An Augmented Reality Framework for Soft Tissue Surgery

Peter Mountney[1], Johannes Fallert[2], Stephane Nicolau[3], Luc Soler[3,4], and Philip W. Mewes[5]

[1] Imaging and Computer Vision, Siemens Corporate Technology, Princeton, NJ, USA
[2] Imaging Technologies Research, Karl Storz, Tuttlingen, Germany
[3] Institut de Recherche contre les Cancers de l'Appareil Digestif (IRCAD), Strasbourg, France
[4] Institut Hospitalo-Universitaire de Strasbourg (IHU Strasbourg), Strasbourg, France
[5] Angiography & Interventional X-Ray Systems, Siemens Healthcare, Germany

Abstract. Augmented reality for soft tissue laparoscopic surgery is a growing topic of interest in the medical community and has potential application in intra-operative planning and image guidance. Delivery of such systems to the operating room remains complex with theoretical challenges related to tissue deformation and the practical limitations of imaging equipment. Current research in this area generally only solves part of the registration pipeline or relies on fiducials, manual model alignment or assumes that tissue is static. This paper proposes a novel augmented reality framework for intra-operative planning: the approach co-registers pre-operative CT with stereo laparoscopic images using cone beam CT and fluoroscopy as bridging modalities. It does not require fiducials or manual alignment and compensates for tissue deformation from insufflation and respiration while allowing the laparoscope to be navigated. The paper's theoretical and practical contributions are validated using simulated, phantom, *ex vivo*, *in vivo* and non medical data.

1 Introduction

Interest in augmented reality (AR) for soft tissue surgery, such as liver resection and partial nephrectomy, has grown steadily within the medical community. The role of AR in this context is procedure- and workflow-dependent. It can be used at the beginning of the surgical procedure for intra-operative planning to rapidly identify target anatomy and critical sub surface vessels, or it can facilitate image guidance to display tumor resection margins and improve dissection accuracy [1].

A number of theoretical and practical challenges remain for the translation of such systems into the operating room. The core challenge is registration of the pre-operative image (CT/MRI) with the intra-operative laparoscopic image. This in itself is challenging due to the lack of cross modality landmarks and the laparoscopic camera's small viewing field. Furthermore, surgical procedures require insufflation of the abdomen causing an initial organ shift and tissue deformation, which must be reconciled. The registration problem is further

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 423–431, 2014.

complicated during the procedure itself due to continuous tissue deformation caused by respiration and tool-tissue interaction.

Due to the complex registration pipeline required to deliver AR to the operating room, current research tends to focus on individual components of the process and do not provide complete solutions. For example, notable work exists in deformable tissue modeling [2,3], dense reconstruction [4,3], non-rigid registration of CT to cone beam CT (CBCT) [5], tissue tracking [6], surface registration [7] and laparoscopic camera pose estimation [8,9].

A handful of end-to-end systems have been proposed for the operating room that rely on additional fiducials, manual registration, or the baseline assumption that tissue is static. Challenges persist in each scenario. Fiducials act as cross modality landmarks and have been attached externally on the patient's skin [10] and to the organ itself [11]. Their use however, can be disruptive to the clinical workflow. Manual registration, on the other hand, requires experts to visually align a 3D model to the laparoscopic image [12]. Accuracy is user dependent even when alignment is constrained with a single cross modality landmark [13]. Finally, as per the static environment assumption, a comprehensive system has been proposed for skull surgery [8], but deformation compromises its accuracy.

This paper proposes an AR framework for intra-operative planning in liver surgery[1]. The novel system registers pre-operative CT and stereo laparoscopic images to a common coordinate system using CBCT and fluoroscopy as bridging modalities. It does not require fiducials or manual model alignment. Tissue deformation caused by insufflation, organ shift and respiration are accounted for along with laparoscopic camera motion. The framework is evaluated on simulated, phantom, *ex vivo*, *in vivo* and non medical data.

2 Method

A key component of the AR system is the introduction of CBCT into the operating room. CBCT machines capture 3D CT-like images and 2D fluoroscopy —in the same coordinate system —while the patient is on the operating table. CBCT and fluoroscopy are used as bridging modalities to co-register pre-operative CT and laparoscopic images. The framework consists of three registration phases: 1) a registration of CT to CBCT (Fig. 1), which takes into account tissue deformation resulting from insufflation 2) a registration of the laparoscope to CT via CBCT coordinate system (Fig. 2), accounting for tissue deformation caused by respiration and 3) a temporal registration of laparoscopic images (Fig. 3), which deals with camera motion and tissue deformation caused by respiration.

2.1 Non Rigid Registration of CT to CBCT

Pre-operative CT and organ segmentation are performed in the days or weeks prior to the operation. With the patient in the supine position, two CT images are captured using a contrast injection at the arterial and venous phases. The images

[1] Not currently commercially available.

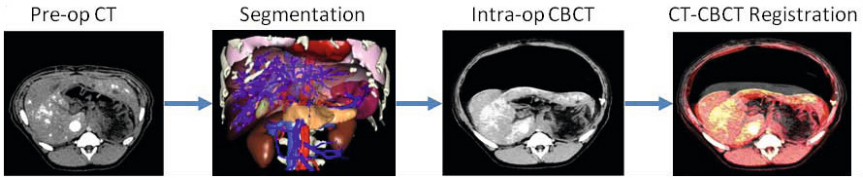

Fig. 1. Registration of pre-operative CT to intra-operative CBCT

are registered together and segmented[2] into 3D anatomical models including the liver, tumor, vessels and abdomen wall as shown in Fig. 1.

During the procedure, the patient is positioned for easy of access (e.g. reverse Trendelenburg) and the abdomen is insufflated with CO_2 causing organ shift and deformation. The tools and laparoscope are removed or positioned safely and a CBCT is acquired during an inhale breath hold. Fig. 1, shows the significant difference between the CT and CBCT images. The CT is registered to the CBCT using a non-rigid biomechanically driven registration technique [5]. This registration approach consists of three steps: 1) rigid alignment of the spine, 2) biomechanical insufflation modeling, and 3) diffeomorphic non-rigid registration. The final deformation field can be applied to the pre-operative planning data and models, thus bringing this information into the CBCT coordinate system.

2.2 Registration of the Laparoscope to CBCT Coordinate System

With the CT to CBCT registration complete, the next task is registering the laparoscope to the CBCT coordinate system. This is challenging due to the lack of cross modality landmarks and the camera's small field of view. A two step registration is proposed- an initial position estimation and local refinement.

The initial position of the laparoscope in the CBCT coordinate system is estimated using fluoroscopic images. A mechanical device holds the laparoscope in position and two mono fluoroscopic images are acquired, each 90 ° apart. A semi-automated method is used to select two points along the shaft which are triangulated to estimate the laparoscope's position and pose with 5 degrees of freedom. The rotation around the laparoscope's optical imaging axis is not estimated due to its symmetrical appearance in the fluoroscopic images. Furthermore, the physical position of the camera center along the shaft is not known, and this introduces additional errors.

A local registration refinement is performed directly between the laparoscopic images and the 3D surface model of the organ in the CBCT coordinate system. At this point in the surgical workflow the patient is not at breath hold. Their breathing is periodic and controlled by a ventilator. This respiration causes the abdominal tissue to deform periodically. The first challenge, therefore, lies in the registration of the laparoscopic images to a 3D model representing the tissue at an inhale breath hold. Registering to any other point in the respiration cycle would introduce error into the system.

[2] `www.visiblepatient.com`

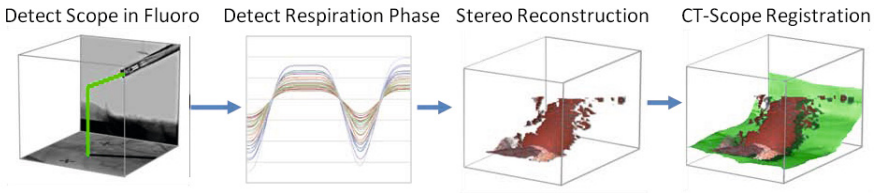

Fig. 2. Registration of laparoscope to CBCT coordinate system

The temporal motion of the tissue in the laparoscopic images is used to estimate the current point in the respiration cycle. Features are detected on the tissue surface and matched in the left and right stereo laparoscope images to estimate their 3D position relative to the camera. The 3D features are transformed into CBCT space using the initial laparoscope alignment and features which are not position near the liver are removed. The features are tracked from frame to frame and their 3D position is computed. Principal Component Analysis (PCA) is applied to extract a 1D respiration signal from the 3D motion of the features [9]. The first component corresponds to respiration, this data is smoothed using a moving average filter to obtain a 1D respiration signal for each feature.

The maximum inhalation position is estimated by fitting a respiration model

$$z(t) = z_0 - bcos^{2n}(\frac{\Pi t}{\tau} - \phi) \tag{1}$$

where z_0 is the position of the liver at the exhale, b is the amplitude, τ is the respiration frequency, ϕ is the phase and n describes the gradient of the model and is empirically set to 4. The parameters of Eq. 1 are estimated using Levenberg-Marquardt minimization algorithm. Before the model is fit, outliers are removed by applying RANSAC to the orientation of the PCA transformation and thresholding the periodicity of the respiration signal which corresponds to τ and ϕ. The remaining inliers are averaged and the model parameters are estimated to identify the point in the respiration cycle corresponding to maximum inhale.

Given the initial estimate of the laparoscope's position and the point in the respiration cycle, the final step remains to perform the direct registration between stereo images and the 3D model. A 3D-3D registration aligns a stereo reconstruction [4] to a point set extracted from the 3D model surface. This point set is extracted using the initial estimate of the laparoscope's position from the previous step, the camera's intrinsic parameters, and z-buffering.

The accurate registration of the 3D model point set and the stereo reconstruction is challenging. At a macro level the point sets represent the same shape, however at a local level they are structurally different because of the way the point sets are generated. The 3D model is continuous, smooth and isotropic. The stereo reconstruction is discretized, contains steps due to pixel level disparity estimates, is anisotropic and may not be a complete surface representation. As a result, even after correct alignment it is impossible to get an exact match for

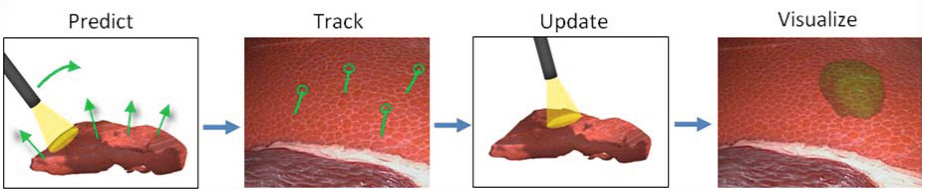

Fig. 3. Temporal registration of laparoscope and tissue

each point. This can cause point-to-point algorithms such as Iterative Closest Point (ICP) to converge a sub-optimal solution as shown in [7].

A probabilistic approach is used [14] that models noise in both the target and source point sets. It makes use of the underlying surface structure while remaining computationally efficient by combines point-to-point and point-to-plane ICP in a single framework. The goal is to align two point sets $A = \{a_i\}_{i=1,\ldots,n}$ and $B = \{b_i\}_{i=1,\ldots,n'}$. The proposed approach replaces the traditional ICP minimization step $\mathbf{T} \leftarrow \text{argmin}_{\mathbf{T}} \sum_i \{\mathbf{T}b_i - m_i\|^2\}$ which finds the optimal transformation $\mathbf{T}$ between point b_i and m_i (the closest corresponding point in A) with

$$\mathbf{T} = \underset{\mathbf{T}}{\text{argmin}} \sum_i \left\{ d_i^{(\mathbf{T})^\top} \left(C_i^B + \mathbf{T} C_i^A \mathbf{T}^\top \right)^{-1} d_i^{\mathbf{T}} \right\} \tag{2}$$

where $d_i^{\mathbf{T}} = b_i - \mathbf{T}a_i$ and C_i^A and C_i^B are the covariance matrices used to model noise in the system. By setting high covariance along the local plane and a low covariance along the surface normal, the registration algorithm is guided to use the surface information in both the 3D model point set and the stereo reconstruction point set. The stereo point set is a subset of the 3D model point set. A maximum correspondence distance is empirically set to account for the fact that some points do not have matches.

2.3 Temporal Alignment

Section 2.2 outlined an approach for registering the laparoscope to the CBCT system where the laparoscope is static and the tissue is temporally static, i.e. at maximum inhale. However, during abdominal surgery, tissue and organs are continuously deforming and the surgeon is free to move the laparoscopic camera.

The position of the laparoscopic camera and tissue deformation are jointly estimated using a modified Simultaneous Localization and Mapping (SLAM) technique [9]. This approach models the position and orientation of the camera in conjunction with a dynamic 3D tissue model which is driven by a respiration model. Within an Extended Kalman Filter (EKF) framework the state vector $\hat{x}$ is comprised of the camera position r^W, its orientation R^{RW}, translational velocity v^W and angular velocity w^R and the respiration model parameters estimated in section 2.2 $\{z_0, b, \tau, \phi\}$. In addition, for each feature, the state contains $\hat{y}_i = (\bar{y}, eig)$ where $\bar{y}$ is the average 3D position of the feature and eig is the PCA

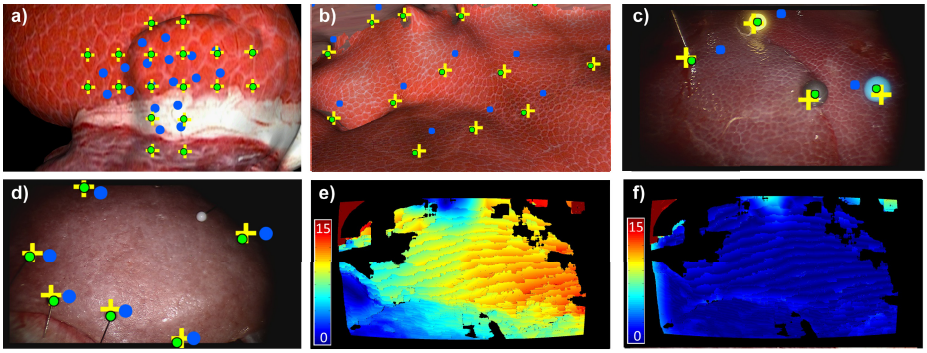

Fig. 4. Laparoscope to CBCT registration: Fiducials shown in green (ground truth), blue (before registration), yellow (after registration). a) non medical, b) sim, c) *ex vivo*, d) phantom. *In vivo* SRE (mm) e) before and f) after registration.

transformation. As shown in Fig. 3, the system iterates between prediction and update steps to estimate the camera's position and tissue deformation. Further details can be found in [9].

The SLAM algorithm initalization follows the registration in section 2.2. As a result, the 3D SLAM features are co-registered to the CBCT coordinate system. In the subsequent image frames, computing the transformation, using singular value decomposition, between the feature positions at time t and time 0 yeilds the estimated 3D model position.

3 Experiments and Results

A range of experiments were performed to validate the proposed framework on simulated, phantom, *ex vivo*, *in vivo* and non medical data. The phases of the pipeline are evaluated separately here, both for clarity and because not all data contain temporal deformation. CT to CBCT obtains accuracy of <1mm on liver, due to space constraints the reader is directed to [5] for evaluation. A description of the datasets follows. **Simulated**: a mesh generated form a CT and textured with laparoscopic images. **Phantom**: a visually realistic silicon liver phantom with surface fiducials for ground truth. ***Ex vivo***: porcine with fiducials for ground truth. ***In vivo***: two porcine without fiducials. **Non Medical**: meshes from Stanford dataset[3] textured with laparoscopic images.

Registration of laparoscopic camera to CBCT. 50 datasets with ground truth were available- simulated (20), phantom (10), *ex vivo* (10) and non medical (10). Random noise (Up to ±20mm) was added to the initial position of the laparoscope in the CBCT system to quantitatively evaluate the registration. 10 noisy datasets were created for each ground truth dataset, making a total of 500 datasets. 11 *in vivo* datasets were evaluated without ground truth fiducials. The results are shown in Table 1 and illustrated in Fig. 4.

[3] `http://graphics.stanford.edu/data/3Dscanrep#bunny`

Table 1. Quantitative validation: Registration of laparoscope to CBCT

Dataset	SRE Before	SRE After	TRE Before	TRE After
Sim	5.3mm	0.8mm	10.4mm, 289.9 px	1.69mm, 56.8 px
Phantom	5.7mm	1.1mm	10.2mm, 90.5 px	4.1mm, 29.9 px
Ex vivo	4.7mm	1.3mm	10.28mm, 136.5 px	3.4mm, 48.7 px
In vivo	5.4mm	0.9mm	N/A	N/A
Non Medical	5.5mm	0.9mm	10.2mm, 321.2 px	0.3mm, 10.6 px

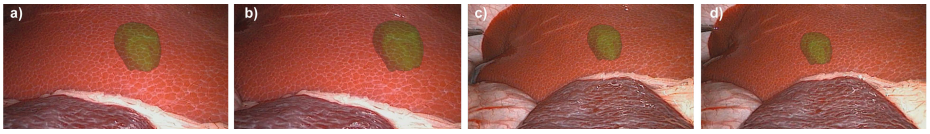

Fig. 5. Augmented reality overlay of a virtual tumor for intra-operative planning

The metrics Surface Registration Error **(SRE)** $= \frac{1}{n}\sum_{i=0}^{n}\{\sqrt{(a_i - m_i)^2}\}$ and Target Registration Error **(TRE)** $= RMSError(Fiducials1 - Fiducials2)$ are used for evaluation. The registration refinement process reduces the TRE for all datasets converging to results of between 0.3-4.1mm. The phantom data has the largest error which is attributed to its homogenous shape. Additional errors may be introduced by manual fiducial annotation. The 2D TRE is dependent on the proximity of the fiducials to the camera and image size. The *in vivo* and *ex vivo* image size is 1280x720 and all others are 1920x1080. The 2D TRE is visualized in Fig. 4. Fig. 4a) shows a successful registration where the added noise is 10° around the optical axis and 10mm along the optical axis. The registration reduces the SRE for all datasets. Fig. 4 e-f) show the SRE for *in vivo* data before and after registration with Fig. 4 f) demonstrates a converged registration. Stereo reconstruction takes 5.1s and registration takes 7.2s however, the proposed surgical workflow does not requrie these step to be real-time.

Temporal registration was quantitatively evaluated on 20 simulated and five *in vivo* datasets. Simulated data was generated by applying a realistic

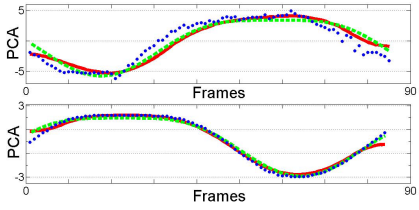

Fig. 6. Respiration model. *in vivo* (top), sim (bot). 1D respiration signal (b), smoothed data (r), model(g).

Table 2. Quantitative evaluation of temporal registration of laparoscopic images

Dataset	TRE Error 3D	Camera Position 3D
Sim	3.6mm	1.9 mm
In vivo	n/a	4.1 mm

biomechanical deformation to the organ model and moving the camera. Evaluation with respect to TRE and camera position are shown in Table 2. For *in vivo* data ground truth was obtained by annotating the position of the scope in fluoro images at the start and end of each sequence. The annotation contains absolute positional errors in the CBCT coordinate system but it can be considered accurate relative to the camera coordinate system. The results are shown in Table 2. Qualitative validation is provided for *in vivo* data in Fig. 5 where a segmented virtual tumor is augmented. This illustrates the accurate estimation of the camera's position and the point in the respiration cycle. The respiration models are visualized in Fig 6. Temporal registration runs at 15fps.

4 Conclusion

In this paper, an augmented reality framework for intra-operative planning is proposed which co-registering pre-operative CT to laparoscope images. It does not require fiducials, manual model alignment and accounts for camera motion and tissue deformation. The framework has been validated on simulated, phantom, *ex vivo* (porcine), *in vivo* (porcine) and non medical data. Future work will focus on improving computational efficiency and more complex tissue modelling.

References

1. Hughes-Hallett, A., Mayer, E.K., Marcus, H.J., Cundy, T.P., Pratt, P.J., Darzi, A.W., Vale, J.A.: Augmented reality partial nephrectomy: Examining the current status and future perspectives. Urology, 266–273 (2013)
2. Allard, J., Cotin, S., Faure, F., Bensoussan, P.J., Poyer, F., Duriez, C., Delingette, H., Grisoni, L.: Sofa-open framework for medical simulation. In: MMVR (2007)
3. Collins, T., Bartoli, A.: Towards live monocular 3d laparoscopy using shading and specularity information. In: Abolmaesumi, P., Joskowicz, L., Navab, N., Jannin, P. (eds.) IPCAI 2012. LNCS, vol. 7330, pp. 11–21. Springer, Heidelberg (2012)
4. Stoyanov, D., Scarzanella, M.V., Pratt, P., Yang, G.-Z.: Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 275–282. Springer, Heidelberg (2010)
5. Oktay, O., Zhang, L., Mansi, T., Mountney, P., Mewes, P., Nicolau, S., Soler, L., Chefd'hotel, C.: Biomechanically driven registration of pre- to intra-operative 3D images for laparoscopic surgery. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part II. LNCS, vol. 8150, pp. 1–9. Springer, Heidelberg (2013)
6. Puerto Souza, G.A., Adibi, M., Cadeddu, J.A., Mariottini, G.L.: Adaptive multi-affine (AMA) feature-matching algorithm and its application to minimally-invasive surgery images. In: IROS, pp. 2371–2376 (2011)
7. Maier-Hein, L., Franz, A., dos Santos, T., Schmidt, M., Fangerau, M., Meinzer, H., Fitzpatrick, J.: Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. PAMI 34(8), 1520–1532 (2012)

8. Mirota, D., Uneri, A., Schafer, S., Nithiananthan, S., Reh, D., Ishii, M., Gallia, G., Taylor, R., Hager, G., Siewerdsen, J.: Evaluation of a system for high-accuracy 3D image-based registration of endoscopic video to c-arm cone-beam CT for image-guided skull base surgery. Transactions on Medical Imaging 32, 1215–1226 (2013)
9. Mountney, P., Yang, G.Z.: Motion compensated slam for image guided surgery. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part II. LNCS, vol. 6362, pp. 496–504. Springer, Heidelberg (2010)
10. Nicolau, S.A., Pennec, X., Soler, L., Buy, X., Gangi, A., Ayache, N., Marescaux, J.: An augmented reality system for liver thermal ablation: Design and evaluation on clinical cases. Medical Image Analysis 13(3), 494–506 (2009)
11. Teber, D., Guven, S., Simpfendrfer, T., Baumhauer, M., Gven, E.O., Yencilek, F., Gzen, A.S., Rassweiler, J.: Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? preliminary in vitro and in vivo results. European Urology 56(2), 332–338 (2009)
12. Su, L.M., Vagvolgyi, B.P., Agarwal, R., Reiley, C.E., Taylor, R.H., Hager, G.D.: Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3D-CT to stereoscopic video registration. Urology 73(4), 896–900 (2009)
13. Pratt, P., Mayer, E., Vale, J., Cohen, D., Edwards, E., Darzi, A., Yang, G.Z.: An effective visualisation and registration system for image-guided robotic partial nephrectomy. Journal of Robotic Surgery 6(1), 23–31 (2012)
14. Segal, A., Haehnel, D., Thrun, S.: Generalized-ICP. In: RSS, p. 4 (2009)