

A FAST CATHETER SEGMENTATION AND TRACKING FROM ECHOCARDIOGRAPHIC SEQUENCES BASED ON CORRESPONDING X-RAY FLUOROSCOPIC IMAGE SEGMENTATION AND HIERARCHICAL GRAPH MODELLING

Xianliang Wu*, James Housden†, YingLiang Ma†, Kawal Rhode†, Daniel Rueckert*

*Department of Computing, Imperial College London, SW7 2AZ, UK

†Division of Imaging Sciences and Biomedical Engineering, King's College London, SE1 7EH, UK

ABSTRACT

3D soft tissue information, which X-ray images cannot provide but 3D echocardiographic imaging can, may be required in cardiac catheter-based interventions. In this paper, we propose a real-time catheter tracking strategy in echocardiographic sequences based on catheter segmentation in 2D X-ray images and registration between these two modalities. Firstly the segmentation from X-ray and the registration between X-ray and ultrasound is computed. The results from these steps reduce the search space in the ultrasound volume to a limited space surrounding a curved surface. This space is straightened and 2D SURF is calculated on the sampled cross-sections. All features are organized as a two level hierarchical graph. The longest path on the top-level graph and shortest paths on the bottom level graphs are solved. This combined path is considered as the potential catheter after B-Spline modelling and growing. The experiments on clinical data (2000 pairs of frames) show a better performance than a previous method and some dominant vesselness filters.

Index Terms— Catheter Tracking, Catheter Ablation, X-ray Fluoroscopy, Echocardiography

1. INTRODUCTION

X-ray fluoroscopy imaging and echocardiography imaging (ultrasound, US) are two modalities which have been widely used in real-time cardiac interventions. X-ray fluoroscopic imaging, which is the standard modality used to monitor cardiac catheter ablation interventions, has some limitations such as lack of 3D soft tissue information and harmful ionising radiation. In contrast, echocardiographic imaging, which overcomes the above limitations, can be seen as a good alternative for catheterization procedures. However, severe acoustics artifacts and small field of view make monitoring by ultrasound alone challenging and sometimes impossible.

For the application of catheter tracking in X-ray images, a great number of studies have been proposed. The methodologies include filter-based segmentation [1], electrode feature detection [2], graph-cut [3] and learning-based methods [4]. The tracking subjects include the catheter tip [1][4], the

electrodes [2] and the soft parts of the catheter or a guide wire [3][5]. For ultrasound images, very limited studies have been proposed for catheter extraction. In [6], a fast catheter segmentation was proposed based on the segmentation from X-ray sequences. A graph model was constructed and the shortest path was solved in this work. However, the approach was tested on very little clinical data.

We are aiming to provide a fast catheter visualization based on a combination of both X-ray and ultrasound imaging for cardiac clinical interventions. This combination includes soft tissue information in the procedure and overcomes the disadvantages of ultrasound imaging. The main contributions are: (1) a new 2D SURF measurement filter in a straightened space for tubular structures; (2) a post-processing growing strategy which can increase the ratio of extracted parts of the catheter; (3) a hierarchical graph modelling strategy consisting of two layers which can preserve the integrity of catheter fragments and link them together as a more likely catheter; (4) more extensive experiments in clinical data.

2. METHODOLOGY

2.1. Projecting the Extracted Catheter from the X-ray Plane to the US Volume

Our system starts with the catheter extraction or tracking from X-ray sequences. The method we use is from [6] which combines a blob detector, local line fitting and Kalman filter based growing. A fast Primal-Dual (Fast-PD) based tracking framework and a post-processing Kalman growing step were also proposed. The reader is referred to [6] for more details. The other information we need is the transformation matrix \mathbf{T} mapping each US volume from the US coordinate system to the X-ray coordinate system. We use the method in [7].

2.2. Search Space Straightening and SURF Feature Extraction

Suppose $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \mathbf{t}_4]^T$, for the k th landmark location $\mathbf{p}_k = (u_k, v_k)$ on the X-ray image, the corresponding ray

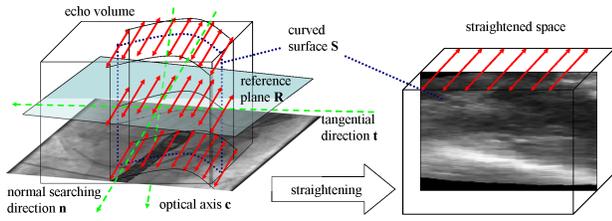


Fig. 1. Space Straightening.

equation in the 3D echo volume is then denoted by:

$$[\mathbf{t}_1 - u_k \mathbf{t}_4, \mathbf{t}_2 - v_k \mathbf{t}_4]^T \mathbf{X}^k = 0 \quad (1)$$

where $\mathbf{X}^k = [x_1^k, x_2^k, x_3^k, 1]^T$ is the coordinates of the voxels in the US volume which correspond to the k th landmark in the X-ray image. Each ray is bounded by two intersections with the border of the echo volume. All of the rays corresponding to locations on the catheter in the X-ray image comprise a curved surface \mathbf{S} . As shown in Fig. 1, a reference plane \mathbf{R} which crosses the centre of the echo volume with a normal vector the same as the direction of the X-ray projection axis, \mathbf{c} , is defined. The projections of the first and last X-ray landmarks of the catheter on reference plane \mathbf{R} are used to define unified tangential direction \mathbf{t} . The unified normal search direction \mathbf{n} is defined as perpendicular to both \mathbf{c} and \mathbf{t} . We then straighten the original echo space surrounding the curved surface into a regular rectangular volume \mathbf{V} . The direction vector \mathbf{v}^k for each ray corresponding to a pixel on the catheter in X-ray, as well as the corresponding intersection \mathbf{X}_0^k with the reference plane \mathbf{R} , are calculated from (1). The ray equation can be reformed as $\mathbf{X}^k = \mathbf{X}_0^k + s\mathbf{v}^k$ and s is bounded by $[s_{min}^k, s_{max}^k]$ based on the intersections with the border of the original echo volume. Given $s_{max} = \max_k s_{max}^k$ and $s_{min} = \min_k s_{min}^k$, the value of each voxel in the straightened volume is defined as:

$$\mathbf{V}_{i,j,k} = I(\mathbf{X}_0^k + (j + s_{min} - 1)\mathbf{v}^k + i * \mathbf{n}) \quad (2)$$

where $I(\mathbf{X})$ is the intensity value of coordinate \mathbf{X} in the original echo volume and set to zero if \mathbf{X} is out of the volume. i, j and k are coordinate indexes for the straightened volume. k also corresponds to the landmark index introduced above. i is bounded by $[-\tau_{normal}, \tau_{normal}]$ where τ_{normal} is the limit of the search range along normal direction \mathbf{n} . By this transform, we straighten the original space surrounding the curved surface to a rectangular volume \mathbf{V} .

For the catheter which has been transformed into the straightened space \mathbf{V} , it is likely that the catheter is along the z -axis of \mathbf{V} , with small changes in the x -axis position, but some variation in the y -axis position. On this assumption, the xy and xz cross-section of \mathbf{V} will show a blob like structure at the catheter position. We calculate the response $Resp(\mathbf{X})$

by selecting the higher SURF value from its two corresponding locations in the xy and the xz cross-sections. By setting up a threshold τ_{surf} , a set of candidate features, with both the coordinates \mathbf{X}_i^k and the responses R_i^k above the threshold for the i th feature on the k th ray, are selected.

2.3. Two Level Hierarchical Graph Construction and Path Finding

Firstly from the 1st to the n th (last) ray, every pair of two feature locations on the same or adjacent ray(s) is checked for their distance. If the distance is below a threshold $\tau_{subgraph}$, then the two feature locations are linked by an edge with a cost equal to the distance. After this closest distance clustering, feature locations are organized as a set of clusters $C_j, j = 1, \dots, m$ (m is the number of clusters) and within each cluster features are linked to each other directly or via other features. For each cluster C_j , the features \mathbf{X}_i^k with no direct linked features with a smaller ray index than k are selected as the candidate starting nodes. Other features that have no direct linked features with a larger index than its ray index are selected as the candidate ending nodes. One starting node is selected among the starting node candidates and one ending node is selected among ending candidates. The principle is to make the distance between the starting and ending nodes as large as possible. This makes each catheter section as complete as possible. Then for every pair of two clusters, if one cluster i 's start or end point A_i is very close to the other cluster j 's start or end point A_j , and their directions are almost co-linear, then these two clusters are combined as one cluster. Suppose the opposite ends of each cluster where they are not linked are B_i and B_j for clusters i and j respectively, then the constraint can be denoted by:

$$e^{|\langle \overrightarrow{B_i A_i}, \overrightarrow{A_i A_j} \rangle|} + e^{|\langle \overrightarrow{A_i A_j}, \overrightarrow{A_j B_j} \rangle|} < \tau_{colinear}, |\overrightarrow{A_i A_j}| < \tau_{cluster} \quad (3)$$

where $\langle \cdot, \cdot \rangle$ denotes the angle between two vectors. $\tau_{colinear}$ and $\tau_{cluster}$ are thresholds set manually. The start and end points are re-allocated by their ray index (the start should have a smaller ray index than the end). To construct the top level graph, the start S_i and end E_i of each cluster i are considered as the nodes. Because a directed acyclic graph is used for the longest path problem, all edges have their own directions. For each cluster i , an edge $edge_{i,i}$ is created from the start to the end points. For any pair of the clusters i, j which are not combined but are likely co-linear, an edge $edge_{i,j}$ is created from the end point of one cluster to the start point of the other cluster. A weight $W_{i,j}$ is allocated to each edge through:

$$W_{i,j} = \begin{cases} dis(i, i) \frac{\overline{R}^i}{\tau_{surf}}, & \text{if } i = j \\ -dis(i, j) \frac{e^{|\theta_j|} + e^{|\theta_j|} + e^{|\omega_{i,j}|}}{3} & \text{if } i \neq j \end{cases} \quad (4)$$

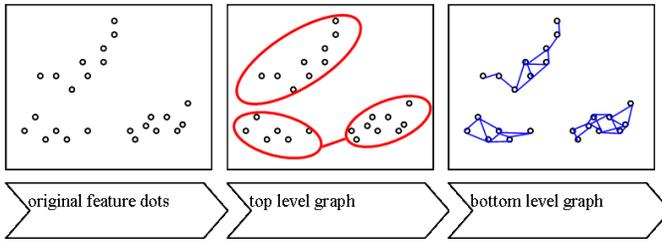


Fig. 2. Hierarchical graph construction.

where $dis(i, i)$ is the length of the vector $\overrightarrow{S_i E_i}$ within the cluster i , and $dis(i, j), i \neq j$ is the distance between the end point of cluster i and the start point of cluster j , along the vector $\overrightarrow{E_i S_j}$. \overline{R}^i is the average value of the SURF responses among the members of the cluster. θ_i is the turning angle from the vector $\overrightarrow{S_i E_i}$ to $\overrightarrow{E_i S_j}$ and θ_j is the turning angle from the vector $\overrightarrow{E_i S_j}$ to $\overrightarrow{S_j E_j}$. $\omega_{i,j}$ is similar to angular acceleration which is defined as $\omega_{i,j} = \theta_j - \theta_i$.

After the top-level graph is complete, a bottom-level graph for each cluster is constructed by allocating each directly linked edge a cost of the distance between its two nodes, as shown in Fig. 2. Then the longest path is selected in the top-level graph using a dynamic programming technique and the shortest path is solved for each cluster. Each intra-cluster edge of the longest path is then replaced by the corresponding shortest path.

2.4. Post Processing

Finally, all landmarks on the longest path are transformed back to the original echo volume. Then the sorted landmarks on the path are used to model a B-Spline curve. The tangential direction \mathbf{d}_i for the i th landmark can also be calculated during B-Spline modelling. Because the overlap of the X-ray and echo images is limited due to alignment errors and this may cause the extracted result to cover only a part of the whole catheter, we use a growing technique to extend the catheter to the un-extracted parts. Starting from the two ends of the already tracked part of the catheter, the growing direction \mathbf{v}_i for each end is along $-\mathbf{d}_0$ for the first landmark and \mathbf{d}_n for the last. Given two adjacent landmarks \mathbf{p}_{i-1} and \mathbf{p}_i , the size of the growth step is determined by:

$$\Delta s_i = (M - N)\mathbf{v}_{i-1}^T \mathbf{v}_i + N \quad (5)$$

where M and N are the maximum and minimum size of growth step set manually ($M = 5, N = 3$ voxels in our application). By setting the size of the steps through this equation, the growth will take smaller steps when a sharp turn occurs. After both the direction and the size of the step are determined, the next candidate position can be acquired by:

$$\mathbf{p}'_{i+1} = \mathbf{p}_i + \Delta s_i(\mathbf{v}_i) \quad (6)$$

The final position of the next landmark is determined by finding the voxel within a defined neighborhood of \mathbf{p}'_{i+1} with the largest Frangi response. If all of the candidates' responses are below the threshold τ_{surf} , the growth will be terminated.

3. EXPERIMENTS

Four long sequences of patient data were used in our experimental evaluation. These comprise more than 2000 frames of X-ray and nearly 1000 frames of corresponding echo data. The number of pairs of related X-ray and echo image is more than 2000 in total. All data contained a guide wire and a catheter. The X-ray data have an image size of 512×512 pixels. The echo volumes have $144 \times 160 \times 208$ voxels of $0.60 \times 0.60 \times 0.29 \text{ mm}^3$ for the first sequence and $144 \times 176 \times 208$ voxels of $0.51 \times 0.50 \times 0.29 \text{ mm}^3$ for the rest of the echo sequences. The synchronization of the X-ray and echo was performed manually. The ground truth catheter locations in the X-ray and echo images were also marked manually. Experiments were performed on an Ubuntu Linux system on a 3.40GHz, 8GB desktop. The methods compared were the proposed strategy with and without growing, Frangi filtering [8] and an ITK filtering method [9], as well as the method proposed in [6].

The following performance metrics were defined to evaluate each algorithm's speed, accuracy and robustness: (1) **Average frame rate**. This evaluates speed, which directly determines performance for real-time applications; (2) **Average tracking error**. For each landmark i , a shortest distance d_i to the ground truth is calculated. Then a threshold ρ is used to select correctly tracked landmarks with $d_i \leq \rho$. The average of d_i among correctly tracked landmarks is defined as the tracking error; (3) **Incorrect tracking percentage (ITP)**. ITP is defined as the number of incorrectly tracked landmarks over the total number of landmarks. ITP indicates the reliability of the tracking results. (4) **Failed tracking percentage (FTP)**. For each landmark on the ground truth, the minimum distance to the tracked curve is calculated. The landmark is considered as successfully tracked if the distance is below ρ . FTP is defined as the number of failed landmarks over the total number of landmarks on the ground truth. It evaluates to what extent the whole catheter can be tracked. (5) **Ratio of failed tracked frames (ROF)**. If the average SURF response for all landmarks is below $\tau_{surf}/3$, we treat it as a failure. The number of frames where a tracking failure is detected over the total number of frames can indicate the need for re-initialisation.

The threshold $\tau_{surf} = 0.01$ is used to select feature locations. The search range τ_{normal} along the normal direction for each voxel on the curved surface depends on how accurate the transformation between X-ray and echo images is. Normally it is set to 10 *pixels* for small registration errors which means the catheter is almost on or very near the curved surface. For some bad registrations or bad synchronizations,

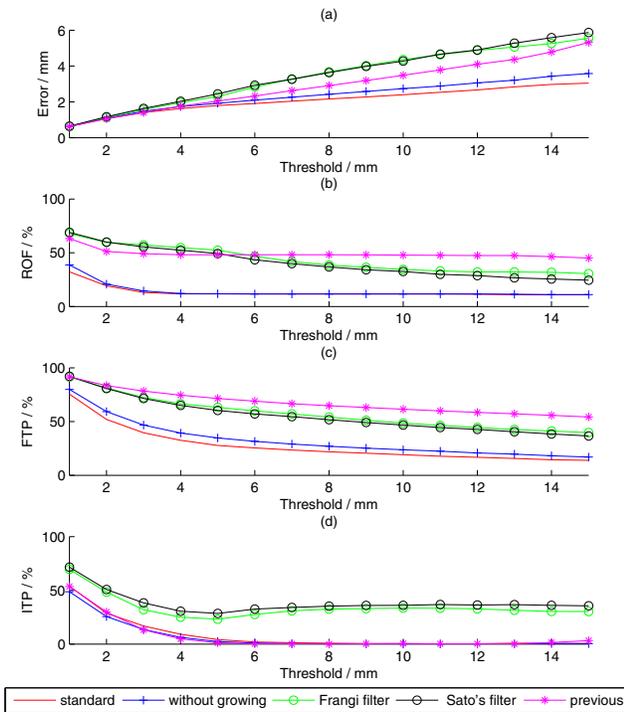


Fig. 3. Comparison Experimental Results. All subfigures have the threshold ρ as the x-axis label.

the value can be set up to 30, with the tradeoff of a larger search range and therefore reduced speed. At the stage of graph construction, the threshold for the minimum distances for subgraph construction or clustering is set to $\tau_{subgraph} = 10$ pixels. The threshold to decide the combination of any two clusters is set to $\tau_{colinear} = 1$ and $\tau_{cluster} = 20$.

Fig. 3 shows the results of error, ROF, FTP and ITP. The average time per frame is 1.3498, 0.7428, 4.2993, 4.3489, 0.5017 seconds for the proposed method with and without growing, Frangi filtering, ITK vesselness filtering, and the method in [6]. Fig. 4 shows a comparison results in which the proposed method spotted the right position whereas [6] failed. We can conclude: (1) The growing strategy can extend the tracking to a longer length, the tradeoff is a larger computation cost and a higher ITP; (2) 2D SURF is more reliable and fast than vesselness filtering techniques in any metric. (3) The proposed strategy is more suitable for clinical application than [6]. Given $\rho = 5mm$, our standard method achieves a speed of 1.35 spf and an error of 1.79 mm, with FTP of 27.7%, ITP of 4.19% and a failure ratio of 11.6%.

4. CONCLUSION

We have proposed a fast X-ray assisted catheter segmentation and tracking strategy for ultrasound imaging in cardiac

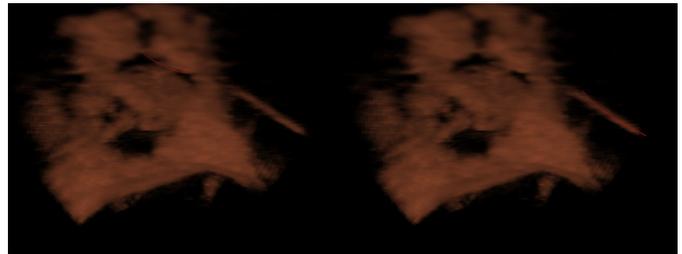


Fig. 4. An example of the tracking results of [6] (left) and proposed method (right) (marked in red).

catheterization. Our method is better than [6] in any performance metric and the proposed SURF measurements outperform some dominant vesselness filtering techniques.

5. REFERENCES

- [1] A. Brost and R. et al Liao, “3-D respiratory motion compensation during EP procedures by image-based 3-D lasso catheter model generation and tracking,” in *MICCAI*, vol. 5761, pp. 394–401. 2009.
- [2] Y. Ma and A. et al King, “Real-time respiratory motion correction for cardiac electrophysiology procedures using image-based coronary sinus catheter tracking,” in *MICCAI*, vol. 6361, pp. 391–399. 2010.
- [3] O. Pauly and H. et al Heibel, “A machine learning approach for deformable guide-wire tracking in fluoroscopic sequences,” in *MICCAI*, vol. 6363, pp. 343–350. 2010.
- [4] W. Wu and T. et al Chen, “Learning-based hypothesis fusion for robust catheter tracking in 2D X-ray fluoroscopy,” in *CVPR*, june 2011, pp. 1097–1104.
- [5] P. Wang and T. et al Chen, “Robust guidewire tracking in fluoroscopy,” in *CVPR*, june 2009, pp. 691–698.
- [6] X. Wu and J. et al Housden, “Catheter tracking in 3D echocardiographic sequences based on tracking in 2D X-ray sequences for cardiac catheterization interventions,” in *ISBI*. 2013.
- [7] G. Gao and G. et al Penney, “Rapid image registration of three-dimensional transesophageal echocardiography and X-ray fluoroscopy for the guidance of cardiac interventions,” in *IPCAI*, vol. 6135, pp. 124–134. 2010.
- [8] A. Frangi and W. et al Niessen, “Multiscale vessel enhancement filtering,” in *MICCAI*, 1998, pp. 130–137.
- [9] Y. Sato and S. et al Nakajima, “3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images,” in *CVRMed-MRCAS’97*, 1997, pp. 213–222.