Paper to Annual Conference of Biomedical Fuzzy System Association

Similarity-Based Method for Tissue Characterization of Coronary Artery Plaque

Gancho VACHKOV* and Eiji UCHINO*,**

*Graduate School of Science and Engineering, Yamaguchi University
**Fuzzy Logic Systems Institute

Abstract: *Tissue characterization of coronary artery plaque plays an important role in diagnosing the acute coronary artery syndrome. This paper presents a similarity-based method for tissue characterization, combined with a moving window computational strategy. At every computation step the dissimilarity degree between two representative models is calculated and estimated.*

Two types of representative models are presented and discussed in the paper, namely the Histogram-based models that do not need learning and Key Point-based models that need unsupervised learning. This leads to two different tissue characterization methods that are explained and analysed in details in the paper. Simulations results by using data from two cross-sections of a coronary artery show that the Key Point model-based method achieves slightly better tissue characterization.

Keywords *Similarity Analysis, Tissue Characterization, Intravasular Ultrasound (IVUS), Classification, Moving Window, Histogram, Key Point Models*

1. Introduction and Problem Statement

Tissue characterization of the plaque that has gradually built in the coronary artery is very important and necessary step for a correct diagnosis of the acute coronary syndrome (ACS) with a possible treatment. Here the Intravascular Ultrasound (IVUS) method has been for a long time the most often used method for diagnosis of ACS.

Many experiments, deep research and analysis have been done until now [1] by using and proposing different classification techniques and algorithms. However, there is still not an "ideal" method for a reliable and correct tissue characterization. In this paper we propose a similarity-based tissue characterization method with moving window computation strategy and evaluate the results obtained.

The IVUS method is a tomographic imaging technology, in which medical doctors use a gray-scale radial-shape *B-mode image* for observation and analysis of the IVUS data. This image corresponds to one cross-section of the coronary artery with a given depth-range in all *256* directions (angles) of rotation of the IVUS probe.

However, in this paper, we use the more classical rectangular type of representation with axis *X* denoting all angles (directions) between *0* and *255*, and the ordinate *Y* denoting the depth of the measurement (the distance between the probe and the current measured signal).

Gancho VACHKOV and Eiji UCHINO
1677-1 Yoshida, Yamaguchi, Yamaguchi, 753-8512, Japan
Phone: +81-83-933-5699, Fax: +81-83-933-5699,
E-mail:{vachkov}{uchino}@yamaguchi-u.ac.jp

As initial information we use the raw (actual) reflected RF signal with a simple preprocessing, as shown in Fig. 1, namely taking the absolute vale of the RF signal after subtracting the central (stationary) value of *2000* from it.

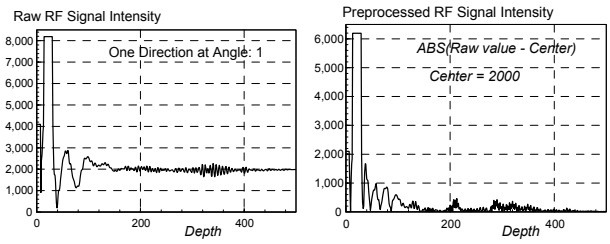


Fig. 1. The Raw and the Preprocessed RF Signal.

In our paper, one cross-section of the coronary artery is represented by a large number of data in the rectangular area *X-Y* (*Directions-Depths*) as shown in Fig. 2. Then the problem is to utilize all these data in the best way, in order to properly find (characterize) the regions of *Lipid* tissue and those of *Fibrous* tissue in the artery plaque. From a research view point this is a typical classification problem. First of all, some proper training data are needed before performing the actual classification.

For such purpose we have used the results from the *true classification,* made by the doctor through microscopic analysis, which is shown in Fig. 2 as two *regions of interest* (ROI), namely the *Lipid* ROI and *Fibrous* ROI. Two other regions, denoted as *Region1* and *Region2* are also shown in the figure as examples of *unknown regions* that have to be classified as *belonging* (or *not belonging*) to either Lipid or Fibrous tissue.

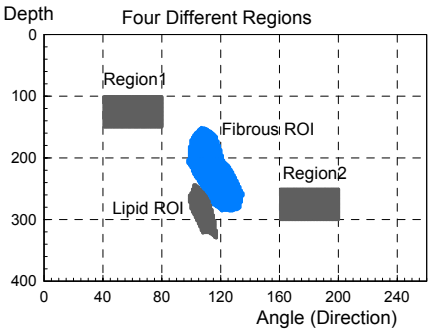


Fig. 2. One Cross-Section with two Regions of Interest (ROI) and two Arbitrary Chosen Regions.

2. The Similarity-Based Moving Window Method for Tissue Characterization

In this paper, we propose the moving window method for classification, based on Similarity Analysis. The general idea of the Similarity analysis was proposed in [2] and has been successfully used for classification of images and machine operating modes. Here we propose its extension version to a moving window method, which is suitable for solving the problem of tissue characterization. It can be summarized briefly as follows:

Preparation Step: The whole examination area from a given cross section is divided into a grid with fixed number of *windows* Nw having an equal predefined size: *Angle_Range* x *Depth_Range*, as shown in Fig. 3. For all further experiments in this paper, a constant window size of *10* x *40* has been used, which means that *400* data from the reflected RF signal can be extracted from each window;

Step 0: The available data from a given ROI (e.g. Lipid ROI or Fibrous ROI) are used for creating the respective *Representative Model* RM*o*; Set *i = 1*;

Step 1: Extract all data in the current *i*-th window from the grid of the examination area;

Step 2: Use the extracted data from this window to create the respective Representative Model RM*i*;

Step 3: Calculate the Difference (the *Dissimilarity Degree*) *DS* between the two models, as a real value, bounded between 0 and 1, namely:

$$DS(RM_0, RM_i) \in [0,1] \tag{1}$$

Step 4 (Decision Step with a given Threshold *Th*):

If $DS(RM_0, RM_i) \le Th$ Then

the *i*-th window is characterized as ***Similar*** to the ROI that has been used for creating the RM0;

Otherwise this window is not characterized.

Step 5: Select the next window $i \leftarrow i+1$ from the grid.

If $i \le N_w$ Then *Go To Step*1 Otherwise *Stop*.

It is easy to notice that the most important part of the above algorithm is the *type* of the Representative Models used in *Step 0* and *Step 2*, as well as the *method* to calculate the Dissimilarity Degree *DS* in (1). The next two sections give two different solutions to these problems.

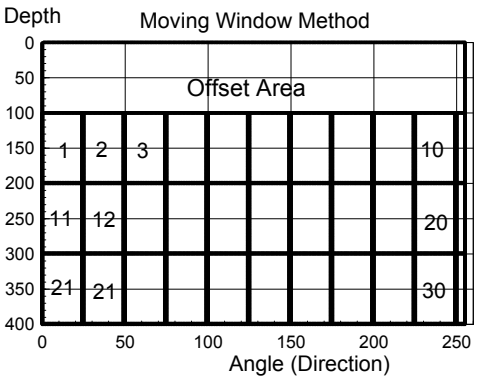


Fig. 3. An Illustration of the Moving Window Method for Tissue Characterization

3. Histogram-Based Similarity Analysis for Tissue Characterization

In this Section we propose the use of normalized Histograms as Representative Models, created for a given ROI and for each of the windows in the grid from Fig. 3. Let us denote with *M* the data number in one ROI or in a given window. Then for a pre-selected number of *N* subsequent Intervals with equal range (width), the Histogram is calculated as follows:

$$h_i = m_i / N \in [0,1]; \quad \sum_{i=1}^{N} h_i = 1 \tag{2}$$

where:

$$m_i \in [0, M]; \quad \sum_{i=1}^{N} m_i = M \tag{3}$$

Here $m_i,\ i = 1,2,...,N$ denote the number of data points that fall into the *i*-th Interval The width of each interval is fixed to *20* in all further calculations.

The next Fig. 4 shows the histogram of the Lipid ROI (containing *885* data) and the histogram of the Fibrous ROI (containing *2896* data), calculated by (2) and (3). Their difference can be easily noticed and such a significant difference is definitely helpful for the proper tissue characterization.

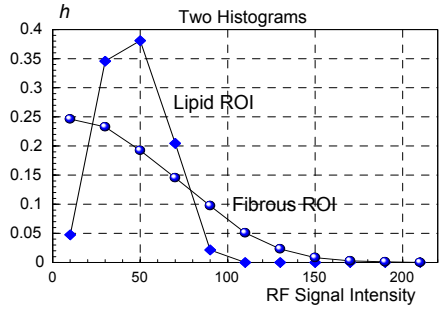


Fig. 4. The Histograms of the two Regions of Interests: Lipid ROI and Fibrous ROI from Fig. 2.

Fig. 5. depicts two other Histograms of the arbitrary chosen regions *Region1* and *Region2* from Fig.1. Each of those regions contains *50* x *40 = 2000* data. It is easy to notice that they look quite different from each other and are also different from the histograms in Fig. 4.

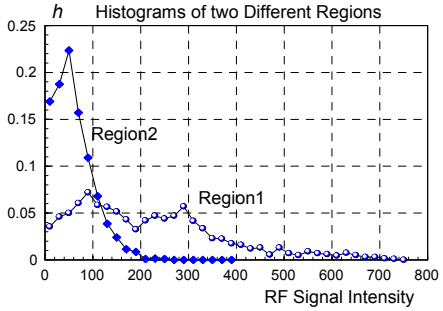


Fig. 5. Histograms of two Arbitrary Chosen Regions, shown in Fig. 2.

Then the Dissimilarity Degree *DS* from (1) uses now the Histograms *Ho* and *Hi* as Representative Models and is calculated as follows:

$$DS(H_0, H_i) = \sum_{j=1}^{N} \left| h_j^0 - h_j^i \right| \Big/ 2 \in [0,1] \qquad (4)$$

The extreme case of "not similar at all" histograms is: *DS = 1*where the two histograms do not overlap.

4. Key Point Model-Based Similarity Analysis for Tissue Characterization

The use of Histograms as Representative Models has the disadvantage that there are "no tuning parameters" in this approach, which could be helpful to change (amplify or attenuate) the dissimilarity degree (4), in order to improve the characterization accuracy. Therefore in this Section we propose another type of Representative Models called Key Point (KP) models.

The proposed KP models here are a kind of modification of the general idea in [2], where all available data *M* are represented (compressed) by a small number of *K* neurons (called also *Key Points*).. This is usually done by a competitive learning algorithm, such as the popular *Neural Gas* or some of its modifications.

In this paper we use a special modification of the basic competitive learning algorithm, called "frequency sensitive competitive learning" FSCL [3]. The idea of FSCL is to gradually suppress the *winner-neuron* that has won "too frequently" during the learning. Because of space limitations, the computational details are omitted in this paper, but can be found in [3].

After the learning is completed, a *post-processing* is performed in order to discover important parameters of the extracted Key Points (neurons), as follows:

- *Relative Weights* of the neurons:

$$w_i = m_i / M \in [0,1]; \quad i = 1,2,...K \qquad (5)$$

Here m_i denotes the number of all data points that have a *minimal* Euclidean *distance* to the *i*-the neuron.

- *Center-of-Gravity (CG) of the Key Point Model.* This is essentially the *Weighted Average* of the locations R_i, $i = 1,2,...,K$ of all *K* neurons in the KP Model:

$$CG = \sum_{i=1}^{K} R_i w_i \qquad (6)$$

This parameter represents an important characteristic of all *M* data, namely the central location of the data set in one-dimensional space of the RF Signal Intensity.

- *Average Distance* of the subset of all data points m_i to their nearest *i*-th neuron:

$$AD_i = \sum_{j=1}^{m_i} D_j^i; \quad i = 1,2,...K \qquad (7)$$

Here D_j^i denotes the Euclidean distance between the *j*-th data point and the nearest *i*-th neuron to it.

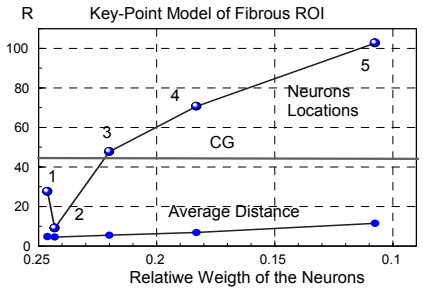
- *Average Spread AS* of the Key Point Model. It is calculated as the weighted average distance of all data points to their respective nearest neurons:

$$AS = \sum_{i=1}^{K} AD_i w_i \qquad (8)$$

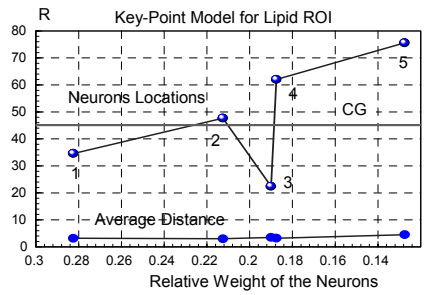
This parameter reveals another important characteristic of the data set of *M* points, which is similar to the data density distribution. A large value of *AS* corresponds to a more sparse data set and a small value of *AS* is an indication for a dense data set.

Among all the parameters of the KP model, further on we use the *Center-of-Gravity CG* (6) and the *Average Spread AS* (8) as two most representative parameters.

The following Fig. 6. shows the parameters of the KP models for Fibrous ROI and for Lipid ROI with the assumption of *K = 5* neurons. These models have been trained with a linearly decreasing Learning Rate of 0.9 for 4000 iterations. The learning time was within *5* seconds for each of the models. The learning of the KP models for the windows was much faster, because of the small number of data (just *400*) in each window.



a)



b)

Fig. 6. Parameters of the KP Models for the Fibrous ROI and Lipid ROI, used for Tissue Charcterization.

Calculation of the Dissimilarity Degree *DS* (1) could be done in several different ways, but we propose here the following calculation of *DS* for the models RM_0 and

RM_i, which has a clear physical meaning:

$$DS = 1 - exp^{-\frac{(CG_0 - CG_i)^2}{2\sigma_C^2}} .exp^{-\frac{(AS_0 - AS_i)^2}{2\sigma_S^2}} \quad (9)$$

Here σ_C and σ_S denote the predefined *width* for the *CG* and width for the *AS*, respectively. Smaller values of σ_C or σ_S mean that we have put a bigger importance (bigger sensitivity) to the *DS*.

The calculated parameters of the KP models for Fibrous ROI and Lipid ROI are as follows:

$$CG_{Fibrous} = 43.50;\ AS_{Fibrous} = 5.95;$$
$$CG_{Lipid} = 45.41;\ AS_{Lipid} = 3.37;$$

It is seen that both *CG*s are quite closed to each other, but both *AS*s are much more apart. Therefore we have put more importance to the AS parameter, in order to achieve a better tissue characterization, by setting the following values for the widths: $\sigma_C = 4.0;\ \sigma_S = 0.5;$ These parameters serve as tuning parameters of the proposed method for KP Model-based similarity analysis.

5. Experimental Results and Analysis

The performance of the above proposed two methods for tissue classification has been tested and compared on a *new data set* that represents another cross-section in the same blood vessel (and the same patient). This test data set was not used for training the Representative Models (the Histograms and KP models) of the Fibrous and Lipid ROI, but the "true solution", i.e. the actual Fibrous and Lipid ROI have been known and used for confirmation of the results. The classification results from both methods are shown in Fig. 7.and Fig. 8.

It is easy to notice that none of the methods can perfectly detect the true ROI, but still the KP Model-based method from Fig. 8. is the better one, because it produces less number of misclassifications for the same assumed threshold of *0.15*.

5. Conclusions and Future Work

In this paper we have proposed a general Similarity-based method for tissue characterization, which uses the moving window computational strategy for comparing the representative model of a given ROI with the respective models of each window in the examination grid. Two kinds of representative models have been proposed in the paper, namely the Histogram-based and the Key Point-based models.

The results show that the KP model-based similarity produces more plausible classification results. The main reason for this is that the KP models can be *tuned* with different number of neurons and also different possible widths for calculating the *DS* in (9) can be applied.

The further research is aimed at improving the classification accuracy by automatic selection of the widths (9), as well as the window sizes and the threshold for classification.

This work was supported by the Grant-in-Aid for Scientific Research (B) of the JSPS under the Contract No. 23300086.

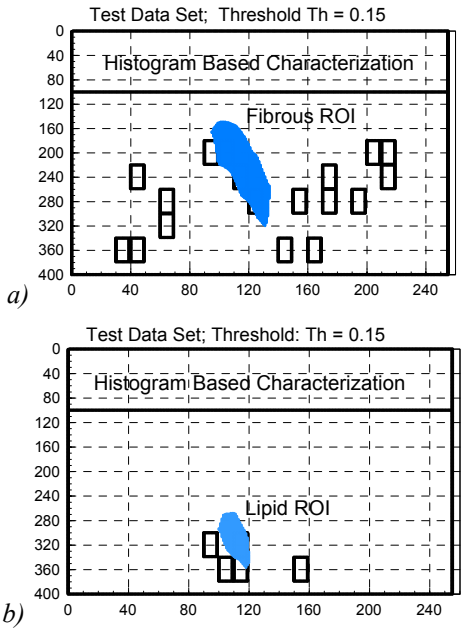


Fig. 7. Results from the Tissue Characterization of Unknown (Test) Data Set by use of Histograms.

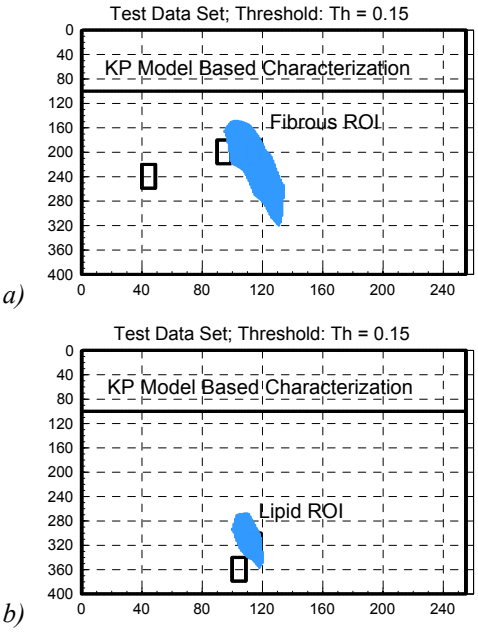


Fig. 8. Results from the Tissue Characterization of Unknown (Test) Data Set by use of KP Models.

References

[1] Koga T., Furukawa S., Uchino E. and Suetake N., "High-Speed Calculation for Tissue Characterization of Coronary Plaque by Employing Parallel Computing Techniques", *Int. Journal of Circuits, Systems and Signal Processing,* Issue 4, Volume 5, pp. 435-442, 2011.

[2] Vachkov G., "Online Incremental Image Classification by use of Human Assisted Fuzzy Similarity", *Proc. of the 2010 IEEE Int. Conf. on Information and Automation,* ICIA-2010, Harbin, China, pp. 834-839, 2010.

[4] Xu, L., Krzyzak A. and Oja, A., "Rival Penalized Competitive Learning for Clustering Analysis, RBF Net and Curve Detection", *IEEE Trans. Neural Networks*, Vol. 4, No. 4, pp. 636-649, 1993.