

Robust Discriminative Wire Structure Modeling with Application to Stent Enhancement in Fluoroscopy

Xiaoguang Lu, Terrence Chen, Dorin Comaniciu
Image Analytics and Informatics, Siemens Corporate Research
Princeton, NJ 08540, USA

xiaoguang.lu, terrence.chen, dorin.comaniciu@siemens.com

Abstract

Learning-based methods have been widely used in detecting landmarks or anatomical structures in various medical imaging applications. The performance of discriminative learning techniques has been demonstrated superior to traditional low-level filtering in robustness and scalability. Nevertheless, some structures and patterns are more difficult to be defined by such methods and complicated and ad-hoc methods still need to be used, e.g. a non-rigid and highly deformable wire structure. In this paper, we propose a novel scheme to train classifiers to detect the markers and guide wire segment anchored by markers. The classifier utilizes the markers as the end point and parameterizes the wire in-between them. The probabilities of the markers and the wire are integrated in a Bayesian framework. As a result, both the marker and the wire detection are improved by such a unified approach. Promising results are demonstrated by quantitative evaluation on 263 fluoroscopic sequences with 12495 frames. Our training scheme can further be generalized to localize longer guidewire with higher degrees of parameterization.

1. Introduction

Stent thrombosis and restenosis are associated with stent under-expansion, which has been shown as a major risk factor for patients undergoing percutaneous coronary intervention (PCI). During the intervention, a stent is deployed via a balloon at the lesion spot inside the coronary artery. This procedure is monitored by X-ray fluoroscopy where the stent visibility is often low because the radiation dose is kept at a minimal level and stents are only slightly radiopaque in typical X-ray fluoroscopy. The low visibility of stent also undermines the assessment of the stent implantation outcome, which increases the risk of incomplete stent expansion. While increasing signal to noise ratio by increasing radiation doses or using stents with radiopaque coatings may

result in negative clinical effects, only a few alternative solutions are available. The most reliable method to assess stent expansion is to observe the stent via intravascular Ultrasounds (IVUS), in which an Ultrasound transducer is put into the target coronary artery via a catheter. Nevertheless, not only that it is an invasive procedure, but IVUS is too expensive for most of current clinics in daily practice.



Figure 1. Left: a frame from a fluoroscopic sequence during stent implantation. Right: stent enhancement through the marker-based motion compensated registration from our algorithm.

Image processing techniques are proposed to improve the image quality for better stent visibility. In [5], a layer decomposition method is used to decompose the X-ray sequence into a sum of moving layers. In [7, 11, 16, 12], a motion-compensated noise reduction via landmark-based registration of multiple images is used. In [2], an elastic registration is proposed to align images through the deformation of guidewire, which shows improvement over marker based alignment. These existing methods utilize conventional image processing techniques such as match filtering or blob detection with automatic scale selection to detect markers, which may have limited capability to cope with large variations and cluttered background presented in real applications. For example, for patients who had previous interventions or surgery, the placed sternal wires, stitches, stents, and other devices introduce locally similar structures to the balloon markers, a significant number of

false alarms in marker detection are often observed. Large variations of marker appearance across time also make it difficult for conventional detection algorithms to consistently differentiate balloon markers, especially when the target markers are overlaid with other structures in the image [2]. Conventional balloon marker detection is heavily dependent upon temporal coherence to compensate detection errors among individual image frames and may require user interactions to achieve desired performance [12].

We propose a novel scheme to detect balloon marker pair as well as the guidewire by a unified scheme using classifiers learned from a large database containing more than ten thousands real fluoroscopic frames during stent implantation. Figure 1 illustrates a fluoroscopic scene and the enhanced stent through motion compensated registration using the balloon markers. The learned classifiers have been evaluated and tested on a database containing 263 real fluoroscopic sequences with 12495 frames. Data is collected world wide including European and Asian sites and covers a large variability. Figure 1 illustrates our marker-based stent enhancement result.

1.1. Discriminative learning

Discriminative learning techniques have been proven effective in many medical imaging applications [4, 3, 17, 8] as well as interventional applications [1, 15] with high accuracy and efficiency. These methods have been quite successful in detecting different landmarks or anatomical structures. Nevertheless, some patterns, such as the guidewire, which is a non-rigid deformable curve, are difficult to be defined by learning-based methods. In [1], classifiers are designed to detect primitive features such as small segments of a guidewire or connectivity between two segments to form a longer curve. However, at the final stage, it relies on dynamic programming to connect all detected segments and curves to form the entire guidewire. Inspired by the application of balloon marker enhancement, we propose to detect the guidewire in conjunction with two landmark points directly. In other words, the wire is obtained directly from the output of the classifier. There are a number of advantages of the proposed scheme:

1. Marker detection is prone to error due to false detections. In interventions, there can be a lot of marker-like patterns inside a scene. It is with high probability that the ground truth marker is not among the top candidates. Therefore, existing methods rely heavily on the coherence analysis along the temporal domain to remove false detections. Nevertheless, false detections can still exist in a cluttered scene with a lot of blob like structures.

In fact, the ground truth that the real marker pair can differentiate from the remaining marker like candidates is that there must be a wire connecting the real marker pair. If we can utilize the fact that the true marker pair should be

connected by the guidewire, we can remove a great deal of false detections even without temporal coherence.

2. The joint probability learned from the classifiers for the marker and the guidewire are more robust than the individual probability of marker detection. Integrating marker and wire probabilities through fusion leads to more robust detection results.

3. In case a guidewire is desired in order to perform a non-rigid matching between frames, the guidewire is obtained inherently from the classifier.

2. Methodology

While learning a balloon marker classifier results in more robust marker detections than conventional methods, false alarms can still be present if only the local context of a single marker is considered. In order to obtain better results, one possible way is to further train another classifier to combine the joint local context of each marker pair to let the classifier selects the best marker pair among all candidates. One example of such approaches was proposed by Lu. et al. [9]. Figure 2 illustrates this idea.

However, there are two limitations: 1. The stent between the two markers can deform wildly and locate outside of the joint context area. 2. The stent can locate in any major coronary and with any types of background. The only semantic ground truth is that both the markers must locate on the guidewire. Learning the features from the entire joint context area may easily lead to overfitting.

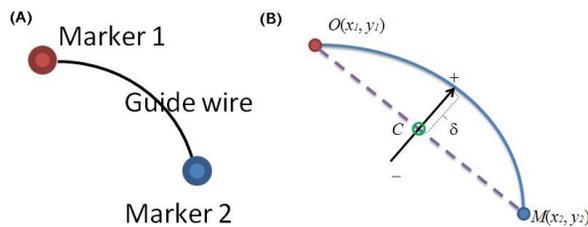


Figure 3. (A) The target detection model χ (B) Parameters of χ .

In order to overcome these limitations, we propose to learn the guidewire directly. We design a classifier which directly locates wire structures as shown in Fig. 3(A). Specifically, we would like to model our target pattern as a marker pair connected by a thin guidewire. Because the balloon is covered by a stent inside the coronary and there is a limitation of the length of the stent, the degree of guidewire deformation is limited. This is based on the observation from around 300 real fluoroscopic sequences collected worldwide. As a result, we model the marker guidewire combination as:

$$\chi(O, M, \delta), \quad (1)$$

where $O = (x_1, y_1)$ and $M = (x_2, y_2)$ are the locations of the two balloon marker, where $x_1 \leq x_2$, and $y_1 \leq y_2$ if

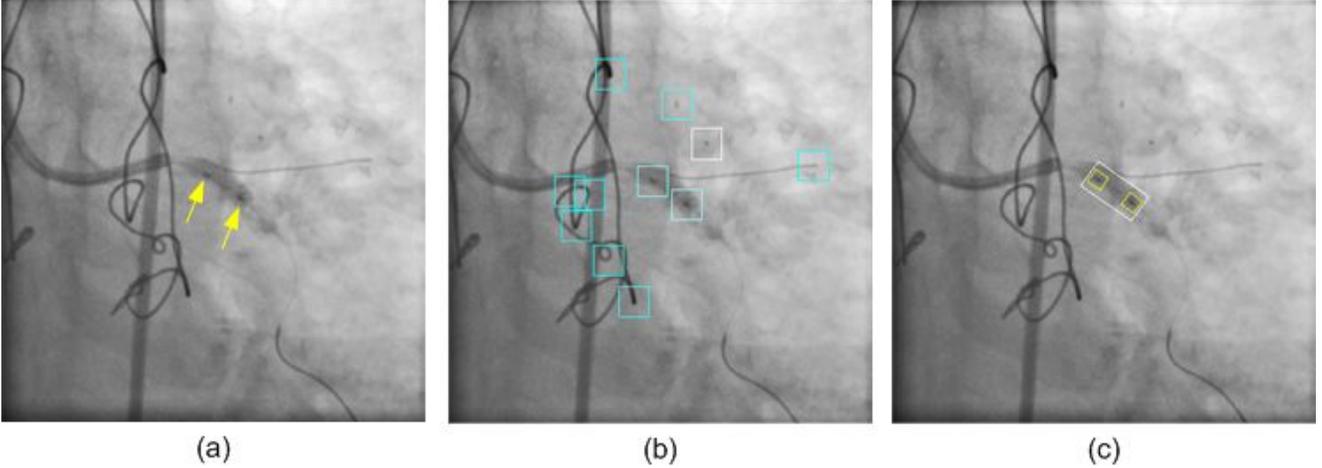


Figure 2. (a) Ground truth balloon marker pair on one fluoroscopic image. (b) Marker detection results from a trained classifier. (c) A joint context region from a pair of marker.

$x_1 = x_2$. δ is defined as the displacement along the normal direction from the center $C = \frac{(O+M)}{2}$ to the guidewire. Fig. 3(B) shows the model. Given an image I , our goal is to train a classifier for $P(O, M, \delta|I)$. The model is then trained in a fashion of marginal space learning [17], where the possible marker locations are identified followed by the possible guidewire locations. And the final probability is calculated in a Bayesian formulation:

$$P(O, M, \delta|I) = P(O, M|I)P(\delta|O, M, I) \quad (2)$$

2.1. Marker detection

We first illustrate how we obtain the possible marker locations $P(O, M|I)$. Marker detection is formulated into a typical object detection framework to solve a two-class (object vs. background) classification problem. A box progresses through an image to extract candidate samples. Each sample is fed to the learned model/detector to be assigned with a likelihood score of being the target object. From another perspective, marker detection is to search the parameter space. For each individual marker, the location parameter space has two parameters, x and y . In our detection scheme, the box based representation is used to include both markers and their context. Given an image, an individual marker detector is applied. Clustered detection results are consolidated into marker candidates.

We use probabilistic boosting trees (PBT) [13] as our learning machine to construct the detectors. The detector is a tree-based structure with which the posterior probabilities of the presence of the marker are calculated from given image data. Therefore, each marker detector not only provides a binary decision for a given sample, but also a confidence value (score) associated with the decision. The nodes in the tree are constructed by a nonlinear combination of simple classifiers using boosting techniques [14, 13].

Each detector selects a set of discriminative features that are used to distinguish the positive (target) from negatives (background) from a large pool of features. Different parameter space utilizes different features computed from image data. For individual marker detectors, we choose Haar wavelet-like features as shown in Figure 4, which are efficiently calculated using integral image-based techniques [14]. Figure 2(B) shows an example of detected markers on a fluoroscopic image.



Figure 4. Haar wavelet-like feature type examples

We apply a bootstrapping strategy to effectively remove the false alarms. We have two stages for individual marker detection. The first stage is trained with target markers against randomly selected background samples. The second stage is trained with the target markers against the false alarms obtained from the first stage detector. The first stage is used to quickly remove negatives and the second one is aimed at pruning out more confusing (difficult) cases. This two-stage approach results in a more robust and efficient solution.

2.2. Guidewire detection

From the marker detection, a list of high probability candidates are extracted. The task is then to select a pair of markers which are most likely to be the target balloon marker pair, i.e, we need to find $P(\delta|O, M, I)$.

2.2.1 Training

For the guidewire, we model it as a 2D cubic spline curve with three control points, that is, the first control point at O , and the last control point at M . The location of the second control point is obtained through the parameter δ as defined in Fig. 3(B).

Given the marker candidates in image I , we obtain $P(O, M|I)$ from every two candidates. One positive and thousands of negatives samples can then be generated for training. The positive sample is the spline curve which locates on the ground truth guidewire (white solid curve in Fig. 5), and negative samples from either the splines other than the ground truth markers or all possible spline curves from marker pair which is not the ground truth pair (red dashed curve in Fig. 5).

In order to search the spline parameter, we discretize the hypothesis space of δ by a ratio defined by the Euclidean distance between the two markers, $d = \|O - M\|_{L_2}$. As a consequence, the hypothesis space of δ becomes:

$$H_\delta = \{-r, -r + \Delta_\delta, -r + 2 \times \Delta_\delta, \dots, 0, \dots, r - \Delta_\delta, r\}, \quad (3)$$

where $r = 0.36 \times d$ is learned from our training database. $\Delta_\delta = 0.15mm$ is the search step size, which should be small enough to capture the groundtruth guidewire at subpixel levels. In order to calculate features along the guidewire, we sample n points along the wire with equal arc length interval, including the two markers. Fig. 6 illustrates two sampling examples for the same O , and M . The left one has larger n , and smaller r and Δ_δ . In our training settings, n is set to 41.

The features we used for training include steerable features [17] extracted at the n sampling points along the wire model, which contain a number of gradient-derived variants, such as magnitude and angle with respect to certain pre-defined orientations w.r.t. the marker pair orientation. Computations of such features is efficient as it does not require image rotation and re-scaling. In addition, magnitude and orientation output of steerable filters [6] are included in the feature pool as well. This is because the steerable filter, as a ridge filter, can capture the guide wire pretty well in general.

The advantages of such classifier design include: 1. It exploits the marginal space learning scheme to reduce search from the large full space (marker position, wire parameters). 2. The wire search stage can be regarded as another bootstrapping of the marker detection, which further improves the marker detection accuracy. 3. A unified probability is obtained. 4. Following the detected markers, the guidewire is automatically localized. This is beneficial when non-rigid matching is preferred and the wire structure is needed.

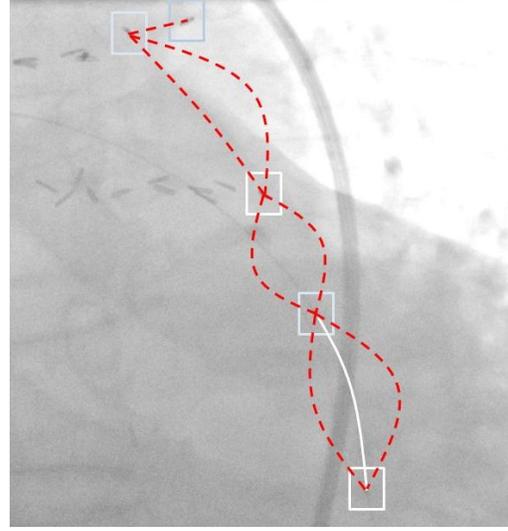


Figure 5. Positive and negative examples in training data generation. Boxes represents marker candidates. The brighter the color, the higher the probability is. The solid (white) curve is a positive wire example. The dash (red) curves are negative wire examples.

2.3. Detection

The individual marker detection with the bootstrapping strategy is applied. Multiple marker pair candidates are formed by the top individual marker candidates. Then for each marker pair candidate, the guidewire classifier is applied to search in the hypothesis space H_δ . The marker pair with the best $P(O, M|I) \times P(\delta|O, M, I)$ is used as the final marker detection results.

3. Stent enhancement

In this section, we present the motion compensated stent enhancement results based on our marker-guidewire detection. As we will illustrate in the next section, with the marker guidewire detection framework, more than 90% of the time the markers are correctly detected. With this high confidence, a heuristic temporal coherence analysis is effectively applied to remove the outliers and generate the stent enhancement image. Our temporal coherence analysis consists of the following steps:

1. Apply the marker pair and guidewire detection on every frame of the sequence, assume $P_i(\chi)$ is the probability at frame i .
2. Calculate a consistency score of each frame i as:

$$C_i = \frac{1}{N-1} \times \sum_{i,j,i \neq j} \exp\left(-\frac{d(i,j)^2}{\sigma^2}\right), \quad (4)$$

where N is the total number of frames, i, j is the index of two individual frames, and $d(i, j) = d_\alpha(i, j) + d_\beta(i, j)$ with the following definitions:

$$d_\alpha(i, j) = \|O_i - O_j\|_{L_2} + \|M_i - M_j\|_{L_2}, \quad (5)$$

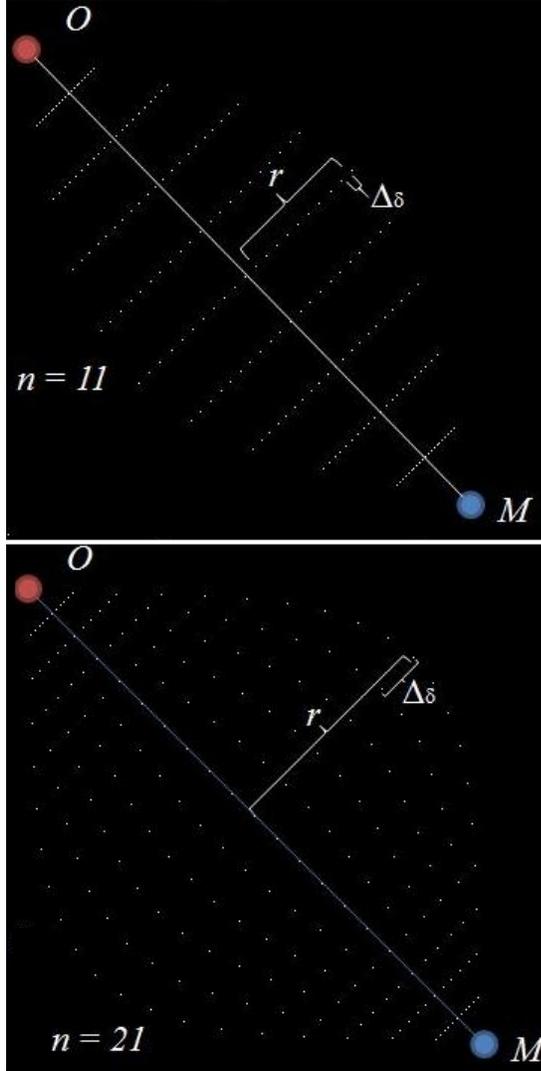


Figure 6. Two different sampling examples due to different r , $\Delta\delta$, and n for the same O , and M .

is the sum of distance of the two detected markers on frame i and frame j .

$$d_\beta(i, j) = |||O_i - M_i||_{L^2} - ||O_j - M_j||_{L^2}||_{L^1}, \quad (6)$$

is the length difference between the detected balloon marker pair O_i, M_i at frame i and the detected balloon marker pair O_j, M_j at frame j .

3. Calculate the reference frame f as $\arg \max_i P_i \times C_i$.

4. From frame f , moving forward and backward to frame k ($k = f - 1, f + 1, f - 2, f + 2, \dots$) to remove detected pair if $d(f, k) > \tau$. τ is a threshold set manually.

4. Experimental results

We collected 263 clinical fluoroscopic sequences during stent implantation containing 12495 frames, which served

as our database for training and evaluation. A sequence has from 10 to 188 frames and each frame contains a pair of balloon markers. Data was acquired from clinics in both Europe and Asia, which covers a wide variety of different stents, balloon markers, angulations, patients, and clinical settings. For each individual frame, we manually annotate the marker positions and the guidewire in between the two markers. The guidewire is annotated by control points of a 2D spline curve.

All image frames are normalized onto the resolution of 0.308mm/pixel. A 4-fold cross-validation scheme was applied for evaluation. That is, the entire dataset was randomly partitioned into four quarters, where no frames from the same sequence appear in different quarters. For each fold evaluation, three quarters were combined for training and the remaining one quarter was used as unseen data for testing. This procedure was repeated four times so that each frame has been used once for testing. For each frame, the average Euclidean distance of the detected marker pair from their groundtruth positions was calculated. We evaluated the individual marker detection scheme and the fusion scheme of marker and wire. For each detected wire along with its two end markers, 41 points were resampled with equal arc length so that a point-to-point distance was computed against the annotated control points. The average distance for all 41 points was used to measure the wire localization precision. Table 1 and Fig. 7 show the marker detection results. Detection examples are provided in Fig. 8. Wire segment localization accuracies are summarized in Table 2.

Table 1. Marker detection accuracies by a 4-fold cross validation.

(a) Marker detection rank-1 accuracy, i.e., the percentage of correct detections, where a correct marker detection for each frame is defined as the two marker candidates with the highest probabilities are within 3 pixels from the ground truth. (b) Marker detection precision. Average distance of the two detected marker candidates with the highest probabilities from ground truth positions. Distances are in unit of pixels.

(a)

	Rank-1
Marker only	83%
Fusion of markers and wire	90%

(b)

	Mean	Std	Median
Marker only	15.1	38.2	0.50
Fusion of markers and wire	8.9	33.6	0.47

When evaluating on the 263 real fluoroscopic sequences, each of which contains between 10 and 188 frames. The stent enhancement is successful in 259 sequences (98.5%)

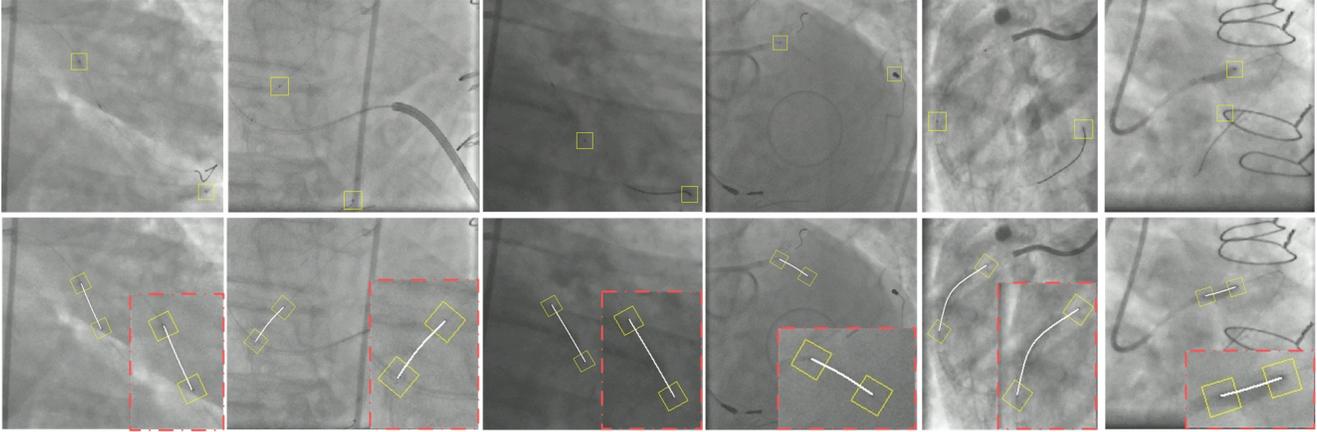


Figure 8. Examples of automatic detection of the marker+wire (bottom) in comparison to marker only (top). Marker detection results are shown as the center of the yellow boxes, while the white wire shows the detected wire by a spline generated from the three control points. A close-up at the bottom-right corner of each image is provided for better visualization.

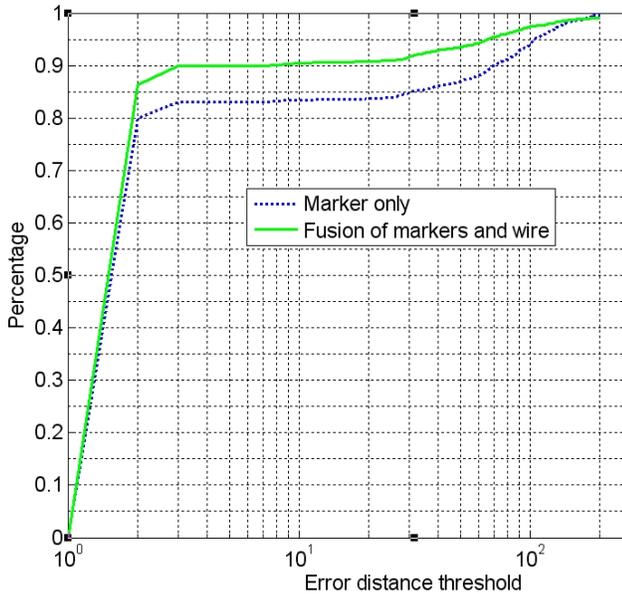


Figure 7. Percentile evaluation results of 4-fold cross-validation. Each curve shows the percentage of cases whose distance of the automatic detection results from ground truth is less than an error distance threshold. Note that the horizontal axis is in log domain.

accuracy).

A successful enhancement is defined by the following two criteria: 1. The algorithm selects at least 80% of the total frames or at least 30 frames from the input sequence. From the 263 sequences, the mean and median percentages of number of frames used for each sequence are 91% and 93%. And in practice, the stent is clearly enhanced if there are 30 frames that have consistent detection results. 2. The balloon marker pair is located correctly on all the selected frames. If any one of the two criteria is not met, the sequence is regarded as a failure case.

Table 2. Wire evaluation results of 4-fold cross validation. The average distance for all 41 control points on the automatically detected wire segment against their counterpart in the ground truth was calculated. Distances are in unit of millimeters. Two scenarios are presented: I. All images are included; II. only images with final marker detection results are within 3 pixels from the ground truth, which corresponds to 90% accuracy in Table 1.

	Mean	Std	Median
I	3.0184	10.3488	0.2556
II	0.44	0.6375	0.231

Using the same criteria and the same temporal coherence analysis, we obtain 60.8% and 87.8% accuracy if we apply conventional image processing method for blob-detection, or if we apply only the trained marker detection without the guidewire. Figure 9 shows examples of our stent enhancement results. The image is filtered by a multi-scale image enhancement algorithm during marker-based registration to remove halo artifacts. This is beyond the scope of this paper and interesting reader may refer to [10] for more detail. Detection on single frame takes about 0.05 seconds. In practice, it takes ~ 5 seconds to detect marker pairs on all the frames on a scene with 100 frames.

5. Conclusion

We have proposed a novel method to detect a non-rigid deformable guidewire segment, and a joint marker-guidewire segmentation fusion scheme under a Bayesian formulation. We utilize the markers to constrain the search of the guidewire parameter space. Quantitative evaluation and cross-validation have been conducted to show significant improvement over marker-only based detection. Applications such as stent enhancement are greatly improved

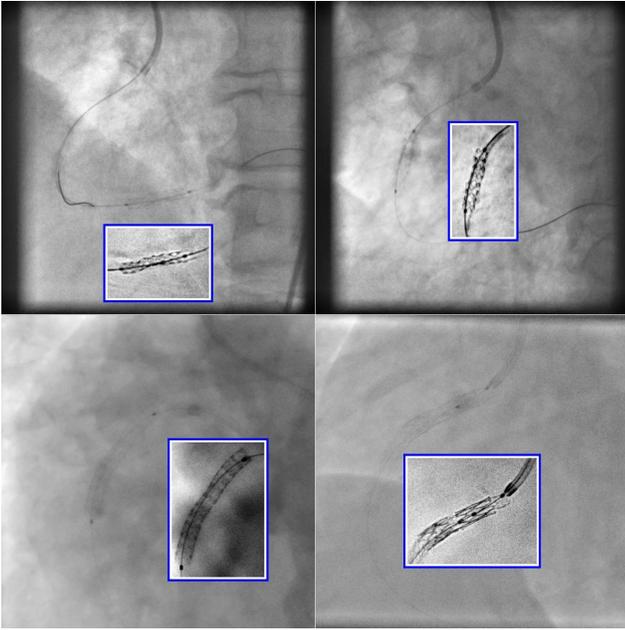


Figure 9. Four examples of the stent enhancement results. The image is one frame from the original sequence. Inside the box is our motion compensated stent enhancement result.

with such design. The same scheme can be generalized to locate longer guidewire or catheter with higher order parameterizations, which is among our future work. For example, a guidewire normally consists of the guiding catheter tip, the wire body, and the guidewire tip. Location of guiding catheter tip and the guidewire tip can be treated as the marker pair, which can be detected by trained classifiers. Another example is that during IVUS pullback, the guiding catheter tip and the IVUS transducer can be located together with the IVUS catheter. Figure 10 illustrates these two examples.

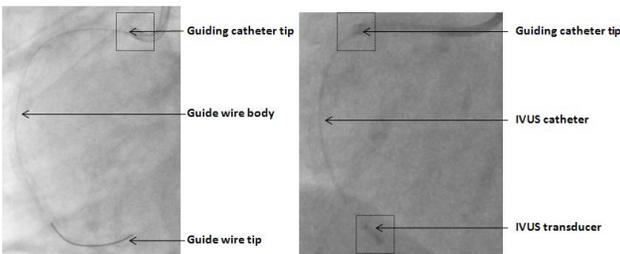


Figure 10. The proposed methodology can be generalized by replacing the balloon markers with other landmarks such as the guiding catheter tip, IVUS transducer, guidewire tip, etc.

References

[1] A. Barbu, V. Athitsos, B. Georgescu, S. Boehm, P. Durlak, and D. Comaniciu. Hierarchical learning of curves: Appli-

cation to guidewire localization in fluoroscopy. *IEEE Conf. Computer Vision and Pattern Recognition*, 2007.

[2] V. Bismuth and R. Vaillant. Elastic registration for stent enhancement in x-ray image sequences. *ICIP*, 2008.

[3] G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu. Automatic fetal measurements in ultrasound using constrained probabilistic boosting tree. *Int. Conf. Medical Image Computing and Computer Assisted Intervention*, 2007.

[4] T. Chen, W. Zhang, S. Good, K. Zhou, and D. Comaniciu. Automatic ovarian follicle quantification from 3d ultrasound data using global/local context with database guided segmentation. *ICCV*, 2009.

[5] R. Cloase, C. K. Abbey, and J. S. Whiting. Improved localization of coronary stents using layer decomposition. *Computer Aided Surgery*, 2002.

[6] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *PAMI*, 13(9):891–906, 1991.

[7] J. J. Koolen, M. V. Veer, and C. E. E. Hanekamp. A stentboost image enhancement: first clinical experience. *MEDICAMUNDI: Clinical Applications*, 2005.

[8] H. Ling, Y. Zheng, B. Georgescu, S. Zhou, M. Suehling, and D. Comaniciu. Hierarchical learning-based automatic liver segmentation. *IEEE Conf. Computer Vision and Pattern Recognition*, 2008.

[9] X. Lu, B. Georgescu, A. Littmann, E. Mueller, and D. Comaniciu. Discriminative joint context for automatic landmark set detection from a single cardiac mr long axis slice. *Functional Imaging and Modeling of the Heart*, 2009.

[10] M. Maitre, Y. Chen, and T. Fang. High-dynamic range compression using a fast multiscale optimization, 2008. ICASSP.

[11] J. Mishell, K. Vakharia, T. Portset, Y. Yeghiazarians, and A. Michaels. Determination of adequate coronary stent expansion using stentboost, a novel fluoroscopic image processing technique. *Catheterization and Cardiovascular Interventions*, 69:84–93, 2007.

[12] G. Schoonenberg, P. Lelong, R. Florent, O. Wink, and B. Romeny. The effect of automated marker detection on in vivo volumetric stent reconstruction. *MICCAI*, 2008.

[13] Z. Tu. Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. *ICCV*, II:1589–1596, 2005.

[14] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.

[15] P. Wang, M. Pfister, T. Chen, and D. Comaniciu. Using needle detection and tracking for motion compensation in abdominal interventions. *IEEE International Symposium on Biomedical Imaging*, 2010.

[16] A. O. Zaid, I. Hadded, W. Belhaj, A. Bouallegue, S. Abdesslem, and R. Mechmeche. Improved localization of coronary stents based on image enhancement. *Intl J. on Biomedical Science*, 2008.

[17] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Four-chamber heart modeling and automatic segmentation for 3d cardiac ct volumes using marginal space learning and steerable features. *IEEE Trans. Medical Imaging*, 27(11):1668–1681, 2008.