Real-Time Methods for Long-Term Tissue Feature Tracking in Endoscopic Scenes

Michael C. Yip[1], David G. Lowe[2], Septimiu E. Salcudean[1], Robert N. Rohling[1], and Christopher Y. Nguan[3]

[1] Electrical and Computer Engineering, University of British Columbia, Canada
[2] Computer Science Department, University of British Columbia, Canada
[3] Vancouver General Hospital, Vancouver, British Columbia, Canada

Abstract. Salient feature tracking for endoscopic images has been investigated in the past for 3D reconstruction of endoscopic scenes as well as tracking of tissue through a video sequence. Recent work in the field has shown success in acquiring dense salient feature profiling of the scene. However, there has been relatively little work in performing long-term feature tracking for capturing tissue deformation. In addition, real-time solutions for tracking tissue features result in sparse densities, rely on restrictive scene and camera assumptions, or are limited in feature distinctiveness. In this paper, we develop a novel framework to enable long-term tracking of image features. We implement two fast and robust feature algorithms, STAR and BRIEF, for application to endoscopic images. We show that we are able to acquire dense sets of salient features at real-time speeds, and are able to track their positions for long periods of time.

1 Introduction

There are many instances in which image guidance for minimally invasive surgery requires tissue tracking in endoscopic video sequences in a robust manner. Tissue tracking provides an evaluation of tissue morphology and deformation *in-vivo* over time, which can be important for tracking regions of interest such as vasculature or lesions. When a medical image is registered to the stereoscopic cameras for augmented reality, tissue tracking provides the knowledge of how the underlying tissues are moving such that the medical image can be moved accordingly. This keeps anatomical features such as nerves and lesions in known locations during surgical intervention, which can be critical in enabling the success of an operation such as laparoscopic radical prostatectomy (LRP) [19]. In LRP, the surgeon images an exposed prostate using transrectal ultrasound (TRUS) and dissects it from the surrounding tissues prior to resection. Since contact is lost between the prostate and the surrounding tissue after mobilization, it is important to maintain an image registration to the mobilized prostate in order to perform resection within adequate surgical margins (1.9 ± 0.8 mm [19]).

While many tissue tracking methods have been employed that use fiducial markers, most recent work uses image-based non-invasive techniques. Unlike natural scenes and urban environments, tissue images have poor color and textural

P. Abolmaesumi et al. (Eds.): IPCAI 2012, LNAI 7330, pp. 33–43, 2012.

distinctiveness, few edges and corners, are poorly illuminated, exhibit a great deal of specular reflection, and exhibit non-rigid deformation due to regular patient motion (e.g. heartbeat, breathing) as well as interactions with surgical instruments. Therefore, simple template-based tracking has been found to be limited in performance[12,15]. Others have used deformation functions [13] to capture real-time deformations of a beating heart, but require predictive models for heartbeat motion, frequent reinitializations to avoid drift, and are only able to capture low-order deformations. The most successful method for densely tracking tissue deformation so far is feature tracking. Work by [6,9,11,18,21] and references therein have shown that it is possible to find stable and uniquely identifiable features in tissue images. However, two areas of difficulty still exists:

Feature Management for Long-Term Tracking. A framework for managing features that enter and exit an endoscopic scene is necessary to maintain long-term stable tracking. Despite the successes of Simultaneous Localization and Mapping methods for long-term tracking, they are suitable for natural scenes and urban environments and predominantly rely on camera pose estimation, assuming that scene features generally only move due to camera motion [20,21]. Due to patient movement (breathing, heartbeat) and surgical instrument interactions with tissues, these methods are limited in their application here. Therefore, a long-term feature tracking method, one that minimizes drift without estimating camera pose, and can track features temporarily lost between frames, is an important contribution.

Speed. Features with high saliency, such as the Scale Invariant Feature Transform (SIFT) [5], have been effective at producing dense, stable features [6,10,20,21], but are unable to run at real-time speeds. Computationally fast features (e.g. Shi-Tomasi [16], FAST [14]) have only been able to generate sparse sets of trackable points in real-time operation[8]. GPU implementations have been limited and have yet to achieve real-time speeds at high-definition, and still rely on some CPU computation where GPU-acceleration is deemed not effective [17]. Therefore, an interesting research question is: are we able to perform high-density feature tracking at real-time speeds?

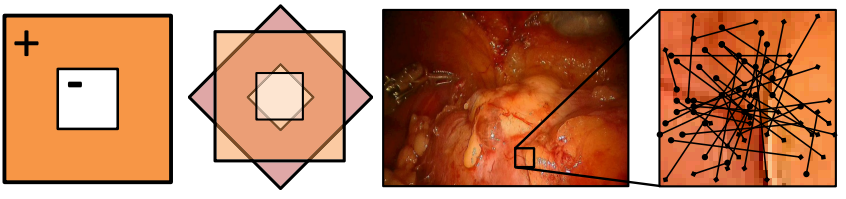
In this paper, we will describe a novel long-term tissue tracking strategy that is capable of capturing a dense map of tissue deformations within endoscopic images. Using two relatively new salient feature algorithms (CenSuRE[1] and BRIEF[2]), we are able to achieve real-time (greater than 10Hz) speeds while maintaining high resolution tracking of tissue deformation for extended periods of time. Furthermore, we show that this can be applied to stereoscopic scenes for describing and tracking 3D tissue surface profiles.

2 Methods

2.1 Salient Features

We propose the use of two relatively new feature detectors: a modified version of the Center Surrounded Extremas (CenSuRE) feature detector [1] called STAR [22], and the Binary Robust Independent Elementary Features (BRIEF) [2].

Mikolojaczyk and Schmid [7] found that the most salient image features found in computer vision literature were derived from the Laplacian of Gaussian (LoG). Where SIFT uses a Difference of Gaussians to estimate the LoG, CenSuRE uses a bi-level center-surrounded square kernel, where pixel values within the kernel are multiplied by either +1 or -1 (Figure 1a). These kernels are applied over pyramidal scale space at all locations to find local extrema that represent salient features. The extrema are then filtered using a scale-adapted Harris measure to eliminate weak responses, and line suppression is performed by evaluating a second moment matrix, and discarding those features that have a large ratio between principle curvatures. The STAR feature detector [22] that we will be using is a modification of the CenSuRE detector by using a star shaped kernel (an overlay of a 0 deg and a 45 deg oriented CenSuRE kernel), in order to better estimate the LoG (Figure 1b). Given that these center-surround kernels are simply addition and subtraction summations, they can be computed extremely efficiently using integral images. The output of the STAR detector will be the salient locations and scales of patches in the image.



(a) (b) (c)

Fig. 1. (a) bi-level center-surround kernel constructed using integral images, and, (b) the STAR kernel, and (c) example of a BRIEF kernel

In order to track features from one frame to the next, we use the BRIEF [2] descriptor to describe a characteristically-scaled patch centered about each STAR location. For each STAR location, N pairs of points in an $S \times S$ square region are chosen at random, with the condition that the probability of choosing a point is equal to an isotropic Gaussian distribution centered about the STAR location (Figure 1c). A descriptor vector of length N is built to describe the patch, where the i^{th} element is either 0 (if, for the i^{th} pair, the pixel intensity of the first point is higher than that of the second point) or 1 (otherwise). This creates a binary vector that can be used to uniquely identify each feature. Feature matching is performed by calculating the Hamming distance between two feature descriptors. Following the formulation provided in [2], we use a patch size of $S = 25$ and vector length of $N = 256$.

There are a number of reasons as to why we chose to combine both STAR and BRIEF (STAR+BRIEF) to describe a single salient feature. Both methods evaluate patch-locations for saliency, and by extracting BRIEF features at characteristic scales of the STAR locations, STAR+BRIEF features can be made scale-invariant. STAR's approximation of the LoG can be well described by the

isotropic Gaussian-distributed pixel comparison vector of a BRIEF feature. Finally, both methods are fast to compute. Performance comparisons of STAR and BRIEF to popular feature detectors can be found in their original papers [1,2].

2.2 Long-Term Feature Tracking

A salient feature tracking framework for long-term tracking is developed. As opposed to other suggested frameworks for feature tracking for endoscopic video [21,20], our method does not use camera pose estimation and feature location reprojection to estimate scene movement. Rather, we allow each feature to move independently and use spatial and temporal filters to remove tracking errors, and we allow features to be lost temporarily and found in subsequent frames.

Temporal Tissue Tracking. A visual flowchart of the proposed tissue tracking framework is presented in Figure 2. We begin by capturing an image from the surgical camera, and performing a pre-processing filter with a 3×3 Gaussian smoothing kernel. We then extract the image features from the current frame (Figure 2B) and match them to features found in previous frames (Figure 2C). A list of features is maintained, where previous features that are matched have their locations and descriptors updated, and features that were not matched in the current frame are added to the list. Finally, features that are matched below a certain percentage of frames are discarded (empirically set to 40%).

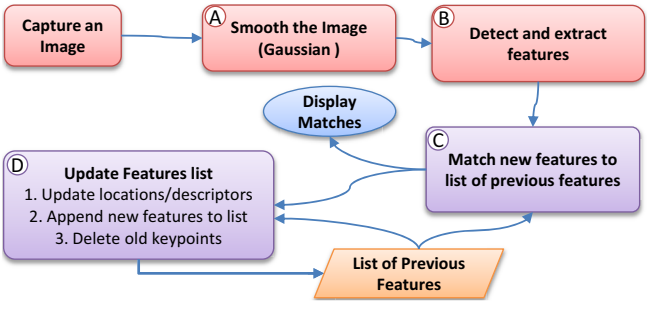


Fig. 2. Flowchart of the proposed feature tracking framework on a single image

There are a number of filters that are applied to improve feature matching accuracy. A feature is defined as $f(x, y, k)$, where x and y represent its pixel location in the original image, and k represents its characteristic scale. Given two feature lists (1 and 2), the i^{th} feature from list 1, f_i, and the j^{th} feature from list 2, f_j, we only perform descriptor comparisons on a subset of features.

– *Physical Proximity:* Since we do not expect large feature movement between consecutive frames, our search space is limited by

$$|x_i - x_j| < \delta_x \text{ and } |y_i - y_j| < \delta_y. \quad (1)$$

where δ_x and δ_y is the range of the search space.

– *Scale Similarity:* We do not expect significant changes in feature sizes between consecutive frames, and therefore limit our search space to

$$|log(k_i/k_j)| < \kappa, \tag{2}$$

where κ is the maximum allowable ratio of feature scales.

– *Descriptor Distance Ratio:* We measure the confidence of feature matches by comparing the descriptor distance of the best match, d_{first}, to the descriptor distance of the second best match, d_{second}, and only take matches if

$$d_{first}/d_{second} < \lambda, \tag{3}$$

where λ is the maximum allowable ratio between the descriptor distances.

A set of matched features will provide a dense mapping of tissue movement within the scene. Matches that move in significantly different directions than the local tissue movement can be rejected. This is performed by checking the movement of each matched feature against its nearest neighbors (matched features within a Euclidean distance of 20% of the image width).

Given a set of two neighboring feature locations in frame n, $Q_{1,n} = \{x_{1,n}, y_{1,n}\}$ and $Q_{2,n} = \{x_{2,n}, y_{2,n}\}$, and their matched locations in the previous frame, $Q_{1,n-1} = \{x_{1,n-1}, y_{1,n-1}\}$ and $Q_{2,n-1} = \{x_{2,n-1}, y_{2,n-1}\}$, we consider their movements to be significantly different if

$$\left|\log\left(\frac{\delta x_1^2 + \delta y_1^2}{\delta x_2^2 + \delta y_2^2}\right)\right| > \gamma, \tag{4}$$

where $\{\delta x_1, \delta y_1\} = \{x_{1,n} - x_{1,n-1}, y_{1,n} - y_{1,n-1}\}$ and $\{\delta x_2, \delta y_2\} = \{x_{2,n} - x_{2,n-1}, y_{2,n} - y_{2,n-1}\}$, and where γ represents the maximum allowable ratio of squared distances of movements between the neighboring features. Furthermore, the directions of their movements are considered to be significantly different if

$$\Delta\theta = \text{acos}\left(\frac{\delta x_1 \cdot \delta x_2 + \delta y_1 \cdot \delta y_2}{\sqrt{\delta x_1^2 + \delta y_1^2}\sqrt{\delta x_2^2 + \delta y_2^2}}\right) > \epsilon, \tag{5}$$

where ϵ is the maximum allowable difference in the direction of movement.

Matched features that move within ϵ and γ for more than 70% of neighboring features are accepted. We check ϵ and γ only if there is a temporal displacement of 5 pixels for each match, since the resolution of ϵ and γ decreases significantly within a Euclidean distance of 5 pixels, making these filters less effective.

Stereoscopic Tracking. Our proposed tracking algorithm can be extended to track features in 3D coordinate space. Features from a left and right channel of a stereo-endoscope are matched in order to triangulate the features' locations in 3D. The matching between features in stereoscopic channels can be filtered using the same methods as above (Equations 1, 2 and 3). Since stereo-triangulation does not take into account temporal movement, neighborhood feature movement (Equations 4 and 5) are unnecessary for stereoscopic matching.

We use one channel for temporal tracking, and the tracked features are then matched against features found in the other channel for 3D localization; we note that in the future, tracking both channels can be used for outlier rejection.

3 Experimental Setup

Table 1 describes the parameters of our tissue tracking framework. $\kappa = \log(2.0)$ is chosen to search within an octave scale of a feature, and $\kappa = \log(\sqrt{2.0})$ reflects smaller changes in scale between stereo channels; δ's are chosen assuming that features move relatively small distances between frames and between stereoscopic channels. $\lambda = 0.5$ chooses feature matches that have under half the descriptor distance of the next best match. $\gamma = 2\log(1.5)$ and $\epsilon = \pi/18$ were chosen to restrict features bundles to smooth, consistent motion, characteristic of tissue.

Four measures were used to evaluate the efficacy of STAR+BRIEF and the long-term feature tracking strategy: (a) the number of features found per frame, (b) the percentage of these features that are matched to ones in previous frames, (c) the persistence of features in subsequent frames (d) and the algorithm speed.

We investigated the tracking algorithm's ability on four endoscopic videos involving different *in-vivo* tissue movement:

Translation: Abdominal cavity after inflation. Surgeon moves the endoscopic camera to approximate translation. (1050 frames, 480×640 pixels).

Rotation: Abdominal cavity after inflation. Surgeon moves the endoscopic camera to approximate rotation. (710 frames, 480×640 pixels).

Series: Abdominal cavity after inflation. Surgeon moves the endoscopic camera to approximate a series of movements involving translation and scaling with slight rotations. (1200 frames, 480×650 pixels)

Heartbeat: Open-chest procedure with an exposed heart. Significant surgical clamps footprint in the image. A stationary camera images a heartbeat. (650 frames, 720×576 pixels).

These videos, acquired by Imperial College London, are available at `http://hamlyn.doc.ic.ac.uk/vision/`

Table 1. Parameters for temporal and stereoscopic matching

Parameter	Symbol	Value (Temporal)	Value(Stereo)
Scale Threshold	κ	$\log(2.0)$	$\log(\sqrt{(2)})$
Local Area Threshold	δ	0.2∗image_width	0.5∗image_width on x-axis 0.05∗image_height in y-axis
Descriptor Distance Ratio	λ	0.5	0.5
Difference in movements	γ	$2 * \log(1.5)$	N/A
Difference in angles	ϵ	$\pi/18$	N/A

4 Results

Figure 3a,b shows two example (Series and Heartbeat videos) of STAR+BRIEF features that are tracked temporally over time. The number of features extracted by the STAR+BRIEF framework fully describe the deformation of the scene. The average number of features tracked are shown in Figure 3c. Figure 3d shows that the STAR+BRIEF framework is able to match 90% of its features from frame to frame for general cases, and over 50% can be matched stereoscopically for 3D localization. Matches were visually inspected and nearly all matches were found to correctly track a correct physical location. We found that fewer features were tracked in the heartbeat video; we believe this can be attributed to the greater degree of specular reflection and significant occlusion from blood, and also the systolic action of the heart, which exhibits frequencies higher than 100Hz that cannot be adequately captured in the endoscopic cameras with conventional framerates. Furthermore, slight discrepancies in triggering during steroscopic frame-grabbing can account for significant movement between stereo channels.

Figure 3e shows a histogram of the percentage of subsequent frames in which features are matched. The histograms are cumulative, where the first column (40% to 50% persistence) is the total number of features being continuously tracked by long-term tracking framework. Features below 40% persistence were not stable and therefore were discarded. The number of features continuously tracked over increasing percentage of frames is fairly linear for general motion. Heartbeat motion shows less features being tracked continuously, and a non-linear dropoff in persistence of features due to the lower stability of the features from dynamic motion, specular reflection and blood occlusion.

Figure 3f shows the speed of our tracking framework with the STAR+BRIEF method. We show that the our long-term tracking framework is able to process sequences above 10Hz and is capable of performing stereoscopic matching and 3D deformation tracking at a fraction of the total tracking time. The heartbeat video, due to the higher resolution and increased specular reflection and occlusion, identified many more features than other videos that were unstable and therefore required significantly higher computation effort than other videos.

Figure 4 shows a sample of the tracking of a feature in the Series video, which involves translation, scaling, and some rotation. A useful method for getting a sense of the aversion to drift, given that we do not have a ground-truth source, is forward-backward tracking [3] shown in Figure 4a, where we track a feature both forwards and backwards in time. Since our long-term feature tracking framework is history-dependent and therefore behaves differently moving in forwards and backwards time, it can be seen that the feature location does not drift. Figure 4b shows the feature location at different timesteps, indicating its stability throughout the sequence of motions.

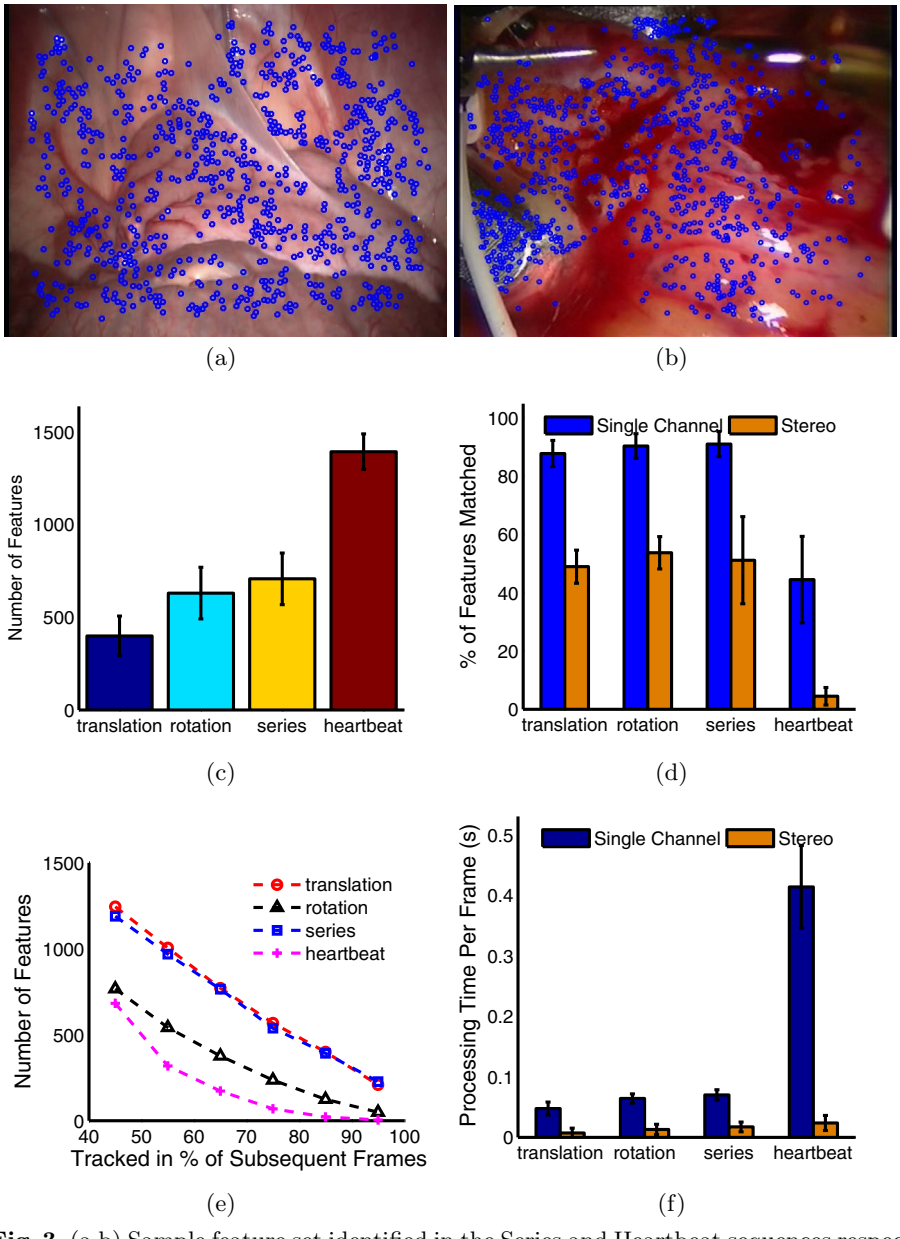


Fig. 3. (a,b) Sample feature set identified in the Series and Heartbeat sequences respectively. (c) Number of Features found per frame; (d) % of features in the current frame that are matched to previous features (blue) and % of features matched stereoscopically (orange); (e)The number of features that are found in % of subsequent frames. The graph is cumulative such that the number of features drops off as the % of matching in subsequent frames increases. (f) Time required for a complete cycle of the feature tracking framework (blue) and the amount of time required for stereoscopic matching (orange).

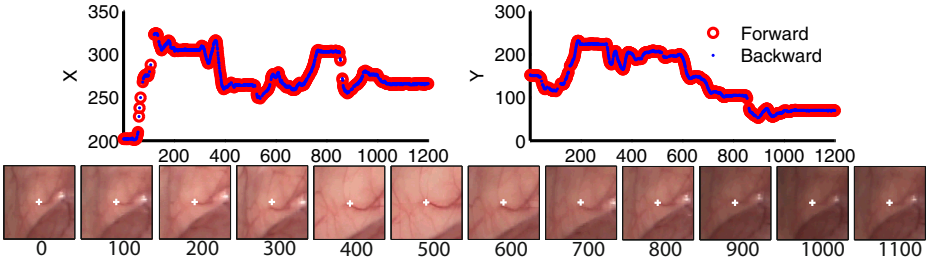


Fig. 4. A STAR+BRIEF feature being tracked over time (1200 frames). Top: tracking performed in both a forward and a backward time direction. Graph shows X and Y coordinates as a function of time. Bottom: a 50×50 pixel window centered about the feature location at every 100 frames.

5 Discussion

We have presented a novel framework for long-term tracking of endoscopic tissue scenes. This framework uses salient features for populating the endoscopic images in order to track the deformation of tissue at high densities. We have shown that by using two new, efficient feature algorithms, CenSuRE/STAR and BRIEF, we are able to achieve both high density tracking on standard definition video for long-periods of time at real-time (greater than 10 Hz) speeds. Other fast feature detectors, as well as GPU-accelerated feature tracking could also be used to achieve fast and dense feature tracking under this framework.

The performance of the algorithm is not reliant on a rigid scene or camera pose estimation and therefore can handle scene deformations, as tissue features are being tracked individually and therefore can describe complex deformations. However, in the case of significant high-frequency deformations (e.g. beating heart), the feature-based approach must widen its search space between frames in expectation of large dynamic movements of individual features. A reduction of the effects of specular reflection and blood occlusion are required to improve tracking, and several strategies can be used such as intensity thresholding and interpolation [4]. In the case of the beating heart, other strategies besides feature tracking, such as the ones presented in [13] can be more effective. Further efforts to handle specular reflection effects and to track instrument occlusion will help enable the long-term tracking strategy to track tissue through an entire surgery.

Given the long-term tissue feature tracking strategy we defined, it may be possible to maintain a registered medical image to endoscopic cameras by tracking the underlying tissue movements; this will provide the surgeon the ability to localize sub-surface tissue features such as nerves and lesions for better surgical guidance. Furthermore, the ability to acquire 3D tracking of dense feature maps may enable a surface-to-feature based registration method, such that the ability to register and maintain a medical image to tissues in the endoscopic cameras can be effectively streamlined. Future work will be to investigate the efficacy of registering and maintaining a medical image to the tissues seen in

the laparoscopic cameras, with the motivation of improving surgical guidance in LRP, where the dissection of the prostate from the surrounding neurovascular bundle must be performed within narrow margins (1.9 ± 0.8 mm [19]).

Acknowledgments. The authors wish to thank NSERC and the CIHR Knowledge Translation Fund for supporting this work, as well as Professor Guang-Zhong Yang and the Hamlyn Centre for Robotic Surgery at Imperial College of London for open access to their *in-vivo* data sets. Research support from the C.A. Laszlo Chair held by Professor Salcudean is gratefully acknowledged.

References

1. Agrawal, M., Konolige, K., Blas, M.: CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 102–115. Springer, Heidelberg (2008)
2. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: BRIEF: Binary Robust Independent Elementary Features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6314, pp. 778–792. Springer, Heidelberg (2010)
3. Kalal, Z., Mikolajczyk, K., Matas, J.: Forward-backward error: Automatic detection of tracking failures. In: IEEE Int. C. Pattern Recognition, pp. 2756–2759 (2010)
4. Lo, B., Chung, A., Stoyanov, D., Mylonas, G., Yang, G.-Z.: Real-time intra-operative 3d tissue deformation recovery. In: I. S. Biomedical Imaging, pp. 1387–1390 (2008)
5. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Computer Vision 60, 91–110 (2004)
6. Luó, X., Feuerstein, M., Reichl, T., Kitasaka, T., Mori, K.: An Application Driven Comparison of Several Feature Extraction Algorithms in Bronchoscope Tracking During Navigated Bronchoscopy. In: Liao, H., Edwards, P.J., Pan, X., Fan, Y., Yang, G.-Z. (eds.) MIAR 2010. LNCS, vol. 6326, pp. 475–484. Springer, Heidelberg (2010)
7. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE T. Pattern Anal. 27, 1615–1630 (2005)
8. Mountney, P., Stoyanov, D., Davison, A., Yang, G.-Z.: Simultaneous Stereoscope Localization and Soft-Tissue Mapping for Minimal Invasive Surgery. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 347–354. Springer, Heidelberg (2006)
9. Mountney, P., Lo, B., Thiemjarus, S., Stoyanov, D., Zhong-Yang, G.: A Probabilistic Framework for Tracking Deformable Soft Tissue in Minimally Invasive Surgery. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 34–41. Springer, Heidelberg (2007)
10. Mountney, P., Yang, G.-Z.: Soft Tissue Tracking for Minimally Invasive Surgery: Learning Local Deformation Online. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 364–372. Springer, Heidelberg (2008)
11. Mountney, P., Stoyanov, D., Yang, G.-Z.: Three-dimensional tissue deformation recovery and tracking. IEEE Signal Processing Magazine 27, 14–24 (2010)

12. Ortmaier, T., Groeger, M., Boehm, D., Falk, V., Hirzinger, G.: Motion estimation in beating heart surgery. IEEE T. Biomed. Eng. 52(10), 1729–1740 (2005)
13. Richa, R., Bo, A. P., Poignet, P.: Towards robust 3D visual tracking for motion compensation in beating heart surgery. Med. Image Anal. 15(3), 302-315 (2011)
14. Rosten, E., Drummond, T.: Machine Learning for High-Speed Corner Detection. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 430–443. Springer, Heidelberg (2006)
15. Sauvee, M., Noce, A., Poignet, P., Triboulet, J., Dombre, E.: Three-dimensional heart motion estimation using endoscopic monocular vision system: From artificial landmarks to texture analysis. Biomed. Signal Process. Control 2(3), 199–207 (2007)
16. Shi, J., Tomasi, C.: Good features to track. In: IEEE Int. C. Computer Vision and Pattern Recognition, pp. 593–600 (1994)
17. Sinha, S. N., Frahm, J.-M., Pollefeys, M., Genc, Y.: Gpu-based video feature tracking and matching. In: Workshop on Edge Computing Using New Commodity Architectures, Technical Report (2006)
18. Stoyanov, D., Mylonas, G.P., Deligianni, F., Darzi, A., Yang, G.-Z.: Soft-Tissue Motion Tracking and Structure Estimation for Robotic Assisted MIS Procedures. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3750, pp. 139–146. Springer, Heidelberg (2005)
19. Ukimura, O., Gill, I., Desai, M.O.: Real-time transrectal ultrasonography during laparoscopic radical prostatectomy. Journal of Urology 172(1), 112–118 (2004)
20. Wang, H., Mirota, D., Ishii, M., Hager, G.: Robust motion estimation and structure recovery from endoscopic image sequences with an adaptive scale kernel consensus estimator. In: Int. C. Computer Vision and Pattern Recognition, pp. 1–7 (2008)
21. Wengert, C., Cattin, P.C., Duff, J.M., Baur, C., Székely, G.: Markerless Endoscopic Registration and Referencing. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 816–823. Springer, Heidelberg (2006)
22. Willow Garage. Star detector, `http://pr.willowgarage.com/wiki/star_detector`