Using Multiple Images and Contours for Deformable 3D-2D Registration of a Preoperative CT in Laparoscopic Liver Surgery

Yamid Espinel[1(✉)], Lilian Calvet[1,2], Karim Botros[1], Emmanuel Buc[2], Christophe Tilmant[1], and Adrien Bartoli[1]

[1] EnCoV, Institut Pascal, UMR 6602 CNRS/Université Clermont-Auvergne, Clermont-Ferrand, France
{yamid.espinel,karim.botros,christophe.tilmant,adrien.bartoli}@uca.fr
[2] University Hospital of Clermont-Ferrand, Clermont-Ferrand, France
ebuc@chu-clermontferrand.fr

Abstract. Deformable registration is required to achieve laparoscopic augmented reality but still is an open problem. Some of the existing methods reconstruct a preoperative model and register it using anatomical landmarks from a single image. This is not accurate due to depth ambiguities. Other methods require of non-standard devices unadapted to the clinical practice. A reasonable way to improve accuracy is to combine multiple images from a monocular laparoscope. We propose three novel registration methods exploiting information from multiple images. The first two are based on rigidly-related images (MV-B and MV-C) and the third one on non-rigidly-related images (MV-D). We evaluated registration accuracy quantitatively on synthetic and phantom data, and qualitatively on patient data, comparing our results with state of the art methods. Our methods outperforms, reducing the partial visibility and depth ambiguity issues of single-view approaches. We characterise the improvement margin, which may be slight or significant, depending on the scenario.

Keywords: Laparoscopy · Liver · Registration · Augmented reality

1 Introduction

Laparoscopic liver surgery is less invasive than open surgery, reducing patient trauma and recovery time. However, it is difficult for the surgeon to localise the inner structures such as vessels and tumours. This is mainly caused by the large size of the liver and its proximity to the laparoscope. Augmented Reality (AR) can provide aid to surgery by showing these internal structures, hence improving resection planning. Concretely, AR overlays a preoperative 3D model reconstructed from CT data onto the laparoscopic images. AR requires one to register the preoperative 3D model to the 2D laparoscopic images. This is done using anatomical landmarks present on the liver's surface. Because the liver deforms

significantly between the preoperative and intraoperative states, the registration must be a 3D-2D deformable one. This is a difficult and currently open problem. Existing works have addressed the cases of monocular laparoscopes [1–3] and of stereoscopic laparoscopes, possibly with external tracking [4–10]. Methods [1,2] are single-view. They use sparse landmarks, leaving ambiguities on the registration. Method [3] is multi-view and reconstructs an intraoperative shape by Structure-from-Motion (SfM), solving registration by ICP and 3D landmarks. This method was however only tested with animal data as the required correspondences are not available in real surgery cases. Finally, methods [4–10] perform rigid registration only. The monocular case is important because it forms the standard in many operating rooms. Therefore, there is a need for an unambiguous monocular registration method compatible with the clinical constraints and the desired clinical outcomes, studied in [13].

We propose to use multiple monocular laparoscopic images to solve deformable registration without unrealistic requirements. Our methods solve registration for cases where images are rigidly-related and non-rigidly-related. We assume that the images show the liver from multiple viewpoints. Our methods are inspired by [1], a method named SV (for Single-View) hereinafter. Concretely, we deform the preoperative 3D model using a particle system with biomechanical properties and position-based dynamics [15]. For the rigidly-related case, we propose two methods. The first one, named MV-B (for Multi-View Base), guides deformation using the landmarks from all the images and the inter-image rigidity as constraints. The second one, named MV-C (for Multi-View Correspondences), extends MV-B by exploiting the liver's texture information via inter-image keypoint correspondences. For the non-rigidly-related case, we propose the MV-D method (for Multi-View Deformable Correspondences) which, compared to MV-C, does not use the inter-image rigidity and produces several registered shapes. These methods are illustrated in Fig. 1. By exploiting multiple images, our methods solve the partial visibility and depth ambiguity issues inherent to SV. We evaluated our methods on synthetic, phantom and patient data. For the synthetic and phantom case, target registration errors (TRE) were measured on uniformly-distributed control points inside the liver models. For the patient case, we measured the liver's 2D reprojection errors on control views. We aim to have a TRE lower than 1 cm, as it is the resection margin advised for Hepato-Cellular Carcinoma (HCC) interventions [17]. We do not measure TRE on patient data as having a reliable ground-truth is not possible with the available devices in the surgery room.

2   Background

SV uses a volumetric deformable model M to predict image contours C. M is modeled using the isotropic Neo-Hookean elastic model [15] with generic values for the human liver's mechanical parameters [16]. We use this model as it can simulate large deformations in a computationally inexpensive way and is validated for the hepatic tissue [11,12]. The constraints given by the elastic model
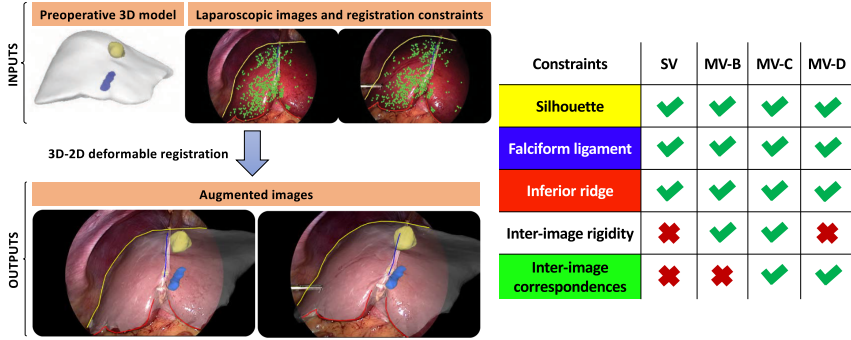
Constraints	SV	MV-B	MV-C	MV-D
Silhouette	✔	✔	✔	✔
Falciform ligament	✔	✔	✔	✔
Inferior ridge	✔	✔	✔	✔
Inter-image rigidity	✖	✔	✔	✖
Inter-image correspondences	✖	✖	✔	✔

Fig. 1. Characteristics of the state-of-the-art and proposed registration methods.

are denoted as Ω_{def}. The predicted contours C should match the observed image contours C^*. As shown in Fig. 1, SV uses three types of contours: the ridge contour, which has a very distinctive profile; the falciform ligament contour; and the silhouette contour which corresponds to the occluding boundaries of the liver. The constraints given by this set of contours are denoted as Ω_{ctr}. SV uses a convergent algorithm of alternating projections to solve the registration problem. As such, it finds the closest model M to the constraints Ω_{def} and Ω_{ctr}:

$$\min_M dist(M, \Omega_{def}) + dist(M, \Omega_{ctr}) \tag{1}$$

where $dist$ is the distance between the model and the constraints.

3   Methodology

The common principle of our methods is to use several laparoscopic images to solve registration. These images should present a noticeable change in camera pose, while keeping some overlap. This can be done by tilting and panning the laparoscope. Views from different trocars can also be used, should they sufficiently overlap. Using multiple images is especially useful for the liver case as it helps to overcome the partial visibility and, thus, the low precision issues inherent to single-view methods.

These methods try to solve registration in two situations: when the liver does not deform across images, e.g. during the exploratory phase (rigidly-related case) and when it does, e.g. during surgical manipulation (non-rigidly-related case). For the rigidly-related case, MV-B and MV-C take advantage of the liver's rigidity to constrain registration. Thus, they produce a single deformed shape, along with the corresponding camera poses for each of the views. In order to have suitable images for this case, the surgeon can pause the artificial ventilation system for a very short period of time (about $10\,\mathrm{s}$), while the scene is filmed. Compared to MV-B, MV-D makes use of inter-image correspondences which helps to improve the registration accuracy. However, MV-B is still useful in

cases where enough valid correspondences cannot be obtained. For the non-rigidly-related case, MV-D also makes use of inter-image correspondences and produces several deformed shapes, according to the number of views used.

Figure 2 illustrates the general pipeline followed by MV-B, MV-C and MV-D to solve registration. The principle is similar for the three methods, taking SV as basis and adding the necessary steps to use the inter-image rigidity and the correspondences as constraints.
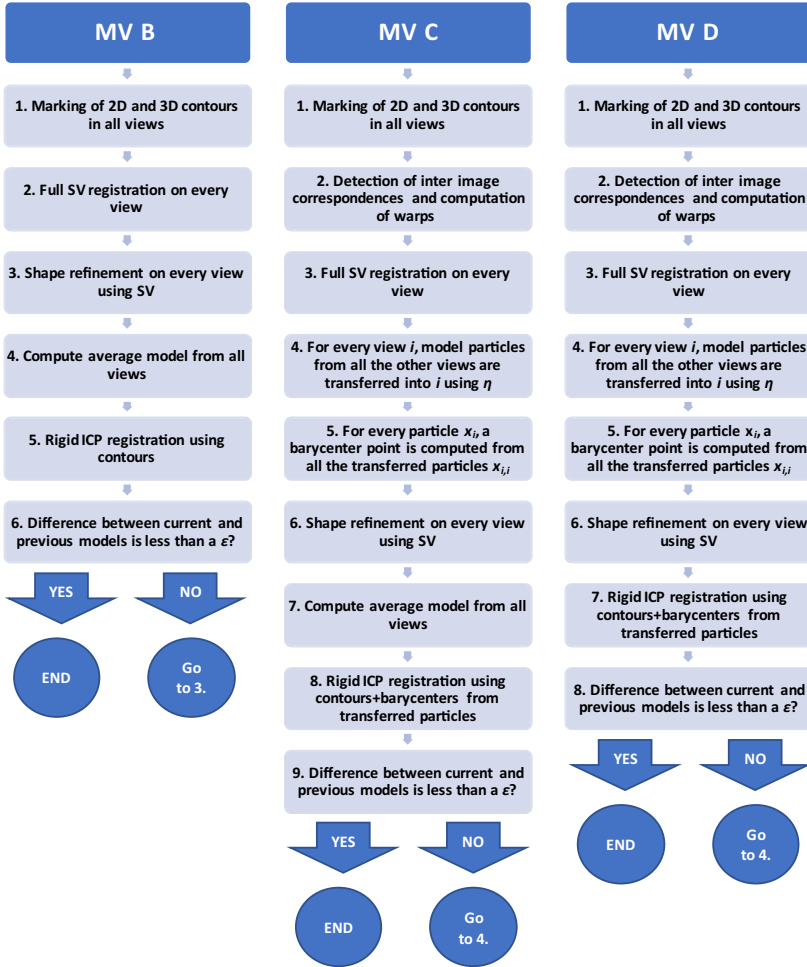

Fig. 2. Pipeline followed by the MV-B, MV-C and MV-D methods. (Color figure online)

3.1   Multi-View Rigid Base (MV-B)

MV-B extends SV to solve registration on N rigidly-related laparoscopic images simultaneously. At every iteration, it computes an average model M from all the views after an SV refinement has been done on every view. A single general model is thus produced from each view's individual contribution. Equation (1) becomes:

$$\min_{M} \sum_{i=1}^{N} dist(M, \Omega_{def_i}) + \sum_{i=1}^{N} dist(M, \Omega_{ctr_i}) \qquad (2)$$

The algorithm that solves this problem is composed by the blue and green stages from Fig. 2. The computation of the average model is the key to impose the inter-image rigidity. It converges when the difference between models from previous and current iterations $\Delta M < 10^{-3}$. This criterion was found after running our methods for a large number of iterations on a variety of synthetic, phantom and patient data.

### 3.2   Multi-View Rigid with Inter-image Correspondences (MV-C)

MV-C extends MV-B to exploit inter-image correspondences. These correspondences are obtained using SIFT and mismatches rejected using FBDSD [18]. Our key idea is to measure the consistency between the registration and the correspondences using the P particles $x_p, p \in [1, P]$, representing the deformable model. These correspondences are related to each other by the warp transformation function η_{ji}, which is based on the Rigid-Perspective Thin-Plate Spline [14]. Such function η_{ji} lets us transfer any point from image j to image i, as shown in Fig. 3. Like MV-B, MV-C generates a single general model from each view's individual contribution by using the inter-image rigidity and the available inter-image correspondences Ω_{corr} as constraints:

$$\min_{M} \sum_{i=1}^{N} dist(M, \Omega_{def_i}) + \sum_{i=1}^{N} dist(M, \Omega_{ctr_i}) + \sum_{i=1}^{N} dist(M, \Omega_{corr_i}) \qquad (3)$$

The algorithm that solves this problem is composed by the blue, yellow and green stages from Fig. 2. It adds the necessary steps to use the keypoint correspondences as constraints, and it also converges when $\Delta M < 10^{-3}$.

3.3   Multi-View Deformable Correspondences (MV-D)

MV-D modifies MV-C to avoid using the inter-image rigidity to solve registration, as well as to use the inter-image correspondences in the SV-refinement stage. In this way, MV-D produces multiple shapes M_N according to the number of images used and the deformations exerted by the liver on each of them. This is especially useful in cases where the liver is being manipulated by tools or other external forces. Consequently, MV-D can be seen as:
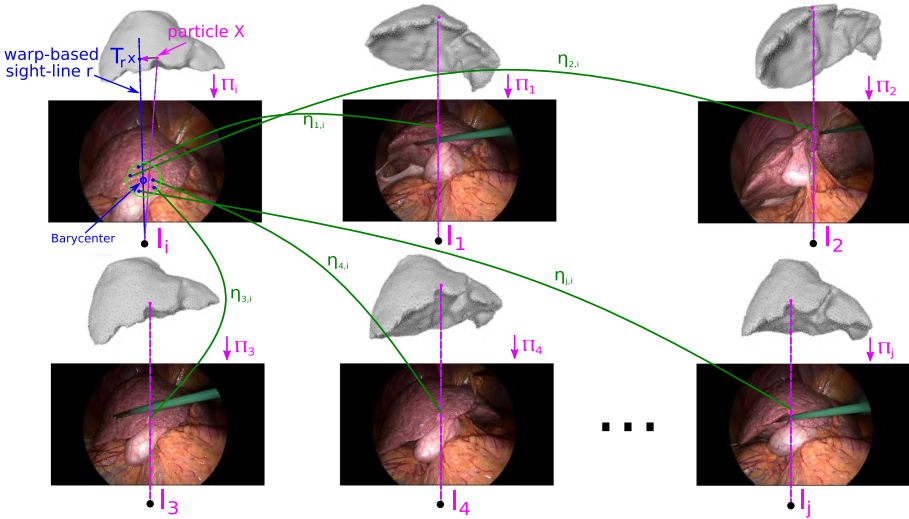
Fig. 3. Warp-based sight-line for particle location update. The particle locations are updated as follows: pixel motion is preliminary computed for all image-pairs i, j and the warps robustly estimated; during an iteration of the optimization algorithm, for every image i, a particle x_p is updated such that its new location corresponds to its orthogonal projection onto its warp-based sight line, namely the backprojections of the barycenter of the imaged particles $\eta_{ji}(\Pi_j(x_p))$ through the warps.

$$\min_{M_1,\ldots,M_N} \sum_{i=1}^{N} dist(M_i, \Omega_{def_i}) + \sum_{i=1}^{N} dist(M_i, \Omega_{ctr_i}) + \sum_{i=1}^{N} dist(M_i, \Omega_{corr_i}) \quad (4)$$

The algorithm that solves this problem is composed by the blue and yellow stages from Fig. 2, avoiding the usage of the inter-image rigidity. It also converges when $\Delta M < 10^{-3}$.

4   Experimental Results

4.1   Rigidly-Related Views

Synthetic Data: We reconstructed a virtual 3D liver model from a patient's CT and synthesized 10 virtual deformations using Abaqus [19], by simulating gravity, the pneumoperitoneum and the action of surgical instruments. We generated 10 images of each deformation by simulating a virtual moving laparoscope. We estimated registration for a varying number of images, going from 1 to 8, and measured TRE as the average prediction error for uniformly sampled points within the virtual liver. It should be noted that, while the contour marking is done manually, it does not take more than 5 min to mark all the 8 views used as the maximum case for our experiments. Thus, the impact on the surgical workflow

is minimal. We use 8 views as a maximum to keep the computation time reasonable. We repeated the estimation 10 times for each number of images, randomly selecting the images being used. The results are shown in Fig. 4(a). We observe that the TRE for SV is steady around (9.57 ± 7.99) mm, while for MV-B and MV-C it consistently decreases as the number of views increases. The decrease is notable in both average and standard deviation, reaching (7.99 ± 7.34) mm and (5.95 ± 4.46) mm respectively for MV-B and MV-C.

Phantom Data: We 3D printed the synthetic deformations generated in Abaqus using PLA (Polylactic Acid). We then used a surgical laparoscope and a pelvitrainer box to take 10 pictures of each printed model. Similarly to the synthetic case, 10 combinations of 8 views were generated per deformation and experiments were run using 1 to 8 views for every combination. Distances were also measured between the registered and ground truth control points, for which the average and standard deviation are shown in Fig. 4(b). As for the synthetic case, SV remains steady around (11.96 ± 7.72) mm, while for MV-B and MV-C it consistently decreases as the number of views increases. We can see a decrease in both average and standard deviation, especially for MV-C. MV-B remains close to SV, reaching (11.59 ± 10.95) mm, while MV-C decreases to and (8.30 ± 4.59) mm.

Patient Data: We collected data for 5 patients, for which we had IRB approval (IRB8526-2019-CE58, CPP Sud-Est VI). We kept 9 images per patient. Out of these images, we singled one out to serve as a control view. The control view is not used to compute registration but as a means to verify registration, using the landmark prediction error expressed in px (number of pixels). The control view error is a weak measure of TRE because it only concerns the visible liver surface. We measure such reprojection errors due to the difficulty of having a reliable groundtruth to evaluate the registration accuracy in 3D. It is shown in Fig. 5(a). We run both MV-B and MV-C with 8 images. We observe a clear benefit of using multiple images and of using the inter-image correspondences.
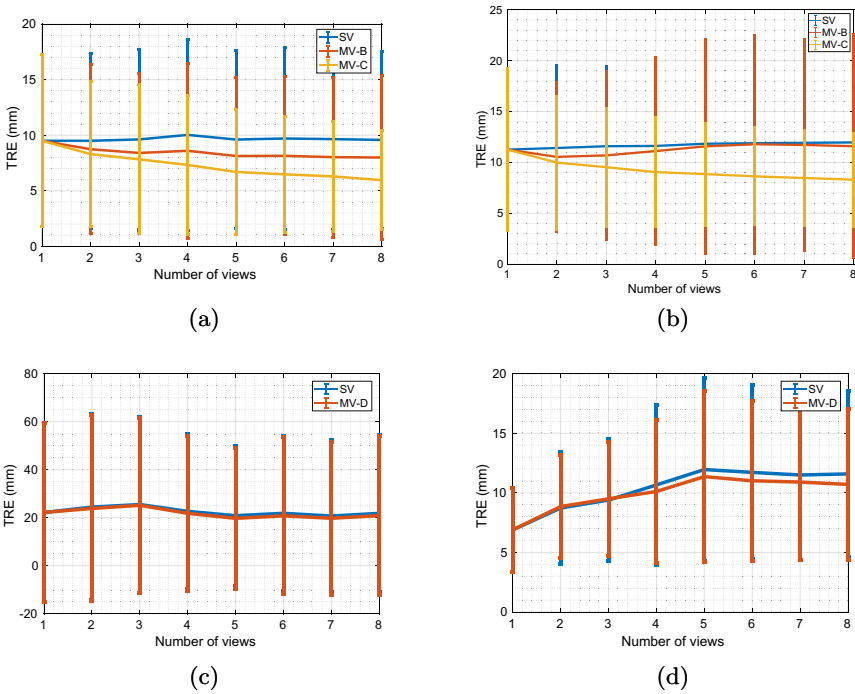
4.2   Non-rigidly-related Views

Synthetic Data: From the previously generated synthetic data, 10 combinations of 8 views were taken to assess MV-D, with every view corresponding to a different deformation. As for the rigidly-related case, we estimated registration for a varying number of images and measured TRE as the average prediction error for uniformly sampled points within the virtual liver. The results are shown in Fig. 4(c). For the experiments using 8 views, SV has a TRE of (21.83 ± 32.61) mm on the control points, while that for MV-D is (20.87 ± 32.96) mm.

Phantom Data: From the generated phantom data, and similarly to the non-rigidly-related synthetic case, experiments were done on 10 combinations of 8

views, with every view corresponding to a different deformation. TRE results on a varying number of images are shown in Fig. 4(d). For the experiments using 8 views, SV has a TRE of (11.59 ± 6.95) mm on the control points, while MV-D's one is (10.71 ± 6.35) mm.

Patient Data: From the previously acquired patient data we have selected 9 images per patient. Here, the liver exerts significant deformation across the images. As for the rigidly-related case, we singled one image out to serve as a control view. Landmark prediction error is computed on the control views and expressed in px (number of pixels). It is shown in Fig. 5(b). We run both MV-B and MV-C with 8 images. We observe a slight benefit of using multiple images to solve registration.

(a)   (b)   (c)   (d)

Fig. 4. Mean TRE and standard deviations on (a) synthetic data usisng the rigidly-related methods, (b) phantom data using the rigidly-related methods, (c) synthetic data using the non-rigidly-related methods, (d) phantom data using the non-rigidly-related methods.

Patient	SV	MV-B	MV-C
1	26.07	<u>21.98</u>	**19.49**
2	58.40	<u>10.28</u>	**09.49**
3	42.78	<u>32.24</u>	**29.75**
4	**17.13**	21.72	<u>21.13</u>
5	26.71	**15.46**	<u>16.95</u>
average	34.28	<u>20.33</u>	**19.36**

(a)

Patient	SV	MV-D
1	32.47	**30.14**
2	63.08	**57.62**
3	**46.24**	49.79
4	25.07	**21.31**
5	31.85	**29.75**
average	39.74	**37.73**

(b)

Fig. 5. Control view errors (px) on patient data for (a) rigidly-related methods with best results in bold and second best underlined and (b) non-rigidly-related methods with best results in bold.

5   Conclusions

We proposed 3 multi-view methods for 3D-2D deformable registration of preoperative CT data into intraoperative images for laparoscopy of liver. They aim to solve registration on rigidly- and non-rigidly-related views. Results on synthetic data show that MV-B improves the mean registration accuracy by 1.58 mm, while MV-C improves it by 3.62 mm compared to SV. MV-D shows a slight improvement of 1 mm compared to SV. On phantom data, MV-B has a similar performance to SV with a difference of 0.37 mm, while MV-C improves registration by 3.66 mm. MV-D also shows a similar performance to SV with a difference of 0.88 mm. On patient data, the reprojection error measured on control views is improved by 13.95 px, 14.92 px and 2.01 px for MV-B, MV-C and MV-D respectively. It means that, for the rigidly-related case, we can see an improvement with respect to SV as we increment the number of views, with registration errors below 1 cm for MV-C. For the non-rigidly-related case, MV-D behaves similarly to SV, with a slight improvement of 1 mm. Visual inspection on the error distribution shows that TRE follows an unimodal distribution with positive skewness. It is worth noting that TRE is measured on the whole 3D model, including the visible and hidden regions. Hidden regions can increase the global TRE as they are less accurately registered than the visible ones. As future work, we will focus on (i) performing extended validation on clinical patient data, (ii) developing a real-time non-rigidly tracking strategy taking the multi-view initial registration as basis, and (iii) using other visual information with improved registration approaches to increase the registration accuracy, such as surgical tools pose and Structure-from-Motion.

References

1. Koo, B., Özgür, E., Le Roy, B., Buc, E., Bartoli, A.: Deformable registration of a preoperative 3D liver volume to a laparoscopy image using contour and shading cues. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10433, pp. 326–334. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_38

2. Adagolodjo, Y., Trivisonne, R., Haouchine, N., Cotin, S., Courtecuisse, H.: Silhouette-based pose estimation for deformable organs application to surgical augmented reality. In: IROS (2017)

3. Modrzejewski, R., Collins, T., Seeliger, B., Bartoli, A., Hostettler, A., Marescaux, J.: An in vivo porcine dataset and evaluation methodology to measure soft-body laparoscopic liver registration accuracy with an extended algorithm that handles collisions. Int. J. Comput. Assist. Radiol. Surg. **14**(7), 1237–1245 (2019). https://doi.org/10.1007/s11548-019-02001-4

4. Chen, L., Tang, W., John, N.W., Wuan, T.R., Zhang, J.J.: SLAM-based dense surface reconstruction in monocular minimally invasive surgery and its application to augmented reality. Comput. Meth. Prog. Biomed. **158**, 135–146 (2018)

5. Haouchine, N., Roy, F., Untereiner, L., Cotin, S.: Using contours as boundary conditions for elastic registration during minimally invasive hepatic surgery. In: IROS (2016)

6. Robu, M.R., et al.: Global rigid registration of CT to video in laparoscopic liver surgery. Int. J. Comput. Assist. Radiol. Surg. **13**(6), 947–956 (2018). https://doi.org/10.1007/s11548-018-1781-z

7. Thompson, S., et al.: Accuracy validation of an image guided laparoscopy system for liver resection. Proc. SPIE - Int. Soc. Opt. Eng. **9415**(09), 1–12 (2015)

8. Plantefeve, R., Peterlik, I., Haouchine, N., Cotin, S.: Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery. Ann. Biomed. Eng. **44**, 139–153 (2016)

9. Clements, L., Collins, J., Weis, J., Simpson, A., Kingham, T., Jarnagin, W., Miga, M.: Deformation correction for image guided liver surgery: an intraoperative fidelity assessment. Surgery **162**(3), 537–547 (2017)

10. Bernhardt, S., Nicolau, S., Bartoli, A., Agnus, V., Soler, L., Doignon, C.: Using shading to register an intraoperative CT scan to a laparoscopic image. In: Computer-Assisted and Robotic Endoscopy, CARE (2015)

11. Chui, C., Kobayashi, E., Chen, X., Hisada, T., Sakuma, I.: Combined compression and elongation experiments and non-linear modelling of liver tissue for surgical simulation. Med. Biol. Eng. Comput. **44**, 787–798 (2004)

12. Shi, H., Farag, A., Fahmi, R., Chen, D.: Validation of finite element models of liver tissue using Micro-CT. IEEE Trans. Biomed. Eng. **55**, 978–984 (2008)

13. Thompson, S., Hu, M., Johnsen, S., Gurusamy, K., Davidson, B., Hawkes, D.: Towards Image Guided Laparoscopic Liver Surgery, Defining the System Requirement. LIVIM (2011)

14. Bartoli, A., Perriollat, M., Chambon, S.: Generalized thin-plate spline warps. Int. J. Comput. Vis. **88**, 85–110 (2010)

15. Bender, J., Koschier, D., Charrier, P., Weber, D.: Position-based simulation of continuous materials. Comput. Graph. **44**, 1–10 (2014)

16. Nava, A., Mazza, E., Furrer, M., Villiger, P., Reinhart, W.H.: In vivo mechanical characterization of human liver. Med. Image Anal. **12**(2), 203–216 (2008)

17. Zhong, F.P., Zhang, Y.J., Liu, Y., Zou, S.B.: Prognostic impact of surgical margin in patients with hepatocellular carcinoma: a meta-analysis. Medicine **96**(37), e8043 (2017)

18. Pizarro, D., Bartoli, A.: Feature-based deformable surface detection with self-occlusion reasoning. Int. J. Comput. Vis. **97**, 54–70 (2010)

19. 3DS Abaqus. http://edu.3ds.com/en/software/abaqus-student-edition. Accessed 2 Mar 2021