

# Registration of Untracked 2D Laparoscopic Ultrasound to CT Images of the Liver using Multi-Labelled Content-Based Image Retrieval

João Ramalinho\*, Henry F. J. Tregidgo, Kurinchi Gurusamy, David J. Hawkes, Brian Davidson and Matthew J. Clarkson

**Abstract**—Laparoscopic Ultrasound (LUS) is recommended as a standard-of-care when performing laparoscopic liver resections as it images sub-surface structures such as tumours and major vessels. Given that LUS probes are difficult to handle and some tumours are iso-echoic, registration of LUS images to a pre-operative CT has been proposed as an image-guidance method. This registration problem is particularly challenging due to the small field of view of LUS, and usually depends on both a manual initialisation and tracking to compose a volume, hindering clinical translation. In this paper, we extend a proposed registration approach using Content-Based Image Retrieval (CBIR), removing the requirement for tracking or manual initialisation. Pre-operatively, a set of possible LUS planes is simulated from CT and a descriptor generated for each image. Then, a Bayesian framework is employed to estimate the most likely sequence of CT simulations that matches a series of LUS images. We extend our CBIR formulation to use multiple labelled objects and constrain the registration by separating liver vessels into portal vein and hepatic vein branches. The value of this new labeled approach is demonstrated in retrospective data from 5 patients. Results show that, by including a series of 5 untracked images in time, a single LUS image can be registered with accuracies ranging from 5.7 to 16.4 mm with a success rate of 78%. Initialisation of the LUS to CT registration with the proposed framework could potentially enable the clinical translation of these image fusion techniques.

**Index Terms**—Multi-modal Registration, Trackerless Registration, Laparoscopic Ultrasound, Content-Based Image Retrieval

## I. INTRODUCTION

**L**APAROSCOPIC Liver Resection (LLR) shows benefits over open surgery in terms of reduced trauma to the patient and consequently shorter hospital stays [1], [2]. However, due to limitations inherent to the laparoscopic setting, only 5-30% of cases are considered for LLR, usually when the tumours are located in easily accessible regions without major vessels [3]. Laparoscopic Ultrasound (LUS) is an imaging tool that can increase the safety of this procedure by imaging sub-surface structures such as vessels and tumours [4]. To enable

use during laparoscopic procedures, LUS probes typically have a relatively small ultrasound transducer, attached to the end of a long shaft, with either a 1 or 2 way pivoting mechanism. This design, combined with the reduced field of view of the transducer, increases the user expertise required to both manipulate the probe inside the abdominal cavity and interpret the resulting images [5]. Additionally, some tumours are iso-echoic and not easy to visualise in the LUS images [6]. To overcome these limitations, registration between LUS and a pre-operative scan such as Computed Tomography (CT) or Magnetic Resonance (MR) based on blood vessel information has been proposed [7]. By aligning LUS with any of these modalities, guidance is enabled by providing the surgeon with spatial context on the relative position between a target tumour and major vessels.

Compared to other ultrasound (US) to CT registration problems, the LUS to CT registration is very poorly constrained due to the difference in imaging field of view and usually depends on either a manual interaction with the images [8] or Electromagnetic (EM) tracking to compose a 3D volume [9], [10]. Such requirements hinder the clinical translation of these methods as they disrupt surgical workflow and increase the hardware complexity in the operating room. In this paper, we extend a novel registration method that provides an accurate initialisation to the problem without requiring tracking information nor a manual interaction with the images. Preliminary results have been previously presented [11].

### A. Background

Several authors have tackled the registration of percutaneous US to CT/MR of the liver, mainly for radiofrequency ablation procedures. Registration for freehand US has been achieved by using an intensity matching between US and a linear combination of CT and simulated US [12], by aligning vessel probability maps derived from US and CT [13] and by matching local orientation description maps of US and MR [14]. However, these intensity-based methods rely on US images that capture large abdominal sections with major liver vessels and liver surface, and not just a small subset of vessels as in LUS. Other authors registered with 3D US probes instead, using vessels as features [15], [16], vessels and liver surface [17], [18], or image intensity gradient information [19], [20]. Since there are no 3D LUS probes, these methods are not easily transferable to the laparoscopic setting.

This work has been supported by the Wellcome/EPSRC grant [203145Z/16/Z] and the NIHR i4i grant [II-LA-1116-20005].

\*J. Ramalinho is with the Wellcome EPSRC Centre for Interventional and Surgical Sciences at University College London, UK (e-mail: joao.ramalinho.15@ucl.ac.uk).

H. F. J. Tregidgo is with the Centre for Medical Image Computing at University College London, UK

B. Davidson and K. Gurusamy are with the Division of Surgery and Interventional Sciences at University College London, UK

M. J. Clarkson and D. J. Hawkes are with the Wellcome EPSRC Centre for Interventional and Surgical Sciences at University College London, UK

Few methods have been demonstrated for the specific registration of LUS to CT. Historically, the first feasibility tests were presented by Bao *et al.* [21] in an isolated phantom and Kruecker *et al.* [22] in a complete laparoscopic setup. Later, Martens *et al.* [23] validated a surface-based rigid registration with the LapAssistent system on an *ex-vivo* animal liver. The first intra-operative solution was demonstrated by Song *et al.* for *in-vivo* animal data, using a locally rigid vessel-based manual alignment [8]. Even though the rigidity assumption is not compatible with the fact that the liver is highly deformed during laparoscopy due to abdominal insufflation (*pneumoperitoneum*) and LUS probe contact [24], the main limitation of these methods is the dependence on a manual match between common vessel landmarks in CT and LUS. This can be attributed to the fact that this is a part-to-whole registration problem where a small subset of vasculature must be matched to a much larger and repetitive vessel tree, making traditional optimisation schemes applicable only if the position of the LUS image is closely initialised to the correct vascular region of the CT. Previously, we approached this problem with a rigid solution, but assumed LUS to be EM tracked [10]. Since tracking devices increase the cost and complexity of the hardware present in the operating room, an untracked registration solution would be highly beneficial for clinical translation. The few current solutions that address untracked US registration are not applicable to laparoscopy as they either rely on 3D US probes [25] or on a very strict probe movement during acquisition [26].

We propose a novel framework that poses registration as a Content-Based Image Retrieval (CBIR) problem. These concepts have been previously applied to the registration of endoscopic video images to a pre-operative CT as a means of enabling image-guidance during lung bronchoscopy [27]. Instead of optimising an alignment, we pre-operatively simulate possible LUS probe poses with their respective imaged vessel content and encode the result to a database. Registration is then globally achieved by finding the pose that best represents the vessel content of an input LUS image without the need for an initialisation. Since the problem is ill-posed and multiple non-unique solutions are expected for a single image, we combine the retrieval results obtained by multiple LUS images acquired closely in time in a Bayesian framework. By assuming that images close in time should also be close in translation and rotation, we construct a discrete Hidden Markov Model (HMM) to estimate the most likely sequence of CT simulations that represent the LUS acquisition. We hypothesise that after including a minimum number of images in the optimisation, a unique registration solution can be obtained. This enables the registration problem to be accurately initialised without tracking data. Previously, we presented preliminary results on a limited sample of synthetic and real LUS sequences, demonstrating the feasibility of this framework without addressing the CBIR system performance comprehensively [11]. In this work, we generalise our CBIR system to include multiple labels in the vessel feature encoding which increases registration performance. In the specific case of liver imaging, this is possible by labelling different vessels as branches of the portal vein or branches of the hepatic vein.

We validate the complete CBIR and HMM framework on clinical data on a larger sample of LUS and CT data from 5 patients.

## B. Contributions

In this paper, we propose a novel CBIR method for the global trackerless registration of LUS to CT and introduce the following contributions:

- We generalise our CBIR registration method to include multiple labels, and provide a comparison of performance between unlabelled and labelled CBIR.
- We provide a comprehensive analysis on the accuracy of the method in the registration of a single LUS image.
- We present insights on what are the minimum requirements in terms of retrieval performance and model complexity in order to obtain a reliable registration.

## II. METHODS

Given a sequence of  $N$  untracked LUS images  $\{I_1, \dots, I_N\}$  with corresponding time stamps  $\{t_1, \dots, t_N\}$ , we pose the registration problem as finding the sequence of pre-operatively simulated CT slices  $\{J_1, \dots, J_N\}$  that best represents the vessel content captured in the LUS acquisition. Our framework comprises two steps:

- A CBIR system that retrieves a set of  $K$  possible slices  $\{J_{1i}, \dots, J_{Ki}\}$  that are candidate solutions for the registration of each LUS image  $I_i$ ;
- A discrete HMM optimisation that estimates the most likely sequence of candidates assuming a kinematic prior on the relative pose between each consecutive LUS slice.

### A. Content-Based Image Retrieval Database Assembly

A set of possible registration solutions is generated by densely sampling 2D planes, bounded by the LUS image geometry, from physically accessible positions  $P$  of the segmented CT volume  $J$ . A virtual reference pose is defined for each point  $P$  within a set of evenly distributed points  $P_S$  along the segmented liver surface likely to be visible during surgery. Each pose is defined by simulating the probe shaft placed orthogonally to the liver surface and aligning the imaging field of view with the sagittal plane. From this reference, multiple combinations of rotations  $R_x$ ,  $R_y$  and  $R_z$  around the probe axes are applied to generate new planes parameterised by  $R = [\vec{x}, \vec{y}, \vec{z}]$ . In order to approximately simulate the case where the probe compresses the liver tissue and captures deeper vessels, a translation  $d$  across the depth of the imaging plane is also sampled.

As illustrated in the top section of Fig.1, for each parameter combination of  $P$ ,  $R$  and  $d$ , a vessel section map is then generated. In order to have a lower dimensional feature representation of each of these maps, we take an approach similar to the one of Petrakis *et al.* [28] and encode them in a feature vector  $\mathbf{f}$  where each of the  $M$  captured vessel sections in a single plane is represented by a feature triplet  $\mathbf{f}_i$  comprised of the corresponding 2D centroid position and area. Then, each triplet is labelled by a class  $c$  from a set of

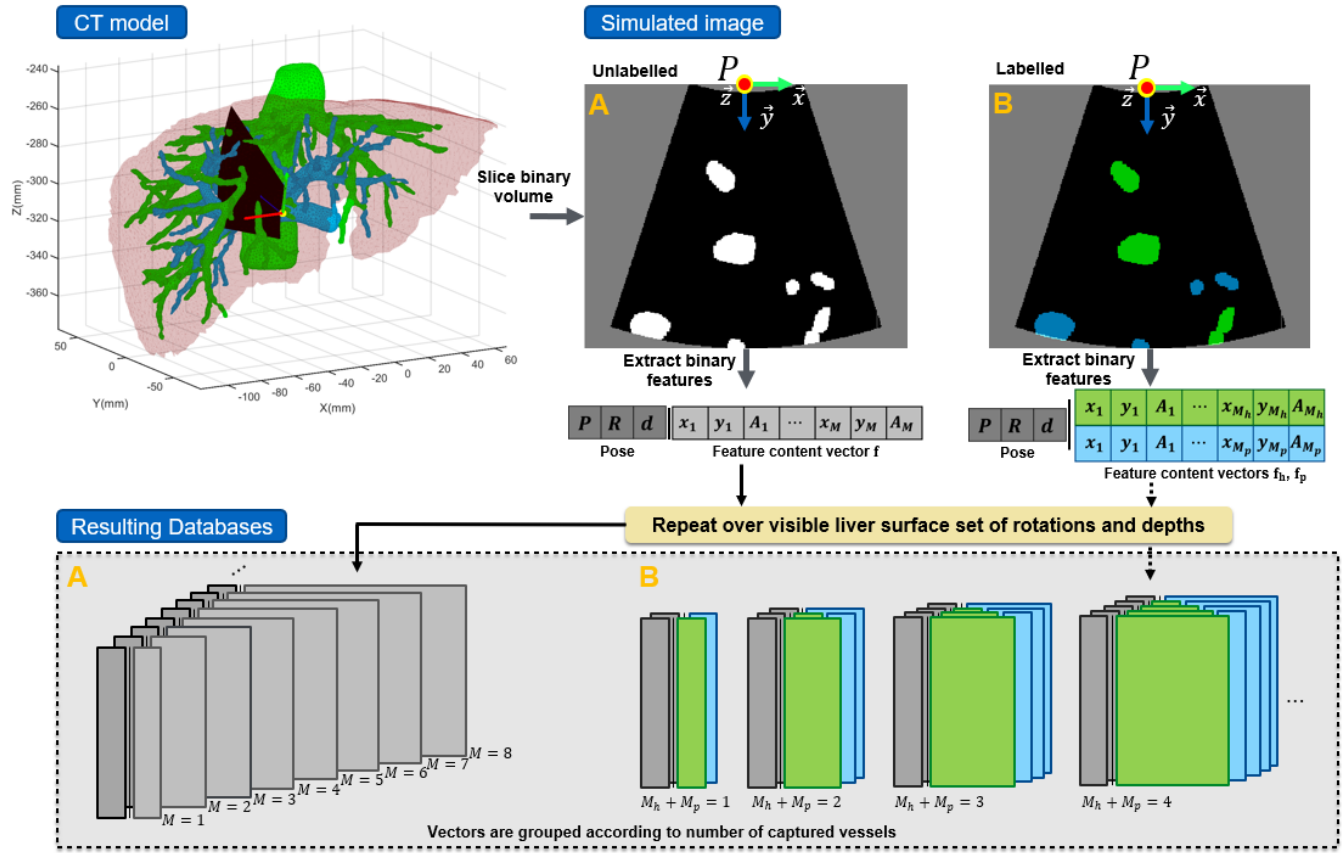


Fig. 1. Database generation process for vessel based image retrieval. For each point  $P$  of the liver surface  $P_S$ , rotation  $R = [\vec{x}, \vec{y}, \vec{z}]$  and translation  $d$  in depth direction, a vessel image is generated. Path (A) illustrates the case where no labels are used and  $\mathbf{f}$  has a single size  $M$ . Path (B) illustrates the case where vessels are labelled as hepatic vein ( $h$ ) and portal vein ( $p$ ) resulting in  $C=2$  and feature vectors  $\mathbf{f}$  have two specific sizes  $M_h$  and  $M_p$ . In the whole figure, green refers to the hepatic vein and blue to the portal vein.

$C$  classes as described in Fig.1 - in (A) no specific labels are considered and  $C=1$ , whereas in (B) portal vein and hepatic vein are considered, resulting in  $C=2$ . Once encoded, vectors  $\mathbf{f}$  are grouped in lists  $F^{M_1, \dots, M_C}$  according to the number  $M_c$  of triplets  $\mathbf{f}_{ci}$  they contain for each class  $c$ . Following the same paths in the bottom section of Fig.1, lists are identified by a single size  $M$  in the unlabelled case (A) and by combinations of two sizes with portal vein and hepatic vein labels (B).

Essentially, a large number of plausible probe poses are sampled, an image plane extracted from CT, and a large lookup table of features extracted, where the lookup table is partitioned according to the number of occurrences of each of the considered vessel classes.

### B. Multi-Labelled Image Retrieval

We formulate the image retrieval process as the computation of a distance measure between an input LUS feature vector  $\mathbf{f}^I$  with  $M^I$  feature triplets and the pre-operatively generated vectors from CT encoded in a database  $F$ . Therefore, we first formalise the problem of measuring the distance between two vectors  $\mathbf{f}^1$  and  $\mathbf{f}^2$ . Taking into account that the size of the vectors may differ, we make the following definitions,

$$\begin{aligned} \mathbf{f}_c^S &:= \arg \min_{\tau \in \{\tau^1, \tau^2\}} (\text{count}(\mathbf{f}, c)), & M_c^S &:= \text{count}(\mathbf{f}_c^S, c), \\ \mathbf{f}_c^L &:= \arg \max_{\tau \in \{\tau^1, \tau^2\}} (\text{count}(\mathbf{f}, c)), & M_c^L &:= \text{count}(\mathbf{f}_c^L, c), \end{aligned} \quad (1)$$

where the subscripts  $L$  and  $S$  refer to the larger and smaller feature vectors in the comparison and  $\text{count}(\mathbf{f}, c)$  operator returns the number of feature triplets in  $\mathbf{f}$  that belong to class  $c$ . We can then define a  $L^2$  distance  $\Delta$  that compares feature vectors of class  $c$  within  $\mathbf{f}^1$  and  $\mathbf{f}^2$ ,

$$\Delta(\mathbf{f}^1, \mathbf{f}^2, c) = \sum_{i=1}^{M_c^S} \|\mathbf{f}_{ci}^S - m(\mathbf{f}_{ci}^S, \mathbf{f}_c^L)\|^2, \quad (2)$$

where the function  $m(\mathbf{f}_{ci}^S, \mathbf{f}_c^L)$  returns the feature triplet in the larger vector  $\mathbf{f}_c^L$  whose centroid is closest to the centroid of  $\mathbf{f}_{ci}^S$ . Intuitively, this function matches all the triplets in the smaller vector  $\mathbf{f}_c^S$  to their closest counterparts in the larger vector  $\mathbf{f}_c^L$ . These class specific costs are subsequently combined to yield the following weighted distance:

$$D(\mathbf{f}^1, \mathbf{f}^2, C) = \frac{A^L}{\sum_{c=1}^C A_c^S} \cdot \sum_{c=1}^C \Delta(\mathbf{f}^1, \mathbf{f}^2, c) \quad (3)$$

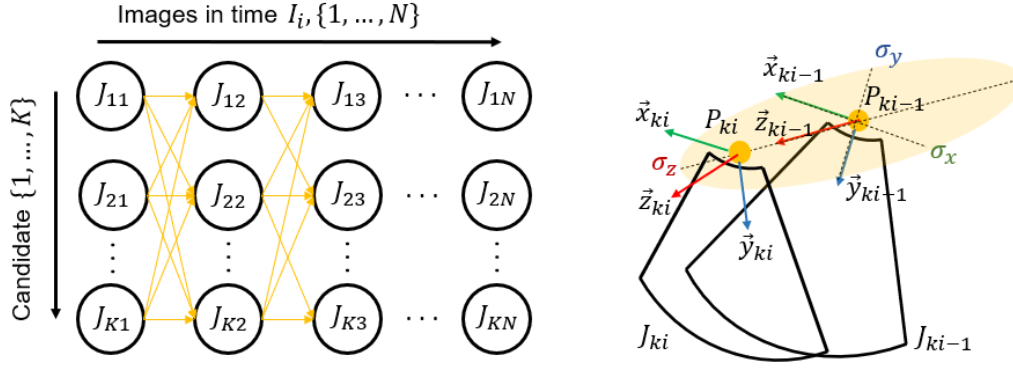


Fig. 2. HMM formulation of the problem. The left-hand side depicts the employed graphical model. The right-hand side illustrates the transition probability  $P(J_{ki}|J_{ki-1})$  of a pose  $J_{ki}$  being followed by  $J_{ki-1}$ . This diagram only displays the translation probability with Gaussian covariances  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$ .

This expression comprises two terms, the sum of  $L^2$  norm for all of the  $C$  classes and a penalty term based on the the area of the vessel sections that were not matched when computing  $\Delta$ . Considering the operator  $A(\mathbf{f}_i)$  that returns the area of triplet  $\mathbf{f}_i$ , and the following definitions,

$$A^L = \sum_{i=1}^{M^L} A(\mathbf{f}_i^L), \quad A_c^S = \sum_{i=1}^{M_c^S} A(m(\mathbf{f}_{ci}^S, \mathbf{f}_c^L)) \quad (4)$$

this penalty term is defined as the ratio between the area  $A^L$  of all triplets in the overall larger vector of the comparison  $\mathbf{f}^L$  and the sum of the areas of the class specific triplets  $\mathbf{f}_{ci}^L$  that were matched in  $\Delta$ . The larger the areas excluded from the matching, the larger the ratio and subsequently the distance  $D$ . In case that  $\mathbf{f}^1$  and  $\mathbf{f}^2$  contain matching numbers of each class, the ratio is 1.

Retrieval is now achieved by searching for the vectors in  $F$  that have a minimal distance  $D$  with the input  $\mathbf{f}^I$ :

$$\mathbf{f}^* = \arg \min_{\mathbf{f}^T \in F^T} \left( \frac{D(\mathbf{f}^I, \mathbf{f}^T, C)}{\min(M^I, M^T)} \right), \quad (5)$$

$$F^T = \{F^{M_1, \dots, M_C} \in F : \sum_{c=1}^C |M_c - M_c^I| \leq r\}$$

Intuitively, this search should only consider lists with a number of vessel sections close to the input. Therefore,  $F^T$  is defined as the target subset of lists whose differences between each of its sizes  $\{M_1, \dots, M_C\}$  and the sizes of the input  $\{M_1^I, \dots, M_C^I\}$  amount to less than an allowable search range  $r$ . In summary, the search only considers lists whose number of vessel occurrences do not differ more than  $r$  from the input. Since the input vector is compared to lists with different sizes, all distances are normalised by the overall smaller number of sections used in each comparison. Assuming that the minimum vector  $\mathbf{f}^*$  may not return an accurate registration, we consider the set of poses providing the  $K$  smallest distances defined in equation (5). These poses are defined as candidate poses for further probabilistic optimisation.

### C. Probabilistic Optimisation

Given the sets of  $K$  candidates  $\{J_{1i}, \dots, J_{Ki}\}$  retrieved for each of the  $N$  LUS images  $I_i$ , we solve the multiple image

registration problem by means of the discrete HMM shown in Fig. 2. In this model where columns and rows refer to images in time and candidates respectively, nodes represent the likelihood  $P(I_i|J_{ki})$  of an image having been acquired with a candidate pose, and edges represent the probability  $P(J_{ki}|J_{ki-1})$  of a pose being followed by another in time. Following discrete HMM theory, the most likely sequence of candidates to represent the acquisition can be obtained through the Maximum A Posteriori (MAP) estimation of the model,

$$\hat{J}_{k1}, \dots, \hat{J}_{kN} = \arg \min_{J_{k1}, \dots, J_{kN}} \left[ - \sum_{i=1}^N \log P(I_i | J_{ki}) - \sum_{i=2}^N \log P(J_{ki} | J_{ki-1}) \right], \quad (6)$$

which can be solved in a closed form by the Viterbi algorithm [29]. In this optimisation, we rely solely on the kinematic prior information in the edges and therefore consider all node probabilities to be 1, giving no priority to any candidate on its own. Assuming the LUS probe to be swept smoothly along the normal to the imaging plane without inverting direction, we define the edge probability similarly to [30], as a multi-variate Gaussian,

$$P(J_{ki}|J_{ki-1}) = \frac{\exp(-\frac{1}{2} \delta_{ki,k_i-1}^T \Sigma^{-1} \delta_{ki,k_i-1})}{\sqrt{2\pi^4 |\Sigma|}} \quad (7)$$

with distance and covariance defined by

$$\delta_{ki,k_i-1} = \begin{bmatrix} R_{ki-1}(\overrightarrow{P_{ki-1}P_{ki}}) \\ \theta_{ki,k_i-1} \end{bmatrix} \quad (8)$$

$$\Sigma = |t_i - t_{i-1}| \text{diag}(\sigma_x, \sigma_y, \sigma_z, \sigma_\theta)$$

As illustrated in the right hand side of Fig. 2, this equation models the transition probability from candidate  $J_{ki-1}$  to  $J_{ki}$  as the distance,  $\delta_{ki,k_i-1}$  between their respective poses conditioned by a Gaussian distribution centred in the pose of  $J_{ki-1}$  with covariance  $\Sigma$ . Four degrees of freedoms are considered in this model, the tridimensional Euclidean difference between the probe contact positions  $P_{ki-1}$  and  $P_{ki}$  projected in the rotation of the starting pose, and the angular difference  $\theta_{ki,k_i-1}$  between the two plane normals  $\vec{z}_{ki}$  and  $\vec{z}_{ki-1}$ . By

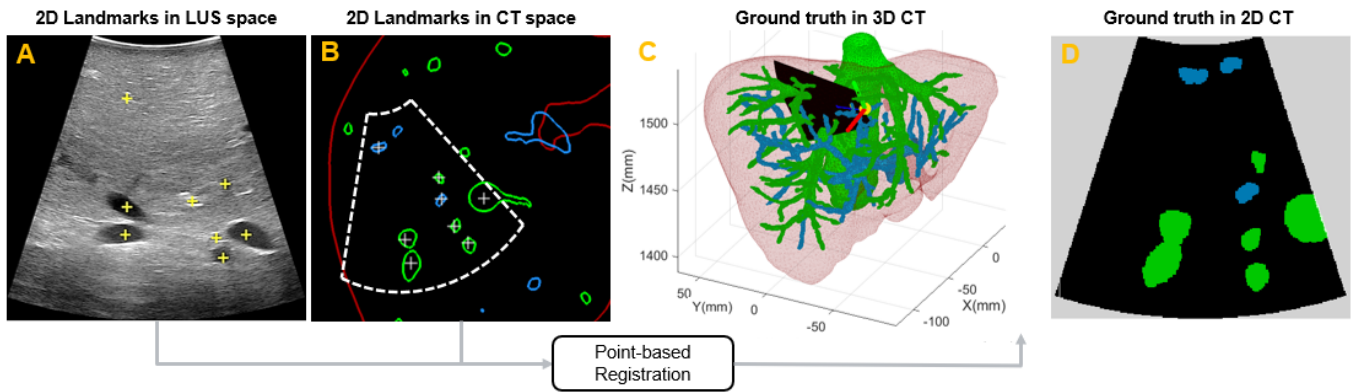


Fig. 3. Manual picking of common vessel landmarks in CT and LUS for posterior accuracy evaluations. By registering the yellow landmarks in LUS (A) and grey landmarks in CT space (B) with a Point-based Registration [31], a ground truth alignment is obtained in 3D (C) and 2D spaces (D). Estimation of the plane represented in (B) is done manually after a careful inspection of the 3D CT models.

defining  $\Sigma$  as a diagonal matrix with values proportional to the time difference between images, this distribution favours transitions where the smaller the time gap, the smaller the translation and rotation difference between them. Additionally, we enforce the forward movement constraint by giving 0 probability to transitions whose direction  $\vec{P}_{ki-1} \vec{P}_{ki}$  have an angular difference above  $90^\circ$  with the direction obtained for the first two images in the sequence.

### III. EXPERIMENTS

We validate the proposed method using retrospective untracked LUS images and contrast-enhanced CT scans from 5 clinical cases. Liver surface, hepatic vein and portal vein models were extracted from each CT using a commercial service ([www.visiblepatient.com](http://www.visiblepatient.com)). LUS images with  $668 \times 544$  pixels and pixel size  $0.12 \text{ mm} \times 0.12 \text{ mm}$  were acquired at a rate of 40 Hz by smoothly sweeping a BK Medical ([www.bkmedical.com](http://www.bkmedical.com)) 4 Way I12C4f laparoscopic transducer over the surface of the right lobe of the liver. Vessel sections are manually segmented and labelled as hepatic vein or portal vein in each LUS image. A future piece of work will consider automating this process, using methods such as the one in [32]. We perform two sets of experiments to assess separately the two components of our registration framework, the CBIR system and the HMM based registration.

#### A. Image Retrieval

In a first experiment, we test the labelled and unlabelled CBIR systems by retrieving  $K = 1000$  candidate poses individually for a set of 63 LUS images distributed among the 5 patients, using an empirically chosen search range  $r = 2$  as in equation (5). Patient-specific unlabelled and labelled databases  $F$  are generated from liver surfaces with an approximate spatial resolution of 3.5 mm, pose rotations within the intervals  $R_x = R_z = [-40, 40]^\circ$ ,  $R_y = [-90, 90]^\circ$  in steps of  $10^\circ$  and depth within the interval  $d = [0, 30]$  mm in steps of 5 mm. A higher amplitude is defined for  $R_y$  as this is the rotation in which the probe can be moved more freely whilst touching the liver surface during acquisition. Given that

the surgeon trivially knows which liver lobe is being scanned during the LUS acquisition, we restrict the translation space of the simulation to the surface of the right lobe.

To measure performance, we adapt the retrieval precision metric that is traditionally used in the evaluation of CBIR systems. Specifically, given an image retrieval task, precision is defined as the percentage of images retrieved that are relevant to an input query image [28]. In the case of registration, an appropriate criterion for relevance should be based on the accuracy of the retrieved poses. Therefore, for each of the 63 tested LUS images, we establish a ground truth alignment by manually picking a set of vessel landmarks with their counterparts in CT, and registering them with a Point-based Registration [31], as illustrated in Fig. 3. For all of the  $K = 1000$  retrieved poses, we project the resulting position of the LUS landmarks (yellow) in CT space, and compute the resulting Target Registration Error (TRE) between them and the CT landmarks (grey). We then define precision as the percentage of poses for which the Root Mean Square (RMS) of this TRE is below 20 mm. This threshold is considered as sufficient for the purpose of a global rigid registration that can be refined further with other US to CT fusion algorithms.

#### B. HMM Registration

We evaluate the HMM optimisation by employing it as a means of registering each of the 63 individual LUS images previously tested for retrieval. Therefore, for every registration task, we consider the image to be registered as  $I_1$  in the optimisation of equation (6), and include the retrieval results of subsequent LUS images that follow in time during acquisition and differ in vascular content.

Firstly, we perform registrations by combining the top 200 retrieved poses for both the image of interest  $I_1$  and the 5 subsequent images  $\{I_2, \dots, I_6\}$ , which yields a HMM with width  $N = 6$  and height  $K = 200$ . Image retrieval for the extra images is also performed using a search range  $r = 2$ , and the HMM probabilistic constraints are parameterised with variances  $\sigma_z = 3 \text{ mm}$ ,  $\sigma_x = \sigma_y = 0.2\sigma_z$  and  $\sigma_\theta = 2^\circ$ . A larger translation variance is given to  $\sigma_z$  in order to prioritise movement along the imaging plane normal, as illustrated in



Fig. 2. Registration accuracy is measured as the TRE of the pose estimated for  $I_1$ .

In a second test, we study the effect of varying the HMM hyperparameters, the width  $N$  and height  $K$ . Therefore, registrations are repeated for all of the 63 LUS images with different numbers of candidate poses  $K$  within  $\{50, 100, 200, 500, 1000\}$ . Instead of considering a fixed width  $N$ , we follow our hypothesis that there is a minimum number of images for which the problem should yield a correct solution, and measure  $N_S$ , the width at which the Viterbi algorithm estimates a pose for  $I_1$  that results in a TRE below the acceptable threshold of 20 mm. This means that for each tested image and value  $K$ , we add retrieval results of subsequent images sequentially until  $I_1$  is successfully registered with a TRE below 20 mm. Retrieval search range  $r$  and probabilistic constraint variances are fixed as in the previous experiment. In practical terms, this experiment aims to provide insight on how many subsequent images and retrieved candidates per image should be used to obtain a reliable registration for a single LUS slice. All experiments are performed for both unlabelled and labelled cases.

### C. Influence of Segmentation Errors

Since we demonstrate our method using manually segmented vessels, we perform a last experiment to evaluate the performance of the CBIR system in the presence of segmentation errors. Traditionally, segmentation errors could be modelled by applying random noise to both the position and outline of the manually segmented vessel sections. However, such errors would not significantly influence the performance

of the retrieval as the system reduces each vessel section to a triplet consisting of a 2D location and area. Additionally, since the database search is constrained by the amount of sections present in the image, the method is most likely affected by topological errors such as mis-segmented vessels. Therefore, we repeat the retrieval experiment of section III-A for all of the 63 LUS images after introducing either fake vessels (False Positives) or missing vessels (False Negatives) in the corresponding manual segmentation.

For each image and corresponding feature vector, we generate augmented vectors with false positives (FP) or false negatives (FN). FN vectors are generated by removing feature triplets whose area does not surpass  $2.5 \text{ mm}^2$ . One FN vector is generated for each possible combination of between one and three removed vessels. FP vectors are generated by adding new feature triplets with centroids adjacent to existing vessels. Each new triplet is created with a fixed area of  $2.5 \text{ mm}^2$  and assigned to the class of the corresponding source vessel. The new centroid is calculated using a Gaussian distribution with a standard deviation of 2.5 mm, and centred on the centroid of the source vessel. One FP vector is generated for each possible combination of between one and three added vessels.

We perform retrieval for all of the resulting FP and FN vectors with mis-segmentations, and group the precision results according to the number of false positives or false negatives in the sampled combination.

## IV. RESULTS

### A. Image Retrieval

To better understand the dimensionality of the solution space considered in the image retrieval task, we first show

TABLE I

CHARACTERISTICS OF SOLUTION SPACES CONSIDERED FOR EACH PATIENT-SPECIFIC DATABASE. LEFT SHOWS DIMENSIONS OF THE SOLUTION SPACE, WHERE #A REFERS TO THE NUMBER OF ELEMENTS IN A. RIGHT SHOWS THE APPROXIMATE MEDIAN AND MAXIMUM TRE OBTAINED ACROSS THE SOLUTION SPACE WITH MEAN AND STANDARD DEVIATION OVER THE NUMBER OF TESTED IMAGES PER PATIENT, LISTED IN THE FIRST COLUMN.

Patient	Dimensions of Solution Space				Expected Errors across Solution Space		
	#Surface Points $P_S$	Surface Area ( $\text{mm}^2$ )	Resolution (mm)	#Feature Vectors in $F$	Number of Images	“Random Retrieval” Median $TRE$ (mm)	Upper Bound Maximum $TRE$ (mm)
Case 1	5025	$3.6 \times 10^4$	4.1	$50.5 \times 10^6$	12	$55.4 \pm 5.6$	$175.1 \pm 19.3$
Case 2	4637	$2.5 \times 10^4$	3.7	$53.7 \times 10^6$	7	$53.9 \pm 2.9$	$110.1 \pm 10.5$
Case 3	3643	$1.6 \times 10^4$	3.3	$47.7 \times 10^6$	16	$50.5 \pm 8.9$	$90.6 \pm 17.0$
Case 4	3467	$1.7 \times 10^4$	3.5	$37.1 \times 10^6$	7	$59.1 \pm 13.0$	$134.5 \pm 21.6$
Case 5	3324	$1.4 \times 10^4$	3.2	$43.4 \times 10^6$	21	$59.8 \pm 4.7$	$107.8 \pm 13.4$

TABLE II

UNLABELLED AND LABELLED IMAGE RETRIEVAL RESULTS FOR 63 LUS IMAGES DISTRIBUTED AMONG 5 PATIENTS USING  $K = 1000$  TOP RETRIEVED POSES AND SEARCH RANGE  $r = 2$  (SEE EQUATION 5). NUMBER OF IMAGES REFERS TO THE NUMBER OF TESTED IMAGES PER PATIENT. OTHER FIELDS ARE PRESENTED AS MEAN AND STANDARD DEVIATION OVER THE NUMBER OF IMAGES PER PATIENT.

Patient	Number of Images	Ground Truth $FRE$ (mm)	Unlabelled Retrieval			Labelled Retrieval		
			Precision (%)	Minimum $TRE$ (mm)	Precision > 0% Images	Precision (%)	Minimum $TRE$ (mm)	Precision > 0% Images
Case 1	12	$7.2 \pm 2.0$	$13.0 \pm 8.4$	$9.7 \pm 1.9$	12/12	$37.6 \pm 24.0$	$9.2 \pm 2.4$	12/12
Case 2	7	$4.2 \pm 2.2$	$20.6 \pm 11.6$	$6.5 \pm 1.6$	7/7	$31.4 \pm 30.3$	$7.5 \pm 2.6$	7/7
Case 3	16	$8.6 \pm 1.1$	$1.7 \pm 2.9$	$17.1 \pm 28.5$	11/16	$12.4 \pm 14.3$	$11.5 \pm 2.6$	16/16
Case 4	7	$8.3 \pm 2.4$	$25.4 \pm 32.3$	$13.0 \pm 5.6$	6/7	$30.1 \pm 27.5$	$11.0 \pm 3.4$	7/7
Case 5	21	$7.2 \pm 2.5$	$12.0 \pm 12.1$	$10.6 \pm 4.6$	20/21	$26.1 \pm 26.2$	$10.1 \pm 5.5$	20/21
<b>Total</b>	<b>63</b>	<b><math>7.3 \pm 2.3</math></b>	<b><math>12.2 \pm 15.9</math></b>	<b><math>12.1 \pm 5.6</math></b>	<b>56/63</b>	<b><math>26.3 \pm 25.6</math></b>	<b><math>10.0 \pm 4.1</math></b>	<b>62/63</b>

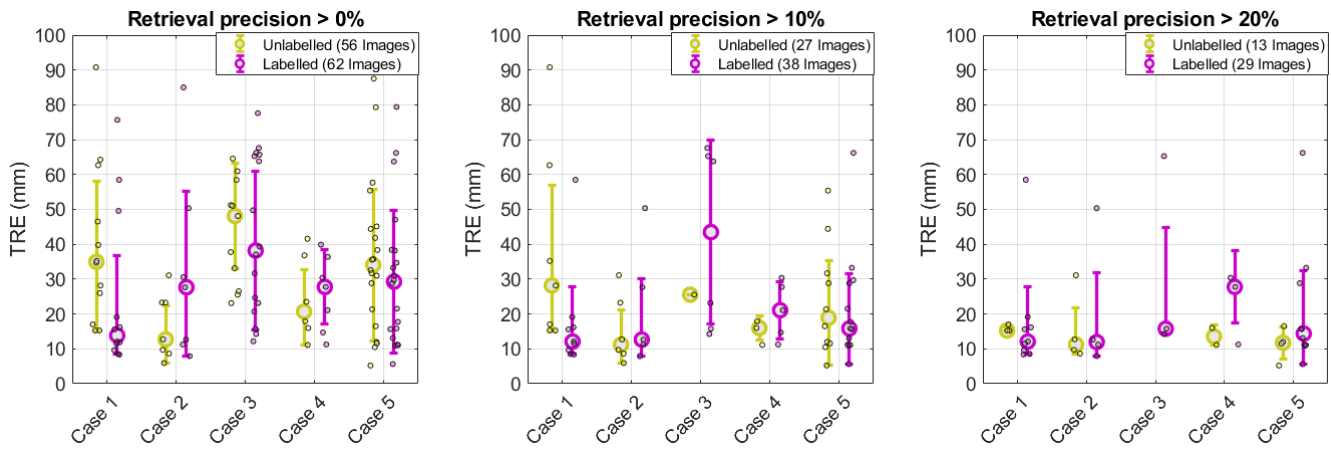


Fig. 4. Registration accuracy on 63 LUS images distributed among 5 patients after employing a HMM registration with 5 subsequent images and number of  $K = 200$  retrieved poses per image. From left to right, results are presented for images with retrieval precision above 0%, 10% and 20%. Median, standard deviation and scatter plots of the accuracies are presented separately for images that belong to the same patient case. Lower deviation limits are presented as the minimum value of the sample. Purple and yellow data series refer to labelled and unlabelled retrieval results, respectively.

an overview of the patient-specific generated databases  $F$  in Table I. In the left-hand side, we present for each database the number of points  $P_S$  that were sampled across the visible surface, the total area covered by them, their spatial resolution, and the total number of generated solutions. In the right-hand side, in order to establish both an upper bound and an expected value of the TRE for a “random retrieval”, we provide an approximation of the maximum TRE and median TRE that can be obtained across all surface points  $P_S$  for each of the 63 tested LUS images. For each patient case, these errors are presented as mean and standard deviation over the number of total tested images per patient listed in the first column of the right section of the table. For the sake of simplicity, these approximations only consider TRE values measured at reference poses ( $R = [0, 0, 0]$ ,  $d = 0$ ).

Retrieval results for the 63 LUS images are presented per patient case in Table II. Similarly to the previous table, results for each case are presented as mean and standard deviation over the number of tested images. For each retrieval method, we present the retrieval precision, the number of images whose precision is above 0%, and the minimum TRE result found in the pool of retrieved poses. In order to set a reference for the TRE values, we use the previously calculated ground truth alignments (Fig. 3) and report the resulting Fiducial Registration Error (FRE). This error represents the maximum accuracy that can be obtained in these landmarks with a rigid registration. Overall, the retrieval systems find a solution with TRE below 20 mm within 1000 retrieved poses for almost every image, except for Case 3, where 5 images have no suitable solutions when the unlabelled method is used. The introduction of vessel labels increases the retrieval precision significantly, as this value ranges from 1.7% to 25.4% for the unlabelled case and ranges from 12.4% to 37.6% for the labelled case. This improvement is consistent for all patient cases and more emphasised in Case 3, where the number of images without a suitable solution decreases from 5 to 0. The obtained minimum TRE values are comparable to the Ground truth FRE and similar between retrieval methods, except for

the unlabelled retrieval of Cases 3 and 4.

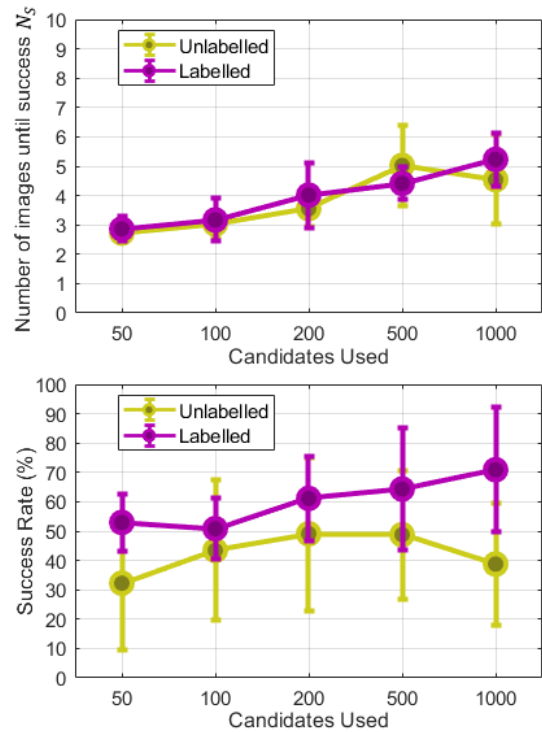


Fig. 5. Number of images  $N_S$  needed to achieve a successful registration of  $I_1$  (top) and corresponding registration success rate (bottom) averaged over 5 patients versus number of candidates  $K$  used in the HMM registration. Each data series represents results with different retrieval methods.

### B. HMM Registration

Registration accuracy results of the first HMM registration experiment are presented per patient and retrieval method in Fig. 4 as a scatter plot with marked median and standard deviations. Results in the left chart show that the median TRE for all images is below the acceptable threshold of 20 mm on only 1 case for each retrieval method. This is possibly

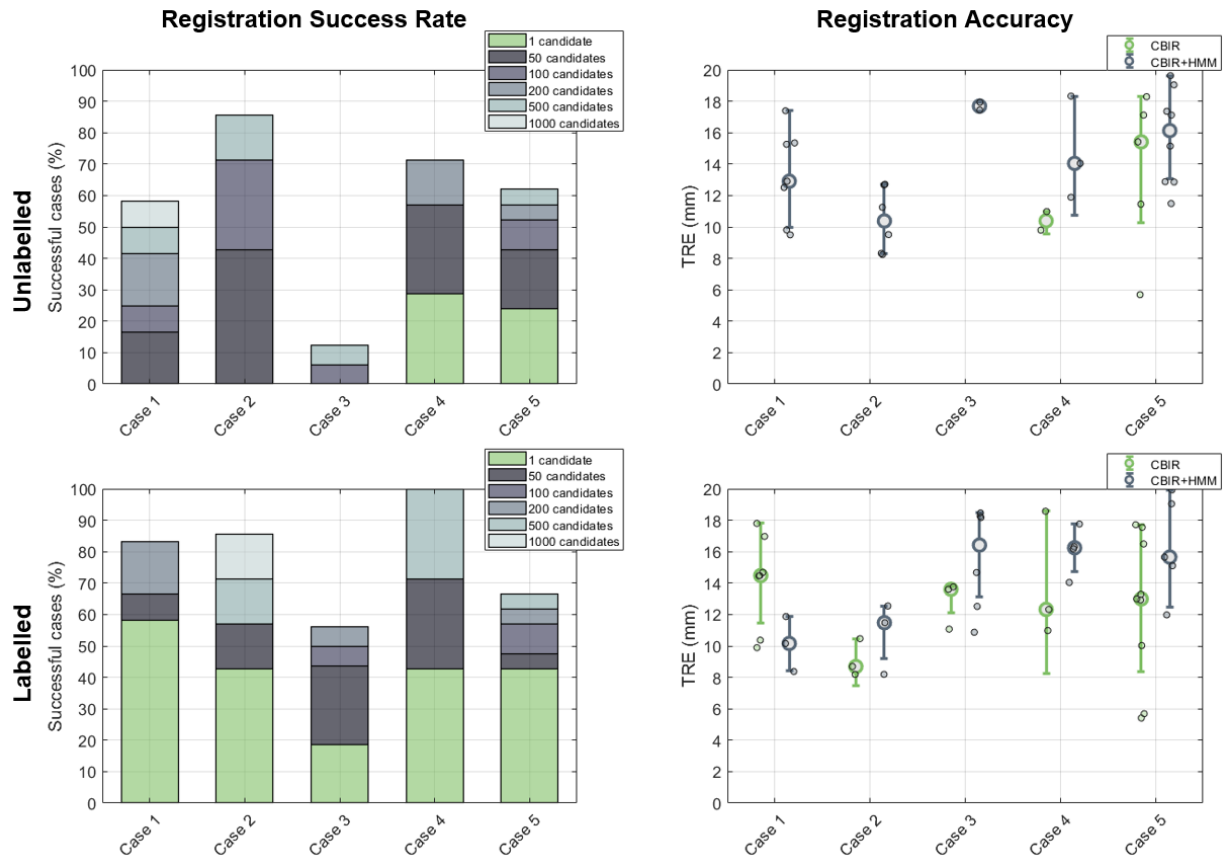


Fig. 6. Distribution of minimum number of candidates  $K$  required for a successful registration of 63 LUS images distributed among 5 patients and corresponding accuracy. Left column charts show these distributions as a set of stacked bars, one for each  $K$  tested value. Bars referring to CBIR only registrations with  $K = 1$  are identified in green, whereas bars referring to HMM registrations are identified with a grey colour map. Right column charts show separate sets of median, standard deviation and scatter plots of the TRE for the CBIR only registrations ( $K = 1$ ) in green and for the HMM registrations in grey.

explained by the fact that images with a lower precision are more difficult to register accurately as the additional LUS images that follow them in time will potentially also have a low retrieval precision. Therefore, we present two additional charts showing the TRE results only for images whose retrieval precision in the previous experiment was above 10% and 20%. These charts show that a higher precision leads to higher registration accuracy, as the median TRE values per patient decrease from left to right, and range from 11.2 mm to 15.7 mm for all cases but one when the precision is above 20%. TRE values do not differ greatly between unlabelled and labelled retrieval methods. However, due to the differences in retrieval precision, the labelled method results in a larger number of registrations with error below 20 mm.

### C. Effect of HMM Width and Height

Two sets of results are presented for the second HMM registration experiment. Firstly, the number of images  $N_S$  required for a successful registration of  $I_1$  with TRE below 20 mm, is presented as a function of the number of used candidates  $K$  in Fig. 5. Additionally, we also show the corresponding percentage of tested images that were successfully registered for each  $K$  value. Since the number of tested images (left column of Table II) varies among patients and each surgical

case has specific physical conditions, we do not consider each image as an independent event, and present these two measurements as the mean and standard deviation over the average of each of the 5 patients. These results indicate that increasing the number of candidates  $K$  potentially increases the success rate of the registration, but also increases the number of images necessary to reach an accurate solution. Specifically, in successful registrations, the mean  $N_S$  ranges from 3 at  $K = 50$  to 5 at  $K = 1000$  for both retrieval approaches. The mean registration success rates are higher when the labelled method is used, with values ranging from 53% at  $K = 50$  to 71% at  $K = 1000$ . For the unlabelled method, the success rates peak at a mean value of 49% at  $K = 200$  and start decreasing with  $K = 1000$ . Such decrease suggests that if  $K$  is too large, the ambiguity of the problem may start increasing.

In a second set of results, instead of looking at the minimum HMM width that leads to an accurate registration of  $I_1$ , we study the minimum HMM height, i.e what is the lowest value of retrieved poses  $K$  required to register each of the tested images. Such analysis is useful as a very large  $K$  may increase not only the ambiguity of the registration in some cases, but also the computational expense of the Viterbi algorithm, whose complexity is quadratic in respect to  $K$ . Results of this analysis



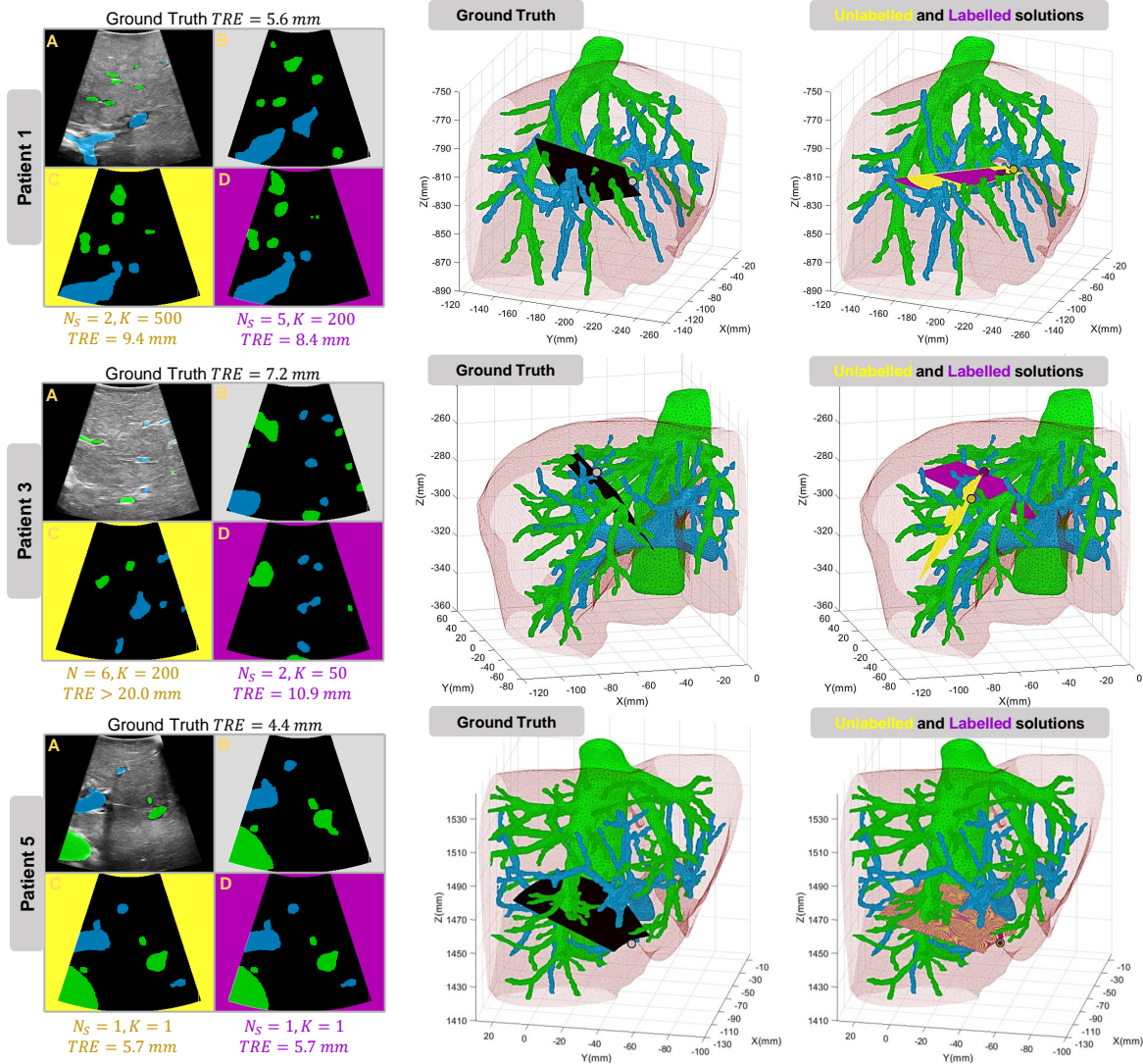


Fig. 7. Visual results of the second HMM registration experiment for 3 LUS images from 3 different patients. Each row refers to a different patient case. Left column shows 2D registration results which include 4 images, the original segmented LUS image (A), and segmented CT projections of the ground truth (B), unlabelled (C) and labelled (D) registration solutions. TRE is displayed for the three solutions, and the minimum number of candidates  $K$  and number of images  $N_S$  required for the registration are displayed for (C) and (D). Middle column shows the 3D position between the resulting plane of the ground truth solution (B) in black and the right column shows the same visualisation for the unlabelled (C) and labelled solutions (D) in yellow and purple, respectively. For an easier interpretation of the plane orientation, the probe contact positions are also highlighted with circular markers. In the whole figure, green refers to the hepatic vein and blue to the portal vein. For an easier interpretation of the 3D results, check the uploaded supplementary material.

are summarised per patient case in Fig. 6. In the left-hand side, we show a stacked bar plot depicting the percentage of tested images per patient whose registration is successful at different minimum  $K$  values. In addition to the  $K$  values tested in the HMM registration, we also consider the cases where the first pose retrieved by the CBIR system (equation (5)) results in a TRE below 20 mm and the minimum  $K$  is 1. The sum of the bars in each stack amounts to the total percentage of tested images for which there is a  $K$  value that results in a registration with TRE below 20 mm. In the right-hand side of Fig. 6, the median, standard deviation and scatter plots of the TRE values of these registrations are presented per patient. CBIR only registrations ( $K = 1$ ) and HMM registrations are shown separately. Visually, CBIR only registration results are highlighted with green whereas HMM registration results are highlighted with a grey colour map.

Overall, the number of candidates required for a successful registration with TRE below 20 mm is lower when using the labelled retrieval - for patient cases 1, 2 and 3, the amount of  $K = 1$  successes significantly increases from 0% to 58%, 43% and 19%, respectively. This improvement is also observed in the percentage of successfully registered images, mainly in the case of patients 1 and 3, where improvements from 58% to 83% and 13% to 56% are observed. In terms of accuracy, the resulting TRE is slightly lower in the labelled retrieval approach and has median values per patient ranging from 8.7 mm in patient 2 to 16.4 mm in patient 3.

Visual examples of these registrations with minimal number of candidates  $K$  are shown in Fig. 7 for 3 LUS images from different patient cases. For each registration example in a row, the left-hand side shows 2D results that include the segmented LUS image (A), the point-based ground truth solution (B),

and the unlabelled (C) and labelled (D) registration solutions. Visualisations of the 3D position between the segmented CT model and resulting imaging planes of the ground truth (black) and the two obtained solutions (yellow and purple) are displayed in the middle and right-hand side of the figure, respectively. To compare our solutions (C and D), in addition to the TRE, we also display the minimum HMM width  $N_S$  and height  $K$  that were required to obtain them. It is important to note that even though in some of these results more than one LUS image was included in the HMM, we focus our attention only in the registration result of the image of interest  $I_1$ .

The registration result of patient 1 shows an example where the accuracy of the labelled and unlabelled methods is comparable, but the minimum number of candidates  $K$  required is smaller for the labelled approach. For patient 3, we show a case where the unlabelled retrieval results never lead to a successful registration with any number of candidates  $K$ , but the labelled method finds a solution for  $K = 50$  candidates and  $N_S = 2$  LUS images. In the registration failure case, we show the solution obtained with the same parameters tested in the first HMM registration experiment (Fig. 4). It is possible to see in this case that the vessel labelling helps the algorithm finding a reliable solution even though there is a large missing portal vein section between the LUS and the ground truth. The registration of patient 5 illustrates the best registration result, where both methods obtain a solution with a TRE of 5.7 mm without using the HMM.

#### D. Effect of Segmentation Errors

Results of labelled retrieval precision averaged over 5 patients versus varying amounts of fake (False Positive) vessels and missing (False Negative) vessels are summarised in Fig. 8. We only show results for the labelled approach since the unlabelled approach showed an overall worse retrieval performance. In this chart, the original mean precision is shown in grey in the centre (see Table II), and the mean precisions in the presence of fake vessels and missing vessels are shown on the bars in the left-hand and right-hand sides, respectively. In the coloured bars, precision is averaged over the median value obtained over all combinations with a fixed number of fake or missing vessels.

As expected, for all cases, the existence of mis-segmented vessels decreases the retrieval precision when compared to the

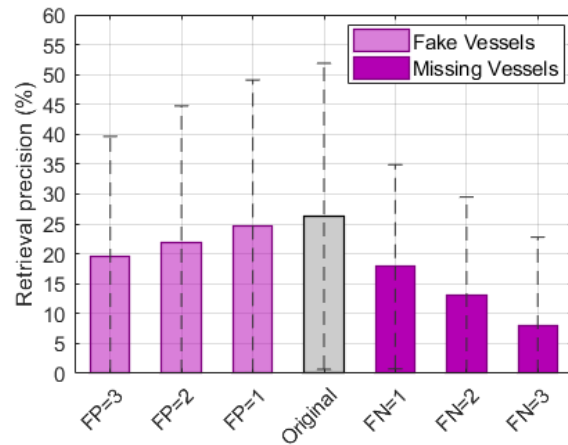


Fig. 8. Mean and standard deviation of labelled retrieval precision of 63 LUS images averaged over 5 patients after introducing varying amounts of fake (False Positive, FP) and missing (False Negative, FN) vessels in the original manual segmentations.

original manual segmentations. However, this effect is more pronounced in the presence of missing vessels, where the mean precision decreases to a value below 20% with only one false negative. In the presence of fake vessels, the precision only reaches a value below this limit when 3 false positives are introduced.

#### E. Computational Expense

An aspect that is crucial to translation of any image-guidance method is the computational requirements. Therefore, we also present a summary of the memory and time expenses of the main components of the proposed method in a set of three tables in Table III. All results were obtained using a machine with an Intel i7 2.8 GHz processor, a NVidia GeForce GTX 1050 graphics card, and a RAM memory of 16 gigabytes (GB).

In Table III. A), the memory required for each of the generated patient-specific databases is presented in GB. In Table III. B), the time expense associated with the simulation and retrieval steps of the CBIR system is presented for both labelled and unlabelled approaches. For simulation, we present the time required for a single image to be simulated and encoded to a feature vector. This step was run using a graphical

TABLE III

THREE TABLES SHOWING APPROXIMATE COMPUTATIONAL TIME AND MEMORY EXPENSE OF DIFFERENT COMPONENTS OF THE METHOD. TIMES ARE PRESENTED IN SECONDS (S) AND MEMORY IN GIGABYTES (GB). IN B), (U) AND (L) REFER TO UNLABELLED AND LABELLED METHODS, RESPECTIVELY. IN C), PER COLUMN REFERS TO ONE ITERATION OF THE HMM OPTIMISATION.

A)	<b>CBIR (Memory)</b>	Case 1 3.9 GB	Case 2 3.9 GB	Case 3 4.0 GB	Case 4 2.7 GB	Case 5 5.7 GB
B)	<b>CBIR (Time)</b>	Simulation (U) 0.02 s per image	Simulation (L) 0.04 s per image	Retrieval (U) 30 s per input image	Retrieval (L) 15 s per input image	
C)	<b>HMM (Time)</b>	$K=50$ 0.5 s per column	$K=100$ 2 s per column	$K=200$ 5 s per column	$K=500$ 75 s per column	$K=1000$ 210 s per column

processing unit (GPU) implementation, and took less than 0.1 seconds for both methods. For retrieval, we present the time required for an input LUS image feature vector to be compared with the database  $F$  using a search range  $r = 2$  (see equation 5). This process was also tested with a GPU implementation, and required 30 seconds for the unlabelled approach and 15 seconds for the labelled one. Such difference can be explained by the database structures - when labels are used, the database is composed of a larger amount of lists with less feature vectors, allowing for a faster search.

Table III. C) shows the time expense of each iteration (computation of a column) of the HMM optimisation versus the number of candidates  $K$  considered in the model. This process was run with a single thread and without GPU acceleration. As expected from the properties of the Viterbi algorithm, these values show an exponential increase in respect to  $K$ .

## V. DISCUSSION

All of the presented results indicate that the proposed registration framework highly benefits from separation of segmented features into specific classes. This is expected since the existence of vessel labels in LUS and CT constrains the registration problem further and helps the CBIR system finding solutions that are more physically plausible. The first results that support this statement were observed in Table II, where retrieval precision significantly improves when labels are used.

The subsequent HMM registration experiment of Fig. 4 demonstrated the importance of retrieval precision in the registration - TRE values below an acceptable threshold of 20 mm were mainly obtained for images whose precision surpassed 20%. Given the higher precision of the labelled approach, it can be concluded that the introduction of labels increases significantly the registration success rate. For the images whose retrieval precision was below 20%, several low accuracy results were obtained. These values may be explained by the fact that the HMM is not robust to the inclusion of an image with poor retrieval precision - if one of the 5 subsequent images did not have any accurate retrieval result, the algorithm is forced to estimate an incorrect alignment.

In the second HMM registration experiment, we investigated the minimum HMM complexity required for the registration to be possible, i.e what is the minimum number of LUS images  $N$  and number of retrieved candidates  $K$  required for a reliable registration of a single LUS image. Compared to the first HMM registration experiment, the results of Fig. 6 indicate that images with a lower retrieval precision can be successfully registered if these parameters are minimised. In this case, the percentage of successfully registered images ranged from 56% to 100% when using labelled retrieval and from 12% to 85% when using unlabelled retrieval. The most striking result of this analysis was the percentage of images for which the CBIR system alone achieved a successful registration - this value was significantly increased with the label constraint. Furthermore, the percentage of images per patient that did not require more than 200 candidates ranged from 53% to 86%. Considering the  $N_S$  measurements of Fig. 5, it is possible to conclude that in most cases, a single LUS image can be registered successfully

if the retrieval results of a maximum of 3 extra images in time are included.

Factors that may explain the poorest retrieval and registration results are intra-operative deformations, the limited pose parameter resolution of the generated database for CBIR and the imaging differences in vessel contrast between LUS and CT. Intra-operative deformations due to insufflation and LUS probe contact are known to significantly compress liver vessels during LUS imaging. In this part-to-whole registration problem, compression can move liver vessels enough so that the field of view of LUS does not capture them anymore but CT does - this is exemplified in the result of patient case 3 in Fig. 7 where the rigid CT ground truth images a large portal vein section that the deformed LUS image does not. In this CBIR system, deformation compensation could be achieved in future work by including deformation parameters in the database simulation step. By simulating vessel content vectors from an insufflated CT model, the formulation of the registration would not require any adaptation.

The database resolution limits the maximum registration accuracy that can be obtained and can hinder the registration performance of images that contain oblique vessel sections. Since simulations assume an infinitely thin plane intersecting the CT vascular model, it is likely that many oblique vessel projections are not properly captured by the limited rotation resolution. This problem could be overcome either by considering the US imaging specific beam width during the simulation, or by increasing the resolution of the parameter space. To avoid an unfeasible increase in the computational expense, it is important to maintain a trade-off between the resolution and size of the covered translation space. For example, a higher resolution database could be generated whilst restricting the translations to a region of interest, such as a spherical domain centred around a target tumour.

The different vessel contrast in LUS and CT leads to missing vessels between both images, affecting directly the problem of matching single 2D images to a 3D volume. Even though we have included an allowable search range  $r$  to account for this problem, substantial topological differences between the LUS vessels and the CT vessels may still compromise the uniqueness of the solutions. This effect is more pronounced when CT and LUS have several mismatching small vessels that are more susceptible to deformation. Therefore, future work could address the pre-operative identification of major vessels that are more likely to be imaged in LUS and guarantee a reliable registration based on them.

Overall, our results are promising for a future clinical application. It has been proposed that image-guidance techniques for LLR should be accurate within 5 mm to have clinical value [33]. Considering this threshold as a target accuracy, the accuracies ranging between 5.7 and 16.4 mm obtained by our method are sufficient as an initialisation for refinement with other registration techniques. To use this initialisation reliably without a tracker, a LUS volume could then be estimated either by separately registering multiple images in time, or by using a single image registration and a freehand US compounding method [34], [35]. Another option would be to estimate the LUS probe position through laparoscopic video

based tracking [36], [37]. However, these methods require a marker to be attached to the LUS probe and a calibration procedure, complicating clinical translation.

In this work, we validated our registration framework with manually segmented vessels. Therefore, to understand segmentation requirements necessary for clinical translation, we tested the performance of the labelled CBIR system in the presence of topological mis-segmentations. Results of this experiment indicate that retrieval performance is more affected by the presence of missing vessels. Such decrease is expected since segmentation failures will reduce the complexity of the input LUS feature vector and possibly increase registration ambiguity, mainly in images with less vessel data. In the case of false positive mis-segmentations, the precision values do not decrease as much, indicating that a future automatic segmentation method should prioritise sensitivity over specificity. For future work, Deep Learning frameworks for segmentation of portal and hepatic veins in liver US images [32] could be integrated.

In terms of computational expense, the proposed approach requires times that are potentially compatible with the clinical workflow. The most time intensive step of the method is the database simulation, as a large number of  $50 \times 10^6$  images is simulated at a rate of 25 images per second. This is not critical for two reasons - firstly, simulation is done pre-operatively without very strict time constraints, and secondly, the expense can be easily reduced by splitting the process across multiple machines using cluster computing. Regarding the intra-operative processes of the method, values in the order of seconds were measured for both image retrieval and the HMM optimisation. Specifically, if we performed registration using  $K = 200$  with  $N = 6$  retrieved LUS images, the resulting time expense would be  $(15 \times 6) + (5 \times 5) = 115$  seconds. Even though these times are not optimal for a real-time application, they can be reduced either by using a higher memory GPU, or by accelerating the HMM column computation with a GPU implementation.

## VI. CONCLUSION

In this paper, we have extended a novel method based on CBIR for the registration of LUS images to CT of the liver without using tracking devices or a manual initialisation, and validated on 63 untracked LUS images. By generalising our CBIR formulation to take into account labels assigned to each extracted feature, the registration performance of the method was improved. Results on clinical data from 5 patients show that if we consider up to 200 retrieval solutions from a sequence of up to 6 LUS images, trackerless registrations with accuracy ranging from 5.7 to 16.4 mm can be obtained in 78% of cases. Registration failures can be explained by the presence of deformations, the differences in the contrast of CT and LUS imaging, and the fact that some liver regions contain vasculature that is not unique enough to constrain the problem. Regardless, our results are promising for the purpose of an initialisation of the LUS to CT registration problem, which is extremely challenging and currently has no solution that is transferable to the clinic. Higher performance can be

achieved in the future by increasing the realism of the CBIR simulation, either by including deformation or simulating more realistic LUS images. To the best of our knowledge, this is the first work in multi-modal image registration of US to a pre-operative scan that makes use of CBIR. The proposed framework has the potential of enabling clinical translation, and could be adapted to other interventional US guidance problems.

## REFERENCES

- [1] S. C. Schiffman, K. H. Kim, A. Tsung, J. W. Marsh, and D. A. Geller, "Laparoscopic Versus Open Liver Resection for Metastatic Colorectal Cancer: a Metaanalysis of 610 Patients," *Surgery*, vol. 157, no. 2, pp. 211–222, 2015.
- [2] R. Ciria, D. Cherqui, D. A. Geller, J. Briceno, and G. Wakabayashi, "Comparative Short-term Benefits of Laparoscopic Liver Resection: 9000 Cases and Climbing," *Ann. Surg.*, vol. 263, no. 4, pp. 761–777, 2016.
- [3] G. Wakabayashi *et al.*, "Recommendations for Laparoscopic Liver Resection: a Report from the Second International Consensus Conference held in Morioka," *Ann. Surg.*, vol. 261, no. 4, pp. 619–629, 2015.
- [4] A. Rethy, T. Langø, and R. Mårvik, "Laparoscopic Ultrasound for Hepatocellular Carcinoma and Colorectal Liver Metastasis: an Overview," *Surg. Laparo. Endo. Per.*, vol. 23, no. 2, pp. 135–144, 2013.
- [5] C. Våpenstad *et al.*, "Laparoscopic Ultrasound: a Survey of its Current and Future Use, Requirements, and Integration with Navigation Technology," *Surg. Endosc.*, vol. 24, no. 12, pp. 2944–2953, 2010.
- [6] M. A. Choti *et al.*, "Patient Variability in Intraoperative Ultrasonographic Characteristics of Colorectal Liver Metastases," *Arch. Surg-Chicago*, vol. 143, no. 1, pp. 29–34, 2008.
- [7] L. Crocetti, R. Lencioni, S. DeBeni, T. C. See, C. Della Pina, and C. Bartolozzi, "Targeting Liver Lesions for Radiofrequency Ablation: an Experimental Feasibility Study Using a CT-US Fusion Imaging System," *Invest. Radiol.*, vol. 43, no. 1, pp. 33–39, 2008.
- [8] Y. Song *et al.*, "Locally Rigid, Vessel-Based Registration for Laparoscopic Liver Surgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 12, pp. 1951–1961, 2015.
- [9] M. Fusaglia, P. Tinguely, V. Banz, S. Weber, and H. Lu, "A Novel Ultrasound-Based Registration for Image-Guided Laparoscopic Liver Ablation," *Surg. Innov.*, vol. 23, no. 4, pp. 397–406, 2016.
- [10] J. Ramalhinho *et al.*, "A Pre-operative Planning Framework for Global Registration of Laparoscopic Ultrasound to CT Images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 8, pp. 1177–1186, 2018.
- [11] J. Ramalhinho *et al.*, "Registration of Untracked 2D Laparoscopic Ultrasound Liver Images to CT Using Content-Based Retrieval and Kinematic Priors," in *Smart Ultrasound Imag. and Perinatal, Preterm and Paediatric Image Anal.* Springer, 2019, pp. 11–19.
- [12] W. Wein, S. Brunke, A. Khamene, M. R. Callstrom, and N. Navab, "Automatic CT-Ultrasound Registration for Diagnostic Imaging and Image-Guided Intervention," *Med. Imag. Anal.*, vol. 12, no. 5, pp. 577–585, 2008.
- [13] G. P. Penney, J. M. Blackall, M. Hamady, T. Sabharwal, A. Adam, and D. J. Hawkes, "Registration of Freehand 3D Ultrasound and Magnetic Resonance Liver Images," *Med. Imag. Anal.*, vol. 8, no. 1, pp. 81–91, 2004.
- [14] M. Yang, H. Ding, J. Kang, L. Cong, L. Zhu, and G. Wang, "Local Structure Orientation Descriptor Based on Intra-Image Similarity for Multimodal Registration of Liver Ultrasound and MR Images," *Comput. Biol. Med.*, vol. 76, pp. 69–79, 2016.
- [15] S. R. Aylward, J. Jomier, J. P. Guyon, and S. Weeks, "Intra-operative 3D Ultrasound Augmentation," in *Proceedings IEEE International Symposium on Biomedical Imaging*. IEEE, 2002, pp. 421–424.
- [16] T. Lange, S. Eulenstein, M. Hünerbein, and P.-M. Schlag, "Vessel-Based Non-Rigid Registration of MR/CT and 3D Ultrasound for Navigation in Liver Surgery," *Comput. Aided Surg.*, vol. 8, no. 5, pp. 228–240, 2003.
- [17] W. H. Nam, D.-G. Kang, D. Lee, J. Y. Lee, and J. B. Ra, "Automatic Registration between 3D Intra-operative Ultrasound and Pre-operative CT Images of the Liver Based on Robust Edge Matching," *Phys. Med. Biol.*, vol. 57, no. 1, p. 69, 2011.
- [18] D. Lee, W. H. Nam, J. Y. Lee, and J. B. Ra, "Non-Rigid Registration Between 3D Ultrasound and CT Images of the Liver Based on Intensity and Gradient Information," *Phys. Med. Biol.*, vol. 56, no. 1, p. 117, 2011.

- [19] T. Lange *et al.*, “3D Ultrasound-CT Registration of the Liver Using Combined Landmark-Intensity Information,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 4, no. 1, pp. 79–88, 2009.
- [20] J. Banerjee *et al.*, “Multiple-Correlation Similarity for Block-Matching Based Fast CT to Ultrasound Registration in Liver Interventions,” *Med. Imag. Anal.*, vol. 53, pp. 132–141, 2019.
- [21] P. Bao, J. Warmath, R. Galloway, and A. Herline, “Ultrasound-to-Computer-Tomography Registration for Image-Guided Laparoscopic Liver Surgery,” *Surg. Endosc.*, vol. 19, no. 3, pp. 424–429, 2005.
- [22] J. Krücker, A. Viswanathan, J. Borgert, N. Glossop, Y. Yang, and B. J. Wood, “An Electro-Magnetically Tracked Laparoscopic Ultrasound for Multi-Modality Minimally Invasive Surgery,” in *Int. Congr. Ser.*, vol. 1281. Elsevier, 2005, pp. 746–751.
- [23] V. Martens, A. Besirevic, O. Shahin, and M. Kleemann, “LapAssistent-Computer Assisted Laparoscopic Liver Surgery,” in *Conference Proceedings of Biomedizinischen Technik (BMT). Rostock, Germany*, 2010.
- [24] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, “Augmented Reality in Laparoscopic Surgical Oncology,” *Surg. Oncol.*, vol. 20, no. 3, pp. 189–201, 2011.
- [25] M. Brudfors *et al.*, “Towards Real-Time, Tracker-less 3D Ultrasound Guidance for Spine Anaesthesia,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 6, pp. 855–865, 2015.
- [26] A. Cifor, L. Risser, M. P. Heinrich, D. Chung, and J. A. Schnabel, “Rigid Registration of Untracked Freehand 2D Ultrasound Sweeps to 3D CT of Liver Tumours,” in *Int. MICCAI Workshop Comput. Clin. Challenges in Abdominal Imag.* Springer, 2013, pp. 155–164.
- [27] S. A. Merritt, R. Khare, R. Bascom, and W. E. Higgins, “Interactive CT-video Registration for the Continuous Guidance of Bronchoscopy,” *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1376–1396, 2013.
- [28] E. G. M. Petrakis and A. Faloutsos, “Similarity Searching in Medical Image Databases,” *IEEE Trans. Knowl. Data Eng.*, vol. 9, no. 3, pp. 435–447, 1997.
- [29] S. J. Prince, *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, 2012.
- [30] G. Nir *et al.*, “Registration of Whole-Mount Histology and Volumetric Imaging of the Prostate Using Particle Filtering,” *IEEE Trans. Med. Imag.*, vol. 33, no. 8, pp. 1601–1613, 2014.
- [31] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-Squares Fitting of Two 3-D Point Sets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 5, pp. 698–700, 1987.
- [32] B. R. Thomson *et al.*, “MR-to-US Registration Using Multiclass Segmentation of Hepatic Vasculature with a Reduced 3D U-Net,” in *Proceedings of Int. Conf. on Med. Image Comp. and Comp. Ass. Intervent.* Springer, 2020, pp. 275–284.
- [33] S. Thompson *et al.*, “In Vivo Estimation of Target Registration Errors during Augmented Reality Laparoscopic Surgery,” *Int. J. Comput. Ass. Rad.*, vol. 13, no. 6, pp. 865–874, 2018.
- [34] R. Prevost *et al.*, “3D Freehand Ultrasound without External Tracking using Deep Learning,” *Med. Imag. Anal.*, vol. 48, pp. 187–202, 2018.
- [35] H. Guo, S. Xu, B. Wood, and P. Yan, “Sensorless Freehand 3D Ultrasound Reconstruction via Deep Contextual Learning,” in *Proceedings of Int. Conf. on Med. Image Comp. and Comp. Ass. Intervent.* Springer, 2020, pp. 463–472.
- [36] L. Zhang, M. Ye, P. L. Chan, and G. Z. Yang, “Real-time Surgical Tool Tracking and Pose Estimation using a Hybrid Cylindrical Marker,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 6, pp. 921–930, 2017.
- [37] U. L. Jayarathne, E. C. Chen, J. Moore, and T. M. Peters, “Robust, Intrinsic Tracking of a Laparoscopic Ultrasound Probe for Ultrasound-Augmented Laparoscopy,” *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 460–469, 2018.