Robust real-time visual odometry for stereo endoscopy using dense quadrifocal tracking

Ping-Lin Chang[1], Ankur Handa[1], Andrew J. Davison[1], Danail Stoyanov[3], and Philip "Eddie" Edwards[1,2]

[1] Department of Computing
[2] Department of Surgery and Cancer
Imperial College London, United Kingdom
{p.chang10, a.handa, a.davison, eddie.edwards}@imperial.ac.uk
[3] Centre for Medical Image Computing and Department of Computer Science
University College London, United Kingdom
danail.stoyanov@ucl.ac.uk

Abstract. Visual tracking in endoscopic scenes is known to be a difficult task due to the lack of textures, tissue deformation and specular reflection. In this paper, we devise a real-time visual odometry framework to robustly track the 6-DoF stereo laparoscope pose using the quadrifocal relationship. The instant motion of a stereo camera creates four views which can be constrained by the quadrifocal geometry. Using the previous stereo pair as a reference frame, the current pair can be warped back by minimising a photometric error function with respect to a camera pose constrained by the quadrifocal geometry. Using a robust estimator can further remove the outliers caused by occlusion, deformation and specular highlights during the optimisation. Since the optimisation uses all pixel data in the images, it results in a very robust pose estimation even for a textureless scene. The quadrifocal geometry is initialised by using real-time stereo reconstruction algorithm which can be efficiently parallelised and run on the GPU together with the proposed tracking framework. Our system is evaluated using a ground truth synthetic sequence with a known model and we also demonstrate the accuracy and robustness of the approach using phantom and real examples of endoscopic augmented reality.

1 Introduction

Visual odometry is the process of determining the position and orientation of a camera moving in 3D space using only the associated image data. In minimally invasive surgery (MIS), visual odometry is an element of surgical vision that enables endoscope/laparoscope tracking without additional hardware such as optical or electromagnetic trackers [16]. Such tracking is crucial for image-guided surgery because the accuracy of camera tracking dominates the stability of applications such as registering a preoperative model to the surgical site [3] or building a mosaic for dynamic view expansion [17]. By using a visual odometry approach it is possible to overcome the hand-eye calibration and to reduce error propagation while simplifying clinical translation.

Camera tracking based on photometrics in endoscopic scenes is difficult because of the homogeneous appearance of certain tissues, tissue deformation and severe specularities caused by the strong illumination intensity. Previous works have adopted a sparse feature based simultaneous localisation and mapping (SLAM) approach to stereo laparoscope tracking [6,11]. In such systems, salient features build a long-term map in order to globally correct for camera drift, but however they are severely affected by large highlights and lack of scene rigidity. Recent dense approaches have shown promising results where the camera tracking benefits from using the entire image data resulting in a very robust motion estimation even without bundle adjustment in a texture-poor or occluded scene [5, 12].

In this paper, we propose a dense approach for real-time stereo laparoscope tracking. Our method uses a combination of stereo reconstruction, which is effective at recovering snapshots of the surgical site geometry [9], and quadrifocal tracking. Benefiting from recent GPU technology and parallelisable optimisation algorithms, the proposed dense visual odometry can reach real-time performance. We validate the proposed approach by a ground truth study using a photo realistic surgical scene rendition. We also demonstrate the robustness of the tracking on a real phantom video as well as *in vivo* clinical MIS sequences.

2 Method

The proposed system for dense stereo visual odometry has two main components: 1) stereo reconstruction and 2) quadrifocal tracking. The first reconstruction step is crucial because it initialises point correspondences for the later quadrifocal warping. Importantly both components rely purely on photometric information.

2.1 Preliminaries

Consider an image function $\mathbf{I}(\mathbf{p}) : \boldsymbol{\Omega}_I \rightarrow \mathbb{R}$ where the $\mathbf{p} = (u, v)$ is the pixel location in the domain $\boldsymbol{\Omega}_I \subseteq \mathbb{R}^2$. In the rectified stereo geometry, point $\mathbf{p}_l$ in the left image has its correspondence $\mathbf{p}_r = (u - \mathbf{d}(\mathbf{p}_l), v)$ in the right image found by the disparity function $\mathbf{d} : \boldsymbol{\Omega}_I \rightarrow \mathbb{R}_{\mathbf{d}}$, where $\mathbb{R}_{\mathbf{d}}$ is the range of the disparity in subpixel accuracy.

To represent variables in the two-view stereo, it is convenient to consider the set of image measurements in a vector form such that $\boldsymbol{\mathcal{I}} = (\mathbf{I}_l, \mathbf{I}_r)^\top$ is a vector of stacked intensity values. The stereo disparity can be represented in a similar way, i.e., $\boldsymbol{\mathcal{D}} = (\mathbf{d}_l, \mathbf{d}_r)^\top$ which also implicitly defines the correspondence set $\boldsymbol{\mathcal{P}}$.

2.2 Dense stereo reconstruction

The task of stereo reconstruction is to optimise the disparity function $\mathbf{d}$ in order to establish point correspondence $\boldsymbol{\mathcal{P}}$ across the stereo pair. We exploit the recent real-time stereo reconstruction algorithm [4], which optimises a variational energy function with respect to $\mathbf{d}$:

$$\mathring{\mathbf{d}} = \arg\min_{\mathbf{d}} E_r(\mathbf{d}), \quad \text{where}$$
$$E_r(\mathbf{d}) = \sum_{\mathbf{p}\in\mathbf{\Omega}_I} \Big\{ \|\gamma(\mathbf{p})\nabla\mathbf{d}(\mathbf{p})\|_\varepsilon + \lambda C(\mathbf{p}, \mathbf{d}(\mathbf{p})) \Big\}. \tag{1}$$

The data term C is a 3D disparity cost-volume which is built up by zero-mean normalised cross-correlation (ZNCC) to save the photometric similarity between left and right pixels within the determined disparity range $\mathbb{R}_{\mathbf{d}}$.

The variational model is regularised by disparity gradient, which takes the assumption that the disparity shall be smooth in areas of homogeneous appearance. To preserve discontinuities, which usually occur along image edges, we adopt the anisotropic diffusion tensor for the weighting function:

$$\gamma(\mathbf{p}) = \exp(-\alpha|\nabla\mathbf{I}(\mathbf{p})|^\beta)nn^\top + n^\perp {n^\perp}^\top,$$

where n is the normalised image gradient $n = \frac{\nabla\mathbf{I}(\mathbf{p})}{|\nabla\mathbf{I}(\mathbf{p})|}$ and $n^\perp$ its perpendicular vector and α and β define the weighting strength [18]. The effects of the data and the regularisation term are controlled by the λ.

The energy function is optimised by a GPU-implemented primal-dual algorithm which provides a linear convergence rate $O(1/N)$ [2]. The optimisation parameters are determined by preconditioning which significantly reduces the number of iterations to converge [13]. Note that the Eq. 1 is a first-order total generalized variation (TGV) model which is only able to reconstruct fronto-parallel structure [14]. However we have observed that instead of applying a rather expensive second-order TGV to reconstruct the affine structure, using the Huber-norm $\|\cdot\|_\varepsilon$ for the regulariser term is a good approximate to avoid the staircasing effect caused by L^1-norm, which is sufficient for reconstructing general endoscopic scenes.

2.3 Dense stereo camera tracking

The camera motion $\mathbf{x}$ is minimally parameterised by $\mathfrak{se}(3)$ Lie algebra. Specifically the 6-vector $\mathbf{x} = (\nu, \omega) \in \mathbb{R}^6$ consists of $\nu \in \mathbb{R}^3$ for the linear velocity and $\omega \in \mathbb{R}^3$ for the angular velocity of the motion. The smooth and invertible rigid-body transformation $\mathbf{T} \in \mathbb{SE}(3)$ based on the 6-vector can be obtained by the exponential map of $g(\mathbf{x})$:

$$\mathbf{T}(\mathbf{x}) = \exp(g(\mathbf{x})) = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix} \in \mathbb{R}^{4\times4},$$

where $\mathbf{R} \in \mathbb{SO}(3)$ and $\mathbf{t} \in \mathbb{R}^3$. Details of the $\mathbb{SE}(3)$ Lie group and its generator function g can be found in [15].

Given a reference frame pair $\boldsymbol{\mathcal{I}}^*$ and the reconstructed disparity $\boldsymbol{\mathcal{D}}$, we can track the camera by continuously registering the current frame pair $\boldsymbol{\mathcal{I}}$ with the reference pair using a generative model called quadrifocal warping $w(\boldsymbol{\mathcal{P}}^*, \mathbf{T}_{rl}, \mathbf{K}_l, \mathbf{K}_r; \mathring{\mathbf{T}})$. The $\mathring{\mathbf{T}} \in \mathbb{SE}(3)$ is the current pose with respect to the reference one in camera coordinate. We assume that the stereo laparoscope is calibrated in advance and the intrinsic matrices $\mathbf{K}_l$, $\mathbf{K}_r$ and the extrinsic matrix $\mathbf{T}_{rl}$ are constant.

The registration warping with respect to the camera motion $\mathbf{x}$ can be obtained by optimising the photometric energy function:

$$\mathring{\mathbf{x}} = \arg\min_{\mathbf{x}} E_t(\mathbf{x}), \quad \text{where}$$

$$E_t(\mathbf{x}) = \sum_{\boldsymbol{\mathcal{P}}^* \in \boldsymbol{\mathcal{R}}^*} \Big(\boldsymbol{\mathcal{I}}\big(w(\boldsymbol{\mathcal{P}}^*; \mathbf{T}(\mathbf{x})\hat{\mathbf{T}})\big) - \boldsymbol{\mathcal{I}}^*\big(\boldsymbol{\mathcal{P}}^*\big) \Big)^2. \tag{2}$$

All the corresponding pixels from the reference frame pair form the set $\boldsymbol{\mathcal{R}}^* = \{\{\mathbf{p}_l^*, \mathbf{p}_r^*\}_1, \{\mathbf{p}_l^*, \mathbf{p}_r^*\}_2, \ldots, \{\mathbf{p}_l^*, \mathbf{p}_r^*\}_n\}$ which mutually includes the left and right matching pair with in total n number of correspondences used for tracking. The optimisation incrementally updates the warping motion $\hat{\mathbf{T}} \leftarrow \mathbf{T}(\mathbf{x})\hat{\mathbf{T}}$ toward the minimum. It is assumed that the truth motion parameter $\mathbf{x}$ exists so that $\exists \mathring{\mathbf{x}} : \mathbf{T}(\mathring{\mathbf{x}})\hat{\mathbf{T}} = \mathring{\mathbf{T}}$.

Quadrifocal geometry To maximally exploit the stereo image data for tracking, the quadrifocal geometry is a constraint for associating geometric entities across the four views. However, instead of adopting the rather complicated quadrifocal tensor, two trifocal tensors are decoupled from the four-view in order to bring the quadrifocal geometry constraint into the optimisation [5]. Fig. 1 shows an example of the trifocal geometry for the left view. Note that we will elaborate only the left trifocal tensor, and the right one is exactly its inverse.

A trifocal tensor $\mathcal{T} = [\mathcal{T}_1(\mathbf{x}), \mathcal{T}_2(\mathbf{x}), \mathcal{T}_3(\mathbf{x})]$ is a $3 \times 3 \times 3$ matrix. Each slice in the tensor is defined by $\mathcal{T}_j = \mathbf{a}_j \mathbf{b}_4^\top(\mathbf{x}) - \mathbf{a}_4 \mathbf{b}_j^\top(\mathbf{x})$ where $\mathbf{a}_j$ are the columns of $\mathbf{T}_{rl}$ and $\mathbf{b}_j(\mathbf{x})$ are the columns of the motion matrix $\mathbf{T}(\mathbf{x})$. We use the point-line-point configuration in which the correspondent line $\mathbf{l}_r = (-1, -1, u+v)$ with each of the three tensor slices form the columns of a homography matrix:

$$\mathcal{H}(\mathbf{x}) = [\mathcal{H}_1(\mathbf{x}), \mathcal{H}_2(\mathbf{x}), \mathcal{H}_3(\mathbf{x})] \quad \text{and} \quad \mathcal{H}_j(\mathbf{x}) = \mathcal{T}_j^\top(\mathbf{x}) \mathbf{K}_r^{-1} \mathbf{l}_r.$$

The corresponding point $\mathbf{p}_l$ in the current image can be simply obtained by the homography transformation of the reference point $\mathbf{p}_l^*$. We can now define the warping function in Eq. 2 for each correspondence as:

$$w(\mathbf{p}_l^*; \mathbf{x}) = \pi\Big(\mathbf{K}_l \mathcal{H}(\mathbf{x}) \mathbf{K}_l^{-1} \begin{bmatrix} \mathbf{p}_l^* \\ 1 \end{bmatrix} \Big), \tag{3}$$

where π is the dehomogenisation function projecting a point to its image coordinate.

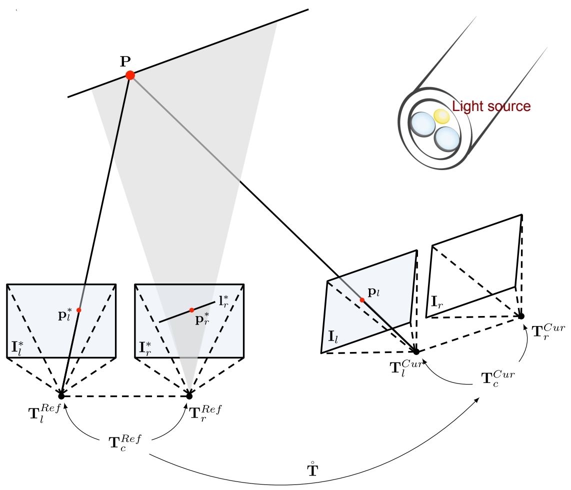


Fig. 1: Point-line-point trifocal geometry: The point $\mathbf{p}_l^*$ in the left reference frame is transformed to the point $\mathbf{p}_l$ in the left current frame using the homography formed by back-projecting the corresponding line $\mathbf{l}_r^*$, which defines an incidence relation $\mathbf{p}_l^* \leftrightarrow \mathbf{l}_r^* \leftrightarrow \mathbf{p}_l$.

Note that the incremental update motion $\mathbf{T}(\mathbf{x})$ is applied to the centralised pose $\mathbf{T}_c$ at the middle of the stereo-rig baseline as shown in Fig. 1. This establishes a canonical coordinate for the stereo geometry, in which the left and right camera poses can be obtained via:

$$\mathbf{T}_c = \exp^{\log(\mathbf{T}_{rl})/2}, \quad \mathbf{T}_l = \mathbf{T}_c^{-1} \quad \text{and} \quad \mathbf{T}_r = \mathbf{T}_c \mathbf{T}_{rl}^{-1}. \tag{4}$$

Robust tracking The original energy function in Eq. 2 is the standard least-square method which assumes the residuals have a zero-mean Gaussian distribution. However, the residual distribution is usually not Gaussian, especially when there are outliers appearing in the scene. For example, occluding objects which do not belong to the original reconstructed model, lighting changes or specularities will generate a considerable number of outliers.

We can instead reformulate Eq. 2 in terms of using a different norm $\rho(r)$. For a least-square norm, $\rho(r) = \frac{1}{2}r^2$:

$$E_{robust}(\mathbf{x}) = \sum_{\boldsymbol{\mathcal{P}}^* \in \boldsymbol{\mathcal{R}}^*} \rho\Big(\boldsymbol{\mathcal{I}}\big(w(\boldsymbol{\mathcal{P}}^*; \mathbf{T}(\mathbf{x})\hat{\mathbf{T}})\big) - \boldsymbol{\mathcal{I}}^*\big(\boldsymbol{\mathcal{P}}^*\big)\Big). \tag{5}$$

We use the non-convex Tukey M-estimator which essentially rejects outliers above the tuning threshold [19]. This results in very robust tracking even with the appearance of instruments occluding the endoscopic scene.

Rapid motion Because the Tukey norm is not a convex function, one cannot expect to find the true global minimum. Furthermore the linearization with respect to the parameters $\mathfrak{se}(3)$ only holds for small camera motions. To make the method more robust towards rapid camera motions we adopt a common coarsetofine scheme.

Optimisation We adopt the efficient second-order minimisation (ESM) algorithm for optimising Eq. 5. ESM is mainly the combination of a forward and an inverse compositional algorithm, which can avoid local minima and takes fewer iterations to converge [10]. The optimisation of quadrifocal warping can be easily framed using ESM due to the fact that the warping is simply two homography transformations in which the warped current image pair and the reference image pair have a linear relationship. Dense tracking by warping a 2.5D surface projection image has no such property and can only use the first-order forward compositional algorithm [1, 12].

The ESM optimisation for solving Eq. 5 is performed with an iteratively reweighted least squares (IRLS) scheme, which will require three Jacobians: $\mathbf{J}_{\mathcal{I}^*}$, $\mathbf{J}_{\mathcal{I}}$ and $\mathbf{J}_w$, the Jacobians of the reference image, the current image and the warping function (Eq. 3) respectively. It can be shown that the overall approximate second-order Jacobian can be derived as:

$$\mathcal{J} = \frac{(\mathbf{J}_{\mathcal{I}} + \mathbf{J}_{\mathcal{I}^*})}{2}\mathbf{J}_w. \tag{6}$$

Derivations of these Jacobians can be found in [5]. Using the common normal equation solver with IRLS, the update parameter $\mathbf{x}$ can be obtained by :

$$\mathbf{x} = -(\mathbf{W}\mathcal{J})^{+}\mathbf{W}(\mathcal{I} - \mathcal{I}^*), \tag{7}$$

where the $\mathbf{W}$ is the diagonal weighting matrix determined by the Tukey M-estimator and $(\cdot)^{+}$ is the pseudo-inverse operator.

2.4 Reference frame selection

The proposed dense stereo visual odometry has the advantage that the reconstruction can be done any time to provide a dense model for the quadrifocal tracking without the need of a bootstrapper. However, reconstructing a model for every frame is unnecessary and in fact frame-to-frame tracking is susceptible to drift. To constrain tracking and prevent drift, we adopt frame-to-model tracking which is essentially the same concept as the keyframe strategy in visual SLAM systems [8, 12].

Whether a subsequent stereo frame pair is selected as the reference frame is based on two criteria: 1) if the overlay between the warping image and the reference image is below a threshold; 2) if the root-mean-square error of Eq. 5

is larger than a threshold. The first criterion occurs when the scope explores a sufficiently large area of the scene, so that there is not enough of the previous reference model in view. The second criterion can also be associated with insufficient overlap one but it is additionally useful that when the scene is invaded by other objects and we have to immediately reconstruct a new reference model for tracking.

3 Empirical Studies

The system is implemented in C/C++ and CUDA running on a Nvidia GeForce GTX 670 with 2GB GPU memory. The real video sequences are acquired from da Vanci robot's stereo endoscopy with size 720×576 and downsampled to 360×288. The stereo reconstruction for two frames takes about 100ms and the tracking for per subsequent stereo pair takes about 40ms.

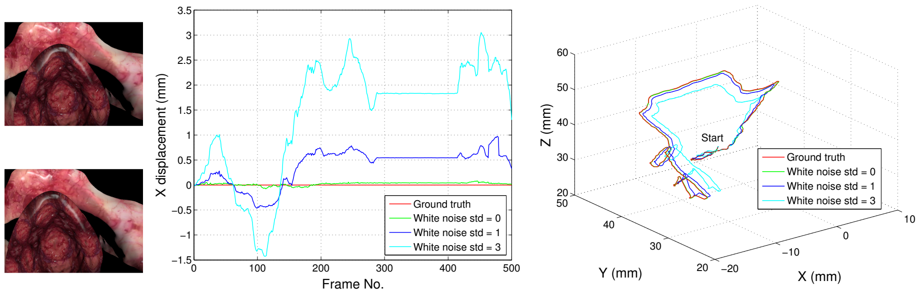
3.1 Sythetic ground truth study

In order to have a ground truth dataset, we use POV-Ray for realistic rendering for a bladder and a pelvis phantom model. The luminance is intentionally set as using a point light source and materials with strong specularity to simulate the real surgical scene where the only light source is at the middle of the endoscopy cameras as shown in Fig. 1. Fig. 2a shows a realistically rendered stereo frames. Following the same methodology in [7], we use the proposed approach to track a real phantom model using the da Vinci robotic platform as shown in Fig. 3b to generate a realistic camera motion. We then use this camera trajectory to render the ground truth sequence. The ground truth trajectory is shown in Fig. 2c.

The first frame is used as the reference frame for tracking the rest. The methodology for the validation is to add white noise to the reference model with different standard deviation and observe how this will affect the tracking. Fig. 2b shows the tracking errors along the x-axis under different level of white noise. It reveals several important results. Firstly, as the green curve shows, tracking with a perfect model gives almost no drift but in practice a perfect reconstruction is never achievable. The blue curve is closer to the real situation where we have a decent reconstruction but not perfect. Due to using the imperfect model, the camera drifts about 0.5mm after tracking for 100 frames. The cyan curve shows that with a very bad reconstructed model, the tracking can still work but with a significant drift.

3.2 Real sequences

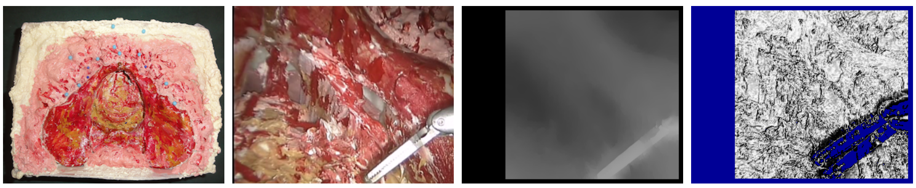
To validate the proposed approach on real data, we use a phantom and a clinical endoscopy sequence to conduct a qualitative evaluation. The phantom is an anatomical pelvis and prostate model from Educational and Scientific Products Limited with added surrounding tissue features made from coloured silicone and outer areas filled with polyeurathane expanding foam to avoid unrealistic sharp



(a) (b) (c)

Fig. 2: The ground truth study. (a) The realistically rendered stereo frames with a pelvis and bladder models. (b) The displacement of the tracked x-translation away from the ground truth. (c) The trajectories. The figures are best viewed on screen with colour and zoomed in.

edges as shown in Fig. 3a. Fig. 3c shows the reconstructed disparity map of the Fig. 3b where the depth discontinuity around the instrument is preserved. With this well-reconstructed model, when the instrument starts to move, the robust estimator assigns low weight for the tracked pixels which do not belong to the model or even completely rejects them, as shown in Fig. 3d.

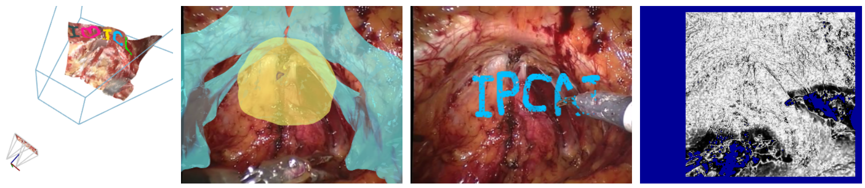


(a) (b) (c) (d)

Fig. 3: (a) The painted plastic phantom. (b) A viewport from the da Vanci robot's endoscopy. (c) The reconstructed disparity map used for quadrifocal tracking. (d) The Tukey M-estimator weighting image where the blue pixels are rejected and gray pixels from black to white corresponds to the weight value from 0.1 to 1.0.

The proposed dense approach can be applied to a variety of applications. We demonstrate augmented reality (AR) using the reconstructed dense model. As shown in Fig. 4a, we can draw text on the dense model and maintain their position on the surface. Note that this is not possible for sparse feature approach in which there is no a dense geometry to be drawn on. This method could be useful as it allows surgeons to tag AR annotation in the endoscopic scenes. Fig. 4b shows another application where we augment the preoperative models into the endoscopic scene.

Another useful function of the robust tracking using a dense model is to detect occlusions. As shown in Fig. 4c and Fig. 4d, the dense reference model provides a strong prior to reject the occluding instrument which is judged directly by Tukey's weight. When a new reference model is added, the occlusion can be also detected by comparing the depths between the tagged markers and the new model. Note that in Fig. 4d, those specularities are also rejected for the quadrifocal tracking. The tracking quality can be observed in the supplementary video[1].



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

Fig. 4: (a) Text drawn on the 3D dense reconstructed model. (b) Preoperative models augmented into the *in vivo* endoscopic scene. (c) An example for occlusion detection. (d) The Tukey weights of (c) showing that the pixels from the invading instrument together with the specularities are mostly rejected.

4 Conclusions

In this paper, we proposed a dense visual odometry method for tracking the motion of the stereo laparoscope in MIS by using quadrifocal constraints. The dense approach has been shown to achieve promising results for synthetic, phantom and clinical data even in sequences with instruments occluding the surgical site. Promising applications of the proposed technique include image-guided surgery with AR overlay onto the laparoscopic images. In our future work we will focus on building a fully dense SLAM system with keyframes refined by pose graph optimisation to accurately maintain a global map while efficiently selecting known keyframes for tracking.

References

1. Baker, S., Patil, R., Cheung, K.M., Matthews, I.: Lucas-kanade 20 years on: Part 5. Tech. Rep. CMU-RI-TR-04-64, Robotics Institute (2004) 6
2. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision (JMIV) 40(1), 150–145 (2011) 3
3. Chang, P.L., Chen, D., Cohen, D., Edwards, P.E.: 2D/3D registration of a preoperative model with endoscopic video using colour-consistency. In: Augmented Environments for Computer-Assisted Interventions (AE-CAI) in Conjunction with MICCAI. vol. 7264, pp. 1–12 (2012) 1

[1] http://www.doc.ic.ac.uk/~pc3509

4. Chang, P.L., Stoyanov, D., Davison, A.J., Edwards, P.E.: Real-time dense stereo reconstruction using convex optimisation with a cost-volume for image-guided robotic surgery. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), vol. 8149, pp. 42–49 (2013) 2
5. Comport, A., Malis, E., Rives, P.: Real-time quadrifocal visual odometry. The International Journal of Robotics 29(2-3), 245–266 (2010) 2, 4, 6
6. Grasa, O., Civera, J., Montiel, J.M.M.: EKF monocular slam with relocalization for laparoscopic sequences. In: IEEE International Conference on Robotics and Automation (ICRA). pp. 4816–4821 (2011) 2
7. Handa, A., Newcombe, R., Angeli, A., Davison, A.: Real-time camera tracking: When is high frame-rate best? In: European Conference on Computer Vision (ECCV), vol. 7578, pp. 222–235 (2012) 7
8. Klein, G., Murray, D.: Parallel tracking and mapping for small AR workspaces. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 225–234 (2007) 6
9. Maier-Hein, L., Mountney, P., Bartoli, A., Elhawary, H., Elson, D., Groch, A., Kolb, A., Rodrigues, M., Sorger, J., Speidel, S., Stoyanov, D.: Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. Medical Image Analysis (MedIA) 17(8), 974 – 996 (2013) 2
10. Malis, E.: Improving vision-based control using efficient second-order minimization techniques. In: IEEE International Conference on Robotics and Automation (ICRA). vol. 2, pp. 1843–1848 (2004) 6
11. Mountney, P., Stoyanov, D., Davison, A., Yang, G.Z.: Simultaneous stereoscope localization and soft-tissue mapping for minimal invasive surgery. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). vol. 9, pp. 347–54 (2006) 2
12. Newcombe, R.A., Lovegrove, S.J., Davison, A.J.: DTAM: Dense tracking and mapping in real-time. In: IEEE International Conference on Computer Vision (ICCV). vol. 1, pp. 2320–2327 (2011) 2, 6
13. Pock, T., Chambolle, A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In: IEEE International Conference on Computer Vision (ICCV). pp. 1762–1769 (2011) 3
14. Ranftl, R., Pock, T., Bischof, H.: Minimizing TGV-based variational models with non-convex data terms. In: Scale Space and Variational Methods in Computer Vision, vol. 7893, pp. 282–293 (2013) 3
15. Stillwell, J.: Naive Lie Theory. Springer (2008) 3
16. Stoyanov, D.: Stereoscopic scene flow for robotic assisted minimally invasive surgery. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). vol. 7510, pp. 479–86 (2012) 1
17. Totz, J., Mountney, P., Stoyanov, D., Yang, G.Z.: Dense surface reconstruction for enhanced navigation in MIS. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). vol. 14, pp. 89–96 (2011) 1
18. Werlberger, M., Trobin, W., Pock, T., Wedel, A., Cremers, D., Bischof, H.: Anisotropic Huber-L1 Optical Flow. In: British Machine Vision Conference (BMVC). pp. 108.1–108.11 (2009) 3
19. Zhang, Z.: Parameter estimation techniques: A tutorial with application to conic fitting. Image and Vision Computing 15, 59–76 (1997) 5