

Automatic Segmentation and Tracking of Interventional Tools in Coronary Angiographies

Petkov S^{1*}, Radeva P¹, Carrillo X² and Gatta C²

¹Centre de Visio per Computador, Universitat de Barcelona, Barcelona, Spain

²Germans Trias I Pujol University Hospital in Badalona (Barcelona), Spain

Abstract

In this paper we present a fully automatic method to detect, segment and track interventional tools in sequences of coronary angiography images. Our work is motivated by several possible applications that would enhance the job of cardiologists and improve the outcome of coronary interventions. Examples are lowering the usage of contrast agent, improving the visualization in 2D and 3D and facilitating tool guidance. Novelty in our proposal are the usage of the Hungarian algorithm for simultaneous tracking of deformable structures and the analysis of movement frequencies to detect catheter and guide wire. After an initial point's selection, based on vessel and centerline masks, the Hungarian method uses a cost function to link points. The cost function is a linear combination of one appearance-based measure and two geometry-based measures. Then the trajectories of linked points are inspected in the Fourier domain to select a trajectory that tracks an interventional tool. The tool segmentation in each image of the sequence uses the selected trajectory plus the vessel and centerline masks for the image. Differentiation between catheter and guide wire makes use of the fact that the catheter is the thicker tool. In addition, we systematize the main challenges in this area of research in a list of issues that need to be addressed for a robust segmentation and tracking of interventional tools. Comprehensive evaluation shows that our method handles these challenges and outperforms the related state of the art.

Keywords: Angiography; Catheter; Detection; Segmentation; Tracking; X-ray

***Corresponding author:** Petkov S Centre de Visió per Computador, Universitat de Barcelona, Bellaterra, Barcelona, 08193, Spain, Tel: +34 693980478; E-mail: simeon.petkov.bg@gmail.com

Received Date: December 12, 2016

Accepted Date: July 20, 2017

Published Date: August 04, 2017

Citation: Petkov S, Radeva P, Carrillo X, Gatta C (2017) Automatic Segmentation and Tracking of Interventional Tools in Coronary Angiographies. J Case Repo Imag 1: 004.

for interventional purposes such as stenting of stenosed vessels or for diagnostic purposes such as assessment of myocardial perfusion or stenosis grading. To perform X-ray angiography during cardiac interventions, physicians insert a catheter in one of the coronary arteries and inject contrast agent into the vascular tree. Figure 1 shows a single image from a coronary angiography. Part of the vascular tree is visible in the left-top, as physicians inject contrast agent through the inserted catheter (starting from the right-top). Depending on the severeness of the case, the physician might insert a guide wire in the catheter and manually guide it to the site of the stenosis, relying on fluoroscopy images (also called frames) and his/her knowledge on the anatomy of the cardiovascular system. Normally, both the catheter and the guide wire are made of radiopaque materials and are visible under X-ray as contrast objects. However, positioning the guide wire correctly is difficult because of the complexity of the vasculature and narrowness of the blood vessels, which causes an increase in interventional time and radiation exposure [1]. Purpose of our work is to develop and evaluate a method that automatically detects, segments and tracks the catheter and the guide wire (if exists) in cardiac X-ray images. In this paper we will be using the term 'interventional tools' to refer both to a catheter and a guide wire.

There has been a significant interest in automatic detection and tracking of interventional tools in fluoroscopy guided cardiac interventions [2-6]. This interest arises from the possible impact in making such interventions more effective and efficient. In their recent paper, Volpi et al., state that "There is an urgent need for computer assistance solutions that support the smooth integration of technological solutions within the surgical workflow" [7]. Although the paper in discusses endovascular aneurysm repair in the abdominal aorta, the cited paragraph is a clear expression what are the current expectations from

Introduction

Minimally invasive cardiac interventions have been established as a better alternative for treating numerous cardiac diseases than conventional open-heart surgery. At the same time, medical image processing has advanced to a level that allows to incorporate semi or fully automatic algorithms in medical imaging systems, in order to improve data visualization and analysis. General enhancements of interventional images include detection, segmentation and tracking of objects or areas of interest and multimodal fusion. Main subfields of research in medical imaging that contribute to these enhancements are analysis of single or follow-up images, entire video sequences and intra/inter modality analysis.

Quantitative Coronary Angiography (QCA) and Percutaneous Coronary Intervention (PCI) are standard procedures to diagnose and restore blood circulation in the cardiovascular system. X-ray angiography is the most popular imaging modality to visualize blood vessels

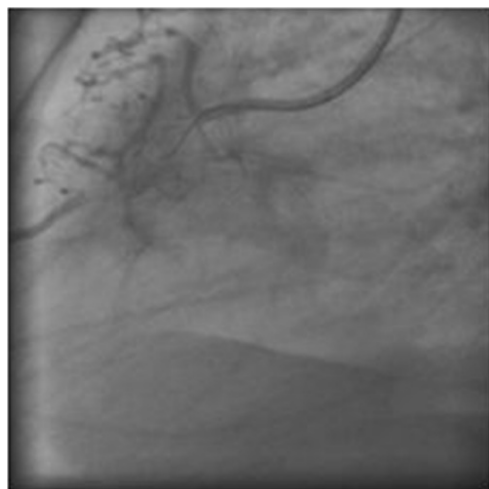


Figure 1: A single image from X-ray angiography. Part of the vascular tree is visible in the left-top, as physicians inject contrast agent through the inserted catheter (starting from the right-top). The thin contrast curve inside the catheter and the part of it that goes outside the catheter, is a guide wire.

medical point of view regarding image processing in minimally-invasive procedures [7].

Usage of contrast agent and exposure to radiation are main issues with cardiovascular interventions. Since contrast agents are toxic and may cause damage to the kidney, it is reasonable to minimize their usage. On the other hand, injecting lower doses of contrast agent hampers image contrast in fluoroscopic images and may result in longer interventions. In general, avoiding prolonged exposure to X-rays and high doses of contrast agent would benefit for the safety of patients and medical staff.

In X-ray angiography images, when contrast agent enters the cardiovascular system, high-contrast blood vessels may overlap or completely hide interventional tools (see section 3). Automatic and real-time annotation of interventional tools in visualizations of endovascular procedures is relevant to improve visualization in 2D and 3D. Detection, segmentation and motion tracking would facilitate physicians to perform tool guidance, lowering the risk of complications and potentially improving success rate [2,3]. Methods for three-dimensional reconstruction of coronary trees or interventional tools from biplane X-ray angiographies have been designed to help pre-interventional planning [4,8,9]. Visualization of such reconstructions could also be improved by our method.

Despite the considerable interest on the topic, the majority of the proposed methods on object tracking in cardiac fluoroscopic sequences tend not to cover the entire complexity regarding automatic detection and tracking of interventional tools; methods either rely on manual initialization or do not consider presence of both catheter and guide wire, neither contrast agent injection (Table 1). However, nowadays contrast agent injection is an essential part during fluoroscopy guided cardiovascular interventions.

Our contributions and main novelty is the use of frequency information for the movement of vessel structures to automatically detect and track interventional tools in coronary angiographies. After tracking a set of objects in the initial frames of a sequence, we perform a Fourier analysis to select a trajectory of interest. In this way, detection and tracking of interventional tools is fully automatic and does not

rely on manual initialization. Hereto, being point-based, the tracking is apt to cover a wide range of non-rigid transformations. To the best of our knowledge this is the first work to dedicate effort on robust vessel tracking between subsequent frames of coronary X-ray angiographies. We specify the main challenges in vessel tracking between frames, like the injection of contrast agent (section 3) and provide visual and quantitative results to show that our method is robust in the presence of these challenges. Our definition for similarity between two points, which combines two different types of measures - appearance-based and geometry-based, is the main contribution for the robustness of our vessel tracking. The method we propose is designed to be executed in online mode, receiving subsequent frames from a streaming data source or in offline mode, processing the entire sequence at once. We also address the issue of differentiating between the two common types of interventional tools, used in cardiac interventions - catheter and guide wire. The data set used to test the method is composed of diagnostic and PCI angiographies and contrast agent is visible in all sequences.

State of the Art

Registration approaches align images, and thus can be used to find the correspondence between them. In X-ray sequences (and especially in coronary angiographies) objects may superimpose other objects. Moreover, structures may appear or disappear within a set of subsequent frames (e.g., contrast agent, parts of controllable interventional tools that are close to image boundaries). In the comprehensive theoretical research authors point out that registration techniques trying to find complete correspondence between two angiography images are not likely to be successful [10]. Usually they obtained transformation attempts to extrapolate structures from the moving frame to represent independent layer movement or injection of contrast agent.

A significant number of papers has been published on detection and tracking of interventional tools and blood vessels in fluoroscopic images of minimally invasive procedures.

In a segment of interventional tool is modeled as a B-spline fitted over control points selected by discrete Maximum a Posteriori (MAP) optimization over Markov Random Field (MRF) [3]. This tracking method relies on manual detection of a curve segment in the first frame. There is no specific information on how the method handles cases when contrast agent enters the arteries.

Wang et al., introduce a Bayesian framework to track guide wire in X-ray sequences of cardiac interventions [11]. Semantic model includes the catheter tip as guide wire starting point. The authors address the case of contrast agent injection but do not provide clear results on how well the method handles it. The method in relies on manual detection and requires labeled training data to perform the tracking [11]. In addition, according to, it "cannot be applied right away to the multitude of different C-arms exhibiting a variety of parameters and thus output images" [3].

In tubular structures are segmented by a novel Graph-cut energy formulation. Supervised learning uses local and contextual information to detect the catheter [5]. Although this publication does not involve tracking, it is the only one we found to contain quantitative results on differentiation between catheter and arteries in coronary angiographies.

Method	Contrast Agent Injection	Cardiac Motion	Detection	Tracking	Differentiate between Cath. and GW	Run-time	Number of Sequences
Volpi '15 [7]	YES	NO	Automatic	YES	NO	1 fps	4
Milletari '14 [12]	NO	YES	Automatic	YES	NO (no GW presence)	0.08 fps	20
Heibel '13 [3]	YES	YES	Manual	YES	NO	16.7 fps	17
Hernandez '12 [5]	YES	N/A	Automatic	NO	NO (no GW presence)	0.05 fps	36
Pauly '10 [2]	NO	NO	Manual	YES	NO (no Cath. presence)	1.5 fps	2
Wang '09 [11]	YES	YES	Manual	YES	only the Cath. tip	2 fps	47
Baert '03 [1]	NO	NO	Manual	YES	NO (No Cath. presence)	0.2 fps	10

Table 1: Overview of methods that segment and/or track interventional tools in cardiac fluoroscopic images and sequences.

Recently, Milletari et al., proposed a fully automatic method to detect and track electrophysiology catheters in fluoroscopy sequences [12]. The method relies on annotated training data and additional meta-data and employs sparse coding to select the best catheter hypothesis.

In the authors propose a method to track only J-tipped guide wires during endovascular interventions in the thorax under X-ray fluoroscopy [1]. The method is based on a two-step procedure. First, the guide wire displacement is roughly estimated by a template matching. Then, the guide wire position is determined by fitting a B-spline to a feature image with enhanced line-like structures.

The assumptions about the shape and the appearance of specific catheters and guide wires are not valid in the case of catheters used in coronary angiographies, where catheter appearance and dynamics resemble much those of an artery filled with contrast agent.

The methods were designed for angiographies of liver chemoembolizations and abdominal aortic aneurysm treatments, where cardiac motion is less prominent and the major motion of the guide wire has lower frequency than in coronary angiographies [2,7]. Although these methods have not been applied to coronary angiographies, we consider them relevant to the topic, because of the potential they have to be extended and to cover complex scenarios.

The supervised learning approach in employs a motion distribution model into a tracking framework based on second-order MAP-MRF optimization [2]. It relies on fully labeled sequences to learn the guide wire motion distribution and on ground truth annotation to initialize the tracking. The method was tested on two sequences of liver chemoembolizations.

The method in based on robust principal component analysis, detects and tracks a stent graft device during endovascular aneurysm repair [7]. The idea of the authors is to decompose fluoroscopic images to background and foreground parts, taking the foreground as a prediction. The method was tested on four clinical cases of abdominal aortic aneurysm treatment.

Table 1 synthesizes the methods that are most related to our work. To the best of our knowledge we present the first fully automatic method to detect and track catheter and guide wire in cardiac fluoroscopic sequences in the presence of contrast agent injection.

Issues in Detection and Tracking of Interventional Tools

The following paragraphs summarize the main challenges in designing and implementing a methodology to detect and track interventional tools in coronary angiographies.

Due to the low signal-to-noise ratio in X-ray images, many points within a tubular structure will be very similar for most image-based descriptors. This is especially an issue in the absence of distinctive curvature, texture or another synchronously moving object near the tracked one. As a consequence, in standard feature-based point-tracking approaches, points would drift alongside tubular structures without keeping their relative positions.

As we pointed out in section 1, injection and presence of contrast agent is one of the main challenges for robust detection and tracking in coronary angiographies. After contrast agent has been injected and enters arteries, the catheter is not visible anymore in the artery, (Figure 2) and this may also be a cause of points drifting from their relative position. Eventually, when the contrast agent washes out and arteries are no longer visible, it is probable that some points return within the borders of the interventional tool. It is not trivial to determine the subset of frames in which tracking loses the object of interest.

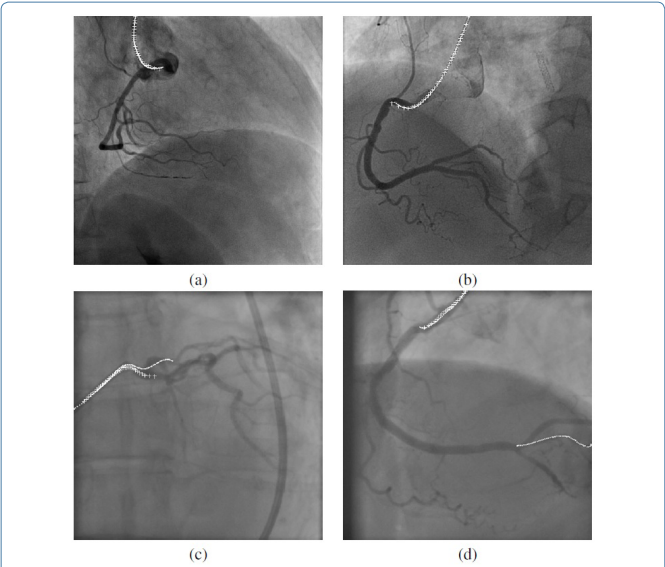


Figure 2: Visual results of our method together with one of the two annotated ground truth curves. Our prediction is marked with sparse '+' symbols and the ground truth is marked with a dotted line.

Structures with high contrast, crossing or passing near the tracked object, would mislead image-based similarity measures that incorporate contextual information. As a result, points could attach to a near structure or drift alongside the tracked object to a point that results in small measure difference, although not corresponding to the point being tracked.

Diaphragm, if present, is visible under X-ray as a darkened area, moving in accordance with patient breathing rate (Figure 2a,b&d). Considering that diaphragm could occlude vessels and breathing rate, in general, differs from heart beating rate, tracking of interventional tools in coronary angiographies may be hampered by diaphragm movement.

If contrast agent injection occurs at the same moment as the heart pumps oxygen-rich blood, a part of the contrast agent could enter the aorta. In coronary angiographies this phenomena looks like a dark area that starts from the catheter tip and is characterized by non-predictable shape, undergoing wide range of deformations before vanishing (Figure 2a). Presence of contrast agent in aorta may mislead point tracking in a similar way as other contrast structures by making points jump away from the tracked object or drift within it.

Depending on the projection, sometimes parts of the interventional tool could move outside of the image acquisition boundaries. In authors emphasize on the need to define events like appearing and disappearing of a structure to register objects, which appear in only one of the registered images [13]. Analogously to registration approaches, tracking approaches that try to handle such cases without explicit definition and detection when a structure appears or disappears, are likely to be suboptimal.

Method

The task of automatically tracking interventional tools in cardiac X-ray images can be divided in two phases - detecting the tool in the initial frame and tracking it through subsequent frames [2]. A naïve method would be to detect the tool in each frame separately and consider the combination of all detections as tracking result. Such simplification is very probable to produce tracking failures since it does not use temporal information; neither imposes any restrictions on the change in shape and position of the object of interest [14].

We propose a point tracking algorithm based on the Hungarian method, which finds the optimal matching between two disjoint sets of points, given a cost function for every pair of points in a bipartite graph. Our definition of the cost function for the Hungarian assignment is a linear combination of three measures, providing complementary information about the similarity between two points (section 4.4). One of the measures is computed over gray level profiles of regions around two points and the other two measures use the similarity between geometrical distributions of points. In our experiments, none of the proposed measures seemed robust enough to be used on its own as a cost function. Considering that each measure is unstable under different conditions, we use a linear combination of them.

In Figure 3 we present a block scheme of our method. The input is a sequence of coronary angiography images. The method is designed to be run online, pre-processing only the frames from the first cardiac cycle to detect the tool and then make predictions for new frames that come from a streaming data source. Another option is to run the

method offline and process a whole sequence at once. For each frame we compute a vessel mask and a centerline mask. To initialize the tracking, we sample a subset of the points resulting by the intersection of the two masks for the first frame. Selection of proper candidate points in next frames also makes use of vessel and centerline masks and additionally restricts the set of candidates to points that are reachable from any tracked point in the previous frame. Based on the Hungarian assignment, tracking a point from the first frame of a sequence to the last frame, produces a trajectory. We select the trajectory J that resembles most an interventional tool movement, by inspecting the Fourier Transform components of all trajectories. In online execution, the trajectory J is selected after pre-processing the frames from the first heart beating cycle and then only the last point from J is tracked to every subsequent frame. Segmenting the interventional tool centerline in each frame uses the centerline mask for the frame and the point of J that belongs to the same frame. The following subsections provide detailed description of each step in our method.

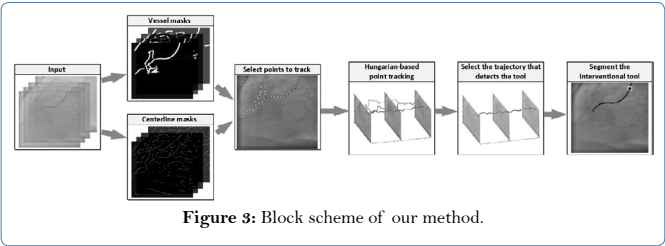


Figure 3: Block scheme of our method.

Vessel masks

The purpose of a binary vessel mask is to determine which pixels belong to vascular-like structures. We use the automatic vessel segmentation proposed in because it is computationally efficient and performs well compared to other state-of-the-art methods to segment vessels [15].

The vessel segmentation takes as input a single frame and computes an EdgeLog signal for it. The computation of EdgeLog signal uses the image, convoluted with Gaussian kernels of multiple scales and the sum of the second order scale-space derivatives for each pixel in both spatial dimensions. The vessel segmentation method also computes an Edgess signal for the frame and uses a region growing process on the EdgeLog and Edgess signals to segments vessels. The initialization of the region growing process introduces two parameters - Otsu threshold ϵ for the EdgeLog signal and a second threshold b , that determines if a seed point belongs to a blob-like structure, after comparison to the ratio of the Hessian matrix eigenvalues for the point.

Considering that the method in is designed for images with bright vessels, we invert the image (already normalized by the maximum pixel value) before segmenting the vessels: $I(x; y) = 1 - I(x; y)$ [15]. To adapt the vessel segmentation for coronary angiographies we define the set of scales $\Phi_{i \in \{1,2,3\}}$ in accordance to the expected size of cardiovascular catheters, which is between 5F(1.66mm) and 8F(2.66mm). We optimized the vessel segmentation on the same coronary angiographies that we used to tune the parameters of our method (see Section 7.1). To do so, we adjusted the blobness threshold parameter b to maximize the F-score of tool segmentation in regards to the ground truth annotations. For the rest of the parameters we used the proposed values in [15].

Computing vessel masks separately for each frame could produce inconsistent masks for consecutive frames - e.g., segmenting the tool in the current frame and failing to do so in the next one because of insufficient number of seed points. To suppress this possibility we introduce the restriction:

$$|M_V(t)| \geq |M_V(1)|, \tag{1}$$

where $M_V(t)$ is the vessel mask for the current frame t and $|M_V(\cdot)|$ is the number of non masked pixels in $M_V(\cdot)$. In cases when the restriction (1) is not satisfied we repeat the vessel mask generation for the frame, by gradually lowering the EdgeLog threshold ϵ and increasing the Blobness threshold b as geometric progressions with ratios 0.9 and 1.1 respectively. For the next frame we start the vessel mask computation with the initial settings of ϵ and b .

Figure 4b shows the vessel mask for the exemplar frame in Figure 4a. All parts of the catheter and the guide wire are within segmented vessel regions. The considerable amount of false positive vessel segmentations is mainly due to the low contrast and low signal to noise ratio in PCI angiographies. Our detection and tracking method is robust enough to tolerate such amount of false positive vessel segmentations (see the video in Section 5.4 that shows tracking a point from the catheter through a whole sequence).

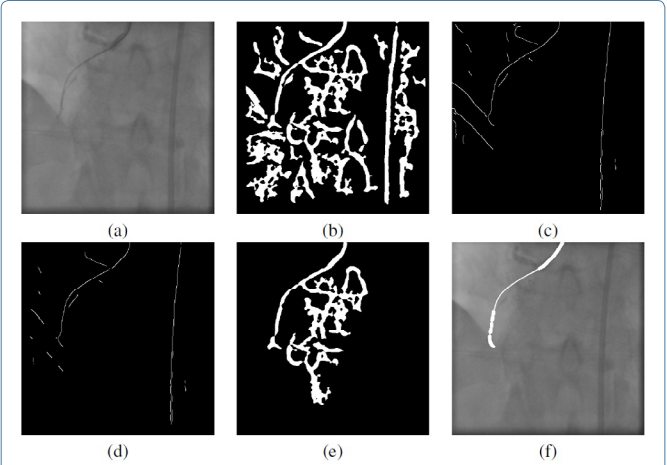


Figure 4: Visual results for each step of our method. (a) is the original frame, (b) is the vessel mask, and (c) is the centerline mask for the frame in (a). (d) is the intersection of (b) and (c). The enclosing vessel polygon for the point that tracks the interventional tool is shown in (e). (f) is the automatic segmentation of catheter and guide wire centerline for the frame in (a). Points predicted as part of the catheter are marked with thick line and thin line marks the guide wire prediction.

Centerline masks

The purpose of a centerline mask $M_C(t)$ is to segment the centerlines of vascular-like structures in frame t of an angiography video sequence. The reported results from the centerline extraction of vascular structures in are quite promising, so we use the same method to automatically extract vessel centerlines [16]. First, static objects are eliminated from the frame by subtracting the median of the first 10 frames of the sequence. Then, the method enhances vessels with the vesselness filter in using the set of scales ϕ and taking the maximum response per scale at every pixel. Subsequently, non-maximum suppression retains ridges in the frame [17]. The ridges signal is convoluted with a Gaussian kernel of standard deviation $1px$ and hysteresis thresholding removes weak ridges. The upper bound of the hysteresis thresholding was set to 0.0116 after optimizing the F-score of the

masks in regards to the annotated ground truth in the same sequences that we used to tune the parameters of our method (see section 7.1). For the lower bound value we use 0.0103 as suggested in [16]. The final centerline mask $M_C(t)$ is obtained after a component analysis removes centerlines shorter than 8 pixels.

Figure 4c shows the centerline mask for the exemplar frame in Figure 4a. Intersecting the centerline mask with the vessel mask removes many of the false positive interventional tool centerlines, by preserving the true positives (Figure 4d). For example, the diaphragm border in Figure 4c is completely segmented by the centerline mask, and only partially by the vessel mask in Figure 4b. This is due to the EdgeLog signal not being affected by edges.

Tracking initialization

To define the set of points to track we use the centerline mask $M_C(1)$ for the first frame of the sequence. By combining vessel and centerline masks, we reduce the amount of centerline points that do not belong to vessels. The intersection $M_V(1) \cap M_C(1)$ gives us a set of possible initialization points. Tracking all points would be computationally inefficient, so we employ a deterministic strategy based on k-means clustering and morphological shrinking to evenly subsample a set of initialization points from $M_V(1) \cap M_C(1)$. The subsampling rate c is a parameter of our method that is determined empirically. In the initialization stage there is no evidence which points belong to an interventional tool and the amount of false positives could be greater than the amount of true positives.

Figure 5 shows the sampled points in the first frame to initialize the tracking for the same sequence as in Figure 4. Due to the absence of contrast agent in arteries and the combination of vessel and centerline masks, most of the initialization points belong to the interventional tools.

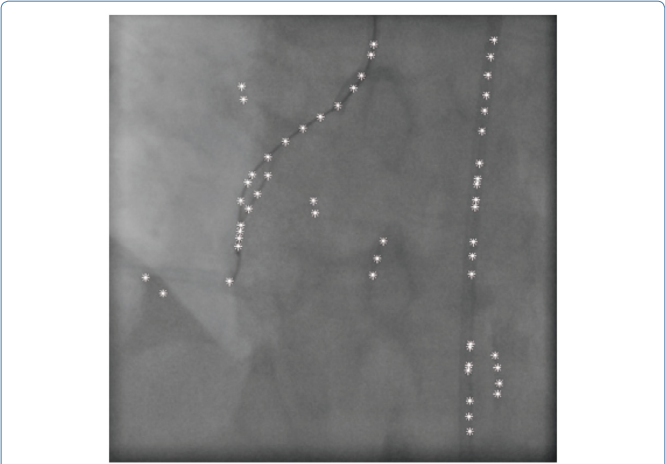


Figure 5: Sampled points to initialize the tracking in the first frame for the same sequence as in Figure 4.

Hungarian tracking

The Hungarian method was developed to find the optimal matching in a weighted bipartite graph that minimizes the total cost of edges (a.k.a. the assignment problem). Thus, finding the optimal tracking for a set of points from one frame to the next frame can be formulated as defining a suitable cost function and selecting proper candidate points in the next frame.

Let P_m be the set of points in the moving frame I_m , which we want to track in the static frame I_s . To define P_s - the set of candidate points in I_s , we use the same subsampling strategy as in Section 4.3, but on I_s . At this step, subsampling by clustering and shrinking may remove the point that we need to track in I_s . For precise point tracking this will be an issue, but for the purposes of our method it is acceptable as long as the best matching point in I_s , after subsampling, belongs to the centerline of the interventional tool (section 4.5). Then segmentation of the whole centerline depends mostly on the selection of seed points (section 4.6). In addition, we remove from P_s the points that theoretically cannot be reached from any point from the moving points P_m . Peak velocity of coronary arteries due to cardiac and breathing motion could reach 180 mm/sec ([18,19]).

The quality of the tracking also depends on how discriminative is the cost provided to the Hungarian algorithm. We propose one cost based on appearance descriptors and two costs based on intra-frame and inter-frame geometrical distribution of candidate points. The following linear combination blends these three contributions into a cost function for connecting two points p_m and p_s :

$$F(p_m, p_s) = w_1 HOD + w_2 ED + w_3 DOC. \quad (2)$$

The weights w_1 , w_2 and w_3 are parameters of our method.

Appearance-based cost:

Histogram of Differences Measure provides information about the texture alignment between two image regions. In authors show this measure to be robust against inflow of contrast agent [10]. Let a region $\Omega(p)$ around a point p be the square area specified by the two corners (p_{x-r+1}, p_{y-r+1}) and (p_{x+r-1}, p_{y+r-1}) . We set r to be slightly bigger than the radius of the largest expected catheter ($4F$). Histogram of differences $H(\delta)$ is computed over the gray level differences of corresponding points from the regions $\Omega(p_m)$ and $\Omega(p_s)$:

$$\delta = I_m(x_m, y_m) - I_s(x_s, y_s) \Big|_{\substack{(x_m, y_m) \in \Omega(p_m) \\ (x_s, y_s) \in \Omega(p_s)}}. \quad (3)$$

The histogram is normalized so that $\sum_{\delta} H(\delta) = 1$. Big value of $v = \sum_{\delta} H^2(\delta)$ corresponds to low dispersion. At the same time low dispersion indicates good alignment, so $HOD(p_m, p_s) = v^{-1}$.

Geometry-based costs:

Euclidean Distance provides information about the spatial distance between two points p_m and p_s :

$$ED(p_m, p_s) = d(p_m, p_s). \quad (4)$$

Since we do not expect interventional tools to make big jumps between subsequent frames, with this cost we penalize large distances between a point from the moving frame to a point in the static frame.

Difference of Correlograms - provides context-based information about the similarity of the coronary tree around two points. This measure uses contextual information by dividing the area around a point in sectors and counting the number of centerline points in each sector. Many candidate points alongside a tubular structure would provide similar values for the other two measures in the cost function (2) and that would make the tracking tolerant to the drifting problem (see section 3). Difference of Correlograms adds additional discrimination

by measuring the difference in the distribution of centerline points, around points p_m and p_s . [20]. To construct a Correlogram for a point p , we quantize the polar coordinates (angle and radius) of a circle area $\Omega(p)$ around p into 6 bins for the angle and 3 bins for the radius. The radius of Ω is set to the maximum displacement of coronary arteries between frames (section 4.4). The number of bins was chosen after optimization on our validation data set (see section 5.1). Counting the centerline points from the mask M_c in each bin produces a vector $o_p = (o_1, o_2, \dots, o_{18})$ for a point p . To normalize o_p we divide each of its elements by the size of the corresponding bin Ω_i . Hence,

$$DCO(p_m, p_s) = d(o_{p_m}, o_{p_s}). \quad (5)$$

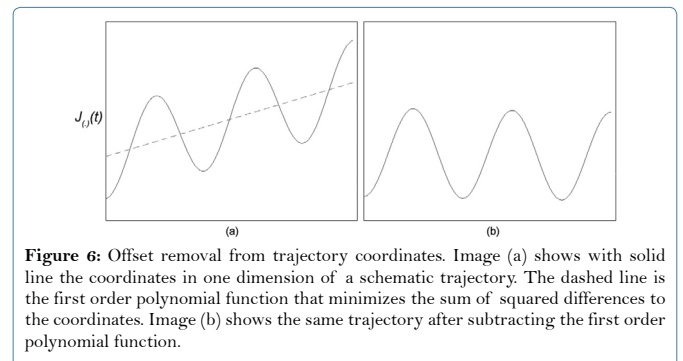
An advantage of using the Hungarian algorithm in tracking is that it optimizes the matching between two frames at once. Given a sparse initialization, we want to avoid one trajectory collapsing into another, which the Hungarian ensures due to the simultaneous (and optimal) pairing of points.

Automatic interventional tool detection

To detect catheter and guide wire, we select the most probable trajectory to follow a moving part of an interventional tool. To do so, for every trajectory we inspect the Fourier components of its movement. Our selection criterion is based on two assumptions:

- 1) Interventional tools move within heart beating frequency range, since physicians insert them in one of the coronary arteries.
- 2) Trajectories, with initial points within an interventional tool, start from a pixel that belongs to a vascular-like structure. This is because catheter and guide wire look similar to arteries filled with contrast agent (Section 3). Also, in the first frames of coronary angiographies, contrast agent is not expected to be injected, which helps automatic detection and segmentation of interventional tools.

Let $J_x(t)$ be the discrete function of a trajectory coordinates in x over time t and $J_y(t)$ the same for the coordinates in y . For tool detection we are mainly interested in the frequency of the movement, regardless of its offset in time. Hence, we subtract from $J_{(x)}(t)$ the first order polynomial function that minimizes the sum of squared differences to the original signal $J_{(x)}(t)$. Figure 6 illustrates the effect of offset removal from trajectory coordinates in one dimension. In image (a) the solid line shows one dimensional coordinates of a schematic trajectory and the dashed line is the first order polynomial function that minimizes the sum of squared differences to the coordinates. Image (b) shows the same trajectory after subtracting the first order polynomial function.



Fourier analysis over time for each trajectory gives the decompositions of component amplitudes across the frequency spectrum for both dimensions - $J_x^{\wedge}(\xi)$ and $J_y^{\wedge}(\xi)$. We formulate the first assumption as maximizing the score:

$$S(J) = \frac{\int_{0.63\text{Hz}}^{1.67\text{Hz}} \sqrt{\hat{J}_x(\xi)^2 d(\xi) + \hat{J}_y(\xi)^2 d(\xi)} d\xi}{\int_{\infty} \sqrt{\hat{J}_x(\xi)^2 d(\xi) + \hat{J}_y(\xi)^2 d(\xi)} d\xi}. \quad (6)$$

The ratio $S \in [0,1]$ measures to what extend the decomposition to frequencies of a trajectory movement coincides with the expected heart beating frequency range during coronary angiographies [0.63Hz, 1.67Hz] [21].

To address the second assumption we consider the output of the vesselness filter for the first frame [17]. Then, we rank all trajectories according to their score S and select the first one that starts from a pixel with bigger vesselness value than the mean vesselness value for all points that start a trajectory. We denote the selected trajectory, which is the most probable to follow an interventional tool, with J^* .

If the method is run online, the selection of trajectory J^* makes use only of the first heart beating cycle of the sequence. Then the method continues only the trajectory J^* to the next streaming frame. The offline mode uses the trajectories from the whole sequence.

Automatic interventional tool segmentation

To segment the centerline of the interventional tool in frame I_v we use the point $J^*(t)$ from the trajectory that follows the tool and find its enclosing vessel polygon $M_v^*(t)$ in the vessel mask $M_v(t)$. We employ a classic region growing algorithm on $M_v(t)$, with $J^*(t)$ as initial seed point. An example of enclosing vessel polygon is shown in Figure 4e. Starting from the point $J^*(t)$ and repeating the process in both directions of the centerline, we iteratively select control points for the final B-spline interpolation of the interventional tool. New control point is added by intersecting the set of centerline points in $M_v^*(t) \cap M_v(t)$ with a circle having as center the previously selected control point and as radius the sampling rate c . Any control points that have been already selected are discarded from the candidates for a next seed point. If there is more than one candidate, we select the point that has the biggest distance to the last selected seed point, excluding the center of the circle. In this way, we lower the chance for our method to be deluded by nearby centerlines belonging to other structures. If there are no candidates, the search for seed points in the chosen direction stops.

As noted in section 3, when contrast agent flows from the catheter into the blood vessel, both structures may look like a single vessel. To avoid the possibility of selecting seed points from centerlines that do not belong to an interventional tool we impose the restriction that the number of selected seed points does not surpass the number of seed points in the first frame. In addition, when selecting seed points we start from the direction that is opposite to the direction of the blood flow. This is usually the direction that leads to the point of the curve that is closest to one of the sides of the frame, since the visible part of interventional tools usually begins from one of the image sides.

By fitting a B-spline curve to all seed points we obtain the centerline of catheter and/or guide wire in the frame (Figure 4f).

Differentiation between catheter and guide wire

For each pixel from the segmented interventional tool centerline we determine if it belongs to a catheter or a guide wire. The guide wire is seen as a thin curve, which would give maximum output of the vesselness filter in for scale $\sigma \approx 1px$. [17]. Hence, we inspect the first order polynomial function that minimizes the sum of squares to the normalized output of for scales $\sigma \in \{1px, \phi_1, \phi_2, \phi_3\}$. [17]. If the function decreases with increasing the scale, we mark the pixel as part of a guide wire. Otherwise we mark it as a part of a catheter. Figure 4f shows a differentiation between catheter and guide wire. At its lower part, the guide wire is thicker due to its bending and is wrongly segmented as catheter. The quantitative results from the evaluation of our differentiation between catheter and guide wire are in section 7.4.

Evaluation

Material

To fine-tune the parameters of our method we used 8 sequences of coronary angiographies and to test the method performance we used additional 39 sequences. Both the validation and the testing data sets contain sequences from Percutaneous Coronary Interventions and Quantitative Coronary Angiographies. A major difference between the two types of sequences is that the amount of injected Contrast agent during PCI is smaller and often it is not trivial to distinguish arteries from other overlapping structures. Also, in PCI, a guide wire is inserted through the catheter in the affected artery, while for the purpose of estimating the healthiness of the myocardium physicians insert only a catheter. Contrast agent injection is present in all of the sequences, 15 of which have been acquired with Philips Allura Xper, and the other 32 with Siemens AXIOM-Artis. The pixel resolution varies from 0.22×0.22 mm to 0.34×0.34 mm and the acquisition frame rate is from 12 to 30 fps. The C-arm primary and secondary angles vary from -39° to 97° and from -40° to 37° respectively. Two experts independently annotated ground truth for the interventional tools in 114 frames totally. For each pixel from the ground truth the experts have put a label if it belongs to a guide wire centerline or to a catheter centerline.

The sequences used in are not available so we couldn't test our method on them [3,11].

Protocol

In our experiments we match each prediction for a tool centerline to the ground truth annotation in several aspects. The ratio $\text{Precision} = \frac{TP}{TP+FP}$ gives information on what part of the predicted curve corresponds to an interventional tool centerline. True Positive prediction (TP) is the number of predicted pixels, having a pixel from the ground truth within a specified threshold distance. To be compatible with the evaluation protocol from we set the threshold distance to $3px$. [3,11]. False positive prediction (FP) is the number of the remaining pixels from the predicted curve, after excluding the true positive prediction. In order to measure what part of the ground truth has been segmented we compute the Sensitivity = TP/GT where GT

is the total number of pixels in a ground truth curve annotation. A single measure that combines Precision and Sensitivity is F-score =
$$2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

To assess the localization error of correctly segmented parts of a curve, for each point of the curve we take the distance to the closest point from the ground truth, as long as the closest point is within the threshold distance of 3px. Then we compute the average distance to estimate how close a correctly segmented centerline is to the manual annotation.

To show the performance of the tool tracking regardless of the final segmentation, we evaluate how well the selected trajectory J* detects the interventional tool. In each frame with ground truth we check if the corresponding point from J* lies within the threshold distance to a point from the ground truth curve.

Testing just the vessel and the centerline segmentation in terms of Sensitivity and Precision gives an oversight on the contribution of the subparts in our method.

We tested the differentiation between catheter and guide wire by computing the percentage of correct labels in successfully predicted tool centerlines.

Parameters tunin

To fine-tune the parameters of our method we adopted the random search strategy proposed in which empirically and theoretically has been shown to be more efficient than grid search or manual search [22]. The parameters that we tuned are the subsampling rate c , the quantization of the polar coordinates for Correlograms construction and the weights w_1 , w_2 and w_3 in the Hungarian cost function. The parameter tuning was performed on two stages, each of 100 random assignments. In the first stage the random value for each parameter was chosen from a broad domain ([0 100] for w_1 and w_2 , [10 50] for c and [1 10] for each of the two polar coordinates) and in the second stage each domain was defined around the optimal value from the first stage, that maximizes the F-Score of the predictions. As a result we set the bins in the Correlograms to 18 (6 for the angle and 3 for the distance), the weights $w_1 = 0.1508$, $w_2 = 0.1321$, $w_3 = 9.95$ and the sampling rate to $c = 20$.

The four plots in Figure 7 illustrate how sensitive is our method to changes in the values for the subsampling rate c and the weights in the Hungarian cost function w_1 , w_2 and w_3 . In each plot we kept the optimal values for the other three parameters. The performance on the testing data is close to the results with the optimal settings, so the method maintains a balance between the need of precise parameter settings and the extent to which the parameter tuning affects the performance. The subsampling rate c affects the performance of our method the least. The most influential parameter is the weight w_2 for the Euclidean distance in the tracking cost function; without it we measured the lowest precision and sensitivity in our experiments.

Results

Table 2 shows the quantitative results for the precision and the sensitivity of our method, both for its online version (only the frames from the first heart beating cycle are pre-processed) and the offline version (the whole sequence is processed at once). The first row is the

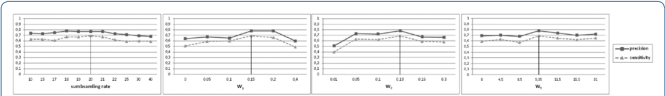


Figure 7: Changes in the performance of our method for different values of the subsampling rate c and the weights in the Hungarian cost function w_1 , w_2 and w_3 . In each plot we kept the optimal values for the rest of the parameters (marked with vertical line in each plot).

inter-observer variability in our testing data set. For overall comparison, the table contains the results for the state of the art method that automatically detects a catheter in the presence of contrast agent in the cardiovascular system [5]. After we executed the method in on our testing data, we observed bad generalization on the new data and the updated performance is significantly lower than the reported performance in [5]. Our method performs better and achieves closer results to the inter-observer variability. Although the results do not match the inter-observer variability, we find them indicative of the potential of our method to automatically detect and segment interventional tools in complex cases like presence of noise, contrast agent and other structures that resemble a catheter or a guide wire. The offline version of the method, when the whole sequence is pre-processed before the tool detection, performs better than the online version, when only the first heart beat from the sequence is pre-processed. This is because the detection is more precise when using information from the whole sequence (Table 3).

Table 4 shows the average localization error (in pixels) for the correctly predicted curve segments. We also computed the inter-observer localization error by using each of the ground truth curves as a prediction and matching it to the other ground truth curve for the same frame. Both the online and the offline version of our method achieve close results to the inter-observer localization error.

To further show the effectiveness of our tracking, regardless the final segmentation, we provide the average precision of the trajectory J* that tracks the tool (Table 3). The offline version of the method, that uses movement frequency information from the whole sequence, detects interventional tools slightly better than the online version, that uses only a part of the sequence to select the trajectory J*. In both cases the tracking precision outperforms the segmentation precision from Table 2.

Table 5 demonstrates the contribution of vessel and centerline masks as subparts of our method and also the effect of combining the two masks. The precision and the sensitivity were computed on our testing data and tolerance of 3 pixels was considered when matching centerlines to ground truth annotations. Both masks, if applied on their own, contain most of the ground truth curves but the precision in both cases is very low. After combining centerline and vessel masks, the sensitivity loses 2% but the precision increases twice - from 6% to 12%. This improvement has a direct positive effect on the computational cost of the method, as it decreases twice the number of points to track (section 4.3) and the number of candidate points in next frames (section 6.4).

Figure 2 shows four visual results of our method together with one of the two annotated ground truth curves. Our prediction is marked with sparse curve of + symbols and the ground truth is marked with dotted curve. To ease the visualization we use identical annotation for catheter and guide wire. Frames (a) and (b) in the first row have been taken from MBG estimation sequences and frames (c) and (d) from

Method	Precision				Sensitivity				F-score			
	Avg	Std	Min	Max	Avg	Std	Min	Max	Avg	Std	Min	Max
O1 vs O2	0.85	0.16	0.16	1	0.82	0.15	0.40	1	0.82	0.15	0.26	1
Our Method (offline)	0.78	0.23	0.10	1	0.69	0.31	0.06	1	0.69	0.26	0.08	1
Our Method (online)	0.71	0.30	0	1	0.62	0.34	0	1	0.63	0.30	0	0.99
Hernandez '12 [5]	0.60		-	-	0.52		-	-	0.56		-	-

Table 2: Quantitative results for the precision and the sensitivity of our method compared to the state of the art method for fully automatic detection of catheter in the presence of contrast agent in the cardiovascular system. The first row is the inter-observer variability in our testing data set.

Method	Precision				Number of sequences
	Avg	Std	Min	Max	
Offline version	0.94	0.20	0	1	39
Online version	0.87	0.33	0	1	39

Table 3: Precision of our fully automatic detection and tracking of interventional tools, without considering segmentation.

	Localization error	
Offline version	1.36	0.27
Online version	1.33	0.24
Inter-observer error	1.27	0.35

Table 4: Average localization errors (in pixels) for the correctly predicted curve segments, both for the offline and the online version of the method. The third row shows the inter-observer localization error.

Masks	Precision	Sensitivity
Vessel masks	0.006 0.01	0.94 0.1
Centerline masks	0.06 0.07	0.91 0.09
Both masks	0.12 0.13	0.89 0.13

Table 5: Contribution of vessel and centerline segmentation and the effect of combining the two masks. The precision and the sensitivity of the masks were computed on our testing data.

PCI sequences. The first two images have higher contrast because of the bigger amount of injected contrast agent. Our prediction in (a) and (b) is almost identical to the ground truth, despite the presence of some of the challenges listed in section 3 - in (a) the tip of the catheter is not clearly visible within the contrast agent filling the artery and part of the aorta (the dark oval area around the catheter tip). Part of a diaphragm border almost coincides with part of the catheter centerline in (b), which has not hampered neither the correct detection nor the correct segmentation. In frame (c) our method is misled by a coronary artery and the prediction does not completely coincide with the ground truth. Frame (d) is an example of suboptimal prediction when contrast agent completely covers part of the guide wire and the ground truth annotation is not trivial. In this case only the catheter curve was segmented. We uploaded a video (<https://drive.google.com/open?id=0B9FghXnnTtCsNUoyZWlmeEFOGc>) of automatic catheter tracking for one of the sequences in our testing data to show how our method performs in the presence of other structures with high contrast that overlap the object of interest. The main contribution for keeping low amount of point drifting is the incorporation of the difference of correlograms cost in the Hungarian tracking.

The disadvantage of our method is its computational cost - a non-optimized MATLAB implementation preprocesses frames with 0.07 fps in offline mode and the online version processes new streaming frames with 0.09 fps. Computation of vessel and centerline masks

takes most of the computational time, while the online tracking on its own runs with 0.68 fps in offline mode and with 1.82 fps in online mode. Once detected, the tool centerline is segmented with 9.32 fps. Regarding the differentiation between interventional tools, our method successfully recognizes 92% of the pixels from the correctly predicted tool centerlines if they belong to a catheter or to a guide wire.

Discussion and Conclusion

In this paper we proposed a method that detects, tracks and segments the centerline of catheter and guide wire through entire coronary angiographies. The method is fully automatic and does not require manual initialization, neither needs training data. All previous methods that have been applied on coronary angiographies rely on manual initialization except the fully automatic method in which detects the catheter, but does not track it [1-3,5,11]. Main novelties in our approach are the usage of the Hungarian algorithm to estimate the optimal tracking of structures between images and the analysis of trajectories in the frequency domain to detect the object of interest.

We explicitly described the main challenges for tracking structures in coronary angiographies (section 3) and designed our method to be capable of handling them. Our evaluation was performed entirely on sequences with injection of contrast agent. Many related methods address cases with contrast agent, but do not provide results to show its effect on algorithm performance. As for the computational time, it is still a challenge for fully automatic methods to process sequences in real-time.

In addition, to the best of our knowledge, our method is the first that differentiates between the two types of interventional tools in PCI - catheter and guide wire.

Among the possible applications, pointed as motivation for our work in section 1, the proposed method can improve registration of images from different modalities. Multi-modal registration between X-ray and other imaging modalities (e.g., CT, MRI, Ultrasound) is a research topic directed to provide physicians with complementary information about the coronary tree [5,6]. The method in is an example of using catheter segmentation in multi-modal registration [6].

The proposed method has a number of limitations. It detects only the moving part of the catheter and the guide wire. Any other parts of the tools that are not inside a coronary artery are not likely to be detected. Images, in which some parts of the tools are not visible due to low contrast may have incomplete centerline or vessel masks, which would result in suboptimal predictions.

We see several directions for future investigation and next steps to improve our method. Imposing proper restriction on the deformation of the object of interest would reduce the amount of false

positive segmentations. Defining and detecting when structures appear or disappear is important to model the spatio-temporal dynamics of x-ray videos, which could be beneficial for simultaneous tracking and segmentation. An essential future work is optimizing the implementation of vessel and centerline masks, in order to achieve run-time performance.

We consider our method to be a step ahead towards applying fully automatic segmentation and tracking of structures in medical practices that use coronary angiography.

Acknowledgement

The authors would like to thank T. van Walsum from Erasmus MC, Rotterdam for his proposals and feedback during the development of the method and the preparation of this article.

References

1. Baert SA, Viergever MA, Niessen WJ (2003) Guide-wire tracking during endovascular interventions. *IEEE Trans Med Imaging* 22: 965-972.
2. Pauly O, Heibel H, Navab N (2010) A machine learning approach for deformable guide-wire tracking in fluoroscopic sequences. *Med Image Comput Comput Assist Interv* 13: 343-350.
3. Heibel H, Glocker B, Groher M, Pfister, Navab N (2013) Interventional tool tracking using discrete optimization. *IEEE Transactions on Medical Imaging* 32: 544-555.
4. Markelj P, Tomaevic D, Likar B, Pernu F (2012) A review of 3d/2d registration methods for image-guided interventions. *Medical Image Analysis* 16: 642-661.
5. Hernandez-Vela A, Gatta C, Escalera S, Igual L, Martin-Yuste V, et al. (2012) Accurate coronary centerline extraction, caliber estimation, and catheter detection in angiographies. *IEEE Transactions on Information Technology in Biomedicine* 16: 1332-1340.
6. Wang P, Chen T, Ecabert O, Prummer S, Ostermeier M (2011) Image-based device tracking for the co-registration of angiography and intravascular ultrasound images. *Med Image Comput Comput Assist Interv* 14: 161-168.
7. Volpi D, Sarhan MH, Ghotbi R, Navab N, Mateus D (2015) Online tracking of interventional devices for endovascular aortic repair. *Int J Comput Assist Radiol Surg* 10: 773-781.
8. Baert SAM, van de Kraats EB, Walsum TV, Viergever MA, Niessen WJ (2003) Three-dimensional guide-wire reconstruction from biplane image sequences for integrated display in 3-d vasculature. *IEEE Transactions on Medical Imaging* 22: 1252-1258.
9. Bender H-J, Mnner R, Poliwoda C, Roth S, Walz M (1999) Reconstruction of 3d catheter paths from 2d x-ray projections. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 981-989.
10. Meijering EHW, Zuiderveld KJ, Viergever MA (1999) Image registration for digital subtraction angiography. *International Journal of Computer Vision* 31: 227-246.
11. Wang P, Chen T, Zhu Y, Zhang W, Zhou S, et al. (2009) Robust guidewire tracking in fluoroscopy. *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* 1063-6919.
12. Milletari F, Belagiannis V, Navab N, Fallavollita P (2014) Fully automatic catheter localization in c-arm images using l1-sparse coding. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014* 8674 570-577.
13. Cootes TF, Marsland S, Twining CJ, Smith K, Taylor CJ (2004) Groupwise Diffeomorphic Non-rigid Registration for Automatic Model Building. *European Conference on Computer Vision* 3024: 316-327.
14. Ma Y, Gogin N, Cathier P, Housden RJ, Gijsbers G, et al. (2013) Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions. *Medical Physics* 40.
15. Romero A, Petkov S, Gatta C, Sabate M, Radeva P (2012) Efficient automatic segmentation of vessels. *Signal* 1: 1-5.
16. Baka N, Metz CT, Schultz CJ, van Geuns RJ, Niessen WJ, et al. (2014) Oriented gaussian mixture models for nonrigid 2d/3d coronary artery registration. *IEEE Trans Med Imaging* 33: 1023-1034.
17. Frangi AF, Niessen WJ, Vincken KL, Viergever MA (1998) Multiscale vessel enhancement filtering. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*. 130-137.
18. Hofman MB, Wickline SA, Lorenz CH (1998) Quantification of in-plane motion of the coronary arteries during the cardiac cycle: Implications for acquisition window duration for mr flow quantification. *J Magn Reson Imaging* 8: 568-576.
19. Shechter G, Resar JR, Mcveigh ER (2006) Displacement and velocity of the coronary arteries: cardiac and respiratory motion. *IEEE Trans Med Imaging* 25: 369-375.
20. Amores J, Sebe N, Radeva P (2007) Context-based object-class recognition and retrieval by generalized correlograms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.
21. Torres FS, Jeddiyan S, Jimnez-Juan L, Nguyen ET (2011) β -Blockers to Control Heart Rate during Coronary CT Angiography. *Radiology* 259: 615-616.
22. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *Journal of Machine Learning Research* 13: 281-305.