Stereoscopic Surface Reconstruction in Minimally Invasive Surgery using Efficient Non-Parametric Image Transforms

Andreas Schoob, Florian Podszus, Dennis Kundrat, Lüder A. Kahrs and Tobias Ortmaier

Abstract— **Intra-operative instrumentation and navigation in minimally invasive surgery is challenging due to soft tissue deformations. Therefore, surgeons make use of stereo endoscopy and adjunct depth perception in order to handle tissue more safely. Moreover, the surgeon can be assisted with visualized information computed from three-dimensionally estimated surgical site. In this study, a review of state-of-the-art methods dealing with surface reconstruction in minimally invasive surgery is presented. In addition, two real-time solutions based on non-parametric image transforms are proposed. A local method using an efficient census transform is compared to model-based tracking of disparity. The algorithms are evaluated on online available image sequences of a deforming heart phantom with known depth. Both approaches show promising results with respect to accuracy and real-time capability. Particularly, the model-based method is able to compute very dense depth maps if the observed surface is smooth.**

I. Introduction

Stereoscopic vision has become a key component in minimally invasive surgery providing three-dimensional visual feedback. The adjunct depth perception of the scene is beneficial when manipulating tissue with surgical instruments such as forceps or scalpels. Moreover, a dual imaging system can be utilized to reconstruct the surgical site in order to facilitate intra-operative guidance or augmented reality based on additional intra- and pre-operative data.

Due to soft tissue characteristics, current research especially addresses tissue motion tracking [1]–[4] as well as registration of pre- to intra-operative data [5]–[7]. These approaches mostly incorporate stereo-based depth estimation. For this purpose, correspondences between the left and right camera image have to be determined. This can be achieved by cost computation based on state-of-the-art methods [8]. In detail, depth of an object point projected to the image planes is computed based on intensity information obtained from the pixel's neighborhood. Common similarity measures for comparing left and right image are sum of squared differences (SSD) and normalized cross-correlation (NCC).

An early NCC based method models depth as a Gaussian distribution in order to detect and remove instruments from the endoscopic view [9]. Aside from observing instruments, the pose of the endoscope with respect to the target can be determined by applying NCC based feature matching on mono camera images [10]. Using a stereo fiberscope instead, NCC based correspondence search can be combined with simultaneous localization and mapping (SLAM) for estimating the tissue's surface as well as the endoscope's pose and motion [11]. A three-dimensional representation of the surgical site can also be computed with a NCC multiple resolution approach applied on images of a stereo camera capsule. Designed for augmented reality in laparoscopy, such a device is deployed inside the patient's abdominal cavity providing an increased field-of-view and depth-of-field [12]. Due to specular highlights on glossy tissue or lighting variations, obtaining an accurate and dense depth map with local similarity measures often leads to non-reliable results. Therefore, one can enhance local metrics by smoothness terms and a dynamic programming framework in order globally estimate depth [5]. In a more recent approach, discriminative matching in especially texture-less tissue regions is achieved by applying adaptive support windows [13]. Further on, depth information obtained at salient features can be propagated into a spatial neighborhood resulting in a semi-dense and smooth depth map of the surgical site [14]. In this particular case, once detected features additionally allow to temporally observe motion of the tissue. Aside from that, one can take both spatial and temporal disparity information into account. Combined with a hybrid CPU-GPU implementation, a powerful real-time reconstruction framework targeting on registration of pre-operative models can be implemented [15].

In contrast to methods combining locally applied similarity measures with spatial or temporal constraints, more global optimization is achieved with model-based methods. In minimally invasive surgery, the surface of the observed soft tissue is generally continuous and smooth. Following this assumption, tissue depth and deformation can be modeled by applying hierarchical free-form registration with piecewise bilinear maps [16]. Especially for cardiac surface deformation, disparity can be described by B-splines and tracked by a first order optimization based on a SSD error function [17]. Even a set of tracked features is sufficient to observe cardiac motion [18]. Despite using features, region-based tracking with elastic deformation modeling and an efficient minimization scheme also guarantees real-time cardiac motion estimation [4]. However, model-based three-dimensional tracking is computationally complex and limited due to depth discontinuities arising at borders of instruments in the surgical field of view.

In general, stereo-based tissue depth estimation is prone to specular highlights, texture-less surface, instrument occlusions, bleeding or fume during laser interventions. If those methods fail, literature provides alternative solutions applicable for intra-operative vision-based navigation. For a comprehensive overview, we recommend [19,20].

All authors are with the Institute of Mechatronic Systems, Leibniz Universität Hannover, Germany. andreas.schoob@imes.uni-hannover.de

Disadvantageously, most of above mentioned methods are evaluated on different image data complicating comparison. In order to provide a structured evaluation framework, a medical data set containing ground truth has been established [6,14]. Unfortunately, only few further methods are verified on this data [15,21]. In addition, literature review has shown that, apart from common similarity measures (i.e. NCC), especially non-parametric metrics based on the efficient rank or census transform are hardly used within image-guided minimally invasive surgery.

In this study, accurate and real-time capable techniques for dense surface reconstruction of surgical scenes are presented. Contrary to the commonly used NCC, a solution comprising a locally applied, more simple and efficient census transform is briefly presented in Sec. II. Providing more dense and smooth depth information, a model-based implementation using thin plate splines (TPS) and fast optimization applied on rank transformed images is introduced subsequently. In Sec. III, both methods are evaluated with respect to accuracy as well as real-time capability using the aforementioned online available image sequences [6,14]. Sec. IV summarizes this contribution.

II. Materials and Methods

In Sec. II-A both rank and census transform are introduced. Subsequently, Sec. II-B and II-C describe our proposed methods for stereoscopic surface reconstruction based on those non-parametric image transforms. To simplify implementation, calibrated and rectified images are used. Thus, disparity computation is reduced to an one-dimensional search problem.

A. Rank and census transform

In contrast to estimating disparity relying on absolute intensities, the rank and census transform of an image $\boldsymbol{I}$ show improved robustness to radiometric differences, lighting changes and noise [22]. For both transforms, the image intensities $\boldsymbol{I}(N(\boldsymbol{p}))$ within a local $M \times N$ neighborhood $N(\boldsymbol{p})$ are compared to the center pixel's $\boldsymbol{p} = (x,y)^{\mathrm{T}}$ intensity $\boldsymbol{I}(\boldsymbol{p})$. The rank transform of image $\boldsymbol{I}$ is defined by

$$\boldsymbol{I}_{\mathrm{R}}(\boldsymbol{p}) = \sum_{j=-N/2}^{N/2} \sum_{i=-M/2}^{M/2} \xi\left(\boldsymbol{I}(\boldsymbol{p}), \boldsymbol{I}(\boldsymbol{p}+(i,j)^{\mathrm{T}})\right). \quad (1)$$

Subsequently, the census transform is given by

$$\boldsymbol{I}_{\mathrm{C}}(\boldsymbol{p}) = \bigotimes_{j=-N/2}^{N/2} \bigotimes_{i=-M/2}^{M/2} \xi\left(\boldsymbol{I}(\boldsymbol{p}), \boldsymbol{I}(\boldsymbol{p}+(i,j)^{\mathrm{T}})\right) \quad (2)$$

with $\bigotimes$ denoting concatenation to a bit string. The function ξ for comparing the two intensities is denoted as

$$\xi(I_1, I_2) = \begin{cases} 0, & \text{if } I_1 \leq I_2 \\ 1, & \text{else.} \end{cases} \quad (3)$$

B. Local census-based disparity computation

For computational efficiency, a sparse census transform is applied resulting in a shortened bit string. Hereby, only every second column and row of the neighborhood $N(\boldsymbol{p})$ are considered. Hamming distance H is used as similarity measure between a pixel $\boldsymbol{p}_\ell = (x_\ell, y_\ell)$ in the left image $\boldsymbol{I}_{\mathrm{C},\ell}$ and a pixel $\boldsymbol{p}_r = (x_r, y_r)^{\mathrm{T}} = \boldsymbol{p}_\ell - (d,0)^{\mathrm{T}}$ in the right image $\boldsymbol{I}_{\mathrm{C,r}}$ (4). Disparity is denoted as d.

$$H(\boldsymbol{p}_\ell, d) = \sum_{i=1}^{M\times N-1} \boldsymbol{I}_{\mathrm{C},\ell}(\boldsymbol{p}_\ell, i) \oplus \boldsymbol{I}_{\mathrm{C,r}}(\boldsymbol{p}_\ell - (d,0)^{\mathrm{T}}, i) \quad (4)$$

Index i defines the appropriate bit in string $\boldsymbol{I}_{\mathrm{C}}(\boldsymbol{p})$ whereas $\oplus$ denotes XOR operation. Smoothness and unambiguous matching are achieved in a cost aggregation by summing up the Hamming distances $H(\boldsymbol{p}_\ell, d)$ within a certain pixel neighborhood. Additionally, our disparity computation comprises a consistency check, a sub-pixel refinement, a removal of disparity speckles and a winner-takes-all (WTA) strategy. Further smoothing is done by bilateral filtering.

C. Model-based disparity computation

An elastic model-based computation is able to provide a dense and reliable disparity map if the observed surface is smooth [17]. Furthermore, such an approach facilitates real-time three-dimensional tissue motion tracking [4]. However, accuracy strongly depends on the number of parameters describing the elastic deformation.

In this study, disparity is estimated by a thin plate spline (TPS) based tracking following the ideas introduced in *Lau et al.* and *Richa et al.* [17,4]. Our algorithm uses rank transformed images and an extended inverse compositional parametrization.

Assuming a rank transformed image region $\boldsymbol{I}_{\mathrm{R}}$ described by n pixels $\boldsymbol{p}_{\ell,i} = (x_{\ell,i}, y_{\ell,i})^{\mathrm{T}}$ in the left and n pixels $\boldsymbol{p}_{\mathrm{r},i} = (x_{\mathrm{r},i}, y_{\mathrm{r},i})^{\mathrm{T}}$ in the right image with $i \in \{1,...,n\}$, mapping between both pixel sets can be formulated by an elastic transformation (5) [23]. Since images are rectified, corresponding points will have the same y-coordinate with $y_{\mathrm{r},i} = y_{\ell,i}$. As a result, disparity is simply defined by $d_i = x_{\ell,i} - x_{\mathrm{r},i}$. One-dimensional mapping for $x_{\mathrm{r},i}$ is then denoted as

$$x_{\mathrm{r},i}(\boldsymbol{p}_{\ell,i}) = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} x_{\ell,i} \\ y_{\ell,i} \\ 1 \end{bmatrix} + \sum_{j=1}^{\alpha} w_j \cdot u(\left\| \boldsymbol{c}_{\ell,j} - \boldsymbol{p}_{\ell,i} \right\|) \quad (5)$$

with TPS basis function $u(r) = r^2 \log r^2$ and parameter vector $\boldsymbol{t} = (w_1, ..., w_\alpha, a_1, a_2, a_3)^{\mathrm{T}}$ [23]. So-called control points $\boldsymbol{c}_{\ell,j} = (\widehat{x}_{\ell,j}, \widehat{y}_{\ell,j})^{\mathrm{T}}$ with $j \in \{1,...,\alpha\}$ are initially set in the left image. According to current depth of the scene, control points $\boldsymbol{c}_{\mathrm{r},j} = (\widehat{x}_{\mathrm{r},j}, \widehat{y}_{\mathrm{r},j})^{\mathrm{T}}$ in the right image need to be estimated . Once correspondence between $\boldsymbol{c}_\ell$ and $\boldsymbol{c}_{\mathrm{r}}$ is known, a mapping for any pixel $\boldsymbol{p}_{\mathrm{r},i}^{\mathrm{T}} = m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_{\mathrm{r}})$ can be formulated with a linear system [4,23]. In general, temporal disparity changes between consecutive frames are described by deviation $\Delta\boldsymbol{c}_{\mathrm{r}}$ with respect to prior control point configuration $\boldsymbol{c}_{\mathrm{r}}$. For real-time estimation of $\boldsymbol{c}_{\mathrm{r}} + \Delta\boldsymbol{c}_{\mathrm{r}}$, an

inverse compositional parametrization is implemented [24]. In detail, a virtual warping of image $\boldsymbol{I}_{\mathrm{R},\ell}(m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\ell + \Delta\boldsymbol{c}_\ell))$ describes disparity changes by shifting $\Delta\boldsymbol{c}_\ell$ with respect to $\boldsymbol{c}_\ell$. Since compositional frameworks cannot be applied to TPS directly, $\boldsymbol{c}_\mathrm{r} = f(\boldsymbol{c}_\ell, \Delta\boldsymbol{c}_\ell)$ is subsequently estimated in a closed-form solution [25]. Thus, our optimization aims on finding $\Delta\boldsymbol{c}_\ell$ instead. As a result, the alignment error to be minimized can be formulated with (6). Before, a sparse rank transform is applied to both left and right image in order to increase robustness to lighting variations without introducing further parameters.

$$\min_{\Delta\boldsymbol{c}_\ell} \epsilon = \sum_{i=1}^{n} \left[\boldsymbol{I}_{\mathrm{R},\ell}(m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\ell + \Delta\boldsymbol{c}_\ell)) - \boldsymbol{I}_{\mathrm{R,r}}(m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\mathrm{r}))\right]^2 \quad (6)$$

For rectified images, one has just to consider the x-components of $\Delta\boldsymbol{c}_\ell = (\Delta\widehat{\boldsymbol{x}}_\ell, \Delta\widehat{\boldsymbol{y}}_\ell)$ where $\Delta\widehat{\boldsymbol{x}}_\ell$ describes a column vector of stacked $\widehat{x}_{\ell,j}$. Performing first order Taylor expansion on (6), this least-squares problem can be iteratively solved by computing $\Delta\widehat{\boldsymbol{x}}_\ell$ as follows

$$\Delta\widehat{\boldsymbol{x}}_\ell = (\boldsymbol{J}(\boldsymbol{I}_{\mathrm{R},\ell}, \boldsymbol{p}_\ell, \boldsymbol{c}_\ell))^+ \begin{bmatrix} \boldsymbol{I}_{\mathrm{R,r}}(m(\boldsymbol{p}_{\ell,1}, \boldsymbol{c}_\mathrm{r})) - \boldsymbol{I}_{\mathrm{R},\ell}(\boldsymbol{p}_{\ell,1}) \\ \vdots \\ \boldsymbol{I}_{\mathrm{R,r}}(m(\boldsymbol{p}_{\ell,n}, \boldsymbol{c}_\mathrm{r})) - \boldsymbol{I}_{\mathrm{R},\ell}(\boldsymbol{p}_{\ell,n}) \end{bmatrix} \quad (7)$$

with pseudoinverse $\boldsymbol{J}^+$ of the Jacobian matrix $\boldsymbol{J}$. Corresponding to pixel $\boldsymbol{p}_{\ell,i}$, the i-th row of $\boldsymbol{J}$ is defined by

$$\boldsymbol{J}_i(\boldsymbol{I}_{\mathrm{R},\ell}, \boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\ell) = \left[\frac{\partial \boldsymbol{I}_{\mathrm{R},\ell}(m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\ell)}{\partial \widehat{x}_{\ell,1}} \cdots \frac{\partial \boldsymbol{I}_{\mathrm{R},\ell}(m(\boldsymbol{p}_{\ell,i}, \boldsymbol{c}_\ell)}{\partial \widehat{x}_{\ell,\alpha}}\right]. \quad (8)$$

Compared to conventional gradient based optimization, Jacobian $\boldsymbol{J}$ and its pseudoinverse $\boldsymbol{J}^+$ has to be computed just once between consecutive frames. As a result, computation time is significantly reduced.

D. Image processing and sequences

The proposed methods are implemented in C++ using NVIDIA CUDA and OpenCV [26]. For highly parallel computing a NVIDIA GTX TITAN is deployed accessing both global and shared GPU memory. Quantitative evaluation is conducted on two online available image sequences of a deforming silicon heart phantom with an image resolution of 320×288 pixels (see Fig. 1) [6,14,27]. Obtained by CT scans and registered to camera frame, there are 20 ground truth depth maps each used for multiple images of a sequence.
In the following, the implemented algorithms will be denoted as *local* (see Sec. II-B) and *TPS* method (see Sec. II-C). The *local* method computes depth for the whole scene. Due to reduced overlapping of the left and right camera image, the *TPS* algorithm cannot be initialized considering the entire region. Here, 200×200 pixels are computed. However, the size of image region being reconstructed strongly depends on the surgical task. During robot-assisted actions, e.g. incisions by surgical tools, depth computation of a target region might be sufficient whereas registration to pre-operative images often requires the whole 3D scene.

III. Results

A. Image sequences with known ground truth

In our experiments, mean disparity and depth errors as well as their standard deviations are determined with respect to ground truth. The results are shown in Fig. 1 and listed in Table I which also contains the percentage of matched points and computation time per frame. Since the heart phantom's surface is smooth, the *TPS* method outperforms the *local* one with respect to accuracy and density of reconstruction. Due to more global optimization, image noise is intrinsically compensated and depth in sparsely textured surface can be estimated by the help of a discriminative neighborhood. Once initialized, tracking of disparity is even robust in dynamically changing tissue surface. Nevertheless, if the scene is sufficiently illuminated, the *local* method is able to compute dense disparity information in real-time, too.
Compared to methods from literature, our errors are within same order of magnitude. For instance, *Stoyanov et al.* estimate disparity with an error of $0.89 \pm 1.13\,px$ for *Heart 1* and $1.22 \pm 1.71\,px$ for *Heart 2* [14]. *Röhl et al.* report a depth error of $1.45\,mm$ for *Heart 1* and $1.64\,mm$ for *Heart 2* [15].

TABLE I
Computational results of the proposed methods

	Heart 1		*Heart 2*	
	local	TPS	local	TPS
Disp. err. $[px]$	1.86 ± 0.94	1.38 ± 0.72	0.96 ± 0.27	0.69 ± 0.26
Depth err. $[mm]$	1.87 ± 0.80	1.74 ± 0.70	1.68 ± 0.47	1.52 ± 0.52
Matched $[\%]$	67.6 ± 9.40	54.4 ± 0.50	67.5 ± 12.3	57.2 ± 0.83
Time $[ms]$	25.6 ± 0.98	34.0 ± 5.77	25.6 ± 0.91	29.7 ± 6.56

B. In vivo image sequence

A qualitative experiment is conducted on an online available *in vivo* porcine sequence [20]. Fig. 2 show that if the number of control points is increased to 4×4, the error between *TPS* and *local* depth estimation is significantly reduced.

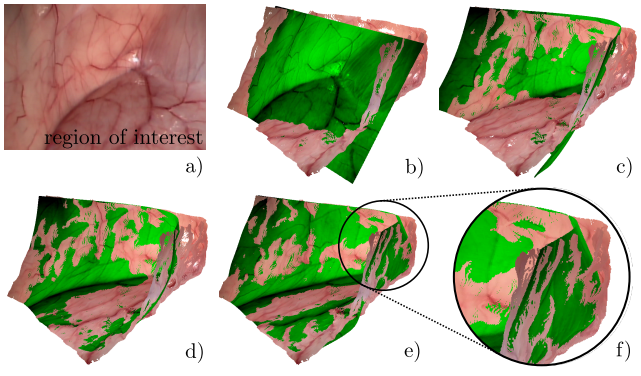

Fig. 2. Qualitative validation on *in vivo* porcine procedure [20]; reconstructed depth with *local* method (natural color) and *TPS* method (green color) with a) selected frame, b) 2×2, c) 3×2, d) 3×3 and e) 4×4 control points with f) improved depth estimation compared to b), c) and d)

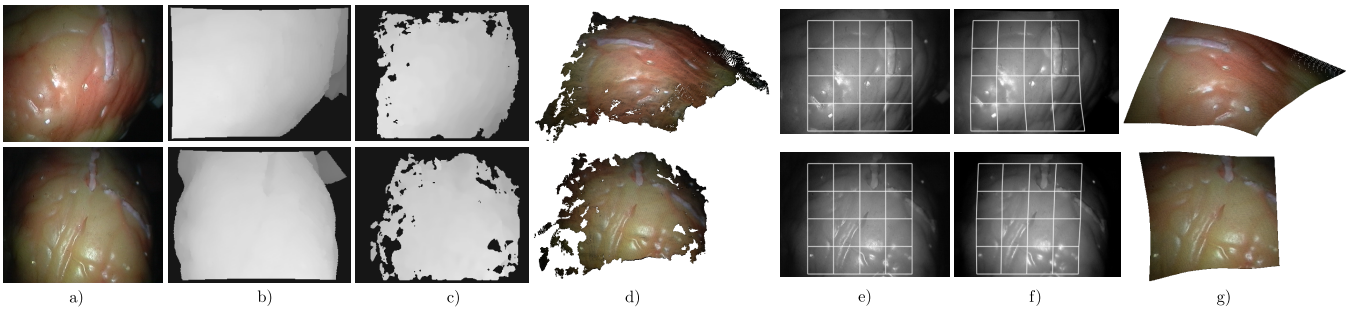

Fig. 1. Results of disparity and depth computation; top: *Heart 1* sequence; bottom: *Heart 2* sequence; a) frame of left camera b) ground truth disparity map; c) computed disparity with *local* method; d) reconstructed depth with *local* method; e) left image with 5×5 control points of *TPS* method; f) corresponding control points in the right image; g) reconstructed depth with *TPS* method

IV. Conclusion and Outlook

In this study, surface reconstruction of surgical scenes based on non-parametric image transforms is evaluated. The proposed methods are fast and provide dense surface estimation. If the observed surface is sufficiently smooth, our model-based algorithm is highly accurate. Although it uses non-deterministic optimization, real-time capability is achieved by an inverse compositional framework. However, the choice between these two methods strongly depends on the surgical task and tissue properties. Future work will deal with incorporation in an intra-operative system for laser phonomicrosurgery.

Acknowledgment

This research has received funding from the European Union FP7 under grant agreement μRALP - n^{o} 288663. The authors thank the Visual Information Processing Group at the Imperial College in London for providing image data.

References

[1] M. Groeger, T. Ortmaier, W. Sepp, and G. Hirzinger, "Tracking local motion on the beating heart," *Proceedings of SPIE Medical Imaging*, pp. 233–241, 2002.

[2] D. Stoyanov, G. Mylonas, F. Deligianni, A. Darzi, and G. Yang, "Soft-tissue motion tracking and structure estimation for robotic assisted mis procedures," in *Proceedings of MICCAI*, 2005, vol. 3750, pp. 139–146.

[3] P. Mountney and G.-Z. Yang, "Soft tissue tracking for minimally invasive surgery: Learning local deformation online," in *Proceedings of MICCAI*, 2008, vol. 5242, pp. 364–372.

[4] R. Richa, P. Poignet, and C. Liu, "Deformable motion tracking of the heart surface," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2008, pp. 3997 –4003.

[5] G. Hager, B. Vagvolgyi, and D. Yuh, "Stereoscopic video overlay with deformable registration," *Medicine Meets Virtual Reality*, 2007.

[6] P. Pratt, D. Stoyanov, M. Visentini-Scarzanella, and G.-Z. Yang, "Dynamic guidance for robotic surgery using image-constrained biomechanical models," in *Proceedings of MICCAI*, vol. 6361, 2010, pp. 77–85.

[7] S. Speidel, S. Roehl, S. Suwelack, R. Dillmann, H. Kenngott, and B. Mueller-Stich, "Intraoperative surface reconstruction and biomechanical modeling for soft tissue registration," in *Proc. Joint Workshop on New Technologies for Computer/Robot Assisted Surgery*, 2011.

[8] J. Banks, M. Bennamoun, and P. Corke, "Non-parametric techniques for fast and robust stereo matching," in *Proceedings of IEEE TENCON*, vol. 1, 1997, pp. 365–368.

[9] F. Mourgues, F. Devernay, and È. Coste-Manière, "3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery," in *IEEE and ACM International Symposium on Augmented Reality (ISAR)*, 2001, p. 191.

[10] T. Thormaehlen, H. Broszio, and P. Meier, "Three-dimensional endoscopy," *Medical imaging in gastroenterology and hepatology, 124th Falk Symposium Hannover, Germany*, 2002.

[11] D. Noonan, P. Mountney, D. Elson, A. Darzi, and G.-Z. Yang, "A stereoscopic fibroscope for camera motion and 3D depth recovery during minimally invasive surgery," in *IEEE International Conference on Robotics and Automation, ICRA*, 2009, pp. 4463 – 4468.

[12] B. Tamadazte, S. Voros, C. Boschet, P. Cinquin, and C. Fouard, "Augmented 3-d view for laparoscopy surgery," in *Augmented Environments for Computer-Assisted Interventions*, 2013, vol. 7815, pp. 117–131.

[13] S. Bernhardt, J. Abi-Nahid, and R. Abugharbieh, "Robust dense endoscopic stereo reconstruction for minimally invasive surgery," in *MICCAI workshop on MCV*, 2012, pp. pp. 198–207.

[14] D. Stoyanov, M. Scarzanella, P. Pratt, and G.-Z. Yang, "Real-time stereo reconstruction in robotically assisted minimally invasive surgery," in *Proceedings of MICCAI*, 2010, vol. 6361, pp. 275 – 282.

[15] S. Rohl, S. Bodenstedt, S. Suwelack, H. Kenngott, B. P. Muller-Stich, R. Dillmann, and S. Speidel, "Dense gpu-enhanced surface reconstruction from stereo endoscopic images for intraoperative registration," *Medical Physics*, vol. 39, p. 1632, 2012.

[16] D. Stoyanov, A. Darzi, and G. Yang, "Dense 3d depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery," in *Proceedings of MICCAI*, 2004, vol. 3217, pp. 41–48.

[17] W. Lau, N. Ramey, J. Corso, N. Thakor, and G. Hager, "Stereo-based endoscopic tracking of cardiac surface deformation," in *Proceedings of MICCAI*, 2004, vol. 3217, pp. 494–501.

[18] D. Stoyanov, A. Darzi, and G. Z. Yang, "A practical approach towards accurate dense 3d depth recovery for robotic laparoscopic surgery," *Computer Aided Surgery*, vol. 10, no. 4, pp. 199–208, 2005.

[19] D. Stoyanov, "Surgical vision," *Annals of Biomedical Engineering*, vol. 40, no. 2, pp. 332–345, 2012.

[20] P. Mountney, D. Stoyanov, and G.-Z. Yang, "Three-dimensional tissue deformation recovery and tracking," *Signal Processing Magazine, IEEE*, vol. 27, no. 4, pp. 14 –24, july 2010.

[21] M. C. Yip, D. G. Lowe, S. E. Salcudean, R. N. Rohling, and C. Y. Nguan, "Tissue tracking and registration for image-guided surgery," *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2169–2182, 2012.

[22] C. Pantilie and S. Nedevschi, "Optimizing the census transform on cuda enabled gpus," in *IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2012, pp. 201–207.

[23] F. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[24] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221 – 255, 2004.

[25] F. Brunet, V. Gay-Bellile, A. Bartoli, N. Navab, and R. Malgouyres, "Feature-driven direct non-rigid image registration," *International Journal of Computer Vision*, vol. 93, pp. 33–52, 2011.

[26] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000. [Online]. Available: http://www.opencv.org

[27] S. Giannarou, D. Stoyanov, D. Noonan, G. Mylonas, J. Clark, M. Visentini-Scarzanella, P. Mountney, and G.-Z. Yang, "Hamlyn Centre Laparoscopic / Endoscopic Video Datasets," 2012. [Online]. Available: http://hamlyn.doc.ic.ac.uk/vision/