Automatic registration of 3D models to laparoscopic video images for guidance during liver surgery

Maria Ruxandra Robu

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Medical Physiscs and Biomedical Engineering

University College London

April 4, 2020

I, Maria Ruxandra Robu, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Acknowledgements

I am extremely grateful to have met so many extraordinary people in these past few years, which were key in shaping my PhD experience.

Firstly, I would like to thank my primary supervisor Dr. Matt Clarkson, who was instrumental in the completion of this work. I am grateful he trusted me to follow my curiosity and enthusiasm in exploring multiple research directions, providing valuable guidance all the way through. Moreover, he was supportive and patient in the difficult parts of this journey. I consider myself very lucky to have been part of his team.

I'd like to thank Prof. Dan Stoyanov for the valuable technical advice during my PhD. He believed in me since I started my studies at UCL in a Master programme and supported my research career with numerous opportunities to grow and learn. I am grateful to Prof. Dave Hawkes, Prof. Kurinchi Gurusamy, Prof. Brian Davidson for the multiple insightful meetings. I've learned so much under your guidance! I'd like to thank Cancer Research UK for funding my PhD research. Moreover, I'd like to thank my examiners, Dr. Giannarou and Dr. Bergeles for taking the time to read my thesis thoroughly and for a stimulating, insightful viva.

I thank my colleagues in the Smart Liver team - working with you has been a great opportunity to grow as a researcher, collaborate and learn. I'd also like to thank Micha for the many insightful research conversation and his friendship. I am so grateful we went from Winter School colleagues to collaborating in each others' research projects.

I am lucky to have met some incredibly talented researchers in the Centre of Doctoral Training at UCL. I am especially grateful to Irina, Danny, Nooshin, Marta

and Alejandro for the many geeky talks, painting sessions, random walks, their support and friendship. I'd also like to thank Diana, Daniela, Stefanos and many more for making these years more enjoyable and fun. I'm so grateful I met you and I can count on you as my friends. Thanks to Dragana, for her patience and wisdom.

I'd also like to thank my boyfriend, Agoston, for his support, patience and for infusing my days with positivity and encouragements.

Ana, tu ai un loc special in inima mea. Ai crezut ca pot sa fac orice imi propun si te-ai asigurat ca nu uit asta - de la surprize in posta si excursii cu povesti pana la lucrat la distanta din Londra. Iti multumesc tare de tot!

Si cel mai important, acest proiect nu ar exista fara familia mea - Catalina, Marcu, George si Ilinca. Le sunt recunoscatoare pentru sprijinul, dragostea si incurajarile din ultimii ani. Sunt extrem de norocoasa sa ii am aproape si sa stiu ca acasa este mereu un loc plin de caldura, bucurie si siguranta.

Abstract

Laparoscopic liver interventions offer significant advantages over open surgery, such as less pain and trauma, and shorter recovery time for the patient. However, they also bring challenges for the surgeons such as the lack of tactile feedback, limited field of view and occluded anatomy.

Augmented reality (AR) can potentially help during laparoscopic liver interventions by displaying sub-surface structures (such as tumours or vasculature). The initial registration between the 3D model extracted from the CT scan and the laparoscopic video feed is essential for an AR system which should be efficient, robust, intuitive to use and with minimal disruption to the surgical procedure.

Several challenges of registration methods in laparoscopic interventions include the deformation of the liver due to gas insufflation in the abdomen, partial visibility of the organ and lack of prominent geometrical or texture-wise landmarks. These challenges are discussed in detail and an overview of the state of the art is provided.

This research project aims to provide the tools to move towards a completely automatic registration. Firstly, the importance of pre-operative planning is discussed along with the characteristics of the liver that can be used in order to constrain a registration method. Secondly, maximising the amount of information obtained before the surgery, a semi-automatic surface based method is proposed to recover the initial rigid registration irrespective of the position of the shapes. Finally, a fully automatic 3D-2D rigid global registration is proposed which estimates a global alignment of the pre-operative 3D model using a single intra-operative image. Moving towards incorporating the different liver contours can help constrain

the registration, especially for partial surfaces.

Having a robust, efficient AR system which requires no manual interaction from the surgeon will aid in the translation of such approaches to the clinics.

Impact

This thesis provides computer vision tools for the registration of a 3D liver mesh extracted from a pre-operative scan and the intra-operative scene for augmented reality (AR) systems in laparoscopic liver surgery. Specifically, three main technical contributions are introduced.

Firstly, a pre-operative planning tool analyses the patient-specific anatomy in a simulation framework. It outputs a selection of surface patches which should be acquired by the surgeon in order to obtain an efficient registration. This could be useful for generating clear protocols for data acquisition for the clinicians, removing redundant data collection since time efficiency is critical during surgery. Moreover, since the proposed pre-operative simulation is independent of the registration technique used, I showed that it can be easily extended to analyse any other registration method. So, it could be used by other researchers as a complementary tool to make their data acquisition collection process more intuitive for surgeons.

Secondly, a semi-automatic 3D-3D global registration method is proposed which demonstrates the importance of introducing anatomical landmarks of the liver as an additional constraint when dealing with partial surfaces. Thirdly, a fully automatic 3D-2D pipeline for global registration is proposed which can align the pre-operative 3D model to a single laparoscopic image of the liver. The pipeline is composed of two steps: i) automatic detection of liver contours using deep learning and ii) automatic global contour-based registration. Both registration techniques are robust to partial visibility ratios and deformation characteristic to laparoscopic interventions and show promising results on retrospective clinical data.

The majority of proposed global registration approaches in the literature re-

quire some degree of manual interaction of a user with the software, which is cumbersome and relies heavily on the user's experience of organ anatomy and the software. Moreover, it is the main bottleneck towards automating a long-term AR system for laparoscopic liver interventions with multiple subsequent stages relying on a good initial alignment. All proposed methods have been designed specifically to increase the usability of AR systems in operating rooms, by making them more intuitive to use, with clear protocols and automation. These characteristics are essential for an easier translation to clinics. The modular approach taken when developing the proposed techniques allows for easy integration within other registration approaches published in the literature in order to automate them or improve their usability. Additionally, the proposed tools could be used for commercially available software aiming to bring AR in laparoscopic liver surgeries.

In order to encourage the contributions of this thesis to be put in practice by other researchers, the work in this thesis has either been published or is in the process of publication. In total, it led to 13 publications including multiple collaborations (among which the winning entry to the MICCAI 2019 ABCD Challenge on predicting fluid intelligence from MRI scans). Moreover, this research inspired the participation in several public engagement activities for the general public (in the NHS Open Day, Science of Surgery in Charles Bell House).

Publications

First Author

- [1] - **M. R. Robu**, B. Koo, M. Allam, M. Pfeiffer, S. Thompson, K. Gurusamy, B. Davidson, S. Speidel, D. Hawkes, D. Stoyanov and M. J. Clarkson, "Automatic, Global Registration in Laparoscopic Liver Surgery" *to be submitted*, 2019

- [2] - **M. R. Robu**, J. Ramalhinho, S. Thompson, K. Gurusamy, B. Davidson, D. Hawkes, D. Stoyanov and M. J. Clarkson, "Global Rigid Registration of CT to Video in Laparoscopic Liver Surgery" International Journal of Computer Assisted Radiology and Surgery, 2018, `https://doi.org/10.1007/s11548-018-1781-z`

- [3] - **M. R. Robu**, P. Edwards, J. Ramalhinho, S. Thompson, B. Davidson, D. Hawkes, D. Stoyanov and M. J. Clarkson, "Intelligent viewpoint selection for efficient CT to video registration in laparoscopic liver surgery" International Journal of Computer Assisted Radiology and Surgery, 2017, `https://doi.org/10.1007/s11548-017-1584-7`

Collaborations

- [4] - S. Thompson, T. Dowrick, G. Xiao, J. Ramalhinho, **M. R. Robu**, M. Ahmad and M. J. Clarkson, "SnappySonic: An Ultrasound Acquisition Simulator", *under review*, Journal of Open Research Software, 2019

- [5] - M. Pfeiffer, I. Funke, **M. R. Robu**, S. Bodenstedt, L. Strenger, S. Engel-

hardt, T. Roß, M. J. Clarkson, K. Gurusamy, B. R. Davidson, L. Maier-Hein, C. Riediger, T. Welsch, J. Weitz and S. Speidel, "Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation" MICCAI 2019, https://arxiv.org/abs/1907.02882

- [6] - Y. Fu, **M. R. Robu**, B. Koo, C. Schneider, S. van Laarhoven, D. Stoyanov, B. Davidson, M. J. Clarkson and Y. Hu, "More unlabelled data or label more data? A study on semi-supervised laparoscopic image segmentation" MIL3ID workshop, MICCAI 2019, https://arxiv.org/abs/1908.08035

- [7] - A. Mihalik, M. Brudfors, **M. R. Robu**, F. S. Ferreira, H. Lin, A. Rau, T. Wu, S. B. Blumberg, B. Kanber, M. Tariq, M. D. M. E. Garcia, C. Zor, D. I. Nikitichev, J. Mourao-Miranda and N. P. Oxtoby, "ABCD Neurocognitive Prediction Challenge 2019: Predicting individual fluid intelligence scores from structural MRI using probabilistic segmentation and kernel ridge regression" MICCAI 2019, **winning entry**, http://arxiv.org/abs/1905.10831

- [8] - N. P. Oxtoby, F. S. Ferreira, A. Mihalik, M. Brudfors, H. Lin, A. Rau, B. Stefano, **M. R. Robu**, C. Zor, M. Tariq, M. Del, M. E. Garcia, B. Kanber and D. I. Nikitichev, "ABCD Neurocognitive Prediction Challenge 2019: Predicting individual residual fluid intelligence scores from cortical grey matter morphology" MICCAI 2019, https://arxiv.org/abs/1905.10834

- [9] - J. Ramalhinho, **M. R. Robu**, S. Thompson, K. Gurusamy, B. Davidson, D. Hawkes, D. Barratt and M. J. Clarkson, "A pre-operative planning framework for global registration of laparoscopic ultrasound to CT images" International Journal of Computer Assisted Radiology and Surgery 2018, https://doi.org/10.1007/s11548-018-1799-2

- [10] - J. Ramalhinho, **M. R. Robu**, S. Thompson, P. Edwards, C. Schneider, K. Gurusamy, D. Hawkes, B. Davidson, D. Barratt and M. J. Clarkson,

"Breathing motion compensated registration of laparoscopic liver ultrasound to CT" in SPIE Medical Imaging 2017, https://doi.org/10.1117/12.2254488

- [11] - E. D. Gibson, **M. R. Robu**, S. A. Thompson, P. E. Edwards, C. Schneider, K. Gurusamy, B. R. Davidson, D. J. Hawkes, D. C. Barrat and M. J. Clarkson, "Deep residual networks for automatic segmentation of laparoscopic videos of the liver" in SPIE 2017, http://dx.doi.org/10.1117/12.2255975

- [12] - S. Thompson, E. Edwards, J. Ramalhinho, **M. R. Robu**, D. Stoyanov, K. Gurusamy, C. Schneider, B. Davidson and M. J. Clarkson, "Calibration Pattern Selection for Laparoscope Calibration" in SPIE 2017, https://doi.org/10.1117/12.2253717

- [13] - G. Jones, N. T. Clancy, X. Du, **M. R. Robu**, S. Arridge, D. S. Elson and D. Stoyanov, "Fast estimation of haemoglobin concentration in tissue via wavelet decomposition" in MICCAI 2017, https://arxiv.org/abs/1706.07263

Contents

1 Introduction **33**

1.1 Liver resection 35

1.2 Laparoscopic liver resection 37

1.3 Augmented reality in laparoscopic liver resection 40

1.4 Problem Statement 43

1.5 Thesis Structure 45

2 Background **47**

2.1 Coordinate systems 47

2.2 Registration overview 51

2.3 3D-3D registration 52

2.3.1 Global alignment 52

2.3.2 Fine alignment 54

2.4 3D-2D registration 55

2.4.1 Camera model 55

2.4.2 Global alignment 57

2.4.3 Fine alignment 58

2.5 Challenges for registration in laparoscopic liver surgery 59

3 Registration in Laparoscopic Liver Surgery **61**

3.1 Pre-operative planning 63

3.2 Global alignment 66

3.3 Fine alignment 69

3.3.1 Rigid alignment . . . 69
3.3.2 Non-rigid alignment . . . 71
3.4 Deep learning in laparoscopic liver interventions . . . 76
3.5 Validation . . . 78
3.5.1 Public datasets . . . 82
3.5.2 Clinical trials . . . 84
3.6 Visualisation . . . 85
3.7 Summary . . . 88

4 Intelligent Viewpoint Selection 91
4.1 Contributions . . . 94
4.2 Methods . . . 95
4.2.1 Search space representation . . . 96
4.2.2 Visible surface simulation . . . 96
4.2.3 Viewpoint scoring strategy . . . 97
4.3 Experiments . . . 97
4.3.1 Viewpoint evaluation . . . 98
4.3.2 View planning for registration . . . 99
4.3.3 Effect of the number of patches on registration accuracy . . 101
4.3.4 Application in laparoscopic liver resection . . . 101
4.4 Discussion . . . 104

5 3D-3D Semi-automatic Global Registration 109
5.1 Contributions . . . 110
5.2 Methods . . . 110
5.2.1 Optimisation . . . 112
5.2.2 Features and Descriptors . . . 115
5.2.3 Distances . . . 115
5.2.4 Estimating the rigid transform . . . 116
5.3 Experiments . . . 116
5.3.1 Synthetic data . . . 117

5.3.2 Liver phantom . . . 119
5.3.3 Application in clinical data . . . 121
5.4 Discussion . . . 121

6 3D-2D Automatic Global Registration 125
6.1 Contributions . . . 127
6.2 Methods . . . 128
6.2.1 Semantic contour detection network . . . 130
6.2.2 3D-2D contour-based registration . . . 132
6.3 Experiments . . . 135
6.3.1 Semantic contour detection . . . 135
6.3.2 Registration . . . 139
6.3.3 Automatic global registration pipeline . . . 141
6.4 Results . . . 141
6.4.1 Semantic contour detection . . . 141
6.4.2 Registration . . . 142
6.4.3 Automatic global registration pipeline . . . 143
6.5 Discussion . . . 145

7 Conclusion 149

Bibliography 154

List of Figures

1.1 Scheme of the abdominal cavity. The liver is situated on the right-upper quadrant of the abdomen. Illustration by Don Bliss. 33

1.2 Incidence and mortality distribution for the 10 most common cancer cases in 2018. Each percentage is computed with respect to the total number of cases. Figure from [14], under the Creative Commons Public License, https://creativecommons.org/licenses/by-nc-sa/2.5/deed.en. 34

1.3 Functional liver segment division proposed by Couinaud [15]. Colour coding: brown - liver, green - gallbladder, blue - falciform ligament, orange - left and right triangular ligaments. 36

1.4 Examples of anatomic resections vs. an example of a non-anatomic resection. Colour coding: brown - liver, green - gallbladder, red - sub-surface lesion, yellow - area to resect 37

1.5 Comparison between scars from: a) Open liver resection and b) laparoscopic liver resection. 38

1.6 Peripheral areas of the liver are considered to be favourable for laparoscopic resections (Fig. from [16]). 39

1.7 Examples showing the advantages of augmented reality in minimally invasive surgery. Left) The position of the blood vessel and tumour indicated by the surgeon with dotted lines. Right) The correct positions provided by an interactive augmented reality system, validated with a laparoscopic ultrasound. A clear difference is observed in the AR visualisation with the correct positions. Out of the 7 clinical cases studied (5 livers, 2 pancreata), the surgeons delineated an erroneous localisation in 2 liver clinical cases and one pancreas surgery. Colour coding: purple - vessel, green - tumour. Images from [17]. 41

2.1 Illustration of the main transforms between the coordinate frames in an OR setup with optical tracking. The tracker represents the world space in which all the laparoscopic views are aligned automatically. Registration consists of estimating T_{CT}^{World} from the CT to the World space. 48

2.2 Illustration of the main transforms between the coordinate frames in a vision-based OR setup. Instead of chaining multiple transforms in a common space, vision-based systems involve directly the estimation of T_{CT}^{Cam}. 49

2.3 Global alignment illustration in 2D. Two shapes X and Y can be successfully aligned given the correspondence pairs $\{(x,y)_i\} \subset X \bigtimes Y$. Illustration inspired from [18]. 52

2.4 Iterative Closest Point. Given a good initial alignment between the shapes X and Y, closest point correspondences are estimated and the alignment is updated. Multiple such iterations ultimately refine the initial alignment leading to an optimal registration between the shapes. Illustration inspired from [18]. 54

2.5 Pinhole camera model. A 3D point in the World space P can be projected to a 2D Image Plane given the intrinsic camera parameters: principal point coordinates c_x, c_y, focal length f, and possibly, skewness (not shown here for clarity). Figure from [19]. 56

2.6 Global 3D-2D alignment illustration. Once the 2D points $\{y_i\}$ and their 3D correspondences $\{x_i\}$ are established, PnP methods can be used to estimate the camera pose T^{Cam}_{World}. 57

2.7 AR example using 3D-2D correspondences. The camera in the head mounted display (left) can analyse and augment the view of the user in real time. In this example, the world consists of a 2D planar marker (middle). The corners are detected by finding the intersection points of the lines fitted to the edges. Once the 2D -3D correspondences are known, a virtual object can be projected on top of the physical plane (right). Figure adapted from [20]. 58

2.8 Bio-mechanical simulation of the abdominal wall under gas insufflation for registration in laparoscopic interventions. Colour coding: abdominal wall - red; liver - green; surrounding organs - purple. Left) Abdominal wall before gas insufflation; Right) Abdominal wall after gas insufflation. Images from [21]. 59

2.9 Variation in the appearance of liver surface for healthy tissue and cirrhosis. Figure from [22]. 60

3.1 Virtual reality tool for pre-operative planning. A virtual screen shows a 3D rendering of the processed CT scan including the liver surface, vessel trees and tumour (yellow). The surgeon can control the transparency of each 3D model with a handheld device. The abdominal CT scan is visible on the right. Figure from [23]. 63

3.2 Pre-operative planning of port placement with augmented reality for a trans-thoracic laparoscopic intervention. A) Testing various port positions. B-C) Marking on the patient skin of the chosen port. Figure from [24]. 65

3.3 Comparison of visible liver surface - a) The falciform ligament was cut and a large portion of both liver lobes is visibile (Figure from [25]); b) The falciform ligament is present and the surface visibility is decreased to a liver lobe or less. 68

3.4 Augmented reality overlay during a porcine liver resection [26]. A global registration was performed manually, followed by ICP for rigid refinement resulting in an accuracy of 8.6 mm. Left) The overlay shows good agreement of the liver surface edges with the projected 3D model. Right) The overlay includes the vasculature and tumours identified in the pre-operative scan. Figure from [26]. . . . 70

3.5 Registration formulation based on the liver contour proposed by Adagolodjo et al. [27]. Firstly (step 1) the deformation due to pneumoperitoneum is simulated before the intervention. Step 2 shows a rigid global alignment performed by a user which aligns the pre-operative 3D model to approximately match the intra-operative image. Finally, in step 3 the annotated 2D organ contour is used as a constraint for a biomechanical simulation of the liver deformation. Figure from [27]. 74

3.6 Partial visibility experiment for the registration framework proposed by Özgür et al. [28]. Left) The field of view circles occlude the liver contours (anterior ridge in red and silhouette in yellow) approximately by 10%, 20% and 30%. Right) Their proposed method [28] is compared against another deformable [25] and rigid registration approach [17] with promising results for the tested levels of occlusion. Figure from [28]. 76

3.7 Image-to-image translation proposed by Pfeiffer et al. [5]. The images obtained from a computer simulation of a laparoscopic liver intervention (domain A) are translated to look like real laparoscopic images (domain B_{syn}) using deep learning. The middle and right columns show various styles for the same simulated scene. Figure from [5]. 77

3.8 Example of liver contours on a phantom (left) and on *in vivo* data (right) which could be used as a measure of registration accuracy, as proposed by Thompson et al. [29]. The yellow lines denote the annotated contours on the images, whereas the white lines are obtained by projecting the corresponding 3D contours given the current registration estimate. Figure from [29]. 81

3.9 OpenCAS dataset. Left) Phantom experiment setup where a metallic rod creates a deformation. Middle) Initial (red) vs. deformed (blue) liver surfaces extracted from CT scans. Right) Initial surface (red) vs. intra-operative acquisition (blue) with a stereo laparoscope. Figure from [30]. 82

3.10 OpenHELP dataset. The CT scan of a healthy male (left) is used to generate 3D surfaces for all the organs in the thoracic and abdominal cavities (middle). These surfaces are then used to obtain synthetic phantoms (right). Figure from [31]. 83

3.11 Examples of image translation from the synthetic environment (left) to photo-realistic laparoscopic images (right) from the public dataset released by [5]. Figure from [5]. 83

3.12 Importance of correctly overlaying lesion information. a) Laparoscopic image with the projected 3D models using transparency blending, b) Tool Access Visualisation - the resection ring is projected with respect to the incision tool port, c) Transparency Blending Visualisation - The centre c of the tumour as it appears on the laparoscopic view does not correspond to the true centre when the incision tool is used, d) Tool Access Visualisation - A safe margin is computed around the tumour and projected to the laparoscopic view from the perspective of the incision tool. Images from [32]. . 86

3.13 Various augmented reality visualisation strategies in a target localisation study [33]. A: No overlay; B: Always-on solid overlay; C) Always-on wire mesh overlay; D) Inverse realism overlay. Figure from [33]. 87

4.1 Example of data collection from the SmartLiver AR system. The liver patches are selected by a user to be generally well distributed over the liver surface. Figure from [26]. 91

4.2 Illustration of a view planning framework for 3D object reconstruction. The arrows represent cameras placed around the bunny, which indicate a sufficient set of best views for surface reconstruction. Figure from [34]. 92

4.3 Overview of the proposed pre-operative planning simulation framework. Using the pre-operative 3D surface model, cameras are simulated around the liver, taking into account the laparoscopic constraints. Then, each camera is used to simulate a surface patch and compute the viewpoint score based on the surface curvature. Finally, a selection of the best surface patches is obtained which can constrain the subsequent alignment and make the data acquisition faster. 94

4.4 Illustration of the simulated abnormality in the liver model. The registration error is calculated between the target and its estimated position given an alignment. 98

4.5 Top-Left) Correlation between the Gaussian curvature scoring strategy $f_{surface}$ and the registration errors. Top-Right) TRE variation over the liver surface. Each sphere represents a camera location. The colour map is applied to $log(mean(TRE))$ over the 100 registrations for each sphere ($mean(TRE) \in [0.7992mm, 13.7048mm]$). Bottom-Left) The size of the spheres is proportional to $std(TRE)$. So, a bigger variation in errors leads to larger spheres. Bottom-Right) The size for the spheres is proportional to the curvature score $f_{surface}$. 99

4.6 Basins of convergence for ICP for the lowest scoring point cloud P_{low}, the highest scoring point cloud P_{high} and the maximum visible surface P_{max} for three scenarios: left) rotation around the X axis; middle) rotation around the Y axis; right) rotation around the z axis. Note that surface patch scoring is computed using $f_{surface}$. The green circles were omitted for clarity for the baseline as it always converges to the desired position. 100

4.7 Visual assessment of ICP progress until convergence. Colour coding: fixed point cloud in white, moving point cloud after the rotation is applied in red, moving point cloud during ICP iterations in green, ideal final position of the moving point cloud in purple. Column a) Success example for a rotation of 40° around the z axis. The images from top to bottom show the iterations 0 (before ICP), 1, 12, 150. Column b) Failure example for a rotation of 18° around the y axis. The images from top to bottom show the iterations 0 (before ICP), 1, 20, 25. 102

4.8 Left) Mean TRE variation (for 100 random transformations) with an increasing number of patches. Middle) Example of top 3 lowest scoring patches chosen by the algorithm. Right) Example of top 3 highest scoring patches. Blue regions highlight flat surfaces, whereas green symbolises positive Gaussian curvature. 103

4.9 Visual assessment of registration with view planning for real data. Colour coding: fixed point cloud in white, moving point cloud in green, ideal final position given by a gold standard registration of the moving point cloud in purple. Left) The previous experiment was re-ran for the constraints given by real data from a liver resection in order to suggest a number of patches to be considered. Middle) Registration results using clinical data after running ICP on a merged point cloud from the top 4 recommended patches by the proposed method. Right) Registration results using clinical data after running ICP on a merged point cloud from 4 random patches. 103

4.10 Example of video frames used for the stereo reconstruction of the surface in a laparoscopic liver resection. 104

4.11 Registration examples highlighting the need for additional constraints in a real world scenario. Colour coding: fixed point cloud in white, moving point cloud in green, ideal final position given by a gold standard registration of the moving point cloud in purple. Left) The moving cloud is built from 2 patches - the highest scoring viewpoint and the patch which is furthest away from it. Right) The moving point cloud is built from the top two highest scoring patches. 106

5.1 An overview of the proposed global alignment framework, showing the pre-operative and intra-operative steps. 111

5.2 Ambiguity in correspondence estimation. The point configuration on the left should be matched to the corresponding red point set on the right. However, on a nearly planar surface, multiple configurations with equal spatial distributions can be found. Figure from [35]. 112

5.3 Pairwise constraints on the moving, M, (blue) and target, T, (pink) models used for pruning the correspondence set. 113

5.4 Experiments on synthetic data. Top: Robustness to reduced partial size in the target model, T, Bottom: Robustness to increasing levels of deformation in T. The target model representing 23% of the total liver surface is used in the bottom experiment with increasing deformation levels. *Colour coding:* moving model, M - blue, target model, T - pink. 118

5.5 Phantom experiment. The proposed global initial alignment is sufficient to allow potentially any fine-alignment method to successfully converge. The global alignment is applied to a partial region of the deformed surface (Top) and a partial surface reconstruction from an intra-operative stereo laparoscopic camera (Bottom), followed by the successful convergence of LM-ICP[36]. The colour map highlights the error distribution (TRE-mm) computed between the ground truth sub-surface markers and their estimated position after alignment. *Colour coding:* moving model, M - blue, target model, T - pink. 120

5.6 Global alignment on clinical data from a dataset acquired during a laparoscopic liver resection. *Colour coding Alignment:* moving model, M - blue, target model, T - pink. The overlay is computed after applying LM-ICP on the proposed alignment. *Colour coding Overlay:* liver tumour - green, vessels - purple, liver contour - yellow. 122

6.1 An overview of the proposed pre-operative and intra-operative stages for a global 3D-2D registration. Colour coding: silhouette - yellow, anterior ridge - blue 128

6.2 Contour labels example: correct annotation for the silhouette (yellow) and the anterior ridge (blue). Around the falciform ligament, the ridge line does not always correspond to the liver boundary - example of an incorrect ridge annotation in the dotted line 129

6.3 CASENet architecture with ResNet [37] as backbone. Solid blue rectangles represent ResNet blocks, where decreasing width translates to a drop in the spatial resolution by a factor of 2. Features are extracted from the bottom layers and concatenated with the top layer classification in the brown block. The fused classification block (green) takes as input the concatenated layers and outputs k maps. The dotted red lines denote the layers where supervision is imposed. More details about the individual blocks can be found in [38]. Illustration adapted from [38] 130

6.4 Qualitative results of ridge (blue) and silhouette (yellow) prediction: a) Original images from the test datasets: *daVinci*: row 1), *lap1*: rows 2)-6), *lap2*:rows 7)-8); b) Ground truth annotations; c) Predictions when CASENet was trained only on a clinical dataset ($I+C$); d) Predictions when CASENet was trained with synthetic and clinical data ($I+S+C$) 137

6.5 Example of F1 score maps for the ridge ($F1_r$) and silhouette ($F1_s$) where green = true positives, blue = false positives and red = false negatives. *a)* Predictions when CASENet was trained on the clinical dataset ($I+C$); *b)* Predictions when CASENet was pre-trained on the synthetic dataset and fine-tuned on the clinical one ($I+S+C$) . 140

6.6 *a)* Registration accuracy(mm) for various surface visibility ratios. The size of the circles is proportional to the standard deviation over 100 registrations. *b)* Error variation (mm) over liver region. Each circle represents a camera location. The circle colour is given by *median(RMSE)* (shown in the colour bar) and their size is proportional to the standard deviation over 100 registrations. The same colours were used in graph a) in order to help match corresponding cameras. Examples of visible ridge (blue) and silhouette (yellow) from chosen camera positions that lead to successful (*c1*) and failed (*c2*) registrations. 142

6.7 Qualitative results of the proposed global registration pipeline on retrospective clinical cases. *a)* Predictions of the ridge (blue) and silhouette (yellow) by the proposed semantic contour detection ($I+S+C$). *b)* Aligned 3D preoperative liver model using the proposed camera pose estimation. *c)* A rendering of the liver model and estimated camera pose in 3D space. 144

6.8 The proposed pipeline is robust to errors in the contour detection: false negatives, false positives or misclassification. *a)* Predictions of the ridge (blue) and silhouette (yellow) by the proposed semantic contour detection ($I+S+C$). *b)* Corresponding error maps for the predicted contours where green = true positives, blue = false positives, red = false negatives and pink = misclassification. *c)* Aligned 3D preoperative liver model using the proposed camera pose estimation. *d)* A rendering of the liver model and estimated camera pose in 3D space. 144

List of Tables

3.1 Fine rigid registration methods in open liver resection (OLR) and laparoscopic liver resection (LLR) 69
3.2 Fine non-rigid registration methods in liver surgery 72
3.3 Validation for rigid methods 78
3.4 Validation for non-rigid methods 80

6.1 Performance of CASENet on the test datasets: F1 score (F1), Recall (R) and Precision (P). The numbers represent the average over all the images in each dataset. 139

Chapter 1

Introduction

The liver is a vital organ in the human body, weighing an average of $1500g$ in adults. It is located under the diaphragm, on the right-upper quadrant of the abdominal cavity [39]. The liver regulates most chemical components in the blood stream, i.e. it breaks down alcohol and drugs, so they can be passed out of the system; it produces certain proteins which help with blood clotting; clears out bilirubin; stores sugars and converts them into energy when needed [40]. Once the harmful substances are broken down, the waste is eliminated either in the blood or through the bile. The waste in the blood will then be filtered by the kidneys and eliminated through urine, whereas the bile enters the intestine and is eliminated as faeces.

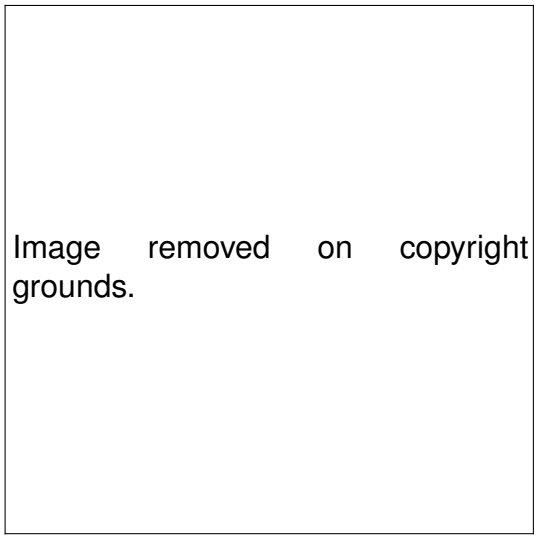

Figure 1.1: Scheme of the abdominal cavity. The liver is situated on the right-upper quadrant of the abdomen. Illustration by Don Bliss.

Liver cancer can be categorised by its origin in primary (starts in the liver cells) and secondary (from cancer cells of other organs). Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer [41]. The worldwide incidence (2018) of liver cancer cases was estimated to be 841 080 (Fig. 1.2), and it is predicted to increase due to the prevalence of hepatitis B and C [42]. Liver cancer is the 4th most common cause of death from cancer worldwide after lung, colorectum and stomach cancer, with 781 631 cases leading to deaths in 2018 [14]. The incidence of new cancer cases in 2018 was estimated at 18.1 million worldwide and it is predicted to reach 20 million by 2025 [43].

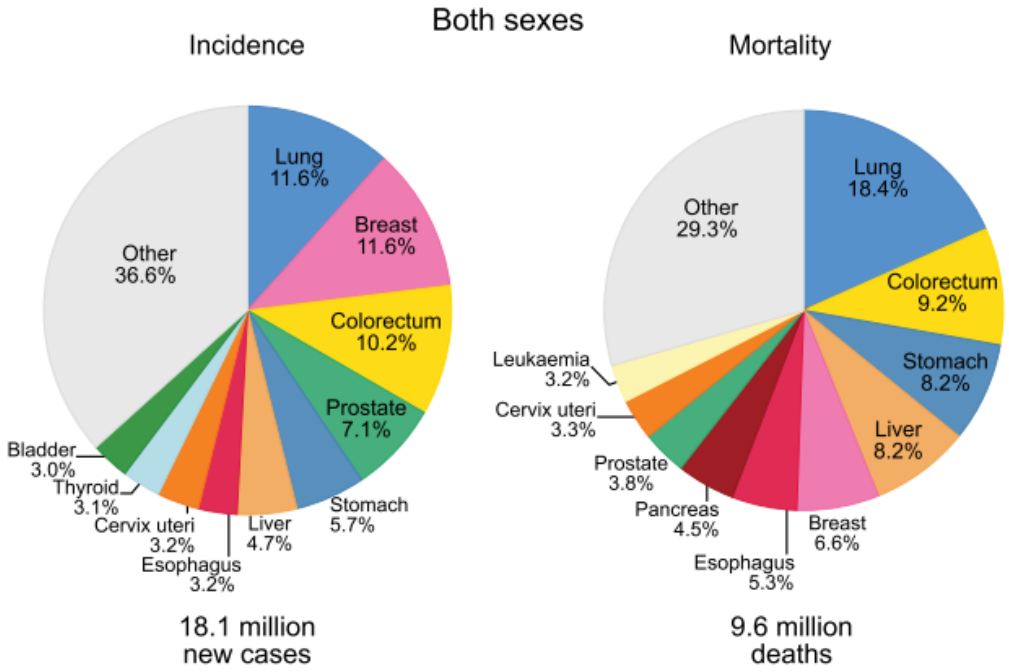

Figure 1.2: Incidence and mortality distribution for the 10 most common cancer cases in 2018. Each percentage is computed with respect to the total number of cases. Figure from [14], under the Creative Commons Public License, `https://creativecommons.org/licenses/by-nc-sa/2.5/deed.en`.

In the late 70s, patients diagnosed with HCC who received no treatment died within one year, irrespective of the diagnosis stage [44]. A median survival rate of 21.6 months was observed for patients undergoing surgical intervention. These prognostics improved over the years, due to a better screening process. A meta-analysis of liver cancer published in 2002 reports 5-year survival rates lower than

20% for untreated patients diagnosed with advanced HCC [45]. On the other hand, surgical interventions were reported to lead to a 5-year survival rate of up to 60%-75% in selected patients [46]. However, the survival data is difficult to compare or extrapolate. If the cancer is detected in an advanced stage, the choice of treatment is usually irrelevant for the final prognosis [45]. On the other hand, if the cancer is detected at an early stage, a course of treatment is recommended, which could be curative. Due to ethical concerns, there are no randomized trials comparing surgical treatment to no treatment in HCC patients.

Surgical resection and liver transplants are curative treatments, with the former being recommended as primary therapy in the very early stages of cancer and no underlying cirrhosis is present. In some very specific cases, resection is also recommended for patients which are not suitable for liver transplant and have liver cirrhosis but with well preserved liver function [46]. The choice for treatment is made based on the number of tumours, their location and sizes and the presence of underlying liver disease. Other neoadjuvant treatments include chemotherapy, chemo-embolisation (a chemical component is delivered to the tumour location to restrict its blood flow), radiotherapy or thermal ablation. However, surgical resection is considered the most efficient type of treatment [46].

1.1 Liver resection

Couinaud [15] proposed to split the liver into 8 functional segments, as in Fig. 1.3. This classification is used by surgeons to plan resections, since each delimitation is created based on branches of the hepatic artery and portal vein [47]. Thus, surgeons need to consider the segment the lesion belongs to along with its neighbouring vessel structure, in order to minimise bleeding. The liver surface is divided into the left and right lobe by the falciform ligament (blue line in Fig. 1.3), which keeps the liver attached to the abdominal wall. The left and the right triangular ligaments attach the liver to the diaphragm.

Resection planning takes into account the aforementioned divisions and their respective major vascular components to minimise possible risks and complications.

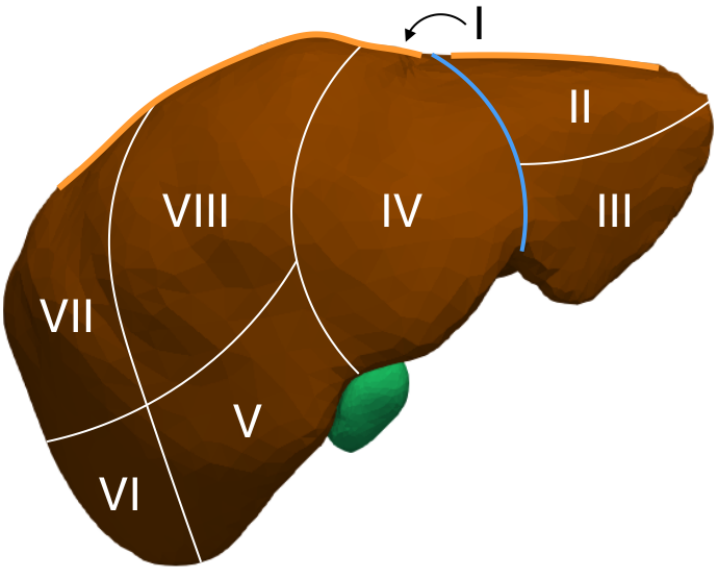

Figure 1.3: Functional liver segment division proposed by Couinaud [15]. Colour coding: brown - liver, green - gallbladder, blue - falciform ligament, orange - left and right triangular ligaments.

While anatomic resections usually involve two or more segments, non-anatomic or wedge resections involve removing the tumour along with a safety margin around it [47] (see Fig. 1.4 for several examples). A negative resection margin is desired, where all the cancerous cells have been removed. On the other hand, a positive resection margin denotes the presence of cancerous cells. The main consideration when planning for surgical resection is the remaining healthy liver volume available.

Contrary to other organs, the liver has a remarkable capacity for regeneration. It can restore the lost mass in order to adjust to body size and recuperate its functions. It has been shown on animals that liver regeneration is possible after two thirds of the liver are removed [48]. Similarly, in cases with no underlying liver disease, it is considered acceptable to perform resection of up to 75% of the normal liver volume in human patients [49]. On the other hand, in cases with cirrhosis, resections of up to 60% of the liver volume can be performed [50].

The main advantage of non-anatomic resections is that more of the liver volume is preserved, which can be extremely important when post-operative treatments such as chemotherapy or recurrent tumours are involved [47]. An analysis over 201 patients with colorectal liver metastases which underwent anatomic vs. non-

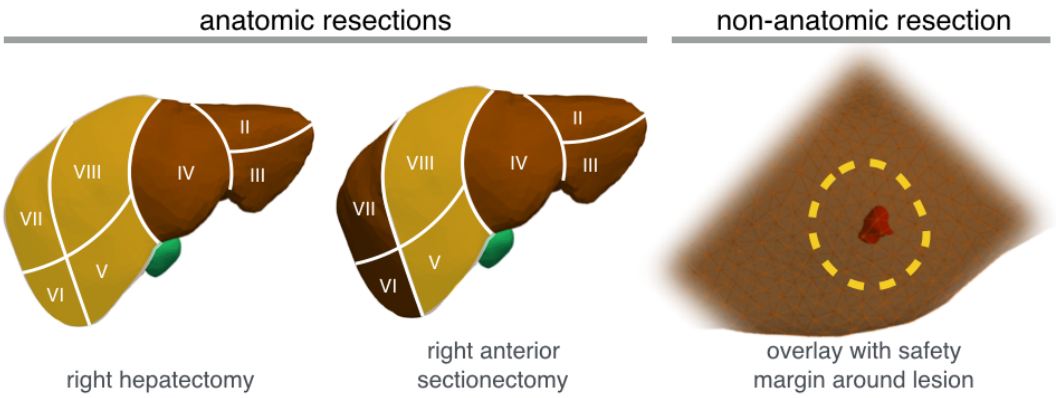

Figure 1.4: Examples of anatomic resections vs. an example of a non-anatomic resection. Colour coding: brown - liver, green - gallbladder, red - sub-surface lesion, yellow - area to resect

anatomic liver resection concluded that the latter group received less blood transfusions and had a shorter hospital stay [51]. The non-anatomic resections were performed for smaller metastases. Moreover, there were no differences between the groups regarding recurrence rates, survival, mortality or morbidity.

1.2 Laparoscopic liver resection

Laparoscopic liver resections (LLR) have been explored in the past decades as a safe alternative to open liver resections (OLR). LLRs consist of making small incisions on the abdomen of the patient, where trocars are placed (Fig. 1.5b). The abdominal cavity is inflated with carbon dioxide (pneumoperitoneum) in order to make space for the movement of the surgical instruments. Then, the laparoscopic camera along with the other instruments can be inserted through the trocars. The surgeon has visual access to the intra-operative scene by looking at a monitor (which is usually placed next to the patient) streaming the laparoscopic video feed. Pre-operative planning on abdominal scans can be used to proceed with the intervention.

A recently published meta-analysis [52] gathered data from 463 manuscripts and studied the effects of laparoscopic liver surgery in over 9000 cases. They confirm safety of LLRs with lower overall rates of complications, estimated blood loss, rate of blood transfusions and hospital stays in both minor and major resections [52]. The use of LLRs can potentially make the surgery more accessible to patients

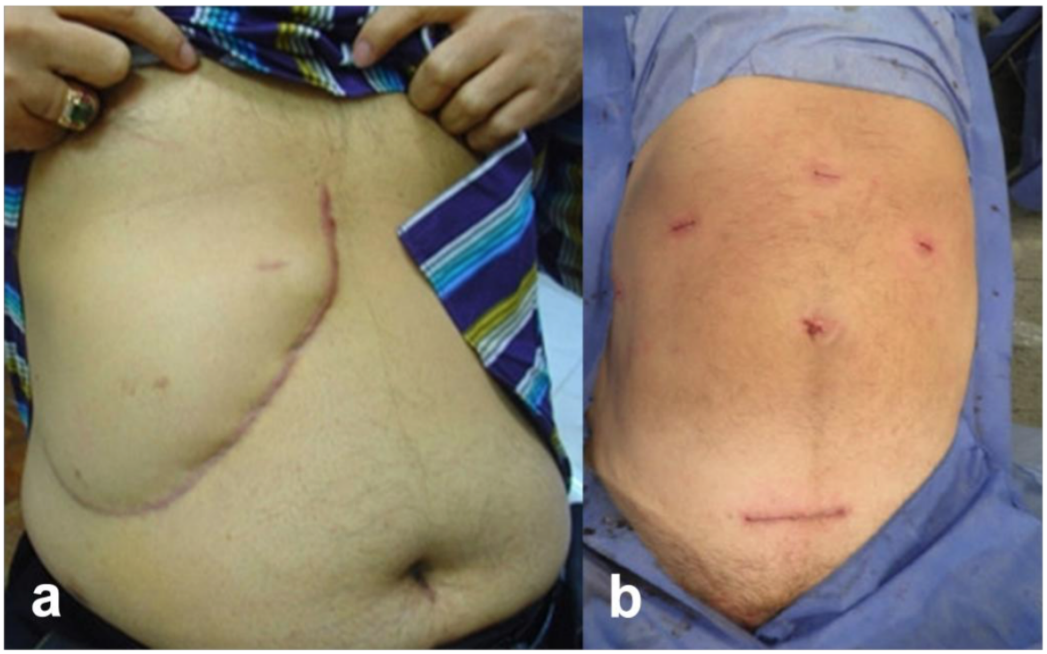

Figure 1.5: Comparison between scars from: a) Open liver resection and b) laparoscopic liver resection.

with underlying liver disease, which might be unsuitable to OLR due to high risk of post-operative complications [53].

In addition, LLRs could be used as bridging therapy for the cases on liver transplant waiting lists [49]. Due to the shortage of donors, patients can be on waiting lists for a prolonged time, during which small lesions might develop. Depending on the case, they can either be removed through open surgery or laparoscopically. Since open interventions usually cause adhesions around the tumour site, the access to the liver is more difficult during the transplant procedure. Laparoscopic liver resections have been found to result in less blood loss, reduced operative time and significantly less blood transfusions needed [54]. Moreover, the recovery time after LLRs is much shorter [49].

An international statement published in 2008 declared LLR is best suited for patients with small single tumours ($<$ 5cm in diameter) in the peripheral liver segments (Fig. 1.6), among other conditions [55]. However, with improved laparoscopic techniques and better visualisation intra-operatively, initial results have shown that LLR for tumours situated in the postsuperior segments in selected pa-

tients is technically feasible and safe [56].

Figure 1.6: Peripheral areas of the liver are considered to be favourable for laparoscopic resections (Fig. from [16]).

Despite the significant benefits and the growing safety of the procedures, LLRs are still not widely adopted in most facilities. A recent worldwide study reported laparoscopic interventions being considered in a small percentage of liver resections (5%-30%), with some centers reaching 80% [57]. Most of these interventions consist of minor resections, however the number of major resections has increased as well.

One of the reasons many surgeons are reluctant to perform laparoscopic liver resections is the high risk of major hemorrhage which might be difficult to control intra-operatively [52]. Moreover, there are several technical challenges characteristic to interventions with a monocular laparoscope, which have been grouped into five main categories [58]:

- **Weak depth perception.** The projection from the 3D surgical environment to a 2D video leads to a loss of depth information. Thus, it takes more training for the surgeons to be able to mentally infer the 3D information of the surgical scene and the movements of the instruments from 2D images.

- **Constrained vantage point.** The laparoscopic camera enters the abdomen through a trocar placed in an incision which is usually performed once during

the surgery. As a result, the surgeon has a restricted view of the surgical scene and it may prevent the accessibility of some parts of the patient anatomy.

- **Limited field of view.** The laparoscopic camera usually has a small field of view and a short focal length. As a result, the surgeon has to mentally create a visual representation of the surgical field for navigation.

- **Limited sense of touch.** Open surgeries allow the surgeon to palpate the organ and differentiate between healthy tissue and the tumour. However, this is not possible in laparoscopic resections where the organ can only be manipulated by the surgical instruments.

- **Hidden anatomy.** Some anatomical structures such as tumours or vessel trees are hidden inside or behind the organ, which is a limiting factor for both laparoscopic and open surgeries. It is essential for the surgeon to locate these hidden structures during the intervention.

The research community aims to merge different imaging techniques to obtain additional information about the surgical scene, thus addressing some of the above challenges. Augmented reality (AR) is one such approach where virtual information extracted from pre-operative data is seamlessly integrated with the live video feed from the laparoscopic camera.

1.3 Augmented reality in laparoscopic liver resection

Augmented reality can potentially provide several advantages during laparoscopic interventions.

Firstly, AR can provide the surgeons with a fast localisation of the sub-surface structures (such as major vessel trees and tumours). Thus, clinicians could potentially avoid hitting unexpected vessels, which would result in additional blood loss (see Fig. 1.7). Moreover, the location of tumours with respect to neighbouring vessels or potential abnormal vasculature is considered critical to the resection plan [49]. So, AR could improve the understanding of complex anatomical configurations.

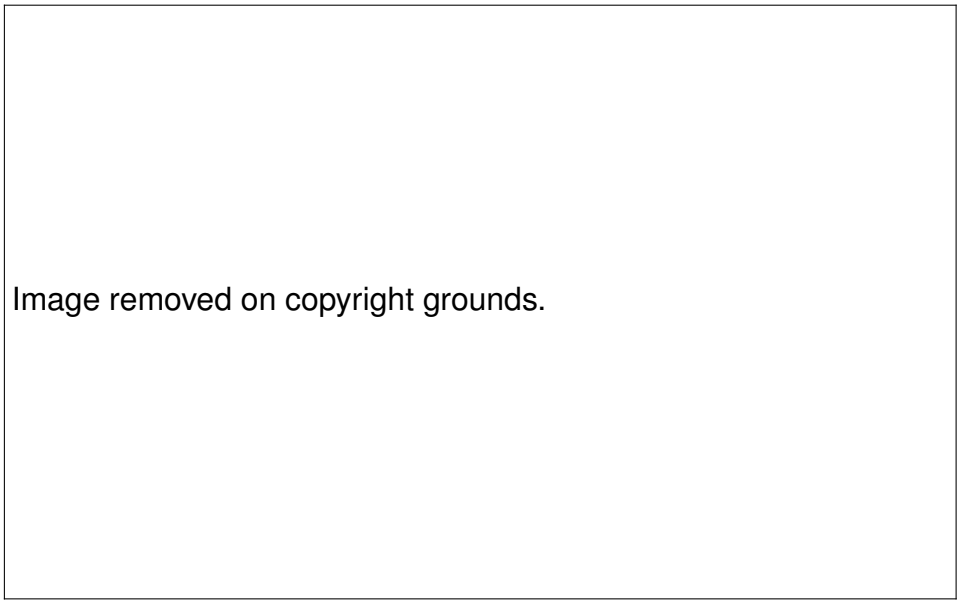

Figure 1.7: Examples showing the advantages of augmented reality in minimally invasive surgery. Left) The position of the blood vessel and tumour indicated by the surgeon with dotted lines. Right) The correct positions provided by an interactive augmented reality system, validated with a laparoscopic ultrasound. A clear difference is observed in the AR visualisation with the correct positions. Out of the 7 clinical cases studied (5 livers, 2 pancreata), the surgeons delineated an erroneous localisation in 2 liver clinical cases and one pancreas surgery. Colour coding: purple - vessel, green - tumour. Images from [17].

At the same time, knowing the location of tumours within a safety margin can help target the resection region more accurately and reduce cancer-positive margins. Any planning performed before the surgery could be displayed for guidance intra-operatively, such as the desired resection margin [32]. So, the surgeons could potentially target the cancerous tissue with more accuracy and spare more healthy tissue, which is important for the patient's recovery and consequent organ functionality. A preliminary study on ex vivo porcine kidneys compared the resection margins in laparoscopic tumour resection [59]. Out of the 30 tumours resected, the rate of tumours with positive resection margins in the AR and non-AR group were 13.8% and 30.3% respectively. They also reported no misses in the AR group, whereas the non-AR group had 4 complete misses of tumour targeting. Such guidance could potentially reduce the risk of complications for the patients through faster decision making, reduced surgical time and less blood loss. Another study proposed an improved AR visualisation method and reported 85.2% negative margin rate with AR

guidance vs. 41.9% without guidance in an ex-vivo experiment (59 pseudo-tumours over 29 porcine kidneys) [32].

Secondly, AR reduces the cognitive load of having to match pre-operative images to the intra-operative scene under significant deformation, a task which depends heavily on the experience of the surgeon. It has been shown that AR can increase the efficiency and accuracy of localisation tasks as well as decrease the time per task and the tool path length [60, 33]. Consequently, augmented reality could also be used in the future to improve training and provide assistance for inexperienced surgeons.

While non-anatomic resections bring important advantages such as preserving more of the liver volume (see Section 1.1), they are difficult to perform in a laparoscopic context. While in open surgery the surgeon can palpate the liver, this is not possible in minimally invasive approaches. Augmented reality could provide the additional guidance in locating tumours identified in the pre-operative images. Given the advantages above, AR could lead to more cases being reconsidered as suitable for laparoscopic liver resections [61].

An area that is considered very difficult to access in laparoscopic liver resections consists of the posterosuperior segments: VII, VIII, IVa (Fig. 1.6). Nonetheless, a trans-thoracic resection guided by an augmented reality system was reported for a patient with a lesion in segment VIII [24].

Another clinical application where AR could bring benefits would be cases where multiple tumours spread over both lobes. In general, such cases can be inoperable due to the difficulty in removing the tumours along with a safety margin and preserving a sufficient volume of healthy liver. Banz et al. [62] report using AR for guiding a treatment plan involving resection of the largest tumour along with ablation for the rest of them. Since some of the tumours were not visible with laparoscopic ultrasound (so called *vanishing lesions*), AR guidance helped with the localisation by overlaying pre-operative CT data on the laparoscopic view. As such, AR navigation may help improve longer-term survival by guiding intra-operative treatment in selected patients considered inoperable or with difficulties in applying

local ablation.

Taking into account the current evidence and clinical experience, AR systems are most likely to be used in complex procedures. Example scenarios could include cases involving tumours in challenging locations or multiple non-anatomic resections which require precise safety margins [63]. Moreover, AR could also be used for educational purposes in helping more inexperienced surgeons gain confidence and operate in laparoscopic contexts. More precise use cases will most likely be defined with the advancement of the technology towards better target localisation and improved usability.

1.4 Problem Statement

In most clinical protocols, the patient undergoes a pre-operative CT scan of the abdomen. Since there is at least a one day gap until the surgery, the CT scan can be segmented and the 3D surfaces of the liver and sub-surface structures can be extracted. Several such services are available for pre-operative planning (i.e. Visible Patient [1], MeVis [2]).

The work presented in this thesis focuses on the registration stage at the beginning of the surgery, before the surgeon starts manipulating the organ. Such registration methods can be formulated in two steps, namely estimating a *global registration* capable of providing an initial alignment of the two models and a *fine alignment* which can refine the registration starting from a good initialisation. Multiple automatic rigid and non-rigid fine alignment methods have been proposed in the literature. However, global registration is generally achieved manually. A user has to rotate and translate a 3D model from a 2D view in order to roughly match it with the current laparoscopic image. This stage is cumbersome, time-consuming and depends highly on the user's experience in the organ's anatomy. Moreover, every interaction with the AR display requires the surgeon to pause the intervention and release any tools or instruct another OR staff to perform the interaction.

This thesis introduces a fully automatic global registration pipeline, each chap-

[1] https://www.visiblepatient.com/en/
[2] https://www.mevis.de/en/

ter providing essential concepts in making the proposed techniques robust to the challenges of laparoscopic liver surgery introduced above. A global alignment technique should provide estimates within several cm [64], as most fine alignment algorithms can successfully converge thereafter.

The overall aim of this thesis is to develop techniques which are compatible with the current clinical workflow. They should be:

- efficient - able to reach a solution in a reasonable time frame, characteristic to the clinical intervention. It is generally considered that the initial alignment should be estimated within several minutes, since it needs to be performed only once. On the other hand, long-term registration update should be real-time in order to cope with breathing motion, laparoscope movement and tissue-tool interaction.

- usable and intuitive - with minimal disruption to the surgical procedure and minimal dependence on user input

- easily translated in surgical operating rooms - using equipment already available in clinical protocols for liver surgery

- reliable - real-time accuracy measurement of the provided alignment. Depending on the visualisation strategy used, various accuracy constraints should be met. If the registered model is viewed side by side with the laparoscopic view, errors of up to $10mm$ have been reported as reasonable since the user can compensate mentally for the discrepancy [65]. However, the use of overlays would require much more accurate alignments with values ranging between 3 to $5mm$ [66, 26]. Equally important is how such alignment errors are conveyed to the clinician in an intuitive way to enable the quick assessment of the system's reliability. The difference between the projected and visible organ contours has shown promising results so far [29]. A more detailed discussion can be found in Chapter 3.

The proposed work focuses on rigid registration due to several advantages. Firstly, while non-rigid registration techniques can improve the accuracy with

promising results, the current main barrier for translation to clinics is the usability of AR systems. The development of the proposed techniques was aimed towards creating an intuitive workflow, with minimal disruptions to the surgical procedure. Secondly, since the majority of non-rigid registration and long-term tracking techniques rely on a good rigid initialisation, providing a robust global rigid registration is critical. Lastly, validation for rigid registration is well understood and standardised. Moreover, a qualitative assessment of a rigid alignment can be easily achieved by comparing the projected organ contour with the one from the laparoscopic view. Such visual cues help the surgeon assess the reliability of the AR system.

Ultimately, the surgeons should be provided with guidance during the intra-operative intervention, by fusing a model extracted from the pre-operative CT scan with the laparoscopic video feed. Such a visualisation could help in the uptake of laparoscopic procedures for liver surgery by making complex procedures easier to perform. Alternately, augmented reality aims to reduce the surgeon's cognitive load of matching pre-operative scans to the intra-operative view. Since the two modalities differ significantly due to a different position of the organ, deformation due to pneumoperitoneum and occlusion, inferring the location of sub-surface structures based on mental mapping relies heavily on the surgeon's expertise. So, having an intra-operative guidance system could also be of use in training more surgeons to safely perform laparoscopic liver resections.

1.5 Thesis Structure

Chapter 2 presents an overview of 3D-3D and 3D-2D registration background as well as the main challenges encountered in liver registration.

Chapter 3 covers the state of the art techniques on the initial registration of the pre-operative 3D volumes and intra-operative scene in laparoscopic liver resections. Moreover, important aspects about validation and visualisation in AR systems for liver surgery are discussed.

The main contributions of the thesis are:

- **Efficient data collection simulation framework** - **Chapter 4** presents a

proof of concept for intelligent collection of intra-operative data for efficient registration. This work was presented at the International Conference on Information Processing in Computer-Assisted Interventions (IPCAI) 2017 in Barcelona and published in the International Journal for Computer Assisted Radiology and Surgery (IJCARS) 2017 [3].

- **Global surface-based registration** - **Chapter 5** details a semi-automatic 3D-3D global rigid registration method. This work was accepted for presentation at IPCAI 2018 in Berlin and for publication in IJCARS 2018.

- **Contour-based automatic, global registration** - **Chapter 6** presents a 3D-2D fully automatic global rigid registration technique. This work is in preparation for submission.

Chapter 2

Background

The alignment between the pre-operative 3D data and the intra-operative scene in the first stages of the surgery is a critical step for augmented reality systems. This chapter introduces theoretical concepts related to 3D-3D and 3D-2D rigid registration. These methods are well-established in the computer vision literature and are considered a gold-standard in many applications. Subsequently, the challenges specific to liver geometry and texture are discussed which break some of the assumptions of traditional registration algorithms.

2.1 Coordinate systems

There are multiple ways to setup an augmented reality system with all its related components [17, 67]. The main variations in hardware stem from different types of laparoscopic camera tracking (i.e. optical, electro-magnetic (EM), vision-based). In the following section, optical and vision-based tracking systems will be discussed, which are used in the following chapters.

Let X and Y be the moving (pre-operative) and the target (intra-operative) models. In the 3D-3D alignment formulation, these models are either meshes or point clouds whereas in 3D-2D, Y can represent points in the image space.

Let T_a^b be a rigid transform from coordinate frame a to frame b:

$$T_a^b = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{2.1}$$

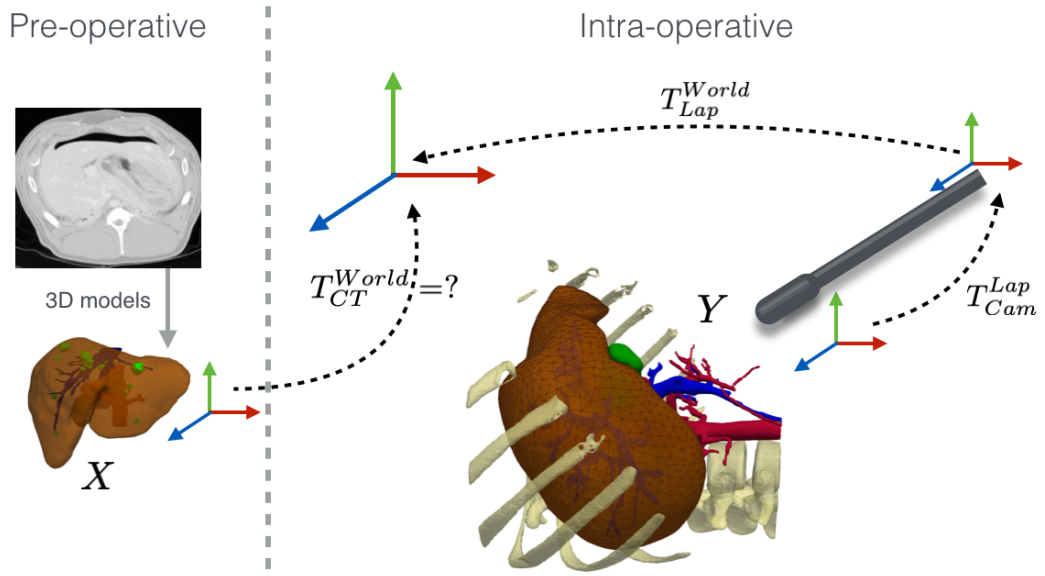

Figure 2.1: Illustration of the main transforms between the coordinate frames in an OR setup with optical tracking. The tracker represents the world space in which all the laparoscopic views are aligned automatically. Registration consists of estimating T_{CT}^{World} from the CT to the World space.

where R is a 3 x 3 rotation matrix and t is a 3 x 1 column vector representing a translation in 3D space.

Various coordinate systems can be seen as either world centric, or camera centric. In Fig. 2.1 the optical tracker can be considered the world space. A set of retroreflective spheres are attached at the end of the laparoscopic camera. The relationship between the camera space and the laparoscope space - T_{Cam}^{Lap} can be obtained using hand-eye calibration [68] after the spheres are attached. The optical tracker detects the 3D position of the spheres with infrared light in real-time T_{Lap}^{World}. The aim of the initial registration stage is to estimate the transform that can align the pre-operative surface models to the intra-operative scene - T_{CT}^{World}:

$$T_{CT}^{World} X = T_{Lap}^{World} T_{Cam}^{Lap} Y \tag{2.2}$$

When using a stereo laparoscope, multiple surface patches obtained through stereo reconstruction [69] can be merged automatically in the world space using optical tracking. Thus, a point cloud of the intra-operative scene can be obtained. Since each surface patch is individually small and unlikely to constrain a registration

algorithm well, such an approach is advantageous because it can provide a larger intra-operative surface quickly.

On the other hand, chaining the transforms results in an accumulation of errors from the hand-eye calibration and the tracking [26, 29]. Moreover, when using stereo reconstruction techniques, a triangulation error is present due to the small distance between the stereo cameras. More details about the individual error sources and their magnitude are discussed in Section 3.5.

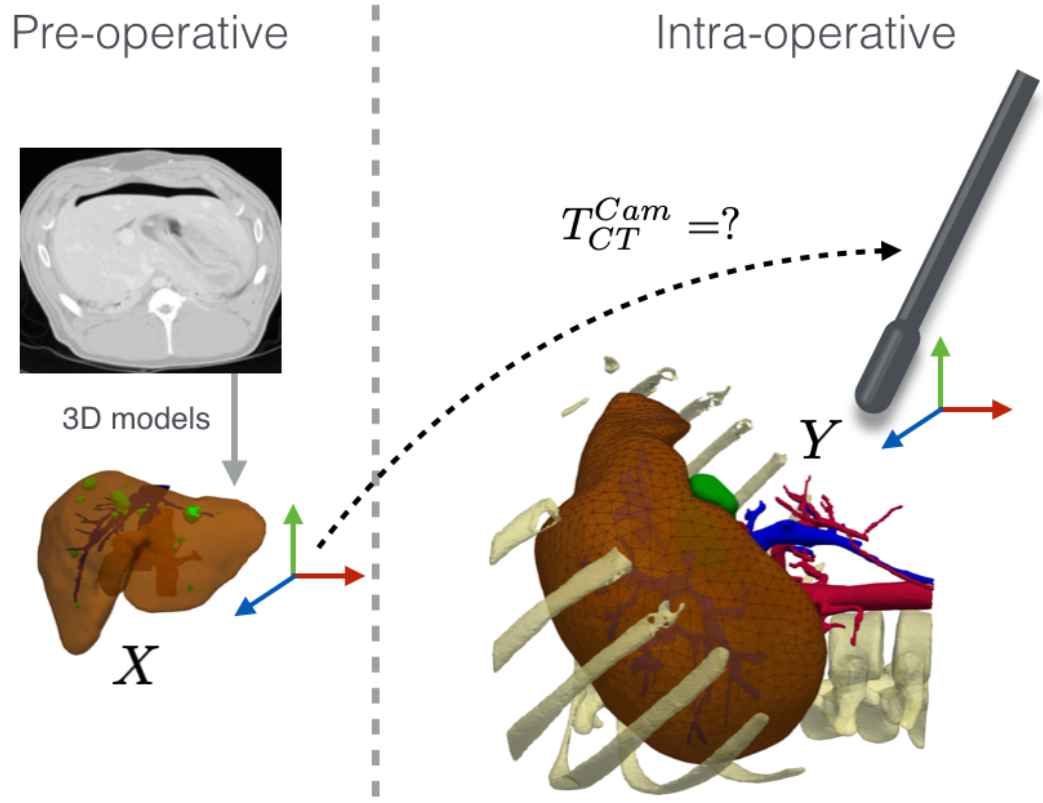

Figure 2.2: Illustration of the main transforms between the coordinate frames in a vision-based OR setup. Instead of chaining multiple transforms in a common space, vision-based systems involve directly the estimation of T_{CT}^{Cam}.

An alternative would be to work directly with a vision-based system, such as the one presented in Fig. 2.2. In this case, the aim of the initial registration is to directly estimate the T_{CT}^{Cam} transform:

$$T_{CT}^{Cam} X = Y \tag{2.3}$$

While the registration accuracy would not contain any of the associated errors from hand-eye calibration or tracking, this is a more challenging problem. In a 3D-3D context, estimating a rigid transform from one single stereo surface reconstruction is in general not possible. Good stereo reconstructions can be obtained if the laparoscope is close to the surface. However, it also means the reconstructed surface corresponds to a small area. Since the liver is locally relatively smooth, a single surface patch does not generally provide sufficient constraints for a registration.

Another approach to obtain a larger surface reconstruction of the liver would be to perform an exploratory video which can reconstruct a 3D map of the world and track the camera simultaneously (such as Simultaneous Localisation and Mapping (SLAM)-based methods). Indirect SLAM relies on extracting features from the input data, whereas direct SLAM techniques use pixel intensities in an optimisation of photometric data and various regularisation terms. Both research directions have reached high accuracy and real-time performance in the vision literature. However, to the best of my knowledge, direct SLAM techniques have not been explored in laparoscopic liver registration. Since the liver surface has a repetitive, homogeneous texture in an environment with challenging illumination, direct SLAM approaches might be difficult to apply.

On the other hand, several attempts have been made to translate indirect SLAM techniques to laparoscopic data [70, 71]. Reichard et al. [70] merge multiple dense stereo reconstructions into one point cloud. They show that using ICP to merge consecutive views fails due to geometrical ambiguity in a small liver patch and to slight changes in illumination that lead to noisy stereo reconstruction. Successful mosaicking results were shown when an external laparoscopic tracker is used to provide a good initial localisation of each individual patch. However, they report difficulties in creating a system that does not rely on additional hardware. Nonetheless, such SLAM-derived frameworks assume a static world with no deformation, which is not the case when the liver undergoes breathing motion. Mahmoud et al. [71] extend a feature-based SLAM technique to estimate dense scene reconstruc-

tions from multiple images taken in an exploratory video. Using multiple images for the scene estimation provides an advantage in dealing with changing lighting conditions over stereo reconstruction due to the larger parallax between frames. While their method can deal with the breathing motion of the organ, they state that it cannot cope with very homogeneous soft-tissue textures. In conclusion, SLAM-based approaches are difficult to apply in laparoscopic contexts with challenging illumination, specular reflections, repetitive texture and homogeneous appearance and organ deformation.

Another alternative would be to estimate T_{CT}^{Cam} from a single laparoscopic image in a 3D-2D formulation. The main difficulty in such a context is to find reliable landmarks which can be identified in the pre-operative 3D surface which lacks texture and the current 2D intra-operative image which lacks geometry.

The next sections will introduce several well-established methods for 3D-3D and 3D-2D registration. They are widely known in the computer vision literature, but challenging to apply to liver surgery.

2.2 Registration overview

In a registration pipeline, the task is to align the moving model X to a target model Y. These models can be represented by the points $\{x_i\} \subset X$ and $\{y_i\} \subset Y$. The solution to align the two sets of points can be obtained by optimising a registration function to estimate the parameters θ of the transform T:

$$\theta = \underset{\theta}{argmin}\, E(T(\theta, X), Y) + E_{reg}(X) \tag{2.4}$$

In non-rigid registration approaches, the θ parameters can represent a rigid transform with an additional deformation model. As a result, more degrees of freedom can be estimated to recover the full deformation between X and Y. Additionally, a regularisation term E_{reg} is usually needed in order to constrain the space of possible solutions and make the optimisation more robust. However, since this thesis focuses on rigid registration, in the rest of this section, θ denotes the six parameters required for a 3D rotation R and translation t.

Finding the optimal set of parameters θ from Eq. 2.4 is not straight-forward, especially when the target point model Y is only partially visible. Due to this ambiguity, registration algorithms are usually split into two stages. Firstly, a global alignment is needed to provide a good initial estimation of θ. Secondly, a fine alignment can be applied to improve the accuracy and find the optimal transform parameters θ. In general, fine alignment methods for surfaces with partial visibility are able to converge if the global alignment errors are less than 20 mm and 20° [64].

2.3 3D-3D registration

An overview of the most popular approaches to solve for the global and fine alignment problems in 3D-3D registration is provided in the next paragraphs.

2.3.1 Global alignment

The global alignment stage or shape matching is based on inferring correspondences from the shape geometry of X and Y. Moreover, such approaches recover the rigid parameters θ irrespective of the models' initial position.

In the 3D-3D shape matching formulation, the moving and target models X and Y are usually represented by point clouds or meshes. Let $\{(x,y)_i\} \subset X \bigtimes Y$ be a set of n given correspondences between the X and Y.

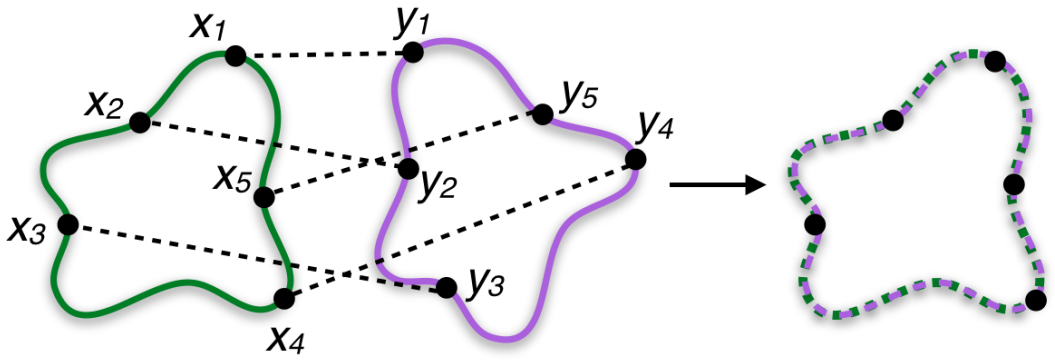

Figure 2.3: Global alignment illustration in 2D. Two shapes X and Y can be successfully aligned given the correspondence pairs $\{(x,y)_i\} \subset X \bigtimes Y$. Illustration inspired from [18].

So, Eq. 2.4 can be re-written as:

$$R,t = \underset{R,t}{argmin} \sum_{i=1}^{n} ||y_i - (Rx_i + t)||^2 \tag{2.5}$$

The aim is to estimate the rotation R and translation t which best align the points x_i to their correspondences on the target model y_i. See Fig. 2.3 for an illustration of shape matching.

First of all, the point sets need to be centred:

$$\overline{x_i} = x_i - \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.6}$$

$$\overline{y_i} = y_i - \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2.7}$$

As derived in [72, 73], the optimal rotation and translation can be obtained using the following equations:

$$R = V \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & det(VU^T) \end{bmatrix} U^T \tag{2.8}$$

where $USV^T = SVD(\overline{X}\,\overline{Y}^T)$. SVD is the singular value decomposition of the $\overline{X}\,\overline{Y}^T$ matrix, resulting in the left and right singular matrices U and V and the singular values on the diagonal of the S matrix. The term $det(VU^T)$ ensures the resulting R is a rotation and not a reflection matrix.

$$t = \frac{1}{n}\sum_{i=1}^{n} y_i - R\frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.9}$$

As can be seen, once a set of correct correspondences is obtained, the following estimation of the rotation and translation is straight-forward. However, obtaining a set of correspondences is not trivial in most cases. In computer vision research, such shape matching problems usually detect unique feature points (i.e. due to variation of curvature) on both shapes. For each feature point, a descriptor could be computed which encodes in a feature vector the local neighbourhood (i.e. normal variations,

point distributions). By estimating the closest neighbour in the descriptor space (i.e. using Euclidean distance, correlation), a set of candidate correspondences can be found. In the case of outliers, techniques such as Random Consensus and Sampling (RANSAC) [74] can be employed to obtain a solution from the correct correspondences.

2.3.2 Fine alignment

Once a global alignment is obtained, the result can be improved with local methods. A popular algorithm in 3D-3D registration is Iterative Closest Point (ICP) [75]. ICP minimises the mean squared error (Euclidean distance) between the corresponding points - the same as in Eq. 2.5. However the correspondences are not known. Due to this ambiguity, an alternating scheme was proposed. Given two point clouds with a good initial alignment, a correspondence set is estimated using closest neighbours. Given the correspondences, an estimate of the current rigid transform is obtained using the shape matching technique described above. After several iterations, ICP is guaranteed to converge to a local minimum. See Fig. 2.4 for an illustration of ICP.

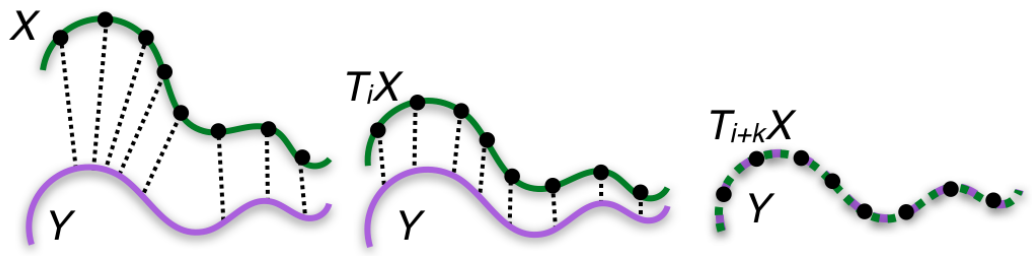

Figure 2.4: Iterative Closest Point. Given a good initial alignment between the shapes X and Y, closest point correspondences are estimated and the alignment is updated. Multiple such iterations ultimately refine the initial alignment leading to an optimal registration between the shapes. Illustration inspired from [18].

The main assumptions in ICP is that a good initial alignment is provided and that the target model Y must overlap or be a subset of the moving model X. Firstly, having a good initialisation is critical since ICP usually gets stuck in local minima. Secondly, if the surfaces do not overlap, computing the closest neighbour correspondence set will result in outliers which could hinder the convergence of the

optimisation. There have been many proposals for improving ICP from a speed or accuracy point of view - i.e. by using different sampling strategies [76], different distance metrics, by taking into account normal differences in the candidate correspondences, dealing with anisotropic and inhomogeneous errors in the point clouds [64]. A more comprehensive review of efficient variants of ICP can be found in [76].

The equations introduced in the global and fine alignment sections above can be applied for 2D-2D registration between point sets in an image. However, in the 3D-2D registration an additional perspective projection step is involved.

2.4 3D-2D registration

In this section, 3D-2D registration refers to the task of recovering the camera pose given correspondences between 3D points and their projection on a single 2D image. Let $\tilde{y}^{2D} = [u, v, 1]^T$ and $\tilde{x}^{3D} = [x, y, z, 1]^T$ be the homogeneous coordinates of a 2D and a 3D point, respectively. The minimisation function for 3D-3D registration (Eq 2.5) can be re-written to include the projection from 3D to 2D.

$$\theta = \underset{\theta}{argmin} \sum_{i=1}^{n} ||\tilde{y}_i^{2D} - K\pi T_{World}^{Cam}(\theta, \tilde{x}_i^{3D})||^2 \tag{2.10}$$

where K is the intrinsic matrix and π is the perspective projection matrix. Similar to the 3D-3D formulation, θ consists of the six parameters needed to define a rigid transform (with rotation R and translation t) which aligns the 3D model to the camera view.

In the next paragraphs, the formulation for the additional matrices is detailed along with the chosen camera model.

2.4.1 Camera model

The point $\tilde{x}^{2D}$ in an image is obtained by perspective projection from the 3D point $\tilde{x}^{3D}$, using:

$$\tilde{x}_i^{2D} = K\pi T_{World}^{Cam} \tilde{x}_i^{3D} \tag{2.11}$$

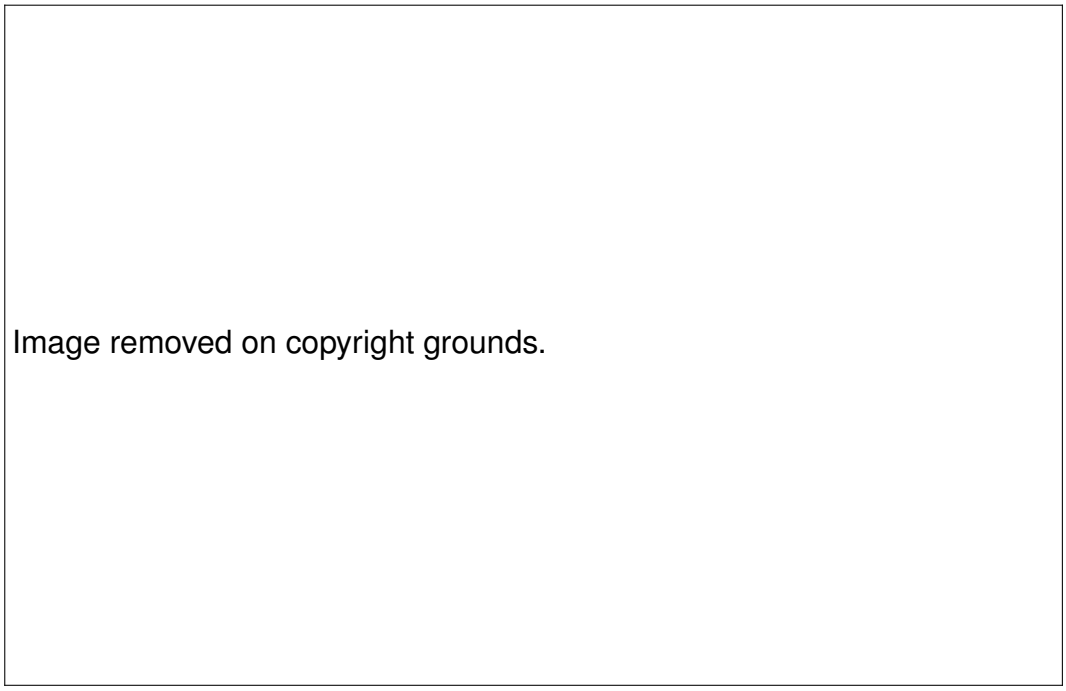

Figure 2.5: Pinhole camera model. A 3D point in the World space P can be projected to a 2D Image Plane given the intrinsic camera parameters: principal point coordinates c_x, c_y, focal length f, and possibly, skewness (not shown here for clarity). Figure from [19].

The intrinsic matrix K is given by:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.12}$$

where $f_\bullet$ is the ratio between the focal length of the lens and the size of a pixel and $c_\bullet$ denote the coordinates of the principal point.

π is the perspective projection matrix:

$$\pi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.13}$$

The intrinsic parameters along with any possible distortions introduced by the lens can be obtained using an offline calibration step [77]. If the focal length is fixed

during the data acquisition step, the intrinsic parameters remain unchanged as they do not depend on the current camera view. This step can be performed before the surgery and is assumed as given in the following chapters.

The only remaining unknown in Eq. 2.10 is the rigid transform T_{World}^{Cam}, which will be the focus of the next sections.

2.4.2 Global alignment

Similarly to the 3D-3D registration problem, a global alignment has to be estimated first, followed by local refinement.

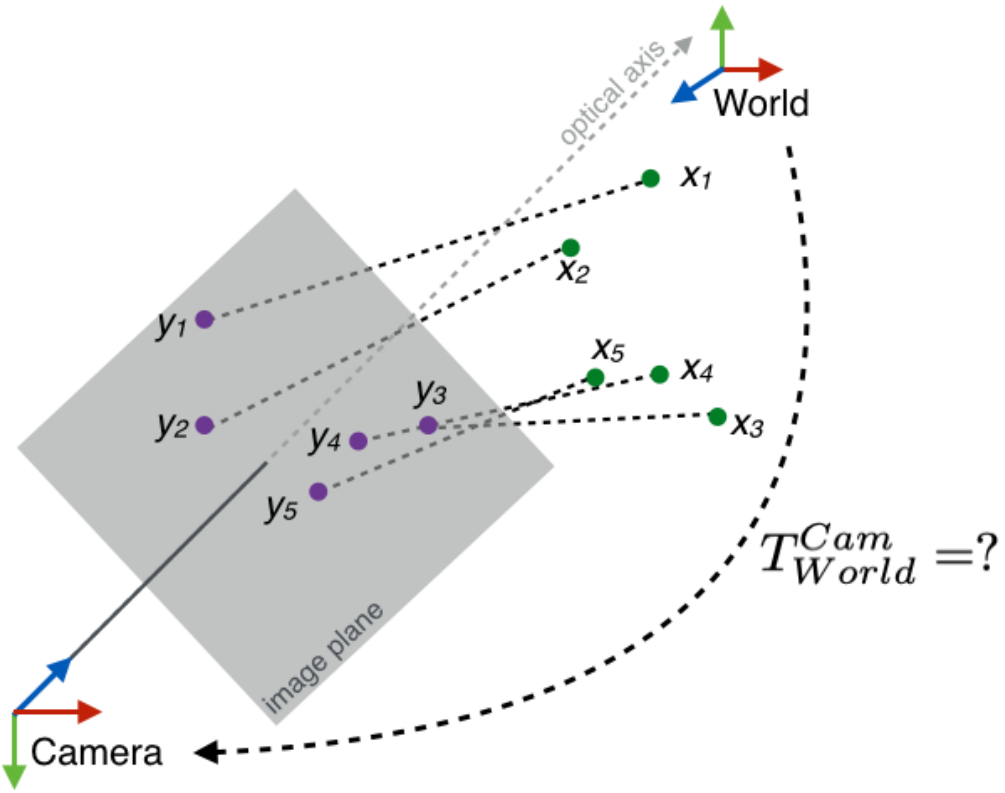

Figure 2.6: Global 3D-2D alignment illustration. Once the 2D points $\{y_i\}$ and their 3D correspondences $\{x_i\}$ are established, PnP methods can be used to estimate the camera pose T_{World}^{Cam}.

The class of solutions for 3D-2D pose estimation from known correspondences is called Perspective from N Points (PnP) [78] (Fig. 2.6). Since R,t can be estimated using six parameters, the minimal solution of this optimisation consists of finding 3 pairs of 3D-2D correspondences $(y_i^{2D}, x_i^{3D}) \subset Y \bigtimes X$, also called *P3P* [79].

However, using 3 point correspondences results in up to four geometrically feasible solutions, which can be disambiguated using an additional correspondence. Other methods have been proposed which are faster or can incorporate various numbers of correspondence pairs [80, 81, 82].

Analogous to the 3D-3D registration problem, obtaining the set of 3D-2D correspondences is difficult. When the moving model X is represented by a texture-less mesh, the neighbourhood of a feature point can be characterised from a geometrical point of view. However, a 2D feature point in an image has no geometrical information of its underlying shape Y, but only texture information. Similarly to Section 2.3.1, a larger correspondence set can be estimated using various heuristics and RANSAC [74] can be used to deal with the outliers, a popular approach in augmented reality applications [78].

2.4.3 Fine alignment

Once a good initialisation is obtained for the rigid parameters R,t, Eq. 2.10 can be refined using a non-linear optimiser to minimise the reprojection errors between the correspondences. Such a two-stage method using global and fine alignment approaches can be considered the ”gold-standard” in terms of both computational efficiency and accuracy [78].

A detailed state of the art overview of PnP methods and non-linear local optimisation approaches is presented in [78].

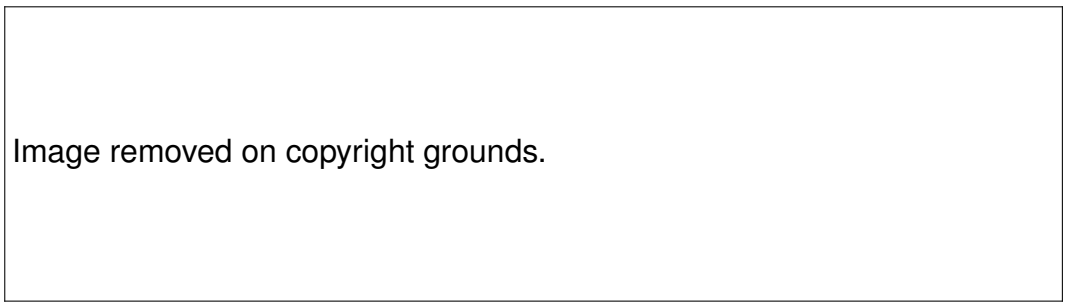

Figure 2.7: AR example using 3D-2D correspondences. The camera in the head mounted display (left) can analyse and augment the view of the user in real time. In this example, the world consists of a 2D planar marker (middle). The corners are detected by finding the intersection points of the lines fitted to the edges. Once the 2D -3D correspondences are known, a virtual object can be projected on top of the physical plane (right). Figure adapted from [20].

2.5 Challenges for registration in laparoscopic liver surgery

Several traditional computer vision registration methods have been introduced. Most of them make strong assumptions on the geometry and position of the shapes i.e. no occlusion, no symmetries, distinctive geometry, no deformation, good initial alignment. Since these algorithms have been developed for general problems in the vision community, they have been heavily used in multiple applications - including the augmented reality example presented in Fig. 2.7. However, initial registration is especially challenging in laparoscopic liver interventions [35, 67] due to:

- **Deformation** - The patient anatomy deforms significantly between the pre-operative scan and the intra-operative acquisition. Firstly, the position of the body is changed between the two time points. Secondly, the pneumoperitoneum changes the location and shape of the liver (see Fig. 2.8). Other sources of deformation are breathing motion and heartbeats. Multiple solutions have been proposed for deformation estimation (more details in Section 3.3). However, they all rely on a good global rigid alignment for initialisation, which is challenging to compute automatically.

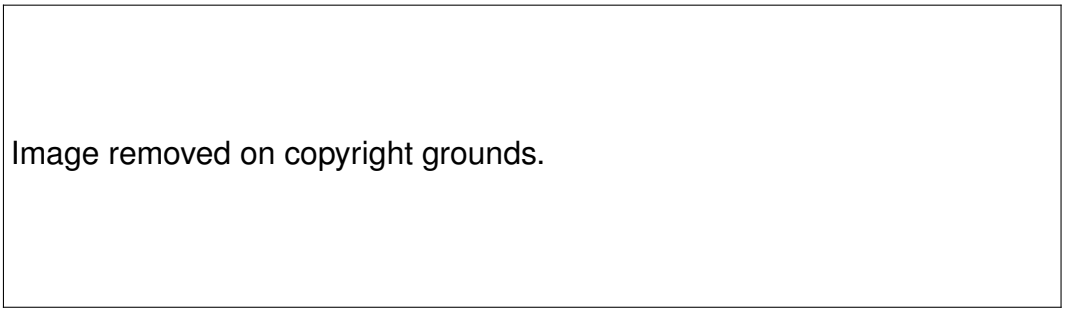

Figure 2.8: Bio-mechanical simulation of the abdominal wall under gas insufflation for registration in laparoscopic interventions. Colour coding: abdominal wall - red; liver - green; surrounding organs - purple. Left) Abdominal wall before gas insufflation; Right) Abdominal wall after gas insufflation. Images from [21].

- **Partial views** - As discussed previously, inherent challenges to laparoscopic interventions are the constrained vantage point and limited field of view,

which result in partial views of the organ of interest. Intra-operative 3D surfaces were observed to represent around 30-50% of the visible liver surface [83]. Furthermore, depending on the intra-operative acquisition method, the liver surface might have missing data in inaccessible regions.

- **Lack of prominent features** - The texture on the surface of the liver is repetitive, making it difficult to detect reliable feature points. Moreover, it varies with patients and disease (see Figure 2.9). Consequently, point cloud reconstruction from the matching of intensity based descriptors might prove difficult for livers with homogeneous appearance. Moreover, small partial views of the liver surface reconstruction are usually almost flat, thus lacking prominent features from a geometrical point of view as well.

Figure 2.9: Variation in the appearance of liver surface for healthy tissue and cirrhosis. Figure from [22].

The main assumptions in the global alignment methods described above (see Sections 2.3.1 and 2.4.2) is that the 3D-3D or 3D-2D point correspondences can be automatically, reliably detected in the same location in both modalities. In cases of deformation or partial visibility, this task becomes extremely challenging. As such, there is no established solution in the literature for the correspondence estimation. Instead, the global registration stage is in the majority of AR systems either partially or fully manual.

A discussion on the most relevant registration solutions in laparoscopic liver registration is presented in the next chapter.

Chapter 3

Registration in Laparoscopic Liver Surgery

Registration for laparoscopic liver surgery can be summarised in the following steps:

Pre-operative

- acquisition of a pre-operative scan
- segmentation of the 3D model of the liver, major vessel trees and any abnormalities from the pre-operative scan
- pre-operative planning of the intervention, based on the extracted 3D models

Intra-operative

- initial registration of the intra-operative and pre-operative images [35, 26, 84, 83]
- visualisation of the fused information [67, 17]

A long-term augmented reality system with real-time updates would also include several additional stages once an initial registration is achieved:

- tracking of surgical instruments - using either optical, electromagnetic or vision based trackers [85, 86]

- registration update in real-time for the rest of the intervention [83, 87, 30, 66, 88, 89, 90, 91, 92]

- modelling the real-time interaction of tissue and surgical tools, along with changes in topology [93, 94, 95]

Literature reviews covering several of these stages have been published [96, 85, 67, 17].

A major step towards automating an augmented reality system is to accurately and automatically estimate the initial registration. More specifically, this thesis focuses on global registration for laparoscopic liver surgery for which there are no solutions which satisfy the clinical requirements outlined in Section 1.4. Moreover, it represents the main bottleneck involving manual intervention in most augmented reality systems. Once a good initialisation is obtained, other fine alignment methods can be employed to refine the registration and to track it in the next few frames potentially using optical/ EM/ vision-based tracking. Alternately, if the automatic initial alignment proves to be fast enough and robust to the presence of instruments, it could be run repeatedly in order to obtain a real time up-to-date augmentation.

This chapter provides an overview of the state of the art of several components relevant to the initial registration stage in augmented reality systems for liver surgery. Firstly, several strategies for pre-operative planning are discussed (Section 3.1). Secondly, if required, the intra-operative surface can be reconstructed from the collected data (i.e. using methods reviewed in [97, 86]). Once both 3D surfaces (pre-operative and intra-operative) are available, the 3D-3D surface based registration component can be investigated. Alternately, 3D-2D registration can be estimated by matching the pre-operative 3D model with a single 2D image. A review of these approaches in the laparoscopic liver surgery field are discussed in Sections 3.2 and 3.3), along with several relevant innovative deep learning techniques in Section 3.4. Finally, since validation is extremely important, Sections 3.5 and 3.6 present an analysis of the current accuracy level of registration methods in the field, along with the role visualisation plays in AR systems. More specifically, the published methods will be analysed with respect to the challenges specific

to laparoscopic interventions introduced in Section 2.5 and the translation of AR systems into a clinical setting (Section 1.4).

3.1 Pre-operative planning

While pre-operative planning of an intervention is extremely important, a fifth of liver surgery cases prove to be more difficult than anticipated [98]. Changes made to the pre-operative plans during the procedure have been shown to increase the risk of adverse clinical events [98]. Computer vision tools have been investigated to improve the understanding of complex anatomies and allow for interaction with the patient-specific models and annotation. A better planning strategy is crucial for reducing the risk of complex laparoscopic liver interventions.

Figure 3.1: Virtual reality tool for pre-operative planning. A virtual screen shows a 3D rendering of the processed CT scan including the liver surface, vessel trees and tumour (yellow). The surgeon can control the transparency of each 3D model with a handheld device. The abdominal CT scan is visible on the right. Figure from [23].

Planning and simulating the procedure in a virtual reality framework [23, 17, 63] have been proposed (Figure 3.1). The surgeons are able to explore the patient-specific anatomy before each intervention. Moreover, multiple information streams can be aggregated and presented simultaneously (i.e. pre-operative CT scan, segmentations, annotations, laboratory results, patient information) in order to facilitate diagnostics, treatment planning, data exploration and education [23]. Such an approach can enhance the understanding of complex patient anatomy and act as an

education tool for less experienced surgeons [23].

Another stream of work has focused on pre-computing an optimal port placement [88] based on the patient specific models. Figure 3.2 shows an overlay of the internal anatomy projected on the patient skin in order to guide the surgeons [24, 63]. The insertion of the trocars for the camera and the instruments can be estimated using these pre-operative 3D models, for optimal triangulation with respect to the lesion [24, 63]. Such a task can be especially useful when the tumour is located in an area which is considered difficult to access in a laparoscopic setting. An AR guided trans-thoracic approach has been successfully reported for resectioning a tumour in segment VIII by using such pre-operative planning of the trocar positions [24].

Alternatively, various computations can be performed on the pre-operative scans and stored. Before the surgery, there are fewer time constraints, since there is at least a day available between obtaining the CT scan and the beginning of the intervention. A structural analysis of the vessels from the CT scan can then be super-imposed intra-operatively [99]. Liver deformation can be pre-computed by simulating multiple liver positions [100] or by taking into account the gravity load and the pneumoperitoneum pressure [28]. The resection area can be precomputed by taking into account the desired surgical margin [88, 32].

Another approach consists of pre-operatively planning for optimal data collection during surgery, such as [101, 102]. David Simon's [101] doctoral work looks into high accuracy registration in computer assisted open surgery. He proposes a framework for intelligent data selection (IDS) based on the observation that specific locations lead to better constraints for registration rather than gathering large and random collections of data. This can be quantified by analysing how much a set of points can constrain a surface based rigid registration algorithm. Ma et al. [102] extended Simon's approach by providing an alternative derivation for his constraint analysis. Ma et al. [103] later incorporated their analysis in a unified method for registration and point selection using a particle filter, which estimates the rigid registration parameters and maximises how well constrained the system is

Figure 3.2: Pre-operative planning of port placement with augmented reality for a transthoracic laparoscopic intervention. A) Testing various port positions. B-C) Marking on the patient skin of the chosen port. Figure from [24].

from a geometrical point of view. Their method is especially interesting because the registration is refined with each new measurement, thus providing the user with suggestions on which data to collect in order to increase the registration accuracy.

A framework was also proposed to minimise the intra-operative registration error by optimising the patient positioning and C-arm orientation [104]. The clinical interest was towards registering planar X-ray images with 3D CT/MR scans for guiding an endoscopic procedure concerning pancreatobiliary ducts. Moreover, localisation errors from selecting 2D and 3D anatomical landmarks were analytically derived in order to asses the error propagation in rigid registration [104]. Their approach can be extended to inform optimal protocols for data collection in order to achieve a desired error during the procedure. Moreover, having an uncertainty measurement on the obtained registration can help surgeons understand when to trust the provided guidance and when to attempt a better alignment.

Other related planning methods for data collection have been proposed in the computer vision community [105, 76, 106]. Low et al. [105] propose an ICP registration predictor which can output absolute error bounds and use it as a selection criterion for view planning in object reconstruction for range acquisition of indoor

environments. Rusinkiewicz and Levoy [76] propose a point selection strategy for ICP which maximises the distribution of normals in a sphere. Gelfand et al. [106] extend the method to maximise both translational and rotation constraints on the transformation.

A different approach to investigate the registration error variation with respect to different data collection strategies was proposed by [107, 108] in CT based navigation on the vertebra and neuronavigation respectively. They repeatedly select points in different configurations and simulate hundreds of registrations to observe the correlation between the regions with the registration accuracy. While they highlight a specific protocol for the surgeons, the recommended configurations are difficult to generalise to other organs.

A recent review [109], including 11 studies with 497 patients on pre-operative planning and simulation, reported that in a third of the studies the planned resection area was changed after seeing the simulation. Moreover, the liver volumes predicted in the simulation along with the surgical margins were correlated highly with the true values in the 8 studies which reported these values. While all studies had a version of pre-operative planning (i.e. either segmentation of the CT scan, virtual resection planned by the surgeon, automatic volume measurement), they are currently not translated to the intra-operative scene effectively [109]. The review concluded that it is essential for more research to be aimed towards providing automatic and intuitive navigation intra-operatively. The current solutions for global and fine alignment in laparoscopic liver surgery are discussed in the next sections.

3.2 Global alignment

The global rigid registration of the pre-operative 3D image and the intra-operative scene has been investigated through methods that rely on fiducials, user interaction or through fully automated methods.

Several approaches propose the use of fiducials, either on the patient skin [110], or on the organ itself [111] in open surgery. These fiducials can be seen in both modalities. Another more robust option, which is applicable in laparoscopic

interventions, would be to attach metabolisable fluorescent markers on the organ [112]. However, these strategies are disruptive to the clinical workflow since they require the acquisition of an additional CT or MRI scan immediately before the intervention.

Surface acquisition of the intra-operative scene has been proposed as an alternative. Several strategies have been developed using laser range scanners [113], optically tracked probes [114], time-of-flight (TOF) data [35] and stereo reconstruction [26, 87]. Once the surface is acquired, the clinician is required to delineate salient anatomical features leading to a point based initial alignment [113, 114] or to a more complex non-rigid optimisation framework [83]. Another option to obtain the rigid alignment is to manually rotate and translate the 3D pre-operative image until it fits the intra-operative data [26, 87] or to use heuristics regarding the location of the trocars and the insertion of the laparoscope in the abdomen [115, 116]. While some level of user interaction is needed for these approaches, it is generally more intuitive and faster to select salient features than to manually set the six degrees of freedom associated with a rigid transform.

Hybrid methods have been proposed using cone beam CT (CBCT) and fluoroscopy [117] as bridging modalities between the laparoscopic camera and the pre-operative CT, which delivers an additional radiation dose to the patient. Bernhardt et al. [118] develop an automatic method of registering an intra-operative CT scan to a single laparoscopic image. Feuerstein et al. [119] propose using intra-operative cone beam CT and optical tracking to register directly to the laparoscopic view without using pre-operative information. While their methods achieve promising results, they are based on advanced hardware which might not be available in most clinical settings.

Finally, fully automated techniques have been proposed for 3D-3D registration when using intra-operative surfaces [84, 35]. Fusaglia et al. [84] develop an exhaustive search over the principal directions of the intra-operative surface acquired with a tracked laser pointer in laparoscopic liver surgery. Dos Santos et al. [35] introduce a new automatic method to establish surface correspondences between the 3D pre-

operative mesh and the intra-operative surface acquired with a TOF camera in open liver surgery. While their validation for deformation under breathing motion can be sufficient for open surgery, livers in laparoscopic interventions undergo significant general deformation due to pneumoperitoneum (Section 2.5). On top of introducing additional tools into the clinical workflow, both methods rely on large surfaces of the liver being visible. As such, it is unclear how they would be translated to laparoscopic interventions.

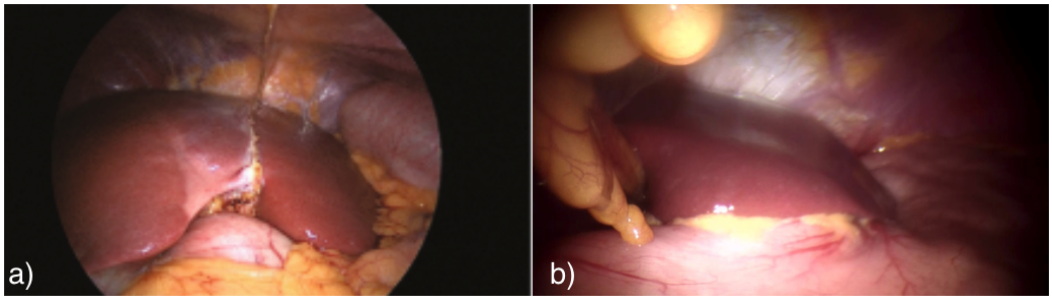

Figure 3.3: Comparison of visible liver surface - a) The falciform ligament was cut and a large portion of both liver lobes is visibile (Figure from [25]); b) The falciform ligament is present and the surface visibility is decreased to a liver lobe or less.

On the other hand, obtaining an automatic global alignment in 3D-2D registration is extremely challenging. In general, this step is performed manually, by having a user specify the 6 degrees of freedom [27, 63]. This could be extremely difficult, depending on the visibility of the organ and the ambiguity in the z direction. A semi-automatic technique has been proposed which incorporates prior knowledge about the organ [25]. Since in most laparoscopic interventions, the camera is inserted through an incision around the belly button, the liver can be transformed to a canonical pose before the surgery. Once such an approximation is made, the clinician has to annotate on the laparoscopic image where the anterior ridge, silhouette and the line where the falciform ligament was cut. These contours can then be matched with the pre-operative 3D surface. Their approach works well if applied after the falciform ligament is cut and a large portion of the liver is visible (Figure 3.3). However, not all interventions require cutting the falciform ligament. In these cases, large portions of the liver are occluded and matching the 3D-2D contours is

not straightforward, leaving manual alignment as the only reliable option.

3.3 Fine alignment

The resulting transformation from global alignment methods can be further improved by local refinements algorithms such as [75, 36] in order to get a rigid [120, 113, 26, 121] or a non-rigid alignment [87, 83, 116] between the two modalities.

3.3.1 Rigid alignment

Table 3.1: Fine rigid registration methods in open liver resection (OLR) and laparoscopic liver resection (LLR)

Paper	Method	Initial alignment	Surgery
Cash 2005 [120]	DIRR[1]	annotations (PBR[2])	OLR
Clements 2008 [113]	wICP[3]	annotations (PBR)	OLR
Rucker 2014 [66]	wICP [113]	annotations (PBR)	OLR
Thompson 2015 [26]	ICP	manual	LLR
Fusaglia 2016 [84]	ICP	automatic (exhaustive search)	LLR

[1] DIRR = Deformation Identifying Rigid Registration
[2] PBR = Point Based Registration
[3] wICP = weighted Iterative Closest Point

A significant amount of work has been done towards achieving good estimates for rigid shape matching, where the two shapes have to be aligned using an Euclidean transform (six degrees of freedom - rotation and translation).

Due to its good performance in practice and convergence properties, ICP has been widely applied in many applications including pre-operative to intra-operative data registration (i.e open [114, 113] or laparoscopic [26, 84] liver surgery - Figure 3.4).

ICP relies on a good initial alignment, which can be provided using the global methods introduced in Section 3.2. Anatomical landmarks can be incorporated in a weighted ICP scheme [113] to make the algorithm more robust. Table 3.1 summarises the main methods proposed in the literature for fine rigid alignment in liver surgery.

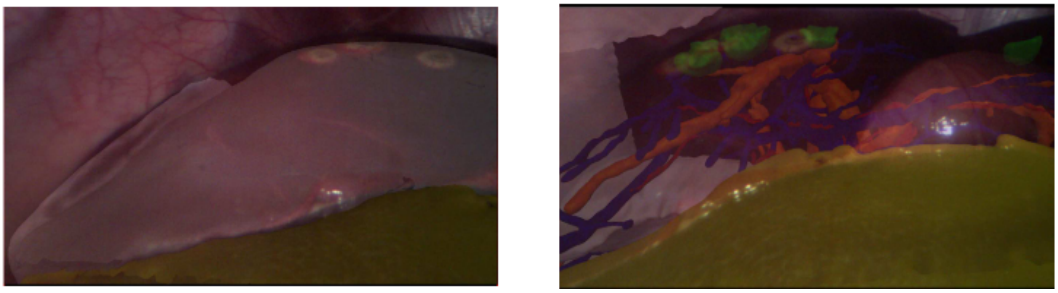

Figure 3.4: Augmented reality overlay during a porcine liver resection [26]. A global registration was performed manually, followed by ICP for rigid refinement resulting in an accuracy of 8.6 mm. Left) The overlay shows good agreement of the liver surface edges with the projected 3D model. Right) The overlay includes the vasculature and tumours identified in the pre-operative scan. Figure from [26].

Rigid transformations are the current *de facto* standard in AR for minimally invasive surgery. However, the liver is known to undergo significant deformation due to pneumoperitoneum, heartbeat and breathing motion in the initial stages of the surgery (see Section 1.3). Even though the respiratory motion varies for each patient, it manifests on the liver as a translational motion in the cranio-caudal direction with magnitudes between 10 mm and 26 mm [122]. Models for breathing motion correction have been proposed [10]. On the other hand, the motion induced by heartbeats is much smaller, on the scale of 1 mm [123]. A recent study [124] measured liver deformation in 25 patients who had to convert from laparoscopic liver surgery to open surgery. They found significant deformation between both pre-operative and intra-operative stages and laparoscopic and open liver surgeries. Several techniques deal with the deformation induced by pneumoperitoneum explicitly [125, 126, 21] by estimating liver deformation from pairs of non-insufflated and insufflated scans of livers [125], or simulating the abdominal wall under pneumoperitoneum for port placement [126] or registration [21]. As the intervention progresses, further deformation occurs in the interaction with the surgical instruments. A review of the most relevant published solutions for the non-rigid registration is presented in the next section.

3.3.2 Non-rigid alignment

While rigid registrations consist of estimating a global transform with six parameters, objects that undergo deformations require more complexity. One approach widely used in the computer vision community is to use parametric models of deformation, where a warping function is estimated from the correspondences found between the two shapes. This kind of deformation parametrisation is usually fast to compute since the non-rigidity is represented by a small number of control points.

The main focus in the augmented reality in laparoscopic surgery is to be able to overlay the subsurface structures present in the CT scan on the laparoscopic video. This step would imply propagating the deformation estimated from the surface of the organ towards the vessel trees and tumours. When using parametric models, one option to propagate of deformation inside the organ is to assume the internal structures are rigidly attached to the surface. For example, an AR non-rigid registration for the uterus has been proposed using an affine warp [115]. A method based on a parametric model (Coherent Point Drift [127]) was also proposed for the non-rigid registration of intra-operative CBCT to stereo reconstruction in liver surgery [128]. However, the liver undergoes significant soft-tissue deformation in laparoscopic interventions (see Section 2.5) and it is unclear how to best propagate the deformation to the sub-surface structures for parametric models. As a result, the research community has mainly focused on using bio-mechanical models of the liver to represent the deformation. Most optimisation schemes have an internal energy term which approximates the elastic deformation of the mechanical model and an external energy term for the registration constraints, given some boundary conditions. These approaches have higher complexity, can be more costly to compute than parametric models and they require prior knowledge of the physical properties of the different tissues in the organ. Nonetheless, they provide a biologically feasible propagation of deformation from the surface to the internal structures with visually coherent results, assuming enough intra-operative data is visible.

Usually, the biomechanical model is built from the 3D model extracted from the pre-operative CT scan. The boundary conditions are set based on prior knowl-

Table 3.2: Fine non-rigid registration methods in liver surgery

Paper	Deformation model	Registration constraints	Initial alignment	Surgery
Cash 2005 [120] , Cash 2007 [114]	homogeneous biomechanical model	support regions, point correspondence	annotations (PBR[1])	OLR
Clements 2007 [100]	homogeneous biomechanical model	support regions, point correspondence	assumed given	OLR
Dumpuri 2010 [129]	homogeneous biomechanical model	extrapolated point correspondence	annotations (PBR [113])	OLR
Rucker 2014 [66]	homogeneous biomechanical model	support regions, extrapolated point correspondence	annotations (PBR [113])	OLR
Suwelack 2014 [30]	homogeneous biomechanical model	support regions, [optional] known landmark correspondences, potential field generated by intraop surface	automatic [35]	OLR
Adagolodjo 2017 [27]	homogeneous biomechanical model	2D liver contour, simulated pneumoperitoneum	manual + annotations	OLR
Allan 2016 [128]	parametric	point correspondence	manual	LLR
Haouchine 2016 [116]	homogeneous biomechanical model	liver contour, abdominal modelling, volume preservation	annotations	LLR
Plantefeve 2016 [83]	heterogeneous biomechanical model	anatomical landmarks, point correspondence	annotations	LLR
Koo 2017 [25]	homogeneous biomechanical model	liver silhouette, falciform, ridge line, point correspondence, shading	annotations	LLR
Heilseman 2017 [124]	homogeneous biomechanical model	falciform, left and right triangular ligaments, organ posterior	annotations (PBR [113])	LLR
Ozgur 2018 [28]	homogeneous biomechanical model	liver silhouette, falciform, ridge line, point correspondence, preop. pneumoperitoneum with gravity load	annotations [25]	LLR

[1] PBR = Point Based Registration

edge of the liver anatomy and certain assumptions for a specific intervention (i.e. prior knowledge of the support region of the liver in open surgery [66, 120, 100]). The registration constraints are usually given by the data observed intra-operatively. As a consequence, most methods differ by their choice of biomechanical model, boundary conditions or registration constraints used. Several techniques chose a homogeneous biomechanical model (liver parenchyma) [66, 30, 116, 25, 27, 66] or a heterogeneous model (liver parenchyma, Glisson's capsule and vascular tree) [83, 87, 130, 131]. A recently published study [132] comparing the two approaches reported no significant improvement in registration results and increased computational time for the heterogeneous model in simulations of varying realistic deformations of the liver. The registration constraints mostly used are anatomical features [83, 116, 25, 27, 124], shading cues [25], correspondences given by the surfaces [83, 87, 130, 66, 120, 129], volume preservation [116]. The liver contours can be manually annotated [83, 25, 28, 27] or automatically estimated from point cloud segmentation [116] in the intra-operative images. In the pre-operative model, the contours and anatomical features can be automatically labelled using atlas based transfer from a database of 3D models of CT scans [83]. Several methods aim to compensate by simulating the pneumoperitoneum deformation [125, 124, 21, 27, 28]. Another important factor to take into account is the gravity load on the pre-operative model, since the pre-operative CT scan is acquired in a different patient position than the laparoscopic intervention [28].

Several of the published methods can only be achieved in open surgery [120, 114, 129, 66, 100] since the intra-operative data is collected with a laser range scanner [120, 114, 129, 66]. While laser range scanners have been proposed for laparoscopic surgery [84], they are prototypes and not widely used. Moreover, these methods make assumptions which are only valid in open surgery, such as access to a large view of the organ, the existence of minimally deformed regions [120], or specific strategies to stabilise the liver by placing support material under its lobes [66, 120]. An interesting approach which uses an atlas built pre-operatively from a range of simulated possible liver deformations was proposed to estimate the intra-

operative deformation [100]. However, since it requires an exhaustive simulation set of plausible deformations specific for each surgical intervention, it is difficult to translate to clinical scenarios. A similar, more plausible approach [124] proposes to build a model of deformations based on random perturbations applied at several control points situated on support regions (i.e. ligaments, organ posterior where the liver makes contact with other organs). Thus, the main task intra-operatively is to find the set of pre-computed perturbation that best matches the intra-operative data.

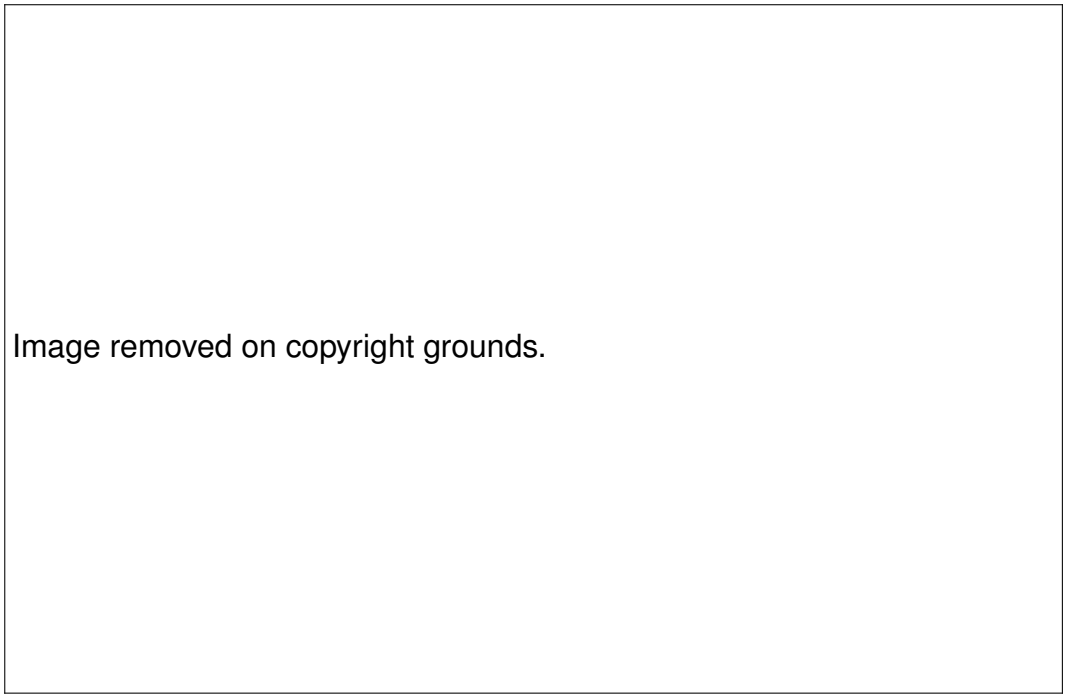

Figure 3.5: Registration formulation based on the liver contour proposed by Adagolodjo et al. [27]. Firstly (step 1) the deformation due to pneumoperitoneum is simulated before the intervention. Step 2 shows a rigid global alignment performed by a user which aligns the pre-operative 3D model to approximately match the intra-operative image. Finally, in step 3 the annotated 2D organ contour is used as a constraint for a biomechanical simulation of the liver deformation. Figure from [27].

A clear trend can be observed in the development of non-rigid registration algorithms presented here. Initial work in defining the registration constraints for the biomechanical modelling started with point correspondences [120, 100, 129, 66]. Since the deformed model is guaranteed to fit mostly around the corresponding points, different strategies to extrapolate a smooth deformation field to unseen regions during surgery have been proposed [129, 66]. Most of these methods were

applied in open surgery, where the intra-operative data acquisition was performed with laser range scanners or optically tracked probes, resulting in large surfaces with enough points. However, non-rigid registration techniques in LLR mostly use monocular or stereo laparoscopic cameras for data acquisition. While point correspondences from sparse correspondences can still be used as constraints [87, 83], the visible surfaces are smaller. Moreover, the liver surface differs greatly in appearance depending on the pathology - which implies difficulties in finding robust correspondences between images of homogeneous liver surface (such as Fig. 2.9). Strategies can be employed to deal with noisy and sparse surface reconstructions, such as using surfaces fitted to the point cloud [116, 83, 131] and using control points [131]. However, if the liver surface lacks clear features, the reconstructed point cloud will not represent the underlying surface accurately, leading to failures in non-rigid registration methods which rely on such representation [116, 83]. This issue was also observed in [133]. As a result, the limited data available in laparoscopic interventions can easily lead to an under-constrained systems when estimating the liver deformation with many degrees of freedom. So, other constraints which are motivated by the organ shape have been proposed - such as anatomical landmarks [83, 25, 28] or liver silhouette [116, 27]. Several of these contour-based approaches [27, 28, 25] show promising results using the 2D contour as projective constraints and a biomechanical model for deformation (Figure 3.5). However, they are validated on synthetic data assuming 70% - 100% of the liver surface visible (Figure 3.6). While such visibility is achievable in open surgery [134], the other methods propose performing the registration in laparoscopic interventions after the falciform ligament is cut and the liver is less occluded [28, 25]. However, not all laparoscopic interventions require the removal of the falciform ligament, in which case approximately 30%-50% of the surface is visible [83]. To date, the only reliable global registration solution for these cases is rigid manual alignment.

Figure 3.6: Partial visibility experiment for the registration framework proposed by Özgür et al. [28]. Left) The field of view circles occlude the liver contours (anterior ridge in red and silhouette in yellow) approximately by 10%, 20% and 30%. Right) Their proposed method [28] is compared against another deformable [25] and rigid registration approach [17] with promising results for the tested levels of occlusion. Figure from [28].

3.4 Deep learning in laparoscopic liver interventions

Recently, deep learning techniques have been heavily explored in medical image analysis [135]. Multiple solutions have been proposed with impressive accuracy for segmentation, object detection, registration and diagnosis from medical images [135, 136, 137, 138]. The main driver behind such methods has been the availability of standardised datasets (i.e. MRI, CT, OCT scans) - detailed lists can be found in [136, 138].

In medical image analysis, the main difficulty in the progression of deep learning techniques is the lack of annotated clinical data [135]. While computer vision tasks can obtain annotations through large crowd-sourcing projects, this is not achievable in medical imaging. Specific expertise is needed in order to provide reliable annotations with clinical relevancy.

Deep learning techniques could in theory fully automate registration of a 3D pre-operative model to laparoscopic videos. However, such registration datasets are not currently available and they are extremely difficult to build. Since each surgery is slightly different (i.e. pathology, organ appearance, organ geometry, patient age, type of intervention), a large dataset with multiple examples for each task is needed. Since liver surgeries are especially challenging from the point of view of inter-subject variability, small datasets will result in the trained network being incapable

of generalising well to new examples. For example, a liver segmentation CNN trained on approx. 2000 images across 13 interventions reported poor generalisation to cases that vary too much from the training data [11].

Since collecting more data requires extensive manual work and possibly years of surgeries, a solution could be provided by synthetic data. Preliminary results in estimating the deformation of the pre-operative 3D model from a 3D intra-operative surface have been published recently [139, 140]. In these cases, the network can be trained with simulated deformations of a 3D liver mesh and a partial point cloud representing a 3D intra-operative surface. Such approaches are encouraging since they can use more data than is generally available.

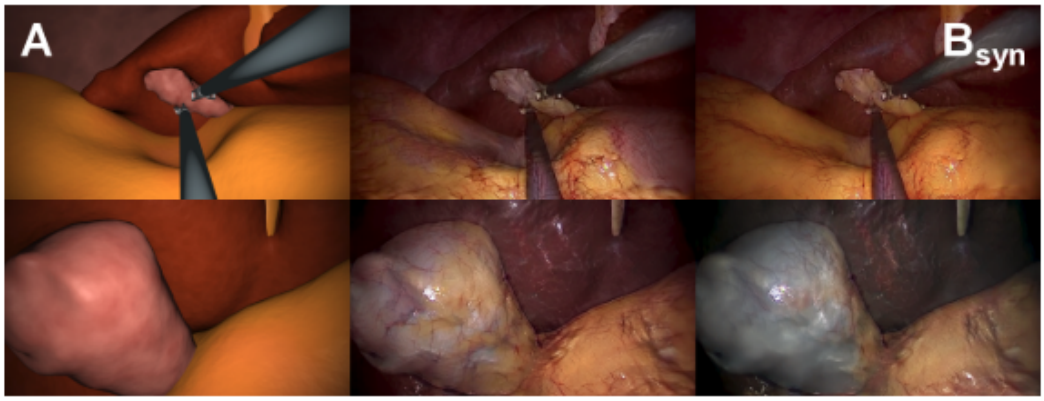

Figure 3.7: Image-to-image translation proposed by Pfeiffer et al. [5]. The images obtained from a computer simulation of a laparoscopic liver intervention (domain A) are translated to look like real laparoscopic images (domain B_{syn}) using deep learning. The middle and right columns show various styles for the same simulated scene. Figure from [5].

Recent studies propose style transfer to enhance the realism of surgical simulations [141, 5]. Figure 3.7 shows an example of such image translation for laparoscopic liver surgery. In [5], they ensure the labels generated for the synthetic data are preserved when the style of the images changes. They achieve high accuracy in liver segmentation without requiring any data annotation and alternatively show that pre-training a model on synthetic data can increase the network's performance significantly. The synthetic dataset is publicly available (more details in Section 3.5.1).

3.5 Validation

Tumour resection is usually performed by taking into account a safety margin between 1 mm and 1 cm [142], such that all cancerous cells are removed and the probability of reoccurrence is minimised. Several studies suggest the registration accuracy should be better than 5 mm [66, 67] or 3 mm [26]. Higher registration errors would imply the surgical margins exceed their limit and the intervention is at risk. However, there are multiple factors that influence the clinician's perception of the alignment accuracy - i.e. misregistration, visualisation. In this section, the various protocols used to validate registration algorithms are reviewed, with their respective errors. Next, the visualisation of registration is discussed along with its impact on the user's perception of accuracy.

Table 3.3: Validation for rigid methods

Paper	Intervention type	Phantom	Animal	Human
Cash 2005[120], Cash 2007[114]	OLR	mean TRE <3mm (6 targets, 2 deformations)		mean RMSE = 7.72 mm (4 cases, LRS) mean RMSE = 5.41 mm (7 cases, tracked probe)
Clements 2008[113]	OLR	mean TRE = 2.9 $\pm$ 0.5mm (250 perturb., no deformation)		mean RMSE = 6.65 mm (6 cases)
Rucker 2014[66]	OLR	mean TRE = 9.87mm (28 targets, 4 deformations) mean TRE = 8.00 mm (58 targets, 1 deformation)		mean RMSE = 4.78 mm (5 cases)
Thompson 2015[26]	LLR	mean TRE = 2.9 mm (9 landmarks, no deformation)	mean RMSE = 9.43 mm (3 porcine cases)	
Fusaglia 2016[84]	LLR	mean TRE = 3.2 $\pm$ 0.57mm (3 phantoms, no deformation)		
Thompson 2018[29]	LLR			4 live registrations (manual): mean RMSE = 25.0$\pm$ 8.8 mm 5 retrospective (manual): mean RMSE = 19.4 $\pm$7.4 mm 1 retrospective (manual+ICP): RMSE = 12.3 mm

An overview of the accuracy values achieved in the literature for rigid and non-rigid registration in laparoscopic liver surgery can be found in Tables 3.3 and 3.4 respectively. Note that these errors are not directly comparable, since they are computed using different techniques on several datasets. The target registration error (TRE) represents the mean distance between the original and estimated position of a landmark not involved in the registration optimisation. So, it is the ideal measurement for the performance of a registration algorithm. However, it cannot be easily computed in a clinical setting, where there is no ground truth available. Instead, a

metric used frequently is the root mean squared error (RMSE) which is computed between the closest corresponding points of the two surfaces. While low values of RMSE represent a good match of the surfaces, they do not necessarily indicate the correct position. It was previously shown that RMSE and TRE are not correlated [143].

In general, rigid registration methods in laparoscopic liver interventions usually achieve approximately 10 mm accuracy in synthetic and phantom experiments. Further refinement of the registration can be achieved using non-rigid alignment. Despite reaching accuracy values under the desired 5 mm, non-rigid methods can be difficult to translate to a clinical setting. In image guidance systems it is critical for the surgeon to be able to confirm whether the registration algorithm provides a feasible solution [26]. The functioning of rigid methods can be comparatively easy to evaluate. On the other hand, with increasing complexity in the deformation model less reliability is observed in the error metrics used, due to the ambiguity introduced by adding more degrees of freedom. In practice, optimisation functions for deformation estimation have several local minima, with multiple deformed configurations leading to an optimal solution. Consequently, non-rigid techniques are highly dependent on a good initial rigid alignment.

Animal models can represent realistic anatomy, similar to human patients [144]. Markers can be implanted [145] or artificial lesions can be created using ablation [26] on the liver surface, in order to serve as landmarks for points based registration or validation. However, due to ethical issues, high costs and licences for animal testing, they are not easy to introduce in the clinical validation needed for translation [146].

Another important factor to consider is the different error sources that might contribute to a registration in a clinical context. Thompson et al. [29] quantify the errors of various alignment strategies (point based registration, manual registration, ICP retrospective alignment) measured in human clinical data (either live during surgery or retrospectively). The errors could accumulate from various sources such as the tracking system, laparoscope calibration, point picking in video, point picking

Table 3.4: Validation for non-rigid methods

Paper	Intervention type	Synthetic	Phantom	Human
Cash 2005[120], Cash 2007[114]	OLR	robustness to initial registration, sensitivity to biomechanical modelling parameters	mean TRE <4.00 mm (6 targets, 2 deformations, partial and full surface)	qualitative (4 cases)
Clements 2007[100]	OLR	mean mesh TRE = 3.51mm (5 deformations)	mean TRE = 3.4 mm (6 targets, 2 deformations)	
Dumpuri 2010[129]	OLR		mean TRE = 1.1mm (full surface) mean TRE = 2.45mm (partial) (43 targets, 2 deformations)	mean RMSE = 2.45 mm (8 cases)
Rucker 2014[66]	OLR	robustness to initial registration, robustness to parameters, sensitivity to support surface, effect of data coverage, effect of gravity, effect of model linearisation	mean TRE = 3.3mm (28 targets, 2 phantoms, 5 deformations)	mean RMSE = 2.74 mm (5 cases)
Suwelack 2014[30]	OLR	robustness to simulated breathing deformation, robustness to model parameters, stability and speed	mean TRE = 2.3mm (full surface) mean TRE = 5.05mm (partial) mean TRE =8.7mm(intra-op recon.) (OpenCAS dataset)	
Adagolodjo 2017[27]	OLR	robustness to initial registration, robustness to partial size up to 70%, computational time	mean 3D err = 1.2 mm (4 deformations)	qualitative (2 cases)
Allan 2016[128]	LLR		mean RMSE = 0.93mm (2 phantoms, no deformation) mean RMSE = 0.75mm (2 phantoms, deformation <2 mm)	
Haouchine 2016[116]	LLR	4 mm <TRE <12 mm (5 tumours, pneumoperitoneum + noise), comparison with two other methods, computational time		mean Hausdorff = 26.21 px (dist-liver contours) (2 cases)
Plantefeve 2016[83]	LLR	mean TRE <3 mm (partial surf >20% pneumoperitoneum), homogenous vs heterogenous models, healthy vs cirrhotic livers, robustness to model parameters, robustness to partial size up to 10%		mean Hausdorff = 1.1mm (dist-surf. mesh and intra-op) (2 cases)
Koo 2017[25]	LLR	mean TRE <10 mm (2 livers, with deformation), robustness to initial registration		qualitative (1 case)
Heilselman 2017 [124]	LLR		mean TRE = 6.4 mm, [66]:mean TRE = 7.9 mm, [113]: mean TRE = 14.7 mm, (1 phantom, 3 states) robustness to surface visibility	local surface correction in MIS & open clinical data (25 cases)
Ozgur 2018[28]	LLR	robustness to initial registration, robustness to contour visibility	mean 3D err = 11.2 mm, [25]: mean 3D err = 13.4 mm [17]: mean 3D err = 18.9 mm	qualitative (2 cases)

in CT, manual registration, OR conditions, static deformation (pneumoperitoneum) or dynamic deformation (breathing). For example, when using a setup similar to Fig. 2.1 with optical tracking and stereo reconstruction of the intra-operative scene, a manual alignment performed during surgery would also include errors in: laparoscope tracking (from merging multiple surface patches together), laparoscope calibration, point picking in the video and the CT, manual registration, OR conditions, pneumoperitoneum and breathing. They report average RMS reprojection errors of 25.0 mm ($\pm$ 8.8) over 4 cases with live manual alignment as opposed to 19.4 mm ($\pm$ 7.4) over 3 cases with retrospective alignment. The estimated errors from point triangulation and laparoscope calibration averaged to 10.2 mm ($\pm$ 3.3) and 7.5 mm ($\pm$ 5.2) respectively, over 9 cases. More details can be found in [29].

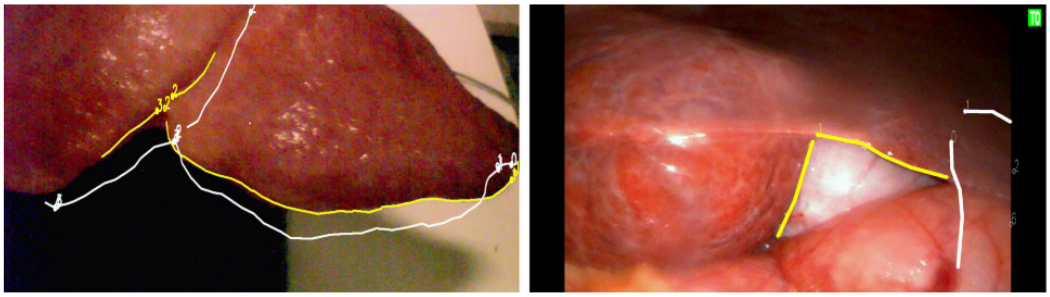

Figure 3.8: Example of liver contours on a phantom (left) and on *in vivo* data (right) which could be used as a measure of registration accuracy, as proposed by Thompson et al. [29]. The yellow lines denote the annotated contours on the images, whereas the white lines are obtained by projecting the corresponding 3D contours given the current registration estimate. Figure from [29].

It is essential to provide the surgeons with an indication of the error of a given overlay or a measurement of confidence in the current registration. In a qualitative way, rigid registrations allow a rapid evaluation of the error by checking if the projected contours of the organ match the ones in the laparoscopic image. Moreover, it has been shown that the projection error between the visible contours shown in Figure 3.8 is a reliable predictor of sub-surface TRE errors in rigid registration [29]. Note that the organ contours do not need to correspond to anatomical landmarks. Such a quantitative measure could allow future AR systems to provide an overlay of the tumour and associated resection area from the pre-operative planning along with a confidence level. Thus, the surgeon would have a better understanding of when

to trust the AR system and when to proceed with alternative strategies (i.e. re-initialising the registration, intra-operative ultrasound). While for rigid alignments, the projection and true organ contours can provide a measure of the registration errors, it is still unclear how this will be achieved with non-rigid approaches where the pre-operative model is deformed.

3.5.1 Public datasets

Given the above challenges in the area of system validation, more reliable validation methods are required.

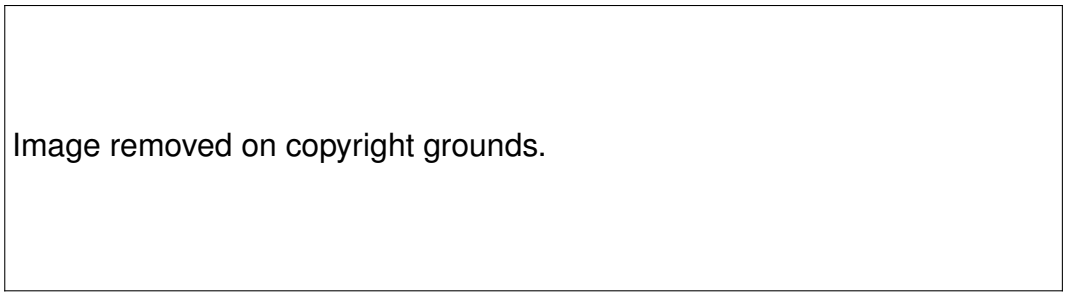

Figure 3.9: OpenCAS dataset. Left) Phantom experiment setup where a metallic rod creates a deformation. Middle) Initial (red) vs. deformed (blue) liver surfaces extracted from CT scans. Right) Initial surface (red) vs. intra-operative acquisition (blue) with a stereo laparoscope. Figure from [30].

OpenCAS - *Liver registration dataset* [30] is an example of a public phantom dataset for non-rigid liver registration with a deformation induced by an indentation (Fig. 3.9). It is extremely useful since it provides CT scans before and after the indentation, with the positions of subsurface targets (small Teflon balls). It also provides partial surfaces, amongst which, one is reconstructed from a stereo endoscopic camera. Thus, it allows for algorithm validation in conditions close to a real clinical scenario. However, the deformation induced to the phantom is unrealistic for laparoscopic interventions.

Another initiative is OpenHELP [31], a torso phantom dataset containing all the organs of the abdominal and thoracic cavities, with realistic material characteristics and anatomy (Fig. 3.10). At the moment, the 3D surfaces of the organs have been released along with information about how to create the phantoms. They can be used to simulate pneumoperitoneum successfully, validate computer-assisted

Figure 3.10: OpenHELP dataset. The CT scan of a healthy male (left) is used to generate 3D surfaces for all the organs in the thoracic and abdominal cavities (middle). These surfaces are then used to obtain synthetic phantoms (right). Figure from [31].

surgery methods and for surgical training. However, the data associated with these experiments has not been made public yet.

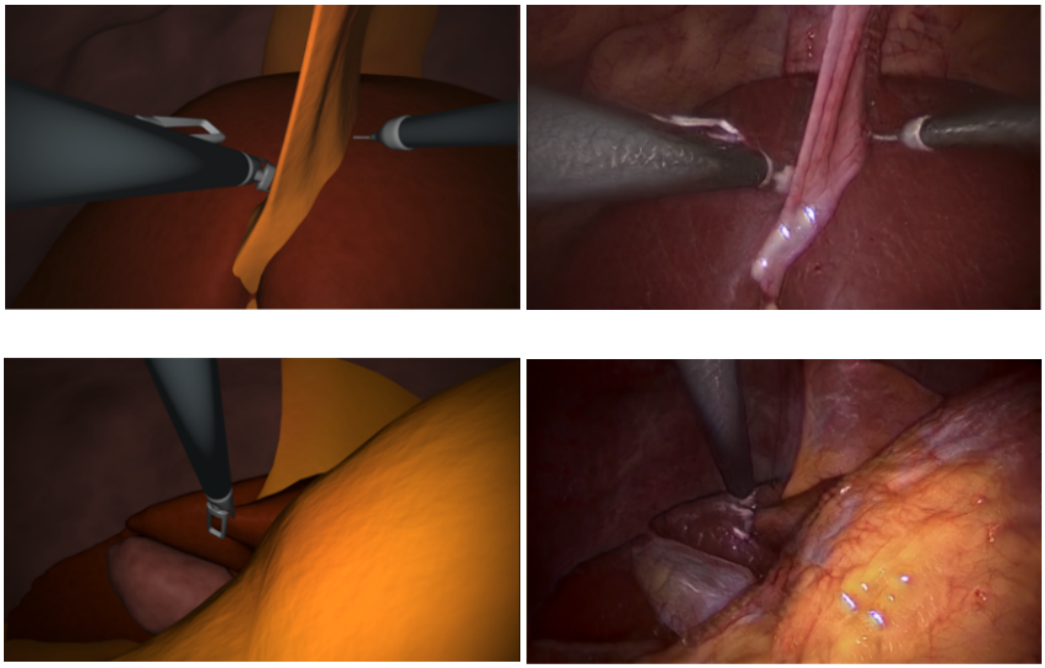

Figure 3.11: Examples of image translation from the synthetic environment (left) to photo-realistic laparoscopic images (right) from the public dataset released by [5]. Figure from [5].

A synthetic dataset of photo-realistic simulations of laparoscopic liver surgery is publicly available[1] [5]. The synthetic images are obtained in a computer-simulated laparoscopic environment and their appearance modified to resemble real laparoscopic images (Fig. 3.11). Two scenarios are considered separately, leading to 100 000 images from various camera viewpoints around the liver and another 100 000 images of complex tool-liver interactions. The synthetic images are generated from 10 different patient cases, thus accounting for varied organ geometry. Moreover, each case contains 5 different appearance styles with 2000 images for each style to account for various liver appearance. Multiple ground truth labels are provided: semantic segmentation (liver, fat, abdominal wall, tool shaft, tool tip, gallbladder), depth images, anterior ridge lines and normal maps. Such a public dataset could potentially advance the development of multiple tasks in laparoscopic interventions in the following years.

An in vivo porcine dataset has been released in order to evaluate surface-based liver registration [147]. They provide 13 deformed states of the liver with corresponding exploratory video and a reconstructed intra-operative surface. While it is a valuable resource, special care has to be taken due to challenges presented by porcine livers which do not appear in the human ones - such as collision and self-occlusion between the lobes. This cannot happen in human livers as the two lobes are fixed with ligaments and do not overlap.

A recent challenge was launched for validating sparse liver surface registration techniques[148]. Such benchmarking datasets can prove very valuable since they provide the same methodology for validation across different algorithms. Hopefully, it will be more straight-forward to compare future algorithms and understand their respective advantages and limitations better.

3.5.2 Clinical trials

Several small clinical studies have reported positive results using AR systems in various interventions - 4 laparoscopic interventions [29], 3 robotic segmentectomies (V, V and VI) [63], 9 laparoscopic left hepatectomies [149], 1 laparoscopic segmen-

[1] http://opencas.dkfz.de/image2image/

tectomy (VI) [150]. In general, the suggested resection area obtained with the AR system was validated with intra-operative ultrasound and in some cases with post-operative histological analysis [150]. Although a semi-automatic approach [25] was used in one intervention [150], the majority of cases rely on a user to perform the initial alignment manually [63, 149, 29]. Such manual alignment is repeated multiple times during the procedure as requested by the surgeon [63]. While these studies present encouraging results, it has been concluded that it is critical to develop automatic and intuitive initial registration techniques, in order to facilitate usability and translation to the OR [63, 118].

There are currently no large clinical trials published regarding the performance and validation of AR systems, as they are difficult to design. While most papers report registration accuracy and various errors, the required accuracy level will most likely depend on the clinical application [151]. In gastric surgery, a registration method was proposed which progressively improves with the additional landmarks visible during the intervention [65]. They report accuracies of approx. 13 mm with rigid alignment. However, their image guidance system shows both views (laparoscopic and aligned pre-operative data) side by side and not overlayed, stating that the surgeons can mentally compensate for the errors along with any deformation and breathing motion. Consequently, another key factor in how well AR systems will be translated into the clinical setting is the visualisation technique used.

3.6 Visualisation

Given an accurate registration, how can the fused information be displayed to the surgeon in order to offer a better understanding of the intra-operative scene without introducing distraction or clutter? Several aspects related to the perception of AR are discussed in the next section, with more comprehensive overviews to be found in [67, 17].

A perfect example of the importance of visualisation was reported in [32]. When the overlay shows the projected 3D model and lesion with transparency blending (Fig. 3.12 a)), the user assumes that the centre of the projected lesion

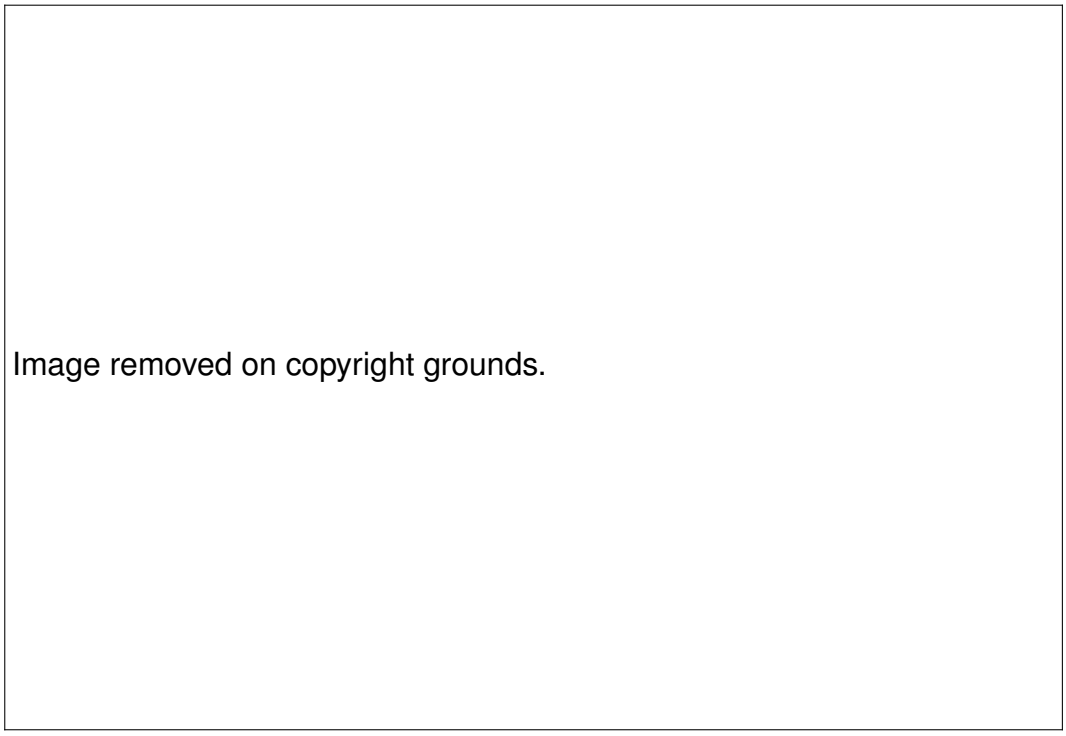

Figure 3.12: Importance of correctly overlaying lesion information. a) Laparoscopic image with the projected 3D models using transparency blending, b) Tool Access Visualisation - the resection ring is projected with respect to the incision tool port, c) Transparency Blending Visualisation - The centre c of the tumour as it appears on the laparoscopic view does not correspond to the true centre when the incision tool is used, d) Tool Access Visualisation - A safe margin is computed around the tumour and projected to the laparoscopic view from the perspective of the incision tool. Images from [32].

corresponds to the real centre of the sub-surface target (Fig. 3.12 c)). However, the projection simply shows the tumour from the perspective of the current camera pose and the incision tool always approaches the organ from a different angle (Fig. 3.12 d)). During a user study, the visualisation with transparency blending caused a lot of surgeons to miss small or deep lesions in the organ [32]. To solve this issue, they propose computing a safe margin around the tumour and estimate two projections: one with respect to the insertion point of the instrument and another one with respect to the laparoscopic camera. In this formulation, the position of the trocar for the incision instrument has to be provided during surgery. Thus, the surgeon would visualise directly the area to be resected as a ring on top of the image (Fig. 3.12 b)).

A critical aspect when discussing overlays is *inattentional blindness* - a phe-

nomenon which appears when people focus on a task and fail to notice events or objects that appear within their field of view. Several studies [33, 152] have investigated the effects of different visualisation strategies of overlays when the users have to perform a task.

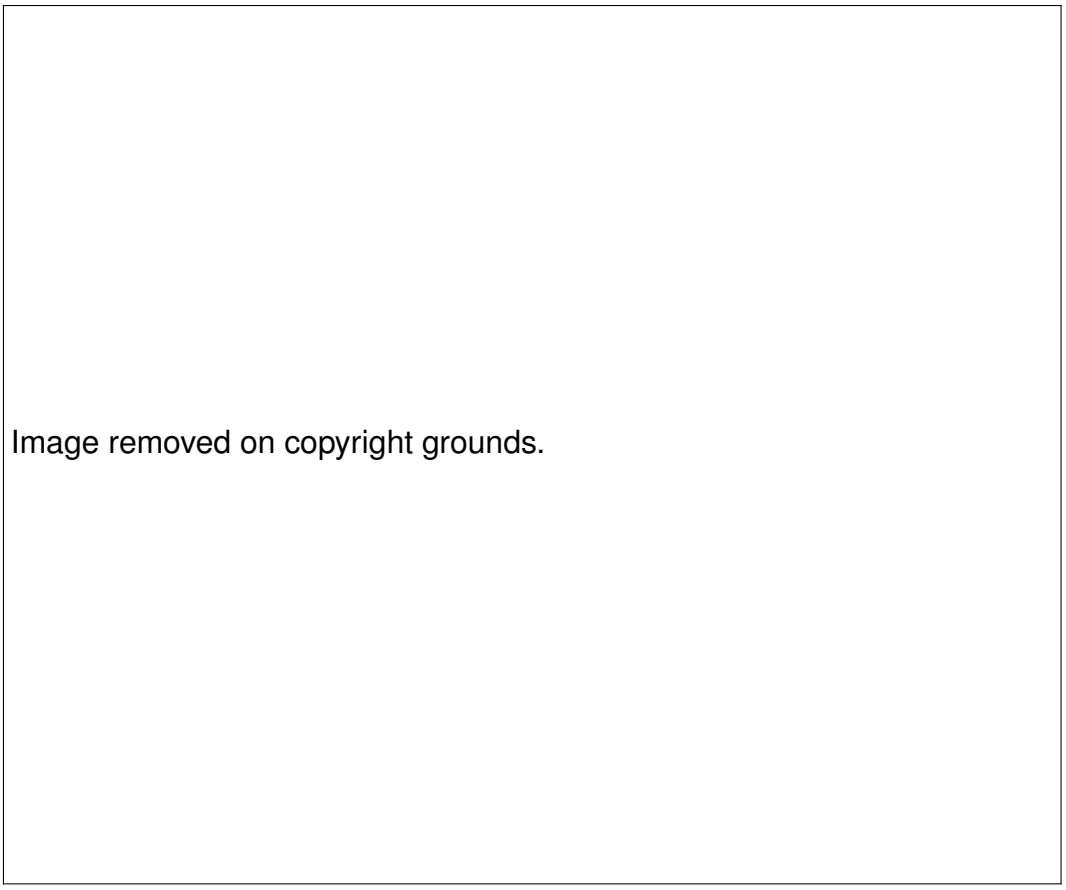

Figure 3.13: Various augmented reality visualisation strategies in a target localisation study [33]. A: No overlay; B: Always-on solid overlay; C) Always-on wire mesh overlay; D) Inverse realism overlay. Figure from [33].

Inattentional blindness was studied when inexperienced surgeons had to locate a target using several visualisation techniques (Fig. 3.13): 1) no image guidance, 2) triplanar display, 3) always-on solid overlay, 4) always-on wire mesh overlay, 5) on-demand inverse realism display [33]. The analysis showed that image guidance (especially overlays) can help with decreasing the time to complete the task and the total tool path length. Such results can be expected since the surgeon has no need to pause the procedure, turn away from the patient and consult the pre-operative scans. Inattentional blindness was found in all types of image guidance with 2),4),5) at a

similar reduced level. However, having an always-on solid overlay (3) lead to only 20% detection of unexpected complications. An interesting finding is that having a triplanar view leads to the same level of inatentional blindness as several overlay methods, which could indicate that other factors than the visualisation strategy are adding to this phenomenon.

Another study [152] investigated the perceptual limits of AR when using head mounted displays (HMDs) for manual tasks that need precision. Consumer HMDs, such as the HoloLens, present the virtual image at a fixed focal distance, generally more than 2m. However, for tasks that involve manual interaction, the real-world objects will be placed roughly at arm's length to the eyes. Consequently, the human eye cannot keep in focus both the virtual image and the real-world objects simultaneously. In such cases, blurring can appear in the periphery when the eye strains to focus on both images.

Visualisation in AR systems is an active area of research, as there is no consensus regarding the best visualisation strategy. Better understanding is needed of how to balance which information to highlight and when, as well as minimising the discomfort and risk introduced. An efficient AR system should show only essential information, which will depend on the clinical application.

However, while the visualisation strategy is important, the translation of an AR system in the clinics depends on many factors. There are currently few registration methods that fit the clinical criteria needed for an efficient AR system described in Section 1.4. As such, developing a clinically relevant, robust, automatic initial registration will be the focus of the following chapters.

3.7 Summary

While promising results have been achieved in the literature, the main aim of this thesis is to describe the tools and theory necessary for an automatic global registration of a pre-operative 3D model to the intra-operative scene.

Obtaining a good global alignment is critical for subsequent rigid and non-rigid registration fine alignment techniques. Multiple fine registration techniques

have been proposed in the literature, some of which are automatic and deal with the specific challenges of liver shapes (Section 2.5). As for estimating a global alignment, the majority of available options generally involve various levels of user interaction. An option is to manually rotate and translate the 3D liver surface model until it roughly aligns with the intra-operative model. Another option is for the user to manually pick corresponding points in both modalities. However, the accuracy and speed of performing these alternatives depends highly on the user's expert knowledge and their experience with the software. Moreover, such tasks would have to be performed during the intervention and possibly repeated multiple times. Consequently, solutions with low computational time become essential. Furthermore, starting with a good initial registration is important for long-term AR since the visualisation is generally updated based on tracking the camera position and not refined with respect to the pre-operative model. So, any initial errors in alignment will be carried over for the rest of the video. Overall, the currently available options might impede the translation of AR systems into the clinics, as they do not comply with the clinical criteria for an efficient and intuitive AR system introduced in Section 1.4).

A general strategy of the work presented in the remaining of this thesis is to maximise the amount of pre-computation before the surgery in order to speed up the intra-operative registration. The 3D models of the liver surface and relevant sub-surface structures are obtained before the surgery and anatomical landmarks can be annotated on the 3D surface models [153]. Additional pre-computation of the surface features, descriptors can also be performed and stored.

Firstly, following the intuition of Simon [101], the hypothesis is that some surface patches will contribute more to the final rigid registration than others, namely those with distinctive geometrical features. A pre-operative planning tool is proposed to select which views of liver surface will lead to a fast and efficient registration (Chapter 4).

Secondly, a fast, semi-automatic 3D-3D global surface registration is proposed which incorporates a novel contour-based constraint in order to deal with the partial

surfaces characteristic to laparoscopic interventions (Chapter 5).

Finally, Chapter 6 presents a fully automatic 3D-2D global registration pipeline. A novel semantic liver contour detection method leveraging the advances of deep learning techniques is followed by a contour-based rigid registration, with promising results in retrospective data from 5 clinical cases.

Chapter 4

Intelligent Viewpoint Selection

Surface based registration (3D-3D) requires an intra-operative surface of the organ of interest. One approach is to capture with the laparoscopic camera as many surface patches as possible in order to get a reasonable reconstruction. In the SmartLiver system [26], a user manually selects surface patches to be non-overlapping and well distributed on the visible part of the liver, as in Figure 4.1. However, these selections are user dependent and there is no specific protocol to follow. Furthermore, this step is currently performed with the assistance of technical staff who have to be present at all the surgical interventions.

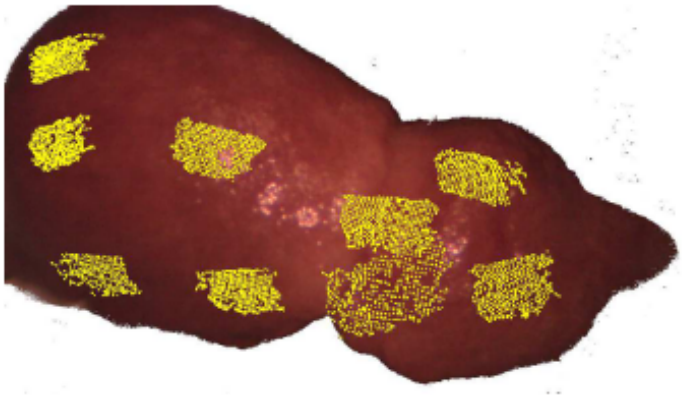

Figure 4.1: Example of data collection from the SmartLiver AR system. The liver patches are selected by a user to be generally well distributed over the liver surface. Figure from [26].

Another approach consists of an initial exploratory video in which the clinician moves the laparoscopic camera for several seconds around the organ of interest. Once all the sharpest images are selected, they can be used to create a dense

intra-operative surface for a Shape from Template formulation of registration [154]. However, not all the collected patches contribute equally to the final data fusion. As stated in Simon's work, increasing the amount of data points collected will typically improve the registration accuracy [101]. However, when the time is limited, the need for an automatic method to generate an efficient registration increases. A clear concise protocol would also reduce user dependent variability in registration performance.

Figure 4.2: Illustration of a view planning framework for 3D object reconstruction. The arrows represent cameras placed around the bunny, which indicate a sufficient set of best views for surface reconstruction. Figure from [34].

View planning was first introduced in the robotics research community, as a way to optimise the 3D reconstruction of a physical object by having large scanning sensors placed around it (see Figure 4.2). In general, view planning techniques output a short list of best views necessary for a reasonable 3D reconstruction, given the constraints of the imaging environment [34]. Techniques related to view planning have been proposed for different applications for an optimal workflow - scene exploration [155], object recognition [156], scene inspection [157]. A good survey of existing techniques in view planning in robotics can be found in [158].

In this chapter, concepts from view planning are introduced in a pre-operative

patient-specific simulation framework. It computes a sufficient set of surface patches, which can potentially lead to a more efficient initial registration. The main motivation behind this work is to explore ways to make the registration stage faster and more robust, while at the same time decreasing the user's interaction with the software. The remainder of this chapter is based on the work titled *"Intelligent viewpoint selection for efficient CT to video registration in laparoscopic liver surgery"* published in IJCARS 2017 [3], under the Creative Commons Attribution 4.0 International License (`http://creativecommons.org/licenses/by/4.0/`).

A comprehensive overview of the state of the art in pre-operative planning is introduced in Section 3.1. The work published by Simon [101] and Ma and Ellis [102] can be considered the most relevant for this chapter. They propose planning for optimal data collection strategies before the surgery, based on the intuition that specific surface areas represent better constraints for registration than random collections. Their work differs from the proposed pre-operative planning tool in several aspects. Both methods use a digitising probe (an optically tracked stylus) to acquire the data points during the surgery. Depending on the user experience, the points acquired with the probe will have a different precision. So, Simon's [101] approach removes from the computation any high curvature points, where the precision of point-picking would be lower. Whereas the chosen filtering technique shows good results for their proposed metric, a method relying exactly the high curvature areas is detailed in the current chapter. Secondly, due to their choice of data acquisition, their methods are heavily influenced if the access to the intra-operative scene is restricted and not enough points are visible, making them difficult to translate to a minimally invasive scenario. However, this work explicitly incorporates the constraints of the laparoscopic imaging environment - i.e. the clinically visible parts of the liver, the range of motion of the laparoscopic camera.

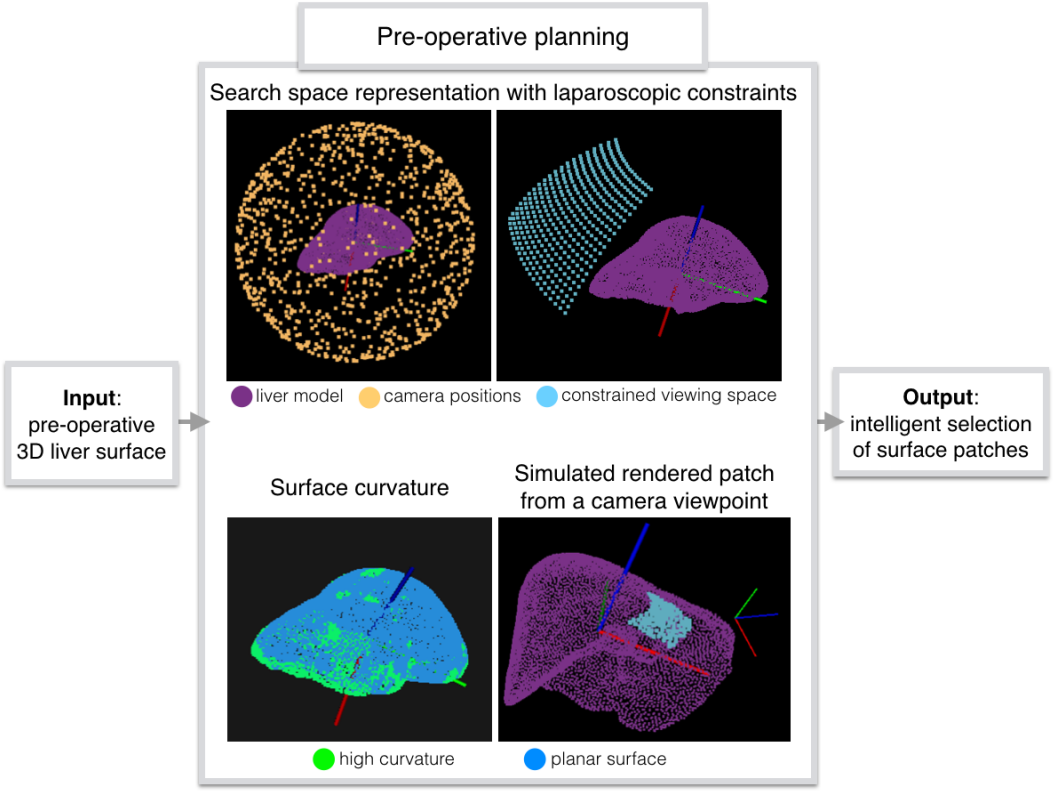

Figure 4.3: Overview of the proposed pre-operative planning simulation framework. Using the pre-operative 3D surface model, cameras are simulated around the liver, taking into account the laparoscopic constraints. Then, each camera is used to simulate a surface patch and compute the viewpoint score based on the surface curvature. Finally, a selection of the best surface patches is obtained which can constrain the subsequent alignment and make the data acquisition faster.

4.1 Contributions

I have developed a simulation framework in order to determine which viewpoints would represent the best constraints for registration. The 3D model of the liver extracted from the pre-operative CT is loaded into the framework. Fig. 4.3 highlights the main components. Given the trocar position, a set of camera positions are sampled and their corresponding views of the liver are simulated. I propose to automatically select these views based on a scoring strategy for efficient registration. This approach aims to remove the need for a time consuming exploratory video or for the reconstruction of a high number of patches. Therefore, it would make the data acquisition process during surgery faster, more intuitive and less disruptive for the surgical workflow. The constraints specific to camera movement and limited field

of view are explicitly included to simulate a realistic surgical scene for minimally invasive surgeries. In order to validate the feasibility of the proposed approach in laparoscopic surgery, the following questions are addressed:

- **How can the viewpoints be evaluated?** (Section 4.3.1) Synthetic data is used to validate if the proposed scoring strategy is correlated to higher registration accuracy. Furthermore, an intuitive visualisation is used to investigate which regions of the liver could potentially lead to better registrations.
- **Can view planning improve registration robustness?**(Section 4.3.2) Simulated intra-operative data is used to demonstrate that the proposed pre-operative planning step can lead to a correct registration convergence for larger initial misalignments.
- **How many patches are needed for a low registration error?** (Section 4.3.3) The effect the number of selected surface patches has on registration accuracy is explored.
- **How can view planning be applied in a real clinical scenario?** (Section 4.3.4) Clinical data from one human patient undergoing a laparoscopic liver resection is used to illustrate the use of the proposed approach in a surgical scenario. I discuss the additional challenges encountered with clinical data and possible future directions.

4.2 Methods

Let P_f denote the fixed point cloud of the liver surface extracted from a pre-operative CT scan. The moving point cloud P_m represents the surface reconstructed from laparoscopic video. Most approaches to registration minimise the distance between P_f and P_m transformed by T:

$$\min_T \sum ||P_f - T(P_m)||^2 \tag{4.1}$$

The focus is on choosing the optimal set of measurements for P_m in order to make the registration faster and more robust, with minimal amount of interaction.

The next sections detail the core components of the simulation framework. The search space of the laparoscopic camera is modelled to simulate the environment of laparoscopic surgery where the field of view is limited (Section 4.2.1). Given a camera position, the visible liver patch as well as the maximum visible area can be generated (Section 4.2.2). Once a liver patch is available, the corresponding viewpoint is evaluated based on a scoring strategy (Section 4.2.3).

4.2.1 Search space representation

The fixed point cloud P_f is placed at the scene origin and it is isotropically scaled to fit in a unit sphere. The search space of a camera looking at the object of interest represents all the possible locations from which it can capture images. A common approach in view planning is to use a virtual sphere surrounding the object on which the cameras can be placed. This approach provides an intuitive way to parametrise the camera location and orientation by working with the spherical coordinates: radius r, longitude θ and colatitude ϕ instead of 3 translations and 3 rotations. The variation in the radius r corresponds to the distance from the camera to the surface of the object. This assumption also restricts the camera orientation by aligning it with the center of the object. The variation in the spherical angles θ, ϕ corresponds to a different camera location on the enclosing sphere given by the radius r.

However, I want to incorporate the reduced surgical invasiveness specific to laparoscopic surgery. Hence, the space of possible camera locations is heavily constrained. The main camera position is considered to be where the trocar is situated in the patient's abdomen, pointing towards the center of the liver: (r_c, θ_c, ϕ_c). Depending on the surgical procedure, the clinicians can provide a range of available motion as $(r_c \pm \Delta_r, \theta_c \pm \Delta_\theta, \phi_c \pm \Delta_\phi)$.

4.2.2 Visible surface simulation

Given a camera position, the liver model is projected into the viewpoint's space. In order to simulate a realistic surface reconstruction, the points corresponding to the back surface of the 3D point cloud are not considered.

Multiple 3D patches can be generated for all the possible camera positions in

the constrained viewing space in order to obtain the maximum visible surface P_{max} by merging them. Due to the large number of points obtained, the point cloud is downsampled before any other computation is done.

4.2.3 Viewpoint scoring strategy

I evaluate each liver patch in the constrained viewing space in order to rank them based on their surface quality. Consequently, the principal curvatures k_1, k_2 are approximated at each 3D point [159]. The score $f_{surface}$ gives us an automatic way to distinguish between ambiguous and distinctive surfaces.

$$f_{surface} = \frac{\sum_i^N K_i}{N} \tag{4.2}$$

where N is the number of points in each rendered scan and $K_i = k_1^i k_2^i$ is the approximated Gaussian curvature at each point.

4.3 Experiments

Experiments were performed to validate the feasibility of view planning for efficient registration in the context of laparoscopic liver surgery. The liver point cloud P_f is extracted from a segmented CT scan of a human patient taken before the surgery[1]. The point clouds P_m used in the following experiments simulate what a camera would see intra-operatively, in a rigid scenario. Anisotropic noise ($N(0, 1mm)$) was added along the normals to P_m in order to simulate the characteristic surface reconstruction noise. All the experiments use the Iterative Closest Point (ICP) [75] algorithm for the registration step. Fig. 4.4 shows a liver abnormality which is simulated inside the model and is used as a landmark for measuring the target registration error (TRE). In the case of a perfect alignment, the registration errors should be 0. Note that the points corresponding to the target landmark are not used during the registration process with ICP.

All the experiments use a fixed radius for the enclosing sphere, thus further reducing the camera parametrisation from 6D to 2D. Consequently, the camera locations are specified using the spherical angles θ, ϕ and the camera is always oriented

[1] www.visiblepatient.com

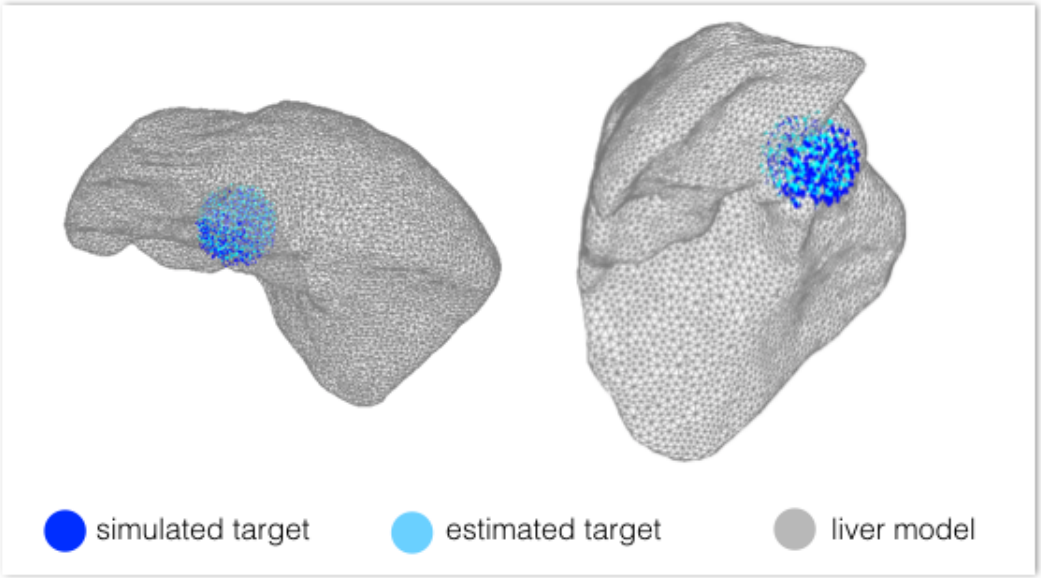

Figure 4.4: Illustration of the simulated abnormality in the liver model. The registration error is calculated between the target and its estimated position given an alignment.

towards the center of the object. Furthermore, starting with Section 4.3.2, a range of $[-30°, 30°]$ is used for the spherical coordinates in order to simulate the constrained search space (as described in 4.2.1).

The simulation framework was implemented in C++ on a MacOS 10.11.2 laptop with an Intel Core i7 3.1 GHz processor. The current running time is of several minutes, well within the requirements for a pre-operative planning stage. The project uses CMake (3.4.2) to enable compilation on multiple platforms. The libraries used as dependencies are: Eigen 3 [160] and the Point Cloud Library 1.2 (PCL) [161].

4.3.1 Viewpoint evaluation

This experiment explores the distribution of registration errors for individual patches of the liver model. Since I am investigating which liver regions could lead to a good registration irrespective of the clinical case, no constraints on the viewing space are used for now. 170 camera positions were randomly picked on a sphere enclosing the liver (with a constant radius). Given the camera location, a liver patch is rendered to simulate what the camera sees. Then, 100 transformations are generated

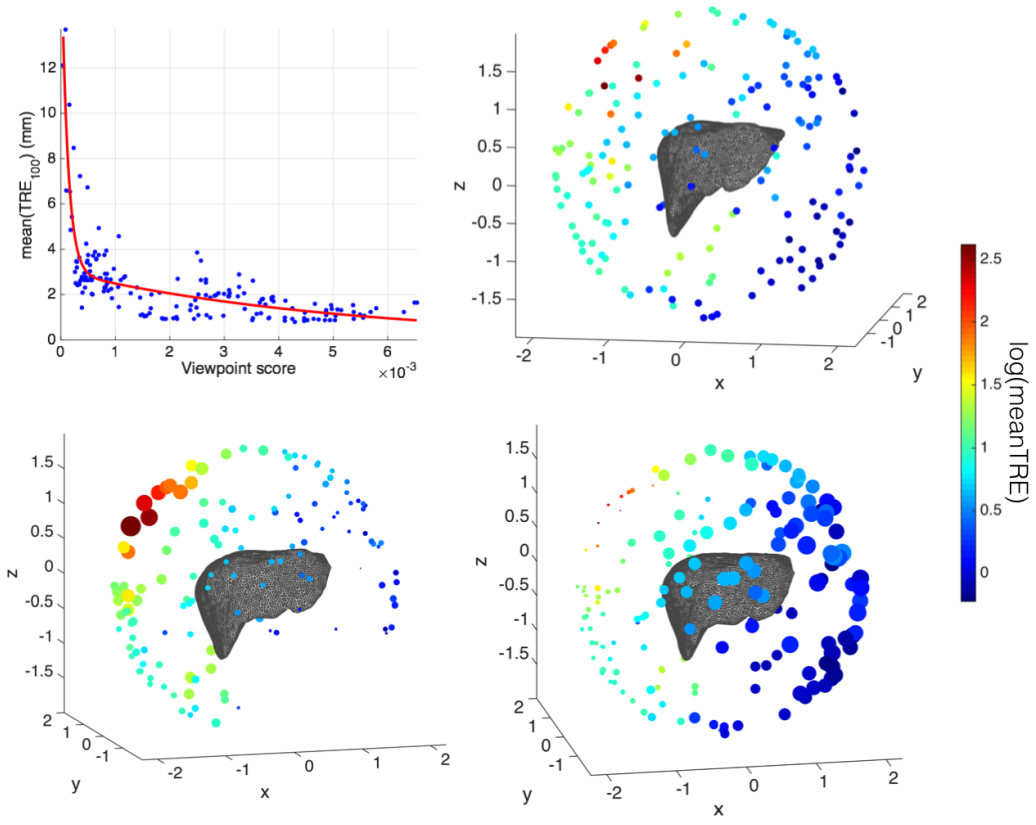

Figure 4.5: Top-Left) Correlation between the Gaussian curvature scoring strategy $f_{surface}$ and the registration errors. Top-Right) TRE variation over the liver surface. Each sphere represents a camera location. The colour map is applied to $log(mean(TRE))$ over the 100 registrations for each sphere ($mean(TRE) \in [0.7992mm, 13.7048mm]$). Bottom-Left) The size of the spheres is proportional to $std(TRE)$. So, a bigger variation in errors leads to larger spheres. Bottom-Right) The size for the spheres is proportional to the curvature score $f_{surface}$.

by randomly choosing rotation and translation parameters $(rot_x, rot_y, rot_z, t_x, t_y, t_z)$ in the ranges $rot \in [-10°, 10°]$ and $t \in [-10mm, 10mm]$. These transformations are applied to each simulated liver patch and the target registration errors are recorded after ICP is used to recover their initial position. Fig. 4.5 shows the distribution and variation of registration errors on the liver model.

4.3.2 View planning for registration

This experiment analyses the basins of convergence of ICP between the original liver model P_f and three different partial point clouds P_m consisting of the patch with the lowest score P_{low}, the patch with the highest score P_{high} and the maximum visible

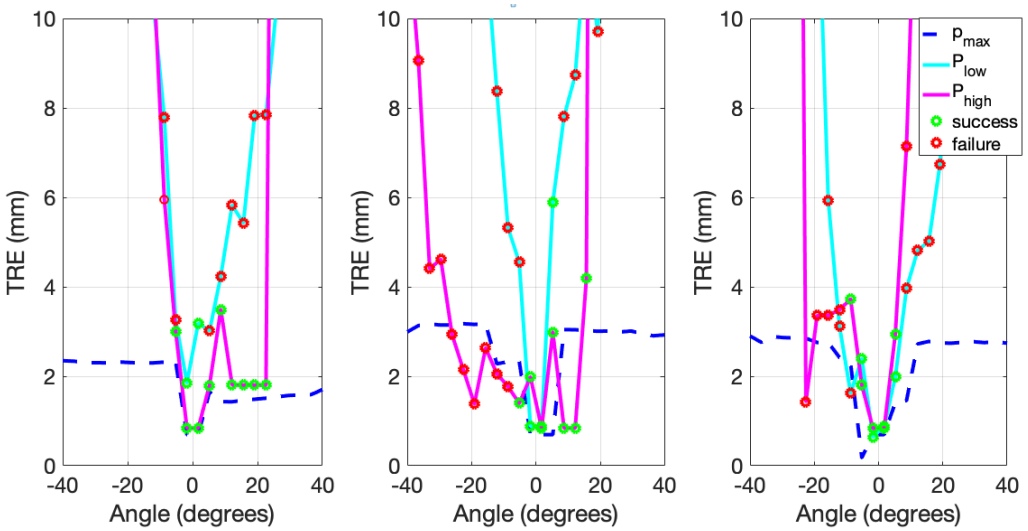

Figure 4.6: Basins of convergence for ICP for the lowest scoring point cloud P_{low}, the highest scoring point cloud P_{high} and the maximum visible surface P_{max} for three scenarios: left) rotation around the X axis; middle) rotation around the Y axis; right) rotation around the z axis. Note that surface patch scoring is computed using $f_{surface}$. The green circles were omitted for clarity for the baseline as it always converges to the desired position.

surface P_{max} respectively. The last point cloud is chosen as a best case scenario when the complete liver surface visible from a trocar position is reconstructed.

The basins of convergence are generated for rotations in the range $[-40°, 40°]$ over 24 steps for each x,y,z axis. Since the visual assessment of the graphs (Fig. 4.6) makes it difficult to draw any clear conclusions about the different scenarios, an interactive registration visualisation application was implemented. A registration was marked as a failure if the moving model was either sliding on the surface of the fixed model or getting locked on a separate part of the liver. These failure modes reflect the main issues with ICP - namely, a poor initial alignment can lead to wrong correspondences and subsequently to the optimisation getting stuck in a local minimum. Figure 4.7 shows examples of successful and failed alignments. After the visual inspection of each case, the qualitative assessment is overlaid on top of the graphs in Fig. 4.6 with red and green circles.

It can be easily observed in Fig. 4.6 that the highest scoring view P_{high} recovers the correct transformation for larger misalignments (with more successful cases). The rotation around the x axis shows successful registrations for P_{low} be-

tween $[-2°, 2°]$, whereas for P_{high} the range goes to $[-6°, 23°]$. The y rotation leads to a convergence in the global maximum for a range of $[-2°, 6°]$ for P_{low} and $[-6°, 16°]$ for P_{high}. Similarly, for the z axis, the ranges are of $[-6°, 6°]$ degrees for P_{low} and $[-9°, 6°]$ degrees for P_{high}. Interestingly, P_{high} has an asymmetrical convergence basin with respect to 0, which could be due to the irregular size of the patch. Furthermore, the plateau observed around 2.5 mm is a characteristic of the registration algorithm chosen (ICP). The planning method could be added as a separate component to a registration pipeline, irrespective of the registration algorithm.

4.3.3 Effect of the number of patches on registration accuracy

This experiment explores the variation of the TRE with an increasing number of patches selected using each of the following strategies: random, top lowest scores and top highest scores. These are all compared to the maximum visible liver surface which is considered the baseline.

Similarly to the previous experiment, 100 random transformations are generated and applied to the point cloud P_m for each selection strategy. Once multiple patches are selected, they are merged into a point cloud which is subsequently used for registration.

Fig. 4.8 shows that the two highest scoring views perform on average as well as the four lowest scoring views. Moreover, their errors are comparable to the ones obtained from the baseline with the registration of the maximum visible surface.

4.3.4 Application in laparoscopic liver resection

In this experiment, the proposed method is validated on real data from one stereo video sequence from a liver resection (Fig. 4.10), which allows for a more realistic representation of intra-operative data. The liver is automatically segmented with the deep learning framework proposed in [11]. 113 surface patches are reconstructed using [69] and merged together to build the maximal surface seen by the laparoscopic camera during surgery. So, two scenarios are compared for registration - the most distinctive patches suggested by the proposed method and a random subset.

Firstly, the proposed view planning pipeline is used to recommend the highest

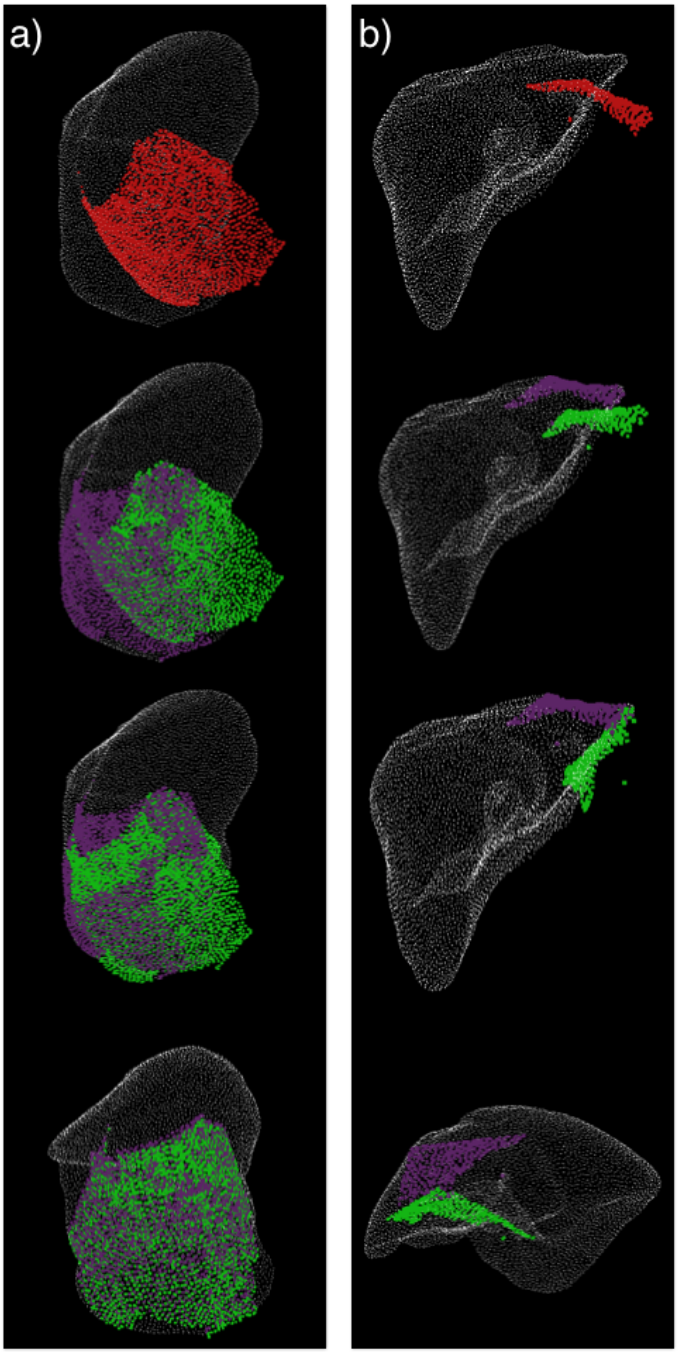

Figure 4.7: Visual assessment of ICP progress until convergence. Colour coding: fixed point cloud in white, moving point cloud after the rotation is applied in red, moving point cloud during ICP iterations in green, ideal final position of the moving point cloud in purple. Column a) Success example for a rotation of 40° around the z axis. The images from top to bottom show the iterations 0 (before ICP), 1, 12, 150. Column b) Failure example for a rotation of 18° around the y axis. The images from top to bottom show the iterations 0 (before ICP), 1, 20, 25.

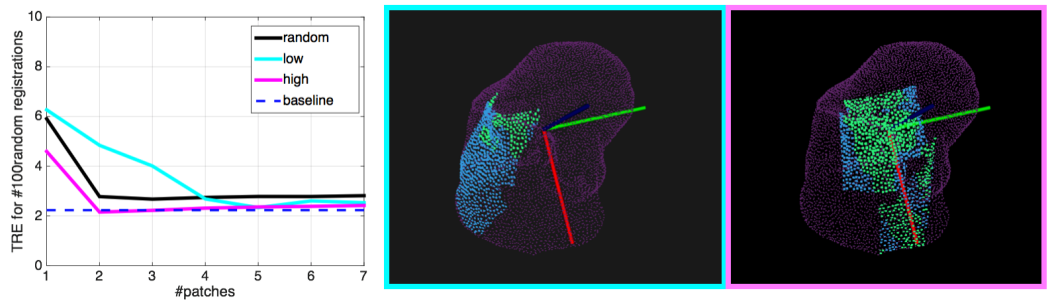

Figure 4.8: Left) Mean TRE variation (for 100 random transformations) with an increasing number of patches. Middle) Example of top 3 lowest scoring patches chosen by the algorithm. Right) Example of top 3 highest scoring patches. Blue regions highlight flat surfaces, whereas green symbolises positive Gaussian curvature.

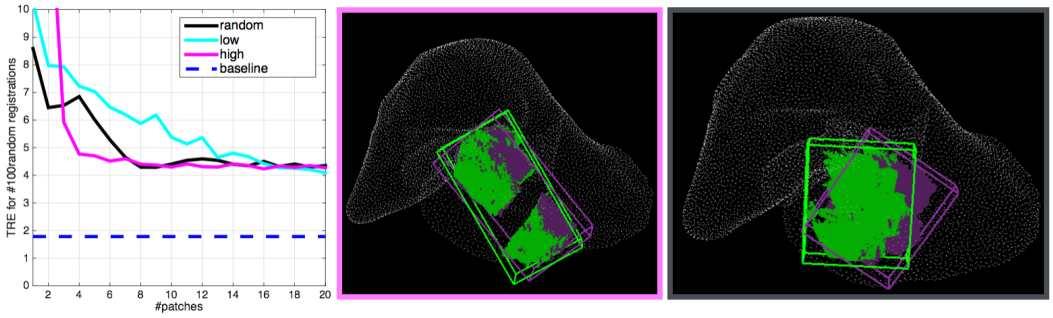

Figure 4.9: Visual assessment of registration with view planning for real data. Colour coding: fixed point cloud in white, moving point cloud in green, ideal final position given by a gold standard registration of the moving point cloud in purple. Left) The previous experiment was re-ran for the constraints given by real data from a liver resection in order to suggest a number of patches to be considered. Middle) Registration results using clinical data after running ICP on a merged point cloud from the top 4 recommended patches by the proposed method. Right) Registration results using clinical data after running ICP on a merged point cloud from 4 random patches.

scoring camera positions based on the CT liver model of the patient. Given that ICP is used as the registration algorithm, a reasonably close initial alignment between the liver model and the camera space is needed. A common approach is employed [26, 89] in which this alignment is estimated by the user through the manual rotation and translation of the 3D model until it approximately matches the intraoperative surface reconstruction.

The constrained viewing sphere was built to match the maximal surface derived from the real surface reconstruction. A fixed radius of 30 mm is used, which

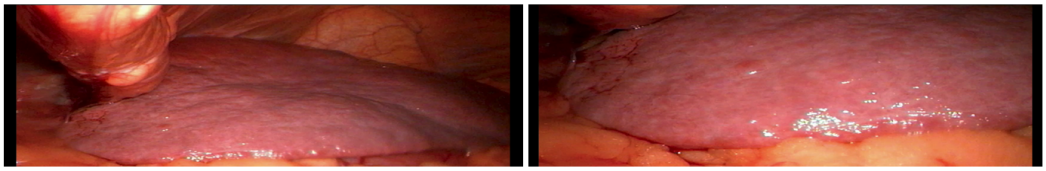

Figure 4.10: Example of video frames used for the stereo reconstruction of the surface in a laparoscopic liver resection.

represents the average distance of the laparoscope from the liver during the video. The previous experiment is used to decide on how many patches to select (see Fig. 4.9 left). The mean of the TRE for the 100 random perturbations applied to P_{high} goes down from 17.5mm for one patch to 4.7 mm at 4 patches and seems to converge afterwards. Consequently, the first 4 suggested liver areas are used as the view planning output and they are merged to form P_m.

Secondly, the proposed computation of the constrained viewing sphere is used to generate the random patches. This step is necessary to ensure that the two scenarios compared have a similar covered surface area.

Once these simulated patches are recommended, I manually selected the real surface patches that match them the most in position, given an initial manual alignment between the CT and the reconstructed liver surface.

Fig. 4.9 shows the visual assessment of the two considered scenarios - P_m suggested by view planning and a randomly selected P_m. The random selection slides away from the correct position, whereas the views suggested by view planning appear to lock the registration close to the desired output.

4.4 Discussion

Fig. 4.5 motivates the use of the curvature score for view planning in registration. With higher scores in the more distinctive views, the registration error decreases rapidly. The scaling associated with the curvature score highlights the camera positions with low registration errors, at the same time reducing significantly the views with higher and more spread apart TRE values. Fig. 4.5 allows an intuitive understanding of which liver regions lead to accurate registrations, as well as how to automatically choose them based on the curvature score. The graphs suggest the

right lobe of the liver as being the most ambiguous for this camera space parametrisation. Similarly, the boundaries and the bottom side of the liver lead to the lowest errors.

Fig. 4.6 points towards the feasibility of using view planning in the registration of intra-operative and pre-operative data. The explicit integration of high scoring patches in a registration algorithm could lead to a faster and more accurate registration for larger rotational offsets. Furthermore, these results are promising because they also show that for small enough rotations, similar registration errors could be achieved by both a high scoring patch and the maximum visible surface. This finding would remove the need for picking multiple patches to cover the whole area and change the focus to selecting the most distinctive patch. Such an approach could potentially automate the selection of useful patches, leading to less user interaction.

Intuitively, a partial point cloud that covers a larger portion of the surface should provide a better registration than a smaller region. However, the third experiment (Fig. 4.8) shows that if the small view is distinctive enough, the target registration errors are comparable. Hence, the time spent collecting data could be significantly decreased by defining a laparoscopic camera guidance system.

Lastly, the proposed view planning method is compared with a random selection of patches on clinical data from a surface reconstruction in liver resection (Fig. 4.9). This experiment highlights the multiple challenges encountered in a real world scenario: the deformation of the liver due to pneumoperitoneum, the flatness of the intra-operative surface and the noisy surface reconstruction. Due to these difficulties, I believe the proposed views in the last experiment perform better also because of the surface covered. Fig. 4.11 highlights two scenarios with just two patches selected based on curvature and surface covered or based only on curvature. After 200 iterations, the selection based only on curvature starts to slide from the desired position, whereas the first selection performs better. Further work could investigate how to incorporate additional constraints to arrive at an optimal selection suited to a laparoscopic scenario.

However, the proposed view planning approach has several limitations. The

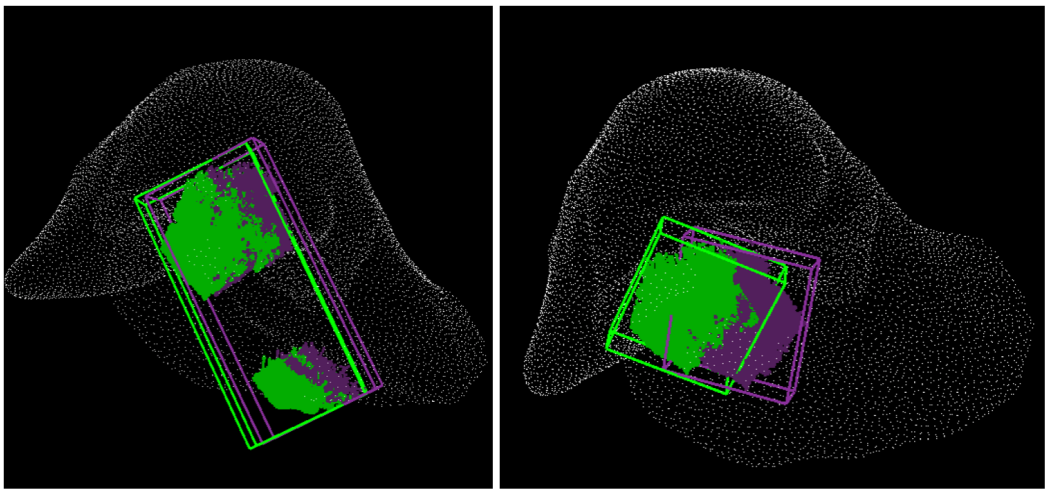

Figure 4.11: Registration examples highlighting the need for additional constraints in a real world scenario. Colour coding: fixed point cloud in white, moving point cloud in green, ideal final position given by a gold standard registration of the moving point cloud in purple. Left) The moving cloud is built from 2 patches - the highest scoring viewpoint and the patch which is furthest away from it. Right) The moving point cloud is built from the top two highest scoring patches.

current implementation makes strong assumptions about the surgical scene (i.e. completely rigid, no pneumoperitoneum deformation) and the camera model (i.e. fixed field of view, fixed focal length, always looks at the center of the liver). More experiments are needed to explore the influence of the parameters (i.e. downsampling) on the computational complexity and robustness of the method. Nevertheless, a simplified yet relevant framework was needed to explore the feasibility of this approach in the context of abdominal image guidance. Given these promising results, more research will be done in this direction for the development of a comprehensive view planning technique for optimal registration.

Finally, one of the strengths of the proposed view planning simulation framework is that it can be easily made specific for surgical procedures. The clinician would have to input the trocar position and the maximum viewing space of the laparoscope and the method would automatically suggest good views for efficient registration. An example of such an approach has been illustrated on real data from a liver resection. Such an intuitive protocol can be easily achieved by clinicians, avoiding the presence of any additional technical personnel during the surgical pro-

cedure. Furthermore, the reduction in the time spent acquiring data could potentially lead to surface reconstruction patches which are not as much influenced by heartbeat or breathing motion deformation.

I proposed the use of view planning for a more efficient registration between a liver model derived from a pre-operative CT scan and laparoscopic surface data. I have shown its feasibility in a simulated environment of laparoscopic liver surgery and demonstrated it can be applied on clinical data. Intelligent viewpoint selection can provide an essential pre-operative planning tool for laparoscopic liver surgery. With further work, reduced operator dependence and registration time should be possible.

While the proposed planning framework could be used with multiple fine-alignment registration algorithms, it still requires a global alignment between the two models. Currently, this step is performed manually. In the next chapter, an alternative semi-automatic method to estimate a global alignment is presented.

Chapter 5

3D-3D Semi-automatic Global Registration

Once an intra-operative surface is available, a global registration is needed. A comprehensive overview of the state of the art in global registration can be found in Section 3.2. A commonly used strategy is the manual alignment of the two surfaces (i.e. by specifying the rotations and translations), which can be difficult to perform quickly and intuitively, as it depends highly on the user expertise. Automatic methods have been proposed to obtain a global 3D-3D alignment [84, 35], with the work proposed by Dos Santos et al. [35] being the most relevant to this chapter. Their 3D-3D global alignment takes into account the geometric fit between two surfaces, the compatibility of the feature descriptors and the reliability of the spatial distribution of the estimated correspondences set. However, both techniques introduce additional hardware such as a tracked laser pointer and a TOF camera. Moreover, they also require large areas of the liver to be visible and it is unclear how they would translate to laparoscopic interventions, with approximately 30%-50% visibility. The remainder of this chapter is based on the work titled *"Global Rigid Registration of CT to Video in Laparoscopic Liver Surgery"* published in IJCARS 2018 [2], under the Creative Commons Attribution 4.0 International License (`http://creativecommons.org/licenses/by/4.0/`).

5.1 Contributions

I propose a semi-automatic method for the initial, global, rigid alignment between a 3D mesh derived from a pre-operative CT of the liver and a surface reconstruction of the intra-operative scene in laparoscopic interventions. The proposed method is able to robustly estimate a correspondence set between the two shapes under deformation, sparse data, partial views and realistic noise levels.

Shape matching is formulated as a quadratic assignment problem which minimises the dissimilarity between feature descriptors while enforcing geometrical consistency between all the feature points. An existing optimisation strategy is extended to incorporate an additional constraint based on the contours of the organ (the ridge line - see Fig. 5.1), which can be identified in both modalities with high confidence. Once the delineation of the liver ridge line is given in the two modalities, no further user interaction is required for the alignment stage. The proposed method outputs a global alignment which can be further refined by fine registration algorithms, such as those in Section 3.3.

The proposed method is validated on synthetic data, on a liver phantom and on retrospective clinical data acquired during a laparoscopic liver resection. Robustness over reduced partial size and increasing levels of deformation is demonstrated. Moreover, I provide quantitative results obtained on a liver phantom and qualitative results on retrospective data from a laparoscopic liver resection to illustrate feasibility in a realistic clinical setting. Good initial alignment is shown, which can successfully converge to the correct position using fine alignment techniques. Furthermore, since the CT scan can be pre-processed before surgery, the proposed method runs faster than current algorithms.

5.2 Methods

Fig.5.1 illustrates the main steps of the proposed pipeline. The input data includes the segmentation of the 3D mesh from the CT scan, a surface reconstruction represented as a point cloud of the intra-operative data and the segmented contours on both surfaces. The liver contour is defined as the ridge line visible in yellow in Fig.

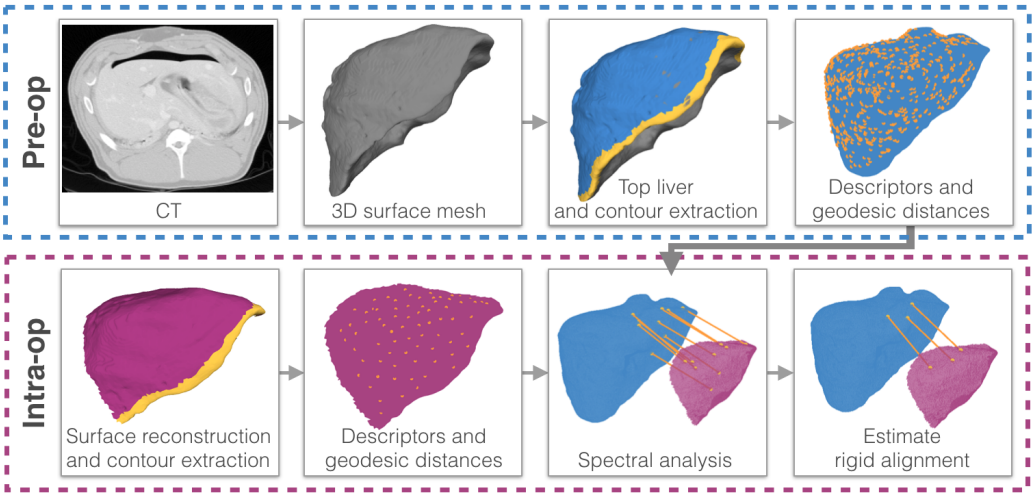

Figure 5.1: An overview of the proposed global alignment framework, showing the pre-operative and intra-operative steps.

5.1 on both the pre-operative and the intra-operative surfaces.

Let M be the moving (pre-operative CT mesh) model and T be the target (intra-operative point cloud) model. Sets of features, $\{m_r\} \subset M$ and $\{t_s\} \subset T$, are selected on both surfaces with $f(\cdot)$ as their corresponding descriptor. Let $d_g(x,y)$ be the geodesic distance between any two points, x and y, on a surface.

Generally, it is difficult to match surfaces in laparoscopic liver surgery only based on descriptors since the surfaces lack prominent, uniquely identifiable features. The use of geometric consistency between the correspondences on both shapes can further constrain the registration problem.

Shape matching can be formulated as a quadratic assignment problem (QAP):

$$E(C) = \sum_{(m,t)_i \in C} d(f(m_i), f(t_i)) + \sum_{(m,t)_i \in C} \sum_{(m,t)_j \in C} (d_g(m_i,m_j) - d_g(t_i,t_j))^2 \qquad (5.1)$$

where $C = \{(m,t)_i\} \subset M \bigtimes T$ is the initial correspondence set composed of candidate pairs of feature points from the two surfaces, $d(f(\cdot), f(\cdot))$, is the distance between the feature descriptors and $d_g(\cdot,\cdot)$ is the geodesic distance between two correspondences on the same surface. This energy function aims to output a set of correspondences for which the dissimilarity between the descriptors is minimised and the geodesic distances between pairs of correspondences are maintained.

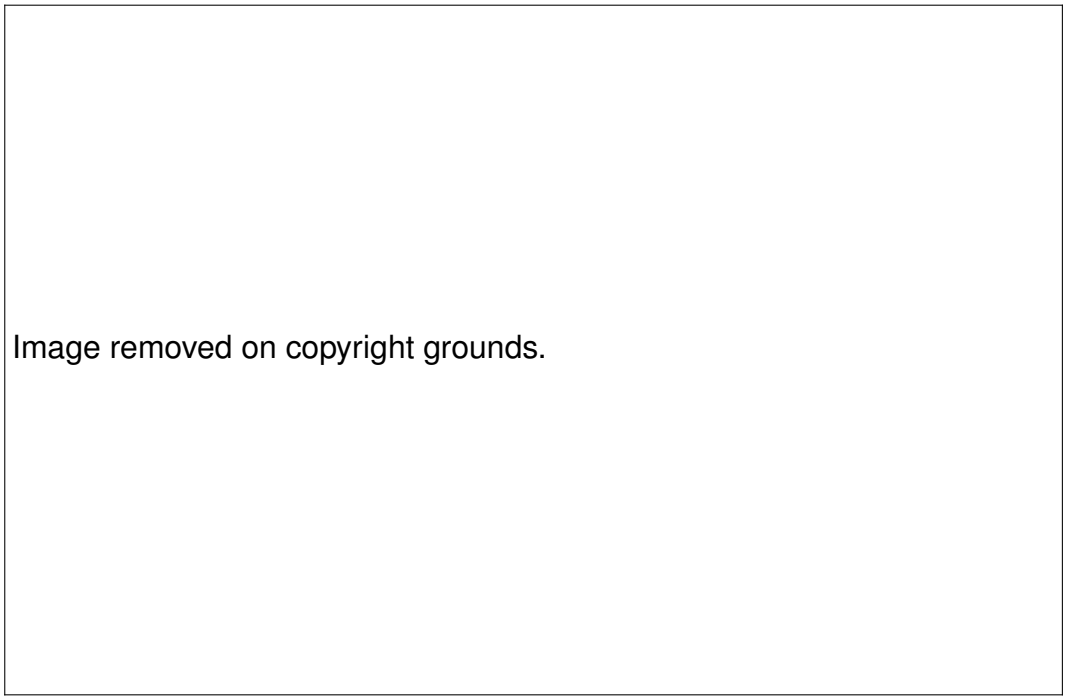

Figure 5.2: Ambiguity in correspondence estimation. The point configuration on the left should be matched to the corresponding red point set on the right. However, on a nearly planar surface, multiple configurations with equal spatial distributions can be found. Figure from [35].

While this approach works well in the vision literature for complex shapes [162], the intra-operative surfaces pose several challenges. It has been previously discussed in [35] that constraining the correspondence set based only on the geodesic distances between them is still ambiguous for almost flat surfaces, in which the same spatial configuration of features can be identified in multiple locations (Figure 5.2). The same behaviour was observed with our data. So, an additional constraint based on the liver contour is proposed, which can be reliably observed on both models. The existing spectral matching framework is extended to incorporate the new term and robustly estimate a set of correspondences.

5.2.1 Optimisation

In order to minimise $E(C)$ from Eq. 5.1, the shape alignment problem is formulated as graph matching [163]. Each node consists of a candidate correspondence (i.e. $(m,t)_i$) and each edge connects two nodes (i.e. $(m,t)_i$ and $(m,t)_j$). Moreover, if pair $(m,t)_i$ corresponds to $(m,t)_j$, the pairwise constraints imposed will quantify

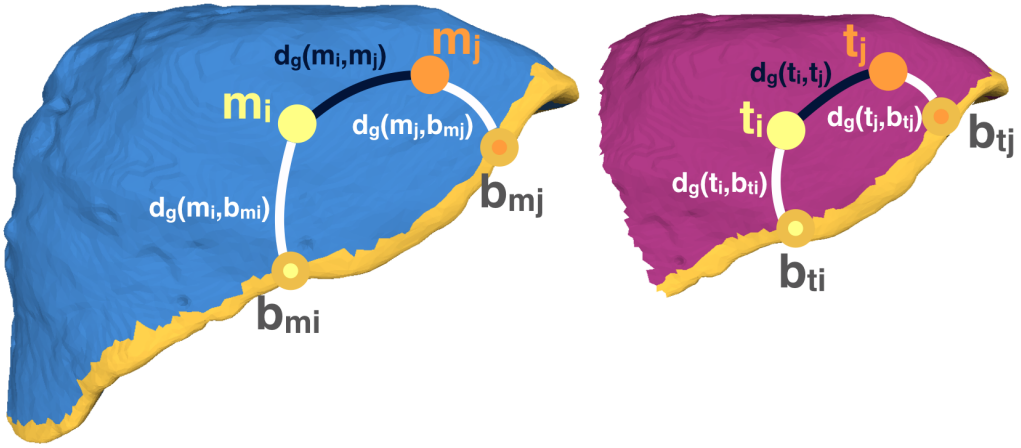

Figure 5.3: Pairwise constraints on the moving, M, (blue) and target, T, (pink) models used for pruning the correspondence set.

how consistent this association is from a geometrical point of view, thus providing weights for the edges. Fig. 5.3 highlights an example of a correct assignment.

An affinity matrix, W, of the graph is built. The weights associated with each node and edge will result in a strongly connected cluster for data with high consistency. On the other hand, outlier nodes will be either weakly linked or linked in an unstructured way. The initial correspondence set C is built by choosing for each target feature point $\{t_s\} \subset T$, the closest k neighbouring descriptors on the moving surface $\{m_r\} \subset M$, which can be performed quickly using *kd trees*.

A spectral analysis method [163] is used to obtain the filtered correspondence set C_p from the matrix W. Leordeanu and Herbert [163] show that the eigenvector associated to the largest eigenvalue of matrix W maximises the inter-cluster score. Consequently, each value in the eigenvector can be seen as the association confidence of each pair of correspondences to the strongest connected cluster. A greedy algorithm [163] is used to extract from the initial correspondence set C the ones with highest confidence C_p. Their approach is generally robust to noise and outliers and is computationally efficient. However, in cases with a high number of outliers, large deformation or symmetry in the data, some wrong correspondences might be included in the main cluster. Additional measures are taken to account for deformation in the data and outliers in the filtered correspondence set C_p.

In the next few paragraphs, the proposed formulation for W is detailed, which

incorporates the additional term based on the liver contours.

The affinity matrix, W, should have values which are non-negative, symmetric and increasing with higher similarity between the correspondences [163]. So, instead of working with distances as in Eq. 5.1, the pairwise terms are parametrised as consistency measures:

$$c(m_i,m_j,t_i,t_j) = min\left[\frac{d_g(m_i,m_j)}{d_g(t_i,t_j)+\varepsilon}, \frac{d_g(t_i,t_j)}{d_g(m_i,m_j)+\varepsilon}\right] \tag{5.2}$$

where ε is a small number to ensure the denominators are not zero, $c \in [0,1]$ and it quantifies how similar the geodesic distances are between the pairs (m_i,m_j) and (t_i,t_j). A pair of correspondences is consistent from a geometric point of view if the ratio of the geodesic distances on each shape is close to 1 [164]. However, in the presence of non isometric deformation of the data, correct correspondences might have consistency values, c, lower than 1. So, the non-rigidity of the data is taken into account by using the following function for c:

$$g(m_i,m_j,t_i,t_j,\sigma) = exp\Big(-\frac{(c(m_i,m_j,t_i,t_j)-1)^2}{2\sigma^2}\Big) \tag{5.3}$$

where the parameter σ sets the amount of non-rigidity allowed for the correspondence set. Furthermore, the function g also helps in separating the outliers by lowering the weights of highly unlikely candidate pairs.

Let $B^M \subset M$ and $B^T \subset T$ be the contour points on each surface. The closest contour point to each feature point x on either M or T is computed as $b_x = min(d_g(x,B))$. The expression used for the proposed contour constraint is:

$$g_b(m_i,m_j,t_i,t_j,\sigma_b) = \frac{1}{2}\Big(g(m_i,b_{mi},t_i,b_{ti},\sigma_b) + g(m_j,b_{mj},t_j,b_{tj},\sigma_b)\Big) \tag{5.4}$$

where σ_b allows some deformation between the candidate pairs and their corresponding contours. In practice, $\sigma_b = \sigma_d$ since they both represent variations in geodesic distances illustrating how restrictive the geometric pairwise constraints will be on the data.

Finally, the affinity matrix, W, is built by placing the unary terms (similarity between descriptors $sim(f(m_i), f(t_i))$) on the main diagonal and the pairwise constraints on the off-diagonal:

$$W(i,j) = \begin{cases} sim(f(m_i), f(t_i)), i = j \\ \alpha g(m_i, m_j, t_i, t_j, \sigma_d) + (1-\alpha) g_b(m_i, m_j, t_i, t_j, \sigma_b), i \neq j \end{cases}$$

where α allows different weights for the importance of the two pairwise constraints.

Spectral analysis on the initial formulation for $E(C)$ (Eq. 5.1) enforces high similarity between nodes $(m,t)_i$ and $(m,t)_j$, as well as approximately equal distances between them ($d_g(m_i, m_j)$ and $d_g(t_i, t_j)$). In addition, the proposed term weighs the edge connecting $(m,t)_i$ and $(m,t)_j$ higher if the distances $d_g(m_i, b_{mi}), d_g(m_j, b_{mj})$ are similar to $d_g(t_i, b_{ti}), d_g(t_j, b_{tj})$ respectively. As a result, the estimated correspondence set, C_p is explicitly constrained to be consistent with the liver contours on both surfaces, M and T.

5.2.2 Features and Descriptors

Reliable landmarks are difficult to identify consistently between the two surfaces. The strategies used are farthest point sampling [165] and normal space sampling [106]. The former approach was chosen for a uniform distribution of the feature points on the surface. The latter aims to select samples such that the normals are distributed as evenly as possible, thus having fewer points in flat regions. The search space is constrained to only select features on the visible surface of the moving mesh, M, in order to eliminate unfeasible solutions (see Fig. 5.1). TOLDI was chosen as a feature descriptor, because it was shown to be robust to data sampled irregularly (which is the case for multiple stereo reconstruction surfaces merged together), robust to varying levels of noise and invariant to rigid transformations [166].

5.2.3 Distances

The geodesic distance represents the shortest distance on the surface between two points. If the surface changes topology through holes or irregularities in the data,

the geodesic distances might become unreliable. Another failure case would be observed if distant parts of the object come into contact and create new shortest paths between feature points. However, it is unlikely the liver shape will change topology in the initial stages of the surgery.

The intra-operative data collected during laparoscopic surgery will most likely have some degree of sparsity - sparse point clouds [83], sparse data collection [66], sparse stereo reconstructed patches [26]. At the same time, intra-operative surfaces are smooth.

So, let S be a smooth interpolated surface of the intra-operative point cloud, T. The target feature points, $\{t_s\}$, and contour points, B^T, can be expressed on the interpolated surface with nearest neighbour computation. The geodesic distances on S are computed using the fast marching algorithm [167, 168]. This step is the most computationally expensive to compute in my implementation. However, faster alternatives can be employed [169].

5.2.4 Estimating the rigid transform

The proposed shape matching technique starts with a large set of correspondences, C, and spectral analysis prunes out the outliers, resulting in C_p. The final set of correspondences is not guaranteed to consist only of correct matches, especially in cases with significant deformation.

Random Sampling and Consensus (RANSAC) [74] (see Fig. 5.1) is used to get the best minimal solution $\{(m,t)_i\} \in C_p$ out of the pruned set of correspondences C_p. The final pairs $\{(m,t)_i\}$ are used for the least squares estimation of rotation and translation. The estimated transformation is considered to be a good fit if the root mean squared error (RMSE) between the target and moving models is less than a threshold d_{RANSAC} and the difference between the normals is less than a specific angle threshold $a_{normals} : dot(n_m, n_t) < cos(a_{normals}) \forall (n_m, n_t)_i \in C_p$.

5.3 Experiments

Three sets of experiments were conducted to validate the proposed method. Firstly, the robustness to the specific challenges present in laparoscopic interventions (par-

tial views, varying degrees of deformation) was tested on synthetic data. Secondly, the proposed initial alignment method was quantitatively validated in a liver phantom experiment. Finally, qualitative results are shown on retrospective clinical data from a dataset acquired during a human liver resection.

The same parameters are used for all the experiments both on synthetic and on clinical data ($\sigma_d = \sigma_b = 0.3$, $\alpha = 0.6$). The maximum number of iterations used for RANSAC is 1000. The difference between the normals is set as $a_{normals} = 60^\circ$ in order to account for some of the deformation. Similarly, $d_{RANSAC} = 5mm$ in the rigid case scenario (see Sec. 5.3.1.1) and $d_{RANSAC} = 10mm$ for all the rest of the experiments. If no solution is found with the RMSE lower than d_{RANSAC}, the transformation which resulted in the smallest error over all the iterations is used. The feature points, descriptors and geodesic distances on the CT mesh are precomputed and stored, in order to minimise the computation time during surgery. The liver contours are currently manually delineated in a matter of seconds and techniques to automate this step will be investigated in the future.

The proposed method was implemented in Matlab and C++, on a MacOS 10.11.2 laptop with an Intel Core i7 3.1 GHz processor. The proposed initial alignment between surfaces can be obtained in approximately 20 seconds. The libraries used as dependencies can be found in [168, 165, 170, 166]. The mesh processing applications MeshMixer [1] and Meshlab [171] were used for visualisation and simulation purposes.

5.3.1 Synthetic data

The robustness of the proposed method to partial views of the liver and to increasing deformation levels is validated. The mesh of a liver phantom (OpenCAS [30] - see Fig. 3.9) is used as the moving model, M. The mean distance between the estimated registration result and ground truth vertex correspondences is measured. Three algorithms are compared:

- (R) RANSAC applied directly to the initial set of correspondences, C. No pruning is applied.

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

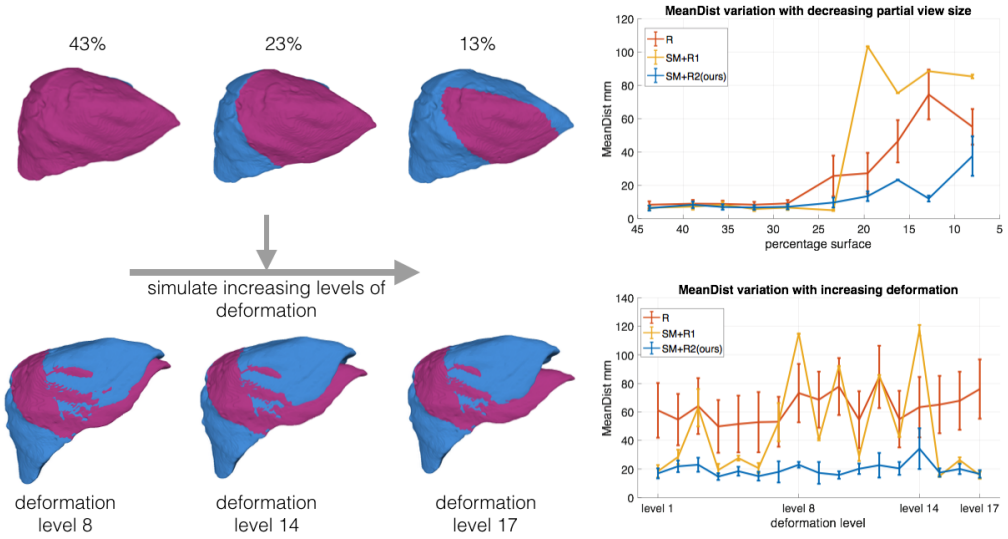

Figure 5.4: Experiments on synthetic data. Top: Robustness to reduced partial size in the target model, T, Bottom: Robustness to increasing levels of deformation in T. The target model representing 23% of the total liver surface is used in the bottom experiment with increasing deformation levels. *Colour coding:* moving model, M - blue, target model, T - pink.

- (SM+R1) The initial set, C, is pruned based on geodesic distances alone, following the spectral analysis technique detailed above. RANSAC is applied on the pruned set, C_p.

- (SM+R2 - ours) The proposed technique with both pairwise constraints.

5.3.1.1 Robustness to reduced partial size

The robustness of the proposed method to reduced partial views of the liver is tested. For this experiment, there is a rigid transformation between the moving (M) and target (T) models, with all the remaining parameters fixed.

In this experiment, the target model, T, is simulated by creating 10 partial views of decreasing sizes (from 43% to 7%) by cropping the original liver mesh M. This step was manually performed in Meshlab. The reduced partial sizes for each mesh in the dataset are computed as a percentage of the top liver surface. Each algorithm is run 500 times for each size. The mean and standard deviation of the resulting errors are reported in Fig. 5.4 - top.

5.3.1.2 Robustness to deformation

The robustness of the proposed algorithm to increasing levels of deformation in the data is tested. A large deformation is applied with control points on the left lobe of the liver mesh M. Intermediate levels of deformation are generated with vertex linear morphing between the original liver shape and the deformed one. These steps were achieved in MeshMixer and Meshlab. Fig. 5.4-bottom) shows examples of the different deformation levels with 17 being the highest.

The patch with the partial view of 23% is chosen (Fig. 5.4-top, middle shape) since this can be registered well by all algorithms tested. Visually, it represents a realistic size for a laparoscopic view.

In this experiment, the deformation level is the only variable. Each algorithm is run 500 times for each deformation level. The mean and standard deviation are presented in Fig. 5.4 - bottom.

5.3.2 Liver phantom

The proposed method is validated using the OpenCAS [30] public dataset, which contains 3D meshes from an experiment in which a silicone liver phantom is deformed by an indentation. The positions of small Teflon marker balls in both the initial and deformed states of the phantom are given. Please refer to Suwelack et al. [30] for more details about how the dataset was built.

The 3D model of the liver phantom in its initial state is used as the moving model, M. The proposed global registration method is tested in two scenarios. Firstly, a partial view of the deformed liver phantom is used as the target model, T (Fig. 5.5-top), which tests the performance under deformation, partial and sparse data. Secondly, a partial surface reconstructed from an intra-operative stereo endoscopic camera (Fig. 5.5-bottom) is used as T. On top of deformation and partial data, this scenario also tests the proposed method in realistic noise levels from a stereo reconstruction. After the global alignment is estimated with the proposed method, Levenberg-Marquardt Iterative Closest Point (LM-ICP)[36] is applied. The distribution of target registration error (TRE) in mm is computed for both cases. So, the mean TRE for the partial deformed surface is 7.94 mm after the proposed global

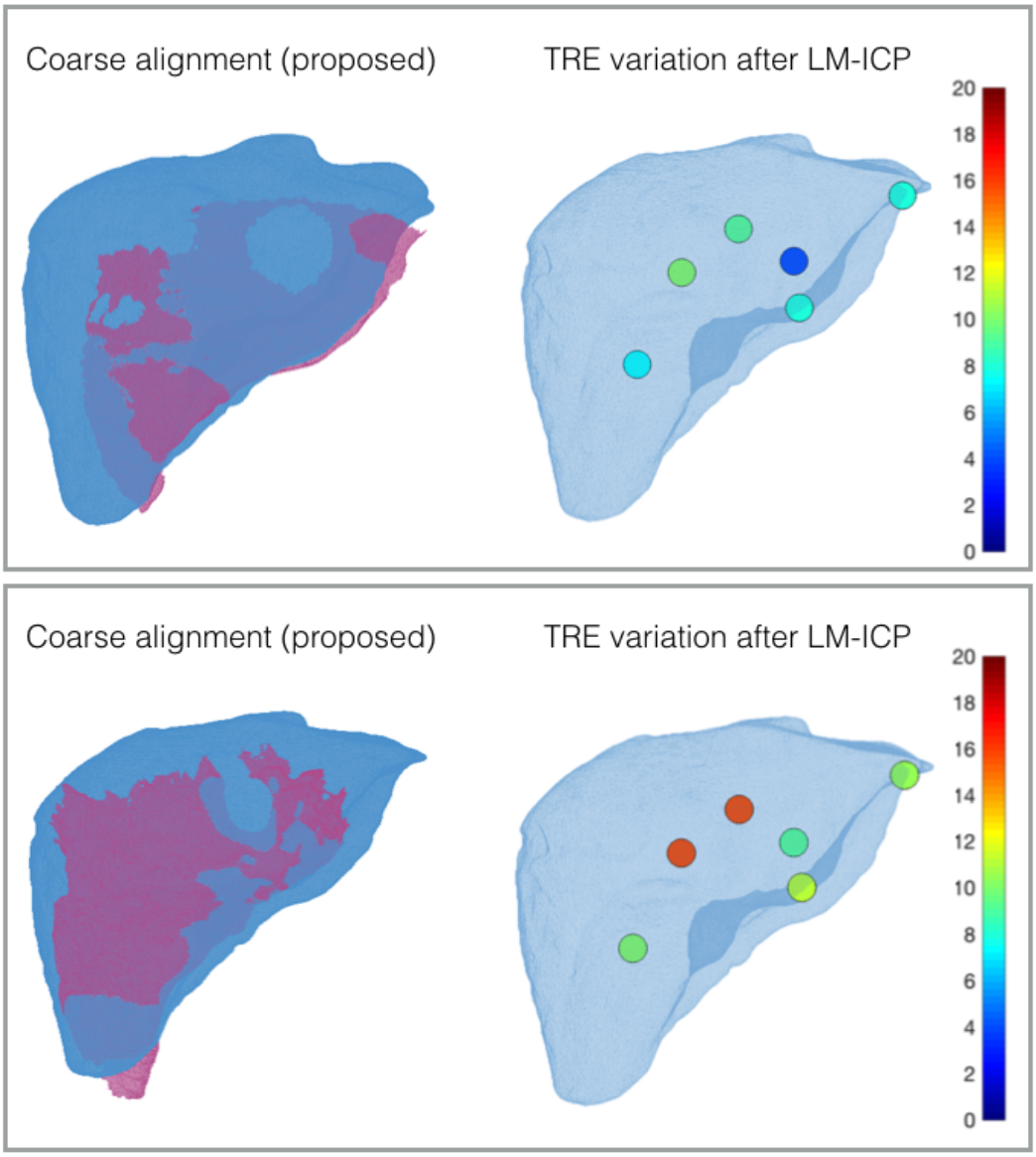

Figure 5.5: Phantom experiment. The proposed global initial alignment is sufficient to allow potentially any fine-alignment method to successfully converge. The global alignment is applied to a partial region of the deformed surface (Top) and a partial surface reconstruction from an intra-operative stereo laparoscopic camera (Bottom), followed by the successful convergence of LM-ICP[36]. The colour map highlights the error distribution (TRE-mm) computed between the ground truth sub-surface markers and their estimated position after alignment. *Colour coding:* moving model, M - blue, target model, T - pink.

alignment, which is further reduced to 7.77 mm after LM-ICP. Similarly, in the case of the intra-operative partial surface, the mean TRE of 28.62 mm after the global alignment is decreased to 12.10 mm after LM-ICP. The best case scenario for rigid registration would be point based alignment of the marker ball positions before and after deformation and the mean TRE is 5.66 mm.

5.3.3 Application in clinical data

The proposed approach is demonstrated on clinical data from a video sequence acquired during a laparoscopic liver resection. The 3D mesh of the liver surface was extracted from a CT scan before surgery[2]. The SmartLiver system [26] is used to process the retrospective data. The liver is automatically segmented in the laparoscopic video with a deep learning framework [11]. Surface patches are collected to cover all the visible surface in each video [69]. They are consequently merged together using optical tracking data.

Fig. 5.6 shows the visual assessment between the manual alignment performed on the SmartLiver GUI and the proposed method. The last column illustrates an example of augmented reality in laparoscopic liver surgery after LM-ICP is applied to the proposed alignment.

5.4 Discussion

The results from the first experiment show that when the intra-operative surface is large enough, all three methods have comparable results. However, having surfaces with size smaller than 23% of the whole mesh becomes challenging for both R and $SM+R1$. From what I noticed in our datasets, having surfaces of 23% or lower is characteristic for videos acquired on the left lobe in laparoscopic interventions with restricted camera movement. Note how the proposed method has less variance in the solutions even for smaller partial shapes. The additional pairwise term which incorporates the boundaries of both M and T makes the problem less ambiguous, as opposed to just using the geodesic distances between pairs of correspondences.

The second experiment illustrates robustness to increasing deformation levels

[2] www.visiblepatient.com

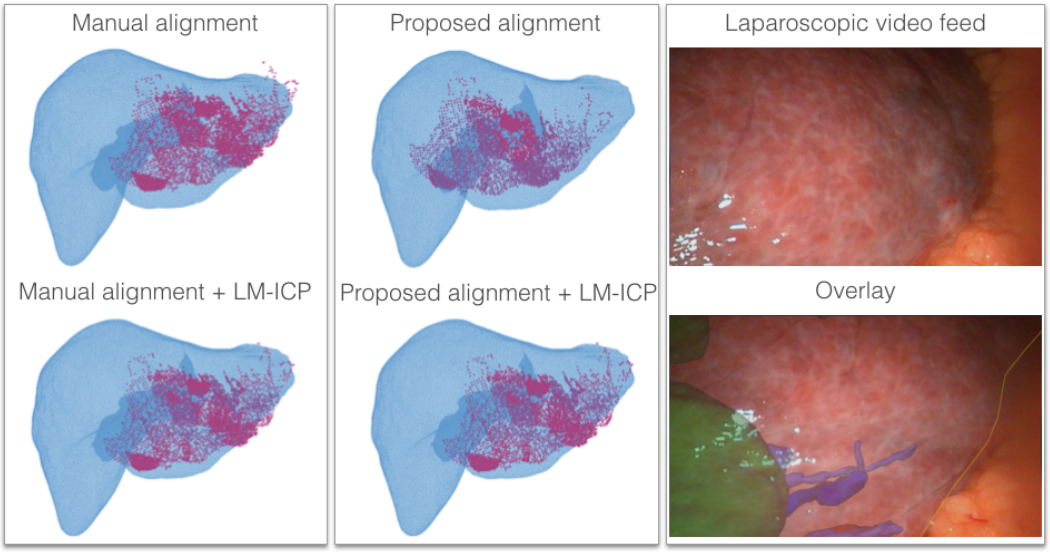

Figure 5.6: Global alignment on clinical data from a dataset acquired during a laparoscopic liver resection. *Colour coding Alignment:* moving model, M - blue, target model, T - pink. The overlay is computed after applying LM-ICP on the proposed alignment. *Colour coding Overlay:* liver tumour - green, vessels - purple, liver contour - yellow.

in the partial views. Similarly to the previous experiment, the proposed method is more consistent across different deformations, with less variation in the solutions it provides. This is mostly due to the fact that the set of correspondences, C_p, obtained from the proposed method contains fewer outliers than the other methods.

These methods are compared with RANSAC because it is a popular algorithm for finding correspondences between point sets related by parametric transformations. If RANSAC is applied directly on the set of correspondences C, it is unable to obtain a good alignment. This is mostly due to the fact that the initial set will contain a high number of outliers, due to the low descriptiveness of the data. Moreover, for partial shape sizes characteristic to laparoscopic surgeries (less than 23% in Fig. 5.4), there are multiple locations on the liver which result in a good fit. However, by allowing for deformation in the proposed pruning technique, a set of correspondences, C_p, is obtained in the correct region of interest, which can be further refined by RANSAC. This approach is also more computationally efficient since RANSAC has to find three good pairs of correspondences from small sets (approximately 10 correspondences, depending on the data).

A quantitative evaluation of the proposed algorithm is performed on a phantom dataset with partial size, deformation and realistic noise levels. The partial surface used in Fig. 5.5-right) is illustrative for an intra-operative scenario, since the data is collected using a stereo laparoscope. The proposed method succeeds in providing a good initial alignment and it is shown that further fine-alignment methods (such as LM-ICP) can successfully converge towards the correct location. The current errors are comparable to the literature in the rigid case scenario [113, 26, 35]. Since most fine registration algorithms can converge successfully if the global alignment is within a few cm [64], the proposed method achieves results within the desired range. In order to decrease the errors further, non-rigid refinement methods could be investigated in the future.

Promising results on a retrospective video acquisitions from a laparoscopic liver surgery are shown. The proposed method is compared against a manual alignment performed on the SmartLiver GUI. Qualitative results are provided to illustrate that the proposed method manages to correctly identify the liver region in a challenging environment with realistic noise levels, significant deformation and small partial views. Note that the proposed method aims to estimate a global surface alignment, which can then be further refined with a local algorithm. For example, Fig. 5.6-right) shows an overlay computed by applying LM-ICP on the proposed global alignment estimate. Furthermore, the current computational time makes it feasible for clinical usage.

The above experiments were all performed using the same choice of parameters, suggesting the proposed method is not very sensitive to variations. However, future validation could investigate their influence. Although the experiments presented here show promising results, future work should include validation on clinical data from more patients to test its robustness with respect to liver surface variations. Furthermore, ways to automate the liver contour selection are needed, such that no user interaction is required and the whole registration process is fully automated.

This chapter introduced a fast, global, semi-automatic surface-based registra-

tion technique of a 3D liver mesh extracted from a pre-operative CT scan of the liver and the surface reconstruction of the intra-operative laparoscopic video feed. I have validated its performance with respect to the challenges characteristic to laparoscopic surfaces on synthetic data, on a phantom dataset and on retrospective clinical data. Moreover, it does not require any advanced hardware, which makes it accessible and comparatively easy to translate to a clinical setting.

Since this chapter is focusing on 3D-3D registration where the intra-operative surface is obtained with stereo reconstruction, the resulting method requires multiple surface patches to be merged in a common space (see Fig. 2.1 for coordinate system setup). Such a task is achieved using optical tracking data of the laparoscope. As discussed in Section 3.5, achieving such registration live during surgery will incorporate multiple error sources (i.e. tracking, triangulation, hand-eye calibration [29]). Such error accumulation effects can be minimised when working with 3D-2D registration (see Fig. 2.2 for coordinate system setup).

While a trend towards incorporating various liver contours in registration frameworks was observed when analysing the state of the art over the past few years (Section 3.3.2), most published methods still assumed large portions of the liver visible intra-operatively (70-100%). An important finding from this chapter is that a global registration can successfully be constrained by the anterior ridge, even in cases with partial visibility levels characteristic to generic laparoscopic interventions (30-50%).

In the next chapter, a novel fully automatic global 3D-2D registration method is proposed which leverages the constraints imposed by the anterior ridge and the liver silhouette for partial visibility.

Chapter 6

3D-2D Automatic Global Registration

In this chapter, an automatic 3D-2D pipeline for global alignment estimation between a pre-operative 3D model and a single laparoscopic image is proposed. It requires no additional user input once the intervention begins. The proposed workflow is composed of two independent modules: i) semantic detection of liver contours on a laparoscopic image using a convolutional neural network; ii) global 3D-2D contour-based alignment. The remainder of this chapter is based on the work titled *"Automatic, Global Registration in Laparoscopic Liver Surgery"* which is in the process of submission at the moment [1].

While stereo laparoscopes are used mostly in robotic systems, monocular scopes are much more commonly available. Several formulations for 3D-2D registration are possible depending on the application. Shape-from-Template proposes to use multiple intra-operative 2D images. A 3D dense model is reconstructed from several key-frames selected in an exploratory video. Then, the pre-operative model is registered to the reconstructed intra-operative surface using 3D-3D registration techniques. The texture on the intra-operative surface can help with long-term tracking by matching it with subsequent frames in the video. However, while such techniques have proven promising in augmented reality systems for the uterus [154], it is difficult to translate them to liver surgery. Several reasons include the presence of breathing motion in the exploratory video and the repetitive texture of the liver

surface. As such, the rest of this chapter focuses on formulating a generally applicable approach by registering the pre-operative 3D model to a single 2D laparoscopic image.

A comprehensive overview of the literature is discussed in Chapter 3. In the next few paragraphs, the most related techniques will be summarised. 3D-2D liver registration requires correspondences to be found between a 3D model and a 2D image of the patient's anatomy under partial visibility. Since this is a difficult problem with multiple solutions, other prior information can be used to constrain the optimisation. The liver boundary [27], the anterior ridge and falciform ligament [25, 28] have been proposed as landmarks. Currently, these techniques either need manual annotation during surgery in order to obtain the liver contours [25, 28, 27], matching end-points [25, 28], or manual global alignment [27]. While the liver boundary could be automated using deep learning [11], a rigid initialisation of the camera pose is still needed which is currently achieved manually [27]. Separating the organ boundary into the anterior ridge and silhouette can lead to automating the registration since they can be matched to the corresponding contours on the 3D model [25], if a large part of both contours is visible in the image. Several approaches have been validated in synthetic experiments for partial data assuming 70%-100% of the liver boundary visible [28, 27]. While such visibility ratios are achievable in open surgery [134], having an unobstructed large view of both liver lobes in laparoscopic intervention requires the cutting of the falciform ligament [25, 28]. However, laparoscopic images from interventions where the falciform ligament is present show only approximately 30% - 50% of the liver surface. For such cases, rigid manual alignment is currently the only reliable option.

The main sources of failure in 3D-2D liver registration stem from the deformation of the 3D model and the partial visibility in the laparoscopic image. Specific characteristics of laparoscopic surgeries can be used to constrain the optimisation. For example, the laparoscopic camera is usually inserted through a trocar placed close to the belly button [25]. Such approximation allows for an initial guess of the camera pose, which can help in finding correspondences between the modali-

ties. Another approach would be to estimate the deformation induced by the pneumoperitoneum [27, 28] and the gravity load [28], which can make the pre-operative 3D model have a more similar shape to the intraoperative image.

Due to the complexity of 3D-2D registration, deep learning techniques can be used for solving well-defined sub-tasks. In the larger computer vision community, the approach proposed in [172] for real-time 3D eyelid tracking from semantic edges is most similar to my work. They use a convolutional neural network (CNN) to detect four edges of the eyelid namely the double-fold, upper eyelid, lower eyelid and lower boundary of the bulge. These detected contours are then used to reconstruct the 3D shape and motion of the eyelids with increased realistic detail. While some similarities do exist, their method assumes the whole eye is visible at all times and explicitly uses the intersection points at the endpoints of the eyelid in the registration formulation. In laparoscopic liver surgery, such an approach would not be possible due to the partial visibility of the organ. Moreover, the variety of liver appearance and illumination fall-off due to the light source being close to the surface makes the laparoscopic environment more complex.

6.1 Contributions

I propose an automatic global 3D-2D registration solution for general laparoscopic interventions. Firstly, a semantic edge detection network is adapted to distinguish between different types of liver contours. I deal with the class imbalance and show how to improve the predictions without performing any additional annotations. Secondly, a traditional pose estimation technique is extended to match corresponding contours, which are only partially visible. Prior information about laparoscopic interventions is explicitly integrated in order to increase the efficiency and reliability of the proposed pipeline.

Furthermore, since the proposed approach estimates the direct transformation between the CT space and the laparoscopic image, it does not contain any of the accumulated errors in the hand-eye, stereo calibrations, optical/EM tracking present in other AR systems [29].

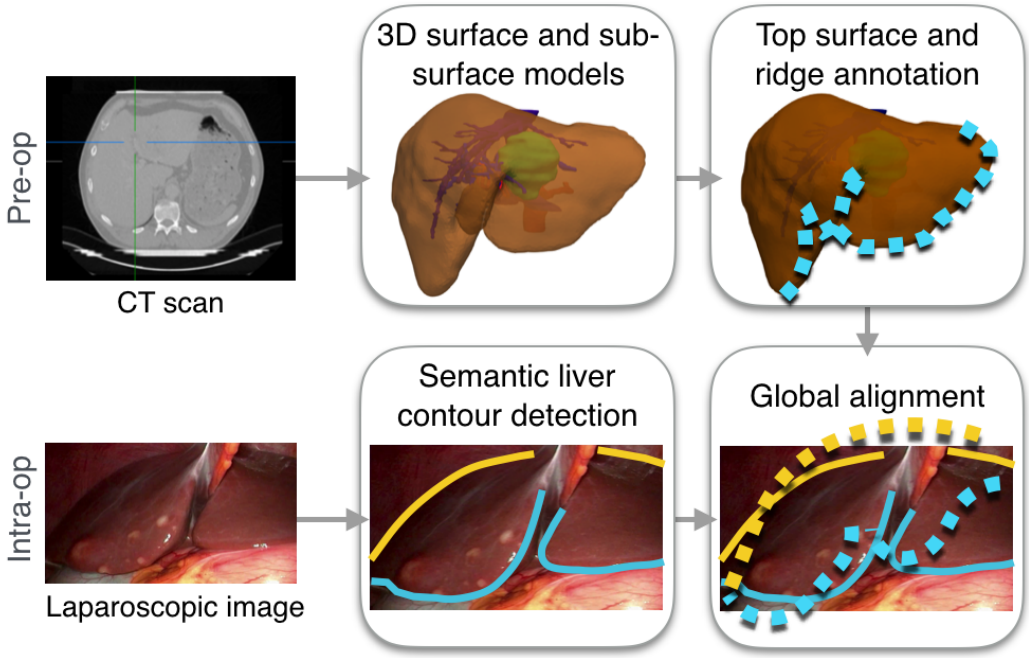

Figure 6.1: An overview of the proposed pre-operative and intra-operative stages for a global 3D-2D registration. Colour coding: silhouette - yellow, anterior ridge - blue

The two proposed components are validated both individually and together as a pipeline. Firstly, I test the semantic liver contour detection on test data from 8 clinical cases with various liver appearances and geometry. Secondly, the simulation framework introduced in Chapter 4 is extended to perform a patient specific analysis of the contour-based registration with respect to partially visible liver surfaces. Such analysis can output pre-operatively a clear protocol for the surgeons of which images to acquire, thus moving towards a more efficient and intuitive initial registration. Finally, the proposed pipeline was validated on 5 of the 8 above clinical cases (for which a pre-operative CT was available) with promising results.

6.2 Methods

An overview of the proposed workflow can be observed in Fig. 6.1. The pipeline takes advantage of no time limit in the pre-operative stage in order to pre-compute several needed components such as: 3D surface and sub-surface extraction from the abdominal CT scan [1], manual annotation of the anterior ridge and the top surface of

[1] www.visiblepatient.com

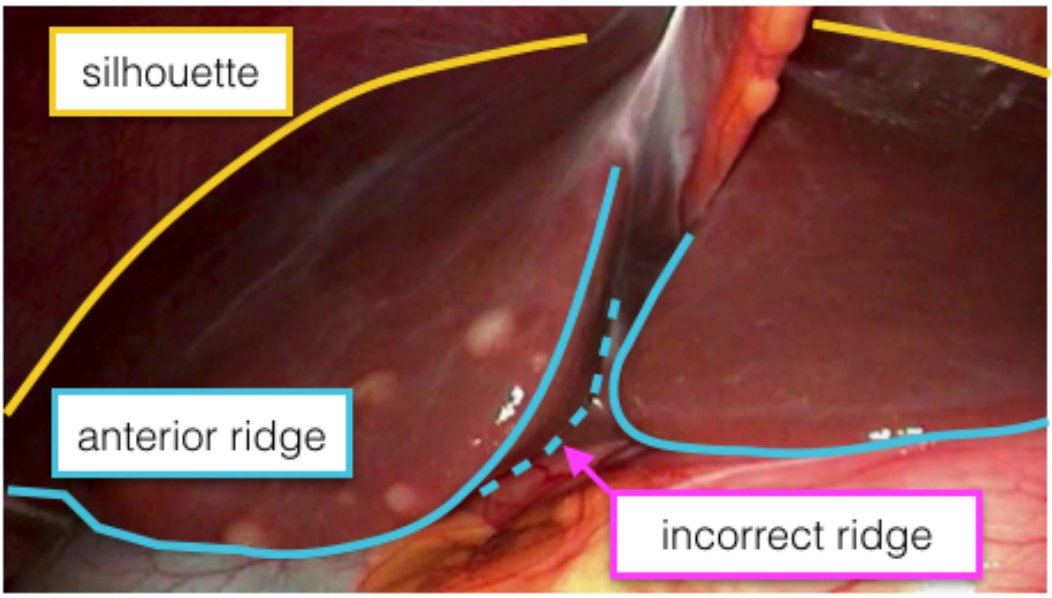

Figure 6.2: Contour labels example: correct annotation for the silhouette (yellow) and the anterior ridge (blue). Around the falciform ligament, the ridge line does not always correspond to the liver boundary - example of an incorrect ridge annotation in the dotted line

the liver. These steps could be easily automated as well [83], however I chose to do it manually due to the variety in liver surface geometry when there are abnormalities present. Moreover, the intrinsic parameters of the laparoscopic camera can also be estimated pre-operatively [77].

I propose to use two types of liver contours: *the anterior ridge* and *silhouette* (Fig. 6.2). The former contour is an anatomical landmark which remains fixed on the organ but can become occluded due to blood or fat. The latter changes depending on the camera position and organ deformation. When used together, these contours can provide complementary constraints to the pose optimisation [25], which becomes essential under partially visibility.

During surgery, there are two main components after the laparoscopic image to be registered is selected:

- semantic liver contour detection
- global 3D-2D contour-based registration

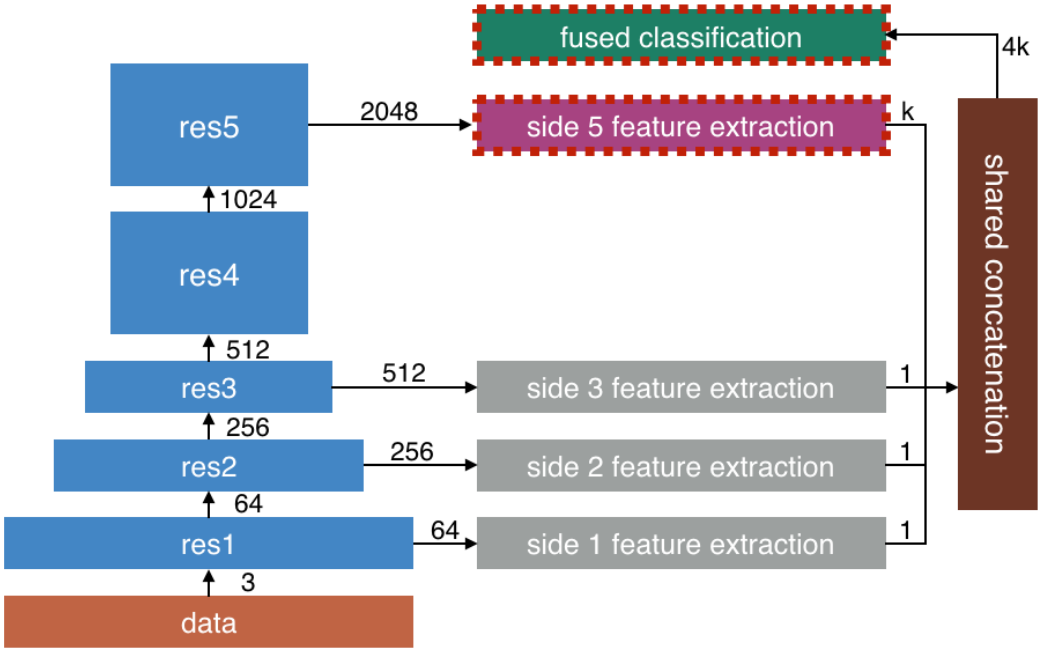

Figure 6.3: CASENet architecture with ResNet [37] as backbone. Solid blue rectangles represent ResNet blocks, where decreasing width translates to a drop in the spatial resolution by a factor of 2. Features are extracted from the bottom layers and concatenated with the top layer classification in the brown block. The fused classification block (green) takes as input the concatenated layers and outputs k maps. The dotted red lines denote the layers where supervision is imposed. More details about the individual blocks can be found in [38]. Illustration adapted from [38]

6.2.1 Semantic contour detection network

In the vision research community, the most recent approaches proposed for semantic edge detection consist of using CNNs to achieve state-of-the-art results in several benchmarks. In particular, CASENet [38] incorporates the intuition that low level features help improve the edge detection, whereas the higher levels might have more semantic information about the overall classes. So, the low-level features extracted from the bottom layers are concatenated with the top layer. Supervision is imposed on the top layer and the fusion of low and high level features. Figure 6.3 illustrates the main components in the architecture. In this work, CASENet is used with ResNet-50 [37] as a backbone. Note that the output number of channels of the *res3* block is 512 (instead of 1024) in the ResNet 50 architecture as well as the official code repository of CASENet.

Semantic contour detection represents a slightly different problem in laparoscopic liver interventions since I do not aim to predict edges belonging to different objects, but contours belonging to different regions of the same object. One of the main challenges is in situations where the ridge line does not coincide with the edge of the organ - see Fig. 6.2 for an incorrect annotation.

The ground truth labels G are represented as three binary maps (i.e. background, silhouette and ridge) where 1 stands for a pixel belonging to the respective class and 0 to the rest. Since the classes should be mutually exclusive for each pixel, semantic liver contour detection is formulated as a multi-class learning problem.

I use a combined loss $\mathscr{L}$ of a weighted categorical cross entropy $\mathscr{L}_{cross}$ and the exponential logarithmic Dice loss $\mathscr{L}_{dice}$:

$$\mathscr{L} = \mathscr{L}_{cross} + \mathscr{L}_{dice} \tag{6.1}$$

$$\mathscr{L}_{cross} = -\mathbf{E}\Big[\sum_{k} w_k * \delta_{gk}(\mathbf{x}) * log(p_k(\mathbf{x}))\Big] \tag{6.2}$$

where $\mathbf{x}$ is the pixel position, k is the class label, g is the ground truth label, $\mathbf{E}[\cdot]$ is the average value with respect to $\mathbf{x}$, $\delta_{gk}(\cdot)$ is the Kronecker delta which is 1 when $g = k$ and 0 otherwise, $p_k(\mathbf{x})$ is the softmax function at pixel $\mathbf{x}$ for class k.

Class imbalance is an inherent issue in edge detection methods due to the large number of background pixels compared to the edges. A proven approach to improve the classification is to balance the influence of the majority and minority classes by introducing weights. So, the weights for the contours $w_{k\neq 0}$ and for the background $w_{k=0}$ are computed as fractions of all the pixels in the ground truth channels $|G|$.

$$\begin{aligned} w_{k\neq 0} &= \frac{\sum \delta_{g0}(\mathbf{x})}{|G|} \\ w_{k=0} &= 1 - w_{k\neq 0} \end{aligned} \tag{6.3}$$

The Dice coefficient can have disproportionately large decreases when few pixels are misclassified in the classification of small structures. The exponential logarithmic Dice loss $\mathscr{L}_{dice}$ was introduced in [173] in order to improve the performance of the coefficient in highly imbalanced datasets:

$$\mathscr{L}_{dice} = \mathbf{E}\Big[(-log(DICE(\delta_{gk}(\mathbf{x}), p_k(\mathbf{x}))))^{0.3}\Big] \tag{6.4}$$

It is shown that the $-log(\cdot)^{0.3}$ function encourages improvements at both small and large values of the Dice coefficient [173] when compared to the linear alternative generally used in segmentation [174].

Once the anterior ridge and silhouette are predicted on the input laparoscopic image, they need to be matched to the corresponding contours on the pre-operative 3D liver model.

6.2.2 3D-2D contour-based registration

Solutions for camera pose estimation from 3D-2D correspondences can be obtained using well established vision techniques such as perspective-n-point (PnP) [78]. Random Sample Consensus (RANSAC) has been proposed to deal with outliers in the correspondence set [74]. A combined PnP-RANSAC approach has been used successfully in multiple AR applications due to its computational efficiency and robustness [78].

In this registration stage, the aim is to find a set of correspondences between the 3D contours estimated on the pre-operative 3D model and the 2D contours predicted by the network on the laparoscopic image. It is essential that the registration algorithm is robust to false positives and negatives in the contour prediction. A PnP technique would be straight-forward to apply once the correspondence point pairs are found. However, this task is not trivial when working with partially visible contours.

Algorithm 1 describes the proposed contour-based PnP-RANSAC extension, which can handle partial occlusion.

Prior information characteristic of general laparoscopic interventions is inte-

Algorithm 1: Global contour-based 3D-2D registration

Input: Pre-op 3D liver model M,
canonical liver transform T_C,
2D predicted contours p^{2D}_{label},
intrinsic matrix K
Parameters: number of initial perturbations m,
max iterations RANSAC j_{max},
min distance contours d_{min},
Result: Global rigid transformation T_{global}

1 $\{T^C_{init}\}_m$= GenerateInitalGuesses(M, T_C, m);
2 $\{T_{init}\}_m$ = CanonicalToOrig($\{T^C_{init}\}_m, T_C^{-1}$);
3 **for** *each* $T_{init} \in \{T_{init}\}_m$ **do**
4 p^{3D}_{label} = Estimate3DVisibleContours(M, T_{init});
5 $\tilde{p}^{2D}_{label}$ = ProjectContours2Image($p^{3D}_{label}, T_{init}, K$);
6 $\{\tilde{p}^{2D}_{label}, p^{2D}_{label}\}$ = EstimCorresp($\tilde{p}^{2D}_{label}, p^{2D}_{label}$);
7 T_{optim} = EstimPose($M, \{\tilde{p}^{2D}_{label}, p^{2D}_{label}\}, K, j_{max}, d_{min}$);
8 $T_{global} \leftarrow$ save the final transformation T_{optim};

grated in order to make the registration robust to partially visible surfaces and constrain the solution space. A transformation of the liver to a canonical space can be pre-computed (T_C). This step is possible due to a common reasonable assumption that the laparoscopic camera will be inserted through a trocar placed approximately around the belly button of the patient [25]. In this work, the liver is aligned with the positive x axis pointing towards the left lobe, the positive y axis towards the liver posterior and the z axis pointing upwards. Such transforms to a predefined local space help in dealing with varying liver geometry and generating candidate camera positions which are relevant for laparoscopic environments. While the camera faces the liver, the movement of the laparoscope is constrained due to the fixed trocar position (see Chapter 4). Let the camera follow the right-handed coordinate system with the x, y and z axes pointing respectively to the right, up, and away from the camera. A range of m initial camera poses are generated in the canonical space $\{T^C_{init}\}_m$ by random perturbations of rotation around the y and z axis ($\in \mathcal{N}(0, 20°)$) of the camera (line 1), which makes the registration robust to smaller areas of the

liver being visible (i.e. only a part of the left lobe). These initial transformation guesses $\{T^C_{init}\}_m$ are brought back to the original space of the 3D liver model (line 2).

For each initial transformation guess (line 3), the visible contours are estimated on the 3D model (line 4) for each label - anterior ridge and silhouette. Firstly, the visible surface M_{vis} is estimated from a given camera position (similar to [27]). Secondly, the visible surface M_{vis} is intersected with the top liver surface M_{top}, which is obtained pre-operatively. This step ensures all the points on the posterior side of the liver are excluded, since they are generally not visible in the initial stages of a laparoscopic intervention. Thirdly, the final 3D ridge points are obtained from the visible surface ($p^{3D}_{ridge} = M_{vis} \cap M_{ridge}$). Lastly, the silhouette $p^{3D}_{silhouette}$ is estimated as all the boundary points of the visible surface M_{vis} that do not belong to the ridge.

Once the visible 3D ridge and silhouette points are estimated, they can be projected on top of the 2D image, using the known intrinsic parameters (line 5). Correspondences between the contours are computed in the image space by searching for the closest neighbour and similar normals (less than 30° difference). The threshold used for normal similarity allows for some deformation between the contours, as shown in previous work [27]. As such, the function *EstimCorresp* on line 6 outputs a set of corresponding points $\{\tilde{p}^{2D}_{label}, p^{2D}_{label}\}$ where the points $\tilde{p}^{2D}_{label}$ belong to the projected 3D contours and the points p^{2D}_{label} to the predicted contours on the laparoscopic image.

The PnP-RANSAC algorithm is then employed to find the optimal camera pose T_{optim} for the current iteration (line 7). Since a rigid transform has six degrees of freedom, the minimal solution of this optimisation consists of finding 3 pairs of 3D-2D correspondences. The P3P technique proposed in [79] is used, which needs an additional correspondence pair for a unique solution. Consequently, the PnP-RANSAC workflow consists of randomly selecting a minimal sample of 4 pairs from the correspondence set $\{\tilde{p}^{2D}_{label}, p^{2D}_{label}\}$ at each iteration j. Since the correspondences for the ridge and the silhouette contours are considered separately, special care is taken to ensure 2 correspondence pairs from each set are selected. PnP is then

employed to estimate the camera pose T_k for the current minimal set of point pairs. In order to improve the computation efficiency, camera poses that are in impossible locations for a laparoscopic intervention are discarded (i.e. inside, behind the liver). In order to measure the agreement of the whole set of correspondences with the current estimate, a distance error is computed between the projected 3D contours (transformed using T_k) and the 2D contours. The chosen error is the modified Hausdorff distance [175] which enforces the corresponding contours to be similar. The stopping criteria for RANSAC consist of the maximum number of iterations j_{max} and a minimal threshold for the distance error d_{min}. Finally, the camera pose estimate T_{optim} with the minimal modified Hausdorff distance over all the iterations is considered the global 3D-2D transformation T_{global} (line 8).

Choosing larger values for the maximum RANSAC iterations j_{max} and the number of perturbations m will generally translate in increased accuracy at the expense of a higher computational time. Additionally, it is not desired to set the contour distance threshold d_{min} to 0, as the contours are not expected to match perfectly due to the deformation induced by the pneumoperitoneum. Consequently, the set of parameters chosen should allow for a good trade-off between accuracy and computational time.

6.3 Experiments

Separate experiments have been setup for the semantic contour detection and the global contour-based registration. Moreover, the whole pipeline was tested on clinical data from 5 laparoscopic interventions. The following section details the experimental setup for each component.

6.3.1 Semantic contour detection

Since in laparoscopic interventions only a partial view of the liver is visible, it is essential to be able to predict as much of the contours as possible. Pre-training on a synthetic dataset of laparoscopic images was shown to improve the generalisation and classification performance of a liver segmentation task [5]. I explore using a similar strategy to increase the recall of the proposed classifier.

6.3.1.1 Training data

The training clinical data (*C*) consists of 133 images extracted from two laparoscopic interventions acquired with a stereo and a monocular laparoscope respectively. A single channel was used from the stereo pairs. The data was annotated by a clinical fellow where polygonal lines were drawn on top of each contour type (Fig. 6.2). In order to cope with any misalignment between the labels and the true edge, the contours are dilated before the training, similar to [38].

The pre-training dataset (*S*) consists of approximately 100 000 synthetic laparoscopic images [5] (more details in Section 3.5.1). The synthetic images were generated to account for various organ geometry and appearance. All the images that did not contain either ridge or silhouette were excluded. Among other labels, the authors provide an anterior ridge mask and a liver mask [5]. I extracted the silhouette as all the boundary liver pixels that are neighbouring the abdominal wall mask. Similar to the clinical dataset, the contours are dilated as a pre-processing step.

Two training scenarios are considered, where weights are first initialised from ResNet50 pre-trained on the ImageNet dataset (*I*) [176]:

- **I+C** : CASENet is trained on the clinical dataset (C).
- **I+S+C** : CASENet is pre-trained on the synthetic dataset(S). Then, the network is fine-tuned on the clinical dataset (C).

Data augmentation is used to make the network invariant to brightness changes, contrast, rotations, translations, scale changes and shear. When training CASENet on the synthetic dataset, I use as data augmentation only brightness changes and contrast in order to make the predictions more insensitive to different liver appearances. The Adam optimiser was used for training the network on the clinical dataset with learning rate 1e-04. A checkpoint is saved at the lowest validation loss, which is used to generate the results presented here.

6.3.1.2 Test data

The performance of the proposed model is evaluated on 3 test datasets:

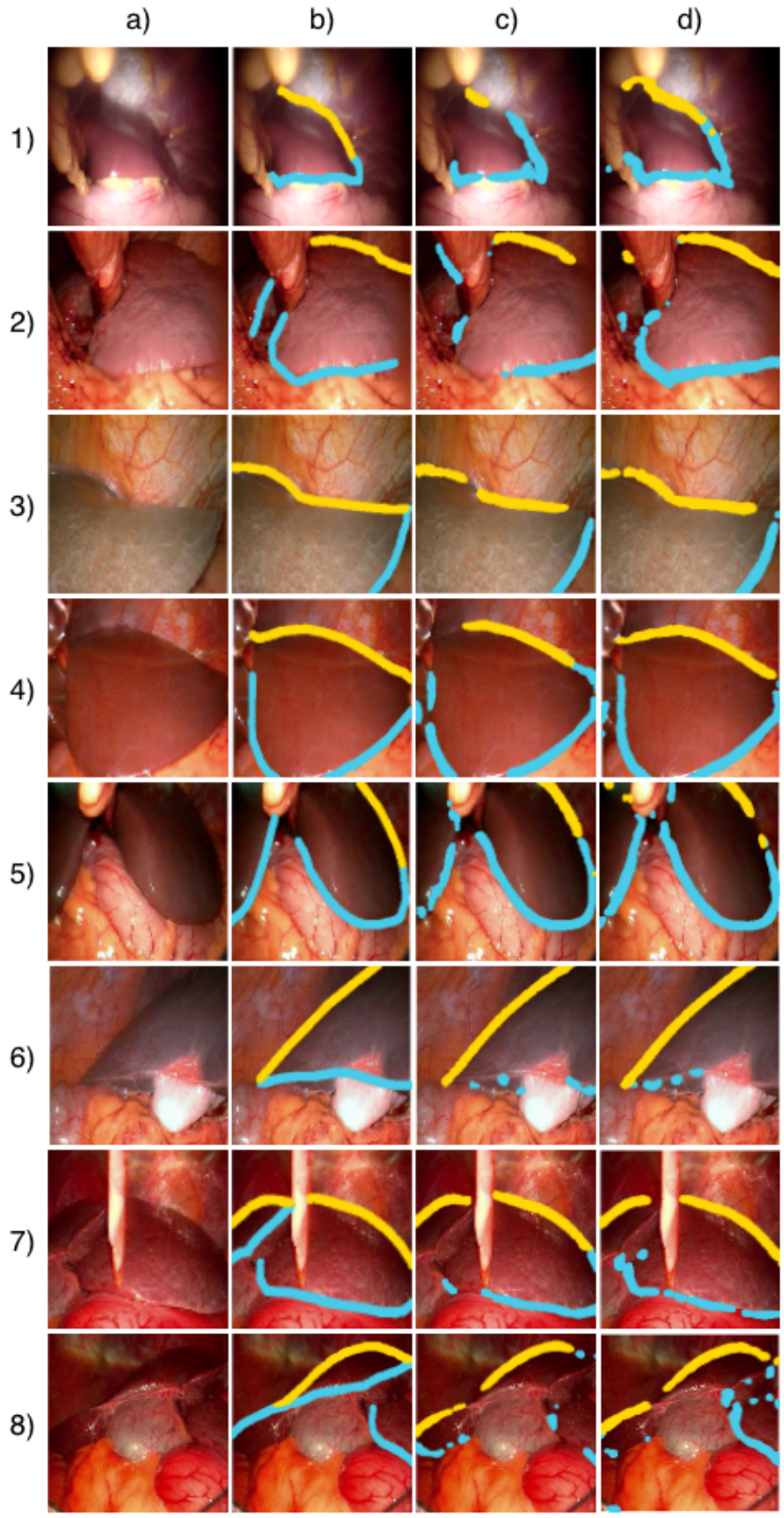

Figure 6.4: Qualitative results of ridge (blue) and silhouette (yellow) prediction: a) Original images from the test datasets: *daVinci*: row 1), *lap1*: rows 2)-6), *lap2*:rows 7)-8); b) Ground truth annotations; c) Predictions when CASENet was trained only on a clinical dataset ($I+C$); d) Predictions when CASENet was trained with synthetic and clinical data ($I+S+C$)

- *daVinci* - 9 images from a daVinci intervention
- *lap1* - 9 images from 6 clinical cases, all acquired with a stereo laparoscope
- *lap2* - 12 images from 1 clinical case, acquired with a stereo laparoscope

The images used in all the test datasets are from different clinical cases than the training data. Fig. 6.4 shows a selection of images from each dataset with the ground truth annotations and the predictions obtained using the two training scenarios. The images have been selected by trying to minimise the overlap of identical views. Any images including blur or imaging artefacts have been excluded.

The *daVinci* dataset aims to test how the model performs when a new acquisition method is used to capture the images, which was not present in the training process. The *lap1* dataset includes a few images from multiple clinical cases covering a wide range of liver appearances and geometry. The *lap2* dataset was included to evaluate the performance on a special case, which has not been encountered in the clinical training data. The right lobe is lifted so much that the silhouette is meeting the ridge in the middle of the contour - see Fig. 6.4 - row 9). In all other images, the intersection points of the silhouette and the ridge coincide with the end of the liver lobe.

Three accuracy measures are used for evaluating the network performance [177]: *precision P* (out of all the predicted contour pixels, how many are correctly labelled?), *recall R* (how many of the ground truth contour pixels are predicted as correct?) and *F1 score* (Eq. 6.5):

$$F1 = \frac{2 \cdot P \cdot R}{P + R + \varepsilon} \tag{6.5}$$

where ε is a small number to avoid the denominator being zero. Table 6.1 summarises the results for each dataset in the two training scenarios. Fig. 6.5 shows quantitative maps with true positives (green), false positives (blue) and false negatives (red) for some of the predictions in the test datasets, along with their associated F1 scores.

Table 6.1: Performance of CASENet on the test datasets: F1 score (F1), Recall (R) and Precision (P). The numbers represent the average over all the images in each dataset.

Dataset		daVinci		lap1		lap2	
Training		I+C	I+S+C	I+C	I+S+C	I+C	I+S+C
F1	*ridge*	0.39	0.46	0.41	0.45	0.29	0.42
	silhouette	0.61	0.75	0.74	0.76	0.73	0.77
R	*ridge*	0.39	0.46	0.42	0.51	0.20	0.32
	silhouette	0.57	0.79	0.74	0.82	0.68	0.77
P	*ridge*	0.43	0.48	0.44	0.45	0.63	0.66
	silhouette	0.70	0.71	0.75	0.72	0.82	0.79

Note that images from both stereo and monocular laparoscopes have been used in the datasets in order to cover different acquisition modalities. However, the proposed method aims to be generally applicable including in monocular laparoscopic interventions. As such, no 3D information is used and a single channel is extracted from the stereo pair.

6.3.2 Registration

The registration performance depends on the uniqueness of the constraints imposed by the contours. Since the characteristics of the contours (such as the curvature) vary greatly depending on the viewing angle, both partial visibility and the liver region need to be taken into consideration.

I adapt the pre-operative simulation framework proposed in Chapter 4 for quantitatively analysing the performance of the proposed 3D-2D contour-based registration. Originally, the method was used for pre-operatively computing a data acquisition protocol in which the surgeon would acquire specific liver surface patches which would ultimately lead to an efficient 3D-3D registration. The approach is appealing because it provides a way to analyse which specific camera views would result in a good registration.

So, the simulation framework loads a 3D liver model from a clinical case. Sev-

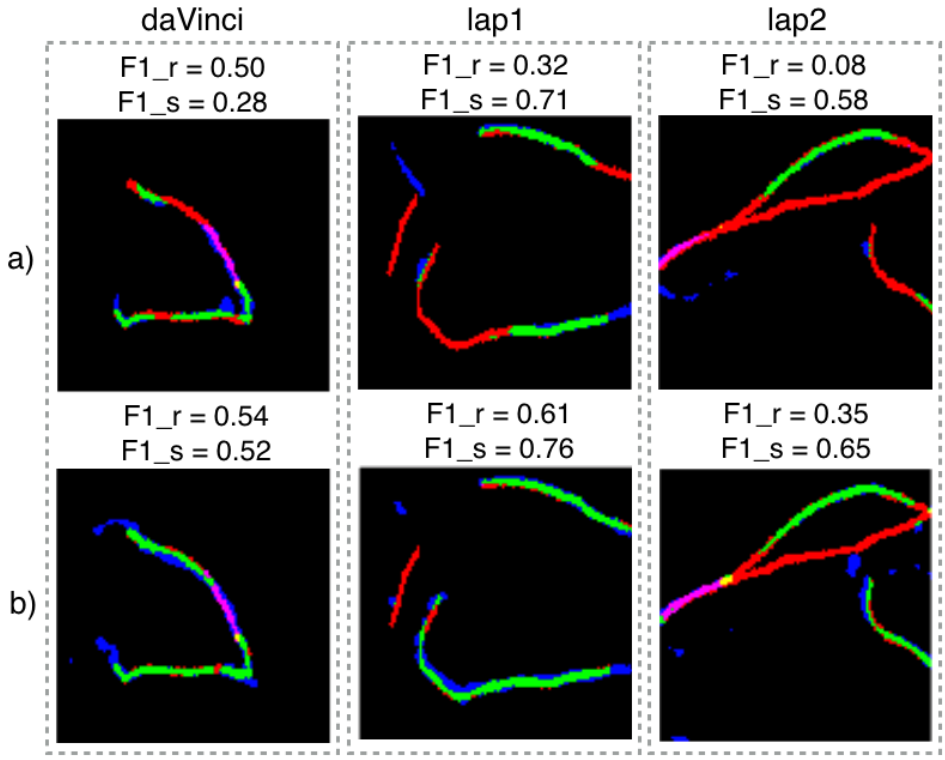

Figure 6.5: Example of F1 score maps for the ridge ($F1_r$) and silhouette ($F1_s$) where green = true positives, blue = false positives and red = false negatives. *a)* Predictions when CASENet was trained on the clinical dataset ($I+C$); *b)* Predictions when CASENet was pre-trained on the synthetic dataset and fine-tuned on the clinical one ($I+S+C$)

eral camera positions are simulated on a sphere around the liver at two distances (Fig. 6.6 b)). For each camera, a synthetic image is obtained by estimating the visible contours (ridge and silhouette) and projecting them to 2D. For each camera position, the proposed contour-based registration is run 100 times between the synthetic image and the 3D liver surface, in order to account for the sources of randomness (Algorithm 1 lines 1,7). The visibility percentage was computed as a ratio between the visible vertices over the total vertices of the front liver surface (similar to previous methods [28, 27]). The root mean square error (RMSE) is measured between the ground truth vertex positions of the 3D liver surface and the estimated vertices obtained after the registration process. See Fig. 6.6 for the results.

On top of analysing the robustness to partial visibility, such a pre-operative planning simulation can provide a clear protocol to clinicians with regards to which portions of the liver provide sufficient constraints for the registration.

6.3.3 Automatic global registration pipeline

On top of separate validation for each component, it is important to asses the pipeline as a whole in order to understand the feasibility of such a method in a laparoscopic liver intervention.

The dataset used to validate the proposed registration pipeline is composed of 14 images from 5 cases. These have been selected from *lap1* and *lap2* such that either the left lobe, right lobe or the middle section are present. The images from the *daVinci* dataset and from 2 of the cases in *lap1* were excluded due to the absence of a pre-operative model derived from CT. The laparoscopic camera was calibrated in order to obtain the intrinsic parameters before each intervention started [77]. The same set of parameters is used across the dataset: the number of perturbations $m = 60$, the maximum number of RANSAC iterations $j_{max} = 60$ and the minimum distance between the contours $d_{min} = 30$. Figures 6.7 and 6.8 illustrate the network predictions and the registration results from each case, along with examples of errors the pipeline is robust to.

6.4 Results

6.4.1 Semantic contour detection

Table 6.1 shows good prediction rates across all datasets. Note that while the F1 scores may not be very high, this is due to the difficulty in predicting contours with exactly the same thickness, a common side-effect of such networks [172]. Additionally, Fig. 6.5 shows that the majority of the false positives are in close proximity to the true positives. Several updated network architectures [178, 179] or post-processing techniques such as curve-fitting [172] could be used to improve the precision. The trade-off between precision and recall generally depends on the application. In this chapter, the predicted contours are needed for correspondence estimation with the 3D liver contours. Consequently, detecting as much of the contour as possible (high recall) is critical, especially for images where the liver is only partially visible.

While in most cases the silhouette is easily detected with a high accuracy, the

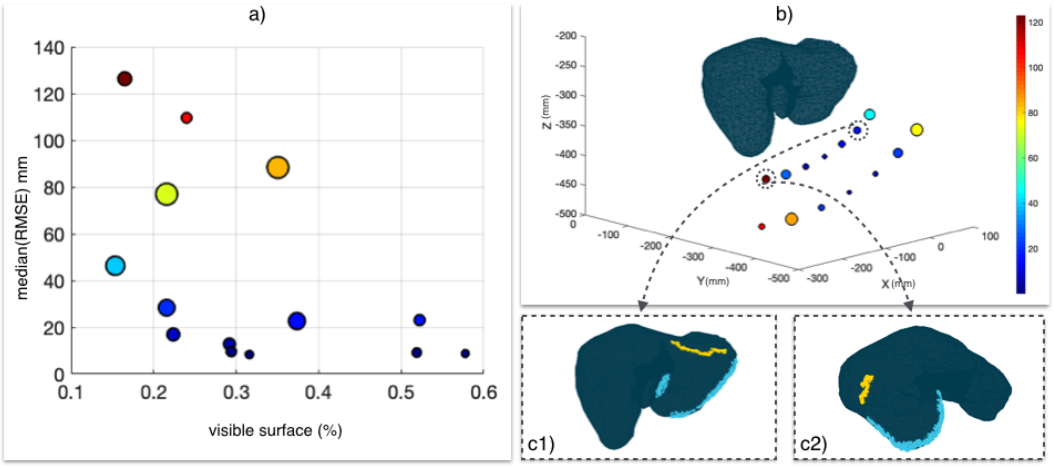

Figure 6.6: *a)* Registration accuracy(mm) for various surface visibility ratios. The size of the circles is proportional to the standard deviation over 100 registrations. *b)* Error variation (mm) over liver region. Each circle represents a camera location. The circle colour is given by *median(RMSE)* (shown in the colour bar) and their size is proportional to the standard deviation over 100 registrations. The same colours were used in graph a) in order to help match corresponding cameras. Examples of visible ridge (blue) and silhouette (yellow) from chosen camera positions that lead to successful (*c1*) and failed (*c2*) registrations.

anterior ridge presents much more variation. The task is also especially difficult where the ridge does not coincide with the liver boundary (Fig. 6.2). However, by pre-training on the synthetic dataset I show an increase in recall with an average of 9.33 % for the ridge and 13 % for the silhouette across all datasets. Such improvement can also be observed in the examples in Fig. 6.5, where the amount of true positives increases with pre-training on the synthetic dataset.

Fig. 6.4 shows excellent generalisation across livers with significant changes in appearance (i.e. rows 4,6,7) and across different image acquisition methods. Notice how pre-training improves how much of the contour gets detected, especially in the last row which presents a case never encountered in the training set C (see Section 6.3.1 for more details).

6.4.2 Registration

Fig. 6.6 a) shows that the proposed registration method copes with severe occlusion of the liver surface. It manages to estimate a good initial alignment (within several cm [64]) with as little as 30% visible liver surface. Since laparoscopic images gen-

erally capture approximately 30-50% of the liver surface, these results are highly encouraging.

Fig. 6.6 b) illustrates the camera locations considered in the pre-operative simulation analysis as circles. Also, each circle colour and size is proportional to the success of a registration from the respective position. Consequently, such visualisation can provide more insight into which areas constrain the registration better, as originally proposed in Chapter 4.

Fig. 6.6 c) shows examples of visible 3D contours from 2 chosen camera positions. The c1) configuration corresponds to a visibility of 22% resulting in *median(RMSE)*=16.78 mm and can be considered a successful global alignment. On the other hand, the c2) configuration represents 16% visibility and leads to a failed registration with *median(RMSE)*=126.42 mm.

What I noticed in this experiment is that even smaller areas (Fig. 6.6 c2) can look like they are successfully registered when comparing the contours on the image (projected from 3D vs. true 2D contours). However, when comparing the ground truth transformation with the estimated one, the error in the z direction was large. This effect disappears in general when the visible area increases or portions of the anterior ridge neighbouring the falciform ligament are included (Fig. 6.6 c1). Such findings also confirm that apart from contour visibility ratio, the area of the liver captured is also important.

6.4.3 Automatic global registration pipeline

The proposed pipeline was successful in estimating a global alignment on all 5 clinical cases in the registration dataset. Without a ground truth available, a global alignment can be considered successful if the estimated pose identifies the correct liver area. As such, the 3D liver model visualised from the estimated camera's perspective should approximately match the laparoscopic image, as can be seen in Fig. 6.7 b)-c). Note that even if the intra-operative liver shape is deformed from the pre-operative model, the proposed pipeline can successfully recover a good global transformation. Moreover, it can cope with various liver geometry and appearance along with different viewpoints of the liver.

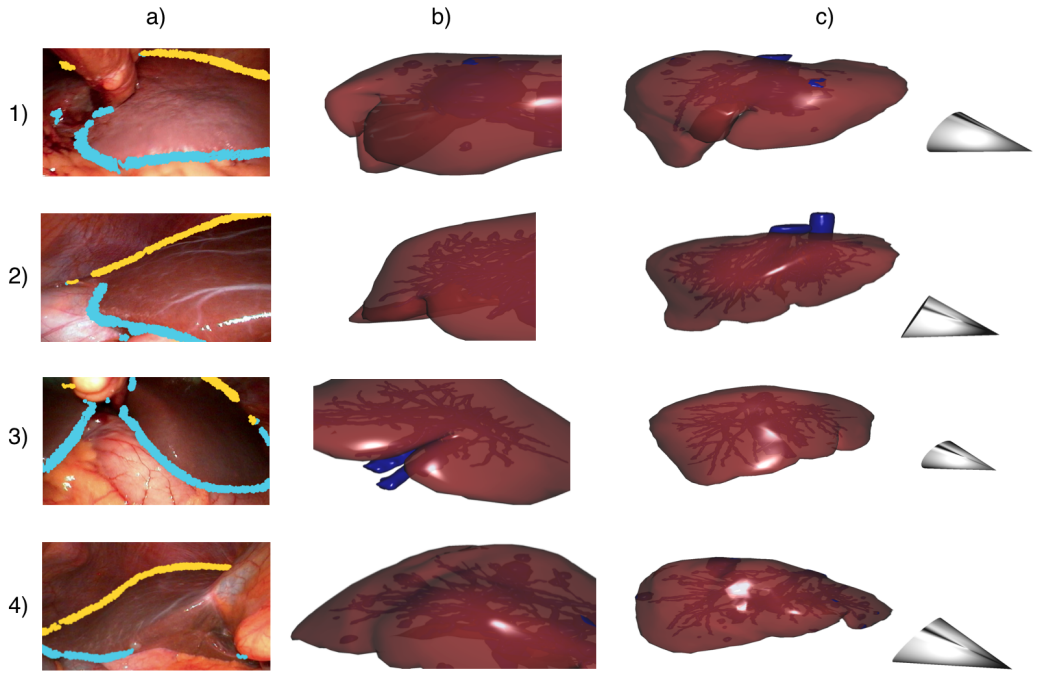

Figure 6.7: Qualitative results of the proposed global registration pipeline on retrospective clinical cases. *a)* Predictions of the ridge (blue) and silhouette (yellow) by the proposed semantic contour detection ($I+S+C$). *b)* Aligned 3D preoperative liver model using the proposed camera pose estimation. *c)* A rendering of the liver model and estimated camera pose in 3D space.

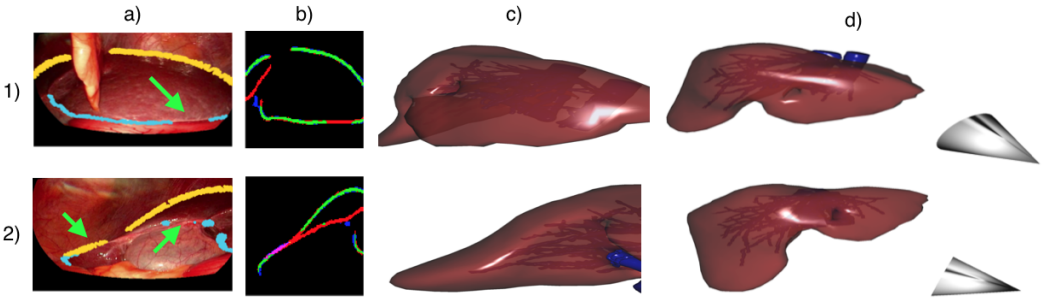

Figure 6.8: The proposed pipeline is robust to errors in the contour detection: false negatives, false positives or misclassification. *a)* Predictions of the ridge (blue) and silhouette (yellow) by the proposed semantic contour detection ($I+S+C$). *b)* Corresponding error maps for the predicted contours where green = true positives, blue = false positives, red = false negatives and pink = misclassification. *c)* Aligned 3D preoperative liver model using the proposed camera pose estimation. *d)* A rendering of the liver model and estimated camera pose in 3D space.

The computational time needed for the 3D-2D registration component is approximately 2 minutes with non-optimised code for the clinical cases in the registration dataset (Fig. 6.7 and 6.8), with the parameters discussed in Section 6.3.3.

Figure 6.8 highlights several errors that might arise in the semantic contour detection stage: false negatives, false positives and misclassification. Fig. 6.8, row 2) shows the ridge line mistaken for the silhouette on the right lobe (see Sec. 6.3.1). Nonetheless, the proposed registration manages to successfully align the pre-operative model even in these cases. Consequently, the increase in generalisation and recall outweighs the occasional errors in false positives and negatives, if the following stages in the pipeline are designed with robustness in mind.

6.5 Discussion

While the proposed pipeline can successfully estimate a global registration without taking into consideration the liver deformation, a straightforward extension of this work would be to perform a deformation simulation before surgery, such as in [28]. They show promising results by simulating the deformation induced by the pneumoperitoneum while also taking into account the gravity load.

Given the current running time of the proposed pipeline, it could be used at the beginning of the surgery to obtain an initial registration based on a laparoscopic image, relying on optical tracking of the laparoscope in order to update the alignment. Since the current version of the code is not optimised, future versions might open up the possibility of running the proposed pipeline every few frames, thus removing the need of external trackers.

Another important point is that the analysis on synthetic data (Fig. 6.6) depends heavily on liver anatomy. As can be seen in Fig. 6.6 a-b), there is no linear relationship between the visibility percentage of the surface and the registration error. Rather, certain contour configurations (i.e. curvature, liver region) lead to better constraints for the registration. However, these configurations will vary depending on liver geometry, as can be seen in Figures 6.7, 6.8 c). While the proposed method can cope with such variety, it would be beneficial to run such a pre-operative plan-

ning analysis (Fig. 6.6) before each intervention, as discussed in Chapter 4. As a result, different camera positions could be explored before surgery, leading to clear indications for the surgeon of which liver region to capture in order to get a reliable registration faster. Having a clear, automatic protocol for initial registration will make AR systems more intuitive and usable in the OR.

While AR generally implies showing an overlay to the surgeon, there is still no consensus as to which visualisation strategy would best benefit a certain clinical application. Some challenges include the need for high accuracy of the alignment and tracking, along with a visualisation that is not distracting or confusing for the surgeon. Nonetheless, surgical guidance can also be achieved by having a separate monitor showing the aligned pre-operative anatomy next to the laparoscopic view (see Fig. 6.7, 6.8). Such a side-by-side guidance was reported to be useful since the surgeons could mentally compensate for the slight misalignment for a rigid registration with approximately 10 mm accuracy [65]. The proposed pipeline can achieve similar results, while being fully automatic and intuitive to perform.

Validation was provided for each component, along with a qualitative assessment of the proposed pipeline on clinical data. Firstly, I show how to improve the detection and generalisation of the proposed semantic contour detection by pre-training on a synthetic dataset. Also, the proposed model is validated in 8 new clinical cases with excellent results. Secondly, I show that the proposed registration can estimate a global alignment with as little as 30% of the liver surface visible by extending a patient-specific pre-operative analysis. Overall, the proposed registration pipeline was successfully applied in 5 retrospective clinical cases and it was robust to the occasional errors in the contour prediction stage.

Several other methods [25, 28, 2] could benefit from the proposed automatic contour prediction by replacing all manual input. Moreover, the estimated global registration can provide a solution for ICP-based methods in 3D-2D or 3D-3D registration since they require a good initial alignment, which is usually performed manually [27, 29]. I hope the proposed automatic global registration pipeline can improve augmented reality systems in laparoscopic interventions to be more effi-

cient and intuitive for surgeons.

Chapter 7

Conclusion

Laparoscopic liver surgery offers benefits for patients, such as less pain, trauma and fewer days in the hospital, when compared to open surgery. However, due to the additional difficulty in localising sub-surface structures and the risk of haemorrhage, most cases are recommended for open surgery. Augmented reality is considered a potential solution for the challenges introduced by laparoscopic interventions, such as: weak depth perception, constrained vantage point, limited field of view, weak haptics and occluded anatomy. This thesis focuses on computer vision tools for the registration of a pre-operative liver surface to the intra-operative scene in the initial stages of the surgery, an essential component of AR systems.

The main methods in the literature concerning registration in laparoscopic liver interventions were analysed with respect to the requirements for an augmented reality system that is compatible with the current clinical workflow. Namely, an AR system in laparoscopic liver surgery should be efficient, intuitive, easy to integrate in the current surgical protocols, reliable.

Several technical contributions have been made leading to 2 first-author journal publications (in IJCARS 2017 [3], IJCARS 2018 [2], a third one in the submission process [1]) and 10 collaborations on other related projects [4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

A pre-operative planning simulation framework was proposed for an improved intra-operative data acquisition strategy (see Chapter 4). The main motivation behind this work was to understand which regions constrain surface-based registration

better, in order to reduce the amount of redundant data collected intra-operatively. Moreover, the framework was designed to output a clear protocol for surgeons such that the subsequent registration is more efficient.

Once the intra-operative surface is acquired (i.e. by merging multiple surface reconstruction patches), most rigid and non-rigid registration solutions in the literature depend on an initial global alignment. This step is performed manually in the majority of published methods. Some semi-automatic methods have been proposed, however, they have been validated on large areas of the liver (which are not available in the initial stages of laparoscopic interventions). A fast, semi-automatic method is proposed to estimate the global alignment, which incorporates a novel contour-based constraint (see Chapter 5). The proposed approach shows promising results on synthetic data, a liver phantom dataset and on retrospective data from one laparoscopic liver intervention.

Moreover, a fully automatic pipeline for 3D-2D global, rigid registration was proposed, which copes with partial surfaces characteristic of laparoscopic interventions (see Chapter 6). Once the surgery starts, no manual intervention is required. Firstly, the anterior ridge and silhouette of the liver are detected on a laparoscopic image using a neural network. The proposed classifier showed excellent generalisation on test data from 8 cases, with variations in liver appearance, anatomical geometry, patient pathology and data acquisition. Secondly, a novel contour-based rigid registration is proposed in order to estimate the camera position and orientation. The pre-operative planning simulation framework introduced in Chapter 4 is used to understand which configurations of liver contours lead to sufficient errors, showing good registration capabilities with as little as 30% of the liver visible. On top of its flexibility, such a pre-operative simulation can prove valuable in analysing the robustness and reliability of a registration technique. Moreover, the proposed automatic global registration pipeline was successfully applied across 14 images from 5 clinical cases with promising results.

Overall, two different approaches have been explored to create a more intuitive global rigid registration: 3D-3D using merged stereo reconstruction patches

as the intra-operative surface (Chapter 5) and 3D-2D using automatically detected contours (Chapter 6). An interesting future avenue of work would be to collect data and compare these two approaches directly. Since the work undertaken here was validated on retrospective data with various protocols, such comparisons could not be made. However, based on the available experiments and registration formulation, the 3D-2D contour based approach from Chapter 6 presents significant advantages: from using widely available monocular cameras, having a completely automatic pipeline to being validated in many more clinical cases. Moreover, the proposed fast liver contour estimation could be used in multiple subsequent non-rigid registration techniques available in the literature which rely on the availability of these landmarks [25, 28].

Finally, the clinical requirements for an augmented reality system have played a critical role in designing and developing the proposed tools. Firstly, *efficiency* is obtained by maximising the use of the pre-operative data in order to decrease the computational load during the surgery. Secondly, the proposed approaches aim to be *intuitive* and clear for the surgeon by: i) computing clear protocols to be followed for intra-operative data acquisition, ii) moving from manual initial alignment (Chapter 4) to a semi-automatic approach which involves brief manual annotation (Chapter 5) and to a completely automatic initial registration (Chapter 6). Thirdly, all the proposed methods could be comparatively *easily translated in the OR* since they follow the current clinical protocol, without introducing additional hardware. Chapters 4 and 5 need a stereo laparoscope whereas Chapter 6 describes a method using a single laparoscopic image and so, can be obtained using a monocular laparoscope. Finally, since it is extremely important for the surgeons to tell if the current alignment is *reliable*, the main focus of this work was on rigid alignment. Visual cues, such as the misalignment of the projected and the image contours can be an important indicator of how accurate the current overlay is.

An important direction for future research would be how to convey to the surgeon the accuracy of the current alignment. For rigid registration, contour-based misalignments (surface based projection errors) have been shown to be a reliable

predictor of sub-surface TRE in clinical data [29]. Such work would allow for a visualisation of the registered sub-surface targets (i.e. tumours, vessels) with confidence levels of accuracy. Thus, the surgeon would have a better understanding of when to trust the AR system and when to proceed with alternative strategies (i.e. re-initialising the registration, intra-operative ultrasound). However, there is still no indication of how to best validate deformable algorithms. Most methods in the literature resort to simulations and phantom experiments. While they are helpful to validate the algorithms in a controlled environment, it is usually extremely difficult to create similar conditions to a real surgery. As such, it is unclear how they translate and generalise to clinical cases. Public synthetic datasets can prove essential in benchmarking and comparing multiple algorithms. A promising direction is currently explored by generating images in a synthetic environment and using deep learning to change their appearance to match real laparoscopic interventions [5]. Having the possibility to use a large dataset, fully annotated with segmentation labels, depth information and camera poses could prove pivotal in how the next few years of research in laparoscopic liver interventions develop. With the release of this data, more challenges or hackathons could be organised as part of larger conferences in order to increase collaborations in the field and generate more ideas. Better validation techniques and more publicly available data would help bridge the gap between research and clinical application.

Currently, augmented reality research in laparoscopic liver surgery focuses on the development of the technology and further increasing its performance. It could be said that the benefits of AR in specific clinical applications will become more clear and easier to validate once the system reaches a certain performance level. However, it would be useful to define smaller steps of a procedure where an overlay could be advantageous. For example, a general aim of AR systems is to localise sub-surface structures and help the surgeon target the lesions while avoiding the vessels. A desired accuracy for an alignment around a lesion would range between 3 and 5 mm. Higher errors might lead to positive surgical margins during resection. However, if the main target is to avoid hitting a vessel, the accuracy required would

be much smaller. Moreover, registration algorithms would need to be validated more thoroughly to account for that, since a rotational error might be more severe than a translational one.

Once better error quantification is available, an interesting direction is to understand the human computer interaction factors at play when surgeons use image guidance. Showing all the pre-operative information at once all the time would be overwhelming and distracting. On the other hand, having more sub-structures be revealed the closer the laparoscope gets to the liver surface might prove more informative, especially with regards to depth perception. Close collaboration with surgeons is needed to explore various visualisation techniques that will ultimately allow them to focus on the intervention, without any disruptions.

Major advances have been made towards registering laparoscopic ultrasound to CT images [9]. Using an ultrasound during surgery provides real-time imaging of vasculature or tumours, thus increasing the safety of laparoscopic procedures. The position of sub-surface structures in the ultrasound plane could inform the deformation correction needed for video-CT registration with improved accuracy. On the other hand, a video-CT registration could provide information about the pose of the ultrasound probe, thus resulting in improved ultrasound-CT registration. On top of providing complementary information to improve the registration stages of such a pipeline, such a guidance system could cope with additional tumours that are not captured in the CT scan. Then, the surgeons would be able to re-evaluate the surgical plan. Consequently, a combined video and ultrasound system for surgical guidance could lead to improved accuracy, robustness and increased clinical relevancy.

Bibliography

[1] Maria R. Robu, Bongjin Koo, Moustafa Allam, Micha Pfeiffer, Stephen Thompson, Kurinchy Gurusamy, Brian Davidson, David Hawkes, Danail Stoyanov, and Matthew J. Clarkson. Automatic, Global Registration in Laparoscopic Liver Surgery, 2019.

[2] Maria R Robu, João Ramalhinho, Stephen Thompson, Kurinchi Gurusamy, Brian Davidson, David Hawkes, Danail Stoyanov, and Matthew J Clarkson. Global Rigid Registration of CT to Video in Laparoscopic Liver Surgery. *International Journal of Computer Assisted Radiology and Surgery*, 2018.

[3] Maria R. Robu, Philip Edwards, João Ramalhinho, Stephen Thompson, Brian Davidson, David Hawkes, Danail Stoyanov, and Matthew J. Clarkson. Intelligent viewpoint selection for efficient CT to video registration in laparoscopic liver surgery. *International Journal of Computer Assisted Radiology and Surgery*, 12(7):1079–1088, Jul 2017.

[4] Stephen Thompson, Thomas Dowrick, Guofang Xiao, Joao Ramalhinho, Maria R. Robu, Mian Ahmad, and Matthew J. Clarkson. scikit-surgery-torso-simulator: An Ultrasound Simulator. *Journal of Open Research Software*, 2019.

[5] Micha Pfeiffer, Isabel Funke, Maria R Robu, Sebastian Bodenstedt, Sandy Engelhardt, Tobias Roß, Matthew J Clarkson, Brian R Davidson, Lena Maier-hein, and Carina Riediger. Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image transla-

tion. *Medical Image Computing and Computer-Assisted Intervention MICCAI*, pages 1–8, 2019.

[6] Yunguan Fu, Maria R Robu, Bongjin Koo, Crispin Schneider, Stijn van Laarhoven, Danail Stoyanov, Brian Davidson, Matthew J. Clarkson, and Yipeng Hu. More unlabelled data or label more data? A study on semi-supervised laparoscopic image segmentation. *Medical Image Learning with Less Labels and Imperfect Data Workshop at MICCAI 2019*, pages 1–8, 2019.

[7] Agoston Mihalik, Mikael Brudfors, Maria Robu, Fabio S. Ferreira, Hongxiang Lin, Anita Rau, Tong Wu, Stefano B. Blumberg, Baris Kanber, Maira Tariq, Maria Del Mar Estarellas Garcia, Cemre Zor, Daniil I. Nikitichev, Janaina Mourao-Miranda, and Neil P. Oxtoby. ABCD Neurocognitive Prediction Challenge 2019: Predicting individual fluid intelligence scores from structural MRI using probabilistic segmentation and kernel ridge regression. *ABCD Neurocognitive Prediction Challenge at MICCAI*, 2019.

[8] Neil P Oxtoby, Fabio S Ferreira, Agoston Mihalik, Mikael Brudfors, Hongxiang Lin, Anita Rau, B Stefano, Maria Robu, Cemre Zor, Maira Tariq, Maria Del, Mar Estarellas Garcia, Baris Kanber, and Daniil I Nikitichev. ABCD Neurocognitive Prediction Challenge 2019: Predicting individual residual fluid intelligence scores from cortical grey matter morphology. *ABCD Neurocognitive Prediction Challenge at MICCAI*, 2019.

[9] João Ramalhinho, Maria R. Robu, Stephen Thompson, Kurinchi Gurusamy, Brian Davidson, David Hawkes, Dean Barratt, and Matthew J. Clarkson. A pre-operative planning framework for global registration of laparoscopic ultrasound to CT images. *International Journal of Computer Assisted Radiology and Surgery*, 13(8):1177–1186, 2018.

[10] João Ramalhinho, Maria Robu, Stephen Thompson, Philip Edwards, Crispin Schneider, Kurinchy Gurusamy, David Hawkes, Brian Davidson, Dean Bar-

ratt, and Matthew J. Clarkson. Breathing motion compensated registration of laparoscopic liver ultrasound to CT. In *SPIE Medical Imaging*, page 101352V, 2017.

[11] Eli D. Gibson, Maria R. Robu, Stephen A. Thompson, Philip E. Edwards, Crispin Schneider, Kurinchi Gurusamy, Brian R. Davidson, David J. Hawkes, Dean C. Barrat, and Matthew J. Clarkson. Deep residual networks for automatic segmentation of laparoscopic videos of the liver. *SPIE*, 10135:101351M–101351M–6, 2017.

[12] Stephen Thompson, Eddie Edwards, Joao Ramalhinho, Maria Robu, Danail Stoyanov, Kurinchi Gurusamy, Crispin Schneider, Brian Davidson, and Matthew J Clarkson. Calibration Pattern Selection for Laparoscope Calibration. *SPIE*, pages 1–4, 2017.

[13] Geoffrey Jones, Neil T. Clancy, Xiaofei Du, Maria Robu, Simon Arridge, Daniel S. Elson, and Danail Stoyanov. Fast estimation of haemoglobin concentration in tissue via wavelet decomposition. In *Medical Image Computing and Computer-Assisted Intervention MICCAI*, 2017.

[14] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68:394–424, 2018.

[15] Claude Couinaud, Andre Delmas, and Jean Patel. *Le foie: etudes anatomiques et chirurgicales*. Masson & Cie, Paris, 1957.

[16] Hanisah Guro, Jai Young Cho, Ho-Seong Han, Yoo-Seok Yoon, YoungRok Choi, and Mohan Periyasamy. Current status of laparoscopic liver resection for hepatocellular carcinoma. *Clinical and Molecular Hepatology*, 22(2):212–218, 2016.

[17] Stéphane Nicolau, Luc Soler, Didier Mutter, and Jacques Marescaux. Augmented reality in laparoscopic surgical oncology. *Surgical Oncology*, 20(3):189–201, sep 2011.

[18] Sofien Bouaziz, Andrea Tagliasacchi, Hao Li, and Mark Pauly. Modern techniques and applications for real-time non-rigid registration. In *SIGGRAPH ASIA 2016 Courses*, SA '16, pages 11:1–11:25, New York, NY, USA, 2016. ACM.

[19] Camera calibration and 3d reconstruction. `https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html`. Accessed: 2019-09-16.

[20] Simon Prince. *Computer Vision: Models, Learning, and Inference*, page 598. 2012.

[21] Ozan Oktay, Li Zhang, Tommaso Mansi, Peter Mountney, Philip Mewes, Stéphane Nicolau, Luc Soler, editor="Mori Kensaku Chefd'hotel, Christophe", Ichiro Sakuma, Yoshinobu Sato, Christian Barillot, and Nassir Navab. Biomechanically driven registration of pre- to intra-operative 3d images for laparoscopic surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 1–9, 2013.

[22] What everyone with diabetes should know about liver disease. `https://www.empoweryourhealth.org/magazine/vol9_issue1/what_everyone_with_diabetes_should_know_about_liver_disease`. Accessed: 2018-02-01.

[23] Micha Pfeiffer, Hannes Kenngott, Anas Preukschas, Matthias Huber, Lisa Bettscheider, and Beat Müller-stich Stefanie Speidel. IMHOTEP : virtual reality framework for surgical applications. *International Journal of Computer Assisted Radiology and Surgery*, 2018.

[24] Julie Hallet, Luc Soler, Michele Diana, Didier Mutter, Thomas F. Baumert, François Habersetzer, Jacques Marescaux, and Patrick Pessaux. Trans-

thoracic minimally invasive liver resection guided by augmented reality. *Journal of the American College of Surgeons*, 220(5):e55–e60, 2015.

[25] Bongjin Koo, Erol Özgür, Bertrand Le Roy, Emmanuel Buc, and Adrien Bartoli. Deformable Registration of a Preoperative 3D Liver Volume to a Laparoscopy Image Using Contour and Shading Cues. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, 2017.

[26] Stephen Thompson, Johannes Totz, Yi Song, Stian Johnsen, Danail Stoyanov, Sébastien Ourselin, Kurinchi Gurusamy, Crispin Scheider, Brian Davidson, David Hawkes, and Matthew J. Clarkson. Accuracy validation of an image guided laparoscopy system for liver resection. In *SPIE*, volume 9415, page 941509, 2015.

[27] Yinoussa Adagolodjo, Raffaella Trivisonne, Nazim Haouchine, Stéphane Cotin, and Hadrien Courtecuisse. Silhouette-based Pose Estimation for Deformable Organs Application to Surgical Augmented Reality. *International Conference on Intelligent Robots and Systems*, pages 1–7, 2017.

[28] Erol Özgür, Bongjin Koo, Bertrand Le Roy, Emmanuel Buc, and Adrien Bartoli. Preoperative liver registration for augmented monocular laparoscopy using backward–forward biomechanical simulation. *International Journal of Computer Assisted Radiology and Surgery*, 2018.

[29] S. Thompson, C. Schneider, M. Bosi, K. Gurusamy, S. Ourselin, B. Davidson, D. Hawkes, and M.J. Clarkson. In vivo estimation of target registration errors during augmented reality laparoscopic surgery. *International Journal of Computer Assisted Radiology and Surgery*, 2018.

[30] Stefan Suwelack, Sebastian Röhl, Sebastian Bodenstedt, Daniel Reichard, Rüdiger Dillmann, Thiago dos Santos, Lena Maier-Hein, Martin Wagner, Josephine Wünscher, Hannes Kenngott, Beat P. Müller, and Stefanie Speidel. Physics-based shape matching for intraoperative image guidance. *Medical Physics*, 41(11):111901, 2014.

[31] H. G. Kenngott, J. J. Wünscher, M. Wagner, A. Preukschas, A. L. Wekerle, P. Neher, S. Suwelack, S. Speidel, F. Nickel, D. Oladokun, L. Maier-Hein, R. Dillmann, H. P. Meinzer, and B. P. Müller-Stich. OpenHELP (Heidelberg laparoscopy phantom): development of an open-source surgical evaluation and training tool. *Surgical Endoscopy and Other Interventional Techniques*, 29(11):3338–3347, 2015.

[32] Toby Collins, Pauline Chauvet, Clément Debize, Daniel Pizarro, Adrien Bartoli, Michel Canis, and Nicolas Bourdel. A system for augmented reality guided laparoscopic tumour resection with quantitative ex-vivo user evaluation. In *Computer-Assisted and Robotic Endoscopy*, pages 114–126, Cham, 2017. Springer International Publishing.

[33] Hani J Marcus, Philip Pratt, Archie Hughes-hallett, Thomas P Cundy, Adam P Marcus, Guang-zhong Yang, Ara Darzi, and Dipankar Nandi. Comparative effectiveness and safety of image guidance systems in neurosurgery: a preclinical randomized study. *Journal of Neurosurgery*, 123:1–7, 2015.

[34] J Irving Vasquez-Gomez, L Enrique Sucar, Rafael Murrieta-Cid, and Efrain Lopez-Damian. Volumetric Next-best-view Planning for 3D Object Reconstruction with Positioning Error. *International Journal of Advanced Robotic Systems*, 11:1, 2014.

[35] Thiago Ramos dos Santos, Alexander Seitel, Thomas Kilgus, Stefan Suwelack, Anna-Laura Wekerle, Hannes Kenngott, Stefanie Speidel, Heinz-Peter Schlemmer, Hans-Peter Meinzer, Tobias Heimann, and Lena Maier-Hein. Pose-independent surface matching for intra-operative soft-tissue marker-less registration. *Medical Image Analysis*, 18(7):1101–1114, 2014.

[36] Andrew W Fitzgibbon. Robust registration of 2D and 3D point sets. *Image and Vision Computing*, 21(13):1145 – 1153, 2003.

[37] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[38] Z. Yu, C. Feng, M.Y. Liu, and S. Ramalingam. CASENet: Deep category-aware semantic edge detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[39] Gerard J Tortora and Bryan H Derrickson. *Principles of anatomy and physiology*. John Wiley & Sons, 2008.

[40] Liver: Anatomy and functions. `https://www.hopkinsmedicine.org/healthlibrary/conditions/liver_biliary_and_pancreatic_disorders/liver_anatomy_and_functions_85,P00676`. Accessed: 2018-02-01.

[41] Hashem B. El-Serag and K. Lenhard Rudolph. Hepatocellular carcinoma: Epidemiology and molecular carcinogenesis. *Gastroenterology*, 132(7):2557 – 2576, 2007.

[42] Josep M Llovet, Andrew Burroughs, and Jordi Bruix. Hepatocellular carcinoma. *The Lancet*, 362:1907 – 1917, 2003.

[43] Freddie Bray and Isabelle Soerjomataram. The Changing Global Burden of Cancer: Transitions in Human Development and Implications for Cancer Prevention and Control. In *Disease Control Priorities, Third Edition (Volume 3): Cancer*, volume 49, pages 23–44. The World Bank, 2015.

[44] Kunio Okuda, Toshio Ohtsuki, Hiroshi Obata, Masahiko Tomimatsu, Nobuo Okazaki, Hiroshi Hasegawa, Yukio Nakajima, and Kunihiko Ohnishi. Natural history of hepatocellular carcinoma and prognosis in relation to treatment study of 850 patients. *Cancer*, 56(4):918–928, 1985.

[45] Jordi Bruix and Josep M. Llovet. Prognostic prediction and treatment strategy in hepatocellular carcinoma. *Hepatology*, 35(3):519–524, 2002.

[46] Josep M Llovet, Myron Schwartz, and Vincenzo Mazzaferro. Resection and liver transplantation for hepatocellular carcinoma. In *Seminars in liver disease*, volume 25, pages 181–200, 2005.

[47] Robert J. Aragon and Naveenraj L. Solomon. Techniques of hepatic resection. *Journal of Gastrointestinal Oncology*, 3(1):28–40, 2012.

[48] Daniel Palmes and Hans-Ullrich Spiegel. Animal models of liver regeneration. *Biomaterials*, 25(9):1601 – 1611, 2004. Animal Models for Tissue Engineering Applications.

[49] Spiros G Delis and Christos Dervenis. Selection criteria for liver resection in patients with hepatocellular carcinoma and chronic liver disease. *World Journal of Gastroenterology : WJG*, 14(22):3452–3460, 2008.

[50] Daniel Azoulay, Denis Castaing, Jinane Krissat, Alloua Smail, Guillermo Marin Hargreaves, Antoinette Lemoine, Jean-François Emile, and Henri Bismuth. Percutaneous portal vein embolization increases the feasibility and safety of major liver resection for hepatocellular carcinoma in injured liver. *Annals of surgery*, 232(5):665, 2000.

[51] Zarina S. Lalmahomed, Ninos Ayez, Anne E. M. van der Pool, Joanne Verheij, Jan N. M. IJzermans, and Cornelis Verhoef. Anatomical versus nonanatomical resection of colorectal liver metastases: Is there a difference in surgical and oncological outcome? *World Journal of Surgery*, 35(3):656–661, Mar 2011.

[52] Ruben Ciria, Daniel Cherqui, David A. Geller, Javier Briceno, and Go Wakabayashi. Comparative short-term benefits of laparoscopic liver resection: 9000 cases and climbing. *Annals of Surgery*, 263(4):761–777, 2016.

[53] Jun-Jie Xiong, Kiran Altaf, Muhammad A Javed, Wei Huang, Rajarshi Mukherjee, Gang Mai, Robert Sutton, Xu-Bao Liu, and Wei-Ming Hu. Meta-analysis of laparoscopic vs open liver resection for hepatocellular carcinoma. *World Journal of Gastroenterology : WJG*, 18(45):6657–6668, 2012.

[54] Alexis Laurent, Claude Tayar, Marion Andréoletti, Jean Yves Lauzet, Jean Claude Merle, and Daniel Cherqui. Laparoscopic liver resection facilitates salvage liver transplantation for hepatocellular carcinoma. *Journal of Hepato-Biliary-Pancreatic Surgery*, 16(3):310–314, 2009.

[55] Joseph F Buell, Daniel Cherqui, David A Geller, Nicholas O'rourke, David Iannitti, Ibrahim Dagher, Alan J Koffron, Mark Thomas, Brice Gayet, Ho Seong Han, et al. The international position on laparoscopic liver surgery: The Louisville Statement, 2008. *Annals of surgery*, 250(5):825–830, 2009.

[56] J Cho, H Han, Y Yoon, and S Shin. Outcomes of laparoscopic liver resection for lesions located in the right side of the liver. *Archives of Surgery*, 144(1):25–29, 2009.

[57] Go Wakabayashi, Daniel Cherqui, David A Geller, Joseph F Buell, Hironori Kaneko, Ho Seong Han, Horacio Asbun, Nicholas O'Rourke, Minoru Tanabe, and Alan J Koffron. Recommendations for laparoscopic liver resection: a report from the second international consensus conference held in Morioka. *Annals of Surgery.*, 261(4):619–629, 2015.

[58] Adrien Bartoli, Toby Collins, Nicolas Bourdel, and Michel Canis. Computer assisted Minimally Invasive Surgery: Is medical Computer Vision the answer to improving laparosurgery? *Medical Hypotheses*, 79(6):858–863, dec 2012.

[59] Pauline Chauvet, Toby Collins, Clement Debize, Lorraine Novais-Gameiro, Bruno Pereira, Adrien Bartoli, Michel Canis, and Nicolas Bourdel. Augmented reality in a tumor resection model. *Surgical Endoscopy and Other Interventional Techniques*, pages 1–10, 2017.

[60] Benjamin J. Dixon, Michael J. Daly, Harley Chan, Allan D. Vescan, Ian J. Witterick, and Jonathan C. Irish. Surgeons blinded by enhanced navigation: The effect of augmented reality on attention. *Surgical Endoscopy and Other Interventional Techniques*, 27(2):454–461, 2013.

[61] Archie Hughes-Hallett, Philip Pratt, James Dilley, Justin Vale, Ara Darzi, and Erik Mayer. Augmented reality: 3D image-guided surgery. *Cancer Imaging*, 15(1):1, 2015.

[62] Vanessa M. Banz, Matthias Baechtold, Stefan Weber, Matthias Peterhans, Daniel Inderbitzin, and Daniel Candinas. Computer planned, image-guided combined resection and ablation for bilobar colorectal liver metastases. *World Journal of Gastroenterology*, 20(40):14992–14996, 2014.

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

[64] L. Maier-Hein, Alfred M Franz, T. R. dos Santos, Mirko Schmidt, Markus Fangerau, Hans-peter Meinzer, and J Michael Fitzpatrick. Convergent Iterative Closest-Point Algorithm to Accomodate Anisotropic and Inhomogenous Localization Error. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(8):1520–1532, 2012.

[65] Yuichiro Hayashi, Kazunari Misawa, David J Hawkes, and Kensaku Mori. Progressive internal landmark registration for surgical navigation in laparoscopic gastrectomy for gastric cancer. *International Journal of Computer Assisted Radiology and Surgery*, 11:837–845, 2016.

[66] D. Caleb Rucker, Yifei Wu, Logan W. Clements, Janet E. Ondrake, Thomas S. Pheiffer, Amber L. Simpson, William R. Jarnagin, and Michael I. Miga. A Mechanics-Based Nonrigid Registration Method for Liver Surgery Using Sparse Intraoperative Data. *IEEE Transactions on Medical Imaging*, 33(1):147–158, 2014.

[67] Sylvain Bernhardt, Stéphane A Nicolau, Luc Soler, and Christophe Doignon. The status of augmented reality in laparoscopic surgery as of 2016. *Medical Image Analysis*, 37:66–90, 2017.

[68] Stephen Thompson, Danail Stoyanov, Brian Davidson, David Hawkes, and Matthew J Clarkson. Hand-Eye Calibration For Rigid Laparoscopes Using An Invariant Point. *International Journal of Computer Assisted Radiology and Surgery*, 2015.

[69] Johannes Totz, Stephen Thompson, Danail Stoyanov, Kurinchi Gurusamy, BrianR. Davidson, David J. Hawkes, and Matthew J. Clarkson. Fast Semi-dense Surface Reconstruction from Stereoscopic Video in Laparoscopic Surgery. In *IPCAI*, volume 8498, pages 206–215, 2014.

[70] Daniel Reichard, Martin Wagner, Stefanie Speidel, Daniel Reichard, Sebastian Bodenstedt, Stefan Suwelack, Benjamin Mayer, and Anas Preukschas. Intraoperative on-the-fly organ mosaicking for laparoscopic surgery. *Journal of Medical Imaging*, 2(4), 2015.

[71] Nader Mahmoud, Toby Collins, Alexandre Hostettler, Luc Soler, Christophe Doignon, and Jose Maria Martinez Montiel. Live Tracking and Dense Reconstruction for Handheld Monocular Endoscopy. *IEEE Transactions on Medical Imaging*, 38(1):79–89, jan 2019.

[72] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, 1987.

[73] Olga Sorkine and Michael Rabinovich. Least-squares rigid motion using svd. *Technical notes*, pages 1–6, 2009.

[74] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications ACM*, 24(6):381–395, 1981.

[75] Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[76] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the ICP algorithm. In *3D Digital Imaging and Modeling*, pages 145–152, 2001.

[77] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

[78] Eric Marchand, Hideaki Uchiyama, and Fabien Spindler. Pose estimation for augmented reality: a hands-on survey. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

[79] Xiao-Shan Gao, Xiao-Rong Hou, Jianliang Tang, and Hang-Fei Cheng. Complete solution classification for the perspective-three-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):930–943, 2003.

[80] L. Kneip, D. Scaramuzza, and R. Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2969–2976, 2011.

[81] Vincent Lepetit, Francesc Moreno-Noguer, and Pascal Fua. Accurate Non-Iterative O(n) Solution to the PnP Problem. *International Journal of Computer Vision*, 81(2):155–166, 2009.

[82] Long Quan and Zhongdan Lan. Linear N-Point Camera Pose Determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):774–780, 1999.

[83] Rosalie Plantefève, Igor Peterlik, Nazim Haouchine, and Stéphane Cotin. Patient-Specific Biomechanical Modeling for Guidance During Minimally-Invasive Hepatic Surgery. *Annals of Biomedical Engineering*, 44(1):139–153, 2016.

[84] Matteo Fusaglia, Hanspeter Hess, Marius Schwalbe, Matthias Peterhans, Pascale Tinguely, Stefan Weber, and Huanxiang Lu. A clinically applicable laser-based image-guided system for laparoscopic liver procedures. *International Journal of Computer Assisted Radiology and Surgery*, 11(8):1499–1513, 2016.

[85] Kevin Cleary and Terry M. Peters. Image-Guided Interventions: Technology Review and Clinical Applications. *Annual Review of Biomedical Engineering*, 12(1):119–142, 2010.

[86] Bingxiong Lin, Yu Sun, Xiaoning Qian, Dmitry Goldgof, Richard Gitlin, and Yuncheng You. Video-based 3D reconstruction, laparoscope localization and deformation recovery for abdominal minimally invasive surgery: a survey. *International Journal of Medical Robotics and Computer Assisted Surgery*, 12(2):158–178, 2016.

[87] Nazim Haouchine, Jeremie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stephane Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 199–208. IEEE, 2013.

[88] M. Feuerstein, T. Mussack, S.M. Heining, and N. Navab. Intraoperative Laparoscope Augmentation for Port Placement and Resection Planning in Minimally Invasive Liver Resection. *IEEE Transactions on Medical Imaging*, 27(3):355–369, mar 2008.

[89] Gustavo A. Puerto-Souza, Jeffrey A Cadeddu, and Gian-luca Mariottini. Toward Long-Term and Accurate Augmented-Reality for Monocular Endoscopic Videos. *IEEE Transactions on Biomedical Engineering*, 61(10):2609–2620, 2014.

[90] Michael C. Yip, David G. Lowe, Septimiu E. Salcudean, Robert N. Rohling, and Christopher Y. Nguan. Tissue Tracking and Registration for Image-

Guided Surgery. *IEEE Transactions on Medical Imaging*, 31(11):2169–2182, 2012.

[91] Danail Stoyanov and Guang-Zhong Yang. Soft tissue deformation tracking for robotic assisted minimally invasive surgery. *IEEE Engineering in Medicine and Biology Society*, pages 254–7, 2009.

[92] Daniel Reichard, Dominik Häntsch, Sebastian Bodenstedt, Stefan Suwelack, Martin Wagner, Hannes Kenngott, Beat Müller-Stich, Lena Maier-Hein, Rüdiger Dillmann, and Stefanie Speidel. Projective biomechanical depth matching for soft tissue registration in laparoscopic surgery. *International Journal of Computer Assisted Radiology and Surgery*, 12(7):1101–1110, 2017.

[93] Christoph J. Paulus, Nazim Haouchine, David Cazier, and Stéphane Cotin. Surgical augmented reality with topological changes. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 413–420, 2015.

[94] Christoph J Paulus, Nazim Haouchine, David Cazier, and Stephane Cotin. Augmented Reality during Cutting and Tearing of Deformable Objects. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 54–59. IEEE, 2015.

[95] Hadrien Courtecuisse, Jérémie Allard, Pierre Kerfriden, Stéphane P.A. Bordas, Stéphane Cotin, and Christian Duriez. Real-time simulation of contact and cutting of heterogeneous soft-tissues. *Medical Image Analysis*, 18(2):394–410, 2014.

[96] Danail Stoyanov. Surgical Vision. *Annals of Biomedical Engineering*, 40(2):332–345, 2012.

[97] L. Maier-Hein, P. Mountney, A. Bartoli, H. Elhawary, D. Elson, A. Groch, A. Kolb, M. Rodrigues, J. Sorger, S. Speidel, and D. Stoyanov. Optical

techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. *Medical Image Analysis*, 17, 2013.

[98] T. Gauss, P. Merckx, C. Brasher, J. Kavafyan, E. Le Bihan, B. Aussilhou, J. Belghiti, and J. Mantz. Deviation from a preoperative surgical and anaesthetic care plan is associated with increased risk of adverse intraoperative events in major abdominal surgery. *Langenbeck's Archives of Surgery*, 398(2):277–285, 2013.

[99] M. Scheuering, A. Schenk, A. Schneider, B. Preim, and G. Greiner. Intraoperative augmented reality for minimally invasive liver interventions. In *SPIE Medical Imaging*, 2002.

[100] Logan W. Clements, Prashanth Dumpuri, William C. Chapman, Robert L. Galloway, and Michael I. Miga. Atlas-based method for model updating in image-guided liver surgery. *SPIE*, 6509:6509 – 6509 – 12, 2007.

[101] David A Simon. *Fast and accurate shape-based registration*. PhD thesis, Carnegie Mellon University Pittsburgh, 1996.

[102] Burton Ma and Randy E. Ellis. Spatial-stiffness analysis of surface-based registration. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 623–630, 2004.

[103] Burton Ma and Randy E. Ellis. Unified point selection and surface-based registration using a particle filter. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 75–82, 2005.

[104] Yipeng Hu, Ester Bonmati, Eli Gibson, John H Hipwell, David J Hawkes, Steven Bandula, Stephen P Pereira, and Dean C Barratt. 2D-3D Registration Accuracy Estimation for Optimised Planning of Image-Guided Pancreatobiliary Interventions. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, 9900:516–524, 2011.

[105] Kok-Lim Low and Anselmo Lastra. Predetermination of ICP registration errors and its application to view planning. In *IEEE 3-D Digital Imaging and Modeling, 3DIM'07*, pages 73–80, 2007.

[106] Natasha Gelfand, Leslie Ikemoto, Szymon Rusinkiewicz, and Marc Levoy. Geometrically stable sampling for the ICP algorithm. In *IEEE 3-D Digital Imaging and Modeling, 3DIM*, pages 260–267, 2003.

[107] Manning Wang and Zhijian Song. Optimal number and distribution of points selected on the vertebra for surface matching in CT-based spinal navigation. *Computer Aided Surgery*, 18(3-4):93–100, 2013.

[108] Manning Wang and Zhijian Song. How does adding anatomical landmarks as fiducial points in the point-matching registration of neuronavigation influence registration accuracy? *Computer Assisted Surgery*, 21(1):39–45, 2016.

[109] Julie Hallet, Brice Gayet, Allan Tsung, Go Wakabayashi, and Patrick Pessaux. Systematic review of the use of pre-operative simulation and navigation for hepatectomy: Current status and future perspectives. *Journal of Hepato-Biliary-Pancreatic Sciences*, 22(5):353–362, 2015.

[110] SA Nicolau, Xavier Pennec, Luc Soler, X Buy, Afshin Gangi, Nicholas Ayache, and Jacques Marescaux. An augmented reality system for liver thermal ablation: design and evaluation on clinical cases. *Medical Image Analysis*, 13(3):494–506, 2009.

[111] Dogu Teber, Selcuk Guven, Tobias Simpfendorfer, Mathias Baumhauer, Esref Oguz Guven, Faruk Yencilek, Ali Serdar Gozen, and Jens Rassweiler. Augmented reality: A new tool to improve surgical accuracy during laparoscopic partial nephrectomy? preliminary in vitro and in vivo results. *European Urology*, 56(2):332 – 338, 2009.

[112] Esther Wild, Dogu Teber, Daniel Schmid, Tobias Simpfendörfer, Michael Müller, Ann-Christin Baranski, Hannes Kenngott, Klaus Kopka, and Lena

Maier-Hein. Robust augmented reality guidance with fluorescent markers in laparoscopic surgery. *International Journal of Computer Assisted Radiology and Surgery*, 11(6):899–907, 2016.

[113] Logan W. Clements, William C. Chapman, Benoit M. Dawant, Robert L. Galloway, and Michael I. Miga. Robust surface registration using salient anatomical features for image-guided liver surgery: Algorithm and validation. *Medical Physics*, 35(6):2528, 2008.

[114] David M. Cash, Michael I. Miga, Sean C. Glasgow, Benoit M. Dawant, Logan W. Clements, Zhujiang Cao, Robert L. Galloway, and William C. Chapman. Concepts and Preliminary Data Toward the Realization of Image-guided Liver Surgery. *Journal of Gastrointestinal Surgery*, 11(7):844–859, 2007.

[115] Toby Collins, Daniel Pizarro, Adrien Bartoli, Michel Canis, and Nicolas Bourdel. Computer-Assisted Laparoscopic Myomectomy by Augmenting the Uterus with Pre-Operative MRI Data. *IEEE International Symposium on Mixed and Augmented Reality*, pages 243–248, 2014.

[116] Nazim Haouchine, Frederick Roy, Lionel Untereiner, and Stéphane Cotin. Using Contours as Boundary Conditions for Elastic Registration during Minimally Invasive Hepatic Surgery. *International Conference on Intelligent Robots and Systems*, 2016.

[117] Peter Mountney, Johannes Fallert, Stephane Nicolau, Luc Soler, and Philip W. Mewes. An augmented reality framework for soft tissue surgery. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pages 423–431, 2014.

[118] Sylvain Bernhardt, Stéphane A Nicolau, Adrien Bartoli, Vincent Agnus, Luc Soler, and Cristophe Doignon. Using Shading to Register an Intraoperative CT Scan to a Laparoscopic Image. *CARE: Computer-Assisted and Robotic Endoscopy*, 2:59–68, 2016.

[119] Marco Feuerstein, Thomas Mussack, Sandro M Heining, Nassir Navab, and Klinikum Der L M U Innenstadt. Registration-Free Laparoscope Augmentation for Intra-Operative Liver Resection Planning. *SPIE Medical Imaging*, 2007.

[120] D.M. Cash, M.I. Miga, T.K. Sinha, R.L. Galloway, and W.C. Chapman. Compensating for intraoperative soft-tissue deformations using incomplete surface data and finite elements. *IEEE Transactions on Medical Imaging*, 24(11):1479–1491, 2005.

[121] Li-Ming Su, Balazs P. Vagvolgyi, Rahul Agarwal, Carol E. Reiley, Russell H. Taylor, and Gregory D. Hager. Augmented Reality During Robot-assisted Laparoscopic Partial Nephrectomy: Toward Real-Time 3D-CT to Stereoscopic Video Registration. *Urology*, 73(4):896 – 900, 2009.

[122] Mark A. Clifford, Filip Banovac, Elliot Levy, and Kevin Cleary. Assessment of hepatic motion secondary to respiration for computer assisted interventions. *Computer Aided Surgery*, 7(5):291–299, 2002.

[123] Alexander F Kolen, Naomi R Miller, Eltayeb E Ahmed, and Jeffrey C Bamber. Characterization of cardiovascular liver motion for the eventual application of elasticity imaging to the liver in vivo. *Physics in Medicine and Biology*, 49(18):4187, 2004.

[124] Jon S Heiselman, Logan W Clements, Jarrod A Collins, Jared A Weis, Amber L Simpson, Sunil K Geevarghese, T Peter Kingham, William R Jarnagin, and Michael I Miga. Characterization and correction of soft tissue deformation in laparoscopic image-guided liver surgery. *Journal of Medical Imaging*, In Press(2), 2017.

[125] S. F. Johnsen, S. Thompson, M. J. Clarkson, M. Modat, Y. Song, J. Totz, K. Gurusamy, B. Davidson, Z. A. Taylor, D. J. Hawkes, and S. Ourselin. *Database-Based Estimation of Liver Deformation under Pneumoperitoneum for Surgical Image-Guidance and Simulation*, pages 450–458. 2015.

[126] J. Bano, A. Hostettler, S. A. Nicolau, C. Doignon, H. S. Wu, M. H. Huang, L. Soler, and J. Marescaux. Simulation of the Abdominal Wall and Its Arteries after Pneumoperitoneum for Guidance of Port Positioning in Laparoscopic Surgery. In *Advances in Visual Computing*, pages 1–11, 2012.

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

[128] Max Allan, Ankur Kapoor, Philip Mewes, and Peter Mountney. Non Rigid Registration of 3D Images to Laparoscopic Video for Image Guided Surgery. In *CARE: Computer-Assisted and Robotic Endoscopy*, volume 8899, pages 109–116, 2016.

[129] Prashanth Dumpuri, Logan W. Clements, Benoit M. Dawant, and Michael I. Miga. Model-updated image-guided liver surgery: Preliminary results using surface characterization. *Progress in Biophysics and Molecular Biology*, 103(2-3):197–207, 2010.

[130] Nazim Haouchine, Stephane Cotin, Igor Peterlik, Jeremie Dequidt, Mario Sanz Lopez, Erwan Kerrien, and Marie-Odile Berger. Impact of Soft Tissue Heterogeneity on Augmented Reality for Liver Surgery. *IEEE Transactions on Visualization and Computer Graphics*, 21(5):584–597, 2015.

[131] Nazim Haouchine, Jeremie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stephane Cotin. Towards an accurate tracking of liver tumors for augmented reality in robotic assisted surgery. In *IEEE International Conference on Robotics and Automation*, pages 4121–4126, 2014.

[132] Rosalie Plantefève, Igor Peterlik, and Stéphane Cotin. Intraoperative Biomechanical Registration of the Liver: Does the Heterogeneity of the Liver Matter? *IRBM*, 1:1–8, 2017.

[133] Rosalie Plantefève. *Augmented Reality and Numerical Simulations for Hepatic Tumors Resection*. PhD thesis, 2016.

[134] Yinoussa Adagolodjo, Nicolas Golse, Eric Vibert, Michel De Mathelin, Stéphane Cotin, and Hadrien Courtecuisse. Marker-based Registration for Large Deformations - Application to Open Liver Surgery. In *International Conference on Robotics and Automation*, pages 1–7, 2018.

[135] Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. A guide to deep learning in healthcare. *Nature Medicine*, 25(1):24–29, 2019.

[136] Fouzia Altaf, Syed M. S. Islam, Naveed Akhtar, and Naeem K. Janjua. Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions. *IEEE Access*, PP(3):1–28, 2019.

[137] Grant Haskins, Uwe Kruger, and Pingkun Yan. Deep Learning in Medical Image Registration: A Survey. 2019.

[138] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

[139] Jean-Nicolas Brunet, Andrea Mendizabal, Antoine Petit, Nicolas Golse, Eric Vibert, and Stephane Cotin. Physics-based Deep Neural Network for Augmented Reality during Liver Surgery. *Medical Image Computing and Computer-Assisted - MICCAI*, 2019.

[140] Micha Pfeiffer, Carina Riediger, Jürgen Weitz, and Stefanie Speidel. Learning soft tissue behavior of organs for surgical navigation with convolutional neural networks. *International Journal of Computer Assisted Radiology and Surgery*, 14(7):1147–1155, 2019.

[141] Imanol Luengo, Evangello Flouty, Petros Giataganas, Piyamate Wisanuvej, Jean Nehme, and Danail Stoyanov. SurReal : enhancing Surgical simulation Realism using style transfer. pages 1–12, 2018.

[142] Mohammad Abu Hilal, Luca Aldrighetti, Ibrahim Dagher, Bjorn Edwin, Roberto Ivan Troisi, Ruslan Alikhanov, Somaiah Aroori, Giulio Belli, Marc Besselink, Javier Briceno, Brice Gayet, Mathieu D'Hondt, Mickael Lesurtel, Krishna Menon, Peter Lodge, Fernando Rotellar, Julio Santoyo, Olivier Scatton, Olivier Soubrane, Robert Sutcliffe, Ronald Van Dam, Steve White, Mark Christopher Halls, Federica Cipriani, Marcel Van Der Poel, Ruben Ciria, Leonid Barkhatov, Yrene Gomez-Luque, Sira Ocana-Garcia, Andrew Cook, Joseph Buell, Pierre Alain Clavien, Christos Dervenis, Giuseppe Fusai, David Geller, Hauke Lang, John Primrose, Mark Taylor, Thomas Van Gulik, Go Wakabayashi, Horacio Asbun, and Daniel Cherqui. The Southampton Consensus Guidelines for Laparoscopic Liver Surgery: From Indication to Implementation. *Annals of Surgery*, 268(1):11–18, 2018.

[143] J Michael Fitzpatrick. Fiducial registration error and target registration error are uncorrelated. *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, 7261:726102, 2009.

[144] Marco A Pereira-Sampaio, Luciano A Favorito, and Francisco JB Sampaio. Pig kidney: anatomical relationships between the intrarenal arteries and the kidney collecting system. applied study for urological research and surgical training. *The Journal of Urology*, 172(5, Part 1):2077 – 2081, 2004.

[145] A J Herline, J D Stefansic, J P Debelak, S L Hartmann, C W Pinson, R L Galloway, and W C Chapman. Image-guided surgery: preliminary feasibility studies of frameless stereotactic liver surgery. *Archives of Surgery*, 134(6):644–9; discussion 649–50, 1999.

[146] Cameron Shelley. Why test animals to treat humans? on the validity of animal models. *Studies in History and Philosophy of Science Part C: Studies*

in History and Philosophy of Biological and Biomedical Sciences, 41(3):292 – 299, 2010.

[147] Richard Modrzejewski, Jacques Marescaux, Toby Collins, Barbara Seeliger, Adrien Bartoli, and Alexandre Hostettler. 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, 2019.

[148] E Lee Brewer, Logan W Clements, Jarrod A Collins, Derek J Doss, Jon S Heiselman, Michael I Miga, Chris D Pavas, and Edward H Wisdom. The Image-to-Physical Liver Registration Sparse Data Challenge. In *SPIE Medical Imaging*, 2018.

[149] Priyanka Phutane, Emmanuel Buc, Karine Poirot, Erol Ozgur, Denis Pezet, Adrien Bartoli, and Bertrand Le Roy. Preliminary trial of augmented reality performed on a laparoscopic left hepatectomy. *Surgical Endoscopy and Other Interventional Techniques*, 32(1):514–515, 2018.

[150] B Le Roy, E Ozgur, B Koo, E Buc, and A Bartoli. Augmented reality guidance in laparoscopic hepatectomy with deformable semi-automatic computed tomography alignment (with video). *Journal of Visceral Surgery*, 156(3):261–262, 2019.

[151] Tomoyoshi Okamoto, Shinji Onda, Katsuhiko Yanaga, Naoki Suzuki, and Asaki Hattori. Clinical application of navigation surgery using augmented reality in the abdominal field. *Surgery Today*, 45(4):397–406, 2015.

[152] Sara Condino, Marina Carbone, Roberta Piazza, Mauro Ferrari, and Vincenzo Ferrari. Perceptual Limits of Optical See-Through Visors for Augmented Reality Guidance of Manual Tasks. *IEEE Transactions on Biomedical Engineering*, 9294:1–1, 2019.

[153] Rosalie Plantefeve, Jean-pierre Radoux, and Stephane Cotin. Atlas-based transfer of boundary conditions for biomechanical simulation. *Medical Image Computing and Computer-Assisted - MICCAI*, pages 33–40, 2014.

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

[155] Gabriele Costante, Christian Forster, Jeffrey Delmerico, Paolo Valigi, and Davide Scaramuzza. Perception-aware Path Planning. *IEEE Transactions on Robotics and Automation*, pages 1–16, 2016.

[156] Tal Arbel and Frank P Ferrie. Entropy-based gaze planning. *Image and vision computing*, 19(11):779–786, 2001.

[157] Weihua Sheng, Ning Xi, Mumin Song, and Yifan Chen. Graph-based surface merging in cad-guided dimensional inspection of automotive parts. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 3127–3132, 2001.

[158] William R. Scott, Gerhard Roth, and Jean-François Rivest. View planning for automated three-dimensional object reconstruction and inspection. *ACM Computing Surveys*, 35(1):64–96, 2003.

[159] J. Berkmann and T. Caelli. Computation of surface geometry and segmentation using covariance techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1114–1116, 1994.

[160] Gaël Guennebaud and Benoît Jacob. Eigen v3. `http://eigen.tuxfamily.org`, 2010.

[161] Radu Bogdan Rusu and Steve Cousins. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation*, 2011.

[162] Dan Raviv, Anastasia Dubrovina, and Ron Kimmel. Hierarchical matching of non-rigid shapes. In *Scale Space and Variational Methods in Computer Vision*, pages 604–615, 2012.

[163] Marius Leordeanu and Martial Hebert. A spectral technique for correspondence problems using pairwise constraints. In *International Conference on Computer Vision*, volume 2, pages 1482–1489, 2005.

[164] Qi-Xing Huang, Bart Adams, Martin Wicke, and Leonidas J. Guibas. Non-Rigid Registration Under Isometric Deformations. *Comput. Graphics Forum*, 27(5):1449–1457, 2008.

[165] Gabriel Peyré. The numerical tours of signal processing. *Advanced Computational Signal and Image Processing,Comput. Sci. Eng.*, 13(4):94–97, 2011.

[166] Jiaqi Yang, Qian Zhang, Yang Xiao, and Zhiguo Cao. TOLDI: An effective and robust approach for 3D local shape description. *Pattern Recognition*, 65:175 – 187, 2017.

[167] J A Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996.

[168] Alec Jacobson et al. gptoolbox: Geometry processing toolbox, 2015.

[169] K. Crane, C. Weischedel, and M. Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Transactions on Graphics*, 32(5):152:1–152:11, 2013.

[170] John D'Errico. Surface fitting using gridfit, 2016.

[171] Paolo Cignoni, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. MeshLab: an Open-Source Mesh Processing Tool. In *Eurographics Italian Chapter Conference*, 2008.

[172] Quan Wen, Feng Xu, Ming Lu, and Yong Jun-Hai. Real-time 3D Eyelids Tracking from Semantic Edges. In *ACM SIGGRAPH Asia*, 2017.

[173] Ken C.L. Wong, Mehdi Moradi, Hui Tang, and Tanveer Syeda-Mahmood. 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes. *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, 2018.

[174] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *Journal of Physics E: Scientific Instruments*, 6(9):925–929, 2016.

[175] Marie-Pierre Dubuisson and Anil K. Jain. A modified hausdorff distance for object matching. *International Conference on Pattern Recognition*, 1:566–568 vol.1, 1994.

[176] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[177] David C. Blair. Information retrieval, 2nd ed. c.j. van rijsbergen. london: Butterworths; 1979: 208 pp. price: $32.50. *Journal of the American Society for Information Science*, 30(6):374–375, 1979.

[178] David Acuna, Amlan Kar, and Sanja Fidler. Devil is in the Edges : Learning Semantic Boundaries from Noisy Annotations. 2019.

[179] Yun Liu, Ming-Ming Cheng, JiaWang Bian, Le Zhang, Peng-Tao Jiang, and Yang Cao. Semantic Edge Detection with Diverse Deep Supervision. 2018.