Image-based Co-Registration of Angiography and Intravascular Ultrasound Images

Peng Wang*, Olivier Ecabert, Terrence Chen, Michael Wels, Johannes Rieber, Martin Ostermeier, and Dorin Comaniciu

Abstract—In image-guided cardiac interventions, X-ray imaging and intravascular ultrasound (IVUS) imaging are two often used modalities. Interventional X-ray images, including angiography and fluoroscopy, are used to assess the lumen of the coronary arteries and to monitor devices in real time. IVUS provides rich intravascular information, such as vessel wall composition, plaque, and stent expansions, but lacks spatial orientations. Since the two imaging modalities are complementary to each other, it is highly desirable to co-register the two modalities to provide a comprehensive picture of the coronaries for interventional cardiologists. In this paper, we present a solution for co-registering 2-D angiography and IVUS through image-based device tracking. The presented framework includes learning-based vessel detection and device detections, model-based tracking, and geodesic distance-based registration. The system first interactively detects the coronary branch under investigation in a reference angiography image. During the pullback of the IVUS transducers, the system acquires both ECG-triggered fluoroscopy and IVUS images, and automatically tracks the position of the medical devices in fluoroscopy. The localization of tracked IVUS transducers and guiding catheter tips is used to associate an IVUS imaging plane to a corresponding location on the vessel branch under investigation. The presented image-based solution can be conveniently integrated into existing cardiology workflow. The system is validated with a set of clinical cases, and achieves good accuracy and robustness.

Index Terms—Angiography, image registration, object detection, object tracking, ultrasonic imaging.

I. INTRODUCTION

INTERVENTIONAL cardiologists rely on medical images to perform minimally invasive procedures. X-ray angiography is the standard imaging modality for the assessment of coronary artery disease. Thereby, a radio-opaque contrast agent

is injected into the circulation system. The propagation of the contrast agent, which is driven by blood flow, is documented under X-ray radiation. These images are the basis to identify stenoses within diseased vessels that restrict blood flow. X-ray imaging provides the spatial and structural information of the coronary arteries at good temporal resolution. However, as angiography is a lumenographic method, changes in the vessel wall cannot be detected. Intravascular ultrasound (IVUS) is an invasive imaging modality used during coronary catheterization to complement the information provided by angiography. In particular, IVUS delivers insights beyond the angiographic vessel lumen. It can be utilized to support the assessment of plaque burden and stent deployment, as well as for assessing the underlying significance of ambiguous coronary lesions [7], [8]. However, IVUS lacks spatial orientation and makes it difficult to fully understand the spatial structure of vessels. It is therefore desirable to combine angiography and IVUS to provide a comprehensive picture of coronaries for interventional cardiologists [2], [11]. In current practice, the relationship between angiographic and IVUS images is left to the experience of the physician using anatomical landmarks such as branches or lumen narrowing. However, early stage diseases or long diffused lesions make it even more challenging to mentally establish a correspondence as no evident anatomical landmarks can be used.

The co-registration of angiography and IVUS or other intravascular imaging modalities has been studied in [2], [5], [9], [10], [13], [14], [18]–[20]. In the work of Wahle *et al.* [18], [19], IVUS images are located along a 3-D vessel branch that is reconstructed from biplanar angiography, and the orientation of IVUS images is further optimized. However, since the aforementioned methods are unable to identify reliable landmarks to fuse angiography and IVUS, they use the motorized pullback and assume a constant moving speed of the IVUS transducer along the vessel. This assumption does not always hold in clinical practice, as the IVUS transducer is sometimes slowed down in stenosed areas. The assumption also limits the usage of manual pullback. An alternative solution is to use manual labeling to identify corresponding landmarks in both angiography and IVUS images, as these preliminary systems [2], [13]. In [13], manually labeled landmarks are used as a baseline to co-register IVUS images to a 3-D reconstructed vessel from two contrasted angiographic images. In [5], [10], IVUS and CTA are co-registered through angiography. However, manual landmark annotations are still used in these methods. Manual labeling is labor intensive, and sometimes good quality landmarks such as stent and major vessel branches are difficult to identify

Fig. 1. A 2-D image-based solution for the co-registration of angiography and IVUS. (a) Angiography provides vessel structure information. The cross indicate a registered IVUS imaging plane. (b) Fluoroscopic image acquired during the pullback. The circle represents the automatically tracked IVUS transducer. (c), (d) Cross-sectional and longitudinal views of IVUS images. The lines in the longitudinal views indicate the IVUS slice corresponding to the position on the vessel in (a).
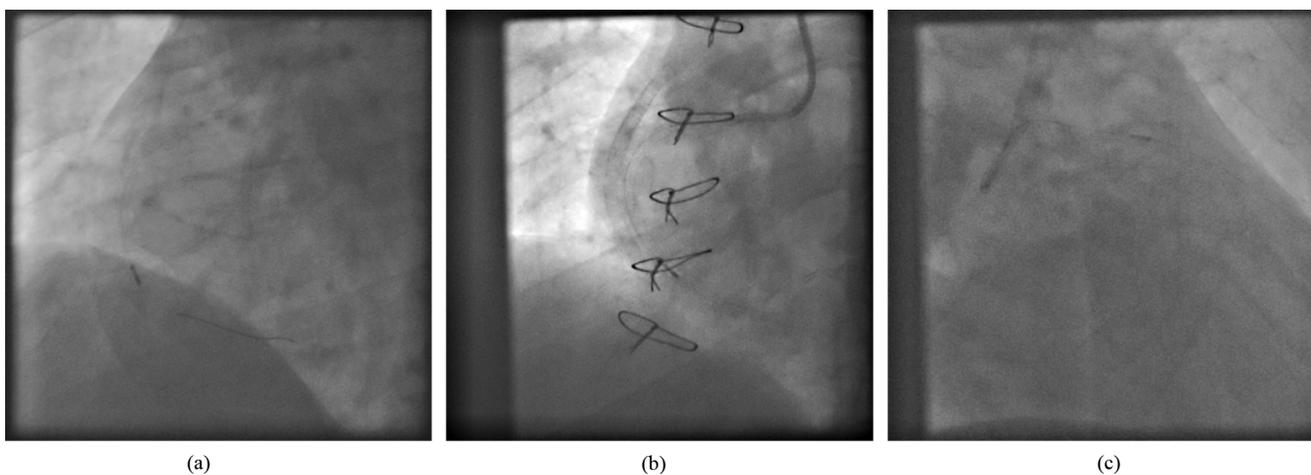
Fig. 2. Challenges for device tracking in fluoroscopy. (a) Low image quality and cluttered background. (b) Overlapping with other surgical structures. (c) Arbitrary projection angle.

or even unavailable. Such limitations make the manual labeling methods not suitable for the usage in clinical practice. In [3], the IVUS transducers are tracked in 2-D, and reconstructed in 3-D space with a prior information of pullback speed. However, the method ignores breathing and cardiac motions in cardiac interventions, and is only validated on synthetic and phantom data.

To handle such challenges in a real-world clinical scenario, we present a novel solution for the co-registration of angiography and IVUS images. The solution is briefly illustrated in Fig. 1. Our solution is image-based, and does not require additional hardware such as electromagnetic sensors [12]. Compared to the approaches cited above, the presented method is purely based on 2-D without requiring a 3-D vessel reconstruction from the biplane angiography, or from sequences acquired from different angulations. The additional interaction required by the cardiologist for the 3-D reconstruction may be a barrier in the regular use in clinical routine, especially in sites of high volumes. Our approach requires just two clicks for the initialization and works independently of the chosen angulation. In addition, it does not assume a constant pullback speed and it is automated in tracking and co-registration. We develop a method to automatically detect and track the movement of IVUS transducers in 2-D fluoroscopy in real time and then map a trans-

ducer imaging plane to a corresponding position in a reference angiography image. The 2-D image-based solution simplifies the clinical workflow compared to the approaches introduced above and meets the requirement of clinicians' daily operations. Finally, the method outlined in this paper is not specific to particular X-Ray or IVUS systems.

Accurate and robust tracking of devices in X-ray images faces great technical challenges. The devices undergo cardiac and breathing motions during interventions. Even with ECG triggering, the breathing motion can be large. X-ray images, especially fluoroscopy, usually have a low signal-to-noise ratio and the appearance of devices can be blurry. There are also artifacts and other surgical devices in X-ray images. Some examples are shown in Fig. 2 to illustrate the challenges that image-based solutions face. Considering all the challenges, conventional tracking methods based on intensity appearance [24] or edge-based energy [4] will encounter problems. This paper presents a computational framework that includes learning-based detection and model-based probabilistic tracking. The learning-based detection identifies devices present in X-ray images. Although only guiding catheter tips and IVUS transducers are used in the final co-registration, the method also detects a guiding catheter body and a guidewire
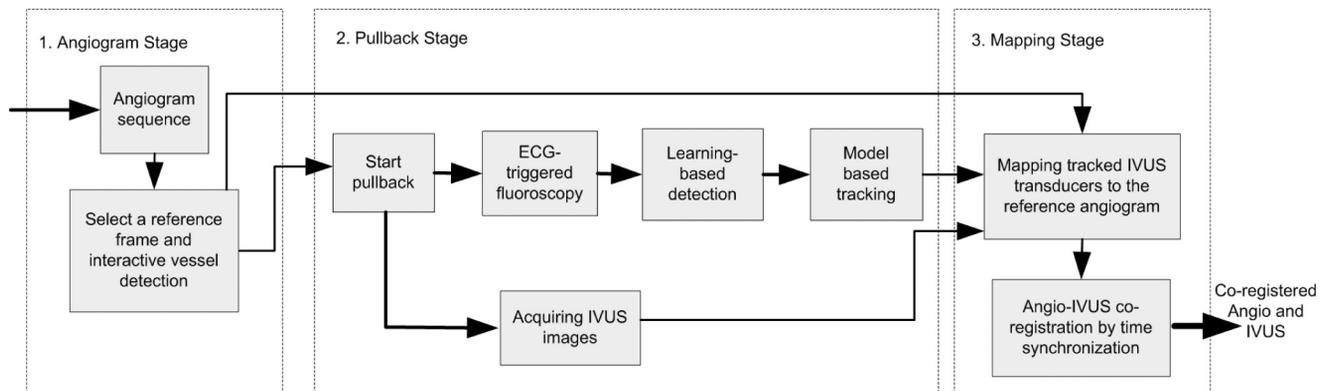
Fig. 3.   Co-registration system workflow.

to help subsequent tracking of guiding catheter tips and IVUS transducers. The detection results are then used in the probabilistic model-based tracking to achieve the robust tracking of the IVUS transducer movement inside vessel. The presented framework does not assume a constant IVUS pullback speed in 2-D. Quantitative evaluation on a set of clinical cases acquired in several hospitals shows that the framework is robust for tracking and accurate for the co-registration of IVUS and angiography images. We believe that the presented co-registration framework can provide a useful tool for many clinical applications, such as advanced vessel quantitative analysis across different imaging modalities.

The rest of the paper is structured as follows. The overall system workflow is outlined in Section II. We present the initialization step in Section III, the learning-based device detection in Section IV, the model-based tracking framework in Section V, and the co-registration based on tracking in Section VI. The quantitative evaluation results are given in Section VII. We conclude the paper with discussions in Section VIII.

II.   SYSTEM WORKFLOW

The system workflow includes three stages: the "angiogram" stage, the "pullback" stage, and the "mapping" stage. The purpose of the angiogram stage is to image the contrast-enhanced coronary arteries, and to identify the vessel branch where the IVUS imaging will take place. In the pullback stage, an IVUS transducer is pulled inside the vessel to obtain a sequence of intravascular images. Meanwhile, ECG-triggered fluoroscopic images are acquired to track the movement of the IVUS transducer. Finally, in the mapping stage, the angiography and IVUS images are co-registered based on tracking and the temporal synchronization between the two imaging modalities. The workflow is summarized in Fig. 3.

A.   Angiogram Stage

In the angiogram stage, contrast agent is injected to visualize the vessel branch under investigation. A frame at the end-diastole (ED) phase is selected as the "angiogram reference frame." In the angiogram reference frame, the vessel branch where the IVUS imaging will be performed should be enhanced by the injected contrast agent. Fig. 4(a) shows a selected angiogram reference frame in the angiogram stage. An interactive method

[6], [22] is applied to detect the vessel branch. Our method further automatically extends the detected vessel toward the guiding catheter to provide more information for the subsequent tracking. Details about the vessel detection algorithm are provided in Section III. The detected branch will be used as the reference path for the co-registration of angiography and IVUS images, and it will also be used in the initialization of device tracking.

B.   Pullback Stage

In the pullback stage, an IVUS transducer is pulled back from the distal end of the target vessel branch toward the proximal end. During the pullback, both the ECG-triggered fluoroscopic and IVUS images are acquired with the time synchronized between the two modalities. Examples of fluoroscopy and IVUS images are shown in Fig. 4(b) and (c). The fluoroscopy image acquisition is triggered by ECG at the ED cardiac phase to minimize the cardiac motion between frames and to reduce radiation dose. Due to the existence of breathing motion, an additional point other than the IVUS transducer is needed as the reference point to compensate for the breathing motion. The guiding catheter tip is selected as the reference point, as it is the most stable and distinguishable point that can be found in the fluoroscopy images. In the pullback stage, the method automatically tracks the movements of IVUS transducers and guiding catheter tips in fluoroscopy.

C.   Mapping Stage

In the mapping stage, the geodesic distances (i.e., the distance along the wire) between IVUS transducers and guiding catheter tips are estimated. The geodesic distances are used to map each IVUS transducer during pullback, therefore registering the corresponding IVUS imaging plane to a point along the segmented vessel branch. Finally the IVUS images are co-registered to the segmented vessel by matching the acquisition time ("time stamps") of corresponding IVUS images and ECG-triggered fluoroscopy images.

III.   INITIALIZATION IN ANGIOGRAM

At the angiogram stage, the vessel branch where the IVUS pullback occurs is identified by a semi-automatic and interactive detection method. Based on our previous work [6], [22],
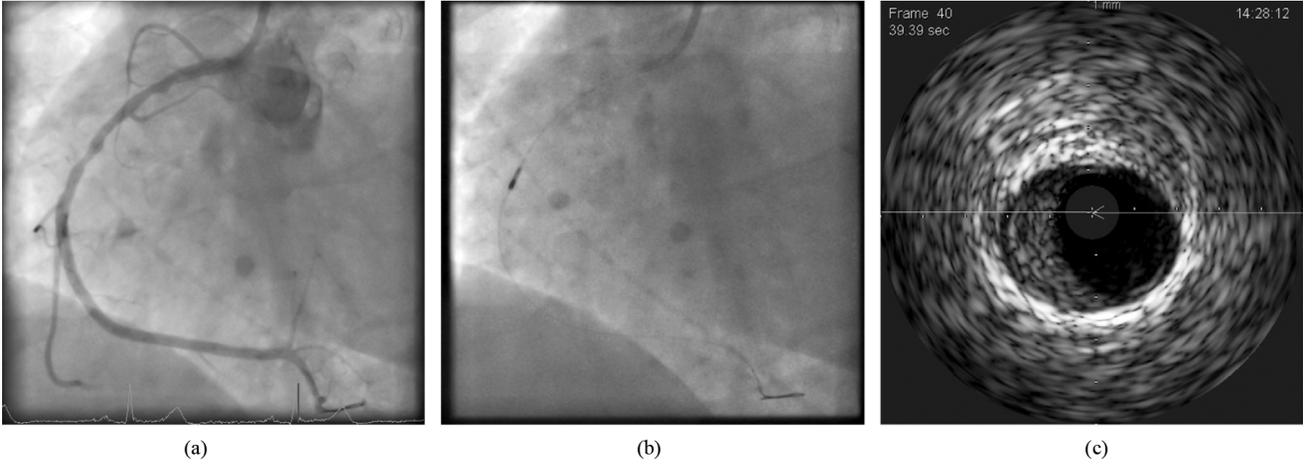
Fig. 4. Image acquisition in the workflow. (a) Angiogram image is selected as the reference frame because of its ECG phase and contrast enhancement. (b) Frame of fluoroscopy acquired in the pullback stage. (c) IVUS image that is acquired at the same time as (b) in the pullback stage.
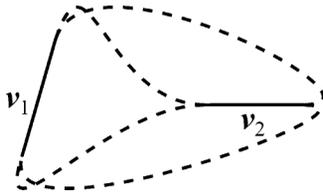
Fig. 5. Two guiding catheter centerline segment candidates and the four possible spline curve connections between them.

the interactive detection method applies learning-based detection and graph-based optimization, to detect a vessel branch. The previous methods [6], [22] are developed mainly to detect guidewires, which are thinner than vessels. Our method down-sized the angiography images to allow the detection of vessel branches.

The interactive detection method requires at least two clicks. The first user click is to specify the proximal end of the vessel, so it is placed at the tip of the guiding catheter into which the IVUS catheter is inserted. A second user click is placed at the distal end of the vessel. The two user clicks should cover the path of IVUS catheter pullback. After obtaining the user inputs and vessel segment detection results, the detector then finds an optimal path, which can best fit the vessel between the source and destination points that are specified by the first two user clicks. In case the vessel detection with two initial points is not satisfactory, additional user clicks can be provided to constrain the algorithm and to obtain refined detection results. The user interaction continues until a satisfactory result is obtained. As shown in [22], more than 82% of points along curves are correctly located with only two clicks. With more clicks, the error decreases quickly. The median error is about 1.0 pixel with five clicks. Fig. 6 shows an interactive detection result.

After the vessel branch is initialized, the guiding catheter's centerline is automatically detected in the angiogram reference frame. It serves as an additional anchor for subsequent transducer tracking initialization. The guiding catheter detection starts from the proximal point of the vessel, denoted as $\mathbf{p}_0$, and stops at an image boundary or collimator boundary. A machine learning-based technique is applied to identify possible
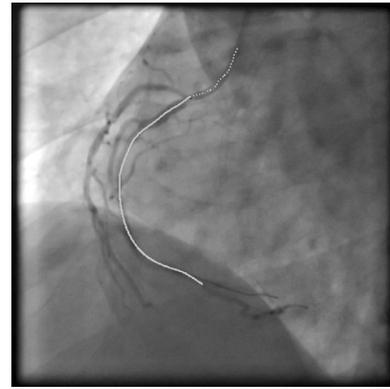
Fig. 6. An example of interactive detection in the angiogram stage. The solid white curve shows detected vessels, and the dotted curve shows detected guiding catheter bodies.

guiding catheter centerline segment candidates. They are represented as triples $(\boldsymbol{p}, \boldsymbol{n}, s)$ where $\boldsymbol{p}$, $\boldsymbol{n}$, and s are the position, orientation and scale of a segment candidate, respectively. A discriminative model, in our case a probabilistic boosting-tree (PBT) [15] using 2-D Haar-like features [16] on rotated input images is used. Segment candidate detection is carried out on a multi-resolution pyramid to get candidates of different scales.

Let $G = (\mathcal{V}, \mathcal{E})$ be an undirected weighted graph, where $\mathcal{V}$ is a set of total N vertices and $\mathcal{E}$ is a set of edges. The weights of edges are defined by a function $c : \mathcal{E} \, ura \, \mathbb{R}$. The vertices include subsets of vertices $\mathcal{V}^{(-)}$ and $\mathcal{V}^{(+)}$, as

$$\mathcal{V} = \{\, \boldsymbol{v}_i^{(-)} = \boldsymbol{p}_i - \frac{1}{2} s_i \boldsymbol{n}_i | i = 1, \ldots, N \,\}$$
$$\cup \{\, \boldsymbol{v}_i^{(+)} = \boldsymbol{p}_i + \frac{1}{2} s_i \boldsymbol{n}_i | i = 1, \ldots, N \,\}. \quad (1)$$

There can be four connections defining a spline curve between two centerline segment candidates v_i and v_j: $(\boldsymbol{p}_i, \boldsymbol{v}_i^{(k)}, \boldsymbol{v}_j^{(l)}, \boldsymbol{p}_j)$, $k, l \in \{-, +\}$ (see Fig. 5). For any edge $e \in \mathcal{E}$ let C_e denote the associated spline curve and let $l(C_e)$ be its curve length. Let $\mathcal{B}_i^{(k)}$ be the set of edges adjacent to $\boldsymbol{v}_i^{(k)}$, $i = 1, \ldots, N$, $k \in \{-, +\}$, without $\{\boldsymbol{v}_i^{(-)}, \boldsymbol{v}_i^{(+)}\}$. Among the nearest neighboring segments in terms of the length of the
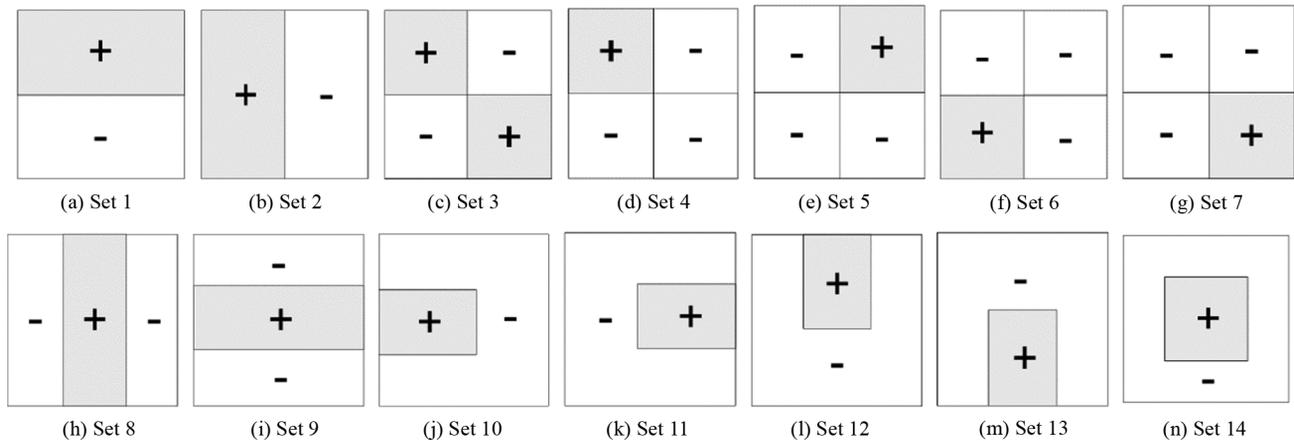
Fig. 7. Haar feature sets used for device detection and tracking.

connecting spline curve we allow a maximum of $M = 12$ adjacent edges approaching a segment candidate from the negative or positive side. This yields the set of edges

$$
\begin{aligned}
\mathcal{E} = &\{\{\boldsymbol{v}_i^{(-)}, \boldsymbol{v}_i^{(+)}\}|i = 1, \ldots, N\} \cup \\
&\{e = \{\boldsymbol{v}_i^{(k)}, \boldsymbol{v}_j^{(l)}\}|i, j = 1, \ldots, N, k, l \in \{-, +\}, \\
&i \neq j, l(C_e) \leq \max\{l(C_f)|f \in \mathcal{B}_i^{(k)}\}, |\mathcal{B}_i^{(k)}| \leq M\}.
\end{aligned}
$$

$$(2)$$

For each $e \in \mathcal{E}$, we define the edge weights $c(e) = \sum_{i=1}^{n-1} |\boldsymbol{I}(c_i) - \boldsymbol{I}(c_{i+1})|$ where $\boldsymbol{I}$ is the image intensity and $(c_1 \ldots c_n)$ is a sequence of discretized pixel positions covered by the curve. By using Dijkstra's algorithm the tree of shortest paths spanning G from $\boldsymbol{p}_0$ can be computed. It serves to generate path candidates connecting $\boldsymbol{p}_0$ with points on the image or collimator boundary. Among them, during path selection, the shortest one in terms of the accumulated edge weights is selected to be the final detection result, as the dotted curves in Fig. 6.

IV. LEARNING-BASED DEVICE DETECTION IN FLUOROSCOPY

In the pullback stage, our method automatically detects the IVUS transducer and the guiding catheter tip in fluoroscopic images. For the purpose of robust tracking, detections of other devices, such as guidewire and guiding catheter body, are also needed in an integrated model-based tracking method (see Section V). We present a unified method to train different types of device detectors from a set of offline collected training data. The training data includes both object samples (positive) and nonobject samples (negative), to learn the decision boundary that separates different classes. The probabilistic boosting tree (PBT) [15] is used to train the learning-based detectors. Here, we denote the probabilistic outputs from PBT as $P(\mathbf{Z}|\mathbf{x})$, where $\mathbf{Z}$ is the observed image, and $\mathbf{x}$ is the object state (i.e., if an image patch belongs to the positive or negative class.) The classifiers are built on the Haar features [17] and our extensions. The detectors are constructed in a hierarchical way to detect object position, orientation, and scale sequentially.

A. Features for 2-D Device Detection

Haar features have been widely used in many types of object detection due to its computational efficiency. We use the original Haar features [17], and further extend conventional Haar features specifically for medical device detections. The Haar feature configurations are shown in Fig. 7.

The feature Sets 1, 2, 3, 8, 9 have been originally used for face detection in [16]. The feature values are the quantitative results of image intensity operations. In each feature configuration, the Haar feature value is the pixel intensity inside the gray boxes minus the pixel intensity outside the boxes. During the Haar feature computation, the image intensity is normalized by subtracting the mean and dividing by the standard deviation of a local window so that features are invariant to linear transformations of image intensity. The image intensity summation and standard deviation can be efficiently computed with the use of integral images [17]. The feature template size is varied and is automatically selected in the feature selection during the training stage. The configuration of Haar features is further extended in the feature Sets 4, 5, 6, 7, 10, 11, 12, 13, 14. The motivation behind the feature configurations is that these features are able to represent the geometries of medical devices such as corners (e.g., Sets 4–7) and tips (e.g., Sets 10–13).

B. Device Detection

Medical devices can appear at different orientations in fluoroscopy images. In this method, images are rotated, and then features are computed in each rotation angle for device detection at the corresponding orientation. The image rotation has been sped up with the use of multi-core CPUs. Furthermore, since usually the region-of-interests-based (ROI-based) operation is used in tracking, only a part of an image needs to be rotated, therefore further improving the speed.

The medical device can appear at different sizes in fluoroscopic images. The size in pixels of the devices in fluoroscopic images is affected by the X-ray detector resolution, the distance between a medical device and an imaging plane, and the imaging angle. To detect devices of different sizes, the Haar features are resized by scaling their corresponding corner coordinates.
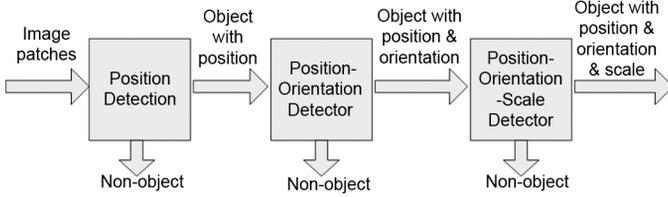
Fig. 8. Hierarchical detector structure.

The detectors are constructed in a hierarchical way, as shown in Fig. 8. In the hierarchy, object states are detected from coarse to fine to reduce the burden of learning a complete set of object states altogether. A position detector is trained as the first layer in the hierarchy to locate the position of objects, regardless of their orientations and scales. An orientation detector is then trained at images rotated at different angles to detect devices at arbitrary orientations. Lastly, a scale detector is trained to search across different scales by varying the size of Haar features.

Four types of detectors are trained. They are denoted as $P_{\text{IVUS\_trans}}$, $P_{\text{cath\_body}}$, $P_{\text{cath\_tip}}$, and P_{wire}, for the IVUS transducer, the guiding catheter body, the guiding catheter tip, and the wire body, respectively. Please note that the guiding catheter body and wire detectors aim at the detection of segments, not the whole structure. Some detection results are shown in Fig. 9. Due to image artifacts and low visibility of devices, there are false detections. To make the tracking robust to detection errors, a tracking framework that integrates multiple detections is introduced in Section V.

V. MODEL-BASED DEVICE TRACKING

In this section, we first introduce a Bayesian framework for device tracking, the fusion of measurements of multiple devices in an integrated model, and then an efficient implementation of the tracking framework.

A. A Bayesian Tracking Framework

In the pullback stage, the fluoroscopic images are acquired by ECG triggering so that all fluoroscopy images are approximately at the same heart phase, and the devices undergo only breathing motion. The breathing motion needs to be compensated for through tracking in order to register the IVUS transducer to the angiogram reference frame. In this work, the breathing motion is approximated by a rigid motion. We found this to be sufficient in tracking devices including both IVUS transducers and guiding catheter tips. The breathing motion at the tth frame is denoted as $\mathbf{M}_t = (m_t^x, m_t^y, m_t^r)$, where m_t^x, m_t^y, and m_t^r are the 2-D translation and rotation parameters. The motion tracking is expressed as the inference of the motion parameters from the fluoroscopic sequences acquired at the pullback stage. We formalize the parameter inference in a sequential Bayesian inference framework. The posterior probability $P(\mathbf{M}_t = (m_t^x, m_t^y, m_t^r)|\mathbf{Z}_t)$ is given by

$$P(\mathbf{M}_t|\mathbf{Z}_t) \propto P(\mathbf{M}_t)P(\mathbf{Z}_t|\mathbf{M}_t) \quad (3)$$

where $\mathbf{Z}_t$ is an observed fluoroscopic sequence at time t. The tracking result is the motion parameter corresponding to the maximal posterior probability, i.e., $\hat{\mathbf{M}}_t = \arg\max_{\mathbf{M}_t} P(\mathbf{M}_t|\mathbf{Z}_t)$.

In (3), $P(\mathbf{M}_t)$ is the motion prior probability. We model the motion prior probability as $P(\mathbf{M}_t) = G(\mathbf{M}_t; \Sigma_{\mathbf{M}})$ where $G(\mathbf{M}_t; \Sigma_{\mathbf{M}})$ is a Gaussian model with a zero mean and the covariance matrix $\Sigma_{\mathbf{M}}$. The Gaussian prior model is used to impose a generic motion smoothness constraint, and is applicable to 2-D motions under a variety of projections. The model parameter is a covariance matrix, which is defined as a diagonal matrix, $\Sigma_{\mathbf{M}} = diag\{\sigma_t^2, \sigma_t^2, \sigma_r^2\}$. To allow for a large motion between successive fluoroscopic frames, the two parameters σ_t and σ_r are empirically set as 60 pixels for translation and 4 degrees for rotation, respectively.

Another component in (3), the likelihood model $P(\mathbf{Z}_t|\mathbf{M}_t)$, measures the likelihood of motion parameters. In our method, the measurement model is a fusion of measurements of multiple devices. More details are provided in Section V-B.

B. Integrated Model and Fusion of Individual Device Measurements

Due to the low image quality of fluoroscopy and cluttered backgrounds, independently tracking each device is prone to detection errors. To achieve robust and accurate tracking, we combine all devices into an "integrated IVUS model." The integrated model to be tracked is initialized from the angiogram phase. As shown in Fig. 10, the integrated model includes four parts: the guiding catheter body, the guiding catheter tip, the wire, and the IVUS transducer. The vessel centerline initialized during the angiogram stage is used to approximate the shape of the wire. The guiding catheter body and tip are also initialized from the angiogram stage. The IVUS transducer is initialized from the detection at the first frame in fluoroscopy. During tracking, the model is propagated from a previous frame and is updated from the motion parameter estimation. The benefit of tracking the integrated model as a whole, instead of tracking individual devices, is that it is robust to false detections and missing detections of individual devices, and can handle some challenging situations where individual tracking would fail, such as poor visibility, occlusion, and large motions.

We denote the integrated model as Γ_t, in which each component is represented by a curve (e.g., a guiding catheter body or a wire) or a point (e.g., guiding catheter tip or an IVUS transducer), as shown in Fig. 10. By defining the integrated model, we can rewrite the likelihood in the form of the curve representation, as $P(\mathbf{Z}_t|\mathbf{M}_t) = P(\mathbf{Z}_t|\Gamma_t)$. Based on the integrated IVUS model, the measurement model is a combination of measurements of individual components. For simplicity, we assume that the measurement of individual component is independent of each other given the curve (i.e., $P(\mathbf{Z}_t|\Gamma_t^k, \Gamma_t) = P(\mathbf{Z}_t|\Gamma_t^k)$ where Γ_t^k denotes a component in the integrated model.) Therefore, we can further decompose the measurement model $P(\mathbf{Z}_t|\Gamma_t^k)$ as

$$P(\mathbf{Z}_t|\mathbf{M}_t) = P(\mathbf{Z}_t|\Gamma_t) = \sum_k P(\mathbf{Z}_t|\Gamma_t^k, \Gamma_t)P(\Gamma_t^k|\Gamma_t)$$

$$= \sum_k P(\mathbf{Z}_t|\Gamma_t^k)P(\Gamma_t^k|\Gamma_t). \quad (4)$$

The component measurements $P(\mathbf{Z}_t|\Gamma_t^k)$ are from the four detectors, i.e., $P_{\text{IVUS\_trans}}$, $P_{\text{cath\_body}}$, $P_{\text{cath\_tip}}$ and P_{wire}.
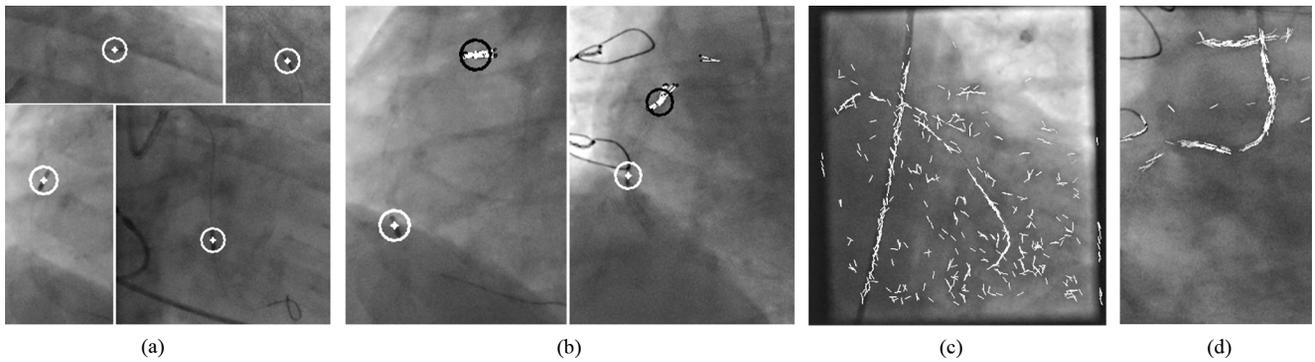
Fig. 9. Learning-based detection examples (see Section VII about the data acquisition). (a) Circles represent detected IVUS transducers. (b) Circles represent detected guiding catheter tips. Multiple detection results that are close to each other are clustered into a single detection result. The points and lines inside the circles represent the orientations of detected guiding catheter tips. (c) Detected wire segment candidate. (d) Detected guiding catheter body segment candidates.
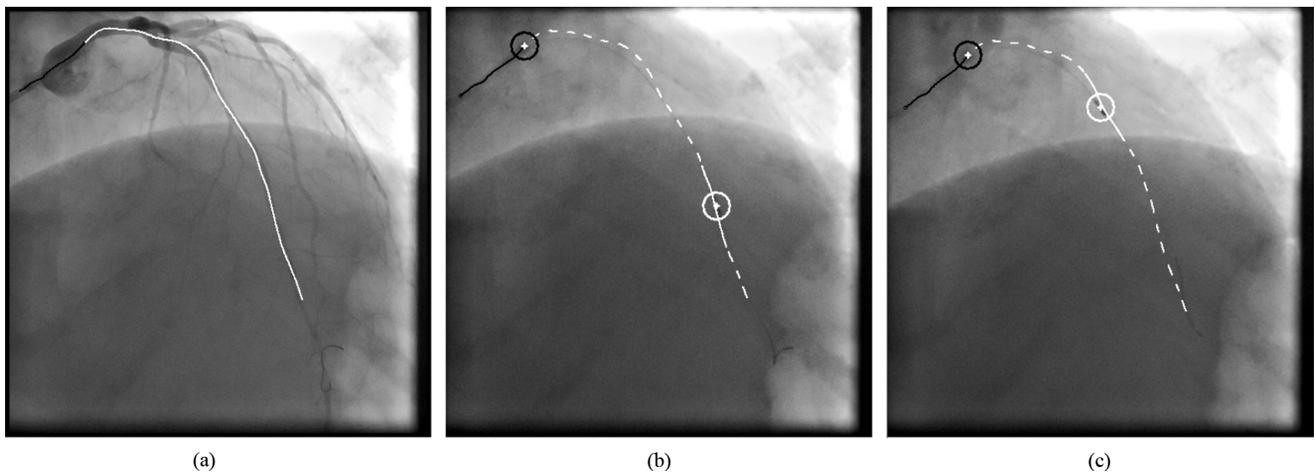
Fig. 10. Model-based tracking. (a) Detection results at an angiogram reference frame will be used in the subsequent model-based tracking. (b), (c) Integrated model, composed of point and curve elements, is initialized from interactive detection, and tracked through fluoroscopy.
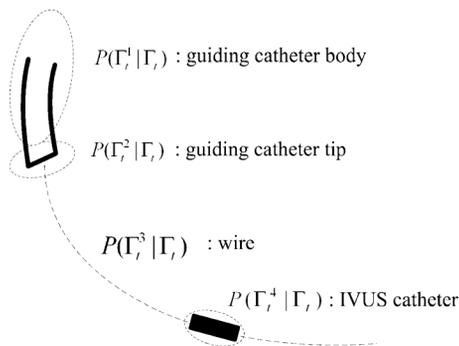
Fig. 11. Measurement fusion in the integrated IVUS model. Dotted curves illustrate the region where each individual measurement is integrated.

$P(\Gamma_t^k|\Gamma_t)$ defines the weight of an individual component in the model and denotes the confidence of individual measurements, as shown in Fig. 11.

C. Efficient Model Tracking

Exhaustively searching the motion parameters to maximize the posterior probability $P(\mathbf{M}_t|\mathbf{Z}_t)$ is computationally expensive, especially when the breathing motion is large. For efficient tracking, we use a kernel-based multi-resolution tracking method, which has been developed originally for the guidewire

tracking [21] and needle tracking [23]. In the multi-resolution tracking, measurements can be computed by a kernel-based estimation from a set of samples. For learning-based measurements, the samples are those points classified as a type of devices in the integrated IVUS model. The kernel-based measurement estimation is represented as

$$P(\mathbf{Z}_t|\Gamma_t^k) \propto \sum_{\mathbf{x}_j \in \Gamma_t^k} P(\mathbf{Z}_t|\mathbf{x}_j^s)G(|\mathbf{x}_j^s - \mathbf{x}_i|;\sigma) \qquad (5)$$

where $P(\mathbf{x}_j^s|\mathbf{x}_i) = G(|\mathbf{x}_j^s - \mathbf{x}_i|;\sigma)$ is a Gaussian kernel with a bandwidth σ. For more details of the hierarchical model tracking, please refer to [21], [23].

The kernel-based measurement estimation can obtain smooth measurements in a neighborhood, and also allow for multi-resolution searching during tracking. An example of multi-resolution searching of translation is shown in Fig. 12. The tracking is performed with decreasing search intervals $\{d_1 > d_2 > \cdots\}$. The corresponding bandwidth in (5) varies accordingly, denoted as σ_i at the ith resolution. At coarse resolution, we use larger kernel bandwidths to avoid missing tracking due to larger searching intervals; and at fine resolution, we use smaller kernel bandwidths to obtain finer tracking results. The Gaussian bandwidths are set as the searching intervals, i.e., $\sigma_i = d_i, i = 1, \dots, T$, to adapt the multi-resolution
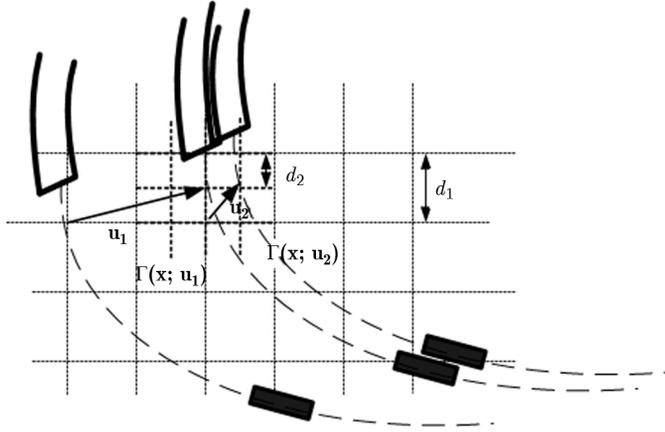
Fig. 12. Multi-resolution tracking with incrementally decreasing searching intervals and kernel bandwidths. The tracking starts from an initial position, and ends at the position close to the true shape. The dotted curves represent intermediate tracking results at different resolutions. u_i is the translation vector inferred from the ith level. $\Gamma(x, u_i)$ is the integrated model updated after the translation from the ith level in the hierarchical tracking.

Fig. 13. Inferring the geodesic distance from Euclidean distance and vessel shape. (a) Geodesic and Euclidean distances can be computed for each point at the vessel. (b) Relationship between geodesic and Euclidean distance is established for each vessel shape, and is used to infer geodesic distances from Euclidean distances in tracking.

tracking. The same strategy is applied to the joint search of translation and rotation. To avoid local optima during tracking, multiple candidates are saved at each resolution and are propagated to the next resolution. The current tracking method runs at 2.0 frames per second in a computer with a Core 2 Duo 2.66-GHz CPU. In a typical case where the image size is 512×512, the image rotation takes about 200 ms, the device detection takes about 150 ms, and the tracking takes about 150 ms. Since the fluoroscopy is acquired by ECG triggering (approximately at the rate of 1 frame per second,) our method is fast enough for online *in vivo* clinical procedures.

VI. CO-REGISTRATION

In the mapping stage, once IVUS transducers and guiding catheter tips are tracked, angiography and IVUS images are co-registered. The mapping stage includes two steps: the spatial mapping, which is to map each tracked transducer in fluoroscopy to a location on the vessel branch in the angiogram reference frame, and the temporal mapping, which is to map a location in the angiography reference frame to an IVUS slice based on the temporal synchronization between the two image modalities. In the spatial mapping step, a geodesic distance $D_g(t)$ is inferred from the corresponding Euclidean distance $D_e(t)$ between a tracked guiding catheter tip and a tracked IVUS catheter at the time t. The relationship $D_g(t) = f(D_e(t))$ is computed from the vessel shape defined in the angiogram reference frame. Fig. 13 shows an example of obtaining $D_g(t) = f(D_e(t))$ from a vessel shape. For some vessel shapes, the function $f(D_e(t))$ may not be a single-valued function, which means that for each computed Euclidean distance, there could be multiple corresponding geodesic distances. To resolve the ambiguity, a smoothness constraint is imposed on the geodesic distance. The geodesic distance should change smoothly between successive frames. If there are multiple geodesic distances that correspond to an Euclidean distance at time t, the kth geodesic distance candidate is denoted as

$D_g^k(t)$. Its weight $w_k(t)$ is voted from all the candidates in a neighborhood, $N(t) = \{D_g^{k'}(t'), \text{ for } t' \neq t, |t' - t| < \epsilon\}$

$$w_k(t) = \sum_{D_g^{k'}(t') \in N(t)} G(D_g^k(t) - D_g^{k'}(t'); \sigma_g)G(t - t'; \epsilon) \quad (6)$$

where σ_g and ϵ are the parameters in Gaussian kernels to smooth votes from the neighborhood. From all possible candidates in each frame, the geodesic distance with the maximal weight is selected, i.e., $D_g(t) = D_g^{\hat{k}(t)}(t)$, where $\hat{k}(t) = \arg\max_k w_k(t)$.

Based on the computed geodesic distance, a pullback model is fitted to estimate the motion of IVUS transducers in a vessel branch. The model fitting serves two purposes: first, we notice that the motion based on geodesic distances without smoothing is noisy, as shown in Fig. 14. Such noise is due to breathing motions, imperfect ECG triggering, possible vessel foreshortening in 2-D images, and/or possible catheter movements inside vessel. A smooth model fitted on estimated geodesic distances can help reduce the impact of such factors on the registration; second, the model fitting can identify false detections and tracking errors as outliers of the fitted model. Such errors can then be removed from motion estimations. The outlier identification and model fitting is iterated, therefore improving the registration accuracy. To handle nonconstant movement of IVUS pullback, we fit a cubic spline to model the geodesic distance during pullback. Assuming that $D_f(t) = D_f(t; d_1, d_2, \ldots, d_k)$ is a spline model defined on the anchor points $d_2, \ldots, d_k$, the model is to minimize the fitting cost C_f

$$C_f = \|\mathbf{D}_g - \mathbf{D}_f\|_2 + \alpha\|\mathbf{D}_f'\|_2 + \beta\|\mathbf{D}_f''\|_2. \quad (7)$$

In (7), $\|\mathbf{D}_g - \mathbf{D}_f\|_2$ is the fitting error, the second and third term are the first-order and the second-order smoothness constraints. α and β are used to balance the two smoothness constraints. The optimization of C_f is obtained by a greedy search. At the beginning, the anchor points are uniformly sampled from the original data $D_g(t)$. During the greedy search, the anchor points are sequentially searched in a local range to minimize the cost function. The iteration continues until the cost function converges. An example is shown in Fig. 14, where there are noises in the
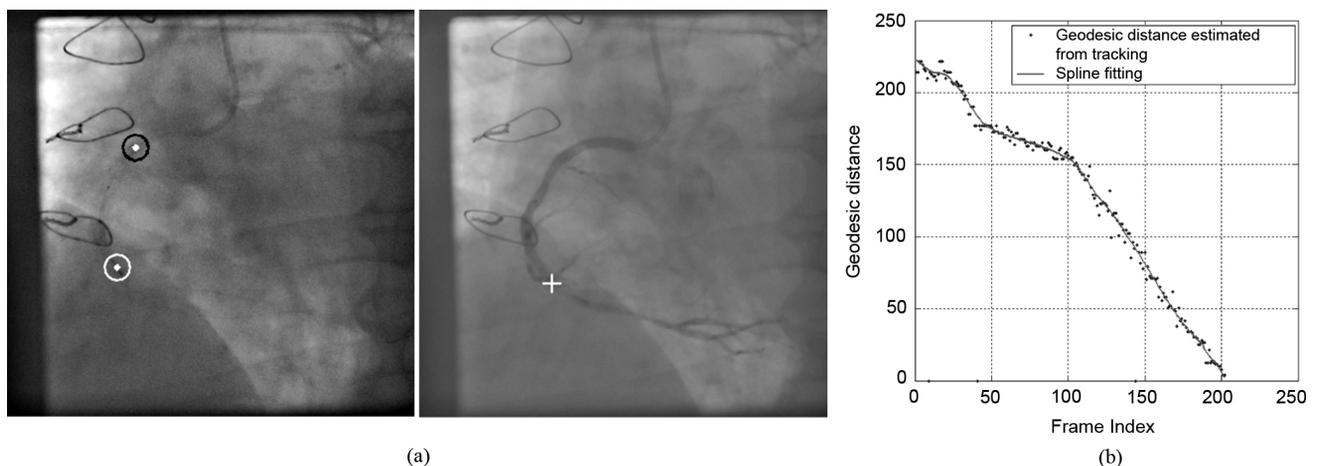
Fig. 14. Fitting the transducer movement with a spline. (a) Tracked transducer in a pullback frame and its spatial mapping in an angiogram reference frame. (b) Geodesic distances estimated from the tracking results and from the fitting.

geodesic distance estimation, and the pullback speed is not constant inside the vessel. The model fitting can remove the noise and model the nonconstant movement of the transducers in the 2-D image plane.

For the purpose of time synchronization, the clocks of the angiography and IVUS systems are manually adjusted. Interventional cardiologists simultaneously start the acquisition of IVUS and ECG-triggered fluoroscopy. The acquisition times of the ECG-triggered fluoroscopy frames and of the gated IVUS images are stored as a frame time vector in the DICOM header of the two sequences. A finer adjustment was achieved by manually looking for the correspondence between the frame time vectors of each respective modality. Finally, based on the estimated geodesic distance, each tracked IVUS catheter is mapped onto a point on the segmented vessel branch. The location is associated with an IVUS slice based on the time synchronization, as explained above. By the spatial and temporal mapping, each IVUS frame is "co-registered" to its corresponding position in the reference angiography image.

VII. EXPERIMENTS

A. Data

The angiography and IVUS co-registration method and system are evaluated on a set of clinical cases acquired from several European hospitals. The acquisition follows the workflow described in Section II. The evaluation set includes 65 cases in total. Each acquired case includes an angiogram sequence and an ECG-triggered fluoroscopic pullback sequence (Artis Zee, Siemens AG, Forchheim, Germany), and corresponding IVUS images (Volcano Corp., San Diego, CA, USA). Although in these experiments, the phased array IVUS catheters are used, the developed method can be directly applied to other types of IVUS catheters, e.g., the mechanical catheters. The X-ray image size is 512×512, and the physical size of each pixel is between 0.2 and 0.3 mm. Each pullback was acquired within a few minutes, usually containing 100 to 300 ECG-triggered fluoroscopy frames. The data set reflects a variety of clinical scenes, including low signal-to-noise ratio,

IVUS pullback in different vessel branches and projection angles, and the presence of other surgical devices in the field of view. Some examples are shown in Fig. 15.

B. Quantitative Evaluations of Device Tracking

To establish ground truth for quantitative evaluation, we manually annotated the positions of IVUS transducers and guiding catheter tips in all the fluoroscopic frames as the ground truth. The tracking accuracy in fluoroscopy is quantitatively measured as the Euclidean distances between automatic tracking results and corresponding manual annotations. The geodesic distance error is also computed as the estimation of the co-registration error. For this purpose, we compute the geodesic distances from the manual annotations as the ground truth, and then compare the ground truth with the geodesic distances computed from corresponding tracking results.

The quantitative validation of the system is performed with two experimental settings: Experiment I and Experiment II. In Experiment I, the detectors are trained with and applied to all the cases. Our method is successful in 64 out of 65 cases, except for one case where the IVUS transducer is occluded by the spine for almost the whole time, thus achieving a 98.46% success rate. In Experiment II, the training and testing sets are separated to test how well our method generalizes to unseen data. We train the detectors in 32 randomly selected cases, and validate the framework in the remaining 32 cases that are excluded from training. Some exemplar tracking results are shown in Fig. 15. Table I summarizes the error statistics from the 64 cases. The millimeters errors are converted from pixel errors based on the physical resolution of the imaging detector, therefore quantifying normalized errors in a 2-D projection space. The tracking error of IVUS transducers is small, with mean errors only around 0.35 mm for both experiments. The error at the guiding catheter tips is slightly larger. The median error is 0.97 mm for Experiment I, and 1.17 mm for Experiment II. The overall small differences between the two experiments demonstrate the good generalization capability of our tracking method. The mean geodesic errors are below 1.20 mm and 1.50 mm from the two validation experiments, respectively.
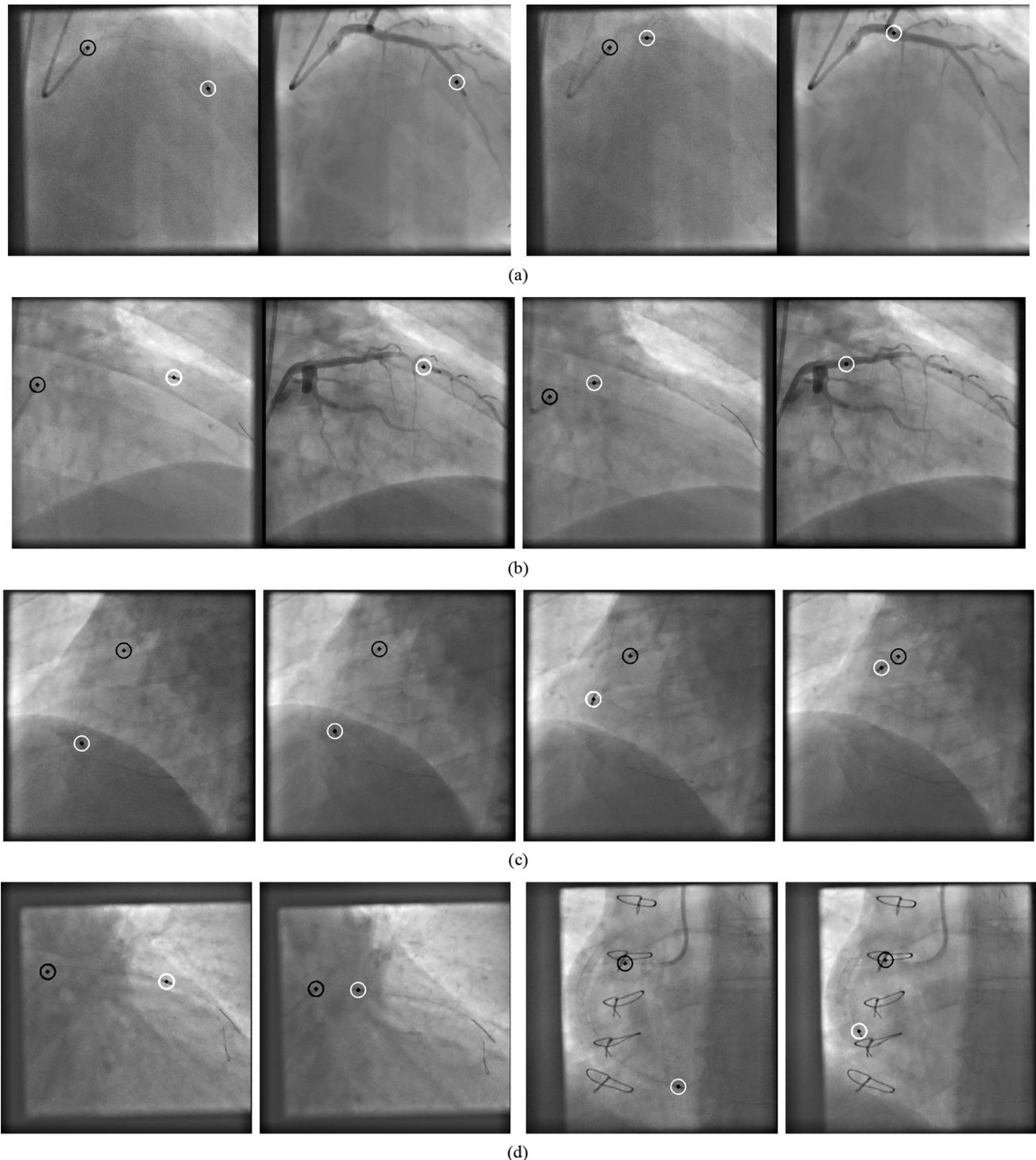
Fig. 15. (a), (b) Tracking and co-registration results. In each frame, the left image shows tracking IVUS transducers and guiding catheter tips, and the right image shows the corresponding mapped IVUS image plane (the red cross) along the vessel in the reference angiogram frame. (c) Tracked IVUS transducers and guiding catheter tips in fluoroscopy in one case. (d) More tracking results in another two cases.

C. Evaluation of Co-Registration With Clinical Ground Truth

Besides evaluating the tracking accuracy, the accuracy is evaluated by comparing the correspondences on the co-registered angiography and IVUS images. For clinical evaluation, a prototype as shown in Fig. 16 was developed. This prototype simultaneously shows the reference angiography image, the IVUS frames and the longitudinal IVUS representation (across the IVUS stack of cross-sections). The position associated with the current IVUS frame is shown by a cross in the reference angiogram and by a line in the IVUS longitudinal display. The reviewing cardiologist first identified a salient landmark such as a bifurcation or a stent termination in the IVUS frames and reported the time corresponding to this frame.

TABLE I
QUANTITATIVE EVALUATION OF THE CO-REGISTRATION SYSTEM

Validation	Experiment I			Experiment II		
Error (in mm)	median	mean	std	median	mean	std
IVUS transducer	0.26	0.34	0.56	0.28	0.35	0.32
Guiding catheter tip	0.97	1.57	1.82	1.17	1.92	2.09
Geodesic distance	0.73	1.19	2.14	1.03	1.41	1.51

TABLE II
COMPARE THE CO-REGISTRATION WITH CLINICAL GROUND TRUTH

	Median	Mean	70 percentile	80 percentile	90 percentile
Error (in mm)	1.08	1.15	1.34	1.52	2.10
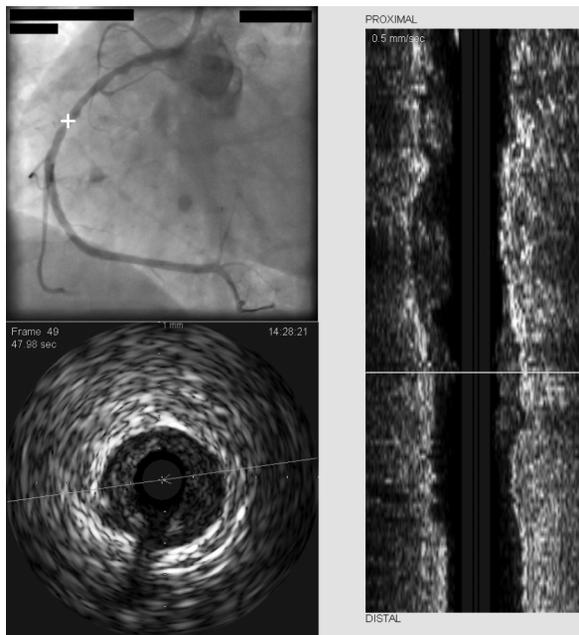
Fig. 16. Frame of co-registered angiography and IVUS images. The left top image shows the registered IVUS image plane on the angiography (i.e., the cross in the upper image). The left bottom image shows the co-registered IVUS image at the cross-sectional view. The right image is the corresponding IVUS images at a longitudinal view.

The cross was then moved to the corresponding landmark in the reference angiography image and the time reported. The co-registration accuracy was calculated as the magnitude of the difference between the two reported times multiplied by the pullback speed. In the clinical evaluation, more factors, such as imperfect time synchronization between angiography and IVUS of the prototype system, affect the final co-registration accuracy.

From all the cases, clinicians identify 60 cases that are suitable for evaluation. The cases where there are no clear landmarks, or image quality is not good enough are excluded from the clinical evaluation. One to three landmarks are manually identified for each case. There are totally 99 landmarks. The comparison between the clinical ground truth and the automatic co-registration is summarized in Table II, where the mean, median, as well as different percentiles are presented. The experiment setting is as the Experiment I in Section VII-B. The results show that although all the affecting factors are taken into account, the mean co-registration error is still small, i.e., only about 1.15 mm. About 90% of the landmarks have less than 2.1 mm co-registration error. The results are obtained from the data acquired from multiple hospitals, and also include all the error sources that can be introduced in a practical setting.

D. Considerations on Radiation Exposure

The reference angiogram can be selected from any already acquired contrasted sequence used for diagnostic purposes and does not require additional radiation exposure. The average number of frames acquired by ECG-triggered fluoroscopy during the 60 pullbacks of the clinical evaluation was 149. It corresponds to 10 seconds of additional fluoroscopy time, assuming an acquisition of 15 frames per second. It is thus very low compared to the overall dose exposure of an intervention.

VIII. CONCLUSION

We report a solution to provide automatic and robust tracking and registration of angiography and IVUS images. The experiments demonstrate its effectiveness in co-registering the two imaging modalities. We hope that this work will establish a useful reference for further research in this exciting field. For example, the co-registration between angiography and IVUS can be a pivotal tool in assessing implanted bioresorbable scaffolds as they are barely visible in angiography.

The presented method is applicable to other intravascular imaging modalities that have been enabled by new technical developments in the recent years. As an example, optical coherence tomography (OCT) has gained much attention, especially due to its high spatial resolution. Other modalities such as near-infrared spectroscopic (NIRS) imaging, intravascular magnetic spectroscopy, intravascular magnetic resonance imaging, Raman spectroscopy, intravascular photoacoustic imaging, near infrared fluorescence imaging, and time resolved fluorescence spectroscopic imaging have contributed to increase knowledge of coronary atherosclerosis in a research setting. The integration of intravascular imaging with other invasive or noninvasive modalities has shown to give a more detailed understanding of coronary artery pathology [1].

The method developed in this paper has the potential to be adapted to other IVUS catheters (e.g., mechanical or other vendor) within a reasonable amount of efforts. For that purpose, a part of training, as described in Sections IV and V, would need to be repeated with fluoroscopic images of these specific catheters. While the tracking and detection of catheters from other modalities is conceivable after some adaptations, the co-registration workflow may need to be adapted for modalities such as OCT due to a significantly higher speed of the pullback.

IX. ACKNOWLEDGEMENT

REFERENCES

[1] C. V. Bourantas, H. M. Garcia-Garcia, K. K. Naka, A. Sakellarios, L. Athanasiou, D. I. Fotiadis, L. K. Michalis, and P. W. Serruys, "Hybrid intravascular imaging: Current applicat. and prospective potential in the study of coronary atherosclerosis," *J. Am. Coll. Cardiol.*, vol. 61, no. 13, pp. 1369–1378, 2013.

[2] N. Bruining, S. d. Winter, and P. W. Serruys, "Intravascular ultrasound registration/integration with coronary angiography," *Cardiol. Clin.*, vol. 27, no. 3, pp. 531–540, 2009.

[3] M. Jourdain, J. Meunier, J. Sequeira, J. M. Boi, and J. C. Tardif, "A robust 3-D IVUS transducer tracking using single-plane cineangiography," *Trans. Info. Tech. Biomed.*, vol. 12, no. 3, pp. 307–314, 2008.

[4] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1987.

[5] H. A. Marquering, J. Dijkstra, Q. J. A. Besnehard, J. P. M. Duthé, J. D. Schuijf, J. J. Bax, and J. H. C. Reiber, "Coronary CT angiography: IVUS image fusion for quantitative plaque and stenosis analyses," in *Proc. SPIE Med. Imag.*, 2008.

[6] P. Mazouer, T. Chen, Y. Zhu, P. Wang, P. Durlak, J.-P. Thiran, and D. Comaniciu, "User-constrained guidewire localization in fluoroscopy," in *Proc. SPIE Med. Imag.: Image Process.*, 2009.

[7] G. S. Mintz, S. E. Nissen, W. D. Anderson, S. R. Bailey, R. Erbel, P. J. Fitzgerald, F. J. Pinto, K. Rosenfield, R. J. Siegel, E. Tuzcu, P. G. Yock, R. A. O'Rourke, J. Abrams, E. R. Bates, B. R. Brodie, P. S. Douglas, G. Gregoratos, M. A. Hlatky, J. S. Hochman, S. Kaul, C. M. Tracy, D. D. Waters, and W. L. Winters, Jr., "American College of Cardiology clinical expert consensus document on standards for acquisition, measurement and reporting of intravascular ultrasound studies (IVUS). A report of the American College of Cardiology task force on clinical expert consensus documents developed in collaboration with the European Society of Cardiology endorsed by the Society of Cardiac Angiography and Interventions," *J. Amer. Coll. Cardiol.*, vol. 37, no. 5, pp. 1478–1492, 2001.

[8] A. Nair, B. D. Kuban, E. M. Tuzcu, P. Schoenhagen, S. E. Nissen, and D. G. Vince, "Coronary plaque classification with intravascular ultrasound radiofrequency data analysis," *Circulation*, vol. 106, pp. 2200–2206, 2002.

[9] O. Pauly, G. Unal, G. Slabaugh, S. Carlie, and T. Fang, "Semi-automatic matching of OCT and IVUS images for image fusion," in *Proc. SPIE Med. Imag.*, 2008.

[10] Z. Qian, G. Vazquez, and S. Voros, "Intermodal registration of CTA and IVUS-VH, and its application on CTA-based plaque composition analysis," in *Proc. IEEE Int. Symp. Biomed. Imag.: From Nano to Macro*, 2011, pp. 419–423.

[11] J. Rieber, S. Prummer, M. Schmidt, and H. Rittger, "A novel method for direct correlation of angiography and endovascular diagnostic tools in human coronary arteries in vivo," *J. Am. Coll. Cardiol.*, vol. 55, pp. A217.E2065–A217.E2065, 2010.

[12] G. Strommer, L. Schwartz, I. Shmarak, M. Y. Flugelman, A. Shiran, L. Sobe, E. Oren, R. Shofti, M. B. Leon, and B. S. Lewis, "Fluoroscopy-free navigation of guided-CStent using medical positioning system and 3-D guided-CIVUS (GIVUS?) image," *Int. Congr. Ser.*, pp. 387–392, 2005.

[13] S. Tu, N. R. Holm, G. Koning, Z. Huang, and J. H. C. Reiber, "Fusion of 3-D QCA and IVUS/OCT," in *Int. J. Cardiovasc. Imag.*, 2011.

[14] S. Tu, L. Xu, J. Ligthart, B. Xu, K. Witberg, Z. Sun, G. Koning, J. H. Reiber, and E. Regar, "In vivo comparison of arterial lumen dimensions assessed by co-registered three-dimensional (3-D) quantitative coronary angiography, intravascular ultrasound and optical coherence tomography," in *Int. J. Cardiovasc. Imag.*, 2012, vol. 28, no. 6, pp. 1315–1327.

[15] Z. Tu, "Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering," in *IEEE Int. Conf. Comput. Vis.*, 2005, pp. 1589–1596.

[16] P. Viola and M. Jones, "Robust real-time object detection," in *Int. Workshop Stat. Comput. Theories Vis.*, 2004, vol. 57, pp. 137–154.

[17] P. Viola and M. Jones, "Robust real-time object detection," *Int. J. Comput. Vis.*, vol. 57 s2, pp. 137–154, 2004.

[18] A. Wahle, G. Prause, C. V. Birgelen, R. Erbel, and M. Sonka, "Fusion of angiography and intravascular ultrasound in vivo: Establishing the absolute 3-D frame orientation," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 10, pp. 1176–1180, Oct. 1999.

[19] A. Wahle, G. Prause, S. DeJong, and M. Sonka, "Geometrically correct 3-D reconstruction of intravascular ultrasound images by fusion with biplane angiography-methods and validation," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 686–699, Aug. 1999.

[20] P. Wang, T. Chen, O. Ecabert, S. Prummer, M. Ostermeier, and D. Comaniciu, "Image-based device tracking for the co-registration of angiography and intravascular ultrasound images," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent.*, 2011, pp. 161–168.

[21] P. Wang, T. Chen, Y. Zhu, S. K. Zhou, and D. Comaniciu, "Robust guidewire tracking in fluoroscopy," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009.

[22] P. Wang, W. Liao, T. Chen, S. K. Zhou, and D. Comaniciu, "Graph based interactive detection of curve structures in 2-D fluoroscopy," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent.*, 2010, pp. 269–277.

[23] P. Wang, P. Marcus, T. Chen, and D. Comaniciu, "Using needle detection and tracking for motion compensation in abdominal interventions," in *Proc. IEEE Int. Symp. Biomed. Imag.: From Nano to Macro*, 2010, pp. 612–615.

[24] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, pp. 13–13, 2006.