Real-time Geometry-Aware Augmented Reality in Minimally Invasive Surgery

Long Chen[1], Wen Tang[1] and Nigel W. John[2]

[1]*Bournemouth University*
[2]*University of Chester*

The potential of Augmented Reality (AR) technology to assist minimally invasive surgeries (MIS) lies in its computational performance and accuracy in dealing with challenging MIS scenes. Even with the latest hardware and software technologies, achieving both real-time and accurate augmented information overlay in MIS is still a formidable task. In this paper, we present a novel real-time AR framework for MIS that achieves interactive geometric aware augmented reality in endoscopic surgery with stereo views. Our framework tracks the movement of the endoscopic camera and simultaneously reconstructs a dense geometric mesh of the MIS scene. The movement of the camera is predicted by minimising the re-projection error to achieve a fast tracking performance, while the 3D mesh is incrementally built by a dense zero mean normalised cross correlation stereo matching method to improve the accuracy of the surface reconstruction. Our proposed system does not require any prior template or pre-operative scan and can infer the geometric information intra-operatively in real-time. With the geometric information available, our proposed AR framework is able to interactively add annotations, localisation of tumours and vessels, and measurement labelling with greater precision and accuracy compared with the state of the art approaches.

1. Introduction: Laparoscopic surgery is a Minimally Invasive Surgical (MIS) procedure using endoscopes with small incisions to carry out internal operations on patients. While MIS offers considerable advantages over open surgeries, it also imposes big challenges on surgeons' performance due to the well known MIS issues associated with the field of view (FOV), hand-eye disalignment and disorientation. Augmented Reality(AR) technology can help overcome the limitations by overlaying additional information with the real scene through augmentation of target surgical locations, annotations [1], labels [2], tumour measurement [3] or even 3D reconstruction of anatomic structures [4] [5].

Despite recent advances in powerful miniaturised AR hardware devices and improvements on vision based software algorithms, many issues in medical AR remain unsolved. Especially, the dramatic changes in tissue surface illumination and tissue deformation as well as the rapid movements of the endoscope during the insertion and the extrusion all give rise to a set of unique challenges that call for innovative approaches.

As with in any technological assisted medical procedures, accuracy of AR in MIS is paramount. With the use of traditional 2D feature based tracking algorithms such as those used in [6] ,[7], [1] and [8], the rapid endoscope movement can easily cause feature points extracted from the vision algorithms to fall out of the field of view, resulting in poor quality visual guidance. The latest visual SLAM (Simultaneous Location and Mapping) based approaches have the potential to overcome this issue by building an entire 3D map of the internal cavity of the MIS environment, but SLAM algorithms are often not robust enough when dealing with tissue deformations and scene illuminations [9] [10] [10] [11] [12]. Furthermore, in order to meet the demand of high computational performance, sparse landmark points are often used in MIS AR, and augmented information are mapped using planar detection algorithms such as Random Sample Consensus (RANSAC) [13] [14]. As a result, AR content is mapped onto planes rather than curved organ surfaces.

In this paper, we introduce a novel MIS AR framework that allows accurate overlay of augmented information onto curved surfaces, i.e. accurate annotation and labelling, and tumour measurements along the curved soft tissue surfaces. There are two key features of our proposed framework: 1) real-time computation of robust 3D tracking through a robust feature-based SLAM approach; 2) an interactive geometry-aware AR environment through incrementally building a geometric mesh via zero mean normalised cross correlation (ZNCC) stereo matching.

2. Related Work: The traditional MIS AR approaches usually employ feature points tracking methods for information overlay. Feature based 2D tracking methods such as Kanade-Lucas Tomasi (KLT) features [6] [7], Scale-Invariant Feature Transform (SIFT) [1], Speeded Up Robust Features (SURF) [15], even those methods specifically designed to cater for the scale, rotation and brightness of soft tissue [8] have a major drawback for AR, because selected feature points extracted from vision algorithms must be within the field of view. Therefore, traditional feature tracking methods can severely affect the precision of procedure guidance, especially in surgical scenes where the accuracy is paramount.

Recently, SLAM (Simultaneous localization and Mapping) algorithms have led to new approaches to endoscopic camera tracking in MIS. Originally designed for robot navigation in unknown environments, the algorithms can be adapted for tracking the pose of endoscopic cameras while simultaneously building landmark maps inside the patient body during MIS procedures. SLAM-enabled AR systems not only improve the usability of AR in MIS due to no optical or magnetic tracking devices to obstruct the surgeons' view, but they also offer greater accuracy and robustness compared with traditional feature-based AR systems.

Direct-based SLAM algorithms compare pixels [16] or reconstructed models [17] [9] of two images to estimate camera poses and reconstruct a dense 3D map by minimising the photometric errors. However, direct methods are more likely to fail when dealing with deformable scenes or when the illumination of the scene is inconsistent. Feature-based SLAM systems [10] [11] only compare a set of sparse feature points that are extracted from images. These methods estimate camera poses by minimising the re-projection error of the feature points. Therefore, feature-based SLAM methods are more suitable for MIS scenes due to it's tolerance to illumination changes and small deformations. Feature-based SLAM such as EKF-SLAM has been used with laparoscopic image sequences [18] [19] [14] and a further motion compensation model [20] and stereo semi-dense reconstruction method [21] were integrated into the EKF-SLAM framework to deal with periodic deformation. However, the accuracy of EKF-SLAM tracking is not guaranteed and prone to inconsistent estimation and drifting due to the linearization of motion and sensor models approximated by a first-order Taylor series expansion.
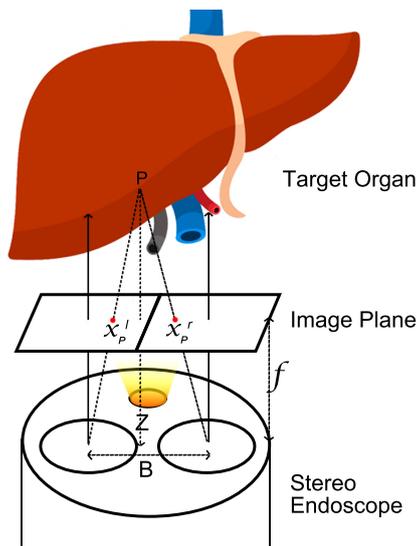
Figure 1 *By using a stereo endoscope, the 3D position of any point in the view can be directly estimated by using stereo triangulation.*

The first keyframe-based SLAM – PTAM (Parallel Tracking and Mapping) [10] was a breakthrough in visual SLAM and has been used in MIS for stereoscope tracking [13]. The extension of PTAM – ORBSLAM [11] has also been tested on endoscope videos with map point densifying modifications [12], but the loss of accuracy still exists. Furthermore, since feature-based SLAM systems can only reconstruct maps based on sparse landmark-points that barely describe the detailed 3D structure of the environment, the augmented AR content has to be mapped onto a plan through planar detection algorithms such as Random Sample Consensus (RANSAC) [13]. Although feature-based SLAM is computationally efficient, different to real-life environments, in MIS scenes, flat surfaces are rare and organs and tissues do have smooth and curved surfaces, hence, resulting in inaccurate AR content registration. One example is the inaccurate labelling and measurement of tumour size without accurate surface fit for information overlay, which can be dangerous and misleading during MIS.

In this paper we present a novel real-time AR framework that provides 3D geometric information for accurate AR content registration and overlay in MIS. We propose a new approach to achieving robust 3D tracking through a feature-based SLAM for real-time performance and accuracy required for endoscopy camera tracking. To obtain accurate geometric information, we incrementally build a dense 3D point cloud by using zero mean normalised cross correlation (ZNCC) stereo matching. Therefore, our framework handles the challenging situations of rapid endoscopy movements with robust real-time tracking, while providing an interactive geometry-aware AR environment.

3. Methods: As can be seen from the flowchart in Figure 2, our proposed framework starts with a SLAM system that can track and estimate the camera pose frame by frame. The following stereo matching algorithm based on ZNCC is used to reconstruct dense surface at each keyframe, which is then transformed and stitched to a global surface based on the inverse transformation of the camera pose. Finally, the global surface is re-projected to 2D based on the camera pose and overlaid on the image frame, serving as an interactive geometric layer. The geometric layer enables the interactive AR applications such as online measurement which will be explained in Section 4.

3.1. Landmark Point Detection and Triangulation: In medical interventions, real-time performance and accuracy are both critical.
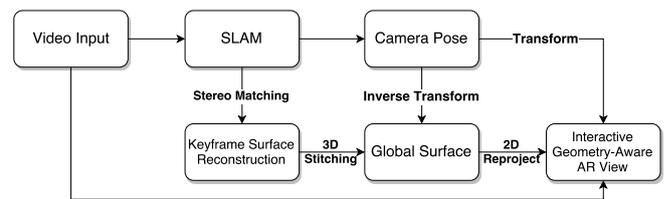

Figure 2. Flowchart describing the whole framework

We adopt Oriented FAST and Rotated BRIEF (ORB) [22] feature descriptors for feature points extraction, encoding and comparison to match landmark points in left and right stereo images. ORB is a binary feature point descriptor that is an order of magnitude faster than SURF [23], more than two orders faster than SIFT [24] and also offers better accuracy than SURF and SIFT [22]. In addition, ORB features are invariant to rotation, illumination and scale, hence, capable of dealing with challenge endoscope camera scenes (rapid rotating, zooming and changing of brightness).

We apply the ORB detector and find the matched keypoints on left and right images. Let x^l_P and x^r_P be the x coordinates on the left and right images, respectively. Assuming the left image and the right image are already rectified, the focal length of both cameras f and the baseline B are known fixed values, by similar triangles, the depth or the perpendicular distance Z between the points and the endoscope can be found according to similar triangles (see Figure 1).

$$\frac{B - \left(x^l_P - x^r_P\right)}{Z - f} = \frac{B}{Z} \;\; \Rightarrow \;\; Z = \frac{f \cdot B}{d_P} \tag{1}$$

Where $x^l_P - x^r_P$ is the disparity d_P of the two corresponding keypoints in the left and the right images detected by the ORB feature.

We then perform a specular reflection detection by removing the keypoints that have intensities above a threshold for efficiency. This could effectively remove the influence of specular reflections from the next stage of computation.

3.2. Frame by Frame Camera Pose Estimation: Any AR application requires the real-time frame by frame tracking to continuously update the overlay positions. To meet the real-time requirement, after initialization, we employ the constant velocity motion model used by MonoSLAM [25] to roughly estimate the position r_{t+1} and quaternion rotation q_{t+1} of the camera position based on the current linear velocity v_t and angular velocity w_t in a small period Δt:

$$\left. \begin{array}{l} r_{t+1} = r_t + v_t \cdot \Delta t \\ q_{t+1} = q_t \times q\left(\omega_t \cdot \Delta t\right) \\ v_t = v_{t-1} + a_{t-1} \cdot \Delta t \\ \omega_t = \omega_{t-1} + \alpha_{t-1} \cdot \Delta t \end{array} \right\} \tag{2}$$

Based on the predicted camera pose (r_{t+1}, q_{t+1}), the potential regions where the feature points may appear on the image are estimate by re-projection of 3D points, hence reducing searching areas and computational cost.

A RANSAC procedure is then performed to obtain the rotation and translation estimations from the set of all the inlier points. During each RANSAC iteration, 3 pairs of corresponding 3D points from current point set p^i_t and point set in next period p^i_{t+1} are selected randomly to calculate the rotation matrix R and the translation t, which minimizes the following objective function:

$$min \sum_{i=1}^{n} \left\| p^i_t - \left(R * p^i_{t+1} + T\right)\right\| \tag{3}$$

From the set with smallest re-projection error, the set of outlier points is rejected and all the inliers are used for a refinement of the final rotation and translation estimations.

During the inlier/outlier identification scheme by RANSAC, false matched ORB feature points, moving specular reflection points and deforming points are effectively rejected. This is a very important step for a MIS scene where the tissue deformation caused by respiration and heartbeat, as well as blood, smoke and surgical instruments can have impact on the tracking stability. Therefore, at this stage, we use the strategy to filter out any influence caused by occlusion and deformation to recover the camera pose. Indeed, the deformable surface is an unsolved challenge in MIS AR, we address this issue by reconstructing a dense 3D map through a more efficient stereo matching method (see Section 3.4).

3.3. Keyframe-based Bundle Adjustment: As our camera pose estimation is only based on the last state, the accumulation of error over time would cause system drifting. However, we cannot perform a global optimization for every frame as this will slow down the system over time. We follow the successful approach of PTAM [10] and ORBSLAM [11] in correcting system drafting, which use the keyframe-based SLAM framework to save "snapshots" of some frames as keyframes to enhance the robustness of the tracking whilst not increasing computational load on the system. Each keyframe is selected based on the criteria that the common keypoints of the two keyframes are less than 80% keypoints but the total number exceeds 50.

Once a keyframe is assigned, bundle adjustment (BA) is applied to refine the 3D positions of each stored keyframe KF_i and the landmark points P_j by minimising the total Huber robust cost function w.r.t the re-projection error between 2D matched keypoints $p_{i,j}$ and camera perspective projections of the 3D positions of keyframes KF_i and the landmark points P_j:

$$\underset{KF_i, P_j}{\arg\min} \sum_{i,j} \rho_h \left(\| p_{i,j} - CamProj\left(KF_i, P_j\right) \| \right) \quad (4)$$

3.4. ZNCC Dense Stereo Matching: We create a feature-based visual odometry system for the endoscopic camera tracking and landmark points mapping, which takes into account of illumination changes, specular reflections and tissue deformations in MIS scenes. However, as the sparse landmark points can barely describe the challenging environment of MIS scenes, we perform a dense stereo matching upon the landmark points to create a dense reconstruction result.

The dissimilarity measure used during the stereo matching is a patch-based ZNCC method. The cost value $C(p, d)$ for a pixel p at disparity d is derived by measuring the ZNCC of the pixel in the left image and the corresponding the pixel $p - d$ in the right image:

$$C(p, d) = \frac{\sum_{q \in N_p} \left(I_L(q) - \bar{I}_L(p)\right) \cdot \left(I_R(q-d) - \bar{I}_R(p-d)\right)}{\sqrt{\sum_{q \in N_p} \left(I_L(q) - \bar{I}_L(p)\right)^2 \cdot \sum_{q \in N_p} \left(I_R(q-d) - \bar{I}_R(p-d)\right)^2}} \quad (5)$$

where $\bar{I}(p) = \frac{1}{N_p} \sum_{q \in N_p} I(q)$ is the mean intensity of the patch N_p centered at p.

ZNCC is proven to be less sensitive to illumination changes and can be parallelised efficiently on a GPU [26]. A WTA (Winner-Takes-All) strategy is applied to choose the best disparity value for each pixel p, followed by a convex optimization to solve the cost volume constructed by Huber-L^1 variational energy function [27] for a smooth disparity map. We used the GPU implement of ZNCC and convex optimization for the efficient disparity map estimation and filtering in real-time.

3.5. Incremental Building of Geometric Mesh : The 3D dense points estimated by stereo matching are transformed to the world coordinate system by the transformation matrix from frame space
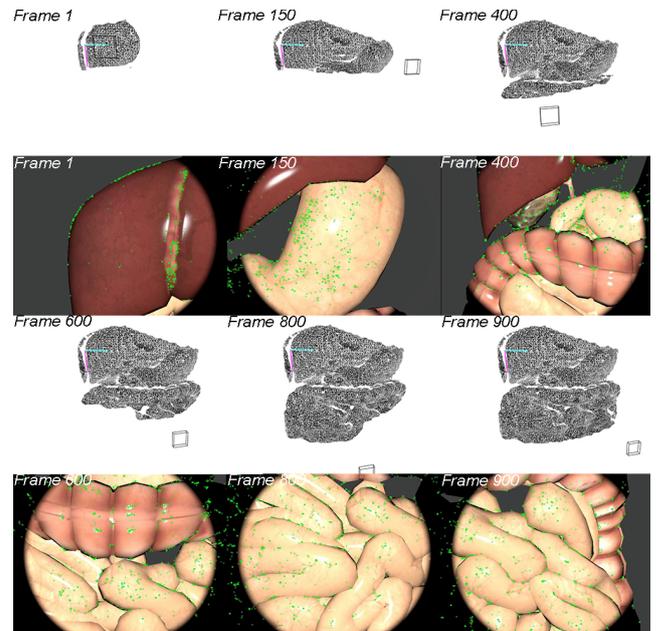

Figure 3 *Incrementally building the geometric mesh. Rectangular boxes are the estimated camera pose; Green points are detected landmark points.*

to the world space T_{f2w} that was estimated by our feature-based SLAM system. A fast triangulation method [28] is then used to incrementally reconstruct the dense points into a surface mesh. Fig. 3 demonstrates the incrementally building process from Frame 1 to 900. The first and third rows are the reconstructed geometric mesh while the second and fourth rows are the current video frames. The geometric mesh can be built incrementally to form a global mesh that can then be re-projected back to the camera's view using the estimated camera pose for the augmented view (see (a) and (c) in Fig. 4 and Fig. 5).

4. Results and Discussion: We have designed a two-parts assessment process to evaluate our AR framework: (i) using a realistic 3D simulated MIS scene as the ground truth study to measure the reconstruction error by measuring the difference between the ground truth values and the reconstructed values; (ii) using a real *in vivo* video acquired from the Hamlyn Centre Laparoscopic/Endoscopic Video Datasets [29] [30] to assess the quality of applications of our proposed framework i.e. measurements, adding AR labels and areas highlighting.

4.1. System setup: Our system is implemented in an Ubuntu 14.04 environment using C/C++. All experiments are conducted on a workstation equipped with Intel Xeon(R) 2.8 GHz quad core CPU, 32G Memory, and one NVIDIA GeForce GTX 970 graphics card. The size of the simulation image sequences and *in vivo* endoscope videos is 840 X 640 pixels. The AR framework and 3D surface reconstruction run in different threads. The 3D surface reconstruction process takes about 200ms to traverse the entire pipeline for each frame. Our proposed AR framework can run in real-time at 26 FPS when the reconstruction only performs at keyframes.

4.2. Ground truth study using simulation data: The performance of our proposed framework is measured in terms of reconstruction accuracy by comparing the reconstructed surface with the 3D model used to render the simulation video.

To quantitatively evaluate the performance of the progressive reconstruction result, we used Blender [31] – an open source 3D software to render realistic image sequences of a simulated abdominal cavity scene using a set of pre-defined endoscopic

Figure 4. Reconstruction error map.

(a)   (b)

(c)   (d)

Figure 5 *Measurement application of our proposed geometry-aware AR framework. Note that the measuring lines (green lines) accurately follow along the curve surface.*
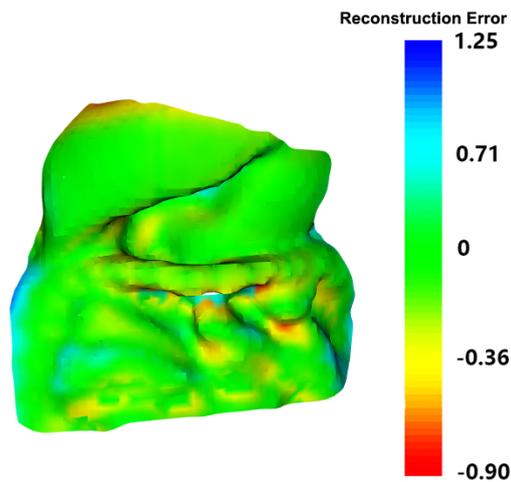
camera movements. The simulated scene contains models scaled to real life size according to an average measured liver diameter of 14.0 cm [32], and the digestive system is rendered with appropriate textures to make the scene as realistic as possible. The material property is set with a strong specularity component to simulate the smooth and reflective liver surface tissue. The luminance is intentionally set high to simulate an endoscope camera as shown in Fig. 5 with a realistic endoscopic lighting condition by using a spot light attached to the main camera. We have designed a camera trajectory that hovers around the 3D models. There are a total of 900 frames of image sequences at a frame-rate of 30 fps being rendered, which is equivalent to a 30 seconds video.

Root Mean Square Distance (RMSD) is used to evaluate the overall distance between the simulated and the reconstructed surfaces. By aligning the surfaces to the real world coordinate system, we apply a grid sample to get a series of x,y coordinate points based on the surface area, and then compared the distance of the z value of the two surfaces.

$$RMSD = \sqrt{\frac{1}{mn}\sum_{x=1}^{m}\sum_{y=1}^{n}(Z_{x,y} - z_{x,y})^2}$$

The RMSD measurement for the two surface alignments has shown a good surface reconstruction results from our proposed methods, compared to the ground truth surface, the RMSD is 2.37 mm. The reconstruction error map can be viewed in Fig. 4.

4.3. Real endoscopic video evaluation: To qualitatively evaluate the performance of our proposed surface reconstruction framework, we applied the proposed approach on *in vivo* videos that we acquired from Hamlyn Centre Laparoscopic / Endoscopic Video Datasets [29] [30]. Fig. 6 (a) shows the reconstruction result from our 3D reconstruction framework with the augmented view of *in vivo* video sequences. By clicking the mesh, augmented objects (colored planes) can be superimposed at corresponding positions with correct poses based on the normals of the points at the click locations. Fig. 6 (b) shows the side-view of the mesh; note that the colored planes (which could be labels) are sticking onto the mesh correctly to create a realistic augmented environment. Fig. 6 (c) shows the area highlighting function of our proposed AR framework. And Fig. 6 (d) is the corresponding mesh view. The area highlighting function can be extended to an area measurement and line measurement (such as shown in Fig. 5) application once the extrinsic parameters of the camera are known.
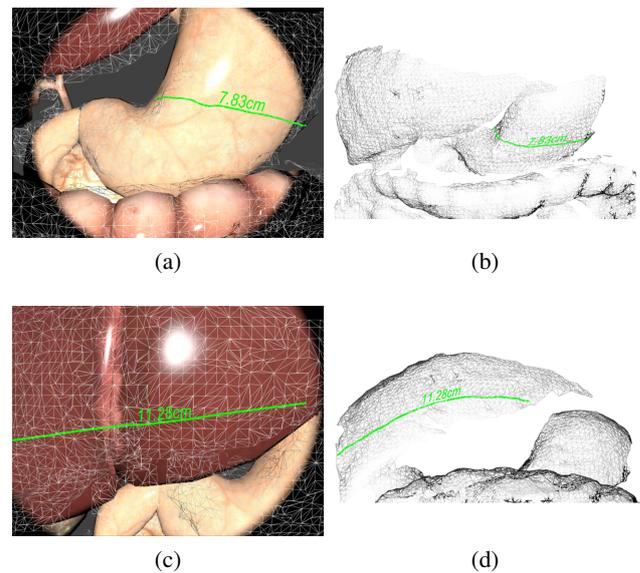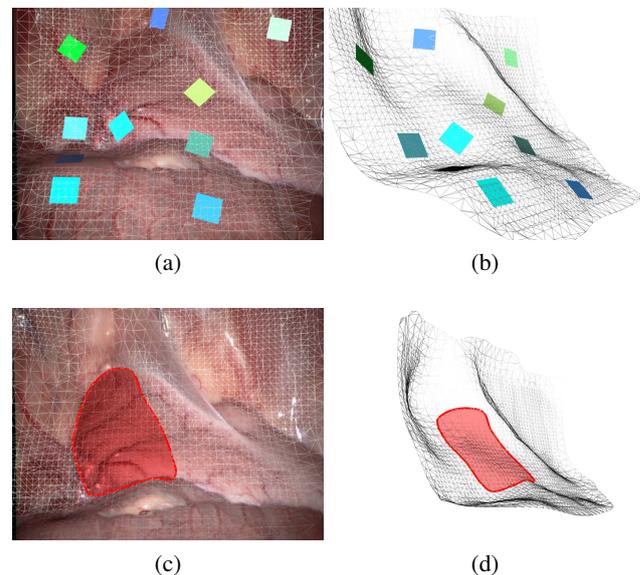

(a)   (b)

(c)   (d)

Figure 6 *Applications of our proposed geometry-aware AR framework: (a) Adding AR labels according to the norm of the geometric surface. (b) The side-view of labels in mesh view. (c) Area highlight and measurement. (b) The side-view of highlighted area in mesh view.*

5. Conclusions: In this paper, we presented a novel AR framework for MIS. Our framework handles the two intertwined issues of tracking the rapid endoscope camera movements and providing accurate information overlay onto the curved surfaces of organs and tissues. By adapting the latest SLAM algorithms, we take a set of innovative approaches at the each stage of the AR process to improve the computational performances and AR registration accuracy. As a result, an interactive real-time geometric aware AR system has been developed. The system is capable of dealing with small soft tissue deformations, rapid endoscope movement and illumination change, which are common challenges in MIS AR. Our proposed system does not require any prior template or pre-operative scan. The system can overlay accurate augmented information such as annotations, labelling, and measurements of a tumour over curved surfaces, greatly improving the quality of AR technology in MIS.

In future work we will carry out a clinical pilot study. A case scenario will be investigated in collaboration with a practicing surgeon, and comparisons will be made as to the effectiveness of our system with the current procedural approach used.

6. Conflict of Interest: None declared.

References

[1] Jae-Hak Kim, Adrien Bartoli, Toby Collins, and Richard Hartley. Tracking by detection for interactive image augmentation in laparoscopy. *Lecture Notes in Computer Science*, pages 246–255, 2012.

[2] Li-Ming Su, Balazs P. Vagvolgyi, Rahul Agarwal, Carol E. Reiley, Russell H. Taylor, and Gregory D. Hager. Augmented reality during robot-assisted laparoscopic partial nephrectomy: Toward real-time 3D-CT to stereoscopic video registration. *Urology*, 73(4):896–900, April 2009.

[3] Nicolas Bourdel, Toby Collins, Daniel Pizarro, Clement Debize, Anne-Sophie Grï£¡meau, Adrien Bartoli, and Michel Canis. Use of augmented reality in laparoscopic gynecology to visualize myomas. *Fertility and sterility*, 107:737–739, March 2017.

[4] Nazim Haouchine, Jeremie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and StÃl'phane Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, pages 199–208. IEEE, 2013.

[5] Nazim Haouchine, Stephane Cotin, Igor Peterlik, Jeremie Dequidt, Mario Sanz Lopez, Erwan Kerrien, and Marie-Odile Berger. Impact of soft tissue heterogeneity on augmented reality for liver surgery. *Visualization and Computer Graphics, IEEE Transactions on*, 21(5):584–597, 2015.

[6] Xiaofei Du, Neil Clancy, Shobhit Arya, George B Hanna, John Kelly, Daniel S Elson, and Danail Stoyanov. Robust surface tracking combining features, intensity and illumination compensation. *International journal of computer assisted radiology and surgery*, 10(12):1915–1926, 2015.

[7] Rosalie Plantefére, Igor Peterlik, Nazim Haouchine, and Stéphane Cotin. Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery. *Ann Biomed Eng*, 44(1):139–153, January 2016.

[8] Peter Mountney and Guang-Zhong Yang. Soft tissue tracking for minimally invasive surgery: learning local deformation online. *Med Image Comput Comput Assist Interv*, 11(Pt 2):364–372, 2008.

[9] M. Turan, Y. Almalioglu, H. Araujo, E. Konukoglu, and M. Sitti. A Non-Rigid Map Fusion-Based RGB-Depth SLAM Method for Endoscopic Capsule Robots. *ArXiv e-prints*, May 2017.

[10] G. Klein and D. Murray. Parallel tracking and mapping for small ar workspaces. In *Proc. 6th IEEE and ACM Int. Symp. Mixed and Augmented Reality*, pages 225–234, November 2007.

[11] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. Orb-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5):114–1163, October 2015.

[12] Alexandre Hostettler Christophe Doignon Luc Soler Jacques Marescaux J.M.M. Montiel Nader Mahmoud, Inigo Cirauqui. Orbslam-based endoscope tracking and 3d reconstruction. In *MICCAI 2016 CARE*, 2016.

[13] Bingxiong Lin, Adrian Johnson, Xiaoning Qian, Jaime Sanchez, and Yu Sun. Simultaneous tracking, 3D reconstruction and deforming point detection for stereoscope guided surgery. *Lecture Notes in Computer Science*, pages 35–44, 2013.

[14] Oscar G Grasa, Ernesto Bernal, Santiago Casado, Iñigo Gil, and JM Montiel. Visual slam for handheld monocular endoscope. *Medical Imaging, IEEE Transactions on*, 33(1):135–146, 2014.

[15] Atul Kumar, Yen-Yu Wang, Ching-Jen Wu, Kai-Che Liu, and Hurng-Sheng Wu. Stereoscopic visualization of laparoscope image using depth information from 3D model. *Computer Methods and Programs in Biomedicine*, 113(3):862–868, March 2014.

[16] Jakob Engel, Thomas Schï£¡ps, and Daniel Cremers. LSD-SLAM: Large-scale direct monocular SLAM. In *Computer Vision – ECCV 2014*, pages 834–849. Springer Nature, 2014.

[17] Ping-Lin Chang, Ankur Handa, Andrew J Davison, Danail Stoyanov, et al. Robust real-time visual odometry for stereo endoscopy using dense quadrifocal tracking. In *International Conference on Information Processing in Computer-Assisted Interventions*, pages 11–20. Springer, 2014.

[18] Peter Mountney, Danail Stoyanov, Andrew Davison, and Guang-Zhong Yang. Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006*. Springer, January 2006.

[19] P. Mountney and G. Z. Yang. Dynamic view expansion for minimally invasive surgery using simultaneous localization and mapping. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1184–1187, September 2009.

[20] Peter Mountney and Guang-Zhong Yang. Motion compensated slam for image guided surgery. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*, January 2010.

[21] Johannes Totz, Peter Mountney, Danail Stoyanov, and Guang-Zhong Yang. Dense surface reconstruction for enhanced navigation in mis. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, January 2011.

[22] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: an efficient alternative to sift or surf. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2564–2571. IEEE, 2011.

[23] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*, pages 404–417. Springer, 2006.

[24] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[25] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse. Monoslam: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067, June 2007.

[26] Danail Stoyanov, Marco Visentini Scarzanella, Philip Pratt, and Guang-Zhong Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*, January 2010.

[27] Ping-Lin Chang, Danail Stoyanov, Andrew J. Davison, and Philip Eddie Edwards. Real-time dense stereo reconstruction using convex optimisation with a cost-volume for image-guided robotic surgery. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2013*, January 2013.

[28] Z. C. Marton, R. B. Rusu, and M. Beetz. On fast surface reconstruction methods for large and noisy point clouds. In *2009 IEEE International Conference on Robotics and Automation*, pages 3218–3223, May 2009.

[29] Imperial College London. Hamlyn centre laparoscopic / endoscopic video datasets, 2016. [Accessed 6 Nov. 2016] http://hamlyn.doc.ic.ac.uk/vision/.

[30] Peter Mountney, Danail Stoyanov, and Guang-Zhong Yang. Three-dimensional tissue deformation recovery and tracking. *IEEE Signal Processing Magazine*, 27(4):14–24, July 2010.

[31] Blender. Blender - free and open 3d creation software, 2016. [Accessed 6 Nov. 2016] https://www.blender.org/.

[32] Wolfgang Kratzer, Violetta Fritz, Richard A Mason, Mark M Haenle, Volker Kaechele, and Roemerstein Study Group. Factors affecting liver size: a sonographic survey of 2080 subjects. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine*, 22:1155–1161, November 2003.