Fully Automatic Catheter Localization in C-Arm Images Using $\ell 1$-Sparse Coding

Fausto Milletari, Vasileios Belagiannis, Nassir Navab, and Pascal Fallavollita

Chair for Computer Aided Medical Procedures,
Technical University of Munich, Germany

Abstract. We propose a method to perform automatic detection and tracking of electrophysiology (EP) catheters in C-arm fluoroscopy sequences. Our approach does not require any initialization, is completely automatic, and can concurrently track an arbitrary number of overlapping catheters. After a pre-processing step, we employ sparse coding to first detect candidate catheter tips, and subsequently detect and track the catheters. The proposed technique is validated on 2835 C-arm images, which include 39,690 manually selected ground-truth catheter electrodes. Results demonstrated sub-millimeter detection accuracy and real-time tracking performances.

1 Introduction

Sudden cardiac death (SDC) is linked to severe disorders of the heart rhythm. In the United States alone, the incidence rate ranges up to 450,000 cases annually [1]. Patients affected by heart beat related diseases can be definitively treated with radio-frequency (RF) catheter ablation. The efficacy of catheter ablation is highly dependent on accurate identification of the site of origin of the arrhythmia. Once this site has been identified, an ablation catheter is positioned in direct contact with it and radio-frequency energy is delivered to ablate it.

Catheter ablation is often a long procedure requiring significant fluoroscopy exposure. It was proved recently [2], that 3D navigation systems contribute to the reduction of the exposure to patients and operators. The common mapping technologies that combine 3D anatomy and electrophysiological data are: CARTO and CARTOMerge (Biosense Webster), NavX (St.Jude Medical), and RPM (Cardiac Pathways-Boston Scientific). Other technologies that provide continuous data of all electrophysiogical events include Ensite 3000 (St. Jude Medical) and Basket (Cardiac Pathways-EP Technologies) [3]. Whether using mapping systems or conventional RF ablation techniques, clinicians still rely on C-arm images to position and guide catheters. Thus, exploiting C-arm image information is crucial for providing additional information to clinicians during cardiac ablation procedures. There are several reasons as to why detecting and tracking the position of ablation catheters relative to the patient anatomy is important. They are related to interventional guidance aspects: (i) accounting for heart motion compensation, (ii) easing positioning & navigation during cardiac ablation, (iii) planning the ablation procedure by (iv) registration to preoperative data such as CT and MRI.

P. Golland et al. (Eds.): MICCAI 2014, Part II, LNCS 8674, pp. 570–577, 2014.

Literature Review (2004-2013): In recent research practice, the medical imaging community has refocused its efforts to localize catheters directly in C-arm images. Fallavollita et al. developed a catheter tip detection algorithm based on thresholds of the fluoroscopic images; this failed in low contrast images [4]. A technique for tracking and detecting the ablation catheter in X-ray images was first proposed by Franken et al. but the computational cost was relatively high making the method not applicable in clinic [5]. Coronary Sinus and ablation catheter detections were first proposed in [6,7]. Multiple user interaction and parameter fine-tunings were necessary to meet the quality of the X-ray image. Employing respiration and motion compensation methods may succeed in overcoming some of the above challenges. Recently, Schenderlein et al. proposed a catheter tracking method using snakes active contour models [8]. Brost et al. developed a model-based lasso catheter tracking algorithm in biplane X-ray fluoroscopy [9]. However, the tracking required re-initialization and user interaction. Wen et al. successfully tracked one catheter in a cardiac cycle and required user-initialization in selecting tip electrodes [10,11]. Multiple catheter-tip detections are presented in [12]. There, authors require user interaction for their detections using a geodesic framework. Finally, methods including fast blob detections, clustering, shape-constrained searching and catheter model-based detection have been proposed [13,14]. A limitation of these is that they assume fixed shape for the catheter and might not cope with different C-arm positions and catheter shape changes due to foreshortening.

Contributions: We propose a unique method that considers all of the key challenges associated with catheter detections. Our method: (i) is fully automatic; (ii) supports the presence of multiple, touching and overlapping catheters; (iii) can detect and track catheters appearing foreshortened or deformed; (iv) is robust to illumination variations and to the sudden motion of the catheters.

2 Methodology

Our catheter tracking and detection pipeline is shown in *Figure* 1. The preprocessing step aims to improve the image signal to noise ratio and to reduce the search space. A further reduction of the search space is obtained in the catheter tips detection stage, where image locations corresponding to catheter tips are selected. In the final step, we detect and track the catheters by the means offered by sparse coding. Catheter hypotheses are formed and associated to a cost, the ones yielding the minimal global cost constitute the output of our algorithm.

L1 Sparse Coding: In order to make our paper self contained, we introduce the main concepts of sparse coding [15]. Let us suppose a signal $\mathbf{y} \in \mathbb{R}^n$ and a dictionary $\mathbf{D} \in \mathbb{R}^{n \times m}$ whose columns, also called words, approximately span $\mathbf{y}$. The signal $\mathbf{y}$ is reconstructed as a linear combination of the words through the weights $\alpha \in \mathbb{R}^m$ by solving the optimization problem

$$\min_{\alpha} \frac{1}{2} \|\mathbf{D}\alpha - \mathbf{y}\|_2^2 + \lambda \|\alpha\|_1 . \tag{1}$$

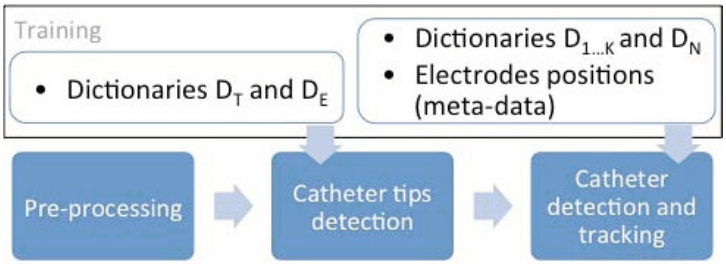

Fig. 1. Proposed pipeline

The weight λ controls the sparsity of the solution establishing a tradeoff between the least squares optimality and the number of words employed for its computation. When the weights α are constrained to be positive, the signal $\mathbf{y}$ can be reconstructed only as a conical combination of words. This is particularly convenient in the tracking by detection scenario, where the words model appearances and therefore should never be subtracted to each other. In our approach, the dictionaries contain the appearance of the catheters and of the electrodes. We rely on the sparsity assumption to match the candidates' appearances with a few, specific ones stored in the dictionaries.

2.1 Pre-processing

In order to cope with the presence of noise and improve the contrast of the fluoroscopic images, we apply to the images an homomorphic filter followed by a bilateral filter, reducing noise artifacts while preserving edges. As a further pre-processing step, we use a determinant of hessian blob detector to obtain the accurate location of electrode-like structures appearing in the images. As demonstrated by [13,14], the electrodes can in this way be localized with sub-millimeter precision, therefore enabling us to effectively limit the search space.

2.2 Training

In our method, we employ two sets of dictionaries to: (i) select image locations corresponding to the tips of the catheters, (ii) reconstruct and associate a cost to each candidate catheter. The dictionaries are obtained in a training stage that makes use of annotated data.

Training Dictionaries for "tips" Detection: In order to detect the catheter tips, we instantiate the dictionaries $\mathbf{D}_T$ and $\mathbf{D}_E$, respectively built from patches depicting catheter tips and electrodes at various orientations. The patches are normalized to have zero mean and unit standard deviation so that illumination invariance and uniform probability of being selected during reconstruction are ensured.

Training Dictionaries for Catheters Detection: In our approach, detection and tracking are coupled tasks. Supposing we want to track K catheters, we train:

1. K dictionaries $\mathbf{D}_{1...K}$ of *positive templates* capturing the appearances of each catheter separately.
2. one dictionary $\mathbf{D}_N$ of *negative templates* capturing typical background appearances.

The words $\mathbf{d}_{jk}$ of each dictionary $\mathbf{D}_k$ are associated with the specific poses assumed by the $k-th$ catheter during training. Furthermore, they are linked to *meta-data* matrices $\mathbf{M}_j$, whose purpose is to store the expected locations of the catheter's electrodes at specific poses. The coordinates stored in $\mathbf{M}_j$ are normalized to a common orientation and expressed with respect to the catheter's tip position. The negative profiles stored in $\mathbf{D}_N$ are used during tracking to penalize candidate catheters whose appearances resemble the background. All the appearances stored in the dictionaries consist of 1D intensity profiles of fixed length r, sampled from training images. The intensity profiles, which are implicitly rotation invariant, are normalized to have zero mean and unit standard deviation.

2.3 Tracking by Detection

We want to detect and track K catheters through a fluoroscopic sequence. The output of the pre-processing step of our algorithm is a set of key-points $\mathbf{X} = \{\mathbf{x}_1...\mathbf{x}_p\}$ (*Figure* 2a). Once small image patches $\mathbf{y}_i$ are extracted around the $\mathbf{x}_i$ (*Figure* 2b), the ones that correspond to catheter tips can be discriminated by solving the following two problems:

$$\hat{\alpha}_t = \min_{a_t} \|\mathbf{D}_T\alpha_t - \mathbf{y}_i\|_2^2 + \lambda_1 \|\alpha_t\|_1 , \; s.t. \; \alpha_t \geq 0 \tag{2}$$

$$\hat{\alpha}_e = \min_{a_e} \|\mathbf{D}_E\alpha_e - \mathbf{y}_i\|_2^2 + \lambda_2 \|\alpha_e\|_1 , \; s.t. \; \alpha_e \geq 0. \tag{3}$$

Key-points associated to patches that have been reconstructed better with $\mathbf{D}_T$ than with $\mathbf{D}_E$, are regarded as catheter "tips" according to

$$\mathbf{T} = \{\mathbf{t}_1...\mathbf{t}_{N\geq K}\} = \left\{\mathbf{x}_i : \|\mathbf{D}_T\hat{\alpha}_t - \mathbf{y}_i\|_2^2 < \|\mathbf{D}_E\hat{\alpha}_e - \mathbf{y}_i\|_2^2\right\}. \tag{4}$$

In the final step of our pipeline, we aim to formulate and score catheter hypotheses (*Figure* 2c). Each catheter tip $\mathbf{t}_n$ yields as many catheter hypotheses as the number of neighboring key-point $\mathbf{x}_i \in \mathbf{X}$ falling within a distance r. The catheter hypotheses are intensity profiles $\mathbf{l}_{ni}$ extracted from lines of length r originated in $\mathbf{t}_n$ and intersected with each $\mathbf{x}_i$ in turn. For each $k = 1...K$ we aim to solve the following problems:

$$\hat{\alpha}_{ni}^k = \min_{\alpha_{ni}^k} \left\|\mathbf{D}_k\alpha_{ni}^k - \mathbf{l}_{ni}\right\|_2^2 + \lambda_3 \left\|\alpha_{ni}^k\right\|_1 , \; s.t. \; \alpha_{ni}^k \geq 0 \tag{5}$$

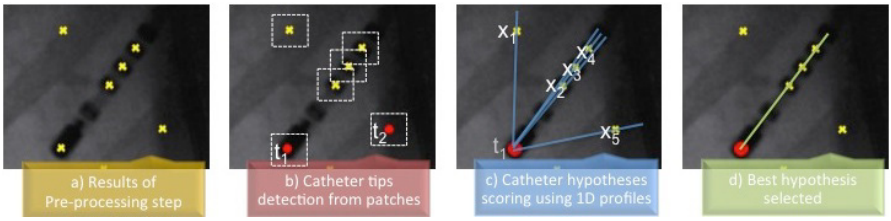

Fig. 2. Main steps of our algorithm. The output of each step is fed into the next.

$$\hat{\beta}^k_{ni} = \min_{\beta^N_{ni}} \left\| [\mathbf{D}_N, \mathbf{D}_{j\neq k}]\beta^k_{ni} - \mathbf{l}_{ni} \right\|_2^2 + \lambda_4 \left\| \beta^k_{ni} \right\|_1, \; s.t. \; \beta^k_{ni} \geq 0. \tag{6}$$

We aim to assess, through 5, the similarity of each catheter hypothesis with the *k-th* catheter and, through 6, its similarity with the background or with catheters having label different than k.

Furthermore, we identify the biggest element of α_j of $\hat{\alpha}^k_{ni}$, and we retrieve the associated meta-data $\mathbf{M}_j = [\mathbf{m}_1 ... \mathbf{m}_Q]$, containing the expected, approximated and pose specific (in terms of out-of-plane rotation of the catheter) coordinates of the electrodes. When a catheter hypothesis corresponds to a true catheter, $\mathbf{m}_j$ and $\mathbf{x}_i$ are spatially close. The minimal distances $d_i = \min_q (\|\mathbf{x}_i - \mathbf{m}_q\|)$ between each point $\mathbf{x}_i$ (after normalization to the orientation of $\mathbf{l}_i$) and the points stored in $\mathbf{M}_j$, are obtained.

The errors $E_P = \left\| \mathbf{D}_k \hat{\alpha}^k_{ni} - \mathbf{l}_{ni} \right\|_2^2$ and $E_N = \left\| [\mathbf{D}_N, \mathbf{D}_{j\neq k}]\hat{\beta}^k_{ni} - \mathbf{l}_{ni} \right\|_2^2$, and the coefficient $d = \sum_i d_i$ determine the cost of a candidate catheter according to

$$E^k_{ni} = \begin{cases} d\, E_P & if \; E_P \geq E_N \\ d\, \frac{E_P}{E_N - E_P} & if \; E_P < E_N \end{cases}. \tag{7}$$

For each tip $\mathbf{t}_i$, the best catheter hypothesis that could be reconstructed using $\mathbf{D}_k$ is retained (*Figure* 2d) and its cost $\hat{E}^k_{ni}$ is stored in a matrix $\mathbf{C} \in \mathbb{R}^{K \times N}$ modeling associations between labels and catheter hypotheses. The hungarian method is employed to select K catheter hypotheses yielding the lowest total cost. Please note that the presence of the mata-data is not only beneficial to score the catheter hypotesis but can be used to effectively recover missed electrodes detections in a meaningful way.

Mild temporal consistency can be enforced to favor catheter hypotheses occurring at similar position over time. This is realized by counting how many consecutive times a catheter k appears in a neighborhood (radius g) of its previous position and dividing the error E^k_{ni} by this number. If the *k-th* catheter moves abruptly, the counter associated with its previous position is decreased until it reaches zero.

3 Results

A total of 2835 C-arm images, belonging to 20 sequences acquired from two views were analyzed. A reference, a pacing and an 8-French ablation/mapping catheter are visible in the sequences. The image sizes are 512×512 with a pixel spacing of 0.44 mm. The X-Ray beam energy was varied between 70-92kV to ensure variability within the data. Ground truth annotation, which included the position of the 39690 electrodes appearing in the sequences, was provided by two observers. The model's parameters were fixed experimentally to be $\lambda_1 = 10$, $\lambda_2 = 150$, $\lambda_3 = \lambda_4 = 1$ for all the experiments. The scale of the blob detector was fixed to $\sigma = 4$. We enforced temporal consistency fixing the quantity g to $8px$ during all the experiments. Since our method requires a training phase, we assessed the performances of our approach when different amount of training data is used. The training images are selected from a sequence that is never used for testing.

Catheter Detection and Tracking: We assessed the performances of our method to detect and track the mapping, pacing and reference catheter respectively. The results are shown in *Table* 1. We evaluated, in particular, the impact of the number of annotated examples used during training on the performances. The pacing and reference catheters that experience little foreshortening and deformations are already well detected using a few training examples while the mapping catheter requires an higher number of training examples due to its frequent out-of-plane rotations. Incrementing the number of training examples the performances improve up to values close to 100%. The computation time increases with the dimension of the dictionaries. When 100 images are used during training, the processing time for one frame is 0.7 seconds using our MATLAB prototype and circa 0.08 seconds using our more optimized C++ implementation.

Table 1. Tracking and detection results. A different number of training examples was used in each test.

	A/P View (%)			Lateral View (%)		
Training set	Mapping	Pacing	Reference	Mapping	Pacing	Reference
3 examples	77.49	98.38	98.17	53.74	96.74	97.52
10 examples	87.13	99.79	98.38	78.86	98.02	99.08
20 examples	88.05	99.79	99.30	89.16	97.95	98.87
50 examples	93.46	99.79	99.36	89.31	**98.09**	99.01
100 examples	**93.95**	**99.79**	**99.51**	**90.02**	97.95	**99.36**

Detection Accuracy: The accuracy of the catheters detections in terms of distance of the electrodes from the ground truth annotation was assessed. The achieved results are shown in *Table* 2.

Table 2. Detection accuracy in pixels and millimeters

	A/P View		Lateral View	
	Pixels	Millimeters	Pixels	Millimeters
mapping	1.17 ± 0.64	0.51 ± 0.28	1.28 ± 0.35	0.56 ± 0.15
pacing	1.48 ± 0.60	0.65 ± 0.26	1.29 ± 0.23	0.56 ± 0.10
reference	1.63 ± 0.75	0.71 ± 0.33	1.49 ± 0.22	0.65 ± 0.09

4 Conclusions and Future Work

We have presented a novel method to detect and track linear EP catheters, that may appear foreshortened or occluded, in fluoroscopic images. The approach, that is based on $\ell 1$-sparse coding is robust to catheter overlap and has great potential in correcting for patient motion when used in conjunction with anatomical overlays. Future work will focus on the development of unique methods to automatically reconstruct catheters from [17,16] single or multi-view C-arm fluoroscopy images . The technique would rely on no user interaction, high clinical accuracy, and real-time performance. Alternatively, the detection of catheter electrodes can be coupled with generative probabilistic models that optimizes correspondence and subsequent 3D reconstructions of the catheters.

We would like to acknowledge Stavroula Timioteraki for the effort spent organizing the training and testing datasets.

References

1. Deo, R., Albert, C.M.: Epidemiology and genetics of sudden cardiac death. Circulation 125(4), 620–637 (2012)
2. D'Silva, A., Wright, M.: Advances in imaging for atrial fibrillation ablation. Radiology Research and Practice 2011 (2011)
3. Casella, M., Pelargonio, G., Russo, A.D., Riva, S., Bartoletti, S., Santangeli, P., Scarà, A., Sanna, T., Proietti, R., Di Biase, L., et al.: Near-zero fluoroscopic exposure in supraventricular arrhythmia ablation using the ensite navx mapping system: personal experience and review of the literature. Journal of interventional cardiac electrophysiology 31(2), 109–118 (2011)
4. Fallavollita, P., Savard, P., Sierra, G.: Fluoroscopic navigation to guide rf catheter ablation of cardiac arrhythmias. In: 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEMBS 2004, vol. 1, pp. 1929–1932. IEEE (2004)
5. Franken, E., Rongen, P., van Almsick, M., ter Haar Romeny, B.M.: Detection of electrophysiology catheters in noisy fluoroscopy images. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 25–32. Springer, Heidelberg (2006)
6. Ma, Y., King, A.P., Gogin, N., Rinaldi, C.A., Gill, J., Razavi, R., Rhode, K.S.: Real-time respiratory motion correction for cardiac electrophysiology procedures using image-based coronary sinus catheter tracking. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 391–399. Springer, Heidelberg (2010)

7. Ma, Y., Gao, G., Gijsbers, G., Rinaldi, C.A., Gill, J., Razavi, R., Rhode, K.S.: Image-based automatic ablation point tagging system with motion correction for cardiac ablation procedures. In: Taylor, R.H., Yang, G.-Z. (eds.) IPCAI 2011. LNCS, vol. 6689, pp. 145–155. Springer, Heidelberg (2011)
8. Schenderlein, M., Stierlin, S., Manzke, R., Rasche, V., Dietmayer, K.: Catheter tracking in asynchronous biplane fluoroscopy images by 3d b-snakes. In: SPIE Medical Imaging, International Society for Optics and Photonics, p. 76251U (2010)
9. Brost, A., Liao, R., Strobel, N., Hornegger, J.: Respiratory motion compensation by model-based catheter tracking during ep procedures. Medical Image Analysis 14(5), 695–706 (2010)
10. Wu, W., Chen, T., Barbu, A., Wang, P., Strobel, N., Zhou, S.K., Comaniciu, D.: Learning-based hypothesis fusion for robust catheter tracking in 2d x-ray fluoroscopy. In: CVPR, pp. 1097–1104. IEEE (2011)
11. Wu, W., Chen, T., Strobel, N., Comaniciu, D.: Fast tracking of catheters in 2d fluoroscopic images using an integrated cpu-gpu framework. In: ISBI, pp. 1184–1187. IEEE (2012)
12. Yatziv, L., Chartouni, M., Datta, S., Sapiro, G.: Toward multiple catheters detection in fluoroscopic image guided interventions. IEEE Transactions on Information Technology in Biomedicine 16(4), 770–781 (2012)
13. Milletari, F., Navab, N., Fallavollita, P.: Automatic detection of multiple and overlapping EP catheters in fluoroscopic sequences. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part III. LNCS, vol. 8151, pp. 371–379. Springer, Heidelberg (2013)
14. Ma, Y., Gogin, N., Cathier, P., Housden, R.J., Gijsbers, G., Cooklin, M., O'Neill, M., Gill, J., Rinaldi, C.A., Razavi, R., et al.: Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions. Medical Physics 40(7), 071902 (2013)
15. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al.: Least angle regression. The Annals of Statistics 32(2), 407–499 (2004)
16. Fallavollita, P.: Is single-view fluoroscopy sufficient in guiding cardiac ablation procedures? Journal of Biomedical Imaging, Jg., S.1 (2010)
17. Fallavollita, P.: Acquiring multiview c-arm images to assist cardiac ablation procedures. Journal on Image and Video Processing, Jg., S.3 (2010)