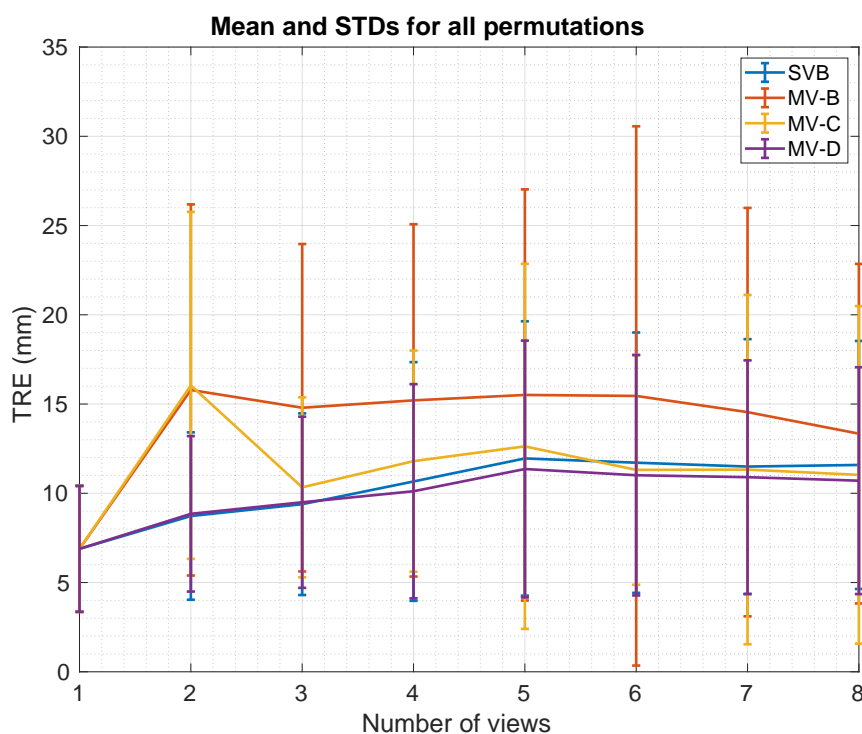


Figure 5.8: Mean TRE and standard deviations using all the methods on the non-rigidly linked views from all the 10 phantoms.

## 5.5 Qualitative Evaluation on Patient Data

### 5.5.1 Rigidly Linked Views

We collected CT and laparoscopic data from 5 patients. We reconstructed the preoperative 3D models from the CT data and selected 9 random representative laparoscopic views for each patient. Out of these views, one was singled out to serve as control view. The control view is not used to perform registration but rather to verify registration, measuring the landmark prediction error expressed in pixels, using the registration computed on other views and rigidly adapted to the control view. Examples of these rigidly linked patient views are shown in figure 5.9. We measure the reprojection errors due to the difficulty of having a reliable groundtruth to compute TRE in 3D. We only run MV-B, MV-C and MV-D for the 8 views. MV-B\* failed to obtain a reliable SfM reconstruction. The results are shown in table 5.1. We observe a clear benefit of using multiple views for registration, especially for the rigidly linked methods.

### 5.5.2 Non-Rigidly Linked Views

We selected 9 random representative views per patient from the previously described data with significant liver deformation across the views. As for the rigidly linked view case, we singled out one view to serve as control view. Examples of such non-rigidly linked patient views are shown in figure 5.10. Landmark prediction error is computed on the control views and expressed in pixels. The results are shown in table 5.2. We ran MV-B, MV-C and MV-D for the 8 views. We observed a 1 mm improvement when using multiple views to solve registration.



Figure 5.9: Example of rigidly linked patient views used in registration. Each row corresponds to a different patient.

Patient	SVB	MV-B	MV-C	MV-D
1	26.07	<u>21.98</u>	<b>19.49</b>	28.61
2	58.40	<u>10.28</u>	<b>09.49</b>	42.73
3	42.78	<u>32.24</u>	<b>29.75</b>	44.15
4	<b>17.13</b>	21.72	<u>21.13</u>	23.79
5	26.71	<b>15.46</b>	<u>16.95</u>	24.28
average	34.28	<u>20.33</u>	<b>19.36</b>	32.71

Table 5.1: Reprojection error average (in pixels) using rigidly linked views of the in-vivo patient data. Best results are in bold and second best are underline.

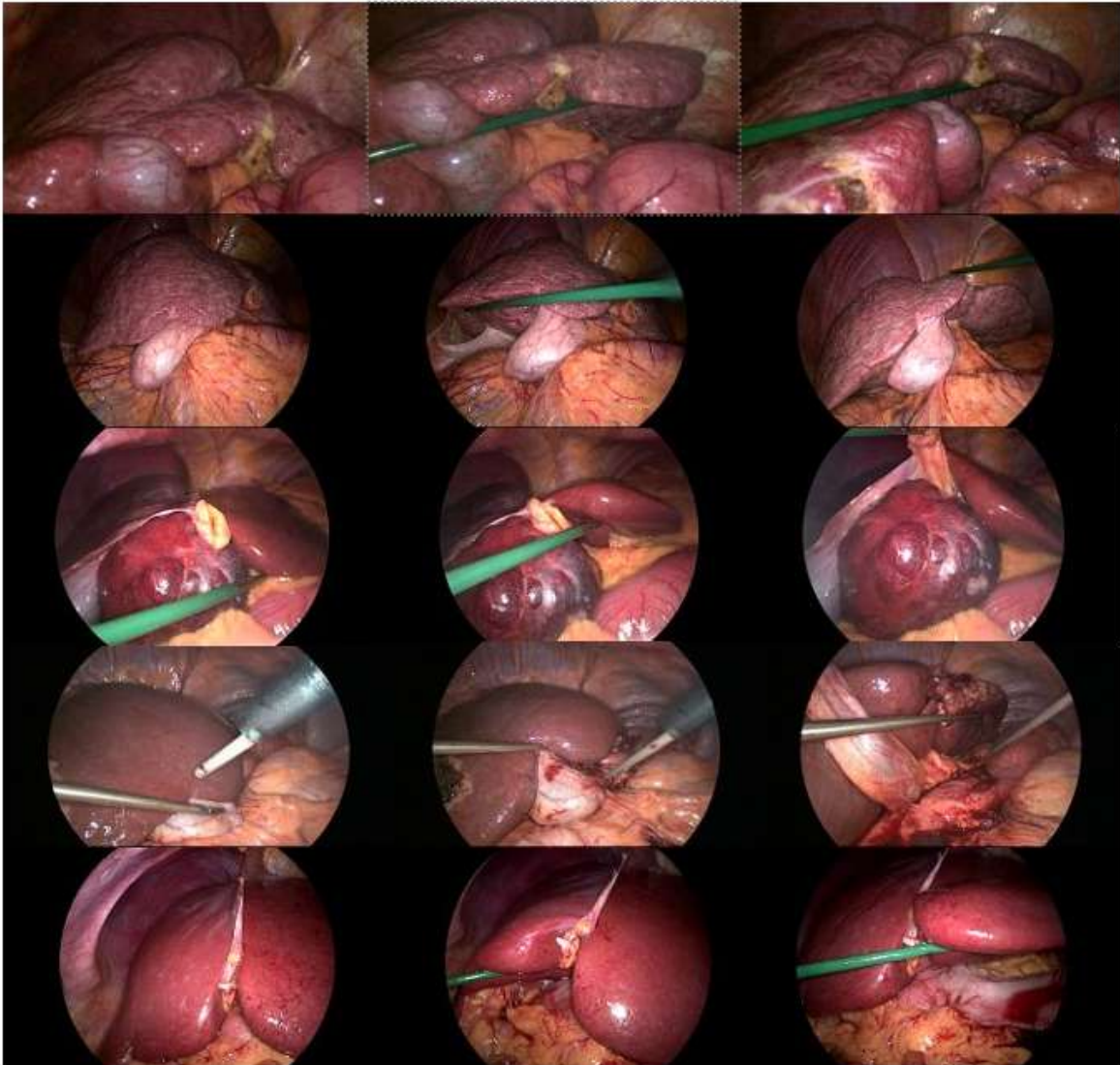


Figure 5.10: Example of non-rigidly linked patient views used in registration. Each row corresponds to a different patient.

Patient	SVB	MV-B	MV-C	MV-D
1	32.47	36.82	<u>30.49</u>	<b>30.14</b>
2	<u>63.08</u>	69.37	66.48	<b>57.62</b>
3	<b>46.24</b>	54.41	49.91	<u>49.79</u>
4	<u>25.07</u>	36.44	29.79	<b>21.31</b>
5	<u>31.85</u>	36.82	32.15	<b>29.75</b>
average	<u>39.74</u>	46.77	41.76	<b>37.73</b>

Table 5.2: Reprojection error average (in pixels) using non-rigidly linked views of the in-vivo patient data. Best results are in bold, second best in underline.

## 5.6 Conclusion

The proposed multi-view methods improve registration accuracy compared to the single-view method, especially for rigidly linked views. The MV-C method is the best performing one in this case, with an improvement of 2.45 mm over SVB on the phantom data. In comparison, MV-D is the worst performing method in the rigidly linked case. These results were confirmed on the patient data, where MV-C improves the reprojection error by 15 pixels over SVB. Its performance also varies with the number of views, although it becomes steady beyond 6 views. The MV-B\* method also works better than SVB if a sufficiently good scale is found. However, it is difficult to obtain a reliable SfM reconstruction on patient data, which limits the usage of this method. When using non-rigidly linked views, MV-D behaves similarly to SVB and MV-C, with a slight improvement of 1 mm.

By using the proposed rigidly linked methods, on top of improving the registration accuracy, the surgeon can perform resection planning while the liver is in rest position. This can be done by looking at several of the augmented images used for registration, which helps to better understand the location of the tumour. Regarding the non-rigidly linked case, it remains a challenging task to improve registration by using multiple images of a deforming liver. Although it is possible to find common structures between the images such as keypoint features and anatomical landmarks, it is still an open problem to find a way on how to use them to better constrain registration. This is especially important for using AR while the liver is being manipulated with the surgical tools.

As future work, we plan to work on *(i)* improving registration accuracy in the non-rigidly linked view case, *(ii)* reducing the registration time by automatically detecting the anatomical landmarks, and *(iii)* using additional information to improve in both the rigidly and non-rigidly linked view cases, such as probing liver points with the tip of the surgical tools. As part of the second goal, we had run the Preoperative to Intraoperative Laparoscopic Fusion challenge (P2ILF) (Ali et al., 2022) in MICCAI 2022, with the objective of automatically detecting the landmark correspondences between the laparoscopic views and the preoperative 3D models. We plan to publish an article with the methods proposed by the participants.



## Chapter 6

# Keyhole-Aware AR with Tumour-Projection Visualisation

---

To improve visualisation in AR, we present a keyhole-aware method that projects the shape of a tumour towards a particular keyhole. The main goal is to give the surgeon an indication of how to reach a tumour using a particular tool, instead of relying solely on camera-based AR. It does so by automatically estimating the keyholes through which the tools are inserted. The keyholes are estimated by first computing the 3D poses of the visible tools in the images, and then by clustering the set of 3D tool poses. We evaluate the robustness of our method with respect to the number of tool poses, as well as noise in the tool poses. We do this on a virtual geometric model that simulates a typical LLR setup. We also evaluate the tumour pointing error while using the camera-based AR and the keyhole-aware AR. We measure this on the virtual geometric model, as well as on a realistic phantom setup. Finally, we perform a preliminary test on one patient case.

---



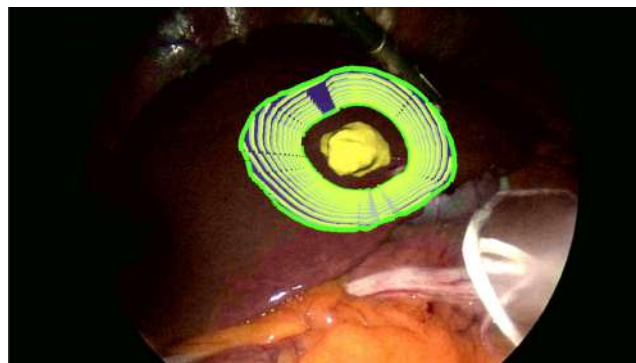
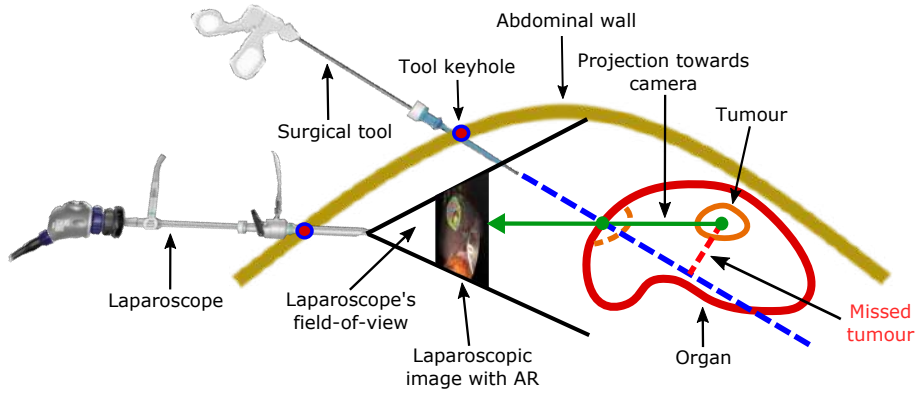
## 6.1 Introduction

Visualisation in AR for LLR is a critical step, as the locations of the internal structures should be accurately represented to enable a better gesture guidance. Surprisingly, the literature dealing with the visualisation step is relatively sparse. Some methods create visual effects to convey information such as oncologic tumour boundaries (Espinell et al., 2020) and a depth scale (Özgür et al., 2017) but most methods just blend the registered structures with the image (Thompson et al., 2015; Bernhardt et al., 2016; Haouchine et al., 2016; Plantefève et al., 2016; Clements et al., 2017; Chen et al., 2018; Robu et al., 2018; Koo et al., 2017; Adagolodjo et al., 2017).

The key common point of such methods is to directly project the internal structures towards the laparoscope. This raises a critical problem when aiming at the tumour with the tools. This problem, shown in figure 6.1, is related to the tools passing through different keyholes than the laparoscope, as identified in (Collins et al., 2017). Quoting (Collins et al., 2017): “In our user study we found this is a significant problem with smaller and/or deeper tumours, and can cause them to be missed.” A simple and sensible solution named Tool Access Visualisation was proposed, which is to project the tumour to the organ anterior boundary towards the tool keyhole and only then to the laparoscope. The impact of this correction was not quantified, but is very important: a simple geometric reasoning shows that the typical aiming error could be of the order of 30 mm for deep tumours. This solution was tested in (Collins et al., 2017) in an experimental setup involving an ex-vivo porcine kidney and a pelvitrainer box. The organ was static and the tool keyhole position was physically measured prior to running the experiment. Tool Access Visualisation was never applied to a real clinical case, because it was not possible to estimate the tool keyhole position with respect to the organ.

We propose a complete pipeline enabling the use of Tool Access Visualisation in the clinical setting in order to realise keyhole-aware AR for LLR. Specifically, we bring two key contributions. Our first contribution is to develop a geometric patient model, from which we precisely evaluate the impact of visualisation for aiming at an internal structure of the liver, as a function of several parameters including the tumour depth. Our second contribution is a computational method to estimate the tool keyhole position relative to the organ intraoperatively. Once the keyhole is computed, a resection path for the tumour is projected towards the keyhole and rendered in the laparoscopic image, as shown in figure 6.1. This is a challenging problem, for the keyhole is never directly seen by the laparoscope, hence must be localised by indirect means. Our pipeline is fully automated. We present results on simulated and clinical data.

**AR using camera-based projection:**



**AR using keyhole-based projection:**

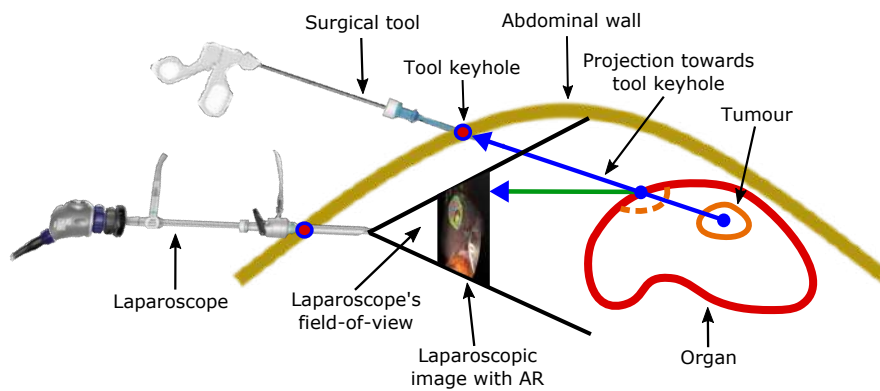


Figure 6.1: AR using camera-based and keyhole-based projections. Existing direct camera projection does not allow the surgeon to aim at the tumour. Keyhole-aware projection takes the tool keyhole into account, allowing the surgeon to directly aim following the AR indication. The green curve represents the 1 cm resection margin advised for HCC interventions. Each yellow curve represents a 1 cm distance through the parenchyma.

## 6.2 Keyhole-Aware AR for LLR

### 6.2.1 General Pipeline

Our system follows 6 steps, illustrated in figure 6.2. We assume the camera is static, which is done in practice using a Martin’s arm holder. We calibrate the camera by filming a checkerboard and using OpenCV. In step 1, the surgeon moves the tools around for a short time of about a minute and the video stream is recorded. In step 2, we detect geometric tool primitives (tool edgelines, midlines and shaft-tip) automatically in the recorded video images using ART-Net (Hasan et al., 2021). In step 3, we compute the tool 3D pose from the primitives using algebraic geometry (Hasan et al., 2021). We save the tool axis  $V$  in camera coordinates for each image, which is represented by a point  $q \in \mathbb{R}^3$  and an orientation vector  $W \in \mathbb{R}^3$ ,  $\|W\| = 1$ . We thus obtain a set of tool 3D axes  $\mathcal{V} = \{V_1, \dots, V_n\}$ . In step 4, described in section 6.2.2, we find the keyhole locations  $\mathcal{X} = \{x_1, \dots, x_c\}$  by clustering the 3D axes. In step 5, we register the preoperative 3D model to the images using SVB (Koo et al., 2017). In step 6, described in section 6.2.3, we visualise the tumours and resection paths with keyhole-aware projection based on the surgeon’s keyhole choice.

### 6.2.2 Keyholes Estimation and Axes Clustering

The process of estimating the keyhole locations works by clustering the tool 3D axes. Our clustering method iteratively finds the dominant cluster within the set of 3D axes  $\mathcal{V}$ . It proceeds by random sampling and has similarities with RANSAC (Fischler and Bolles, 1981). There are key differences however. First, RANSAC works for one cluster only. Second, we give an improved sampling scheme, which preemptively discards erroneous axes from the axis set. Our approach is given in algorithm 6.1. Concretely, **KeyholesEstimationByAxesClustering** has a main loop to find the dominant cluster and keyhole position. It stores the keyhole position if the cluster is larger than the constant **min\_cluster\_size** and terminates otherwise. The dominant cluster is found by subroutine **FindDominantCluster**. This finds the cluster with maximum support from the axes and proceeds by sampling keyhole positions and computing their support with subroutine **FindSupportingAxes**. An axis supports the keyhole hypothesis if their distance is lower than the constant **max\_dist**. Everytime a larger cluster is found, the number of required iterations is updated, as in modern RANSAC implementations, from subroutine **MaxTrials**. This simply implements equation (4.18) from (Hartley and Zisserman, 2003) with  $s = 2$  and  $p = 0.99$ . **FindDominantCluster** terminates by removing the clustered axes from the axis set. Finally, the subroutine **SampleKeyholePosition** creates random keyhole hypotheses. It works by sampling pairs of axes and finding their mid-point, defined as their least-squares closest point. More specifically, our optimised sampling exploits the fact that if the mid-point is too far from the axes, then it is necessarily a false hypothesis, which does not require further consideration from the algorithm. For that, we first sample an initial axis using subroutine **SampleOneAxisRandomly**. We then randomly search for an axis which intersects up to a mid-point distance residual of **max\_dist**, with subroutine **FindIntersectingAxisRandomly**. If such an intersecting axis is not found, it means that the initial axis can never be clustered, hence is an erroneous measurement not pertaining to any keyhole, and is thus removed from the axis set. The two global variables, the minimum number of axes required to support a valid keyhole and the maximum clustering axis to keyhole distance, are fixed to which we fix to **min\_cluster\_size = 4** and **max\_dist = 20 mm**.

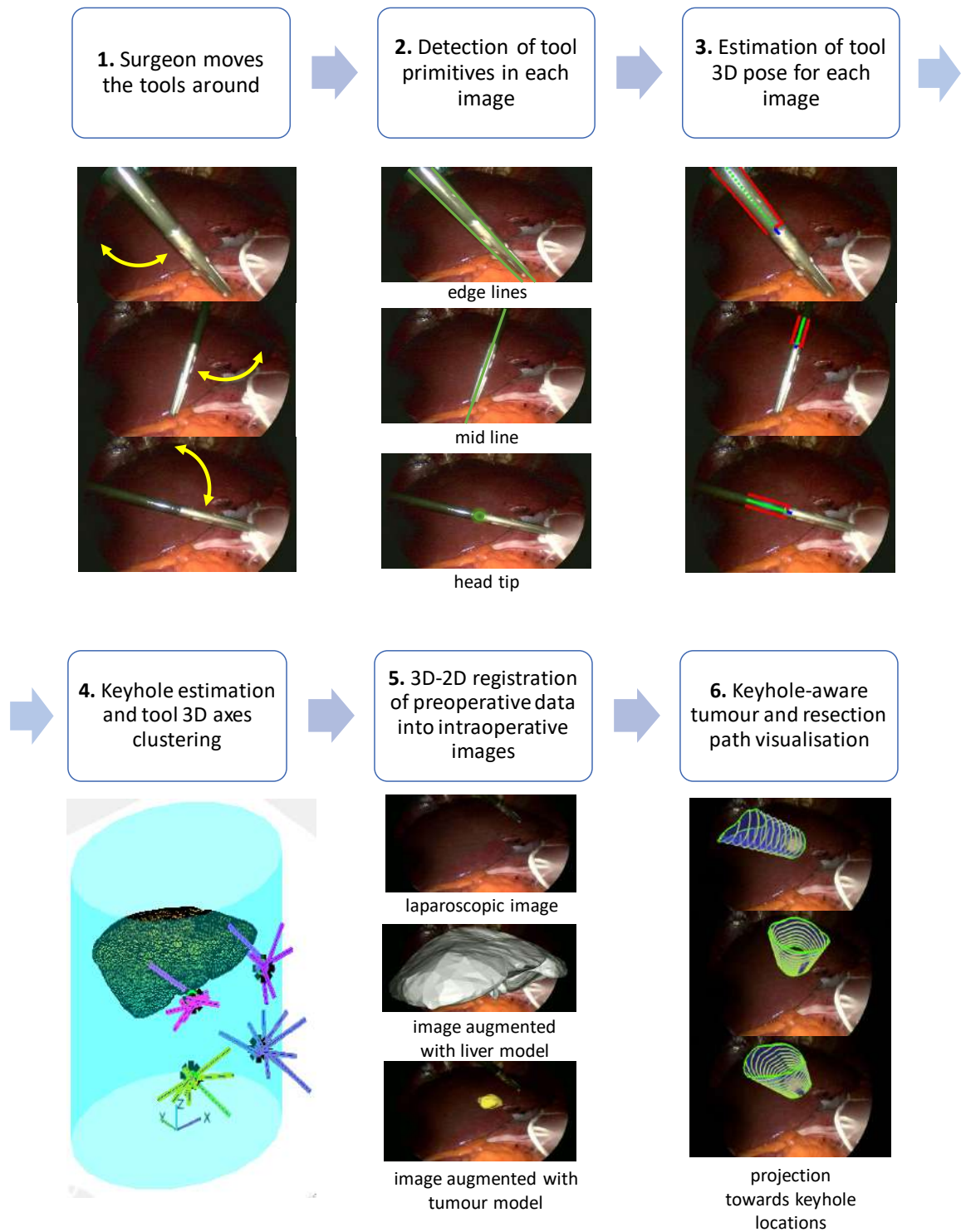


Figure 6.2: Pipeline of the proposed keyhole-aware AR method on a patient case.

**Algorithm 6.1:** Estimation of 3D keyhole locations through clustering of tool 3D axes.

---

```

1 Function SampleKeyholePosition
   Input:  $\mathcal{V} = \{V_1, \dots, V_m\}$  // Set of 3D axes
   Output:  $\mathcal{V}, x$  // Set of remaining 3D axes, 3D keyhole location
2   found  $\leftarrow$  false
   // Main sampling loop:
3   while not found and  $\text{size}(\mathcal{V}) \geq \text{min\_cluster\_size}$  do
4      $A \leftarrow \text{SampleOneAxisRandomly}(\mathcal{V})$  // Random sample the initial axis
5      $[\text{found}, x] \leftarrow \text{FindIntersectingAxisRandomly}(\mathcal{V}, A)$  // Attempt to find an axis that intersects
       the initial one
6     if not found then
7        $\mathcal{V} \leftarrow \mathcal{V} \setminus A$  // If no intersecting axis is found, remove the initial axis from the set of axes

8 Function FindDominantCluster
   Input:  $\mathcal{V} = \{V_1, \dots, V_m\}$  // Set of 3D axes
   Output:  $\mathcal{S} = \{V'_1, \dots, V'_p\}, x, \mathcal{V}$  // Set of 3D axes in the dominant cluster, 3D keyhole location, set
       of remaining 3D axes
9   max_trials  $\leftarrow \infty$  // Maximum number of trials
10  nb_trials  $\leftarrow$  0 // Trial count
11   $\mathcal{S} \leftarrow \emptyset$ 
12  repeat
13     $[x, \mathcal{V}] \leftarrow \text{SampleKeyholePosition}(\mathcal{V})$  // Randomly sample a keyhole
14     $\mathcal{S}' \leftarrow \text{FindSupportingAxes}(\mathcal{V}, x)$  // Find supporting axes for keyhole
15    if  $\text{size}(\mathcal{S}') > \text{size}(\mathcal{S})$  // If new support is greater than current best
16      then
17         $\mathcal{S} \leftarrow \mathcal{S}'$  // Set the new cluster as the current best
18        max_trials  $\leftarrow \text{MaxTrials}(\text{size}(\mathcal{V}), \text{size}(\mathcal{S}'))$  // Update the required number of trials
19  until nb_trials < max_trials and  $\text{size}(\mathcal{V}) \geq \text{min\_cluster\_size}$ 
20   $\mathcal{V} \leftarrow \mathcal{V} \setminus \mathcal{S}$  // Remove the dominant cluster from the set of axes

21 Function KeyholesEstimationByAxesClustering
   Input:  $\mathcal{V} = \{V_1, \dots, V_n\}$  // Set of 3D axes
   Output:  $\mathcal{X} = \{x_1, \dots, x_c\}$  // Set of 3D keyholes
22  converged  $\leftarrow$  false,  $\mathcal{X} \leftarrow \emptyset$ 
23  while not converged do
24    if  $\text{size}(\mathcal{V}) < \text{min\_cluster\_size}$  then
25      converged  $\leftarrow$  true // Too few axes for a new cluster
26       $[\mathcal{S}, x, \mathcal{V}] \leftarrow \text{FindDominantCluster}(\mathcal{V})$ 
27      if  $\text{size}(\mathcal{S}) < \text{min\_cluster\_size}$  then
28        converged  $\leftarrow$  true // Dominant cluster too small for a new keyhole
29      else
30         $\mathcal{X} \leftarrow \mathcal{X} \cup \{x\}$  // Dominant cluster forms a new keyhole

```

---

This algorithm implements a general form of clustering in order to avoid manual surgeon interactions with the computer. Two obvious simplifications could be introduced by requiring surgeon interactions. First, the number of keyholes, which is found automatically by Algorithm 6.1, could be given by the surgeon and used to stop the main clustering loop. Second, the video could be manually cut into clips, each showing only one tool for a specific keyhole, so that the keyholes could be directly found by calling subroutine **FindDominantCluster** once for each clip. We propose that the surgeon interactions thus remain optional and be only requested upon failure of the general Algorithm 6.1. Our experimental results show that, in spite of its generality, Algorithm 6.1 performs well.

### 6.2.3 Keyhole-Aware Tumour and Resection Path Visualisation

Keyhole-aware visualisation is a sequence of two projections, as shown in figure 6.3: first, orthographic projection of the internal structure to the liver surface, second, perspective projection to the camera. Concretely, we implemented two visualisation modes: a simple mode where only the tumour is visualised, and an advanced mode where the full resection path is visualised. An ex-

ample of such visualisations can be seen in figure 6.14. In 3D this path is constructed from a line joining the keyhole to the tumour centroid and clip to the parenchyma and hence corresponds to the first projection implemented in keyhole-aware visualisation. We graphically represent the path by a generalised cylinder whose axis is the projection line, following five steps. First, we generate a clipping plane oriented towards the keyhole and with origin at the tumour centroid. Second, we construct the tumour boundary contour as the intersection between the tumour and the clipping plane. Third, we expand this contour by 1 cm within the clipping plane to add the resection margin advised for HCC interventions (Zhong et al., 2017). The size of this margin can also be specified by the surgeon. Fourth, we project the margin curve orthographically towards the keyhole, creating the generalised cylinder. The intersection between the cylinder and the liver surface is used to draw the margin projection on the liver surface. Fifth, we sample the cylinder in planes parallel to the clipping plane and 1 cm apart, providing depth visualisation. Finally, we use perspective projection to render the generated elements on the camera, forming the augmented laparoscopic image with the keyhole-aware tumour and resection path. Both the proposed simple and advanced keyhole-aware visualisation modes can be used with any projection direction chosen by the surgeon. A tool keyhole position is particularly adapted, as it allows them to directly aim at the tumour. However, we have also implemented a mode where the surgeon simply clicks on the liver to define a custom projection direction.

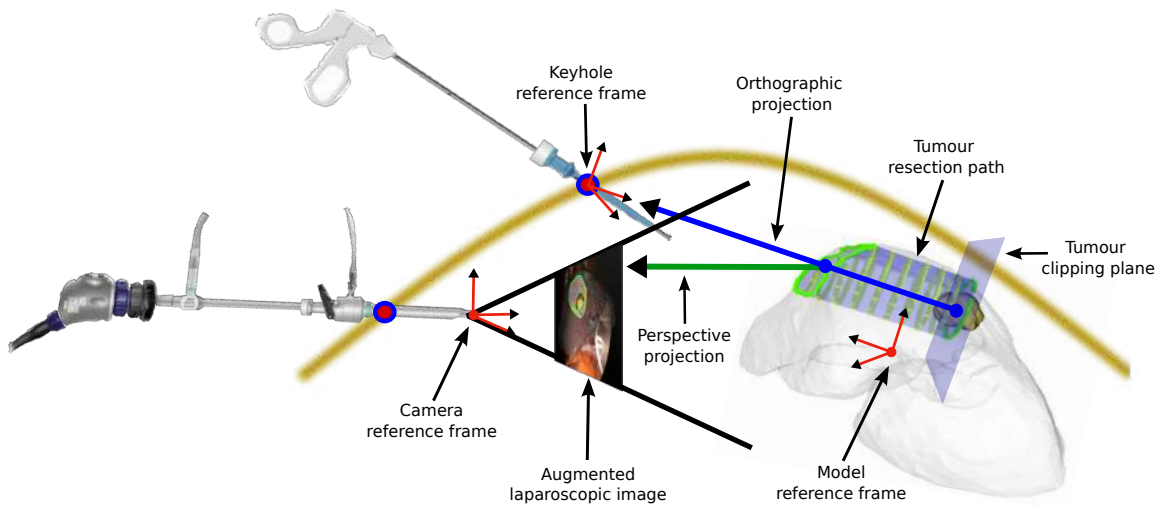


Figure 6.3: Generation of keyhole-aware resection path.

## 6.3 Quantitative Evaluation on a Geometric Model

### 6.3.1 A Geometric Model of LLR

We use a geometric 3D model shown in figure 6.4 to simulate the laparoscopic setting and measure the clustering, keyhole and pointing errors. Our model uses a cylinder to represent a human abdomen, with a radius  $r = 16.5$  cm, which is a common size of an insufflated abdominal cavity (Malbrain et al., 2014). We position a liver 3D model at a distance  $d_{liver} = 5$  cm from the abdominal wall and a virtual tumour  $a$  with its centre located 3 cm away from the liver surface. We use 4 keyholes distributed in a rectangular fashion with an inter-keyhole distance  $d_{keyhole} = 15$  cm and a laparoscope keyhole located beneath the tool keyholes.

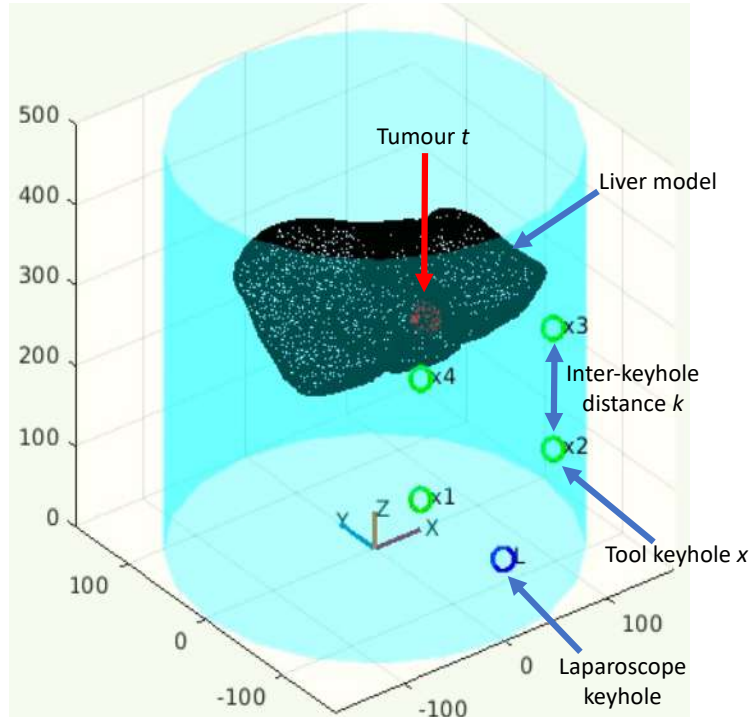


Figure 6.4: Geometric 3D model of the laparoscopic settings used to generate synthetic data.

### 6.3.2 Assessment of the Clustering Success and Keyhole Estimation Error

We use our geometric 3D model to randomly generate multiple tool axes passing through each of the keyholes. Because the real estimated tool axes are affected by detection errors, calibration errors and tissue deformations, we add random noise to the generated axes. The tool axes are thus simulated as follows: for a given keyhole  $x$ , we sample a line  $V = (x', W)$  where  $x'$  is obtained by perturbing  $x$  with random noise and  $W$  is a random unit vector. The perturbed keyhole is defined as  $x' = x + \gamma W'$ , where  $W'$  is a unit vector perpendicular to  $W$  and  $\gamma$  is the noise magnitude. This process is repeated for every generated tool axis. We create both valid and erroneous axes by changing the noise distribution. For valid axes, we use a normal distribution  $\gamma \sim \mathcal{N}(0, 2)$  with  $\sigma = 5$  mm, while for erroneous axes we use a uniform distribution  $\gamma \sim U(\psi, \omega)$  with  $\psi = 1.965$  mm and  $\omega = 150$  mm. We generate a total of 1000 tool axes, from which 10% are erroneous. These axes are drawn at 40% for keyhole #1, 30% for keyhole #2, 20% for keyhole #3 and 10% for keyhole #4. We measure the clustering success and keyhole error against varying valid noise magnitude from 0 to 20 mm in steps of 2 mm, number of observed tool axes from 100 to 2000 in steps of 100, and rate of erroneous tool axes from 0% to 50% in steps of 5%. At every step, we run the clustering algorithm 100 times. We measure the error mean and standard deviation across all runs. These results are shown in figure 6.5.

We observe that increasing tool axis noise decreases clustering success. Starting at a nearly perfect result of  $98.2 \pm 0.5\%$  without noise, it decreases to  $62.6 \pm 1.5\%$  at 20 mm noise. In spite of this decrease in clustering success, we observe that the keyhole error, starting at  $1.95 \pm 0.44$  mm increases to  $4.21 \pm 1.11$  mm only, which represents a reasonable error. Its value is stable for noise values lower than 5 mm. Increasing the number of observed tools does not affect the average clustering success but dramatically reduces its spread, from  $96.3 \pm 2.5\%$  for 100 tools to

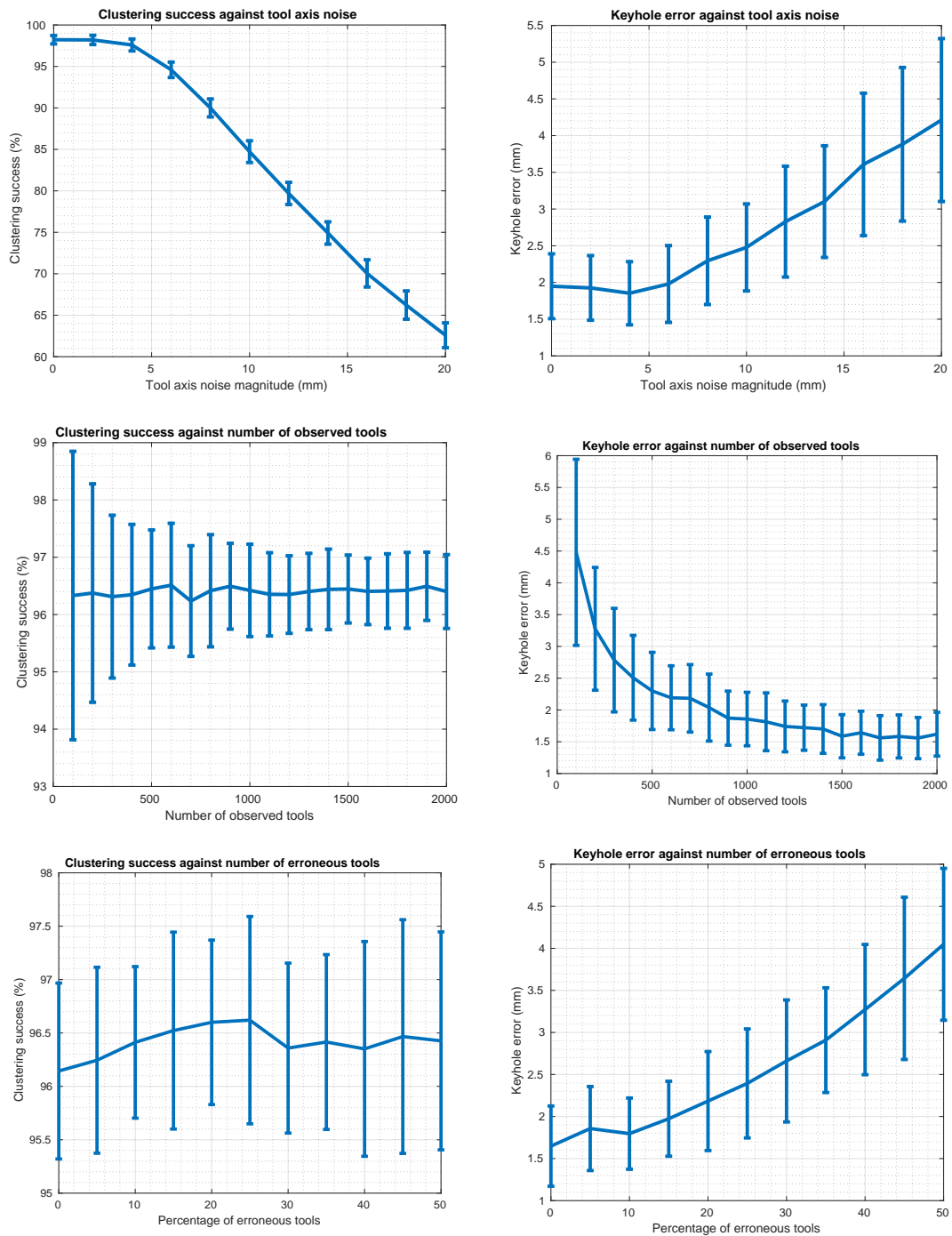


Figure 6.5: Clustering success (left) and keyhole error (right) against tool axis noise, number of observed tool axes and number of erroneous tool axes.



$96.4 \pm 0.6\%$  for 2000 tools, while substantially decreasing the keyhole error from  $4.48 \pm 1.46$  mm to  $1.61 \pm 0.34$  mm, stabilising beyond 1500 observed tool axes, representing a minute of video at 25 fps. Varying the percentage of erroneous tools from 0% to 50% does not have an effect on the clustering accuracy either, remaining at a good rate of  $96.5 \pm 0.9\%$ , while the keyhole error increases from  $1.64 \pm 0.47$  mm to  $4.04 \pm 0.90$  mm, which is still a reasonable error. From these observations and the typical settings of laparoscopy, we define a keyhole error average as  $\mu_0 = 1.9$  mm and standard deviation  $\sigma_0 = 0.45$  mm.

### 6.3.3 Assessment of Pointing Errors

We use our geometric 3D model to measure the influence of keyhole-aware visualisation, compared to direct camera projection, on tumour aiming with a surgical tool, as illustrated in figure 6.6. To achieve this, we first generate a keyhole  $x' = x + \gamma W$  from one of the true keyholes  $x$ , where  $W$  is a random unit vector and  $\gamma \sim \mathcal{N}(\mu_0, \sigma_0^2)$  is the noise magnitude. We define the pointing error as the distance between the tool axis directed to the tumour from the surgeon's perception and the tumour. The direct-camera pointing error, or type-A error, occurs when the surgeon follows the existing direct camera projection. We measure it by first projecting the tumour  $a$  towards the camera centre  $k$ . Then, this projection is intersected with the liver surface  $J$ , which gives the point  $w_k = (a, k) \cap J$ . The type-A error will be the distance between the line  $(x, w)$  and the tumour  $a$ . The keyhole-aware pointing error, or type-B error, occurs when the surgeon follows the proposed keyhole-aware visualisation. We measure it by first projecting the tumour  $a$  towards the keyhole  $x'$  and intersecting the liver surface  $J$ , giving point  $w_x = (a, x') \cap J$ . The type-B error will be the distance between the line  $(x, w_x)$  and the tumour  $a$ . We measure the pointing errors against tumour depth between 0 and 100 mm in steps of 5 mm, inter-keyhole distance between 100 and 200 mm in steps of 5 mm, and keyhole precision between 1.84 and 4.2 mm in steps of 0.1 mm. At every step, we estimate a keyhole  $x'$  and measure the pointing errors 100 times. These results are shown in figure 6.7.

We observe that the type-A error is always larger than the type-B error. The type-A error is strongly affected by the tumour depth, with a critical error beyond 35 mm. The variations are due to the liver's irregular shape. In contrast the type-B error is not significantly affected by tumour depth, being always lower than 1 mm. The type-A error linearly increases with inter-keyhole distance, from 21.16 mm to 37.74 mm. This was expected, as closer tool and camera keyholes reduce the direct visualisation bias. The type-B error is unaffected at 0.48 mm. The type-A error is obviously unaffected by keyhole uncertainty, as it does not depend on the keyhole position, while the type-B error linearly increases from 0.36 mm to 0.79 mm, remaining largely lower than the type-A error.



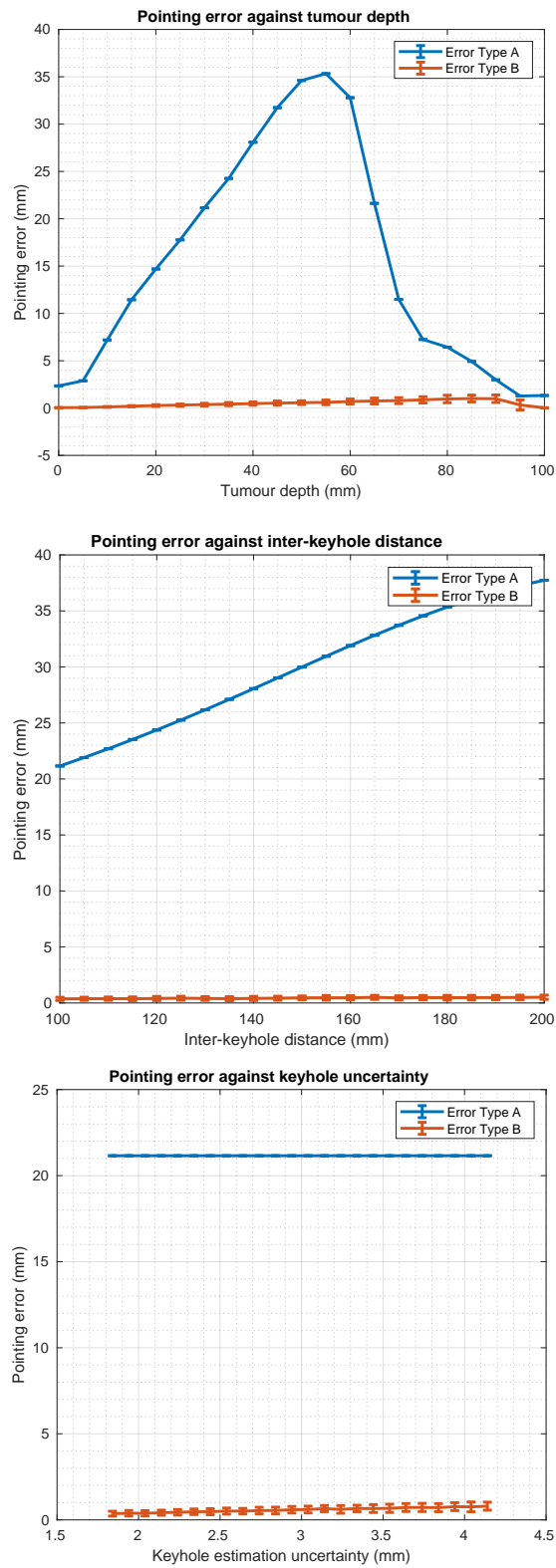


Figure 6.7: Type A-and type-B pointing errors against tumour depth, inter-keyhole distance and keyhole precision.

## 6.4 Quantitative Evaluation on Phantom Data

### 6.4.1 Experimental Setup

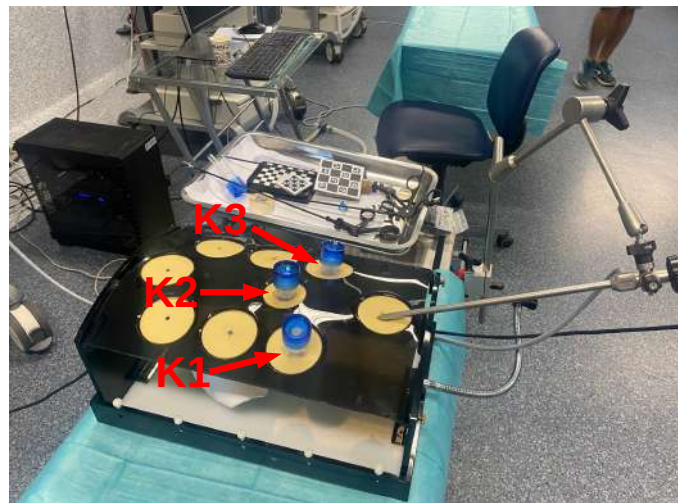
We design a realistic experimental pipeline that uses one of the deformed liver phantoms generated in chapter 5, and a STORZ monocular endoscopy system. As for the geometric model case, the goal of this setup is to measure the keyhole estimation errors and the pointing errors by using the estimated keyholes. A desktop PC with the AR software is connected to the video output port of the endoscopy system. The liver phantom is positioned inside a pelvitrainer and the laparoscope is fixed in place using a Martin's arm holder, in a way that we have a global view of the phantom. Three keyholes are used in our experiments, as shown in figure 6.9(a). To compute the keyhole errors, the groundtruth keyhole locations are obtained by finding the pose of a ChArUco board in the laparoscopic image, as shown in figure 6.9(b). To measure the pointing errors, an external camera is mounted on a tripod, looking towards the pelvitrainer and the liver phantom, as shown in figure 6.9(c). A ChArUco board is also attached to a surgical tool, in order to find the tool poses during the pointing experiments. Both the laparoscope and the external camera were properly calibrated using a checkerboard pattern. A view of the whole setup is shown in figure 6.8.



Figure 6.8: Setup for keyhole estimation experiments using phantom data. At the left of the image: The external camera used to measure the pointing errors. Center: The pelvitrainer with the liver phantom and the laparoscope fixed with a mechanical arm. Right: the PC with the AR software connected to the endoscopy system.

### 6.4.2 Clustering Success and Keyhole Estimation Error

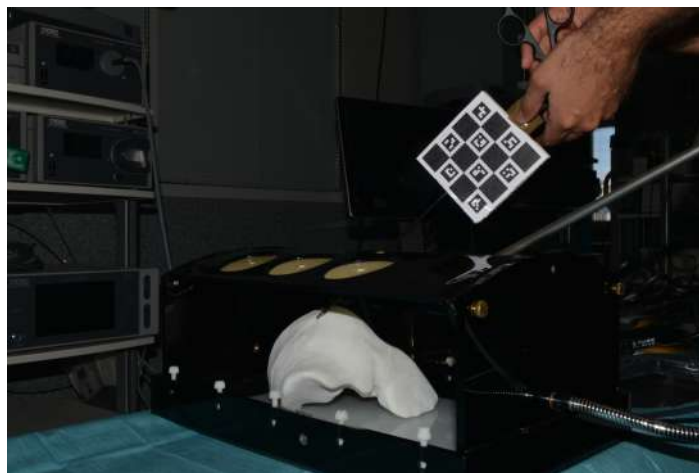
Three surgical tools were inserted simultaneously through the aforementioned keyholes, and a video stream was captured of the tools making circular movements. From the captured video, 30 frames were selected and the tool primitives were manually annotated for each frame. An example of the extracted frames is shown in figure 6.10(a), with the corresponding primitives shown in



(a)



(b)



(c)

Figure 6.9: Details of the phantom experimental setup: (a) The 3 keyholes used in the experiments; (b) Laparoscopic image showing a ChArUco board to estimate a groundtruth keyhole location; (c) View from the external camera showing a tool with the attached ChArUco board pointing towards the liver phantom.

figure 6.10(b). Then, the 3D tool poses were estimated from the annotated primitives using the method from (Hasan et al., 2021), as shown in figure 6.11(a). We applied our method on this unordered set of tool poses to cluster them and estimate the keyhole locations. By setting the global parameters to **min\_cluster\_size = 15** and **max\_dist = 10 mm**, we were able to cluster the 3 keyholes and estimate their locations in the laparoscope’s reference frame, as shown in figure 6.11(b). The number of clustered tools per keyhole and the estimated keyhole locations are shown in table 6.1.

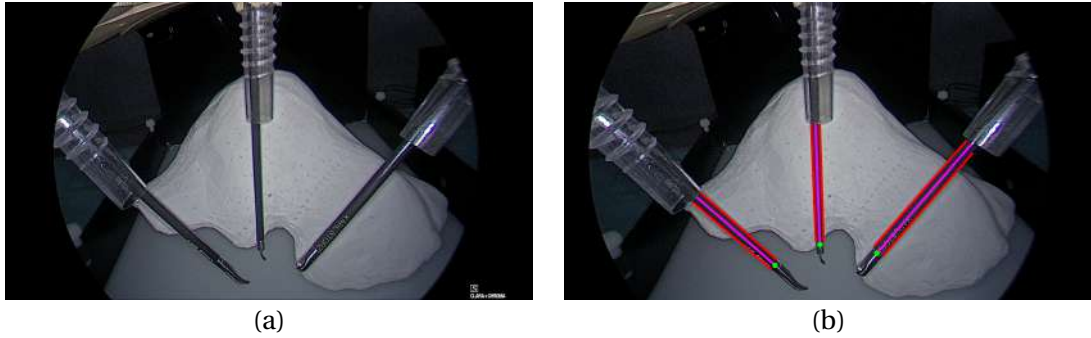


Figure 6.10: Tools used to estimate the keyhole positions: (a) Laparoscopic image of the tools being inserted in the 3 keyholes; (b) Annotated tool primitives corresponding to the edge lines, middle line, and tool tip.

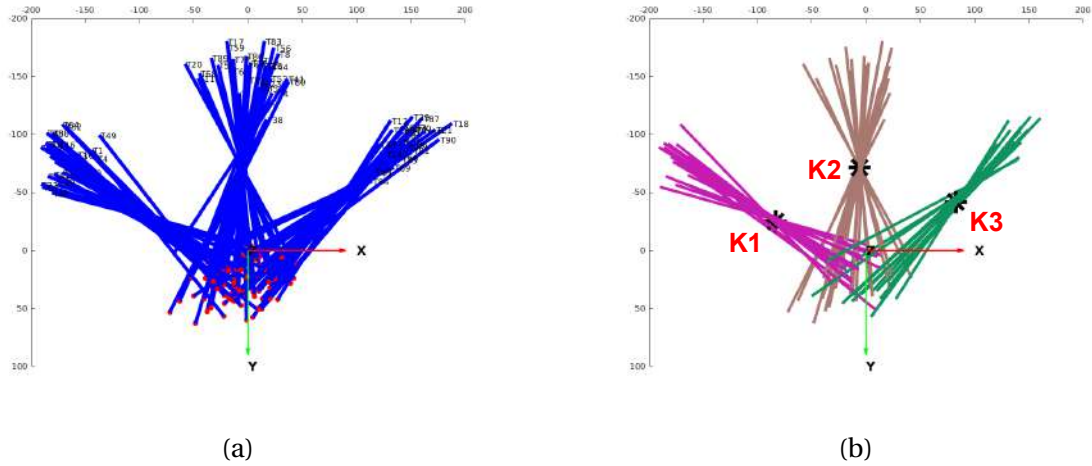


Figure 6.11: Tool poses and estimated keyholes: (a) Reconstructed tool poses in laparoscope’s reference frame; (b) Clustered tools and estimated keyholes.

Keyhole	Number of detected tools	Keyhole estimation error (mm)
#1	21	16.68
#2	17	05.37
#3	16	06.15

Table 6.1: Clustering results and keyhole estimation errors.

### 6.4.3 Assessment of Pointing Errors

We make use of the previously estimated keyholes to perform keyhole-aware AR on the liver phantom and measure the pointing accuracy on virtual tumours. A preoperative 3D model of the liver phantom containing 8 virtual tumours is registered to the laparoscopic image using SVB. Then, we point with the surgical tool from every keyhole towards each of the tumours using their direct camera projections, as well as their projections towards the keyholes. This means that, if the instrument is inserted through keyhole #1, we point to the camera projection of a tumour, and to the projection of that tumour towards the estimated location of keyhole #1. An example of such pointings is shown in figure 6.12. For every pointing experiment, we take a photo of the tool and the phantom with the external camera. This gives a total of 48 pointings: 24 for the camera projections and 24 for the keyhole projections. To measure the pointing errors, we register the deformed 3D model used to print the liver phantom, which also contains the groundtruth locations of the deformed tumours, to the corresponding external camera image using PnP. We also find the 3D pose of the surgical tool in every image from the attached ChArUco board. Finally, we compute the average type-A error for the camera projections and the average type-B error for the keyhole projections, by measuring the distance between every tumour centre and the tool axes. These errors are reported in table 6.2. We observe a reduction of 20.25 mm in the pointing error when using the keyhole-aware AR compared to the traditional AR. It is worth noticing that the pointing errors include inaccuracies coming from the preoperative-to-intraoperative registration, the keyhole estimations, and the pointing gestures done by the user. For example, we can observe that the type-B errors per keyhole are proportional to the keyhole estimation errors, which is an expected outcome of the experiments.

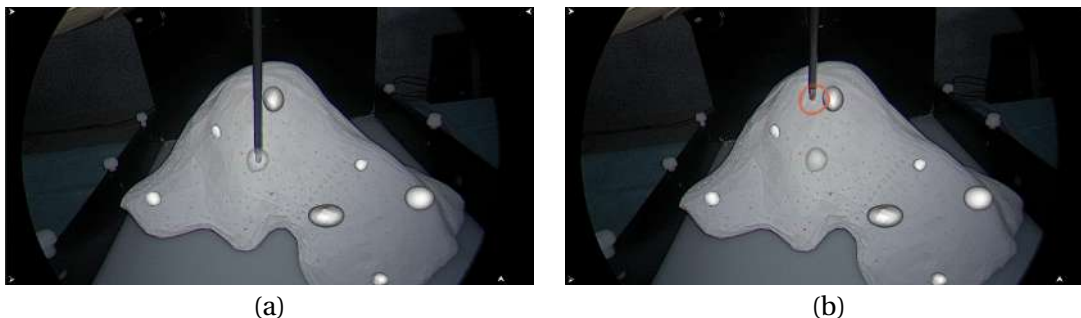


Figure 6.12: Example of tumour pointing using AR on a liver phantom: (a) Pointing to the camera-based projection of the tumour; (b) Pointing to the keyhole-based projection of the tumour.

Keyhole	Type-A error (mm)	Type-B error (mm)
#1	$27.52 \pm 14.54$	$12.71 \pm 05.82$
#2	$31.70 \pm 16.73$	$07.43 \pm 05.17$
#3	$30.65 \pm 15.23$	$08.98 \pm 06.76$
average	$29.96 \pm 14.94$	$09.71 \pm 06.12$

Table 6.2: Pointing errors for camera projections (type A) and keyhole projections (type B).

## 6.5 Qualitative Evaluation on Patient Data

We did a preliminary test on a patient case, for which the surgeon fixed the laparoscope using a Martin’s arm holder. The surgeon opened four keyholes in the abdominal cavity and, following the pipeline of section 6.2.1, performed camera calibration and started moving the tools in each of the keyholes while the video stream was being recorded. After the surgery, we extracted 23 images from the recorded video, 8 correspond to keyhole #1, 5 to keyhole #2, 5 to keyhole #3, and 5 to keyhole #4. We used the ART-Net method (Hasan et al., 2021), whose model was kindly provided to us by the authors, to compute the tool 3D poses. However, we observed a poor performance, which we explain by the fact that ART-Net was originally trained on a limited dataset with a different distribution compared to our images. We thus fine-tuned ART-Net with 71 additional images which we manually annotated, split in 56 images for training and 15 for validation. We applied augmentation to the training data including random flipping, scaling, blurring and rotation. The fine-tuning improved ART-Net’s results as shown in figure 6.13, and we could obtain the geometric primitives and the tool 3D poses for most of the images. Specifically, we obtained reasonable poses in 3 images for keyhole #1, 2 images for keyhole #2, 2 images for keyhole #3 and 3 images for keyhole #4. We then used our method to estimate the corresponding keyhole locations, and proceed to project the tumours towards the keyholes. An example of such a projection to one keyhole using the simple and advanced visualisation methods is shown in figure 6.14. From the keyhole-aware renderings we observe that the tumour projections are oriented towards the tool. Compared to existing direct camera projection, which offers ambiguous information on the tumour’s real localisation, keyhole-aware rendering makes it easier for the surgeon to access the tumour with the surgical tool. This could potentially increase the precision of laparoscopic resections, and reduce the trauma to the patient.

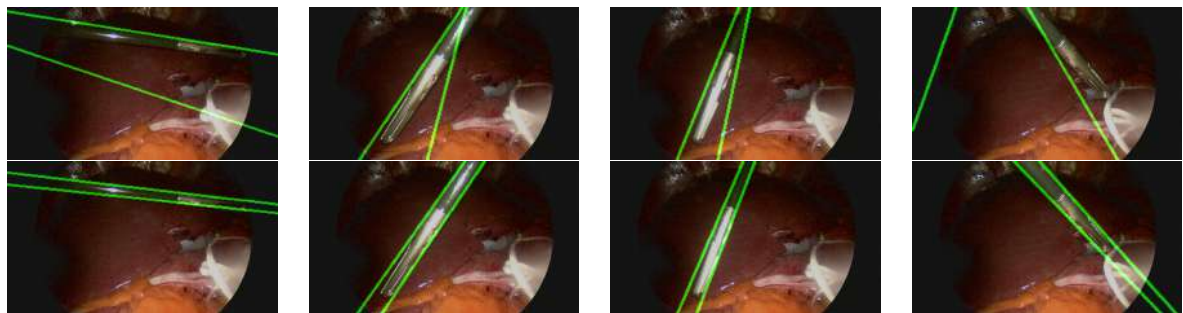


Figure 6.13: Detection of tool edge primitives before tuning (top) and after fine tuning (bottom) of ART-Net.



Figure 6.14: AR visualisation, keyhole #2. Original camera-based visualisation (left), simple keyhole-aware visualisation showing the projected tumour shape on the liver surface (middle), advanced keyhole-aware visualisation with full resection path (right).



## 6.6 Conclusion

We have presented a visualisation method for surgical AR in laparoscopic liver surgery, which takes into account the location of the keyholes of the tools used to resect internal tumours. Specifically, we have proposed the keyhole-aware visualisation, where the tumour is not directly rendered to the image plane, but first projected to the liver surface towards the aiming keyhole. We have also proposed to estimate the keyhole position by reconstructing the tool 3D pose for several images and computing their shaft axis intersection point. Our method copes with noise and erroneous tool axes. We have shown that keyhole-aware visualisation fixes a gap in existing AR systems and may have a dramatic impact on AR usability and surgical safety.

Our experiments show that, estimating the keyhole locations and using them in a keyhole-aware AR system improve the perception of the tumour location inside the liver and, therefore, increases the probability of reaching the tumour with the surgical tool. While using an increasing number of tool poses improves the keyhole estimation accuracy, our method is able to estimate the keyholes with less than 100 tool poses in way that the tumour pointing still respects the 1 cm security margin advised for HCC interventions. According to the feedback given by the surgeon collaborating in this work, our method does help to have a better understanding of the tumour's location, which can improve the surgical gesture to reach the tumour effectively.

We plan to extend our system in three ways. First, we plan to handle a mobile organ such as the kidney, in which case, the keyholes must be dynamically updated from new tool observations available during resection. Second, we plan to handle the presence of multiple tools visible at once, for which the tool 3D pose method which we re-used from the literature will have to be extended with an initial step of instance segmentation. Third, we plan to extend our clinical evaluation and include both monocular and stereo laparoscopy. We also plan to strengthen our results by asking several surgeons to perform tumour pointing with the help of the keyhole-aware AR, as well as by applying the method on more patients.

## Chapter 7

# Conclusions and Future Work

Laparoscopy of the liver remains a difficult technique due to the big dimensions of the liver, the reduced FoV of the camera, and the lack of tactile feedback. While AR may help to overcome these issues by showing the internal anatomy of the liver, it is still an open and challenging research problem that requires custom-designed techniques to register and visualise such inner structures. These techniques should deal with problems inherent to the liver such as the partial visibility in the images and the high elasticity. In this thesis we have proposed several registration methods that can take advantage of the information available in the intraoperative scene, beyond a single laparoscopic image. This has been done by refining the non-visible parts of the liver through manual tactile interactions as proposed in chapter 4 and by combining the information coming from multiple laparoscopic images as proposed in chapter 5. Even though these methods target the initial registration step for AR, the resulting registrations are a critical basis for the subsequent tracking step. In that sense, an accurate initial registration is important to guarantee reliable real-time augmentations.

While the multi-view methods proposed in chapter 5 are suitable for cases where the liver is in a rest state, without noticeable deformations between images, the augmentations can be used to perform resection planning before starting the surgery. This shows that, by merging the liver landmarks and texture information spread around several images, the registration accuracy can be improved. However, using multiple images of a deforming liver did not improve accuracy, even though it was possible to find common information between them. Also, using views from different keyholes only works if there is enough overlapping information between them. Experiments on patient data showed that the warp-based methods were able to provide reasonable results, while the SfM-based method was not even able to reconstruct an intraoperative shape. This is mainly due to the presence of slight liver deformations across views, for which SfM techniques are not ready.

We have also demonstrated that a clear understanding of the augmentations is crucial to accurately locate and reach the inner structures. By using a keyhole-aware method like the one proposed in chapter 6, the surgeon can instantly know how to reach a tumour from a particular keyhole, and perform a resection plan accordingly. While our method was only tested on static monocular images, it could be adapted to work in real-time tracking, as well as in stereo endoscopic systems.

According to the results presented in this work, several of the proposed methods can be already transferred to the industry and applied on clinical cases. This includes the SVH method, which

offers improved registration accuracy compared to single-view methods like SVB, while remaining easy to use. For the multi-view methods, the rigidly linked approaches MV-B and MV-C can also be applied to clinical cases. This is thanks to the improved accuracy and the possibility of seeing the augmentations from multiple view points. The MV-B\* method cannot be used in practice due to the difficulty of reconstructing an intraoperative shape using SfM on real patient data. The MV-D method should not be used in practice either due to the little benefit it gives over the SVB method. For the keyhole-aware visualisation case, even if the clustering algorithm can be used in practice, there is still work to be done in the automatic detection of the tool primitives. Once the primitives can be accurately detected, the keyhole-aware AR can be directly used in the OR. All the methods were implemented using unoptimised code, which means that the execution times can still be reduced.

The proposed methods can be improved in several ways, both to increase the registration accuracy and the execution times. Specifically, the Delaunay tetrahedrisation method used to generate the tetrahedra on which the biomechanical model relies converts the non-convex preoperative model into a convex one, as shown in figure 7.1. This has a negative impact on the registration accuracy due to the extra tetrahedra that join sections of the liver that should not be connected. Therefore, a triangulation that respects the topology of the liver should be implemented. Second, the PBD deformation method does not reduce completely the strain energy at the end of registration (Bender et al., 2014), which also decreases the final accuracy. More robust optimisation techniques should be tested like Gauss-Newton or Euler. Regarding registration using non-rigidly linked images, a method that can cope with the liver deformations across images while improving registration accuracy should be implemented. This would be useful to ensure a reliable real-time tracking with a deforming liver across time. Another way in which both the registration accuracy and the execution times could be improved is by automating the landmark marking on both the laparoscopic images and the preoperative 3D models. This can be done using deep learning approaches that automatically detect the 3D-2D landmark correspondences prior to registration. Regarding the automatic detection of the tool primitives for keyhole-aware AR, the high variability of the tools, textures, and environment lighting makes it necessary to collect more data for deep learning training purposes. This would also require designing a neural network that could cope with such inter-patient variability.

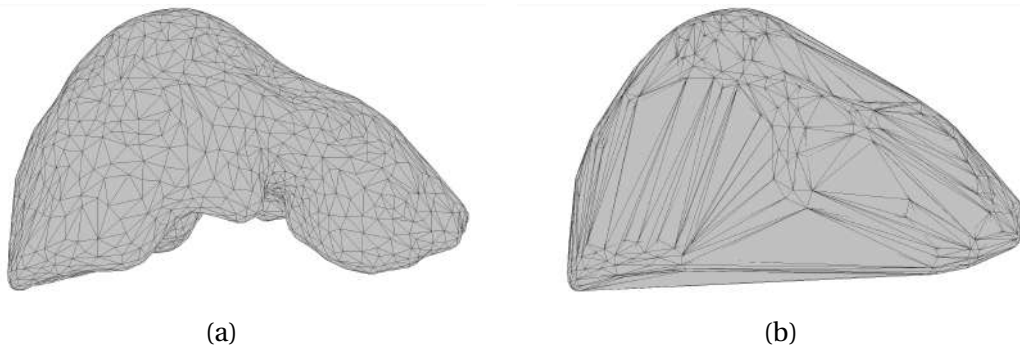


Figure 7.1: Triangulation of preoperative 3D model: (a) Original preoperative liver model; (b) Preoperative volumetric liver model after Delaunay tetrahedrisation.

# Bibliography

- Abergel, V., Jacquot, K., Luca, L., and Véron, P. (2019). Towards a slam-based augmented reality application for the 3d annotation of rock art. *International Journal of Virtual Reality*, 19. [50](#)
- Adagolodjo, Y., Trivisonne, R., Haouchine, N., Cotin, S., and Courtecuisse, H. (2017). Silhouette-based pose estimation for deformable organs application to surgical augmented reality. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 539–544. [29](#), [30](#), [33](#), [34](#), [36](#), [58](#), [74](#), [92](#)
- Adams, F., Qiu, T., Mark, A., Fritz, B., Kramer, L., Schlager, D., Wetterauer, U., Miernik, A., and Fischer, P. (2017). Soft 3d-printed phantom of the human kidney with collecting system. *Annals of Biomedical Engineering*, 45(4):963–972. [63](#)
- Agasthian, T. (2010). Revisiting the prone position in video-assisted thoracoscopic surgery. *Asian Cardiovascular and Thoracic Annals*, 18:364 – 367. [5](#)
- Agisoft LLC (2022). Agisoft Metashape. <https://www.agisoft.com/>. [Online; accessed 8-August-2022]. [63](#)
- Ahad, S. and Figueredo, E. J. (2007). Laparoscopic colectomy. *MedGenMed: Medscape general medicine*, 9(2). [6](#)
- Alander, J. T., Kaartinen, I., Laakso, A., Pättilä, T., Spillmann, T., Tuchin, V. V., Venermo, M., and Välisuo, P. (2012). A review of indocyanine green fluorescent imaging in surgery. *International Journal of Biomedical Imaging*, 2012. [10](#)
- Ali, S., Espinel, Y., Jin, Y., and Bartoli, A. (2022). Preoperative to Intraoperative Laparoscopy Fusion Challenge (P2ILF) - MICCAI 2022. <https://p2ilf.grand-challenge.org/>. [Online; accessed 12-August-2022]. [89](#)
- Alvelo-Rivera, M., De Hoyos, A., and Luketich, J. D. (2004). Laparoscopic and thoracoscopic esophagectomy. *Operative Techniques in Thoracic and Cardiovascular Surgery*, 9(2):157–176. [4](#)
- Amberg, B., Romdhani, S., and Vetter, T. (2007). Optimal step nonrigid icp algorithms for surface registration. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. [22](#)
- Andreou, A., Struecker, B., Raschzok, N., Krenzien, F., Haber, P., Wabitsch, S., Waldbaur, C., Touet, E.-M., Eichelberg, A.-C., Atanasov, G., Biebl, M., Bahra, M., Öllinger, R., Schmelzle, M., and Pratschke, J. (2018). Minimal-invasive versus open hepatectomy for hepatocellular carcinoma:

- Comparison of postoperative outcomes and long-term survivals using propensity score matching analysis. *Surgical Oncology*, 27(4):751–758. 14
- Angelico, R., Guzzo, I., Pellicciaro, M., Saffioti, M. C., Grimaldi, C., Mourani, C., Smedile, F., Pariente, R., Semprini, A., Monti, L., Candusso, M., Dello Strologo, L., and Spada, M. (2019). Same donor laparoscopic liver and kidney procurement for sequential living donor liver–kidney transplantation in primary hyperoxaluria type i. *Journal of Laparoendoscopic & Advanced Surgical Techniques*, 29(12):1616–1622. 8
- Aoki, T., Murakami, M., Koizumi, T., Matsuda, K., Fujimori, A., Kusano, T., Enami, Y., Goto, S., Watanabe, M., and Otsuka, K. (2018). Determination of the surgical margin in laparoscopic liver resections using infrared indocyanine green fluorescence. *Langenbeck's Archives of Surgery*, 403(5):671–680. 10
- Aragon, R. and Solomon, N. (2012). Techniques of hepatic resection. *Journal of Gastrointestinal Oncology*, 3(1). 15
- Asrani, S. K., Devarbhavi, H., Eaton, J., and Kamath, P. S. (2019). Burden of liver diseases in the world. *Journal of Hepatology*, 70(1):151–171. 13
- B. Braun (2021). Minimally invasive surgery. <https://www.bbraun.com/en/products-and-solutions/therapies/minimally-invasive-surgery.html>. [Online; accessed 13-July-2022]. 2
- Bae, H., Walker, M., White, J., Pan, Y., Sun, Y., and Golparvar-Fard, M. (2016). Fast and scalable structure-from-motion based localization for high-precision mobile augmented reality systems. *mUX: The Journal of Mobile User Experience*, 5(1):4. 48
- Barr, A. H. (1984). Global and local deformations of solid primitives. *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*. 59
- Bartoli, A., Perriollat, M., and Chambon, S. (2010). Generalized thin-plate spline warps. *International Journal of Computer Vision*, 88(1):85–110. 78
- Bechmann, D. and Gerber, D. (2003). Arbitrary shaped deformations with dogme. *The Visual Computer*, 19(2):175–186. 59
- Bellekens, B., Spruyt, V., and Maarten Weyn, R. B. (2014). A survey of rigid 3D pointcloud registration algorithms. In *Fourth International Conference on Ambient Computing, Applications, Services and Technologies, Proceedings*, pages 8–13. IARA. 38, 42
- Belli, G., Gayet, B., Han, H.-S., Wakabayashi, G., Kim, K.-h., Cannon, R., Kaneko, H., Gamblin, T., Koffron, A., Dagher, I., and Buell, J. F. (2013). Laparoscopic left hemihepatectomy a consideration for acceptance as standard of care. *Surgical Endoscopy*, 27(8):2721–2726. 14
- Bender, J., Koschier, D., Charrier, P., and Weber, D. (2014). Position-based simulation of continuous materials. *Computers Graphics*, 44:1–10. 53, 54, 59, 110
- Benjemaa, R. and Schmitt, F. (1999). Fast global registration of 3d sampled surfaces using a multi-z-buffer technique. *Image and Vision Computing*, 17(2):113–123. 41

- Benlice, C., Aytac, E., Costedio, M., Kessler, H., Abbas, M. A., Remzi, F. H., and Gorgun, E. (2017). Robotic, laparoscopic, and open colectomy: a case-matched comparison from the acs-nsqip. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(3). [6](#)
- Berber, E. and Siperstein, A. (2004). Laparoscopic ultrasound. *Surgical Clinics of North America*, 84(4):1061–1084. [10](#)
- Bernhardt, S., Nicolau, S. A., Bartoli, A., Agnus, V., Soler, L., and Doignon, C. (2016). Using shading to register an intraoperative ct scan to a laparoscopic image. In Luo, X., Reichl, T., Reiter, A., and Mariottini, G.-L., editors, *Computer-Assisted and Robotic Endoscopy*, pages 59–68. [24](#), [28](#), [29](#), [33](#), [35](#), [92](#)
- Bernhardt, S., Nicolau, S. A., Soler, L., and Doignon, C. (2017). The status of augmented reality in laparoscopic surgery as of 2016. *Medical Image Analysis*, 37:66–90. [16](#)
- Berryhill, R., Jhaveri, J., Yadav, R., Leung, R., Rao, S., El-Hakim, A., and Tewari, A. (2008). Robotic prostatectomy: A review of outcomes compared with laparoscopic and open approaches. *Urology*, 72(1):15–23. [7](#)
- Besl, P. and McKay, N. D. (1992). A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256. [39](#), [40](#)
- Bhave Chittawar, P., Franik, S., Pouwer, A., and Farquhar, C. (2014). Minimally invasive surgical techniques versus open myomectomy for uterine fibroids. *Cochrane Database of Systematic Reviews*, (10). [6](#)
- Biber, P. and Strasser, W. (2003). The normal distributions transform: a new approach to laser scan matching. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, volume 3, pages 2743–2748 vol.3. [40](#)
- Bilal, O. and Samah, K. (2013). Segmental oriented liver surgery. In Abdeldayem, H., editor, *Hepatic Surgery*, chapter 9. IntechOpen, Rijeka. [15](#)
- Billings, S. D., Boctor, E. M., and Taylor, R. H. (2015). Iterative most-likely point registration (implp): A robust algorithm for computing optimal shape alignment. *PLOS ONE*, 10:1–45. [41](#)
- Bouré, L. (2005). General principles of laparoscopy. *Veterinary Clinics of North America: Food Animal Practice*, 21(1):227–249. [2](#)
- Bove, P., Asimakopoulos, A., Kim, F., and Vespasiani, G. (2009). Laparoscopic radical prostatectomy: a review. *International Brazilian Journal of Urology*, 35(2):125–139. [7](#)
- Caruso, R. P., Phillips, C. K., Kau, E., Taneja, S. S., and Stifelman, M. D. (2006). Robot assisted laparoscopic partial nephrectomy: Initial experience. *Journal of Urology*, 176(1):36–39. [7](#)
- Caruso, S., Patrìti, A., Roviello, F., De Franco, L., Franceschini, E., Coratti, A., and Ceccarelli, G. (2016). Laparoscopic and robot-assisted gastrectomy for gastric cancer: Current considerations. *World Journal of Gastroenterology*, 22(25):5694–5717. [5](#)
- Caruso, S. and Scatizzi, M. (2022). Laparoscopic gastrectomy for gastric cancer: has the time come for considered it a standard procedure? *Surgical Oncology*, 40. [5](#)

- Castellani, U. and Bartoli, A. (2012). *3D Shape Registration*, pages 221–264. Springer London. 38, 40
- Censi, A. (2008). An icp variant using a point-to-line metric. In *2008 IEEE International Conference on Robotics and Automation*, pages 19–25. 41
- Chang, W. and Zwicker, M. (2009). Range scan registration using reduced deformable models. *Computer Graphics Forum*, 28(2):447–456. 39
- Charalabopoulos, A., Lorenzi, B., Kordzadeh, A., Tang, C.-B., Kadiramanathan, S., and Jayanthi, N. V. (2017). Role of 3d in minimally invasive esophagectomy. *Langenbeck's Archives of Surgery*, 402(3):555–561. 4
- Cheema, M. N., Nazir, A., Sheng, B., Li, P., Qin, J., Kim, J., and Feng, D. D. (2019). Image-aligned dynamic liver reconstruction using intra-operative field of views for minimal invasive surgery. *IEEE Transactions on Biomedical Engineering*, 66(8):2163–2173. 22, 23, 24, 33, 34, 35
- Chen, L., Tang, W., John, N. W., Wan, T. R., and Zhang, J. J. (2018). Slam-based dense surface reconstruction in monocular minimally invasive surgery and its application to augmented reality. *Computer Methods and Programs in Biomedicine*, 158:135–146. 24, 25, 33, 35, 92
- Chen, X., M.D., R. G., and Fei, B. (2007). Automatic intensity-based 3D-to-2D registration of CT volume and dual-energy digital radiography for the detection of cardiac calcification. In *Medical Imaging 2007: Image Processing*, volume 6512, page 65120A. International Society for Optics and Photonics, SPIE. 46
- Chen, X., Varley, M., Shark, L.-K., Shentall, G., and Kirby, M. (2006). An extension of iterative closest point algorithm for 3d-2d registration for pre-treatment validation in radiotherapy. In *International Conference on Medical Information Visualisation - BioMedical Visualisation*, pages 3–8. 45
- Chen, X.-Z., Wang, S.-Y., Wang, Y.-S., Jiang, Z.-H., Zhang, W.-H., Liu, K., Yang, K., Chen, X.-L., Zhao, L.-Y., Qiu, M., Gou, H.-F., Zhou, Z.-G., and Hu, J.-K. (2017). Comparisons of short-term and survival outcomes of laparoscopy-assisted versus open total gastrectomy for gastric cancer patients. *Oncotarget*, 8(32):52366–52380. 9
- Chen, Y. and Medioni, G. (1992). Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155. Range Image Understanding. 39
- Chin, T.-J. and Suter, D. (2017). The maximum consensus problem: Recent algorithmic advances. In *Synthesis Lectures on Computer Vision*. 39
- Choi, G. H., Chong, J. U., Han, D. H., Choi, J. S., and Lee, W. J. (2017). Robotic hepatectomy: the korean experience and perspective. *Hepatobiliary Surgery and Nutrition*, 6(4):230–238. 8
- Chow, C. K., Tsui, H. T., and Lee, T. (2004). Surface registration using a dynamic genetic algorithm. *Pattern Recognition*, 37(1):105–117. 40
- Chui, H. and Rangarajan, A. (2003). A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141. Nonrigid Image Registration. 39

- Clements, L., Collins, J., Weis, J., Simpson, A., Kingham, T., Jarnagin, W., and Miga, M. (2017). Deformation correction for image guided liver surgery: An intraoperative fidelity assessment. *Surgery*, 162. [92](#)
- Collins, T., Chauvet, P., Debize, C., Pizarro, D., Bartoli, A., Canis, M., and Bourdel, N. (2017). A system for augmented reality guided laparoscopic tumour resection with quantitative ex-vivo user evaluation. In Peters, T., Yang, G.-Z., Navab, N., Mori, K., Luo, X., Reichl, T., and McLeod, J., editors, *Computer-Assisted and Robotic Endoscopy*, pages 114–126. Springer International Publishing. [92](#)
- Collins, T., Pizarro, D., Gasparini, S., Bourdel, N., Chauvet, P., Canis, M., Calvet, L., and Bartoli, A. (2021). Augmented reality guided laparoscopic surgery of the uterus. *IEEE Transactions on Medical Imaging*, 40(1):371–380. [11](#), [32](#), [33](#), [55](#)
- Crombez, N., Seulin, R., Morel, O., Fofi, D., and Demonceaux, C. (2018). Multimodal 2d image to 3d model registration via a mutual alignment of sparse and dense visual features. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6316–6322. [43](#), [45](#)
- Dassault Systèmes (2022). Abaqus. <https://www.3ds.com/products-services/simulia/products/abaqus/>. [Online; accessed 12-August-2022]. [83](#)
- Database Center for Life Science (DBCLS) (2015). Liver segment classification by couinaud. [https://commons.wikimedia.org/wiki/File:Liver\\_04\\_Couinaud\\_classification\\_anterior\\_view.png](https://commons.wikimedia.org/wiki/File:Liver_04_Couinaud_classification_anterior_view.png). [Online; accessed 17-July-2022]. [13](#)
- de Rooij, T., Sitarz, R., Busch, O., Besselink, M., and Abu Hilal, M. (2015). Technical aspects of laparoscopic distal pancreatectomy for benign and malignant disease: Review of the literature. *Gastroenterology Research and Practice*, 2015. [8](#)
- de Vries, J. (2014). Lean OpenGL: Coordinate Systems. <https://www.learnopengl.com/Getting-started/Coordinate-Systems/>. [Online; accessed 15-August-2022]. [52](#)
- Dey, J. and Napel, S. (2006). Targeted 2d/3d registration using ray normalization and a hybrid optimizer. *Medical Physics*, 33(12):4730–4738. [45](#)
- Dili, A. and Bertrand, C. (2017). Laparoscopic ultrasonography as an alternative to intraoperative cholangiography during laparoscopic cholecystectomy. *World Journal of Gastroenterology*, 23(29):5438–5450. [10](#)
- Doddamani, R., Hiremath, S., Ahmed, Z., and Surapaneni, L. (2018). Complications of laparoscopic surgery in general surgical practice and their management. *International Surgery Journal*, 5(4). [9](#)
- Driessen, S., Sandberg, E., Rodrigues, S., van Zwet, E., and Jansen, F. (2017). Identification of risk factors in minimally invasive surgery: a prospective multicenter study. *Surgical Endoscopy*, 31:2467–2473. [9](#)
- Duriez, C., Cotin, S., Lenoir, J., and Neumann, P. (2006). New approaches to catheter navigation for interventional radiology simulation. *Computer Aided Surgery*, 11(6):300–308. [25](#)



- Eglinton, T. W. and Frizelle, F. A. (2019). *Laparoscopic Sigmoid Colectomy for Diverticular Disease*, pages 285–306. Springer International Publishing. 6
- Espinel, Y., Calvet, L., Botros, K., Buc, E., Tilmant, C., and Bartoli, A. (2021). Using multiple images and contours for deformable 3d-2d registration of a preoperative ct in laparoscopic liver surgery. In de Bruijne, M., Cattin, P. C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., and Essert, C., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pages 657–666. Springer International Publishing. 19
- Espinel, Y., Özgür, E., Calvet, L., Le Roy, B., Buc, E., and Bartoli, A. (2020). Combining visual cues with interactions for 3d–2d registration in liver laparoscopy. *Annals of Biomedical Engineering*, 48(6):1712–1727. 19, 74, 92
- Estépar, R. S. J., Brun, A., and Westin, C.-F. (2004). Robust generalized total least squares iterative closest point registration. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, pages 234–241. 41
- Felippa, C. and Haugen, B. (2005). A unified formulation of small-strain corotational finite elements: I. theory. *Computer Methods in Applied Mechanics and Engineering*, 194(21):2285–2335. 25
- Feng, J., Shao, J., Jin, X., Peng, Q., and Forrest, A. R. (2006). Multiresolution free-form deformation with subdivision surface of arbitrary topology. *The Visual Computer*, 22(1):28–42. 59
- Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395. 39, 94
- Fukuda, T., Nada, H., Adachi, H., Shimizu, S., Takei, C., Sato, Y., Yabuki, N., and Motamedi, A. (2017). Integration of a structure from motion into virtual and augmented reality for architectural and urban simulation. In Çağdaş, G., Özkar, M., Gül, L. F., and Gürer, E., editors, *Computer-Aided Architectural Design. Future Trajectories*. Springer Singapore. 47
- Gadjiev, N., Pogosyan, R., Ostanin, M., Petrov, S., and Semenyakin, I. (2020). Laparoscopic partial nephrectomy using a technology of augmented reality: new vector of the development of operative urology? *Urologiia*, 5:37–40. 11
- Gao, Q. H., Wan, T. R., Tang, W., Chen, L., and Zhang, K. B. (2017). An improved augmented reality registration method based on visual slam. In *E-Learning and Games*, pages 11–19. Springer International Publishing. 49
- Gay-Bellile, V., Lothe, P., Bourgeois, S., Royer, E., and Collette, S. N. (2010). Augmented reality in large environments: Application to aided navigation in urban context. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 225–226. IEEE Computer Society. 48
- Gelfand, N., Guibas, L., and Pottmann, H. (2005). Robust global registration. In *Proceedings of the 3rd Eurographics Symposium on Geometry Processing*. 41

- Gelfand, N., Ikemoto, L., Rusinkiewicz, S., and Levoy, M. (2003). Geometrically stable sampling for the icp algorithm. In *Proceedings of the 3DIM*, volume 2003, pages 260–267. [41](#)
- Germain, T., Favelier, S., Cercueil, J.-P., Denys, A., Krausé, D., and Guiu, B. (2014). Liver segmentation: Practical tips. *Diagnostic and Interventional Imaging*, 95(11):1003–1016. [12](#)
- German Cancer Research Center (DKFZ) (2008). The medical imaging interaction toolkit (mitk). [https://www.mitk.org/wiki/The\\_Medical\\_Imaging\\_Interaction\\_Toolkit\\_\(MITK\)](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK)). [Online; accessed 1-August-2022]. [26](#), [50](#)
- Giannarou, S., Zhang, Z., and Yang, G.-Z. (2012). Deformable structure from motion by fusing visual and inertial measurement data. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4816–4821. [48](#)
- Gingold, J. A., Gueye, N.-A., and Falcone, T. (2018). Minimally invasive approaches to myoma management. *Journal of Minimally Invasive Gynecology*, 25(2):237–250. [6](#)
- Godard, C., Aodha, O. M., and Brostow, G. J. (2017). Unsupervised monocular depth estimation with left-right consistency. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6602–6611. [26](#)
- Grant, D., Bethel, J., and Crawford, M. (2012). Point-to-plane registration of terrestrial laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing*, 72:16–26. [41](#)
- Gray, D., Kim, E., Cotero, V. E., Siclovan, T. M., Zhang, R., Bajaj, A., LaPlante, N., Staudinger, V. P., Hehir, C. T., and Yazdanfar, S. (2012). Fluorescence image guided surgical instruments and contrast agents for intraoperative visualization of nerves. In *Conference on Lasers and Electro-Optics*. [10](#)
- Greenspan, M. and Godin, G. (2001). A nearest neighbor method for efficient icp. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pages 161–168. [41](#)
- Gressin, A., Mallet, C., Demantké, J., and David, N. (2013). Towards 3d lidar point cloud registration improvement using optimal neighborhood knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79:240–251. [41](#)
- Guha, D., Alotaibi, N. M., Nguyen, N., Gupta, S., McFaul, C., and Yang, V. X. D. (2017). Augmented reality in neurosurgery: A review of current concepts and emerging applications. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*, 44:235–245. [11](#)
- Hageman, D., Caillet, V., Kostohryz, J., and Madick, S. (2008). Laparoscopic-assisted colon surgery. *AORN Journal*, 88:403–416. [2](#)
- Hamabe, A., Ogino, T., Tanida, T., Noura, S., Morita, S., and Dono, K. (2019). Indocyanine green fluorescence-guided laparoscopic surgery, with omental appendices as fluorescent markers for colorectal cancer resection: a pilot study. *Surgical Endoscopy*, 33(2):669–678. [10](#)
- Han, J., Wang, F., Guo, Y., Zhang, C., and He, Y. (2015). An improved ransac registration algorithm based on region covariance descriptor. *2015 Chinese Automation Congress (CAC)*, pages 746–751. [39](#)

- Han, J., Yin, P., He, Y., and Gu, F. (2016). Enhanced icp for the registration of large-scale 3d environment models: An experimental study. *Sensors*, 16(2). 40
- Haouchine, N., Dequidt, J., Berger, M. O., and Cotin, S. (2013). Deformation-based augmented reality for hepatic surgery. *Studies in Health Technology and Informatics*, 184:182–188. 18, 24, 26, 33, 35, 55
- Haouchine, N., Roy, F., Untereiner, L., and Cotin, S. (2016). Using contours as boundary conditions for elastic registration during minimally invasive hepatic surgery. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 495–500. 92
- Hartley, R. and Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*. Cambridge University Press. 39, 47, 94
- Harvard Health Publishing (2022). Hysterectomy. <https://www.drugs.com/health-guide/hysterectomy.html>. [Online; accessed 1-October-2022]. 7
- Hasan, M. K., Calvet, L., Rabbani, N., and Bartoli, A. (2021). Detection, segmentation, and 3d pose estimation of surgical tools using convolutional neural networks and algebraic geometry. *Medical Image Analysis*, 70:101994. 94, 105, 107
- Himal, H. (2002). Minimally invasive (laparoscopic) surgery. *Surgical Endoscopy*, 16:1647–1652. 2
- Hipwell, J., Penney, G., McLaughlin, R., Rhode, K., Summers, P., Cox, T., Byrne, J., Noble, J., and Hawkes, D. (2003). Intensity-based 2-d - 3-d registration of cerebral angiograms. *IEEE Transactions on Medical Imaging*, 22(11):1417–1426. 45
- Hockstein, N. G., O'Malley Jr, B. W., and Weinstein, G. S. (2006). Assessment of intraoperative safety in transoral robotic surgery. *The Laryngoscope*, 116(2):165–168. 3
- Jans, H.-S., Syme, A. M., Rathee, S., and Fallone, B. G. (2006). 3d interfractional patient position verification using 2d-3d registration of orthogonal images. *Medical Physics*, 33(5):1420–1439. 45
- Jha, S. K., Dhamija, N., Kumar, A., and Rawat, S. (2020). Robotic-assisted esophagectomy: A literature review and our experience at a tertiary care centre. *Laparoscopic, Endoscopic and Robotic Surgery*, 3(3):74–79. 4
- Jia, C., Li, H., Wen, N., Chen, J., Wei, Y., and Li, B. (2018). Laparoscopic liver resection: a review of current indications and surgical techniques. *Hepatobiliary Surgery and Nutrition*, 7(4). 7, 12
- Jian, B. and Vemuri, B. C. (2011). Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645. 38
- Johnston, W. K. and Wolf, J. S. (2005). Laparoscopic partial nephrectomy: Technique, oncologic efficacy, and safety. *Current Urology Reports*, 6(1):19–28. 7
- Jud, L., Fotouhi, J., Andronic, O., Aichmair, A., Osgood, G., Navab, N., and Farshad, M. (2020). Applicability of augmented reality in orthopedic surgery – a systematic review. *BMC Musculoskeletal Disorders*, 21. 11

- Kamen, A., Bloch, P., Wein, W., Svatos, M., and Sauer, F. (2006). Automatic registration of portal images and volumetric ct for patient positioning in radiation therapy. *Medical Image Analysis*, 10(1):96–112. [45](#)
- Kang, X., Azizian, M., Wilson, E., Wu, K., Martin, A. D., Kane, T. D., Peters, C. A., Cleary, K., and Shekhar, R. (2014). Stereoscopic augmented reality for laparoscopic surgery. *Surgical Endoscopy*, 28(7):2227–2235. [11](#)
- Kaplan, J., Lee, Z., Eun, D., and Reese, A. (2016). Complications of minimally invasive surgery and their management. *Current Urology Reports*, 17(6):47. [9](#)
- Kauffmann, E. F., Napoli, N., Menonna, F., Cacace, C., Genovese, V., Vistoli, F., and Boggi, U. (2020). Robot-assisted spleen preserving distal pancreatectomy: case report. *Annals of Laparoscopic and Endoscopic Surgery*, 6(0). [8](#)
- Khronos Group (2022). OpenGL. <https://www.opengl.org/>. [Online; accessed 25-August-2022]. [51](#)
- Kikinis, R. (2012). 3d slicer. <https://www.slicer.org/>. [Online; accessed 1-August-2022]. [50](#)
- Koguciuk, D. (2017). Parallel ransac for point cloud registration. *Foundations of Computing and Decision Sciences*, 42(3):203–217. [39](#)
- Koo, B., Özgür, E., Le Roy, B., Buc, E., and Bartoli, A. (2017). Deformable registration of a preoperative 3d liver volume to a laparoscopy image using contour and shading cues. In *Medical Image Computing and Computer Assisted Intervention MICCAI 2017*, pages 326–334. [19](#), [29](#), [30](#), [31](#), [33](#), [34](#), [36](#), [52](#), [54](#), [58](#), [59](#), [74](#), [92](#), [94](#)
- Koo, B., Robu, M. R., Allam, M., Pfeiffer, M., Thompson, S., Gurusamy, K., Davidson, B., Speidel, S., Hawkes, D., Stoyanov, D., and Clarkson, M. J. (2022). Automatic, global registration in laparoscopic liver surgery. *International Journal of Computer Assisted Radiology and Surgery*, 17(1):167–176. [29](#), [30](#), [32](#), [33](#), [36](#)
- Labrunie, M., Ribeiro, M., Mourthadhoi, F., Tilmant, C., Le Roy, B., Buc, E., and Bartoli, A. (2022). Automatic preoperative 3d model registration in laparoscopic liver resection. *International Journal of Computer Assisted Radiology and Surgery*, 17(8):1429–1436. [29](#), [32](#), [33](#), [36](#)
- Lamata, P., Lamata, F., Sojar, V., Makowski, P., Massoptier, L., Casciaro, S., Ali, W., Stüdeli, T., Declerck, J., Elle, O. J., and Edwin, B. (2010). Use of the resection map system as guidance during hepatectomy. *Surgical Endoscopy*, 24(9):2327–2337. [58](#)
- Lazzara, C., Navarra, G., Lazzara, S., Barbera, A., Saitta, C., Raimondo, G., Latteri, S., and Curro, G. (2017). Does the margin width influence recurrence rate in liver surgery for hepatocellular carcinoma smaller than 5 cm? *European Review for Medical and Pharmacological Sciences*, 21(3). [12](#)
- Lee, W.-J., Chan, C.-P., and Wang, B.-Y. (2013). Recent advances in laparoscopic surgery. *Asian Journal of Endoscopic Surgery*, 6(1):1–8. [3](#)

- Leonardis, R., Duvvuri, U., and Mehta, D. (2013). Transoral robotic surgery in the pediatric patient. *Operative Techniques in Otolaryngology-Head and Neck Surgery*, 24(2):111–114. [3](#)
- Leordeanu, M. and Hebert, M. (2005). A spectral technique for correspondence problems using pairwise constraints. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1482–1489 Vol. 2. [27](#)
- Levin, D. (2004). Mesh-independent surface interpolation. In Brunnett, G., Hamann, B., Müller, H., and Linsen, L., editors, *Geometric Modeling for Scientific Visualization*, pages 37–49. [24](#)
- Li, F., Stoddart, D., and Hitchens, C. (2017). Method to automatically register scattered point clouds based on principal pose estimation. *Optical Engineering*, 56:044107. [41](#), [42](#)
- Li, H. and Hartley, R. (2007). The 3d-3d registration problem revisited. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. [38](#)
- Li, J., Hu, Q., and Ai, M. (2020a). Gesac: Robust graph enhanced sample consensus for point cloud registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167:363–374. [40](#)
- Li, P., Wang, R., Wang, Y., and Tao, W. (2020b). Evaluation of the icp algorithm in 3d point cloud registration. *IEEE Access*, 8:68030–68048. [41](#)
- Li, S., Wang, J., Liang, Z., and Su, L. (2016). Tree point clouds registration using an improved icp algorithm based on kd-tree. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4545–4548. [41](#)
- Li, Y., Kong, Q., Wei, H., and Wang, Y. (2021). Comparison of the complications between minimally invasive surgery and open surgical treatments for early-stage cervical cancer: A systematic review and meta-analysis. *PLOS ONE*, 16(7):1–16. [9](#)
- Liang, S., Hameed, U., and Jayaraman, S. (2014). Laparoscopic pancreatectomy: Indications and outcomes. *World Journal of Gastroenterology*, 20(39). [8](#)
- Litynski, G. (1999). Profiles in laparoscopy: Mouret, Dubois, and Perissat: the laparoscopic breakthrough in Europe (1987-1988). *Journal of the Society of Laparoendoscopic Surgeons*, 3(2):163–167. [2](#)
- Liu, J., Xie, Y., Gu, S., and Chen, X. (2020). A slam-based mobile augmented reality tracking registration algorithm. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(01):2054005. [49](#)
- Liu, J., Zhu, J., Yang, J., Meng, X., and Zhang, H. (2016). Three-dimensional point cloud registration based on ICP algorithm employing K-D tree optimization. In Falco, C. M. and Jiang, X., editors, *Eighth International Conference on Digital Image Processing (ICDIP 2016)*, volume 10033, page 100334D. International Society for Optics and Photonics, SPIE. [41](#)
- Liu, R., Liu, Q., Zhao, Z.-M., Tan, X.-L., Gao, Y.-X., and Zhao, G.-D. (2017). Robotic versus laparoscopic distal pancreatectomy: A propensity score-matched study. *Journal of Surgical Oncology*, 116(4):461–469. [8](#)

- Livyatan, H., Yaniv, Z., and Joskowicz, L. (2003). Gradient-based 2-d/3-d rigid registration of fluoroscopic x-ray to ct. *IEEE Transactions on Medical Imaging*, 22(11):1395–1406. [45](#), [46](#)
- Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3d surface construction algorithm. *Computer Graphics*, 21(4):163–169. [50](#)
- Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, page 674–679. [25](#)
- Luo, H., Yin, D., Zhang, S., Xiao, D., He, B., Meng, F., Zhang, Y., Cai, W., He, S., Zhang, W., Hu, Q., Guo, H., Liang, S., Zhou, S., Liu, S., Sun, L., Guo, X., Fang, C., Liu, L., and Jia, F. (2020). Augmented reality navigation for liver resection with a stereoscopic laparoscope. *Computer Methods and Programs in Biomedicine*, 187. [11](#), [24](#), [25](#), [26](#), [33](#), [35](#)
- Magnusson, M. (2009). *The Three-Dimensional Normal-Distributions Transform — an Efficient Representation for Registration, Surface Analysis, and Loop Detection*. PhD thesis, Orebro University. [40](#)
- Majbar, A., Aalala, M., Elalaoui, M., Sabbah, F., Raiss, M., Hrorra, A., and Ahallat, M. (2014). Asymptomatic intra-peritoneal rupture of hydatid cyst of the liver: case report. *BMC Research Notes*, 7. [10](#)
- Majno, P., Mentha, G., Toso, C., Morel, P., Peitgen, H. O., and Fasel, J. H. (2014). Anatomy of the liver: an outline with three levels of complexity—a further step towards tailored territorial liver resections. *Journal of Hepatology*, 60(3):654–662. [13](#)
- Malbrain, M., Roberts, D., Laet, I., De Waele, J., Sugrue, M., Schachtrupp, A., Duchesne, J., van Ramshorst, G., Keulenaer, B., Kirkpatrick, A., Ahmadi-Noorbakhsh, S., Mulier, J., Ivatury, R., Pracca, F., Wise, R., and Pelosi, P. (2014). The role of abdominal compliance, the neglected parameter in critically ill patients - a consensus review of 16. part 1: Definitions and pathophysiology. *Anaesthesiology intensive therapy*, 46:392–405. [97](#)
- Markelj, P., Tomaževič, D., Likar, B., and Pernuš, F. (2012). A review of 3d/2d registration methods for image-guided interventions. *Medical Image Analysis*, 16(3):642–661. Computer Assisted Interventions. [43](#), [44](#)
- Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571. [26](#)
- Modrzejewski, R., Collins, T., Seeliger, B., Bartoli, A., Hostettler, A., and Marescaux, J. (2019). An in vivo porcine dataset and evaluation methodology to measure soft-body laparoscopic liver registration accuracy with an extended algorithm that handles collisions. *International Journal of Computer Assisted Radiology and Surgery*, 14(7):1237–1245. [22](#), [23](#), [33](#), [34](#), [35](#)
- Moghari, M. H., Ma, B., and Abolmaesumi, P. (2008). A theoretical comparison of different target registration error estimators. In Metaxas, D., Axel, L., Fichtinger, G., and Székely, G., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*, pages 1032–1040. Springer Berlin Heidelberg. [64](#)

- Mooser, J., You, S., Neumann, U., and Wang, Q. (2009). Applying robust structure from motion to markerless augmented reality. In *2009 Workshop on Applications of Computer Vision (WACV)*, pages 1–8. [48](#)
- Müller, M., Heidelberger, B., Hennix, M., and Ratcliff, J. (2007). Position based dynamics. *Journal of Visual Communication and Image Representation*, 18(2):109–118. [53](#), [54](#)
- Munbodh, R., Jaffray, D. A., Moseley, D. J., Chen, Z., Knisely, J. P. S., Cathier, P., and Duncan, J. S. (2006). Automated 2d-3d registration of a radiograph and a cone beam ct using line-segment enhancement). *Medical Physics*, 33(5):1398–1411. [45](#)
- Munoz-Montoya, F., Juan, M.-C., Mendez-Lopez, M., Molla, R., Abad, E., and Fidalgo, C. (2021). Slam-based augmented reality for the assessment of short-term spatial memory. a comparative study of visual versus tactile stimuli. *PLOS ONE*, 16(2):1–30. [50](#)
- Mur-Artal, R., Montiel, J. M. M., and Tardós, J. D. (2015). Orb-slam: A versatile and accurate monocular slam system. *IEEE Transactions on Robotics*, 31(5):1147–1163. [49](#)
- Nagaya, T., Nakamura, Y. A., Choyke, P. L., and Kobayashi, H. (2017). Fluorescence-guided surgery. *Frontiers in Oncology*, 7(314). [9](#)
- Nicolau, S., Soler, L., Mutter, D., and Marescaux, J. (2011). Augmented reality in laparoscopic surgical oncology. *Surgical Oncology*, 20(3):189–201. Special Issue: Education for Cancer Surgeons. [58](#)
- Njei, B., Rotman, Y., Ditah, I., and Lim, J. K. (2015). Emerging trends in hepatocellular carcinoma incidence and mortality. *Hepatology*, 61(1):191–199. [7](#)
- Ocuin, L. M. and Tsung, A. (2016). Minimally invasive hepatic surgery. *Surgical Clinics of North America*, 96(2):299–313. [14](#)
- Okunrintemi, V., Gani, F., and Pawlik, T. (2016). National trends in postoperative outcomes and cost comparing minimally invasive versus open liver and pancreatic surgery. *Gastrointestinal Surgery*, 20:1836–1843. [9](#)
- Otake, Y., Wang, A. S., Uneri, A., Kleinszig, G., Vogt, S., Aygun, N., fu L Lo, S., Wolinsky, J.-P., Gokaslan, Z. L., and Siewerdsen, J. H. (2015). 3d-2d registration in mobile radiographs: algorithm development and preliminary clinical evaluation. *Physics in Medicine and Biology*, 60(5):2075–2090. [43](#)
- Özgür, E., Koo, B., Le Roy, B., Buc, E., and Bartoli, A. (2018). Preoperative liver registration for augmented monocular laparoscopy using backward-forward biomechanical simulation. *International Journal of Computer Assisted Radiology and Surgery*, 13(10):1629–1640. [18](#), [52](#), [58](#)
- Patzak, M., Porzner, M., Oetzuerk, S., Mason, R. A., Wilhelm, M., Graeter, T., Kratzer, W., Haenle, M. M., Akinli, A. S., , and the EMIL Study Group (2014). Assessment of liver size by ultrasonography. *Journal of Clinical Ultrasound*, 42(7):399–404. [15](#)
- Penney, G., Weese, J., Little, J., Desmedt, P., Hill, D., and hawkes, D. (1998). A comparison of similarity measures for use in 2-d-3-d medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–595. [45](#)

- Peterlík, I., Duriez, C., and Cotin, S. (2012). Modeling and real-time simulation of a vascularized liver tissue. In Ayache, N., Delingette, H., Golland, P., and Mori, K., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, pages 50–57. [25](#)
- Pizarro, D. and Bartoli, A. (2010). Feature-based deformable surface detection with self-occlusion reasoning. *International Journal of Computer Vision*, 97:54–70. [78](#)
- Plantefève, R., Peterlik, I., Haouchine, N., and Cotin, S. (2016). Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery. *Annals of Biomedical Engineering*, 44(1):139–153. [92](#)
- Puerto-Souza, G. A., Cadeddu, J. A., and Mariottini, G.-L. (2014). Toward long-term and accurate augmented-reality for monocular endoscopic videos. *IEEE Transactions on Biomedical Engineering*, 61(10):2609–2620. [55](#)
- Rabbani, N., Calvet, L., Espinel, Y., Roy, B. L., Ribeiro, M., Buc, E., and Bartoli, A. (2022). A methodology and clinical dataset with ground-truth to evaluate registration accuracy quantitatively in computer-assisted laparoscopic liver resection. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 10(4):441–450. [11](#)
- Rashid, P., Goad, J., Aron, M., Gianduzzo, T., and Gill, I. S. (2008). Laparoscopic partial nephrectomy: Integration of an advanced laparoscopic technique. *ANZ Journal of Surgery*, 78(6):471–475. [7](#)
- Robinson, T. and Stiegmann, G. (2004). Minimally invasive surgery. *Endoscopy*, 36(1):48–51. [2](#)
- Robu, M. R., Ramalhinho, J., Thompson, S., Gurusamy, K., Davidson, B., Hawkes, D., Stoyanov, D., and Clarkson, M. J. (2018). Global rigid registration of ct to video in laparoscopic liver surgery. *International Journal of Computer Assisted Radiology and Surgery*, 13(6):947–956. [24](#), [26](#), [27](#), [28](#), [33](#), [35](#), [92](#)
- Rosenfeld, J. V. (1996). Minimally invasive surgery. *Australian and New Zealand Journal of Surgery*, 66(8):553–559. [3](#)
- Rusinkiewicz, S. and Levoy, M. (2001). Efficient variants of the icp algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pages 145–152. [40](#)
- Russakoff, D. B., Rohlfing, T., Adler, J. R., and Maurer, C. R. (2005). Intensity-based 2d-3d spine image registration incorporating a single fiducial marker<sup>1</sup>. *Academic Radiology*, 12(1):37–50. [46](#)
- Rusu, R. B., Blodow, N., Marton, Z. C., and Beetz, M. (2008). Aligning point cloud views using persistent feature histograms. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3384–3391. [41](#)
- Sacht, L. K., Vouga, E., and Jacobson, A. (2015). Nested cages. *ACM Transactions on Graphics (TOG)*, 34:1–14. [59](#), [60](#)
- Sattler, T., Leibe, B., and Kobbelt, L. (2011). Fast image-based localization using direct 2d-to-3d matching. In *2011 International Conference on Computer Vision*, pages 667–674. [44](#)



- Schmelzle, M., Krenzien, F., Schöning, W., and Pratschke, J. (2020). Laparoscopic liver resection: indications, limitations, and economic aspects. *Langenbeck's Archives of Surgery*, 405(6):725–735. 14
- Sederberg, T. and Parry, S. (1986). Free-form deformation of solid geometric models. In *ACM Siggraph Computer Graphics*, volume 20, pages 151–160. 59
- Sheasby, T. (2011). Liver diagram. [https://commons.wikimedia.org/wiki/File:Liver\\_Diagram.svg](https://commons.wikimedia.org/wiki/File:Liver_Diagram.svg). [Online; accessed 17-July-2022]. 12
- Sheetz, K. H., Claflin, J., and Dimick, J. B. (2020). Trends in the Adoption of Robotic Surgery for Common Surgical Procedures. *JAMA Network Open*, 3(1). 15
- Shen, J., Zemiti, N., Taoum, C., Aiche, G., Dillenseger, J. L., Rouanet, P., and Poignet, P. (2020). Transrectal ultrasound image-based real-time augmented reality guidance in robot-assisted laparoscopic rectal surgery: a proof-of-concept study. *International Journal of Computer Assisted Radiology and Surgery*, 15(3):531–543. 11
- Shibata, K., Kawamura, H., Ichikawa, N., Shibuya, K., Yoshida, T., Ohno, Y., Homma, S., and Take-tomi, A. (2018). Laparoscopic total gastrectomy for advanced gastric cancer in a patient with situs inversus totalis. *Asian Journal of Endoscopic Surgery*, 11(1):39–42. 5
- Siddaiah-Subramanya, M., Tiang, K. W., and Nyandowe, M. (2017). A new era of minimally invasive surgery: Progress and development of major technical innovations in general surgery over the last decade. *The Surgery Journal*, 3(4):e163–e166. 9
- Silva, L., Bellon, O., Gotardo, P., and Boyer, K. (2003). Range image registration using enhanced genetic algorithms. In *IEEE International Conference on Image Processing*, volume 2, pages II – 711. 40
- Soler, L., Nicolau, S., Pessaux, P., Mutter, D., and Marescaux, J. (2014). Real-time 3d image reconstruction guidance in liver resection surgery. *Hepatobiliary Surgery and Nutrition*, 3(2). 24, 28, 29, 33, 35
- St John, A., Caturegli, I., Kubicki, N., and Kavic, S. (2020). The rise of minimally invasive surgery: 16 year analysis of the progressive replacement of open surgery with laparoscopy. *Journal of the Society of Laparoendoscopic Surgeons*, 24(4). 9
- Stachniss, C., Leonard, J. J., and Thrun, S. (2016). *Simultaneous Localization and Mapping*, pages 1153–1176. Springer International Publishing. 48
- Striewski, P. and Wirth, B. (2022). Elastic 3d–2d image registration. *Journal of Mathematical Imaging and Vision*, 64(5):443–462. 46
- Sun, M. R. M., Brennan, D. D., Kruskal, J. B., and Kane, R. A. (2010). Intraoperative ultrasonography of the pancreas. *RadioGraphics*, 30(7):1935–1953. 10
- Sun, V. and Fong, Y. (2017). Minimally invasive cancer surgery: Indications and outcomes. *Seminars in Oncology Nursing*, 33(1):23–36. 3, 4

- Takeuchi, E. and Tsubouchi, T. (2006). A 3-d scan matching using improved 3-d normal distributions transform for mobile robotic mapping. *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3068–3073. 40
- Tao, W., Hua, X., Yu, K., He, X., and Chen, X. (2018). An improved point-to-plane registration method for terrestrial laser scanning data. *IEEE Access*, 6:48062–48073. 41
- The Blender Foundation (1995). Blender. <https://www.blender.org/>. [Online; accessed 1-August-2022]. 50
- Thiruchelvam, N., Lee, S., and Chiow, A. (2021). Patient and port positioning in laparoscopic liver resections. *Hepatoma Research*, 7(22). 7
- Thompson, S., Hu, M., Johnsen, S. F., Gurusamy, K., Davidson, B., and Hawkes, D. (2011). Towards image guided laparoscopic liver surgery, defining the system requirement. *LIVIM (Living Imaging) 2011*. 74
- Thompson, S., Tetz, J., Song, Y., Johnsen, S., Stoyanov, D., Ourselin, S., Gurusamy, K., Schneider, C., Davidson, B., Hawkes, D., and Clarkson, M. J. (2015). Accuracy validation of an image guided laparoscopy system for liver resection. In *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9415, pages 52 – 63. SPIE. 24, 27, 28, 33, 35, 92
- Tomaževič, D., Likar, B., and Pernuš, F. (2001). Gradient-based registration of 3d mr and 2d x-ray images. *International Congress Series*, 1230:338–345. Computer Assisted Radiology and Surgery. 46
- Torr, P. and Zisserman, A. (2000). Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156. 39
- Toth, D., Miao, S., Kurzendorfer, T., Rinaldi, C. A., Liao, R., Mansi, T., Rhode, K., and Mountney, P. (2018). 3d/2d model-to-image registration by imitation learning for cardiac procedures. *International Journal of Computer Assisted Radiology and Surgery*, 13(8):1141–1149. 46
- Touijer, A. and Guillonneau, B. (2004). Laparoscopic radical prostatectomy. *Urologic Oncology: Seminars and Original Investigations*, 22(2):133–138. 7
- Troisi, R., Montalti, R., Smeets, P., Van Huysse, J., Van Vlierberghe, H., Colle, I., De Gendt, S., and de Hemptinne, B. (2008). The value of laparoscopic liver surgery for solid benign hepatic tumors. *Surgical Endoscopy*, 22(1):38–44. 14
- Tsai, F. S., Johnson, D., Francis, C. S., and Cho, S.-H. (2010). Fluidic lens laparoscopic zoom camera for minimally invasive surgery. *Journal of Biomedical Optics*, 15(3):1–3. 15
- Unity (2022). Unity Technologies. <https://www.unity.com/>. [Online; accessed 25-August-2022]. 51
- Urade, T., Sawa, H., Iwatani, Y., Abe, T., Fujinaka, R., Murata, K., Mii, Y., Man-i, M., Oka, S., and Kuroda, D. (2020). Laparoscopic anatomical liver resection using indocyanine green fluorescence imaging. *Asian Journal of Surgery*, 43(1):362–368. 10

- Visual Computing Lab - ISTI-CNR (2005). Meshlab. <https://www.meshlab.net/>. [Online; accessed 1-August-2022]. 50
- Vitrani, M.-A., Marx, A., Iordache, R., Muller, S., and Morel, G. (2015). Robot guidance of an ultrasound probe toward a 3d region of interest detected through x-ray mammography. *International Journal of Computer Assisted Radiology and Surgery*, 10(12):1893–1903. 41
- Vávra, P., Roman, J., Zonča, P., Ihnát, P., Němec, M., Kumar, J., Habib, N., and El-Gendi, A. (2017). Recent development of augmented reality in surgery: A review. *Journal of Healthcare Engineering*, 2017. 16
- Wabitsch, S., Kästner, A., Haber, P. K., Benzing, C., Krenzien, F., Andreou, A., Kamali, C., Lenz, K., Pratschke, J., and Schmelzle, M. (2019). Laparoscopic versus open liver resection for benign tumors and lesions: A case matched study with propensity score matching. *Journal of Laparoscopic & Advanced Surgical Techniques*, 29(12):1518–1525. 14
- Wang, J.-M., Li, J.-F., Yuan, G.-D., and He, S.-Q. (2021). Robot-assisted versus laparoscopic minor hepatectomy: A systematic review and meta-analysis. *Medicine*, 100(17). 15
- Watanabe, J., Ota, M., Suwa, Y., Ishibe, A., Masui, H., and Nagahori, K. (2015). Real-time indocyanine green fluorescence imaging-guided complete mesocolic excision in laparoscopic flexural colon cancer surgery. *Diseases of the Colon and Rectum*, 59(7):701–705. 10
- Weierstrass Institute for Applied Analysis and Stochastics (2015). Tetgen: A quality tetrahedral mesh generator and a 3d delaunay triangulator. <https://www.wias-berlin.de/software/index.jsp?id=TetGen>. [Online; accessed 19-July-2022]. 22
- Wein, W., Röper, B., and Navab, N. (2005). 2d/3d registration based on volume gradients. In *SPIE Medical Imaging*. 46
- Wong, K., Yee, H. M., Xavier, B. A., and Grillone, G. A. (2018). Applications of augmented reality in otolaryngology: A systematic review. *Otolaryngology–Head and Neck Surgery*, 159(6):956–967. 11
- Xian, Y., Xiao, J., Wang, Y., Shan, M., and Zhou, C. (2016). A review of fine registration for 3d point clouds. In *Proceedings of the 2016 5th International Conference on Advanced Materials and Computer Science*, pages 108–113. Atlantis Press. 40
- Xiang, L., Xiao, L., Li, J., Chen, J., Fan, Y., and Zheng, S. (2015). Safety and feasibility of laparoscopic hepatectomy for hepatocellular carcinoma in the posterosuperior liver segments. *World Journal of Surgery*, 39(5):1202–1209. 14
- Xin, M., Li, B., Yan, X., Chen, L., and Wei, X. (2018). A robust cloud registration method based on redundant data reduction using backpropagation neural network and shift window. *Review of Scientific Instruments*, 89(2):024704. 43
- Xue, S., Zhang, Z., Lv, Q., Meng, X., and Tu, X. (2019). Point cloud registration method for pipeline workpieces based on PCA and improved ICP algorithms. *IOP Conference Series: Materials Science and Engineering*, 612(3):032188. 42

- Yang, J., Li, H., Campbell, D., and Jia, Y. (2016). Go-icp: A globally optimal solution to 3d icp point-set registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11):2241–2254. [26](#)
- Yang, J., Zhang, Q., Xiao, Y., and Cao, Z. (2017). Toldi: An effective and robust approach for 3d local shape description. *Pattern Recognition*, 65:175–187. [26](#)
- Yasui, K. and Shimizu, Y. (2005). Surgical treatment for metastatic malignancies. anatomical resection of liver metastasis: indications and outcomes. *International Journal of Clinical Oncology*, 10(2):86–96. [15](#)
- Yoon, Y.-I., Kim, K.-H., Kang, S.-H., Kim, W.-J., Shin, M.-H., Lee, S.-K., Jung, D.-H., Park, G.-C., Ahn, C.-S., Moon, D.-B., Ha, T.-Y., Song, G.-W., Hwang, S., and Lee, S.-G. (2017). Pure laparoscopic versus open right hepatectomy for hepatocellular carcinoma in patients with cirrhosis: A propensity score matched analysis. *Annals of Surgery*, 265(5):856–863. [14](#)
- Yoshida, H., Taniai, N., Yoshioka, M., Hirakata, A., Kawano, Y., Shimizu, T., Ueda, J., Takata, H., Nakamura, Y., and Mamada, Y. (2019). Current status of laparoscopic hepatectomy. *Journal of Nippon Medical School*, 86(4):201–206. [15](#)
- Yu, Z., Feng, C., Liu, M.-Y., and Ramalingam, S. (2017). Casenet: Deep category-aware semantic edge detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1761–1770. [31](#)
- Zhang, X.-L., Liu, R.-F., Zhang, D., Zhang, Y.-S., and Wang, T. (2017). Laparoscopic versus open liver resection for colorectal liver metastases: A systematic review and meta-analysis of studies with propensity score-based analysis. *International Journal of Surgery*, 44:191–203. [14](#)
- Zheng, G. (2007). Unifying energy minimization and mutual information maximization for robust 2d/3d registration of x-ray and ct images. In Hamprecht, F. A., Schnörr, C., and Jähne, B., editors, *Pattern Recognition*, pages 547–557, Berlin, Heidelberg. Springer Berlin Heidelberg. [45](#), [46](#)
- Zheng, Y., Yang, H., Wang, H., Kang, K., Zhang, W., Ma, G., and Du, S. (2019). Fluorescence-guided surgery in cancer treatment: current status and future perspectives. *Annals of Translational Medicine*, 7. [9](#)
- Zhong, J.-H., Peng, N.-F., You, X.-M., Ma, L., Xiang, X., Wang, Y.-Y., Gong, W.-F., Wu, F.-X., Xiang, B.-D., and Li, L.-Q. (2017). Tumor stage and primary treatment of hepatocellular carcinoma at a large tertiary hospital in china: A real-world study. *Oncotarget*, 8(11):18296–18302. [19](#), [74](#), [97](#)
- Zhong, Z., Guo, X., Cai, Y., Yang, Y., Wang, J., Jia, X., and Mao, W. (2016). 3d-2d deformable image registration using feature-based nonuniform meshes. *BioMed Research International*, 2016. [44](#)
- Zhu, M., Liu, F., Chai, G., Pan, J., Jiang, T., Lin, L., Xin, Y., Zhang, Y., and Li, Q. (2017). A novel augmented reality system for displaying inferior alveolar nerve bundles in maxillofacial surgery. *Scientific Reports*, 7. [11](#)
- Özgür, E., Lafont, A., and Bartoli, A. (2017). Visualizing in-organ tumors in augmented monocular laparoscopy. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 46–51. [92](#)







# Appendix A

## Acronyms

**AALD** Alcohol-associated Liver Disease. [13](#)

**ABSOR** Absolute Orientation. [39](#)

**AR** Augmented Reality. [10](#)

**BLT** Benign Liver Tumours. [14](#)

**CNN** Convolutional Neural Network. [26](#)

**CRLM** Colorectal Liver Metastases. [14](#)

**CT** Computed Tomography. [11](#)

**FBDS** Feature-Based Deformable Surface Detection. [78](#)

**FGS** Fluorescence-guided Surgery. [9](#)

**FoV** Field-of-View. [15](#)

**GUI** Graphical User Interface. [58](#)

**HCC** Hepatocellular Carcinoma. [7](#), [14](#)

**ICG** Indocyanine Green. [9](#)

**ICP** Iterative Closest Point. [24](#), [40](#)

**LLR** Laparoscopic Liver Resection. [7](#)

**LUS** Laparoscopic Ultrasound. [10](#)

**MIS** Minimally Invasive Surgery. [2](#)

**MR** Magnetic Resonance. [11](#)

**MV-B** Multi-View Rigid Base. [74](#)



- MV-B\*** Variant of Multi-View Rigid Base. [74](#)
- MV-C** Multi-View Rigid with Correspondences. [74](#)
- MV-D** Multi-View Deformable with Correspondences. [74](#)
- NAFLD** Non-alcoholic Fatty Liver Disease. [13](#)
- OR** Operating Room. [62](#)
- PBD** Position-Based Dynamics. [30](#), [53](#), [59](#)
- PCA** Principal Component Analysis. [42](#)
- PLA** Polylactic Acid. [83](#)
- PnP** Perspective-n-Point. [31](#)
- RANSAC** Random Sample Consensus. [27](#), [39](#)
- RMS** Root Mean Square. [31](#)
- SfM** Structure from Motion. [22](#), [47](#)
- SIFT** Scale-invariant Feature Transform. [23](#)
- SLAM** Simultaneous Localisation and Mapping. [24](#), [48](#)
- SURF** Speeded Up Robust Features. [25](#)
- SVB** Single-View Biomechanical. [19](#), [58](#)
- SVH** Single-View Hybrid. [58](#)
- SVR** Single-View Rigid. [58](#)
- TRE** Target Registration Error. [22](#)
- TVE** Target Visualisation Error. [28](#)
- US** Ultrasound. [10](#)

# Appendix B

## Notation

$A$  Global anisometry. 77

$a$  Tumour model used in the geometric model. 97

$\alpha_g$  Anisometry measure for view combination  $g$ . 77

$B$  Number of phantoms used to compute anisometry. 77

$b$  Index of phantom used to compute anisometry. 77

$\beta$  Baricentre point used to solve the inter-image correspondence constraint. 78

$c_{image}$  Image contours. 54

$C_{model}$  Model contours in 3D. 54

$c_{model}$  Model contours projected in 2D. 54

$\mathcal{C}$  Set of 2D and 3D contours for multiple views. 76

$\delta_b$  Mean of anisometry measures for phantom  $b$ . 77

$d_{keyhole}$  Distance between keyholes in geometric model. 97

$d_{liver}$  Distance between liver and abdominal wall in geometric model. 97

$E$  Error function for registration. 39

$\epsilon$  Minimum distance threshold for registration. 54

$\eta$  Warp transformation function. 78

$f_x$  Focal length in x axis. 51

$f_y$  Focal length in y axis. 51

$G$  Number of view combinations per phantom to compute anisometry. 77

$g$  Index of view combination used to compute anisometry. 77

- 
- $\gamma$  Noise magnitude to simulate keyhole uncertainty. 98
- $H$  Number of vertex pairs used to compute anisometry. 77
- $h$  Index of vertex pair used to compute anisometry. 77
- $I$  Laparoscopic image. 54
- $i$  Iteration counter. 54
- $\mathcal{I}$  Set of input images for registration. 76
- $I_{Xray}$  X-ray image. 45
- $J$  Liver surface that intersects the keyhole-aware tumour projection. 100
- $K$  3D-2D projection matrix. 52
- $k$  Camera centre in 3D. 100
- $l$  SfM scale. 76
- $L_n$  Registered model in view  $n$ . 76
- $M$  Moving pointcloud/model in registration. 38
- $M_0$  Input preoperative model in registration. 54, 76
- $m_{clip}$  Model in clip coordinated. 52
- max\_iter Maximum number of registration iterations. 54
- $M_{CT}$  CT volume. 45
- $N$  Number of views in registration. 76
- $n$  Counter of views for registration. 76
- $\Omega$  Keypoint correspondences used to compute warp functions  $\eta$ . 81
- $\omega$  Upper interval of uniform distribution. 98
- $P$  3D points. 44
- $p$  2D points. 44
- $\Phi$  Number of points in the moving pointcloud  $M$ . 38
- $\pi$  Projection function. 44
- $\pi^\dagger$  Back-projection function. 44
- $P_{image}$  Image contour correspondences in 3D. 54

- $p_{image}$  Image contour correspondences in 2D. 54
- $\Psi$  Number of points in the reference pointcloud  $S$ . 38
- $\psi$  Lower interval of uniform distribution. 98
- $Q$  SfM pointcloud reconstruction. 76
- $q$  Tip point of the tool axis. 94
- $R$  Rotation matrix. 39
- $r$  Radius of the cylinder for the geometric model. 97
- $S$  Static reference pointcloud/model in registration. 38
- $s$  Similarity measure for registration. 45
- $\mathcal{S}$  Number of tool axes that support a dominant cluster in the keyhole estimation algorithm. 96
- $T$  Spatial transformation function. 38
- $t$  Translation vector. 39
- $T^*$  Optimal transformation function. 38
- $T_{model}$  Model transformation matrix. 51
- $\mathcal{T}$  Set of SfM poses. 78
- $T_{view}$  View transformation matrix. 51
- $U$  Number of model particles used to solve the inter-image correspondence constraint. 78
- $u$  Index of model particle used to solve the inter-image correspondence constraint. 78
- $u_0$  Principal point in x axis. 51
- $V$  Tool axis for keyhole estimation. 94
- $v_0$  Principal point in y axis. 51
- $\mathcal{V}$  Set of tool axes for keyhole estimation. 94
- $W$  Orientation vector of the tool axis. 94
- $w_k$  Intersection point in the liver surface of the tumour projection towards the camera centre  $k$ .  
100
- $w_x$  Intersection point in the liver surface of the tumour projection towards the estimated keyhole  
 $x'$ . 100
- $x$  Keyhole location. 94
- $\mathcal{X}$  Set of keyhole locations. 94