

Registration of Untracked 2D Laparoscopic Ultrasound Liver Images to CT using Content-based Retrieval and Kinematic Priors

João Ramalhinho^{1,2}, Henry Tregidgo^{1,2}, Moustafa Allam³, Nikolina Travlou³,
Kurinchi Gurusamy³, Brian Davidson³, David Hawkes^{1,2}, Dean Barratt^{1,2},
Matthew J. Clarkson^{1,2}

¹ Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University
College London, London, UK

² Centre for Medical Image Computing, University College London, London, UK

³ Division of Surgery and Interventional Science, University College London, London,
UK

joao.ramalhinho.15@ucl.ac.uk

Abstract. Laparoscopic Ultrasound (LUS) can enhance the safety of laparoscopic liver resection by providing information on the location of major blood vessels and tumours. Since many tumours are not visible in ultrasound, registration to a pre-operative CT has been proposed as a guidance method. In addition to being multi-modal, this registration problem is greatly affected by the differences in field of view between CT and LUS, and thus requires an accurate initialisation. We propose a novel method of registering smaller field of view slices to a larger volume globally using a Content-based retrieval framework. This problem is under-constrained for a single slice registration, resulting in non-unique solutions. Therefore, we introduce kinematic priors in a Bayesian framework in order to jointly register groups of ultrasound images. Our method then produces an estimate of the most likely sequence of CT images to represent the ultrasound acquisition and does not require tracking information nor an accurate initialisation. We demonstrate the feasibility of this approach in multiple LUS acquisitions taken from three sets of clinical data.

Keywords: Laparoscopic Ultrasound · Multi-modal Registration · Bayesian models.

1 Introduction

There are well known advantages of laparoscopic liver resection over open surgery. However, globally, only 5-30% of cases are deemed suitable for the laparoscopic approach, typically low risk cases in which tumours are small and far from major vessels [1]. Laparoscopic Ultrasound (LUS) can reduce this risk by imaging sub-surface structures of the liver. However, tumours are not always visible in these images. Therefore, registration of LUS to CT based on vessel information has been proposed as a guidance method.

Registering LUS images to CT is a very challenging problem in part due to the limited field of view of the LUS probe. The majority of US to CT registration methods rely on the acquisition of complete liver sections using transabdominal US [2, 3]. However, LUS captures much smaller sections of the liver, providing less information to constrain the registration problem. Therefore, current methods either require manual point-picking of vessel bifurcations [4] or a fairly accurate initialisation [5]. During laparoscopic surgery, none of these options are desirable as they require a challenging and time consuming manual interaction with the intra-operative data. It has been demonstrated that a globally optimal registration can be obtained if enough LUS data is acquired [6]. However, tracking information is required to compose the LUS images in 3D space.

We propose a novel registration method that is globally optimal and uses a Content-Based Image Retrieval (CBIR) approach to register a group of untracked ultrasound slices. By simulating a finite number of ultrasound planes in the pre-operative model, we generate a discrete set of possible solutions for the alignment. For this to be feasible, we encode the vessel content of each image to a single feature vector [7]. A registration is then obtained by comparing the feature vector of a segmented ultrasound image with the pre-computed retrieval database. Due to differences between the pre-operative and intra-operative data, the simulated CT will not contain exact matches with the LUS input. Thus a set of possible matches must be considered. Assuming consecutive images are close in rotation and translation space, we employ a probabilistic model to estimate the most likely sequence of simulated images in CT space that represents the LUS acquisition. We formulate a discrete Hidden Markov Model (HMM) in which each state represents the probability of a CT match representing the original LUS and incorporate a kinematic prior as a boundary condition. The most likely combination of images is then estimated using the Viterbi algorithm. We hypothesise that this sequence estimation converges to a correct solution after enough LUS images hence states are combined in the algorithm.

Even though untracked ultrasound registration methods have been presented [8, 9], the initialisation problem has not been solved. Our novel approach is the first to tackle the problem of initialisation and to work without a tracker. We perform tests in both synthetic and clinical LUS data from 3 different patients and show the feasibility of this method.

2 Methods

Given a set of N , 2D ultrasound images $\{I_1, \dots, I_N\}$ and corresponding acquisition time stamps $\{t_1, \dots, t_N\}$, we aim to recover the sequence of LUS slices, simulated from pre-operative CT slices $\{J_1, \dots, J_N\}$ that most closely represent the LUS acquisition in terms of vascular content. In a first step we use CBIR to obtain a set of K possible slices $\{J_{1i}, \dots, J_{Ki}\}$ as candidates for each image I_i . We then apply the Viterbi algorithm with kinematic prior information in order to find the optimal sequence.

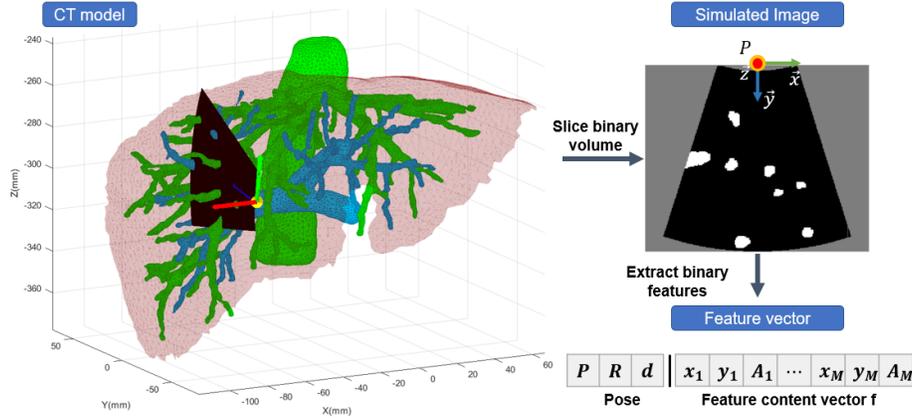


Fig. 1. CBIR database generation. For each position on the liver surface P determined by the surface mesh faces, each rotation $R = [\vec{x}, \vec{y}, \vec{z}]$, and each depth d along the surface normal, a binary image is simulated from the CT, capturing the vessels, which are encoded as a feature vector \mathbf{f} .

2.1 Retrieval based candidate selection

The pool of possible solutions J is generated by intersecting the CT segmented vascular model with 2D planes, bounded by an LUS field of view and parameterised by a set of evenly distributed points P_S along the segmented liver surface. At each of these points P_S , we create a virtual reference orientation of the LUS probe by placing it orthogonal to the liver surface normal and aligning its imaging plane with the sagittal plane. Several combinations of rotations R_x , R_y , and R_z are applied to this reference to generate rotated projections parameterised by $R = [\vec{x}, \vec{y}, \vec{z}]$. Additionally, we apply a translation d along the liver surface normal, simulating the case in which the probe compresses the liver tissue and images deeper structures. For each combination of P_S, R and d a binary image containing vessel sections is generated, as illustrated in Figure 1. The 2D Centroid position and area are extracted from each of the M binary vessel sections and stored in a feature vector \mathbf{f} as a single *feature triplet*, \mathbf{f}_i . Therefore, we establish a retrieval system in which \mathbf{f} holds the content that encodes an image J and corresponding probe configuration $[P_S, R, d]$.

Image Retrieval: Assuming prior segmentation of LUS vessel lumens, we retrieve feasible candidate poses for an input LUS image, I , by comparing its feature vector \mathbf{f}^I to all the generated pre-computed vectors \mathbf{f} . Comparison between vectors is performed by calculating the weighted L^2 distance,

$$D(\mathbf{f}^S, M^S, \mathbf{f}^L, M^L) = \left(\frac{\sum_1^{M^L} A(\mathbf{f}_i^L)}{\sum_1^{M^S} A(m(\mathbf{f}_i^S, \mathbf{f}^L))} \right) \cdot \sum_{i=1}^{M^S} \|\mathbf{f}_i^S - m(\mathbf{f}_i^S, \mathbf{f}^L)\|^2 \quad (1)$$

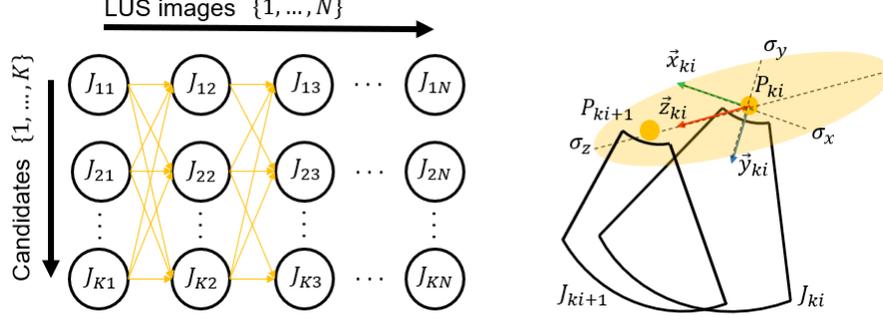


Fig. 2. HMM formulation of the problem. Left depicts the graphical model of the HMM that is optimised. Right is a visual representation of the translation probability density function of an image J_{ki} being followed by image J_{ki+1} .

where \mathbf{f}^S and \mathbf{f}^L are feature vectors with a smaller number M^S and larger number M^L of vessel sections respectively. In equation 1, the function $m(\mathbf{f}_i^S, \mathbf{f}^L)$ returns the feature triplet values in \mathbf{f}^L with the closest lumen centroid to that of triplet \mathbf{f}_i^S , and the function $A(\cdot)$ returns the area value from a triplet. We introduce an area ratio to penalise the exclusion of triplets from the longer vector \mathbf{f}^L - the total area of all vessels in \mathbf{f}^L is divided by the sum of the ones that were included in the matching. The larger the excluded areas, the larger D becomes and the less similar the vectors.

To perform an efficient search over the database, we only search for vectors that have a number of sections similar to the input \mathbf{f}^I . For this reason, we group feature vectors in lookup tables F_M according to their size M . The search for the best candidates \mathbf{f}^* is expressed in equation 2:

$$\mathbf{f}^* = \arg \min_{\mathbf{f}^I \in F_T} \frac{D(\mathbf{f}^I, M^I, \mathbf{f}_T, M^T)}{\min(M^I, M^T)}, \quad F_T = \bigcup_{l=-r}^r F_{M^I+l} \quad (2)$$

Here, the distance D is computed between \mathbf{f}^I and members of the lookup tables of size $M^I - r$ to $M^I + r$, where r is the allowable limit on feature vector length differences. The results are normalised by the minimum number of sections used in each comparison, and a set of lowest K candidate \mathbf{f}^* vectors picked. Using the CBIR encoding, these vectors then become a set of CT images $\{J_{1i}, \dots, J_{Ki}\}$ with corresponding probe poses.

2.2 Viterbi Algorithm Kinematic Constrained Optimisation

Once we obtain a pool of K possible matches $\{J_{1i}, \dots, J_{Ki}\}$ for an image I_i we introduce kinematic prior information to select the optimal candidate. We hypothesise that, given enough LUS images and prior knowledge on the kinematics of the acquisition, we can pick the sequence of candidates $\{J_{k1}, \dots, J_{kN}\}$ that best

represents $\{I_1, \dots, I_N\}$ in CT space. We formulate this problem as the optimisation of a discrete HMM as described in figure 2. In this model, nodes represent probabilities of images I_i matching a candidate J_{ki} and edges represent a probability associated with an input kinematic prior. We assume two main priors in the acquisition: firstly, there is smoothness in the acquisition; secondly, the probe follows a continuous path along the direction normal to the imaging plane without moving backwards. We then define the transition probability between 2 candidates with the following multivariate Gaussian:

$$P(J_{ki+1}|J_{ki}) = \frac{\exp(-\frac{1}{2}\delta_{ki+1,ki}^T \Sigma^{-1} \delta_{ki+1,ki})}{\sqrt{2\pi^4|\Sigma|}}, \quad \Sigma = (t_{i+1} - t_i) \begin{bmatrix} \Sigma_t & 0 \\ 0 & \sigma_\theta \end{bmatrix} \quad (3)$$

where $\delta_{ki+1,ki}$ is a vector containing the differences in rotation and translation between the 2 candidates. Specifically this is the 3D difference between probe contact points P_{ki+1} and P_{ki} projected along the orientation R_{ki} of J_{ki} and the angle between the imaging plane normals \vec{z}_{ki+1} and \vec{z}_{ki} . The covariance matrix Σ is expressed in block matrix notation and holds a variance σ_θ and a diagonal translation covariance Σ_t with three terms σ_x , σ_y and σ_z . This equation models a Gaussian distribution centered at the pose of the previous image J_{ki} with variance proportional to the time difference. The lower the time difference, the lower the pose difference should be. Since we expect the probe to move along the imaging plane normal, we define the variance σ_z to be larger than σ_x and σ_y , favouring differences in that direction (see Figure 2).

To find the optimal sequence of candidates we find the lowest cost path of this graphical model by applying the Viterbi algorithm. Since we are mainly interested in constraining the problem with kinematic information, we assume node probabilities $P(I_i|J_{ki})$ to be 1. During optimisation, we introduce a hard constraint in order to obtain trajectories that fulfil the forward movement prior. For every path in the graphical model, a sweep direction is defined as the difference between the two first probe contact positions. The probability $P(J_{ki+1}|J_{ki})$ is set to 0 if the angle between $P_{ki+1} - P_{ki}$ and that direction is above 90° .

3 Experiments

We apply our method to data from 3 patients. Pre-operative models of the liver surface and vasculature are segmented⁴ and respective databases generated using rotation angles in the intervals $R_x = R_z = [-40, 40^\circ]$, $R_y = [-90, 90^\circ]$ with steps of 10° and depth values in the interval $d = [0, 20\text{mm}]$ with steps of 5mm.

Initially, we test the validity of our model by registering synthetic sweeps generated from a CT model to itself. For each of the 3 patients we generate 3 smooth trajectories of 20 images with time stamps $t = [1, \dots, 20s]$. Retrieval with search limit $r = 0$ is applied to find $K = 200$ candidates for each image and

⁴ www.visiblepatient.com

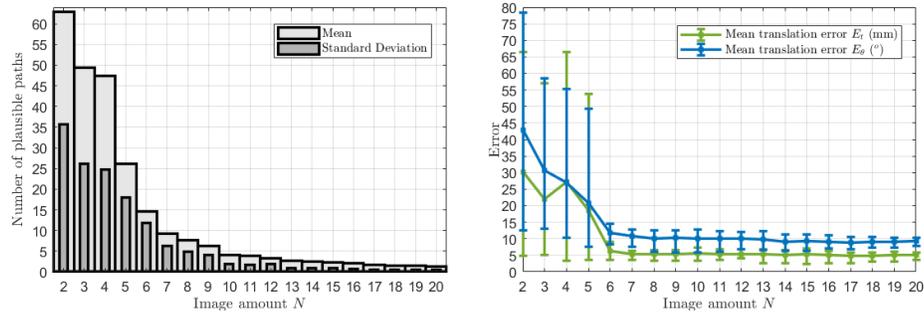


Fig. 3. Registration results for 9 synthetic sweeps. Left plot shows the number of plausible sequences found by the Viterbi algorithm. Right plot displays the translation and rotation error for each paths optimal registration. For visualisation purposes, the mean error for these nine optimal paths is displayed. Bars have been placed to show minimum achieved error and one standard deviation above the mean.

registrations are performed using model variances $\sigma_z = 1.5\text{mm}$, $\sigma_x = \sigma_y = 0.2\sigma_z$ and $\sigma_\theta = 2^\circ$.

Mean results over the nine sweep registrations are summarised in Figure 3. Since the Viterbi algorithm is recursive on the number of columns of the HMM, results are displayed as a function of the number of images used so far in the optimisation (from 2 to 20). The left hand graph shows the number of kinematically possible paths for N images. As expected, the number of plausible trajectories found by the algorithm converges to 1 if enough images are used ($N = 17$ in this case). The right hand graph shows the mean translation error E_t and mean rotation error E_θ across all N registered images for the lowest cost path. Here the convergence is observed in terms of error: at an average number of 7 images, these errors converge to 5mm and 10° . Such values are expected since the rotation resolution of the database is 10° and the used liver surfaces have a spatial resolution in the range [3-4mm].

Table 1. Results of registration of 6 sweeps of clinical LUS. N is the number of images in the sweep. N_C the number of images the Viterbi algorithm required to converge in error. E_t and E_θ are the mean translation and rotation error respectively. TRE represents the Root Mean Square (RMS) of the TRE of manually picked landmarks found in the sweep.

Dataset	Patient 1		Patient 2		Patient 3	
	Sweep 1	Sweep 2	Sweep 1	Sweep 2	Sweep 1	Sweep 2
N	42	12	25	11	23	19
N_C	6	8	16	6	19	16
\bar{E}_t (mm)	14.7	18.7	11.2	10.3	19.1	15.3
\bar{E}_θ ($^\circ$)	14.8	33.3	17.2	13.9	44.1	32.3
TRE(mm)	18.8	3.7	14.2	11.4	25.3	21.9

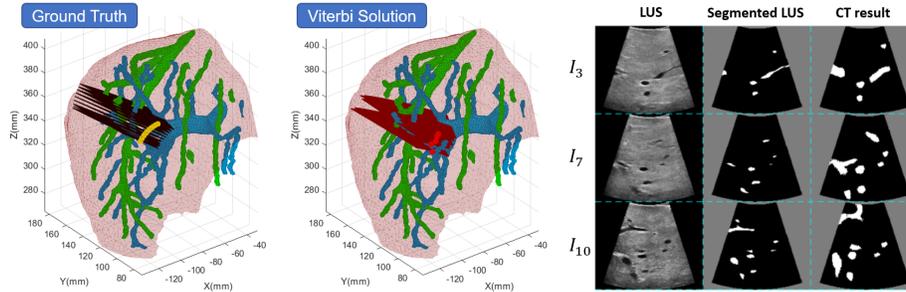


Fig. 4. Registration result of Sweep 2 of Patient 2. Left shows the 3D registration of the LUS planes using the Ground Truth (black planes with yellow dots) and the algorithm solution (red planes with red dots). Right shows LUS and segmented CT alignment results of 3 images in the sweep.

To test the feasibility of the method on real data, we retrospectively register LUS acquired intra-operatively with a BK Medical 8666-RF⁵ probe at a frame rate of 40Hz. From each patient, we select 2 sequences of contiguous images that do not contain shadowing artefacts or large non-tubular vessel sections, and manually segment their vessels. To avoid redundancy, inside each dataset we pick evenly spaced images in time that differ in content. We apply the algorithm with a wider search of $r = 2$ to find $K = 1000$ candidates. We double the translation variance values ($\sigma_z = 3\text{mm}$) and keep the rotation the same. For each sweep, we manually register LUS images to CT and interpolate the result with a cubic polynomial to generate a ground truth trajectory. After obtaining a solution, we again measure the errors E_t and E_θ and assess the Target Registration Error (TRE) of a set of manually picked vessel bifurcations found in the path. Since these bifurcations may land in images in between the sequence that were not registered, we perform a cubic polynomial fit to predict their position given the algorithm solution.

Results on the six sweeps are summarised in Table 1. The best trajectory registration results are found in the sweeps of patient 2, with translation errors around 10mm. A visual result with the registration of Sweep 2 is shown in Figure 4. Lowest accuracies are obtained for patient 3, but these do not surpass 20mm. This value is still usable as we are performing a globally optimised alignment. We also display the number of images N_C at which the errors converge as in the previous experiment. Since these vary greatly, we assume that this value depends on the uniqueness of the registered images that is specific to each dataset. TRE results are in the range of [3.7-25.3mm] and are in reasonable agreement with the other errors.

⁵ www.bkmedical.com

4 Conclusions

Our results show that our framework can register smaller field of view images to a larger volume globally and without tracking information. While the proposed method does not perfectly register each LUS frame, the accuracy is sufficient to act as an initialisation for local registration methods such as [5]. We see this as a great step forward in this field. It poses both a reduction in manual interaction and less interruption to the clinical workflow as a tracking device is not required. Furthermore, our simulation was purely rigid: by increasing the realism of our database simulations with deformation, higher accuracies can be achieved.

Although the method used manually segmented vessels in ultrasound, we believe that automatic segmentation results can be obtained using state-of-the-art Deep Learning frameworks [10]. We did not include ultrasound images with large non-tubular sections in our validation. This is due to the fact that our retrieval based in position and area is not specific enough to identify such structures. We intend to tackle this problem by both refining our encoding and including modelling of physics in the simulation step.

It is worth noting that this framework can be translated to other registration problems. The only requirements are suitable priors on the acquisition and a robust image-to-feature encoding that describes the target anatomy.

Acknowledgements JR was supported by the EPSRC CDT in Medical Imaging [EP/L016478/1] and EPSRC grant [EP/P034454/1]. MJC, DH and DB were supported by the Wellcome/EPSRC [203145Z/16/Z]. BD was supported by the NIHR Biomedical Research Centre at University College London Hospitals NHS Foundations Trust and University College London. The imaging data used for this work was obtained with funding from the Health Innovation Challenge Fund (HICF-T4-317), a parallel funding partnership between the Wellcome Trust and the Department of Health. The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health.

References

1. Wakabayashi, G., Cherqui, D., Geller, D.A., Buell, J.F., Kaneko, H., et al.: Recommendations for laparoscopic liver resection: a report from the second international consensus conference held in Morioka. *Ann. Surg.* **261**(4), 619–629 (2015)
2. Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Med. Image. Anal.* **12**(5), 577–585 (2008)
3. Penney, G.P., Blackall, J.M., Hamady, M.S., Sabharwal, T., Adam, A., Hawkes, D.J.: Registration of freehand 3D ultrasound and magnetic resonance liver images. *Med. Image. Anal.* **8**(1), 81–91 (2004)
4. Song, Y., Totz, J., Thompson, S., Johnsen, S., Barratt, D., et al.: Locally rigid, vessel-based registration for laparoscopic liver surgery. *IJCARS* pp. 1–11 (2015)
5. Ramalhinho, J., Robu, M., Thompson, S., Edwards, P., Schneider, C., et al.: Breathing motion compensated registration of laparoscopic liver ultrasound to CT. In: *Proc. SPIE Medical Imaging* (2017)

6. Ramalhinho, J., Robu, M.R., Thompson, S., Gurusamy, K., Davidson, B., et al.: A pre-operative planning framework for global registration of laparoscopic ultrasound to CT images. *IJCARS* **13**(8), 1177–1186 (2018)
7. Petrakis, E.G.M., Faloutsos, A.: Similarity searching in medical image databases. *IEEE Trans Knowl Data Eng* **9**(3), 435–447 (1997)
8. Cifor, A., Risser, L., Heinrich, M.P., Chung, D., Schnabel, J.A.: Rigid registration of untracked freehand 2D ultrasound sweeps to 3D CT of liver tumours. In: Yoshida, H., Warfield, S., Vannier, M.W. (eds.) *Abdominal Imaging. Computation and Clinical Applications*. pp. 155–164. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
9. Nir, G., Sahebjavaher, R.S., Kozlowski, P., Chang, S.D., Jones, E.C., et al.: Registration of whole-mount histology and volumetric imaging of the prostate using particle filtering. *IEEE Trans Med Imaging* **33**(8), 1601–1613 (2014)
10. Smistad, E., Løvstakken, L.: Vessel detection in ultrasound images using deep convolutional neural networks. In: *Deep Learning and Data Labeling for Medical Applications*, pp. 30–38. Springer (2016)