Classification of Intravascular Ultrasound Signal by Kernel Density Estimation and Bayes Theorem for Identification of Coronary Plaque Tissue

Hiroki Tanaka, Kazuhiro Tokunaga, Eiji Uchino, and Noriaki Suetake

Abstract—We propose a method for tissue characterization of coronary plaque using IntraVascular UltraSound (IVUS) data. In the proposed algorithm, the probability density function of each plaque is calculated by the kernel density estimation and neural gas. The probability distributions vary continuously and overlap one another in the feature space. The class labels of tissue are determined by using Bayes estimation. This algorithm enables to give more delicate information to a medical doctor, which matches with the real situation of plaque composition. In this paper, the proposed algorithm is presented brief and its effectiveness is confirmed by applying it to the real IVUS data.

Index Terms—IntraVascular UltraSound (IVUS), Tissue Characterization, Kernel Density Estimation, Neural Gas, Bayes Estimation.

I. INTRODUCTION

A coronary plaque consists of three types of tissue. There are Fibrous, Fatty and Fibrofatty. The plaque is classified into the stable plaque or the unstable plaque depending on the difference in the regions of interest. Therefore, it helps medical doctors to indicate the delicate difference in the regions.

In this paper, the proposed algorithm is presented brief and its effectiveness is confirmed by applying it to the real IVUS data.

Therefore, a tissue characterization system is needed to assist medical doctors.

In the previous studies, some researchers indicated that the frequency analysis is effective in the tissue characterization [2] [3] [4]. Furthermore, we have proposed the Multiple k- Nearest Neighbor (MkNN) method using the power spectra as feature vectors [5]. In those methods, each tissue is characterized as one specific tissue.

However, the actual tissues of the coronary plaque blend with each other. That is, the region characterized as one kind of tissue must include a combination of different tissues. A medical doctor diagnoses the plaque as the stable plaque or the unstable plaque based on the difference in the regions of interest. Therefore, it helps medical doctors to indicate the delicate difference in the regions.

In this paper, we propose a tissue characterization method and representation method of the tissues included in the regions of interest as continual colors.

The details are as follows; firstly, the feature vectors are extracted from RF signals by using Fast Fourier Transform (FFT) after performing some preprocessing. Then, probability density functions of each kind of tissue in the feature space are estimated by using Kernel Density Estimation (KDE) and Neural Gas (NG).

When using NG, the probabilities of tissue are calculated more efficiently in KDE. The existence probabilities of the tissues are characterized by using Bayes estimation. The color of tissue is determined on the values of probabilities. Consequently these colors represent the delicate differences of the tissues.

II. METHOD

A. IVUS Method

IntraVascular UltraSound (IVUS) method gives the tomographic images of coronary arteries. In IVUS method, a catheter with an ultrasound probe attached to its end is inserted in a blood vessel.

Figure 1 shows the catheter inserted in the blood vessel. The probe emits the Radio Frequency (RF) ultrasound signals continuously, and then the probe also receives the signals reflected from the vascular wall. Depending on the kind of tissue the RF signal reflects onto, the received signal characteristics are different. The tomographic images (B-mode images) are made by transforming the amplitude values of RF signals into the luminosity values. Figure 2 shows a B-mode image.
Kernel function can be any function if it satisfies the following:

\[ \phi(u) \geq 0, \quad (2) \]

\[ \int \phi(u) du = 1. \quad (3) \]

The Gaussian function is generally used as a kernel function. In this paper, we use Gaussian kernel function. The probability density function \( p(x) \) is thus expressed as follows:

\[ p(x) = \frac{1}{N} \frac{1}{(2\pi h^2)^{D/2}} \exp \left\{ -\frac{||x - x_i||^2}{2h^2} \right\}, \quad (4) \]

where the bandwidth is a standard deviation of a Gaussian function.

In kernel density estimation, this probability density function is equivalent to the true probability density function if the sufficient number of data vectors is provided. However, the computing time is increased greatly.

### C. Neural Gas

In the proposed method, Neural Gas (NG) [6] is used for calculating the probability density function of tissue efficiently. The NG is one of the vector quantization methods. It has been reported that the quantization error of NG is smaller than the other methods.

The following shows the algorithm for a batch type NG. The algorithm of the batch type NG repeats the following processes, i.e., evaluation process, competitive process, and adaptation process.

1. **Initialization:**
   Suppose that the D-dimensional data vectors \( x_i (i = 1, 2, ..., N) \) are given. Moreover, suppose that the NG is composed of the reference vectors \( w_k(k = 1, 2, ..., K) \). In the first stage of learning, \( w_k \) are initialized by random numbers.

2. **Evaluation process:**
   The distances \( d_{i,k} \) between all the reference vectors \( w_k \) and all the data vectors \( x_i \) are calculated by:

\[ d_{i,k} = ||x_i - w_k|| \forall i, k. \quad (5) \]

3. **Competitive process:**
   The ranking of the reference vectors for each data vector \( x_i \) is evaluated by the distances calculated by (5). After that, the learning rate of each reference vector is determined according to its ranking. Here, the learning rate \( \psi_{i,k} \) of the \( k \)-th reference vector \( w_k \) for the \( i \)-th data vector \( x_i \) is defined as follows:

\[ \psi_{i,k} = \frac{\phi}{\sum_i \phi_{i,k}}, \quad (6) \]

\[ \phi = \exp \left\{ -\frac{r_{i,k}}{\lambda(t)} \right\}, \quad (7) \]

where \( r_{i,k} \) is a rank of the \( k \)-th reference vector \( w_k \) for the \( i \)-th data vector \( x_i \). In addition, \( \lambda(t) \) is a monotonically decreasing function for the learning step \( t \). In this work, \( \lambda(t) \) is defined as follows:

\[ \lambda(t) = (\lambda_{max} - \lambda_{min}) \exp \left\{ -\frac{t}{\tau} \right\} + \lambda_{min}, \quad (8) \]

where \( \tau \) is a time constant.
4) Adaptation process:
The update rule of all the reference vectors \( \mathbf{w}_k \) is defined as follows:

\[
\mathbf{w}_k = \sum_{i=1}^{N} \psi_{i,k} \mathbf{x}_i. \tag{9}
\]

In this paper, the semi-batch type NG is employed to provide the consistent results and to process much data.

III. TISSUE CHARACTERIZATION BY IVUS DATA USING KERNEL DENSITY ESTIMATION

The proposed method is divided into the following 3 steps, i.e., feature vector extraction, estimating probability density functions, and characterization.

A. Feature Vector Extraction by Using Fast Fourier Transform

It is reported that the frequency analysis is valid for tissue characterization [2]. In the present method, the power spectra of RF signals are used as feature vectors. Specifically, the RF signals before and after the points of interest are processed by a Fast Fourier Transform (FFT) with a hammering window. As a preprocessing, removing of bias on the RF signal and Time-Gain Compensation (TGC) [7] are performed. TGC is generally done for compensating the attenuation of the ultrasound signal in IVUS method.

B. Estimating Probability Density Function

In the present method, the probability density functions of each tissue are estimated by the using Kernel Density Estimation (KDE) [8]. However, KDE takes much time to calculate the probability density function \( p(x) \) if there are a lot of data vectors. The reference vectors \( \mathbf{\mu}_j (j = 1, ..., M) \) are thus used as the data vectors \( \mathbf{x}_i (i = 1, ..., N) \) in KDE \((N > M)\).

In order to estimate the probability density function with high accuracy and the small number of data vectors, Neural Gas (NG) is used for quantization. It is reported that there are less quantization errors of NG than the other methods. NG is applied to the feature vectors of each tissue. A conditional probability density function \( p(x|\omega) \) is calculated by using (4) and reference vectors \( \mathbf{\mu}_j \), \( \omega \) is a class label of tissue.

C. Characterization by Using Bayes Estimation

The proposed method characterizes and represents the delicate and sophisticated differences of the tissue for better helping a medical doctor for diagnosis of ACS. In characterizing, the class label \( \omega \) of a tissue is determined by:

\[
\omega = \arg \max_{\omega_j} p(\omega_j|x), \tag{10}
\]

\[
p(\omega_j|x) = \frac{p(x|\omega_j)p(\omega_j)}{\sum_{i=1}^{L} p(x|\omega_i)p(\omega_i)}, \tag{11}
\]

\[
p(x, \omega_i) = p(x|\omega_i)p(\omega_i), \tag{12}
\]

where \( p(\omega) \) is a-priori probability. The conditional probability is calculated by (4). Equations (10) and (11) are the maximum a-posteriori probability estimation.

Equation (12) is a joint probability density function. This joint probability density function expresses a probability of characterizing a tissue in the point of interest as each class \( \omega_i \). We call this joint probability density existing probability. In this paper, each value of \( p(x, \omega_i) \) is allocated to the luminosity value for representing the delicate and sophisticated differences of tissue. The values of joint probability density functions \( p(x, \omega_1) \), \( p(x, \omega_2) \) and \( p(x, \omega_3) \) are allocated to R, G, and B values, respectively.

IV. EXPERIMENTS

In the experiments, the proposed method is compared with the Multiple k-Nearest Neighbor (MkNN) method proposed by the authors in the past [5]. Firstly, the results of the proposed method and MkNN are compared qualitatively. And then, the accuracy of tissue characterization is numerically evaluated in order to confirm the validity of the proposed method.

A. Experimental Settings

In the experiments, three different sections of the blood vessel were used. The findings in the sections are obtained by a medical doctor. There are three types of tissue, Fibrous, Fatty and Fibrifatty.

The present method is compared with Multiple k-Nearest Neighbor (MkNN) method [5]. MkNN is a method proposed by the authors for tissue characterization [5].

In MkNN method, the class label is determined by considering not only a feature space but also an observation space, under the presumption that adjacent feature vectors in the observation space should be in the same class.

In order to quantitatively evaluate the performance of tissue characterization, each method was compared by using True Positive Rate (TPR) and True Negative Rate (TNR). Leave-one-out cross-validation was also used. TPR and TNR is calculated by the following:

\[
TPR = \frac{\text{the number of TP}}{\text{the number of TP} + \text{the number of FN}}, \tag{13}
\]

\[
TNR = \frac{\text{the number of TN}}{\text{the number of TN} + \text{the number of FP}}, \tag{14}
\]

where \( TP \) (True Positive) means correct identification, \( FN \) (False Negative) means incorrect rejection, \( TN \) (True Negative) means correct rejection, and \( FP \) (False Positive) means incorrect identification. For using leave-one-out cross-validation, the data are classified into two groups. One is the test data and the other is the training data. Every possible combination of the data is evaluated.

The width of the hammering window for FFT was 64 points. In both the proposed method and the MkNN method, 40 reference vectors were calculated for each tissue by NG. In MkNN method, the number of browsed teaching vectors in the feature space is 9. The bandwidth of the kernel function is 1.7. The experiments were tried 20 times.
The existence probabilities of unknown input vectors are estimated by using Bayes estimation. The class labels are determined based on the existence probabilities. Furthermore, allocating the existence probabilities to R, G, and B values, the delicate and sophisticated information in the tissue characterization could be visually expressed.

In the experiments, the proposed method was compared with MkNN. The proposed method is more stable than MkNN and it enables to give more delicate information to a medical doctor for a diagnosis.

VI. Acknowledgements

Many thanks are due to Prof. Takafumi Hiro of Nihon University, School of Medicine, for providing the IVUS data and for his helpful discussions.

REFERENCES


Fig. 4. The tissue characterization results by the conventional method (MkNN) and the proposed method. (a) The tissue composition given by a medical doctor by examining the dyed tissue using microscope. (b) The tissue characterization results by the MkNN. (c) The tissue characterization results by the proposed method. (b) and (c) give almost the same results.

Fig. 5. The probability of each class of tissue. (a), (b), and (c) show the probabilities of each class; Fibrous, Fatty and Fibrofatty. Brighter points have higher probabilities. In (d), the probability values of each tissue, i.e., the probabilities of (a) Fibrous, (b) Fatty, and (c) Fibrofatty, are allocated to R, G, and B values, respectively. The general and continuous changes of tissue can be seen.

TABLE I
QUANTITATIVE EVALUATION OF THE PROPOSED METHOD.

<table>
<thead>
<tr>
<th>Method</th>
<th>Fibrous TPR</th>
<th>Fibrous TNR</th>
<th>Fatty TPR</th>
<th>Fatty TNR</th>
<th>Fibrofatty TPR</th>
<th>Fibrofatty TNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Method</td>
<td>77.8 ± 0.8</td>
<td>72.4 ± 2.3</td>
<td>54.6 ± 4.7</td>
<td>94.2 ± 0.3</td>
<td>67.8 ± 2.4</td>
<td>83.6 ± 0.8</td>
</tr>
<tr>
<td>Conventional Method</td>
<td>74.6 ± 1.4</td>
<td>80.7 ± 2.8</td>
<td>72.5 ± 3.2</td>
<td>92.2 ± 0.4</td>
<td>74.2 ± 3.1</td>
<td>82.8 ± 1.3</td>
</tr>
</tbody>
</table>