

Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Catheter segmentation in three-dimensional ultrasound images by feature fusion and model fitting

Hongxu Yang
Caifeng Shan
Arash Pourtaherian
Alexander F. Kolen
Peter H. N. de With

Catheter segmentation in three-dimensional ultrasound images by feature fusion and model fitting

Hongxu Yang,^{a,*} Caifeng Shan,^b Arash Pourtaherian,^a Alexander F. Kolen,^b and Peter H. N. de With^a

^aEindhoven University of Technology, VCA Research Group, Eindhoven, The Netherlands

^bPhilips Research, In-Body Systems, Eindhoven, The Netherlands

Abstract. Ultrasound (US) has been increasingly used during interventions, such as cardiac catheterization. To accurately identify the catheter inside US images, extra training for physicians and sonographers is needed. As a consequence, automated segmentation of the catheter in US images and optimized presentation viewing to the physician can be beneficial to accelerate the efficiency and safety of interventions and improve their outcome. For cardiac catheterization, a three-dimensional (3-D) US image is potentially attractive because of no radiation modality and richer spatial information. However, due to a limited spatial resolution of 3-D cardiac US and complex anatomical structures inside the heart, image-based catheter segmentation is challenging. We propose a cardiac catheter segmentation method in 3-D US data through image processing techniques. Our method first applies a voxel-based classification through newly designed multiscale and multidefinition features, which provide a robust catheter voxel segmentation in 3-D US. Second, a modified catheter model fitting is applied to segment the curved catheter in 3-D US images. The proposed method is validated with extensive experiments, using different *in-vitro*, *ex-vivo*, and *in-vivo* datasets. The proposed method can segment the catheter within an average tip-point error that is smaller than the catheter diameter (1.9 mm) in the volumetric images. Based on automated catheter segmentation and combined with optimal viewing, physicians do not have to interpret US images and can focus on the procedure itself to improve the quality of cardiac intervention.

© 2019 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.6.1.015001]

Keywords: catheter segmentation; three-dimensional ultrasound; intervention guidance; feature fusion; model fitting.

Paper 18153RR received Jul. 17, 2018; accepted for publication Dec. 14, 2018; published online Jan. 14, 2019.

1 Introduction

Intervention therapy, also called interventional cardiology, is a type of cardiology based on catheter treatment for structural heart diseases or electrophysiology, which has been widely used during the past decades due to its minimized incision and shorter recovery time. Many cardiology procedures are performed by catheterization, which involves insertion of a catheter into the heart through a femoral artery or any large vein. During the intervention without open surgery, the obstructed tools and organs have to be visualized through imaging modalities, such as x-ray and ultrasound (US). However, using x-ray has drawbacks such as radiation dose and harmful contrast agents make many researchers focus on US-based catheterization. Applying US has many benefits, including mobility, lower cost, and the absence of radiation, for both surgeon and patient. Recently, three-dimensional (3-D) US imaging has achieved fast and real-time performance, which offers great potential for image-guided interventions and therapies, offering more spatial information derived from direct 3-D sensing.

Despite its advantages, 3-D US suffers drawbacks such as low signal-to-noise ratio in the images, lower resolution than two-dimensional (2-D) US images and degraded instrument visibility. Consequently, the target area is hard to recognize in images, requiring surgeons with sufficient navigation skills, while consuming extra effort during the operation. Figure 1 demonstrates an example of the 3-D US image quality for cardiac catheterization. To improve the quality of the intervention and to help the clinician to focus on the procedure itself rather

than finding the instruments, automatic catheter segmentation in 3-D US images becomes beneficial, because it makes it easier to better identify the catheter in the correct heart chamber in the US images. Many researchers have concentrated on catheter identification in US imaging with the aid of robots,¹ or by adding active sensors inside the catheter.² Although these approaches have achieved attractive results, the high cost of equipment and complicated system set up in the operation room have hampered their broad acceptance.

In this work, we focus on cardiac catheter segmentation in 3-D US images for image-guided cardiac intervention therapy. With the segmented catheter in the 3-D US volumetric data, a better visualization, and perception can be provided to physicians that they can interpret the instrument with less effort. Figure 1 shows an example of the application of US-guided intervention therapy. The US image with a catheter is acquired from a catheterization experiment. As can be observed, the catheter is hardly recognized in the low-resolution 3-D US image, because the surrounding anatomical structures may resemble a catheter. Since segmenting the catheter in 3-D cardiac US is challenging, a more in-depth study on discriminative catheter features is required for supervised learning. Also, a better catheter model should be defined for an improved segmentation accuracy.

1.1 Related Work

Many studies have recently focused on image-based medical instrument segmentation or identification in the 3-D US, but

*Address all correspondence to Hongxu Yang, E-mail: h.yang@tue.nl

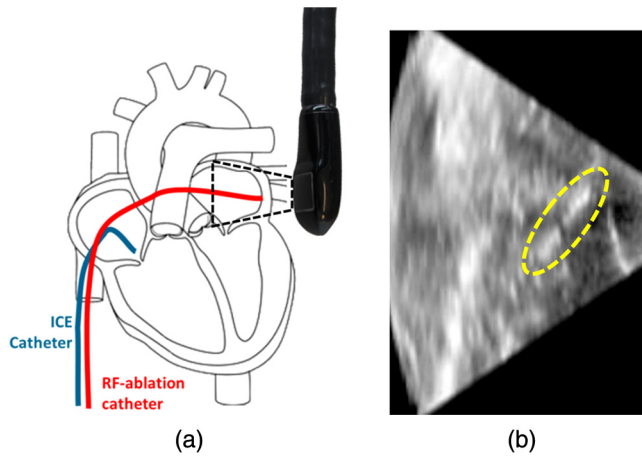


Fig. 1 (a) Example of cardiac intervention therapy with catheters in the heart. The black bar is the US probe; (b) 2-D slice from a 3-D image, where the RF-ablation catheter is marked in yellow circle.

their approaches are not suitable for catheter segmentation in cardiac imaging. Methods such as principle component analysis (PCA)³, Hough transformation,⁴ and parallel integral projection transformation⁵ were proposed to detect straight electrodes in 3-D images. However, these transformation-based methods are not stable when the background includes high-intensity values, as is the case with instruments. This instability results from the fact that image transformation cannot extract the discriminant shape information of tools to distinguish them from bright tissues or noise. Cao et al.⁶ proposed a template matching method to segment a catheter with *a-priori* knowledge of direction and diameter. Nevertheless, this method is limited by the *a-priori* knowledge of the shape and orientation of the catheter. In addition, the carefully designed template is not only unstable to catheter appearance variations but also lacking discriminating information. Alternatively, Uherčík et al.^{7,8} applied Frangi et al.⁹ vesselness features to classify instrument's voxels using supervised learning algorithms. The model-fitting based on RANdom SAMple Consensus (RANSAC) was applied to determine straight tubular-like instruments. Meanwhile, Zhao et al.¹⁰ used a similar method to track the needle on an ROI-based Kalman filter. Although the ROI-based algorithm decreases computation complexity, there are still some limitations. First, the ROI-based algorithm requires a fixed view of images, which introduces the extra consideration of avoiding ultrasound transducer movement during operation. Furthermore, both Uherčík et al.⁸ and Zhao et al.¹⁰ only considered a predefined Frangi feature as discriminating information, which is not only less robust to diameter variation but also considers a small amount of information only, i.e., information in ultrasound volume is not fully used for discriminating classification. Recently, Pourtaherian et al.^{11–13} have intensively studied needle detection algorithms based on the 3-D US. Their method segments the candidate needle-like voxels by incorporating the Gabor-based feature. This feature introduces more discriminating information on local orientation distribution, which is similar to the histogram of gradient. After the voxel-based classification, the two-point RANSAC algorithm is applied to estimate the axis of the needle. However, their proposed method is specifically designed for a thin needle with a large length versus diameter ratio in a high quality US image, which is not the case in cardiac catheter segmentation. Although they did an experiment on catheter

segmentation in an *in-vitro* dataset, their results showed that further studies on segmenting the catheter on *ex-vivo* or *in-vivo* dataset are necessary.

Although the methods discussed above have shown successful results in 3-D ultrasound-based instrument segmentation, there are still many challenges concerning cardiac intervention. First of all, the spatial resolution of the transducer in cardiac intervention is far lower than the standard one used in the anesthesia (response frequency from 5 to 13 MHz in anesthesia probe¹³ to 2 to 7 MHz in the cardiac probe). This low spatial resolution leads to lower image contrast and fewer details at instrument boundaries. Second, the transducer, e.g., phase-array transducer, is optimized designed for cardiac tissue visualization that the plastic catheter (not like a metal needle) cannot be perfectly visualized in the 3-D US. This phenomenon makes catheter deformed in shape and diameter. Third, in the 3-D volume, the cardiac tissue occupies a larger space than the catheter, which contains anatomical structures that make it more challenging to segment the catheter from complex anatomical structures using traditional methods. As a result, the methods above may not be suitable for catheter segmentation in 3-D US images to facilitate cardiac intervention therapy.

1.2 Our Work

In this paper, we present an extensive study on various features for catheter segmentation in the 3-D US data. With the observation that catheters and tissues have different responses under different scale window, we extend the Frangi filter to multiscale operation.¹⁴ This approach can better handle the diameter variation than filtering with a predefined scale.^{8,10} In addition, instead of only describing the tubular structure as the traditional Frangi filter, a multidimensional objectness feature is derived.¹⁵ Considering the information loss in objectness features and inspired by the circular shifting in the existing methods,¹³ we propose Hessian-based features to fully describe the 3-D information. Next to this, log-Gabor filters are considered to add more orientation information. Last, statistical features are defined to further extract local information about the voxels. Using the state-of-the-art classifiers,¹⁶ these features are compared on multiple datasets for the catheter-like voxel classification. Our experiments show that the best voxel classification can be achieved by fusing these features. Based on catheter-like voxel classification, we present a modified spares-plus-dense (SPD) RANSAC model fitting for catheter segmentation, which employs cubic spline fitting to identify the curvilinear catheter inside the 3-D US images. With successful catheter segmentation in the volumetric US data, physicians would have less effort to identify the instrument in the images and easier to interpret the images for operation.

Our contributions are summarized as follows. First, we present an extensive study on various features extracted from the 3-D US for catheter-like voxel classification, where it will be shown that the best voxel classification is achieved by fusing these features. Second, a modified model fitting algorithm is introduced for catheter segmentation in the noisy voxel-classification output data. Third, we have collected multiple datasets (*in-vitro*, *ex-vivo*, and *in-vivo*), which are used to extensively validate our method on these datasets.

The paper is structured in the following way. Section 2 describes the proposed method in detail, including the various features for voxel classification and the modified SPD-RANSAC model-fitting algorithm. The collected datasets and

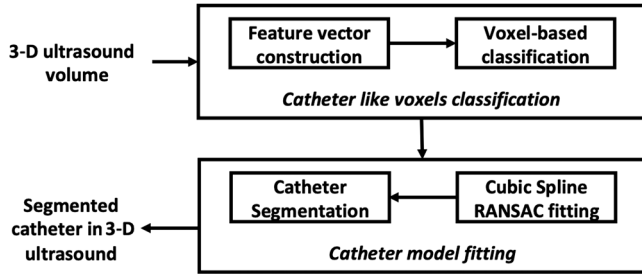


Fig. 2 Block diagram of the proposed system for catheter segmentation method.

experimental results are discussed in Sec. 3. Finally, Sec. 4 concludes the paper and presents some discussions on possible refinements.

2 Methodology

Figure 2 shows the block diagram of our catheter segmentation system. In the first step, the 3-D volumetric image is processed to extract features from each voxel. The voxels are then classified by supervised learning methods into catheter-like voxels and noncatheter voxels. In the second step, in the noisy output of voxel classification, the modified SPD-RANSAC model fitting is applied to segment the catheter. We describe each step in the following sections.

2.1 Catheter-Like Voxel Classification

The procedure of catheter-like voxel classification consists of two steps. First, 3-D discriminating features are extracted from each voxel in the 3-D US image and then the supervised learning classifier is applied to classify the voxels. The discriminating features employed for voxel classification are described in the following paragraphs.

2.1.1 Objectness feature

Multidimensional objectness was first introduced by Antiga,¹⁵ who extended the traditional definition of vesselness filter into the different shape descriptions for multidimensional images, see Fig. 3 as an example. For a 3-D image, the Hessian matrix is defined below, where the f^σ is a Gaussian-filtered image with the standard deviation σ , while f_{xx}, \dots, f_{zz} represents the second-order derivative in the x -, y -, or z -directions. This leads to Eq. (1), specifying the matrix as

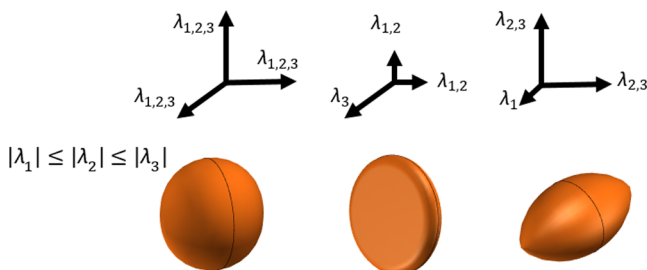


Fig. 3 Objectness descriptors based on Eigenvalues of the Hessian matrix, showing different structures.

$$H_\sigma = \begin{bmatrix} f_{xx}^\sigma & f_{xy}^\sigma & f_{xz}^\sigma \\ f_{yx}^\sigma & f_{yy}^\sigma & f_{yz}^\sigma \\ f_{zx}^\sigma & f_{zy}^\sigma & f_{zz}^\sigma \end{bmatrix}. \quad (1)$$

The Eigenvalues of Eq. (1) are ranked by $|\lambda_1| \leq \dots \leq |\lambda_3|$. Using Eigenvalues, the M -dimensional ($M \leq 3$) shape structures are described by Eq. (3) based on the parameters from Eq. (2) ($M = 0$ for blob, $M = 1$ for vessel, and $M = 2$ for plate in shapes, i.e., the Frangi vesselness filter equals to $M = 1$). Parameters R_A and R_B have two special cases. When $M = 2$, parameter $R_A = \infty$, and when $M = 0$, parameter R_B is set to zero:

$$\begin{cases} R_A = \frac{|\lambda_{M+1}|}{\prod_{i=M+2}^3 |\lambda_i|^{1/(3-M-1)}} \\ R_B = \frac{|\lambda_M|}{\prod_{i=M+1}^3 |\lambda_i|^{1/(3-M)}} \\ S = \sqrt{\sum_{j=1}^3 \lambda_j^2} \end{cases}. \quad (2)$$

Similar to Frangi's vesselness feature,⁹ when $\lambda_j < 0$ for $M < j \leq 3$, the objectness measurement is defined as

$$O_\sigma^M = (1 - e^{-R_A^2/2\alpha^2}) \cdot e^{-R_B^2/2\beta^2} \cdot (1 - e^{-S^2/2\gamma^2}). \quad (3)$$

For the other cases of j , the value of $O_\sigma^M = 0$. The parameters α , β , and γ are empirically determined, which defines the sensitivity of the response.¹⁴

From the original definition, both Antiga and Frangi select the maximum response per pixel among a range of spatial scale, e.g., the maximum value among the scale range with $\sigma \in [1, \dots, 5]$. However, this maximizing step loses scale-distribution information. As a result, we propose to exploit all the scale responses as features. Meanwhile, we calculate three different shape measurements, i.e., $M = 0, 1, 2$, instead of the tube descriptor used for needle detection in Ref. 8. Based on the definitions above, for each voxel v in 3-D volume V , the final feature vector is $\mathcal{O}(v) = [O_{\sigma=1}^{M=0}(v), O_{\sigma=1}^{M=1}(v), O_{\sigma=1}^{M=2}(v), \dots, O_{\sigma=3}^M(v), \dots]$ in the multiscale approach, where σ represents a standard deviation of a Gaussian filter and M is the type of objectness feature ranging from 1 to 3.

2.1.2 Hessian features

Essentially, the Eigenvalue analysis in the objectness feature space is to extract the direction information of edge distributions through the Hessian matrix and to remove the noise. However, the predefined descriptors in objectness may lose some information because of the low signal-to-noise ratio and due to the projection of nine features into three Eigenvalues. To preserve more information from a low-contrast image, we consider the elements of the Hessian matrix, as given in Eq. (1). Due to the symmetric structure of the Hessian matrix and to preserve the orientation response, we use six elements from the upper right of the Hessian matrix, and the shift the maximum response to the first position via circular shifting. As a result, the feature vector length is six for a specific scale. The Hessian feature is noted as $\mathcal{H}(v) = [\mathcal{H}(v, \sigma = 1), \dots, \mathcal{H}(v, \sigma = 3), \dots]$ in the multiscale case with shifted elements of Eq. (4) in each scale:

$$\mathcal{H}(v, \sigma) = [f_{xx}^\sigma, f_{xy}^\sigma, f_{xz}^\sigma, f_{yz}^\sigma, f_{zz}^\sigma, f_{yy}^\sigma]. \quad (4)$$

2.1.3 Log-Gabor filter

Pourtaherian et al.^{11–13} introduced Gabor features as an attractive discriminative feature for needle detection. The conventional Gabor-based features can be influenced by the DC components of the images.¹⁷ To stabilize the performance for varying DC components and related image gray-value variations in different US images, we adopt 3-D log-Gabor features.¹⁷ The 3-D log-Gabor filter in the frequency domain is defined as

$$\mathcal{L}(\omega, \phi, \theta) = \exp\left[\frac{\left(\log \frac{\omega}{\omega_0}\right)^2}{2 \log \frac{B}{\omega_0}}\right] \times \exp\left[-\frac{\alpha(\phi, \theta)^2}{2\sigma_\alpha^2}\right], \quad (5)$$

where B is the bandwidth of the filter in polar coordinates, and ω_0 is the filter's central response frequency. The term B/ω_0 is set to a constant value to keep constant shape ratio filters. The direction of the filter is defined by azimuth angle ϕ and elevation angle θ . The position vector $\alpha(\phi, \theta)$ at frequency f is defined as $\alpha(\phi, \theta) = ar \cos[(f \cdot d)/|f|]$, where the unit direction vector is $d = (\cos \phi \cos \theta, \cos \phi \sin \theta, \sin \phi)$. The bandwidth of angular direction is defined by σ_α . Discriminative features are extracted using the real parts of the response in the spatial domain, due to their symmetric response. The circular operation is performed to shift the maximum response to the center and is denoted by $\mathcal{L}(v, \omega)$ at specific frequency ω . The log-Gabor feature is denoted by $\mathcal{L}(v) = [\mathcal{L}(v, \omega = 2\pi), \dots, \mathcal{L}(v, \omega = 6\pi), \dots]$ for multiple frequencies with unit 2π . We have chosen both angle parameters, i.e., ϕ and θ , to {15 deg, 65 deg, 115 deg, 165 deg} after several experiments.

2.1.4 Statistical features

To extract more local information in 3-D cube, we propose to introduce a feature type, i.e., local statistical features. For a voxel v at the center point, we extract a 3-D cube with specific sizes, such as $3 \times 3 \times 3$ voxels. The statistical features are obtained by calculating the mean, standard deviation, maximum, minimum, and local entropy of this cube. The statistical feature $\mathcal{S}(v)$ in the multiscale case is denoted as

$$\mathcal{S}(v) = [I(v), \text{mean}_{s=3}(v), \text{std}_{s=3}(v), \text{max}_{s=3}(v), \text{min}_{s=3}(v), \text{en}_{s=3}(v), \dots], \quad (6)$$

where the s is the size of cubes expressed in voxels using v as center point.

Table 1 summarizes our proposed 3-D features with their symbols, scale variables, and feature lengths for each scale.

To enhance the performance of voxel classification, we apply a feature fusion strategy, which combines four different types of

Table 1 Summary of 3-D features for catheter voxel classification.

Name	Symbol	Scale	Length
Objectness	\mathcal{O}	σ	3/scale
Hessian	\mathcal{H}	σ	6/scale
log-Gabor	\mathcal{L}	ω	16/scale
Statistic	\mathcal{S}	s	1+5/scale

features in a multiscale approach. The fused feature $\mathcal{C}(v)$ is defined as $\mathcal{C}(v) = [\mathcal{O}(v), \mathcal{H}(v), \mathcal{L}(v), \mathcal{S}(v)]$, in which each component is in multiscale.

2.2 Supervised Classifiers for Voxel Classification

To achieve the best performance of the proposed features, we perform the classification under linear discriminant analysis (LDA), linear support vector machine (LSVM), random forest (RF), and adaptive boosting (AdaBoost). Typically, the kernel-based SVM performs better than LSVM, but fine-tuning of kernel parameters requires a large computation cost and its performance is no better than RF and AdaBoost from empirical experience.¹⁶ As a consequence, we consider only LSVM as the SVM classifier, which has box constraint equal to unity. The RF is set to generate 50 trees. For the AdaBoost, the weak learner is set to be decision stump with 50 learning cycles. During the training stage, due to an imbalanced class ratio, we randomly resample the noncatheter voxels to have the same size as the catheter voxels. For testing, the whole volume of imbalanced classes is processed to generate the classified volume. Because of the class imbalance in the testing image, we use precision (P), recall (R), specificity (SP), and F_1 score as evaluation metrics for classification performance after the supervised classification on each 3-D US image. The definitions are shown in Eq. (7). The TP , TN , FP , and FN represent true positive, true negative, false positive, and false negative, respectively. Specifically, the positive voxels are defined as voxels from the catheter, while the negative voxels represent the remaining voxels in the whole image, which typically involves an amount of voxels that is thousands times larger than the amount of catheter voxels:

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN},$$

$$SP = \frac{TN}{TN + FP}, \quad F_1 = 2 \cdot \frac{P \cdot R}{P + R}. \quad (7)$$

2.3 Catheter Model Fitting

Misclassified voxels commonly occur, due to the complex local information from anatomical structures inside the heart and nonperfect description of 3-D features. As a result, after the voxel classification, there are many outlier blobs inside the US volumes. Figure 4 shows example results from AdaBoost.

To correctly segment the catheter in the noisy 3-D image, we apply catheter model fitting based on the *a-priori* knowledge that the shape of the catheter is a curved cylinder. The medical instrument model is conventionally reconstructed by fitting its skeleton together with instrument body voxels surrounding it.^{7,13} However, this method is not stable and inaccurate when assuming a straight-line model in our challenging and noisy classified images. To segment the curved catheter in 3-D US, we first modified the sparse-plus-dense-RANSAC (SPD-RANSAC)¹⁸ from the high-contrast x-ray image into our 3-D US to reduce the complexity of segmentation. Meanwhile, we also modify the instrument model into a three-point curvature line to improve the segmentation accuracy. In the following paragraphs, we first describe the generation of a sparse volume, which reduces the complexity of the RANSAC algorithm. After this, a more complex catheter model is introduced to improve the segmentation accuracy, based on modified SPD-RANSAC.

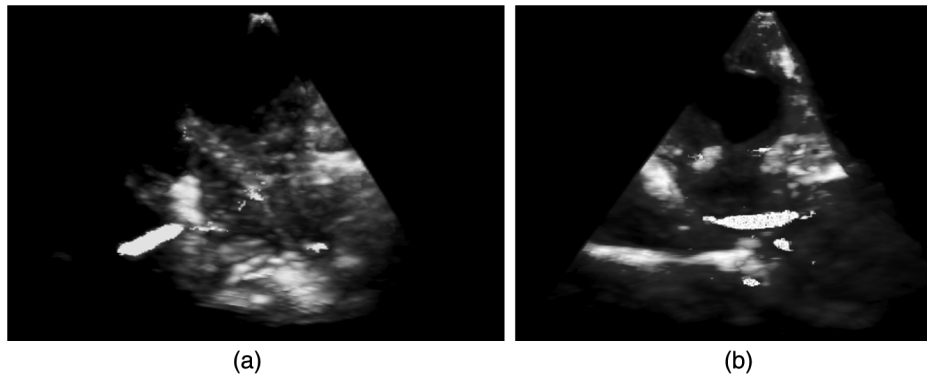


Fig. 4 Examples of classified volumes. The classified catheter voxels are highlighted in the images by high intensity. (a) A classified volume from *in-vivo* dataset and (b) a classified volume from *ex-vivo* dataset.

2.3.1 Sparse volume generation for 3-D US

After the voxel-level classification, the resulting binary image is called dense volume V_d . We then apply a connectivity analysis to cluster the voxels, which are assumed to be part of catheter or tissue with the catheter-like shape. The voxels from the same cluster are considered to belong to the same model. This means that the RANSAC algorithm includes many redundant processes, if it is applied directly to dense data.¹⁸ As a result, the centerline along the skeleton in each cluster is extracted to construct the sparse volume V_s , which reduces the model-fitting iterations. The centerlines in original SPD-RANSAC are generated directly by filtering the x-ray image, which is the benefit from using the high-contrast imaging. However, the centerlines are hard to extract directly in a coarsely classified 3-D US image. As a result, we propose a new method to extract the centerline for each classified cluster in 3-D US. The steps of centerline extraction are described in Algorithm 1, see Fig. 5 as a result example.

Algorithm 1 Sparse volume generation from a dense volume.

Input: dense volume V_d and empty V_s

find connected clusters in V_d

for each cluster in V_d **do**

PCA analysis is applied to find dominant axis among lat., az., and ax.

for each 2-D slice along dominant axis **do**

find connected 2-D areas in the slice

for each 2-D area in the slice **do**

find center point of the area

project center point to V_s

end for

end for

end for

Output: sparse volume V_s

2.3.2 Model fitting based on sparse and dense volumes

In our method, the catheter is modeled as a curved cylinder, which relies on centerline fitting of the catheter.⁷ Since we are looking for the catheter skeleton, the curved skeleton k can be modified as

$$k = \{\mathbf{r} \in v, \mathbf{r}_0 \in v, t \in \mathbb{R}, \mathbf{h}_0, \mathbf{h}_1 \in \mathbb{R}^3 : \mathbf{r} = \mathbf{r}_0 + t\mathbf{h}_0 + t^2\mathbf{h}_1\}, \quad (8)$$

where the v is the voxel group from the 3-D images, t is the real number, and \mathbf{h}_0 and \mathbf{h}_1 are vectors in 3-D space. For catheter segmentation in 3-D US, the model is fitted by a cubic spline interpolation, which is controlled by three control points.¹⁹ For each RANSAC iteration, three control points are randomly selected from V_s , which are ranked by PCA analysis to define the interpolation order and to model the skeleton. The skeleton with the highest number of inliers in V_d is chosen to be catheter skeleton. The outliers are determined by computing its Euclidean distance to the skeleton. Finally, the inliers together with the skeleton in V_d are regarded as the segmented catheter. With *a-priori* knowledge that the RF-ablation catheter cannot be heavily curved inside the blood chamber, we constrain the curvature by controlling the distance between the middle point to the straight line constructed from the endpoints. The maximum distance is selected as 10 voxels in this paper.

2.3.3 Accuracy of the segmentation

Our method starts with finding the voxels and identifying the catheter inside those voxels. The following major step is the previously discussed model fitting to the classified voxels. The accuracy of the method can be defined as absolute accuracy or a relative accuracy. The definition of absolute accuracy would require a completely calibrated physical setup with predefined phantoms or tissues and reference catheters. In our case, we will define the accuracy as the deviation of the visual ground truth, where the catheter is manually annotated within the image. All the annotations are made by clinical experts.

In order to define the deviation as a distance, we define the skeleton of the catheter in the form of the center line of the catheter. The deviation is then the distance between the annotated center line and the model-fitted catheter. From the model, we obtain a limited set of key points, so that a spline function is used to construct a smooth curve going through the key points.

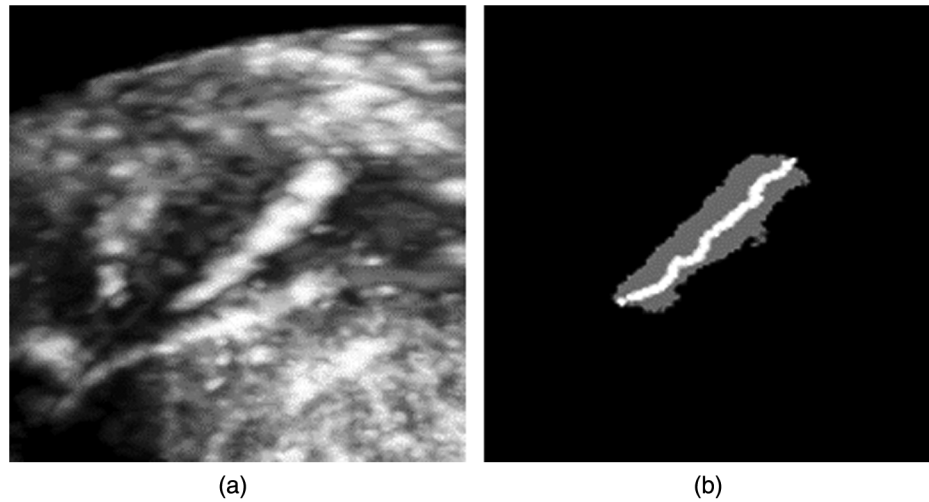


Fig. 5 (a) Example of a catheter in a 3-D image and (b) resulting dense cluster and sparse centerline.

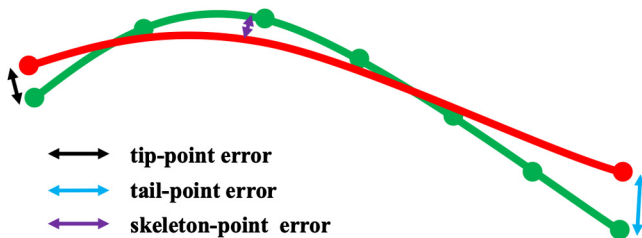


Fig. 6 Example of three errors: tip-point error, tail-point error, and skeleton-point error. The red curve is ground-truth skeleton and green curve is the segmented catheter skeleton.

This approach makes the model-fitted catheter well defined between the end points.

In our case, we define three types of errors: skeleton-point error, and two errors concerning the beginning and ending of the model, i.e., tip-point error and end-point error (the average of tip-point error and tail-point error). These errors are visualized in Fig. 6. The latter two errors are defined as the distance between the detected point with the corresponding ground-truth point, either at the tip or at the tail of the catheter. The skeleton-point error is the distance between the sampled points from the segmented catheter to the ground truth skeleton. All errors are measured visually in the images and are initially expressed in voxels, which can be translated to distance using the voxel resolution. Further details and outcomes can be found in the experiments in Sec. 3.3.

3 Experimental Results

For the experiments, we start with the different datasets from Sec. 3.1. The results on the voxel classification using different features and classifiers are reported in Sec. 3.2. Section 3.3 shows the performance on catheter segmentation using the modified SPD-RANSAC.

3.1 Datasets

To validate the stability of our system, we have collected 3-D US datasets under different recording conditions and performed the experiments. As for the *in-vitro* dataset, a polyvinyl alcohol (PVA) rubber heart was placed into a water tank. The images

were captured by a 3-D transesophageal echocardiography probe (TEE) while an RF-ablation catheter is inserted into it. Due to the less complex structure inside the rubber heart and the absence of anatomical material from a real heart, a clear contrast between catheter and background or phantom wall is shown. For the *ex-vivo* datasets, porcine hearts were placed in several water tanks and images were captured through TEE (HT-CX-TEE and HT-EP-TEE) or a transthoracic echocardiogram probe (TTE, HT-EP-TTE). During the recording, the catheters were inserted into the ventricle or atrium. As for TEE-based images, although they were obtained from a different US system, we obtained a similar US quality due to using the same US probe. However, the HT-EP-TTE was collected by employing a TTE probe, which has lower response frequency leading to a noisy image with low-contrast appearance. Finally, we also collected an *in-vivo* (LH-EP-TTE) dataset on a live porcine. During the recording, the TEE probe was placed next to the beating heart through the open chest, while the RF-ablation catheter was inserted through the vein to approach the heart. Because of challenging recording conditions and an unstable environment, the *in-vivo* dataset has the worst image quality. More detailed meta-data about our datasets are presented in Table 2. All the datasets are manually annotated for catheter locations and confirmed by both medical and technical experts as the groundtruth. In the following experiment, to fully make use of limited datasets, the leave-one-out-cross-validation (LOOCV) is performed on each dataset. Some 2-D slices from different datasets are shown in Fig. 7.

3.2 Voxel-Based Classification

For the voxel classification, both feature and classifier can influence the performance of candidate voxels segmentation. To evaluate the discriminative power of the proposed features, we exploit their performance applying both a single-scale approach and a multiscale approach. Conventional methods, e.g., needle segmentation in 3-D US,^{8,10} only considered a predefined scale size, i.e., single-scale and denoted this by SS- N for a predefined single-scale size N , based on *a-priori* knowledge of the instrument diameter. However, these predefined scales only extract discriminating information of tools while ignoring the information from the anatomical background, such as a heart wall or microvalve inside the heart. To extract more

Table 2 Characterization of 3-D ultrasound volumes for experiments.

Dataset	Recording condition	Catheter diameter	Number of acquisitions	US system	Transducer type and frequency range	Voxel size per dimension	Volume size (lat. × az. × ax.)
PVA-EP-TEE (<i>in-vitro</i>)	Rubber (PVA) heart phantom	2.3 mm ^a	20	EPIQ 7	3-D phase array 2 to 7 MHz	0.4 mm	141 × 168 × 101 to 145 × 185 × 101
HT-CX-TEE (<i>ex-vivo</i>)	Porcine heart	2.3 mm ^a	10	CX50	3-D phase array 2 to 7 MHz	0.4 mm	179 × 175 × 92
TH-EP-TEE (<i>ex-vivo</i>)	Porcine heart	2.3 mm ^b	10	EPIQ 7	3-D phase array 2 to 7 MHz	0.6 mm	120 × 69 × 92 to 193 × 284 × 190
TH-EP-TTE (<i>ex-vivo</i>)	Porcine heart	2.3 mm ^c	12	EPIQ 7	3-D phase array 1 to 5 MHz	0.7 mm	137 × 130 × 122
LH-EP-TTE (<i>in-vivo</i>)	Live porcine	2.3 mm ^c	8	EPIQ 7	3-D phase array 2 to 7 MHz	0.4 mm	146 × 76 × 153 to 172 × 88 × 178

^aAvailable from Chilli II.
^bAvailable from Biosense.
^cAvailable from OSYPKA.

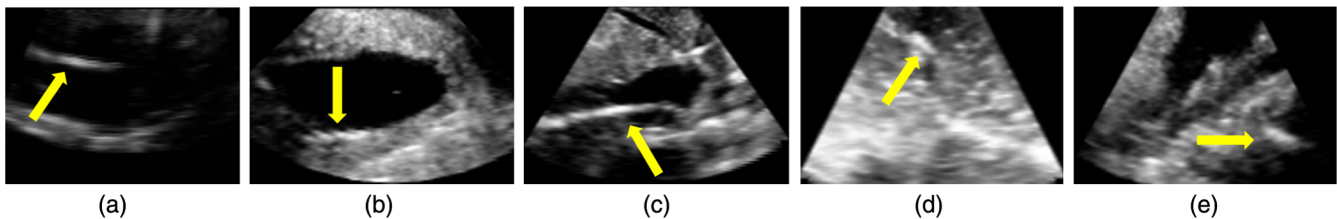


Fig. 7 Examples of 2-D slices from different datasets, the catheter locations are indicated by yellow arrows. (a) PVA-EP-TEE, (b) HT-CX-TEE, (c) HT-EP-TEE, (d) HT-EP-TTE, and (e) LH-EP-TEE.

discriminative information for a better and stable classification, we also employ the multiscale approach which involves different scales simultaneously, e.g., the scale ranges from 1 to N , denoted as MS- N . In the following section, all comparisons are based on AdaBoost classification, due to its optimized performance which is shown in Fig. 12.

3.2.1 Single-scale versus multiscale

Using the features objectness (\mathcal{O}) and Hessian (\mathcal{H}), we have performed experiments with σ ranging from 3 to 15 and step size 4. To measure the scale influence on the features in a simple way, we only employ the precision as a metric, while fixing recall at 75% in each volume. The experimental results are shown in

Figs. 8 and 9 separately. These experiments lead to the following conclusions. (1) The multiscale approach in objectness achieves a higher performance, due to different shape information is contained within the different scale sizes. When considering more scales, the features become more discriminating. (2) When comparing Frangi and objectness features with Hessian features, the latter one has better performance, due to preserving more spatial information without PCA analysis. However, in the PVA-EP-TEE, the objectness gives a higher precision, which can be explained by the high-contrast image quality when compared with real tissue. Meanwhile, in all cases, the Frangi feature achieves a lower precision than objectness.¹⁴

For features \mathcal{S} and \mathcal{L} , similar results are obtained, i.e., when multiscale range is increasing, the classification performance

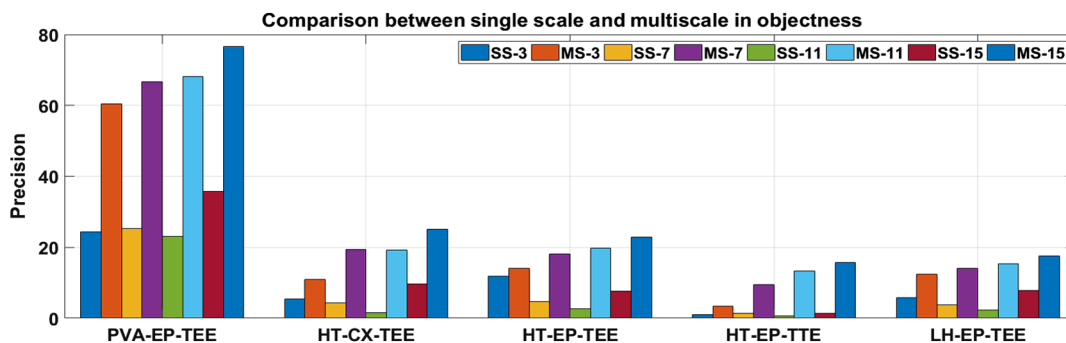


Fig. 8 Average precision of single-scale (SS) and multiscale (MS) objectness.

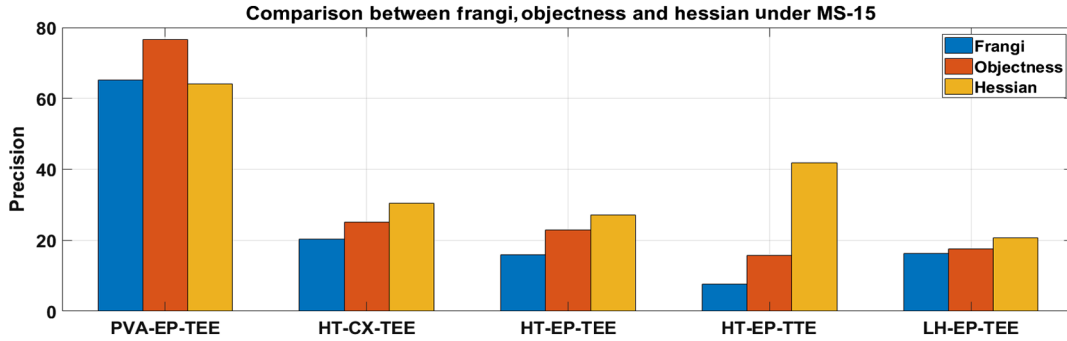


Fig. 9 Average precision of Frangi, objectness, and Hessian features in MS mode.

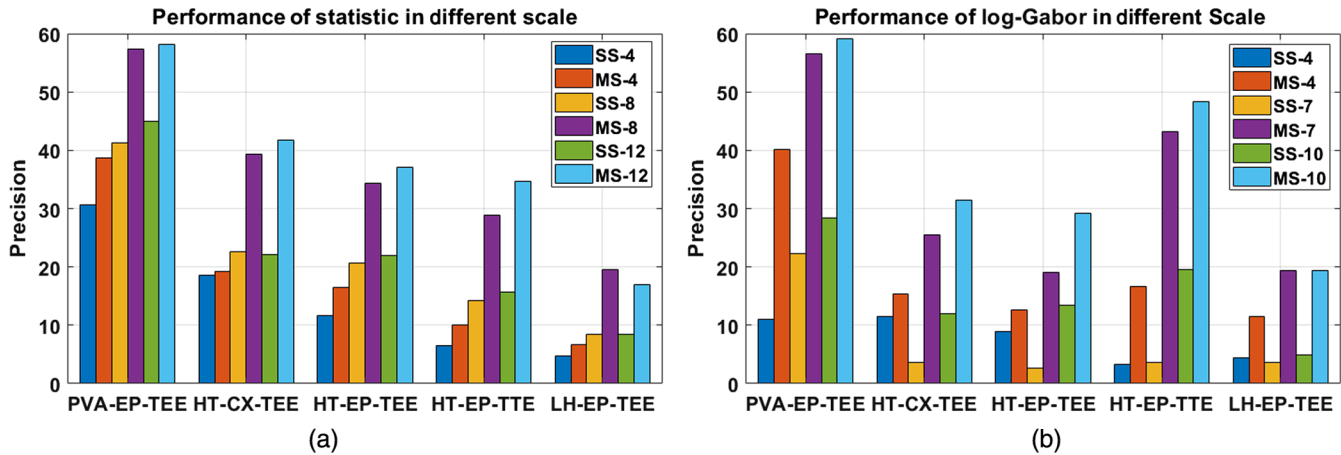


Fig. 10 Average precision of (a) statistic feature and (b) log-Gabor features.

improves and the multiscale approach achieves a higher performance than single-scale operation. We have performed the experiments with S as scale ranging from 4 to 12 with step size 4 and the experiments with L as scale ranging from 4 to 10 with step size 3. Experimental results are shown in Fig. 10. From the experiments, we conclude that the single-scale approach of the Gabor feature in needle detection does not offer sufficient performance for our catheter segmentation in tissue-based images.¹³

Based on the comparison between single scale and multiscale in different feature types, we fixed scale range as MS-15 for

objectness and Hessian feature, MS-12 for statistic feature, and MS-10 for log-Gabor feature in the following section.

3.2.2 Feature comparison and fusion

Based on multiscale approach in different features, their individual and fusion performances in each dataset are shown in Fig. 11. The results are demonstrated under AdaBoost, due to it achieved the best performance when compared with other classifiers (under C and shown in Fig. 12). All the results are obtained by LOOCV and thresholds are tuned to achieve

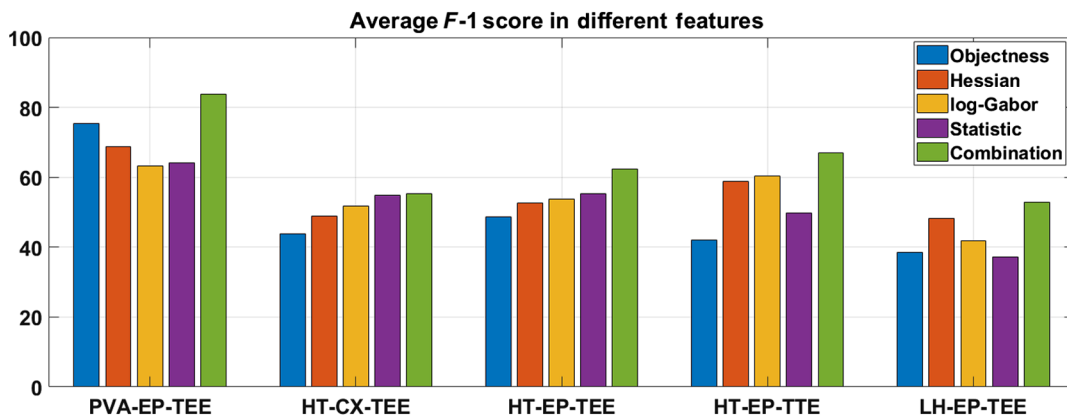


Fig. 11 Optimizing the $F-1$ scores when tuning the thresholds. The combination is the best choice. Corresponding to Table 3.

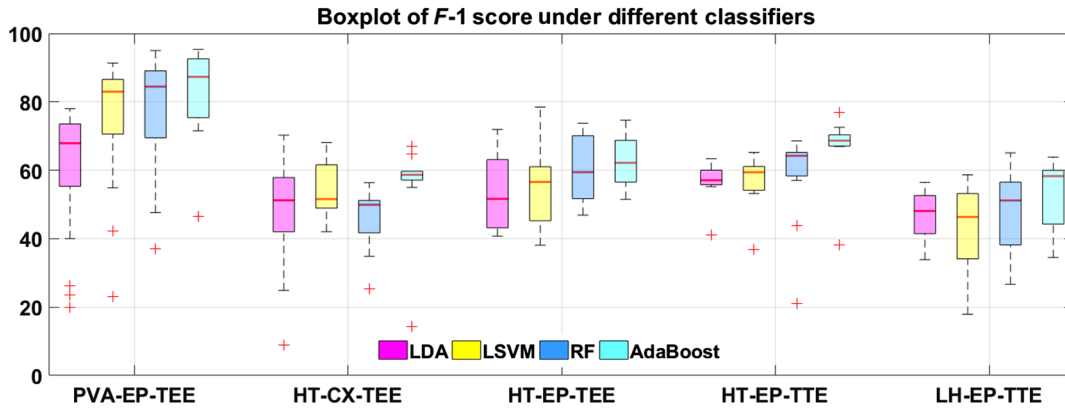


Fig. 12 Distributions of $F-1$ score for different classifiers (LDA, LSVM, RF, and AdaBoost).

the best $F-1$ score on average. More detailed performance information can be referred to Table 3.

From the performances in the table and figures, some observations can be made. For the phantom dataset, having less complexity and higher image contrast, the objectness feature is able

to achieve a promising result with *a-priori*-defined descriptors. For *ex-vivo* datasets using different recording probes and US machines, the complex anatomical structure which has a similar appearance as catheters makes it difficult for the objectness feature to describe the 3-D space information. Moreover, when

Table 3 Average classification performance under best $F-1$ score achieved. Numbers are mean and (standard deviation).

Dataset		Adaptive boosting				
		\mathcal{O}	\mathcal{H}	\mathcal{L}	\mathcal{S}	\mathcal{C}
PVA-EP-TEE (<i>in-vitro</i>)	Recall	80.06 (11.70)	71.05(24.94)	66.82(29.79)	73.16(22.78)	89.77(12.51)
	Precision	75.70 (19.39)	74.25(12.51)	66.90(13.64)	63.48(15.89)	81.58(16.10)
	$F-1$ score	75.44 (12.54)	68.78(15.39)	63.35(21.30)	64.07(14.44)	83.70(11.85)
	Specificity	99.98 (3.3e-4)	99.98(1.6e-4)	99.98(1.5e-4)	99.96(2.7e-4)	99.98(2.0e-4)
HT-CX-TEE (<i>ex-vivo</i>)	Recall	50.31(25.73)	56.40(24.91)	58.46(25.30)	66.36(18.43)	64.12(25.17)
	Precision	48.55(18.92)	51.01(11.77)	53.18 (9.15)	50.22(11.16)	53.76(16.28)
	$F-1$ score	43.80(14.34)	48.87(11.01)	51.84(11.84)	54.80(7.45)	55.24(14.82)
	Specificity	99.93(6.5e-4)	99.94(5.2e-4)	99.95(4.04-4)	99.93(4.4e-4)	99.94(4.3e-4)
HT-EP-TEE (<i>ex-vivo</i>)	Recall	48.97 (9.87)	55.15(16.27)	58.79 (8.84)	60.53(10.62)	70.62(11.70)
	Precision	51.35(15.75)	52.87(14.92)	50.47(12.72)	51.91(9.22)	57.96(11.34)
	$F-1$ score	48.66(10.03)	52.72(13.39)	53.75(10.24)	55.38 (7.75)	62.45 (7.55)
	Specificity	99.91(9.1e-4)	99.94(2.1e-4)	99.92(3.5e-4)	99.92(3.7e-4)	99.93(3.2e-4)
HT-EP-TTE (<i>ex-vivo</i>)	Recall	47.65(11.05)	71.58(19.07)	69.59(19.67)	66.35(11.63)	75.59(17.75)
	Precision	38.27(5.59)	51.14(8.88)	56.49(7.83)	40.75(6.78)	63.79(7.77)
	$F-1$ score	41.99(6.93)	58.78(12.34)	60.47(12.40)	49.72(5.71)	66.95(9.48)
	Specificity	99.95(1.2e-4)	99.96(1.3e-4)	99.97(1.5e-4)	99.94(2.0e-4)	99.97(1.2e-4)
LH-EP-TTE (<i>in-vivo</i>)	Recall	37.67(19.21)	51.22(17.96)	43.25(13.63)	43.82(16.21)	60.64(23.76)
	Precision	43.24(23.54)	47.43(11.44)	41.99(13.43)	32.84(14.32)	52.86(9.42)
	$F-1$ score	38.52(17.96)	48.32(14.72)	41.91(11.92)	37.27(14.77)	52.93(10.65)
	Specificity	99.95(4.0e-4)	99.95(1.4e-4)	99.94(3.7e-4)	99.92(2.3e-4)	99.94(3.6e-4)

Note: Bold characters represent the best performance on average.

Table 4 Average performance of catheter segmentation error mean \pm std. (mm). TE, tip-point error; EE, end-point error; SE, skeleton-point error; R-2, two-point RANSAC; R-3, three-point RANSAC; SR-3, three-point SPD-RANSAC.

Dataset	R-2			R-3			SR-3		
	TE	EE	SE	TE	EE	SE	TE	EE	SE
PVA-EP-TEE	4.0 \pm 2.6	3.9 \pm 1.9	3.0 \pm 1.5	1.8 \pm 0.6	2.1 \pm 0.5	1.8 \pm 0.4	1.4 \pm 0.8	1.4 \pm 0.6	1.5 \pm 0.5
HT-CX-TEE	4.0 \pm 1.8	4.8 \pm 0.9	3.1 \pm 0.7	1.9 \pm 0.4	2.5 \pm 1.3	2.1 \pm 1.1	1.2 \pm 0.3	1.7 \pm 1.0	1.5 \pm 0.6
HT-EP-TEE	9.6 \pm 6.0	10.7 \pm 6.1	6.7 \pm 6.4	3.3 \pm 1.3	3.5 \pm 1.6	3.1 \pm 1.4	3.0 \pm 1.6	3.3 \pm 1.8	3.0 \pm 1.8
HT-EP-TTE	4.0 \pm 1.3	4.3 \pm 1.2	2.9 \pm 0.6	2.1 \pm 0.4	2.2 \pm 0.3	2.0 \pm 0.1	2.1 \pm 0.5	1.9 \pm 0.4	1.8 \pm 0.2
LH-EP-TTE	3.9 \pm 2.8	5.4 \pm 1.2	3.7 \pm 0.6	2.0 \pm 0.9	2.0 \pm 0.5	1.7 \pm 0.3	2.4 \pm 2.8	2.4 \pm 1.4	1.9 \pm 0.8
Average error	5.0 \pm 3.7	5.5 \pm 3.6	3.7 \pm 2.2	2.1 \pm 0.9	2.4 \pm 1.0	2.1 \pm 0.9	1.9 \pm 1.4	2.0 \pm 1.2	1.9 \pm 1.0

Note: Bold characters represent the best performance on average.

PCA is introduced, more spatial details are lost. Both Hessian features and log-Gabor features perform similarly in *ex-vivo* datasets, which may be explained by exploiting orientation and scale-sensitive features to describe the spatial information. For the statistic feature, although it can extract 3-D local intensity distribution information, the performance has less stability when compared with Hessian/log-Gabor features. For the *in-vivo* dataset, due to challenging recording conditions and low-contrast image quality from real blood in the blood pool, the performances of all features are decreased. Although the log-Gabor feature introduces more orientation information, due to the low image contrast and the blurry boundary of the catheter, the orientation information cannot improve the classification performance. In all datasets, the feature combination is able to further improve the classification performance and appears to be the best choice.

3.3 Catheter Segmentation by Model Fitting

After voxel-based classification, the SPD-RANSAC is applied to the binary images to segment the catheter in the noisy segmented images. The RANSAC algorithm generates the end points and the skeleton of the catheter, which is used to analyze the segmentation error when compared to the ground truth. To evaluate the segmentation accuracy, we consider three types of errors: tip-point error (TE), end-point error (the average of tip-point error and tail-point error, EE), and skeleton-point error (SE). Like commonly considered, we regard the farthest point from the image border between the two end-points of

the catheter as the tip.²⁰ The skeleton error is the average distance of five equally sampled points (except the endpoints) on the identified skeleton to the annotated skeleton. For each sampled point, its distance to the ground truth is measured. An example of the three different error types is shown in Fig. 6

The segmentation performances from Table 4 are expressed in millimeter (mm) and involve three different model-fitting methods: (1) RANSAC with the two-point catheter model (R-2), (2) RANSAC with the three-point model (R-3), and (3) SPD-RANSAC with the three-point model (SR-3). Several of slices from tissue images are visualized in Fig. 13. The segmented catheters are overlaid with colored annotations. To directly visualize in a 3-D volume, the corresponding 3-D images are shown in Fig. 14. Furthermore, Fig. 15 shows an example of comparing our three-point SPD-RANSAC with a two-point RANSAC model fitting.^{7,13}

As shown in Table 4, the three-point model (R-3 and SR-3) is able to segment the catheters accurately when compared with the two-point methods (R-2). This is evident because almost every catheter is curved in the image, even a slightly curvature occurs when compared with the needle segmentation in the image. Meanwhile, the SPD-RANSAC is able to improve the segmentation accuracy when compared with R-3, which directly applied the model fitting to the classified volume. As a result, our three-point SPD-RANSAC can achieve a higher segmentation performance giving an average segmentation tip-point error of only 1.9 mm.

From the results in classification performance, multiscale together with feature fusion is robust to classify the catheter

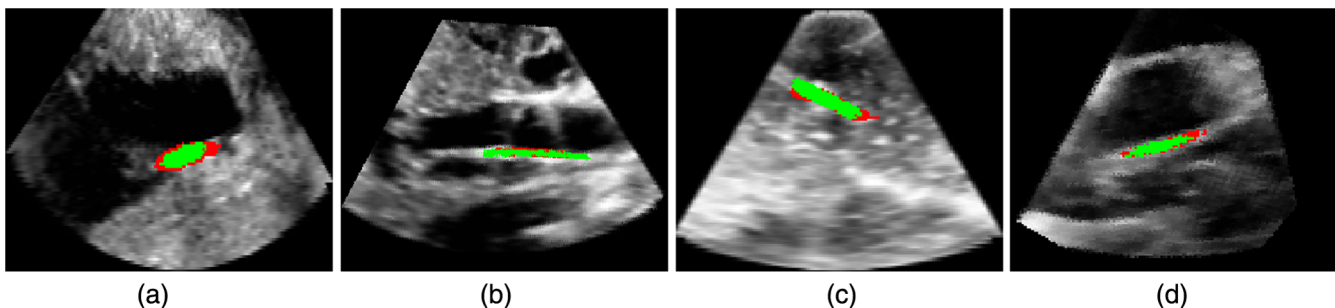


Fig. 13 Slices (cropped) from real heart volumes, red represents annotation, and green represents fitted catheter. (a) HT-CX-TEE, (b) HT-EP-TEE, (c) HT-EP-TTE, and (d) LH-EP-TTE.

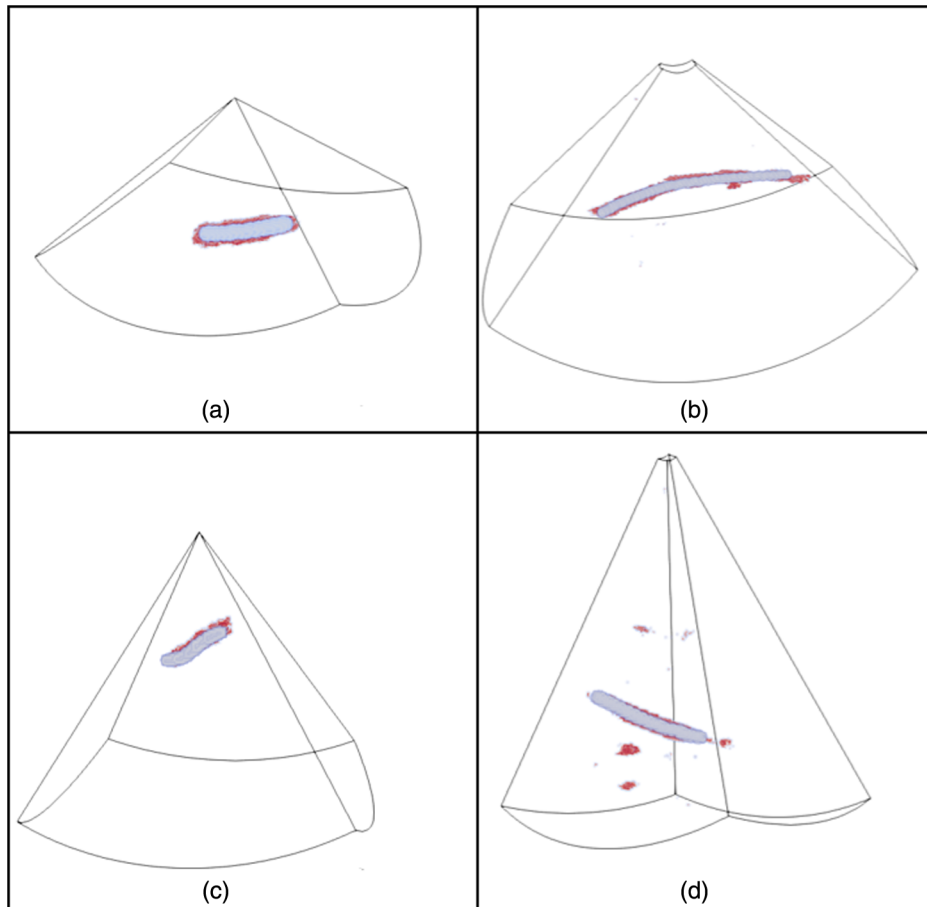


Fig. 14 Classification results in tissue volumes, prediction (red) versus annotation (white). (a) HT-CX-TEE, (b) HT-EP-TEE, (c) HT-EP-TTE, and (d) LH-EP-TTE.

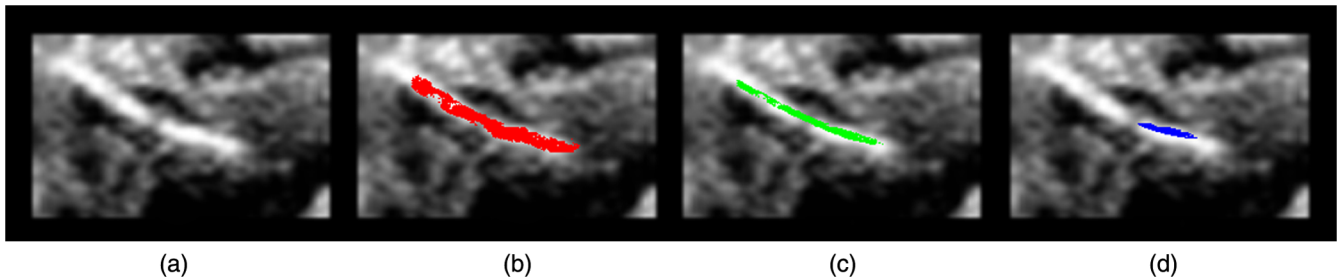


Fig. 15 Comparison between SPD-RANSAC and a simple model-based RANSAC. (a) Original image, (b) annotation, (c) SPD-RANSAC, and (d) two-points RANSAC.

voxels using AdaBoost classifiers. Although the classified volumes include some false positive (as shown in Fig. 14), *a-priori* knowledge of the catheter shape leads to a correct segmentation result. When an optimized 3-D view generation would be implemented and added to our algorithms or alternatively, a 2-D view on slices would be created, the catheter can be easily found and annotated for surgeons such that the cardiac intervention becomes easier and obtains a higher safety.

4 Conclusion and Discussion

In this paper, we have jointly studied different features in-depth and combined them with different classifiers. Based on a model-fitting method, our quantitative analysis on accuracy shows that the catheter segmentation errors are smaller than the

catheter diameter. With the proposed method, it is possible to automatically segment the catheter in a 3-D US image and demonstrate the optimal viewing of the catheter for surgeons during an intervention.

It was found that all features are interesting but capture only a part of the catheter's discriminating information, due to the challenging imaging conditions, noise, and anatomical structures. Therefore, the combination of the proposed features yields a highest performance in finding the catheter when using the AdaBoost classifier. With model fitting, three-point SPD-RANSAC achieves an average segmentation error of three to four voxels at the voxel-level. Moreover, the total execution time of the whole processing chain is ranging from 2 to 30 min (around nine minutes on the average), based on the volume

size. The experiments were performed on a Xeon CPU running at 3.6 GHz without any code optimization or acceleration.

Based on the accurate segmentation of the catheter in the 3-D US image, there are several ways to exploit this catheter segmentation method for enhancing and facilitating the operation during cardiac interventions.

First, the advantage of having a 3-D US image available can provide a richer spatial information than conventional 2-D x-ray. As a result, an accurate catheter segmentation enables to enhance the clinical procedures requiring more spatial information. For example, for percutaneous aortic valve replacement (so-called TAVI), it is difficult for surgeons to insert the guide wires to pass through the aortic valve under the guidance of x-ray images. Alternatively, if the catheter would be accurately segmented in 3-D US and could be reconstructed into a 3-D heart model, the richer and more accurate spatial information between instrument and tissue would help the physicians to better understand the procedure. Such an approach would definitely enhance the intervention navigation.

Second, there may be new solutions to visualize surgery in the future using US, such as constructing a real-time heart model with a catheter inside it. With accurate catheter segmentation, the segmented catheter with tips can be reconstructed into the model, which can enhance the doctors' understanding of the operation.

Moreover, the catheter segmentation in 3-D US can be also beneficial to 2-D US visualization. The 2-D slices are commonly used during the procedure to guide the catheter. However, it would enforce sonographers to spend too much time of tuning the slice to visualize the instrument. With an accurate catheter segmentation in 3-D US, which can provide accurate positioning information in the 3-D image, an automated slicing technique can be performed to extract the slices containing the catheter. As a result, sonographers can easily perceive these slices with the catheter inside. The 3D image may have lower spatial resolution and more noise than the 2-D image, so that automatic slicing and visualization into 2-D slices would increase the visual perception and benefit to surgeons.

Further improvements are possible for higher segmentation performance and creating a real-time application. For example, tuning the US system to address varying recording conditions, e.g., adapting image gain or focal depth of the US array, may lead to better segmentation performance and higher robustness. Moreover, for different US resolution and catheter appearance, the multiscale with feature fusion (e.g., more features²¹) approach may be simplified or extended to achieve a better and robust segmentation accuracy. With respect to the real-time application, the main challenge is coming from complex feature extraction during the voxel-level classification, which takes more than 85% of whole processing time. Some possible solutions for enhancing the processing speed are (1) embedding the feature extraction steps in a parallel manner on a GPU, which accelerates the computation efficiency, (2) classification at voxel level can be accelerated by a coarse-to-fine strategy to reduce the calculation complexity, and (3) a 3-D US video stream using a high US frame rate system enables to combine the catheter segmentation algorithm with frame skipping for reducing the computational complexity.

In future work, more datasets especially images from human bodies are required for a robust evaluation for a possible clinical application. With more datasets, developed deep learning might

be combined with an ensemble classification scheme to enhance the classification accuracy.²²

Disclosures

The research and this paper have no conflict of commercial interest.

Acknowledgments

This research was conducted in the framework of "Impulse-2 for the healthcare flagship" at Eindhoven University of Technology in collaboration with Catharina Hospital Eindhoven and Royal Philips. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. This article does not contain patient data.

References

1. C. Nadeau et al., "Intensity-based visual serving for instrument and tissue tracking in 3D ultrasound volumes," *IEEE Trans. Autom. Sci. Eng.* **12**(1), 367–371 (2015).
2. X. Guo et al., "Photoacoustic active ultrasound element for catheter tracking," *Proc. SPIE* **8943**, 89435M (2014).
3. M. G. Linguraru et al., "Statistical segmentation of surgical instruments in 3-D ultrasound images," *Ultrasound Med. Biol.* **33**(9), 1428–1437 (2007).
4. M. Abofazel et al., "A new scheme for curved needle segmentation in three-dimensional ultrasound images," in *IEEE Int. Symp. Biomedical Imaging: From Nano to Macro*, IEEE, pp. 1067–1070 (2009).
5. M. Barva et al., "Parallel integral projection transform for straight electrode localization in 3-D ultrasound images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **55**(7), 1559–1569 (2008).
6. K. Cao, D. Mills, and K. A. Patwardhan, "Automated catheter detection in volumetric ultrasound," in *IEEE 10th Int. Symp. Biomedical Imaging*, IEEE, pp. 37–40 (2013).
7. M. Uherčík et al., "Model fitting using RANSAC for surgical tool localization in 3-D ultrasound images," *IEEE Trans. Biomed. Eng.* **57**(8), 1907–1916 (2010).
8. M. Uherčík et al., "Line filtering for surgical tool localization in 3D ultrasound images," *Comput. Biol. Med.* **43**(12), 2036–2045 (2013).
9. A. F. Frangi et al., "Multiscale vessel enhancement filtering," *Lect. Notes Comput. Sci.* **1496**, 130–137 (1998).
10. Y. Zhao, C. Cachard, and H. Liebgott, "Automatic needle detection and tracking in 3D ultrasound using an ROI-based RANSAC and Kalman method," *Ultrason. Imaging* **35**(4), 283–306 (2013).
11. A. Pourtaherian et al., "Gabor-based needle detection and tracking in three-dimensional ultrasound data volumes," in *IEEE Int. Conf. Image Processing*, IEEE, pp. 3602–3606 (2014).
12. A. Pourtaherian et al., "Multi-resolution Gabor wavelet feature extraction for needle detection in 3D ultrasound," *Proc. SPIE* **9875**, 987513 (2015).
13. A. Pourtaherian et al., "Medical instrument detection in 3-dimensional ultrasound data volumes," *IEEE Trans. Med. Imaging* **36**, 1664–1675 (2017).
14. H. Yang et al., "Feature study on catheter detection in three-dimensional ultrasound," *Proc. SPIE* **10576**, 105760V (2018).
15. L. Antiga, "Generalizing vesselness with respect to dimensionality and shape," *Insight J.* **3**, 1–14 (2007).
16. R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proc. 23rd Int. Conf. Machine Learning*, ACM, pp. 161–168 (2006).
17. I. Hacıhaliloglu et al., "Bone segmentation and fracture detection in ultrasound using 3D local phase features," *Lect. Notes Comput. Sci.* **5241**, 287–295 (2008).
18. C. Papalazarou, P. H. de With, and P. Rongen, "Sparse-plus-dense-RANSAC for estimation of multiple complex curvilinear models in 2D and 3D," *Pattern Recognit.* **46**(3), 925–935 (2013).
19. C. De Boor et al., *A Practical Guide to Splines*, Vol. **27**, Springer-Verlag, New York (1978).

20. P. Ambrosini et al., "Fully automatic and real-time catheter segmentation in x-ray fluoroscopy," *Lect. Notes Comput. Sci.* **10434**, 577–585 (2017).
21. T. Tan et al., "Computer-aided detection of breast cancers using Haar-like features in automated 3D breast ultrasound," *Med. Phys.* **42**(4), 1498–1504 (2015).
22. T. Tan et al., "Computer-aided detection of cancer in automated 3-D breast ultrasound," *IEEE Trans. Med. Imaging* **32**(9), 1698–1706 (2013).

Hongxu Yang received his bachelor's degree in electrical engineering from Tianjin University and Nankai University, Tianjin, China, in 2014. In 2016, he received his master's degree in electrical engineering from Eindhoven University of Technology (TU/e), Eindhoven, The Netherlands. He is currently pursuing his PhD in the signal processing systems and video coding and architectures research group (SPS-VCA) at TU/e.

Caifeng Shan is a senior scientist and project leader with Philips Research, Eindhoven, The Netherlands. His research interests include computer vision, pattern recognition, image and video analysis, machine learning, and biomedical imaging. He has authored more than 80 scientific publications and 50 patent applications. He has been associate editor and guest editor of many scientific journals, and served as a program committee member and reviewer for numerous international conferences and journals.

Arash Pourtaherian received his BSc degree in electrical engineering jointly from Indiana University-Purdue University Indianapolis,

Indiana, USA, and the University of Tehran, Iran, in 2010. He then moved to the Netherlands, where he studied his MSc degree in electrical engineering at the Eindhoven University of Technology (TU/e). He received his PhD in 2018 from the video coding and architectures, signal processing systems group at the Electrical Engineering faculty.

Alexander F. Kolen received his PhD from the Institute of Cancer Research from the University of London, United Kingdom, in 2003. His research work involved the ultrasound diagnosis of liver tumors, followed by monitoring high intensity focused ultrasound treatment based on ultrasound elastography. In 2004 to 2005, he held a post-doc position at Maastricht University, the Netherlands, focusing on the analysis of cardiac mechanic using tissue Doppler ultrasound. In 2003, he joined Philips Research, leading several cardiac ultrasound-related projects.

Peter H. N. de With received his PhD from Delft University of Technology in 1992. From 1984 to 1997, he was a senior scientist at Philips Research. From 1997 to 2000, he was a full professor at the University of Mannheim in Germany. He joined LogicaCMG in Eindhoven as a principal consultant in 2000. In 2011, he was appointed full professor at TU/e and scientific director of the Centre for Care and Cure Technologies and Board member of the SA Health team. He is the fellow of the IEEE and (co-)recipient of multiple papers awards of the IEEE CES, VCIP and Transactions papers and a Eurasip Signal Processing award.