

One-Class Acoustic Characterization Applied to Blood Detection in IVUS

Sean M. O'Malley¹, Morteza Naghavi², and Ioannis A. Kakadiaris¹

¹ Computational Biomedicine Lab, University of Houston; Houston, TX

² Association for Eradication of Heart Attack; Houston, TX

Abstract. Intravascular ultrasound (IVUS) is an invasive imaging modality capable of providing cross-sectional images of the interior of a blood vessel in real time and at normal video framerates (10-30 frames/s). Low contrast between the features of interest in the IVUS imagery remains a confounding factor in IVUS analysis; it would be beneficial therefore to have a method capable of detecting certain physical features imaged under IVUS in an automated manner. We present such a method and apply it to the detection of blood. While blood detection algorithms are not new in this field, we deviate from traditional approaches to IVUS signal characterization in our use of 1-class learning. This eliminates certain problems surrounding the need to provide “foreground” and “background” (or, more generally, n -class) samples to a learner. Applied to the blood-detection problem on 40 MHz recordings made *in vivo* in swine, we are able to achieve $\sim 95\%$ sensitivity with $\sim 90\%$ specificity at a radial resolution of $\sim 600 \mu\text{m}$.

1 Introduction

Intravascular ultrasound (IVUS) is currently the gold-standard modality for intravascular imaging. However, the use of this imagery has been hampered by the fact that some tasks remain difficult as a result of the indistinguishability of certain features under IVUS. To help alleviate these problems, a number of computational methods have been proposed over the last decade which aim to detect or characterize one or more features in the IVUS image (e.g., plaque components [1], contrast agents [2], or blood [3]). Digitally enhancing the contrast between a feature and its background allows easier manual interpretation as well as improved computer-aided analysis. For instance, blood detection may serve as a pre-processing step for segmentation of the luminal border.

Our contribution in this paper is to investigate the feasibility of 1-class learning to the problem of distinguishing a single feature imaged under IVUS. In particular, we apply this to the problem of blood detection. The primary advantage to our approach is the fact that “background” samples need never be provided. In our case, as well as others in the field, the background can consist of a wide variety of other imaged tissues. As such, providing suitable background samples for training may be labor-intensive and subjective. With 1-class learning, we circumvent this problem by ignoring the background during training. Instead,

training only requires samples of the foreground class which, in general, can be obtained relatively easily from expert annotations. In practical terms, our method has an advantage over luminal-border segmentation methods in that it is not inherently tied to only detecting luminal blood. For example, we intend to apply this framework to detect extra-luminal blood as well, specifically in the microvascular network known to feed the coronary arteries and plaques [4].

An apparent disadvantage to our approach is that it need not necessarily work at all: as the learner is never exposed to negative examples, it could naively classify every sample as positive and give a 100% true-positive rate along with a 100% false-positive rate. Features which provide little distinguishing power between foreground and background will also result in a high false-positive rate. This being the case, our study has two goals: to describe how the recognizer framework may be applied to problems such as that of blood detection under ultrasound, and to examine specific features for accomplishing this. In Sec. 2 we provide background on the problems surrounding our task. In Sec. 3, we discuss our contribution. We conclude with results (Sec. 4) and a discussion (Sec. 5).

2 Background

Intravascular ultrasound: The IVUS catheter consists of either a solid-state or a mechanically-rotated transducer which transmits a pulse and receives an acoustic signal at a discrete set of angles over each radial scan. Commonly, 240 to 360 such one-dimensional signals are obtained per (digital or mechanical) rotation. The envelopes of these signals are computed, log-compressed, and then geometrically transformed to obtain the familiar disc-shaped IVUS image (Fig. 1). However, most of our discussion will revolve around the original polar representation of the data. That is, stacking the 1-D signals we obtain a 2-D frame in polar coordinates. Stacking these frames over time, we obtain a 3-D volume $I(r, \theta, t)$ where r indicates radial distance from the transducer, θ the angle with respect to an arbitrary origin, and t the time since the start of recording (i.e., frame number). The envelope and log-compressed envelope signals we will represent by I_e and I_l respectively. Note that I contains real values while I_e and I_l are strictly non-negative. The I_l signal represents the traditional method of visualizing ultrasound data, in which log compression is used to reduce the dynamic range of the signal in order for it to be viewable on standard hardware. This signal is the basis for texture-based characterization of IVUS imagery. The signal I has a large dynamic range and retains far more information, including the frequency-domain information lost during envelope calculation. This “raw” signal is the basis for more recent radiofrequency-domain IVUS studies [1, 5].

One-class learning: The backbone of our method is the 1-class support vector machine (SVM); a widely-studied 1-class learner or “recognizer.” The problem of developing a recognizer for a certain class of objects can be stated as a problem of estimating the (possibly high-dimensional) PDF of the features characterizing those objects, then setting a probability threshold which separates in-class

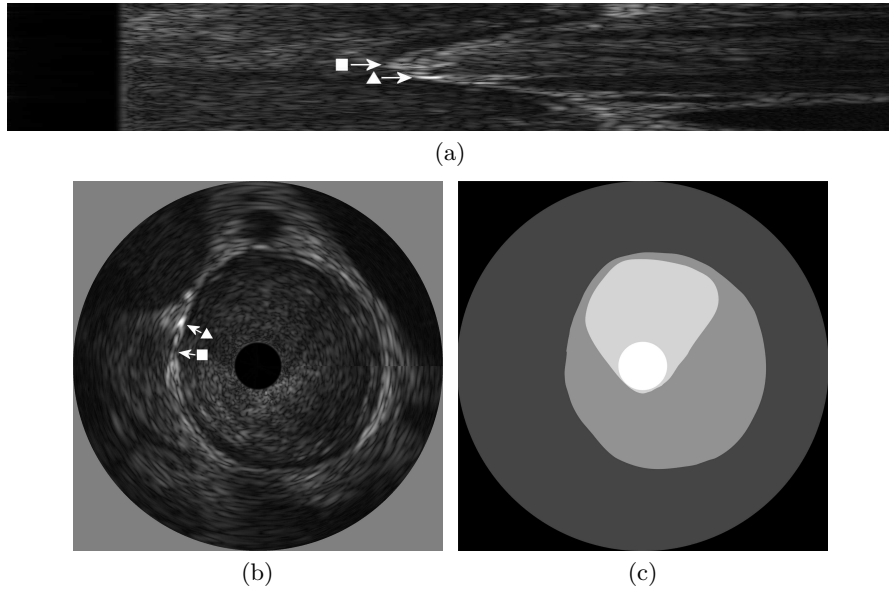


Fig. 1. (a) The log-compressed envelope of the IVUS signal in polar format. The r axis is horizontal (the origin being at the left, at the catheter) and the θ axis vertical (of arbitrary origin). (b) The same signal after Cartesian transformation. The arrows marked \triangle and \square (provided for orientation only) are positioned similarly in the polar and Cartesian spaces. (c) Diagram of the features of interest (from the center outward: the catheter, blood, plaque, and adventitia and surrounding tissues).

objects from all other out-of-class objects. This threshold is necessary since, as learning does not make use of out-of-class examples, the in-class decision region could simply cover the entire feature space, resulting in 100% true- and false-positive rates. Following the approach of Schölkopf *et al.* [6], we denote this threshold as $\nu \in (0, 1)$. We note that as the learner is never penalized for false positives (due to its ignorance of the negative class), it is essential that the PDF's of the positive and negative classes be well-separated in the feature space.

The other parameter of interest is the width, γ , of the SVM radial basis function (i.e., $k(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2)$ for a pair of feature vectors \mathbf{x} and \mathbf{x}'). Properties of a good SVM solution include an acceptable classification rate as well as a low number of resulting support vectors. A high number of support vectors relative to the number of training examples is not only indicative of overfitting, but is computationally expensive when it comes to later recognizing a sample of unknown class. A further discussion of the details of SVM operation is outside the scope of this paper; the unfamiliar reader is encouraged to consult the introduction by Hsu *et al.* [7].

3 Materials & Methods

3.1 Data Acquisition and Ground Truth

Ungated intravascular ultrasound sequences were recorded at 30 frames/s *in vivo* in the coronary arteries of five atherosclerotic swine. The IVUS catheter’s center frequency was 40 MHz. Each raw digitized frame set $I(r, \theta, t)$ consists of 1794 samples along the r axis, 256 angles along the θ axis, and a variable number of frames along t (usually several thousand). The envelope I_e and log-envelope I_l signals were computed offline for each frame.

For training and testing purposes, a human expert manually delineates three boundaries in each image: one surrounding the IVUS catheter, one surrounding the lumen, and one surrounding the outer border of the plaque (as in Fig. 1(c)). The blood within the lumen is used as the positive class in training and testing. As our goal is to separate blood from all other physical features, we use the relatively blood-free tissue of the plaque as the negative class in testing. For most of these studies we have ignored the adventitia and surrounding tissues; in many cases this region contains blood-containing vessels and/or is difficult to reliably interpret.

3.2 Features

We investigate two classes of features: those intended to quantify speckle (i.e., signal randomness in space and time) and those based on frequency-domain spectral characterization. The former are traditionally used for blood detection and the latter for tissue characterization. These features³ are defined for a 3-D signal window of dimensions $r_0 \times \theta_0 \times t_0$ as follows:

$$F_\alpha = \frac{1}{r_0\theta_0} \sum_{i=1}^{r_0} \sum_{j=1}^{\theta_0} \text{stddev}[I(i, j, \cdot)] \quad (1)$$

$$F_\beta = \frac{1}{r_0\theta_0t_0} \sum_{i=1}^{r_0} \sum_{j=1}^{\theta_0} \sum_{k=1}^{t_0} |I(i, j, k)| \quad (2)$$

$$F_\delta = \frac{1}{r_0\theta_0} \sum_{i=1}^{r_0} \sum_{j=1}^{\theta_0} \text{corr}[I(i, j, \cdot)] \quad (3)$$

$$F_\epsilon = \frac{1}{t_0} \sum_{k=1}^{t_0} \text{stddev}[I(\cdot, \cdot, k)] \quad (4)$$

$$F_\zeta = \sum_{i=1}^{\lceil r_0/2 \rceil} \sum_{j=1}^{\lceil \theta_0/2 \rceil} \sum_{k=1}^{\lceil t_0/2 \rceil} ijk \hat{I}(i, j, k) \quad (5)$$

$$F_\eta = \frac{F_\zeta}{\sum_{i=1}^{\lceil r_0/2 \rceil} \sum_{j=1}^{\lceil \theta_0/2 \rceil} \sum_{k=1}^{\lceil t_0/2 \rceil} \hat{I}(i, j, k)} \quad (6)$$

$$F_\iota = \text{FFT}\{\text{mean_signal}[I]\}, \quad (7)$$

³ These features were inspired by the principles behind temporal averaging in ultrasound (e.g., [8]) and tissue characterization (e.g., [1]).

where $\text{stddev}(\cdot)$ returns the sample standard deviation of the samples in its argument and $\text{corr}(\cdot)$ returns the correlation coefficient of its argument compared to a linear function (e.g., a constant signal), returning a value on $[-1, +1]$. The function \hat{I} indicates the magnitude of the Fourier spectrum of I . Function $\text{FFT}(\cdot)$ computes the magnitude of the Fourier spectrum of its vector input (the vector result will be half the length of the input due to symmetry) and $\text{mean_signal}(\cdot)$ takes the mean of the θt IVUS signals in the window, producing one averaged 1-D signal.

The features represent measures of temporal (F_α and F_δ) and spatial (F_ϵ) speckle, a measure of signal strength (F_β), measures of high-frequency signal strength (F_ζ and, normalized by total signal strength, F_η), and a vector feature consisting of the raw backscatter spectrum (F_l). In practice, this final feature is windowed to retain only those frequencies within the catheter bandwidth (~ 20 -60 MHz in our case). Each feature, with the exceptions of (F_ζ , F_η , F_l), are computed on I_e and I_l in addition to I . Hence, features (F_α , F_β , F_δ , F_ϵ) actually consist of vectors of three values. Feature (F_l) consists of a vector that varies according to the sampling rate and bandwidth of the IVUS system.

Samples are extracted by setting a fixed window size (r_0 , θ_0 , t_0) and, from a set of consecutive IVUS frames (i.e., a volume) for which associated manually-created masks are available, placing the 3-D window around each sample in the volume. If this window does not overlap more than one class, the above features are computed for that window and associated with the class contained by it. To improve the scaling of the feature space, each feature of the samples used for training are normalized to zero mean and unit variance. The normalization values are retained for use in testing and deployment.

3.3 Training & Testing Scheme

In general, given a set of positive S_+ and negative S_- samples (from the lumen and plaque respectively), which typically represent some subset of our seven features, a grid search over γ and ν is performed to optimize a one-class SVM (Fig. 2). Optimization in this case aims to obtain an acceptable true positive rate on S_+ , true negative rate on S_- , and low number of support vectors. In order to avoid bias, at every (γ, ν) point on the grid, 5-fold cross-validation is used. That is, the recognizer is trained on one-fifth of S_+ and tested on the remaining four-fifths of S_+ and all of S_- (the negative class is never used in training).

As feature selection is especially critical in a one-class training scenario, we will gauge the performance of each feature individually. More elaborate feature selection schemes (e.g., genetic algorithms) could be employed, but as one of our goals is to determine which features individually best characterize the blood, we will not investigate this issue here.

4 Results

In each case described here, thirty frames were manually segmented into “blood” and “non-blood” for training and testing purposes. (This number was deemed

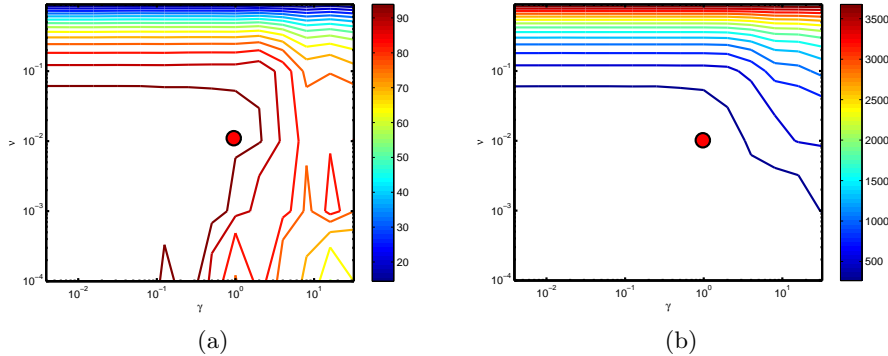


Fig. 2. SVM optimization over ν and γ . Contour maps represent (a) blood true-positive rate and (b) support vector count. The marker at $(\gamma = 1, \nu = 0.01)$ indicates a true-positive rate for blood of 97.1% and 101 support vectors. Similar plots (not shown) are made to show the false-positive rate in the plaque, in order to aid optimization.

sufficient as the backscatter properties of the blood should not change significantly over the course of recording.) For each of our seven features, we obtained the best possible results using the training method described previously. Specifically, we chose the parameters ν and γ such that there was a true-positive rate (sensitivity) of $\sim 98\%$, where possible, and a minimal false-positive rate. (The 98% value was chosen arbitrarily to be a reasonable level.) The number of support vectors at this point is indicative of the generalization power of the feature; however, this information was not used for optimization purposes except as a sanity check, as a large number of support vectors is indicative of overfitting. A final parameter to be mentioned is the window size for feature extraction. In previous experiments we determined a reasonable tradeoff between window size and spatial accuracy to be $(r_0, \theta_0, t_0) = (255, 13, 13)$; this equates to a radial resolution of $\sim 600 \mu\text{m}$, angular resolution of $\sim 18^\circ$, and temporal resolution of $\sim 0.4 \text{ s}$. In all experiments presented here we use this window size.

We describe the results from one case in detail and two in summary. For Case 1, Table 1 summarizes the results for each feature for a typical sequence. To determine whether the performance of a particular feature was mainly due to that feature's application to a specific form of the data (i.e., either the raw signal, its envelope, or its log-compressed envelope), this table also lists the results of subdividing three of the highest-accuracy features into their components and performing experiments on these alone. Results from additional cases were similar as we employed the same hardware and the same type of subject (i.e., swine). In Case 2 as in Case 1, we classified blood against non-blood (i.e., plaque tissue) to obtain results of 97%/82.3%/2.8% (sensitivity, specificity, support vector fraction). In Case 3, we classified blood against manually-selected adventitial tissue to obtain results of 95.3%/100%/4%.

Feature	TP	FP	TN	FN	Sensitivity	Specificity	SV (%)
F_α	8644	705	8334	93	98.9	92.2	106 (1.2)
F_β	8727	3868	5171	10	99.9	57.2	17 (0.2)
F_δ	8649	8796	243	88	99.0	2.69	102 (1.2)
F_ϵ	8716	2	9037	21	99.8	100	33 (0.4)
F_ζ	8653	1264	7775	84	99.0	86.0	98 (1.1)
F_η	8600	2334	6705	137	98.4	74.2	246 (2.8)
F_l	5010	27	9012	3727	57.3	99.7	8083 (92.5)
F_α^*	8404	3446	5593	333	96.2	61.9	271 (3.1)
F_α^\dagger	8064	2811	6228	673	92.3	68.9	391 (4.5)
F_α^\ddagger	8094	2552	6487	643	92.6	71.8	373 (4.3)
F_ϵ^*	7838	3488	5551	899	89.7	61.4	271 (3.1)
F_ϵ^\dagger	7623	2962	6077	1114	87.2	67.2	187 (2.1)
F_ϵ^\ddagger	7542	1860	7179	1195	86.3	79.4	191 (2.2)
F_ζ^*	8576	3241	5798	161	98.2	64.1	163 (1.9)
F_ζ^\dagger	8591	3264	5775	146	98.3	63.9	160 (1.8)
F_ζ^\ddagger	8417	3277	5762	320	96.3	63.7	147 (1.7)

Table 1. Statistics relating the classification accuracy obtained by each feature with respect to true/false (T/F) positives/negatives (P/N). Positive/negative examples used: 8737/9039. Sensitivity is defined as $TP/(TP + FN)$; specificity as $TN/(TN + FP)$. Support vectors (SV) are listed as an absolute value and as a percentage of the number of (positive) examples used for training. Also shown are statistics relating the classification accuracy obtained by features F_α , F_ϵ , and F_ζ when they are applied to only one type of signal: the original*, envelope[†], and log-envelope[‡].

It should be noted that in all cases, the blood content of the extra-luminal tissue we compare against can affect specificity in an unknown manner. The high specificity of Case 3 is probably due to our choosing as our non-blood class extra-luminal tissues which were likely to be blood-free. In Cases 1 and 2 we used the entire plaque region as our negative class, in spite of the fact that this region may in reality contain detectable blood.

5 Discussion and Conclusion

Our highest performance was obtained using features which attempt to directly measure the amount of variability (“speckle”) present in the signal, either temporally (F_α), spatially (F_ϵ), or in the frequency domain (F_ζ , F_η). Direct learning from the Fourier spectrum tended to perform poorly (F_l). This is likely because one-class learning is ill-suited to determining the subtle differences in frequency spectra between the backscatter of various features imaged under ultrasound.

The performance of these features as applied to a single signal type (e.g., F_α^*) tended to be poorer than the results obtained otherwise (e.g., F_α). However, this trend does not extend to increased performance when a larger number of features

are combined during training. For instance, we found that using all features except F_ζ together results in prohibitively poor specificity (<20%). This is an expected result for one-class SVMs, as their performance will degrade with the inclusion of features in whose spaces the objects of interest are poorly separated.

In the experiments described here, training and testing were performed on each sequence independently. A topic of future investigation is whether a recognizer trained on one sequence will have similar accuracy when applied to another (for instance, a sequence recorded in a different subject). We also plan to perform histological validation to determine the true accuracy of our approach when applied to the detection of extra-luminal blood. To achieve our ultimate goal of visualizing the microvasculature, we will also attempt to increase the resolution of our method to near the diameter of the vasa vasorum ($\sim 50\text{-}200\ \mu\text{m}$) [4].

Acknowledgements: We would like to thank S. Carlier (CRF; New York, NY) and E. Falk (Aarhus Univ. Hospital; Aarhus, Denmark) for their valuable assistance with data acquisition, and all the members of the Ultimate IVUS team for their input. This work was supported in part by NSF Grant IIS-0431144 (IAK) and a NSF Graduate Research Fellowship (SMO). Any opinions, findings, conclusions or recommendations expressed in this material are the authors' and may not reflect the views of the NSF.

References

1. Nair, A., Kuban, B.D., Tuzcu, E.M., Schoenhagen, P., Nissen, S.E., Vince, D.G.: Coronary plaque classification with intravascular ultrasound radiofrequency data analysis. *Circulation* **106** (2002) 2200–2206
2. Goertz, D.E., Frijlink, M.E., Tempel, D., van Damme, L.C.A., Krams, R., Schaar, J.A., ten Cate, F.J., Serruys, P.W., de Jong, N., van der Steen, A.F.W.: Contrast harmonic intravascular ultrasound: A feasibility study for vasa vasorum imaging. *Invest Radiol* **41** (2006) 631–638
3. Hibi, K., Takagi, A., Zhang, X., Teo, T.J., Bonneau, H.N., Yock, P.G., Fitzgerald, P.J.: Feasibility of a novel blood noise reduction algorithm to enhance reproducibility of ultra-high-frequency intravascular ultrasound images. *Circulation* **102** (2000) 1657–1663
4. Gössl, M., Malyar, N.M., Rosol, M., Beighley, P.E., Ritman, E.L.: Impact of coronary vasa vasorum functional structure on coronary vessel wall perfusion distribution. *Am J Physiol Heart Circ Physiol* **285** (2003) H2019–H2026
5. Nair, A., Kuban, B.D., Obuchowski, N., Vince, D.G.: Assessing spectral algorithms to predict atherosclerotic plaque composition with normalized and raw intravascular ultrasound data. *Ultrasound Med Biol* **27** (2001) 1319–31
6. Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. *Neural Comput* **13** (2001) 1443–1471
7. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification. Technical report, Dept. of Computer Science and Information Engineering, National Taiwan University (2004)
8. Pasterkamp, G., van der Heiden, M.S., Post, M.J., ter Haar Romeny, B.M., Mali, W.P.T.M., Borst, C.: Discrimination of the intravascular lumen and dissections in a single 30-MHz US image: Use of “confounding” blood backscatter to advantage. *Radiology* **187** (1993) 871–872