Catheter Tracking:
Filter-Based vs. Learning-Based

Alexander Brost[1], Andreas Wimmer[1], Rui Liao[2], Joachim Hornegger[1] and
Norbert Strobel[3]

[1] Pattern Recognition Lab, Department of Computer Science,
Friedrich-Alexander-University of Erlangen-Nuremberg, Erlangen, Germany
[2] Siemens Corporate Research, Princeton, NJ, USA
[3] Siemens AG, Forchheim, Germany

Abstract. Atrial fibrillation is the most common sustained arrhythmia. One important treatment option is radio-frequency catheter ablation (RFCA) of the pulmonary veins attached to the left atrium. RFCA is usually performed under fluoroscopic (X-ray) image guidance. Overlay images computed from pre-operative 3-D volumetric data can be used to add anatomical detail otherwise not visible under X-ray. Unfortunately, current fluoro overlay images are static, i.e., they do not move synchronously with respiratory and cardiac motion. A filter-based catheter tracking approach using simultaneous biplane fluoroscopy was previously presented. It requires localization of a circumferential tracking catheter, though. Unfortunately, the initially proposed method may fail to accommodate catheters of different size. It may also detect wrong structures in the presence of high background clutter. We developed a new learning-based approach to overcome both problems. First, a 3-D model of the catheter is reconstructed. A cascade of boosted classifiers is then used to segment the circumferential mapping catheter. Finally, the 3-D motion at the site of ablation is estimated by tracking the reconstructed model in 3-D from biplane fluoroscopy. We compared our method to the previous approach using 13 clinical data sets and found that the 2-D tracking error improved from 1.0 mm to 0.8 mm. The 3-D tracking error was reduced from 0.8 mm to 0.7 mm.

1 Motivation

Recent research in the area of X-ray guidance for electrophysiology (EP) procedures found that augmented fluoroscopy using overlay images rendered from 3-D images (CT, MRI, C-Arm CT) facilitates more precise catheter navigation and a reduction in fluoroscopy time [1–3]. Critical structures like the esophagus and the left atrial appendage are invisible under regular fluoroscopy unless contrast agent is applied. Thus, rendering overlays of such structures for visual procedure guidance further improves safety. Unfortunately, current image overlay methods still lack motion compensation. A first approach tracking a commonly used mapping catheter has been proposed in [4]. This circumferential mapping catheter measures the electrical potentials at the ostium of the pulmonary vein (PV)

(a) ROI   (b) Filtering   (c) Classification
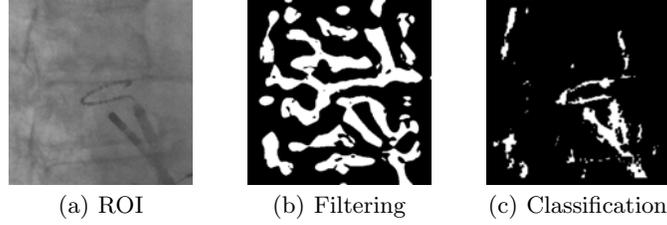
Fig. 1. (a) A region-of-interest around the catheter is presented on the left. (b) The result of the segmentation by filter methods [4, 7, 8]. (c) Segmentation using a boosted classifier cascade.

considered for ablation. Catheter tracking was accomplished by calculating a 3-D model of the catheter, a filter-based segmentation and 2-D/3-D registration of the catheter model to biplane fluoroscopic images. Various methods for catheter tracking have been proposed. They involve filter-based techniques [5] as well as template-matching and learning-based approaches [6]. For electrophysiology procedures, different types of catheters are available. They differ in width, number of electrodes and electrode spacing. These parameters have to be considered explicitly for a filter-based catheter tracking approach. This adds additional complexity often reduces the robustness of filter-based approaches. This is why learning-based methods are often preferable for more complicated pattern recognition problems. They can reach a better performance if the training set is sufficiently comprehensive to capture all relevant catheter features encountered in clinical practice. Learning-based algorithms are usually superior with respect to suppressing interfering structures that are not of interest. This is demonstrated in Fig. 1 presenting a comparison between segmentation results obtained using filtering [4, 7, 8] and by classification. We used a boosted classifier cascade to segment the circumferential mapping catheter. In the next step, the 3-D catheter model is generated as proposed in [4, 7, 8]. Tracking itself is performed by a 2-D/3-D registration. These steps are explained in more detail below.

2   Catheter Model Generation

Our method requires the generation of a 3-D catheter model, which is based on the assumption that the perspective projection of the circumferential mapping catheter, when fit to the pulmonary veins, can be approximated as a 3-D ellipse. The circumferential mapping catheter can also be approximated as an ellipse in 2-D, because a 3-D elliptical object remains elliptical when perspectively projected onto a 2-D imaging plane. The ellipses on the 2-D fluoroscopic images are denoted as $\boldsymbol{C}_{\mathrm{A/B}} \in \mathbb{R}^{3\times3}$, with the index A or B indicating the corresponding imaging plane of the C-arm. A 3-D elliptical cone can then be spanned with the projection matrix $P_{\mathrm{A/B}} \in \mathbb{R}^{3\times4}$ and the ellipse within the imaging plane. The base of the elliptical cone is the ellipse in the imaging plane and the vertex is

the optical center. It can be shown that the elliptical cone can be represented as

$$\boldsymbol{Q}_{\mathrm{A/B}} = \boldsymbol{P}_{\mathrm{A/B}}^{T}\boldsymbol{C}_{\mathrm{A/B}}\boldsymbol{P}_{\mathrm{A/B}} \tag{1}$$

in matrix presentation [9]. The 3-D ellipse representing the 3-D mapping catheter is reconstructed by intersecting the two elliptical cones $\boldsymbol{Q}_A$ and $\boldsymbol{Q}_B$ corresponding to plane A and plane B of a biplane system respectively. The solution is found by calculating η such that the quadric

$$\boldsymbol{Q}(\eta) = \boldsymbol{Q}_A + \eta\boldsymbol{Q}_B \tag{2}$$

is of rank 2 [9]. As pointed out in [4, 7, 8], there are two possible solutions. Prior knowledge about the pseudo-circular shape of the mapping catheter is used and the result that is more circular is chosen.

3   Classifier Cascade

The catheter segmentation method not only has to be reliable, but it needs to be fast as well. Speed is necessary to ensure that the catheter can be tracked in real-time at the frame rate set at the X-ray acquisition system. We found that a combination of Haar-like features and a cascade of boosted classifiers met both requirements to differentiate the live fluoroscopic images into *catheter* and *background*. Haar-like features [10] calculate various patterns of intensity differences. Several feature prototypes are listed in Fig. 2(a). Some features detect edges, whereas others focus on line structures. Especially the latter are useful for detecting the circumferential mapping catheter, which often appears as a thin, elongated object with a loop at its end, see Fig. 1(a). Actual features are obtained by shifting and scaling the prototypes within a predefined window. In our case, a window size of 15×15 was found to be sufficient for good results. Thereby, contextual information around the center pixel is considered, which is important to differentiate between catheter and background structures. However, even for moderate window sizes, the resulting number of features is large and easily amounts to several hundreds of thousands. Features are calculated efficiently through integral images [10]. To achieve reliable and fast segmentation, the most suitable features for discriminating between catheter and background have to be chosen and integrated into a classifier in a suitable manner. This is carried out by the AdaBoost algorithm [11]. The idea is to combine several weak classifiers, to form a strong classifier. The classifier minimizing the classification error is added to a linear combination of weak classifiers until the overall error is below the desired threshold. After each training iteration, the importance of individual samples is re-weighted to put more emphasis on misclassifications for the next evaluation. Instead of single features and intensity thresholds, we use classification and regression trees (CARTs) [12] as weak classifiers. A CART is a small tree of fixed size. At each node, a threshold θ_j associated with a feature partitions the feature space. This way, flexibility is increased and objects with complex feature distributions can be handled. The result of a CART is the

(a)   Feature   (b) Classification and Regression Tree   (c) Cascade
Types
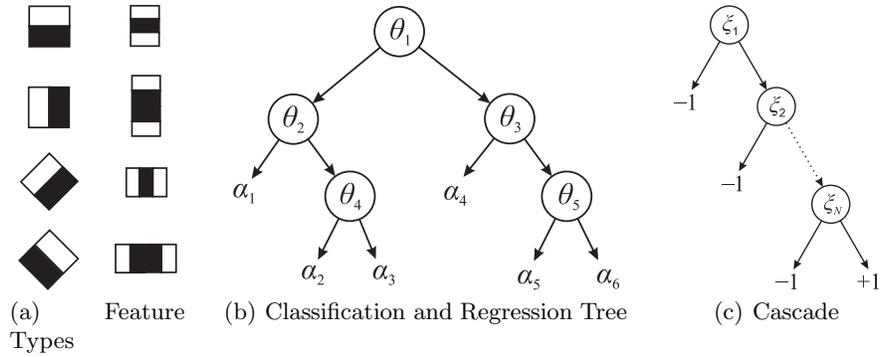
Fig. 2. Features types and classifier structure for catheter segmentation. (a) Several prototypes of Haar-like features. (b) Exemplary classification and regression tree (CART) with five feature nodes $\theta_1, \ldots, \theta_5$ and six leaves $\alpha_1, \ldots, \alpha_6$. (c) Classifier cascade consisting of N stages with strong classifiers $\xi_1, \ldots \xi_N$; each strong classifier ξ_i consists of a linear combination of weak classifiers, here CARTs.

value α_k of the classifier reached as leave node. An exemplary CART is shown in Fig. 2(b). We organize N strong classifiers $\xi_i, \ldots, \xi_N$ composed of weighted combinations of CARTs into a cascade, which is illustrated in Fig. 2(c). In our case, four strong classifiers ($N = 4$) yielded good results. At each stage, a sample is either rejected (-1) or passed on to the next stage. Only if the sample is accepted ($+1$) at the final stage, it is accepted as part of the object. Thus during training, the focus is on maintaining a high true positive rate while successively reducing the false positive rate, either by adding more weak classifiers to a stage or by adding an entirely new stage. The training data set consisted of 13 clinical data sets with a total of 938 monoplane frames. For evaluation, the classifier cascade was trained on a leave-one-out basis, i.e., 12 sequences were used for training and the remaining sequence was used for segmentation and tracking.

4   Tracking by Registration

The elliptical shape of the circumferential mapping catheter is used for tracking. Catheter tracking itself is performed by rigid registration [13] of the catheter model to the segmentation result derived from the previous step. To this end, the same ROI as for the classification is used. As the size of the mapping catheter may not be available beforehand, the previous approach in [4, 7, 8] used normals to the ellipse to simulate the width of the catheter. As the length and the sampling of the normals depends on further parameters that need to be adjusted, we use a thinning algorithm as proposed in [14]. By thinning, we generate a skeleton of the catheter, which involves far fewer parameters than needed otherwise. A distance map $\boldsymbol{I}_{\mathrm{DT,A/B}}$ is calculated from the skeleton as proposed in [15] for each imaging plane. It encodes the absolute distance from a pixel to its closest

segmented catheter pixel. It also provides a smooth representation of the fluoroscopic image with a pronounced minimum around the shape of the mapping catheter to increase the capture range. Model-based catheter tracking in 3-D is achieved by performing 2-D/3-D registration. Hence, the reconstructed catheter model is rotated by $\boldsymbol{R} \in \mathbb{R}^{4 \times 4}$ and translated by $\boldsymbol{T} \in \mathbb{R}^{4 \times 4}$ first. It is then projected onto the two imaging planes of the bi-plane C-arm system. The average distance between the projected points and the closest feature point (i.e. the circumferential mapping catheter) in fluoroscopic images is efficiently calculated using the distance map introduced above. A suitable rotation and translation is found by optimizing

$$\hat{\boldsymbol{R}}, \hat{\boldsymbol{T}} = \arg \min_{\boldsymbol{R}, \boldsymbol{T}} \sum_i \boldsymbol{I}_{\mathrm{DT}}(\boldsymbol{P}_A \cdot \boldsymbol{T} \cdot \boldsymbol{R} \cdot \boldsymbol{w}_i) + \sum_i \boldsymbol{I}_{\boldsymbol{DT}}(\boldsymbol{P}_B \cdot \boldsymbol{T} \cdot \boldsymbol{R} \cdot \boldsymbol{w}_i) \qquad (3)$$

with the 3-D catheter model points $\boldsymbol{w}_i \in \mathbb{R}^4$ in homogeneous coordinates. The projection matrices $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ do not need to be identical to the ones in Eq. 1. The parameters used for optimizing are three rotation angles around the main axes in 3-D, combined in $\boldsymbol{R}$, as well as a three-dimensional translation, represented in $\boldsymbol{T}$. As optimization strategy, a nearest neighbor search [16] is used, i.e., the position of the local optimum on a large scale is taken as starting point for the optimization on a smaller scale. The estimated 3-D rotation and translation can be directly applied to the 2-D overlay to move it in sync with the tracked device. An overview of the reference method [4, 7, 8] is presented in Fig. 3 and an overview of our proposed algorithm is given in Fig. 4.

5   Evaluation and Results

Our approach was evaluated on 13 clinical data sets, collected from 6 different patients at one clinical site. Three different circumferential mapping catheters were used. For evaluation, we calculate the 2-D tracking error as the average 2-D distance between the projection of the 3-D catheter model and a 2-D gold-standard segmentation of the circumferential mapping catheter provided by a cardiologist. We compare our results with those in [4, 7, 8], see Fig. 5. Clinical data set No. 6 contains a barium swallow of the patient to outline the esophagus, a critical structure during ablation. Unfortunately, in one frame of this sequence, the barium hides the mapping catheter, resulting in a rather high maximum error. This single frame was not excluded. Since catheter tracking is performed in 3-D, we follow the evaluation in [7, 8] to estimate the 3-D motion correction. Therefore, the tip of the mapping catheter was manually localized throughout all sequences by triangulating its 3-D position from bi-plane frames to get a reference point. In the next step, we applied our motion estimation approach to the catheter tip to move it from its 3-D position in the previous frame to the next frame. Because of that, we can compare the 3-D position reached by applying the estimated motion to the actual 3-D reference point obtained by triangulation [17]. Finally, the error was calculated as the Euclidean distance in 3-D space. Moreover, an error without performing motion compensation can be
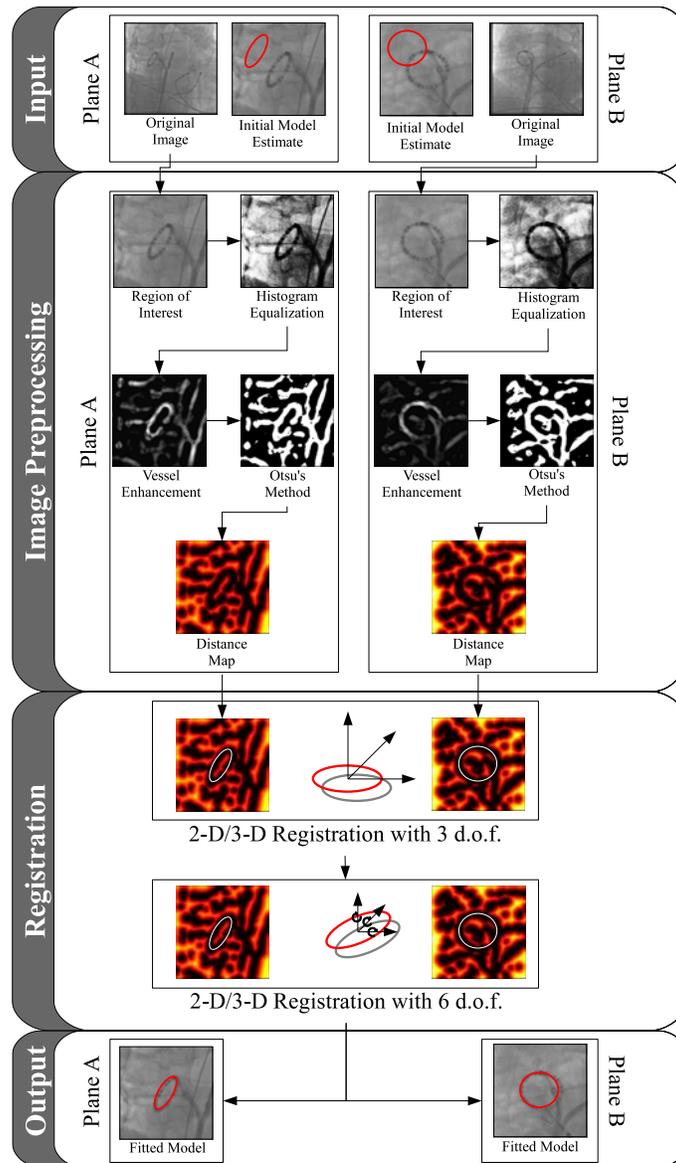
Fig. 3. Flow diagram of the filter-based catheter tracking approach for motion compensation [8]. The objective is to obtain a dynamic fluoroscopic overlay image for improved catheter navigation.
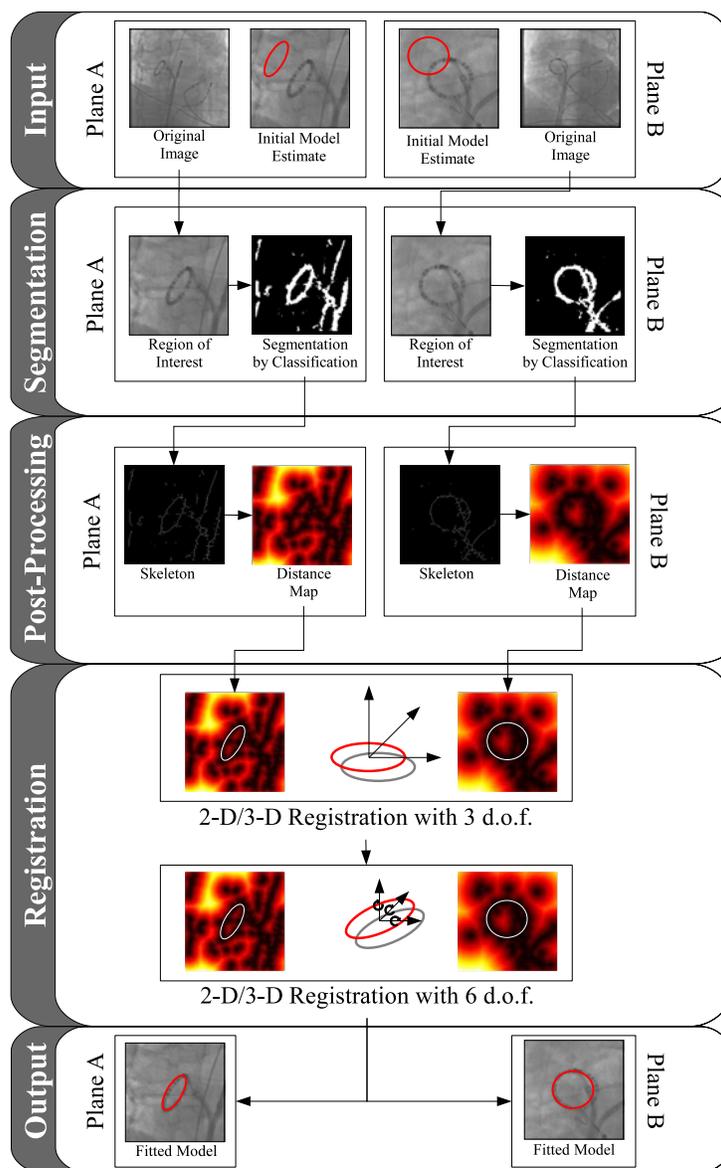
Fig. 4. Flow diagram of our learning-based motion compensation approach. Image pre-processing has been replaced by classification-based segmentation followed by a post-processing step. The goal is to obtain a distance map.
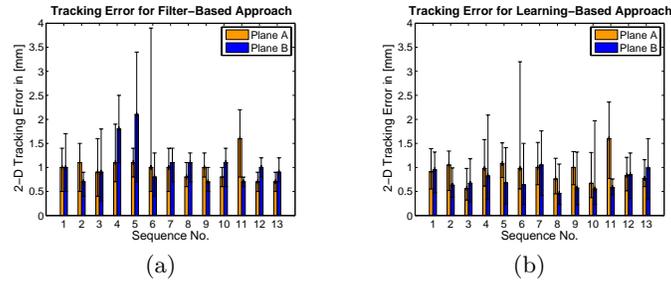
Fig. 5. Comparison of the different catheter tracking approaches. (a) The result of the reference method presented in [4, 7, 8]. This method yielded an average 2-D tracking error of 1.0 mm $\pm$ 0.4 mm. (b) Our improved algorithm using a boosted classifier cascade to segment the circumferential mapping catheter yielded a 2-D error of 0.8 mm $\pm$ 0.4 mm.

calculated was well. To this end, the 3-D distance between the first frame to all remaining frames is used to estimate the observed 3-D motion. A comparison of the observed 3-D motion to both motion correction approaches is shown in Fig. 6(a). A direct comparison of the filter-based approach versus the learning-based method is given in Fig. 6(b).

6   Discussion and Conclusions

We presented a method for 3-D motion estimation for radio-frequency catheter ablation of atrial fibrillation. It is based on tracking of a circumferential mapping catheter in biplane fluoroscopy imaging. Catheter tracking is performed by 2-D/3-D registration of a 3-D elliptical catheter model to 2-D biplane images. The method assumes that the circumferential mapping catheter remains anchored at the pulmonary vein during ablation. Our clinical data suggests that the circumferential mapping catheter indeed moves very little with respect to the PV ostia when used to measure the electrical signals at the pulmonary ostia. When comparing the two tracking approaches, it can be seen, that the learning-based approach performs better than the filter-based approach overall. The 2-D tracking error could be reduced from 1.0 mm $\pm$ 0.4 mm to 0.8 mm $\pm$ 0.4 mm and the 3-D error from 0.8 mm $\pm$ 0.5 mm to 0.7 mm $\pm$ 0.4 mm. This has been displayed in Fig. 1 showing one frame of Seq. 5. Here, the learning-based method reduced the average 3-D tracking error from 2.0 mm to 1.2 mm. The benefit of motion compensation for X-ray fluoroscopy guidance can be appreciated by looking at contrast-enhanced anatomical structures. An example is presented in Fig. 7. The main advantage of the proposed learning-based approach is its robustness. In particular, it generates fewer false positives as obtained when using the filter-based approach. The advantage of 2-D/3-D registration with biplane images is that it provides a more robust 3-D tracking compared to a monoplane situation. The quality of the segmentation could be further improved by using either more

(a)   (b)
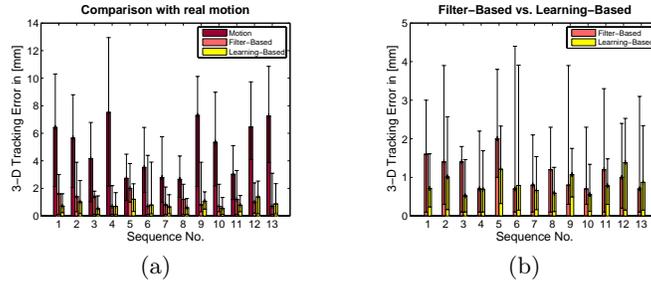
Fig. 6. Comparison of the different catheter tracking approaches. (a) The result of the filter-based method [4, 7, 8] and the learning-based method compared to the actual 3-D motion. An average 3-D motion of 4.5 mm ± 2.4 mm has been observed in our clinical data. (b) Direct comparison of the 3-D tracking error of the filter-based and the learning-based method. The filter-based method yielded an average 3-D tracking error of 0.8 mm ± 0.5 mm, whereas the learning-based approach yielded an average error of 0.7 mm ± 0.4 mm.

(a)   (b)   (c)
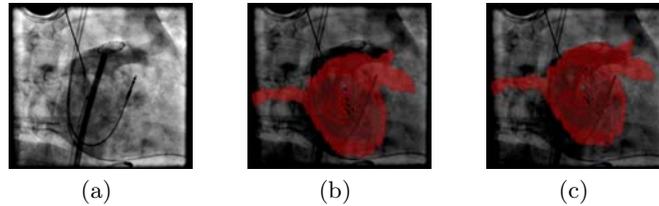
Fig. 7. Motion compensation can be visually assessed by contrast injection into a pulmonary vein. (a) One image of a fluoroscopic sequence showing the administration of contrast agent. (b) Misfit of a static overlay due to motion. (c) Visual overlay with improved positioning of the (red) fluoroscopic overlay image thanks to motion compensation.

cascades or an increased number of features, but this goes hand in hand with higher computational time. Our current implementation reaches a processing speed of one frame-per-second (fps), as this frame rate is used at our clinical partner for simultaneously biplane fluoroscopy. The presented method has not yet been optimized for multi-core CPUs or GPUs. Further speed improvements can be expected by taking advantage of parallel processing options. Moreover, a classifier cascade was used to segment the catheter, but other segmentation approaches could be applied as well. Beyond improving the technology, future work will also focus on clinical evaluation of this method.

References

1. Ector, J., De Buck, S., Huybrechts, W., Nuyens, D., Dymarkowski, S., Bogaert, J., Maes, F., Heidbüchel, H.: Biplane three-dimensional augmented fluoroscopy as

single navigation tool for ablation of atrial fibrillation: Accuracy and clinical value. Heart Rhythm **5**(7) (March 2008) 957–964

2. Sra, J., Narayan, G., Krum, D., Malloy, A., Cooley, R., Bhatia, A., Dhala, A., Blanck, Z., Nangia, V., Akhtar, M.: Computed Tomography-Fluoroscopy Image Integration-Guided Catheter Ablation of Atrial Fibrillation. J Cardiovasc Electrophysiol **18**(4) (April 2007) 409–414

3. De Buck, S., Maes, F., Ector, J., Bogaert, J., Dymarkowski, S., Heidbüchel, H., Suetens, P.: An augmented reality system for patient-specific guidance of cardiac catheter ablation procedures. IEEE Transactions on Medical Imaging **24**(11) (November 2005) 1512–1524

4. Brost, A., Liao, R., Hornegger, J., Strobel, N.: 3-D Respiratory Motion Compensation during EP Procedures by Image-Based 3-D Lasso Catheter Model Generation and Tracking. In: Lect. Notes Comput. Sci. Volume 5761., London (2009) 394–401

5. Palti-Wassermann, D., Brukstein, A., Beyar, R.: Identifying and Tracking a Guide Wire in the Coronary Arteries During Angioplasty from X-Ray Images. IEEE Transactions on Biomedical Engineering **44**(2) (February 1997) 152–164

6. Barbu, A., Athitsos, V., Georgescu, B., Boehm, S., Durlak, P., Comaniciu, D.: Hierarchical Learning of Curves: Application to Guidewire Localization in Fluoroscopy. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007. (2007) 1–8

7. Brost, A., Liao, R., Hornegger, J., Strobel, N.: 3D model-based catheter tracking for motion compensation in EP procedures. In: Medical Imaging 2010: Visualization, Image-Guided Procedures, and Modeling. Volume 7625., SPIE (2010) 762507

8. Brost, A., Liao, R., Strobel, N., Hornegger, J.: Respiratory motion compensation by model-based catheter tracking during EP procedures. Medical Image Analysis (In Press, Corrected Proof) (2010)

9. Quan, L.: Conic Reconstruction and Correspondence From Two Views. IEEE Trans Pattern Anal Mach Intell **18**(2) (February 1996) 151–160

10. Viola, P., Jones, M.: Robust real-time face detection. Int. J. Comput. Vision **57**(2) (2004) 137–154

11. Freund, Y., Schapire, R.: A decision-theoretic generalization of on-line learning and an application to boosting. JCSS **55**(1) (1997) 119–139

12. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Chapman & Hall, New York, USA (1984)

13. Hill, D., Batchelor, P., Holden, M., Hawkes, D.: Medical image registration. Phys. Med. Biol. **46**(3) (2001) R1–R45

14. Cychosz, J.: Efficient Binary Image Thinning using Neighborhood Maps. Graphics Gems IV (1994) 465–473

15. Breu, H., Gil, J., Kirkpatrick, D., Werman, M.: Linear time Euclidean distance transform algorithms. IEEE Trans Pattern Anal Mach Intell **17** (May 1995) 529–533

16. Duda, R., Hart, P.: Pattern Classification. Second edn. John Wiley & Sons, Inc (Aug. 2000)

17. Brost, A., Strobel, N., Yatziv, L., Gilson, W., Meyer, B., Hornegger, J., Lewin, J., Wacker, F.: Geometric Accuracy of 3-D X-Ray Image-Based Localization from Two C-Arm Views. In: Workshop on Geometric Accuracy In Image Guided Interventions - Medical Image Computing and Computer Assisted Interventions 2009, London UK, MICCAI (September 2009) 12–19