



Framework for augmented reality in Minimally Invasive laparoscopic surgery

Bruno Marques, Rosalie Plantefeve, Frédérick Roy, Nazim Haouchine,
Emmanuel Jeanvoine, Igor Peterlik, Stéphane Cotin

► To cite this version:

Bruno Marques, Rosalie Plantefeve, Frédérick Roy, Nazim Haouchine, Emmanuel Jeanvoine, et al.. Framework for augmented reality in Minimally Invasive laparoscopic surgery. HealthCom 2015, Oct 2015, Boston, United States. 10.1109/HealthCom.2015.7454467 . hal-01315574

HAL Id: hal-01315574

<https://inria.hal.science/hal-01315574>

Submitted on 13 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Framework for Augmented Reality in Minimally Invasive Laparoscopic Surgery

Bruno Marques
Frédéric Roy
Nazim Haouchine
Emmanuel Jeanvoine
and Stéphane Cotin
Inria (Mimesis Team)
Strasbourg France
Email: stephane.cotin@inria.fr

Rosalie Plantefève
ALTRAN Medic@ and
Inria (Mimesis Team)
Strasbourg
France
Email: rosalie.plantefeve@altran.com

Igor Peterlik
Masaryk University
Brno
Czech Republic
Email: peterlik@gmail.com

Abstract—This article presents a framework for fusing pre-operative data and intra-operative data for surgery guidance. This framework is employed in the context of Minimally Invasive Surgery (MIS) of the liver. From stereoscopic images a three dimensional point cloud is reconstructed in real-time. This point cloud is then used to register a patient-specific biomechanical model derived from Computed Tomography images onto the laparoscopic view. In this way internal structures such as vessels and tumors can be visualized to help the surgeon during the procedure. This is particularly relevant since abdominal organs undergo large deformations in the course of the surgery, making it difficult for surgeons to correlate the laparoscopic view with the pre-operative images. Our method has the potential to reduce the duration of the operation as the biomechanical model makes it possible to estimate the in-depth position of tumors and vessels at any time of the surgery, which is essential to the surgical decision process. Results show that our method can be successfully applied during laparoscopic procedure without interfering with the surgical work flow.

I. INTRODUCTION

In the last decades, advances in medicine have seen the emergence of MIS. In this procedure, the surgeon manipulates the organs with instruments inserted through trocars placed in small abdominal incisions. The view of the operating field is provided by a laparoscopic camera, inserted through another port in the abdomen, allowing the surgeon to perform the surgery by watching a high-definition monitor placed above the patient. MIS provides real benefits to the patient, by reducing pain, bleeding and risks of infection and therefore also shortening recovery time. However, since the manipulation of instruments is indirect and the visual feedback is limited, it remains quite challenging from a surgical standpoint and increases the duration of operations. Allowing surgeons to see through organs surface and to visualize cutting planes defined during planning facilitate the orientation and navigation during the surgery. For this purpose we need to register the pre-operative data onto the laparoscopic view. The objective is to reduce the operating time, to locate tumors with greater accuracy and preserve vessels which are needed for the regeneration of the liver tissue post-operatively.

The interest of the research community in surgical planning tools have been growing in the last decade [1]. The

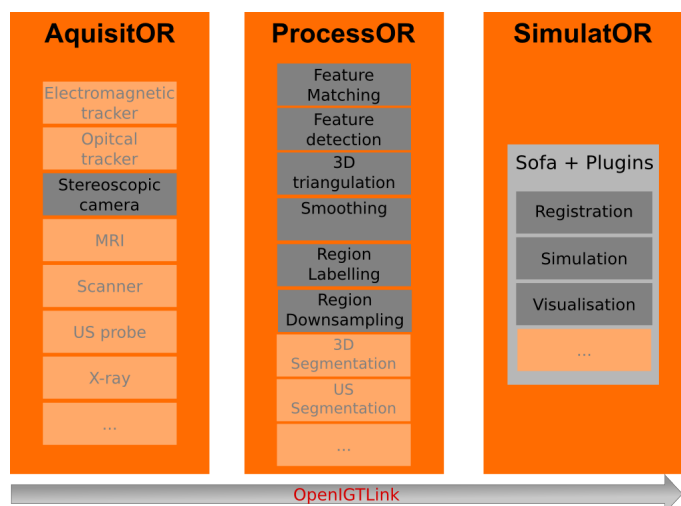


Fig. 1. Our framework architecture.

benefit of such tools are obvious in liver surgery where the surgeons must resect tumors while preserving the vascular tree. However, during MIS the abdominal organs undergo important deformation and the best planning tool may become useless to surgeons. Therefore intraoperative navigation system have to be used in conjunction with the planning tools. Nevertheless, most existing systems used for navigation are limited to rigid registration [2], [3]. Hence, these systems do not provide real time tracking of organs deformation and the position of internal structure may not be accurate. In [4] the authors propose to use C-arm system to align the intra-operative and pre-operative data. Although this method provides accurate registration, it relies on intraoperative scans, which are currently not available during clinical routines. The use of Ultrasound (US) intraoperative data seems promising [5] but still required additional hardware. To reduce the impact on the usual clinical workflow the use of the stereo-endoscopic camera as the intraoperative data source for MIS is optimal.

Nonetheless, processing stereoscopic data is computationally expensive. A solution is proposed in [6] the author suggest to distribute processes through different computers. In one

dimensional shape of the liver from intra-operative images can be performed in different ways depending on the image modality [11]. Let us assume the stereoscopic pair of images \mathcal{I}_l and \mathcal{I}_r . We extract from this stereo pair points of interest (i.e. features) that are sufficiently reliable for the 3D reconstruction. Several detectors have been reported in the literature (e.g. [12]), we employ the *Speeded-up Robust Features* (SURF); our choice is motivated by its particular suitability for robotic-guided endoscopy applications [13]. When using this detector on the image pair $(\mathcal{I}_l, \mathcal{I}_r)$ with an appropriate threshold, we obtain two sets of features $\mathbf{E}_l = (x_{l_1}, \dots, x_{l_n})$ and $\mathbf{E}_r = (x_{r_1}, \dots, x_{r_m})$ where it is necessary to estimate for each feature $x_i = (u_i, v_i)$ the 3D point $X_i = (U_i, V_i, W_i)$. This is done by establishing a correspondence between image points $x_l \longleftrightarrow x_r$, using a descriptor-based matching method [14]. Once a correspondence is found, a sparse set of m 3D points, denoted $\mathbf{T} = (X_1, \dots, X_m)$, is reconstructed using triangulation algorithm [15] and a surface mesh S is interpolated using Moving Least-Square surface approximation [16], [17]. Two different point cloud are using for our pipeline. One for the initial registration witch is labelled as explained in the following section and a down-sampled point cloud without label for the temporal registration.

Labeling In order to improve the accuracy and robustness of the initial registration process, we augment the pre-operative liver model and the intraoperative point cloud with anatomical features. We label corresponding areas on both the pre-operative mesh and the point cloud to ensure that this point cloud, corresponding to the visible part of the liver, is accurately matched on the pre-operative surface mesh. As in [18], two anatomical features are used as landmarks: the umbilical notch where the right margin of falciform ligament is attached to the liver and the anterior margin of the liver.

For the pre-operative liver model the anterior margin is automatically detected using a curvature analysis. Edges separating two triangles with sufficiently different directions of normals are selected as seeds. Then, a ridge line is extended from the seed edges; if no extension is found, the seeds are removed. The definition of the two thresholds, one for the seeds and one for the extensions, is based on statistics on the normals for the whole mesh. Iteratively, the anterior margin is reconstructed. Depending on the image modality, the umbilical notch is either difficult or simply impossible to segment from the pre-operative data. To solve this issue, these features are transferred onto the three-dimensional mesh via a statistical atlas hand-built by experts. Details regarding the construction of the atlas are described in [19].

On intraoperative images the labels are set manually by clicking around the region of interest. Three point clouds are then reconstructed: one for the anterior margin, one for the umbilical notch and one for the rest of the visible surface of the liver.

Registration For the initial registration, we employ the iterative process described in [18]. The biomechanical model is used to simulate the behaviour of the tissue while the registration constraints imposed to the deformable object are modelled with penalty forces \mathbf{f}_{ext} . Each step of the registration is decomposed in three parts: (i) matching between the target points and their projections, (ii) definition of external forces

and (iii) resolution of the system:

$$\mathbf{K}(\mathbf{u})\mathbf{u} = \mathbf{f}_{ext} \quad (1)$$

where \mathbf{u} is the displacement vector and \mathbf{K} the non-linear stiffness provided by the FE formulation. Each labeled point p_{ti} of the target point cloud is projected onto its closest point p_{si} in the corresponding labeled area on the source surface to obtain a set of matching pairs $\{p_t, p_s\}_i$. For each pair we define a penalty force:

$$\mathbf{f}_{ext} = \sum_i^m \mathbf{f}_i, \quad (2) \quad \text{with} \quad \mathbf{f}_i = \sum_j^{p_i} f_i \arctan(k\mathbf{x}_{ij}) \quad (3)$$

where m is the number of labelled area, p_i the number of corresponding pairs in the area i , f_i a penalty factor, k a scale factor and \mathbf{x}_{ij} the vector between the j -th target point and its projection. The k value is related to the point cloud noise and is used to avoid over-fitting as a lower value decrease the force intensity for small distances. The f_i are different among the three labelled area and evolve during the registration process. At the beginning of the process the f_i of the umbilical notch and the anterior margin are greater than the one of the remaining surface. This helps the matching to converge toward the accurate solution. If the system reaches an equilibrium, the value of all f_i are increased unless they exceed their maximum value: $\max(f_i)$. Typically, we have $\max(f_{un}) \leq \max(f_{am}) \leq \max(f_{surf})$ where un stand for the umbilical notch, am for the anterior margin and $surf$ for the remaining surface.

After attaining convergence, the temporal registration step is performed. The temporal registration permits to register the biomechanical model at each frame acquired from the endoscopic camera, following the motion of the liver (which moves due to the heart beating and/or respiratory motion). The motion of the liver is captured using an optical flow algorithm [20] that tracks image-points from the liver surface based on a brightness consistency constraint. The combination of optical flow and SURF features has proven to be robust to track heart motion in laparoscopic images, where real-time is needed [13] and was successfully translated to liver motion estimation [21]. Indeed, at this stage, a real-time simulation is mandatory since video frames are acquired at least at 30 Hz and the registration w. r. t. the n -th frame must to be performed before the frame $n+1$ is available.

B. Framework architecture

The implementation described in this paper is divided into three distinct processes: AcquisitOR, ProcessOR, and SimulatOR (see Fig.1). This architecture allows to distribute the workload on different machines without suffering from high latency as long as it runs on a Local Area or ad Hoc Network. AcquisitOR retrieves data from specialized hardware devices, ProcessOR processes this data for further usage in the simulation, and SimulatOR runs the simulation according to the patient-specific data retrieved and processed in the previous steps.

Using a standard protocol to perform communications between processes is very important since it allows third party programs to use this solution. Furthermore, it allows to replace some parts of the processing line with external tools. The solution presented here implements the Open Network

Interface for Image-Guided Therapy (OpenIGTLink) protocol. Therefore, instead of implementing multiple drivers for the many devices possibly present in the OR, the Plus Toolkit [22] can be used to send data to ProcessOR. It would also be possible to use 3DSlicer [23] or IGSTK [24] to visualize the processed data generated from ProcessOR, as well as any software relying on the OpenIGTLink protocol to exchange data through the network.

AcquisitOR is the process responsible for connecting all the available devices in the Operating Room, and stream their data to ProcessOR. It streams data received from medical devices to a server for processing. One instance of AcquisitOR is launched per IGT device, allowing distributing the workload on multiple cores or workstations.

ProcessOR has the charge of processing the data retrieved from the OR. It offers a generic API allowing the execution of different filters to process the data retrieved from an IGT client. ProcessOR offers an API to develop operation-specific programs. Specific programs will have to be developed according to the different use cases, as they depend on what is taken as input, and what is needed as an output for SimulatOR.

SimulatOR is based on the Simulation Open Framework Architecture (SOFA). It allows us to easily create simulated scene where different types of objects can interact with each other. Several state-of-the-art algorithms necessary for realistic simulations are implemented. Bio-mechanical behavior modeling is especially necessary, as deformations will be a key part of this project. Finally, extending and implementing new algorithms and methods are possible through plugins mechanism. It allows us to easily add components to include the features needed for our project, e.g augmented reality rendering.

C. Hardware setup

We present in this section the hardware setup used for our application: intraoperative guidance during laparoscopic liver surgery. An overview of this setup is presented in Fig.3. Our approach relies on a 3D point extraction from stereoscopic frames, and therefore, a calibrated stereoscopic camera is required. We chose to interface our framework with Da Vinci's surgical robot, but a stereoscopic camera from Storz like the Tipcaml could be used instead of the Da Vinci's endoscope. In order to retrieve the video streams from the Da Vinci console, we need two video editing cards (Blackmagic Intensity shuttle). We also need two video recorders to retrieve the raw video streams from the operation, for research and post operation processing purposes (Blackmagic Hyperdeck shuttle). The video editing cards are connected to a MacBook Pro 13' as they are connected through thunderbolt. This very compact setup allows us to go in the operating room without taking too much space since operating rooms are very confined rooms, containing sometimes more than 20 persons along with the Da Vinci robot.

D. Implementation

AcquisitOR

The video streams fetched from the stereoscopic endoscope are the input of our augmented reality pipeline. To retrieve

these streams from intuitive's Da Vinci endoscope, the Left and Right video output from the Da Vinci are connected to video editing cards (Blackmagic intensity). These video editing cards have no driver implementation in Plus Toolkit. Therefore, it was relevant to implement our own device acquisition server that would retrieve the video streams from the video editing card and convert its raw data into computer-readable RGB frames.

In order to connect the video editing card to AcquisitOR, the Blackmagic driver is implemented as a separate library, loaded at runtime according to the arguments passed to the command line.

ProcessOR

In the case of liver registration, our ProcessOR needs to generate a point cloud from the acquired stereoscopic frames.

Extracting a good point cloud from endoscopic frames is a complicated matter, as image processing relies on the quality of the frames acquired, the quality of the environment inside the abdominal cavity (lighting, smoke blurry frames, dirty lenses etc.), and the quality of the liver's surface (if too smooth, the extraction will perform badly). Thus, we need to apply many filters (feature detection and matching, triangulation, region labeling, downsampling etc.).

Each of these filters require specific settings, which is why we decided to rely on a semi-automatic point extraction process that allows a manual tuning through a graphical user interface of the many settings that needs to be taken into consideration during a point cloud extraction. This interface also allows us to manually label each region of the image.

As an output, a labeled point cloud is sent to SimulatOR, through OpenIGTLink.

SimulatOR

It combines pre-operative data (segmented liver, potential tumors, and vascular network) we got previously with a CT-Scan, with per-operative data (point clouds extracted by ProcessOR) through registration algorithms. The registration algorithm needs a realistic mechanical behavior to perform a correct registration. Thus, we are using a corotational tetrahedral finite element model for the biomechanical behavior of the liver and its inner components. Moreover, constraints are already included, allowing us to implement the complete registration algorithm quickly.

III. RESULTS

We implemented a framework to interface medical devices with medical simulation using SOFA, named SOFA-OR.

This framework is designed to allow the development of a complete software pipeline and has been tested through a demonstrator for augmented reality provided during minimally invasive laparoscopic surgery.

A command-line AcquisitOR provides a robust and flexible interface with the Operating Room, retrieving and streaming HD stereoscopic frames in real time, as well as any data that could potentially come from the OR, through the standard OpenIGTLink protocol.

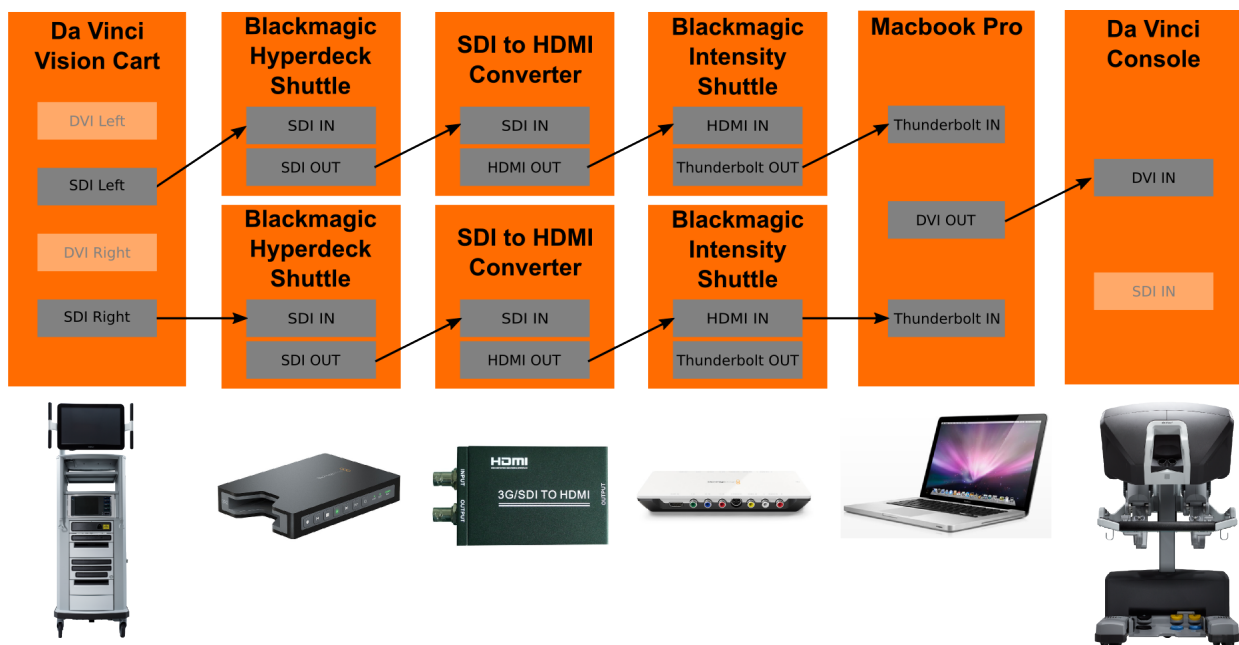


Fig. 3. Hardware setup for Per-operative Liver registration.

Our ProcessOR provides a complete graphical user interface to which the stereo stream is sent through a Local Area Network in real time with OpenIGTLink, to precisely tune image processing filters, manually segmentate regions, extract labeled point clouds and stream them over OpenIGTLink.

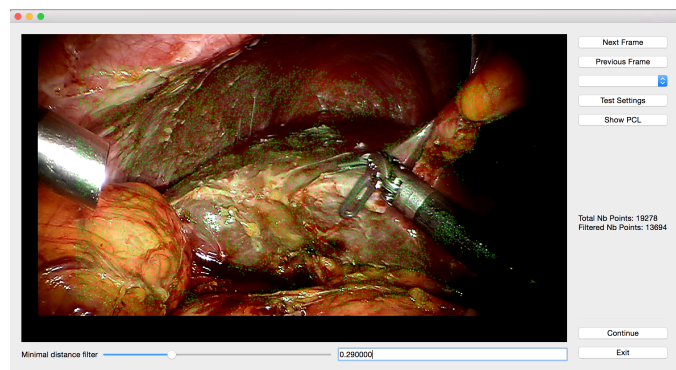


Fig. 4. ProcessOR's manual mask edition window

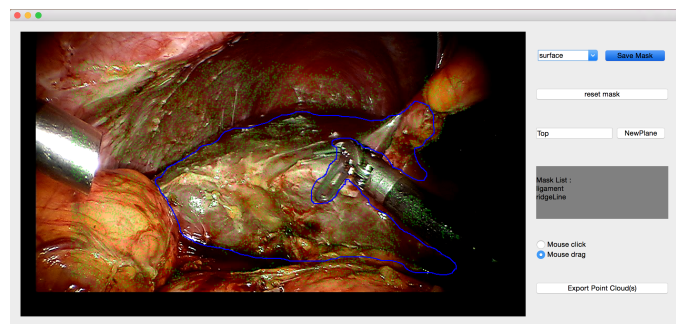


Fig. 5. ProcessOR's overview of the 3D point cloud

The SimulatOR benefits from SOFA's accuracy and stabil-

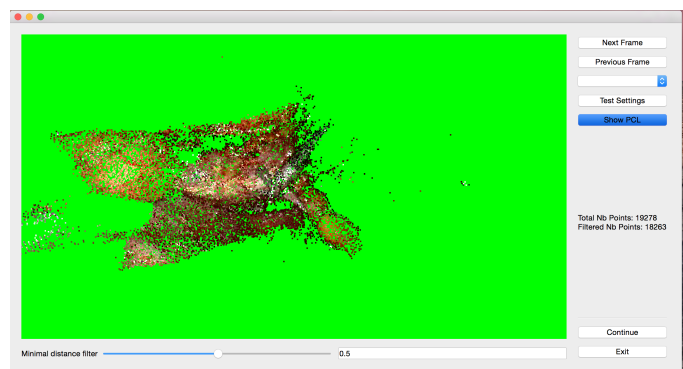


Fig. 6. SimulatOR's augmented view of the registered liver and its internal structures

ity to implement a complex simulation of the patient-specific model of the liver, and renders through OpenGL an augmented view of the operating field that is sent back to the surgeon's console.



Fig. 7. Final result.

The whole pipeline has been tested on human patients data

during laparoscopic hepatectomies. Results are presented in Fig.4, 5, 6, 7. It has demonstrated that our framework was able to perform in real time the point cloud reconstruction and to display the result of the registration onto the laparoscopic view. However, as the ground truth is unknown we do not have pertinent quantitative results for the real cases. Therefore, we performed *in silico* experiments for the initial registration in [18] as well as phantom experiments for the temporal registration in [17].

One of the current limitations of our algorithm concerns the loss of points during a persistent tracking. If too many points are lost because of an obstruction caused by smoke or surgical tools, the registration quality also decreases. Also, our method does not account for topological changes but recent results on this topic have been acquired [25].

CONCLUSION

This article presents a complete framework for simulation-assisted IGT procedures. It implements a standard protocol (OpenIGTLink) ensuring interoperability between the framework and potential third party softwares, as well as providing a standard interface with medical devices. It allows distribution on multiple computational units over a Local Area Network for better performances, and relies on the robust, open and largely used simulation framework SOFA.

We applied successfully this framework in the context of liver MIS guidance in Operating Room.

In future work we plan to further test our framework in different context and to carry out more advanced analysis on its performances. To consider a systematic use in OR where our framework would be embedded in a medical device, special efforts will also have to be carried out in order to conform the development process to a specific certification process, like IEC 62304 for instance.

REFERENCES

- [1] "Mint liver," <http://www.mint-medical.de/productsolutions/mintliver/mintliver/>, accessed: 2015-07-15.
- [2] "Cascination," <http://www.cascination.com/>, accessed: 2015-07-15.
- [3] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, "Augmented reality in laparoscopic surgical oncology," *Surgical Oncology*, vol. 20, no. 3, pp. 189 – 201, 2011.
- [4] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, "Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 3, pp. 355–369, 2008.
- [5] T. Langø, S. Vijayan, A. Rethy, C. Våpenstad, O. V. Solberg, R. Mårvik, G. Johnsen, and T. N. Hernes, "Navigated laparoscopic ultrasound in abdominal soft tissue surgery: technological overview and perspectives," *International journal of computer assisted radiology and surgery*, vol. 7, no. 4, pp. 585–599, 2012.
- [6] S. Suwelack, D. Katic, S. Wagner, P. Spengler, S. Bodenstedt, S. Rhl, R. Dillmann, and S. Speidel, "Lightweight distributed computing for intraoperative real-time image guidance," D. R. Holmes III and K. H. Wong, Eds., Feb. 2012, p. 831621. [Online]. Available: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.911404>
- [7] T. Sielhorst, M. Feuerstein, J. Traub, O. Kutter, and N. Navab, "Campar: A software framework guaranteeing quality for medical augmented reality," *Int. J. Comput. Assist. Radiol. Surg.*, pp. 29–30, 2006.
- [8] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.
- [9] D. Boltcheva, M. Yvinec, and J.-D. Boissonnat, "Mesh generation from 3d multi-material images," in *Proceedings of the 12th MICCAI - Volume Part II*. Springer, 2009, pp. 283–290.
- [10] M. Nesme, Y. Payan, and F. Faure, "Efficient, physically plausible finite elements," in *Eurographics 2005, Short papers, August, 2005*, J. Dingliana and F. Ganovelli, Eds., Trinity College, Dublin, Ireland, 2005, p. 7780.
- [11] L. Maier-Hein, P. Mountney, A. Bartoli, H. Elhawary, D. Elson, A. Groch, A. Kolb, M. Rodrigues, J. Sorger, S. Speidel, and D. Stoyanov, "Optical techniques for 3d surface reconstruction in computer-assisted laparoscopic surgery," *MEDICAL IMAGE ANALYSIS*, vol. 17, pp. 974–996, 2013.
- [12] S. Gauglitz, T. Hllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *International Journal of Computer Vision*, vol. 94, no. 3, pp. 335–360, 2011.
- [13] H. Elhawary and A. Popovic, "Robust feature tracking on the beating heart for a robotic-guided endoscope," *Int J Med Robot.*, pp. 459–468, Oct. 2010.
- [14] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pp. 1615–1630, Oct 2005.
- [15] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [16] N. Haouchine, J. Dequidt, I. Peterlik, E. Kerrien, M.-O. Berger, and S. Cotin, "Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery," in *ISMAR 2013*, pp. 199–208.
- [17] N. Haouchine, I. Peterlik, J. Dequidt, M. Sanz-Lopez, E. Kerrien, M.-O. Berger, and S. Cotin, "Impact of Soft Tissue Heterogeneity on Augmented Reality for Liver Surgery," *IEEE TVCG*, pp. 584 – 597, 2015, 14 pages.
- [18] R. Plantefeve et al., "Automatic alignment of pre and intraoperative data using anatomical landmarks for augmented laparoscopic liver surgery," in *Biomedical Simulation*, ser. LNCS, F. Bello and S. Cotin, Eds., 2014, vol. 8789, pp. 58–66. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-12057-7_7
- [19] R. Plantefeve, I. Peterlik, H. Courtecuisse, R. Trivisonne, J.-P. Radoux, and S. Cotin, "Atlas-based transfer of boundary conditions for biomechanical simulation," in *Proceedings of the 17th MICCAI: Part III*. Springer, 2014, pp. 33–40.
- [20] J. Y. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the algorithm," 2002.
- [21] N. Haouchine, J. Dequidt, I. Peterlik, E. Kerrien, M.-O. Berger, and S. Cotin, "Towards an accurate tracking of liver tumors for augmented reality in robotic assisted surgery," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 2014, pp. 4121–4126.
- [22] A. Lasso, T. Heffter, C. Pinter, T. Ungi, and G. Fichtinger, "Implementation of the plus open-source toolkit for translational research of ultrasound-guided intervention systems," *MICCAI-Systems and Architectures for Computer Assisted Interventions*, pp. 1–12, 2012. [Online]. Available: <http://www.plustoolkit.org>
- [23] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka et al., "3d slicer as an image computing platform for the quantitative imaging network," *Magnetic resonance imaging*, vol. 30, no. 9, pp. 1323–1341, 2012. [Online]. Available: <http://www.slicer.org>
- [24] K. Gary, L. Ibanez, S. Aylward, D. Gobbi, M. B. Blake, and K. Cleary, "Igstk: an open source software toolkit for image-guided surgery," *Computer*, vol. 39, no. 4, pp. 46–53, 2006. [Online]. Available: <http://www.igstk.org>
- [25] C. Paulus, N. Haouchine, D. Cazier, and S. Cotin, "Augmented reality during cutting and tearing of deformable objects," in *The 14th IEEE International Symposium on Mixed and Augmented Reality*, 2015.