IVUS-Based Coronary Plaque Tissue Characterization Using Weighted Multiple k-Nearest Neighbor

Eiji Uchino, Kazuhiro Tokunaga, Hiroki Tanaka, Noriaki Suetake

***Abstract*—In this paper, we propose an extended algorithm of the multiple k-nearest neighbor (MkNN), which is applied to an intravascular ultrasound (IVUS)-based tissue characterization of coronary plaque. In the proposed algorithm, a weighted decision based on the distances between the input vector and the prototype vectors is employed instead of the majority decision in the labeling process of k-nearest neighbor (kNN). The fibrous and lipid tissues were thus characterized more accurately than ever. Furthermore, the accuracy of tissue characterization was improved even in the case where the number of prototype vectors in MkNN is smaller. The effectiveness of the proposed method has been examined by the actual experiments using the artificial data and the true IVUS data.**

***Index Terms*—intravascular ultrasound (IVUS), tissue characterization, multiple k-nearest neighbor.**

I. Introduction

A Myocardial infarction is caused by a failure of plaque built inside the coronary artery. It is very important to characterize the tissue of plaque in order to early prevent myocardial infarction [1].

In general, the tissue characterization is carried out by analysing the radio frequency (RF) signal obtained from the intravascular ultrasound (IVUS) method [2] using catheter. In our past works [3], [4], [5], we proposed a multiple k-nearest neighbor (MkNN) method, that is an extension of the k-nearest neighbor (kNN) method, in order to get better results when applied to the tissue characterization of coronary plaque.

Although a good tissue characterization by MkNN had been obtained in our previous works [3], [4], [5], it still remained a problem that MkNN takes a lot of computing time for tissue characterization. The computing time for characterization depends mainly on the number of prototype vectors in kNN, since the kNN has to calculate the distances between the input vector and all the prototype vectors.

On the other hand, it is necessary to adjust the number of prototype vectors to keep the accuracy of tissue characterization, since the characterization results of MkNN are influenced by the distribution of the prototype vectors.

Hence, we propose in this paper an extended algorithm of a multiple k-nearest neighbor (MkNN) method for an intravascular ultrasound (IVUS)-based tissue characterization. In the proposed algorithm, a weighted decision based on the distances between the input vector and all the prototype vectors is employed instead of the majority decision in the labeling process of k-nearest neighbor (kNN).

The experiments show that the accuracy of tissue characterization by the extended MkNN is good, even if a small number of the prototype vectors are selected at random. The effectiveness of the proposed method is verified by the artificial data and by the real IVUS data.

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

E. Uchino is with the Graduate School of Science and Engineering, Yamaguchi University, 1677-1 Yosida, Yamaguchi 753-8512, Japan (phone: +81-83-933-5699; e-mail: uchino@yamaguchi-u.ac.jp), and also with the Fuzzy Logic Systems Institute (phone: +81-948-24-2771; e-mail: uchino@flsi.or.jp).

K. Tokunaga is with the Fuzzy Logic Systems Institute, 680-41 Kawazu, Iizuka, Fukuoka 820-0067, Japan (phone: +81-948-24-2771; e-mail: tokunaga@flsi.or.jp).

H. Tanaka is with the Graduate School of Science and Engineering, Yamaguchi University (e-mail: s015vc@yamaguchi-u.ac.jp).

N. Suetake is with the Graduate School of Science and Engineering, Yamaguchi University (e-mail: nsuetake@yamaguchi-u.ac.jp).

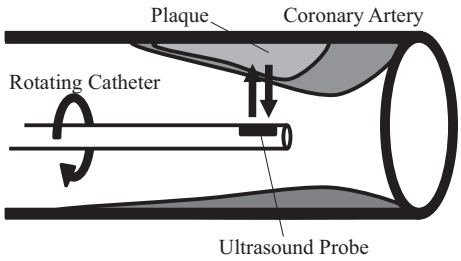


Fig. 1 An ultrasound probe attached to the distal end of a catheter. The ultrasound signal is transmitted from the probe and the reflected signal from the tissue is observed also by the probe.

II. Tissue Characterization by IVUS Method Using Multiple k-Nearest Neighbor

This chapter shows briefly the intravascular ultrasound (IVUS) method, the multiple k-nearest neighbor (MkNN) method, and the tissue characterization by the IVUS method using MkNN.

A. IVUS method

The ultrasound probe attached at the tip of the catheter is inserted into a blood vessel. After that, the ultrasound signal is transmitted forward from the probe and then the reflected signal from the tissue is received while rotating the probe [6]. Fig. 1 shows the catheter in the blood vessel.

The ultrasound signal transmitted from the probe is called a radio frequency (RF) signal. The intensity of the reflected RF signal from the tissue depends on the characteristics of the tissue and also on the location of the probe.

The RF signals are observed in all directions in the blood vessel. Concretely, the absolute value of the sampled RF

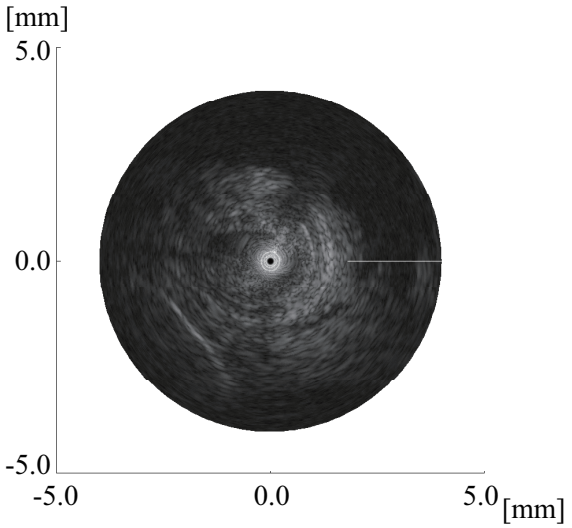


Fig. 2 An example of the B-mode image obtained by the IVUS method. This is a real time ultrasound cross-sectional image of a blood vessel where a catheter probe is currently rotating.

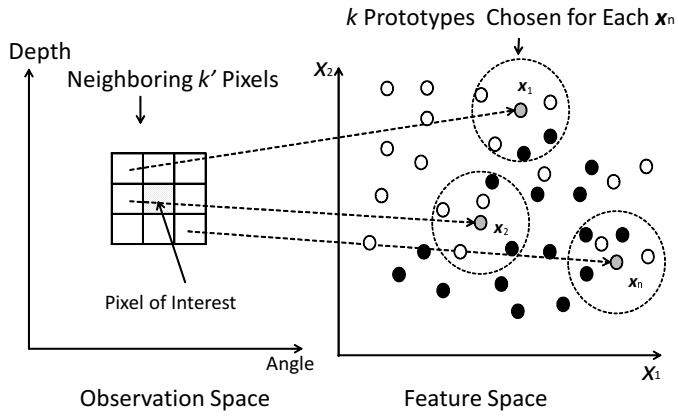


Fig. 3 Basic scheme for classification by MkNN using the information not only in the feature space but also in the observation space [3]. kNNs both in the feature space and in the observation space are performed.

signal is first taken, and its envelop, and then finally its logarithmic value is calculated. This transformed signal is converted into 8-bit luminosity values. Those luminosity values in all radial directions of RF signals are used to obtain a tomographic cross sectional image of a coronary artery. This image is called a B-mode image. Fig. 2 shows a sample image of this.

The B-mode image is a real time ultrasound cross-sectional image of a thin section of a blood vessel where a catheter probe is currently rotating. In this study, the B-mode image is constructed with 1,024 pixels in depth, and 256 lines in radial direction.

It is difficult however to see the conditions inside the blood vessel only from this B-mode image.

B. Multiple k-nearest neighbor

The multiple k-nearest neighbor (MkNN) method is an extension of the traditional k-nearest neighbor (kNN) method. In MkNN, the class label is determined based on the information not only in the feature space but also in the observation space.

Fig. 3 shows a basic scheme for classification by MkNN using both information in the feature space and in the observation space. In the algorithm of MkNN, the classification by kNN in the feature space is first performed, and then the classification by kNN in the observation space is followed.

1) kNN in the feature space

Suppose that the feature vectors $\boldsymbol{w}_i (i = 1, 2, .., N)$ are given, and let the class label ω_i of each feature vector be known. We use the feature vectors as the prototype vectors in MkNN.

When the input vector $\boldsymbol{x}$, whose class label is unknown, is applied to MkNN, the class label of the input vector $\boldsymbol{x}$ is determined as follows [7]:

$$l = \arg\max_{\omega} \sum_{i=1}^{N} \delta\left(\boldsymbol{x}; \boldsymbol{w}_i \mid \omega_i = \omega\right), \tag{1}$$

$$\delta\left(\boldsymbol{x}; \boldsymbol{w}_i \mid \omega_i = \omega\right) = \begin{cases} 1 & \text{if } \|\boldsymbol{w}_i - \boldsymbol{x}\| \leq r\left(k\right), \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $r(k)$ represents the Euclidean distance between the input vector $\boldsymbol{x}$ and the k-th nearest prototype vector.

In the first step, the prototype vectors in the k-th nearest neighbor are determined by calculating the distances of Eq.(1). After that, the class label of the input vector $\boldsymbol{x}$ is determined by a majority vote for the class labels of the prototype vectors in the k-th nearest neighbor.

2) kNN in the observation space

The kNN is used also in the observation space after determining the class labels of the input vectors $\{\boldsymbol{x}_n\}$ in the feature space. Except for the calculations of the distances between the input vector and the prototype vectors in the observation space, other calculations are the same as Eqs.(1) and (2).

C. Tissue characterization by the multiple k-nearest neighbor

In the tissue characterization based on IVUS method by using MkNN, the power spectra calculated for the RF signals by the short-time Fourier transform are employed as the prototype vectors in MkNN [8], [9].

The class label of each prototype vector is known from the findings of a medical doctor examining the corresponding dyed tissue looking through a microscope.

In this work, the fibrous tissue, the lipid tissue, and the fibrofatty tissue are classified. Concretely, the power spectra are calculated for each of the corresponding RF signal, i.e., for the fibrous tissue, for the lipid tissue, and for the fibrofatty tissue.

Fig. 4 shows the distributions of power spectral feature vectors for each tissue. These distributions are overlapping, and it is difficult to separate them into 3 classes linearly.

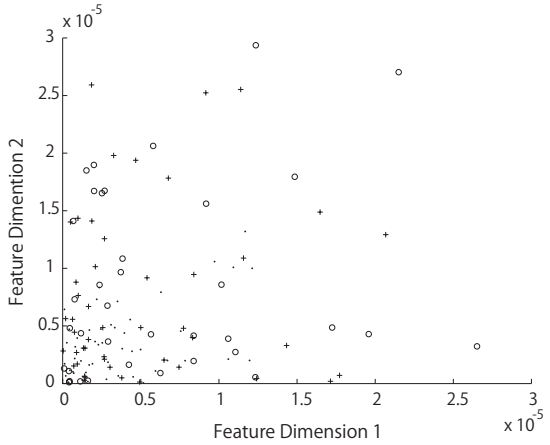


Fig. 4 The distributions of power spectral feature vectors of each tissue. The feature vectors are obtained by Fourier analysis for the corresponding RF signals. ○: Fibrous tissue, ●: Lipid tissue, +: Fibrofatty tissue.

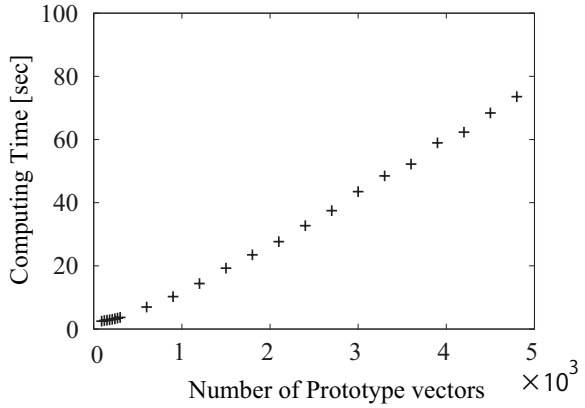


Fig. 5 The computing time for classification by MkNN versus the number of prototype vectors. The computing time increases in proportion to the number of prototype vectors.

The computing time for classification depends mainly on the number of prototype vectors in MkNN. Fig. 5 shows the computing time for classification versus the number of prototype vectors.

On the other hand, the accuracy of classification also depends on the number of prototype vectors, since the classification results by MkNN are influenced by the distribution of the data.

The purpose of this study is thus to innovate the algorithm in which the classification accuracy is not affected by the number of prototype vectors.

III. Tissue characterization using weighted MkNN

In our past tissue characterization works by MkNN [3], [4], [5], all the feature vectors observed from each tissue were used as the candidates of the prototype vectors. The representative vectors that were used as the prototype vectors were selected manually or selected according to the simple rules. The characterization results thus depended on the selected prototype vectors.

Furthermore, the accuracy of tissue characterization is decreased if the number of prototype vectors becomes small, because the distribution of the prototype vectors in feature space becomes sparse. Hence, it was needed to adjust the number of prototype vectors keeping the accuracy of the tissue characterization.

In the extended MkNN, a weighted decision, based on the distances between the input vector and the prototype (reference) vectors, is employed instead of the majority decision [10], [11]. As a result, it is expected that the tissue characterization dose not depend on the number and the distribution of the prototype vectors.

In the classification of the input vector, the distances between the input vector and the prototype vectors are calculated, and they are employed as the weights for the weighted decision.

In the extended MkNN, Eqs.(1) and (2) are replaced by the following equations:

$$l = \arg\max_{\omega} \sum_{i=1}^{N} \phi\left(\boldsymbol{x}; \boldsymbol{w}_i \mid \omega_i = \omega\right), \tag{3}$$

$$\phi\left(\boldsymbol{x}; \boldsymbol{w}_i \mid \omega_i = \omega\right) = \begin{cases} \exp\left(-\|\boldsymbol{x}_i - \boldsymbol{w}_k\|^2/\epsilon^2\right), \\ \qquad \text{if } \|\boldsymbol{x}_i - \boldsymbol{w}_k\| \le r\left(k\right) \\ 0 \qquad \text{otherwise,} \end{cases} \tag{4}$$

where ϵ is a decay parameter that is adjusted according to the distribution of the prototype vectors.

IV. Experiments

We apply the normal MkNN and the extended (proposed) MkNN to two test data sets. One is the artificial data and the other is the real IVUS data for coronary plaque.

A. Application to artificial data

1) Experimental settings: Fig. 6 shows the observation space and data structure. Observation space is a two-dimensional space with 500×500 pixels. Each pixel in the observation space is assigned as one of the three classes (i.e., red, blue, and green classes), each of which has a two-dimensional feature vector. Feature vectors are distributed as follows;

Class 1 (red): Gaussian distribution with mean (1.0, 1.0) and standard deviation (0.05, 0.05).

Class 2 (green): Gaussian distribution with mean (1.0, 1.0) and standard deviation (0.20, 0.20).

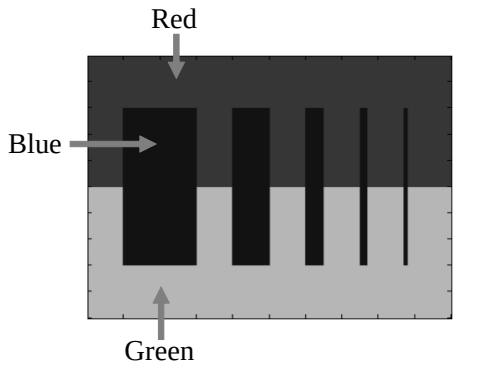
Class 3 (blue): F distribution with $F(100, 10)$ for both axes.

The distribution of each class (red, blue, and green) overlaps with each other as shown in Fig. 7. The probability density of each distribution is different. Prototype (reference) vectors are selected at random.

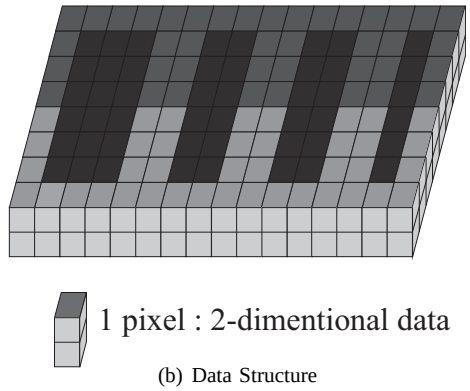
We compare the results by MkNN and wMkNN (extended MkNN) for these data set. The correct classification rate (CCR) is defined by:

$$\mathrm{CCR} = \frac{\text{number of correctly classified data}}{\text{number of total classified data}} \times 100. \tag{5}$$

The mean of CCR is calculated for 100 trials. CCR with 100 reference vectors and with 10 reference vectors for each class are evaluated.



(a) Observation Space



(b) Data Structure

Fig. 6 Observation space and data structure.

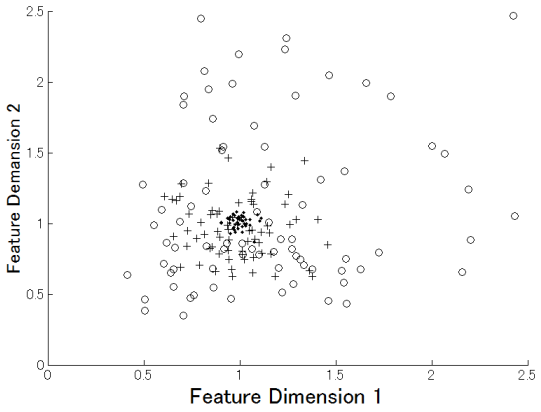


Fig. 7 Distribution of feature vectors for the artificial data used in the experiments. •: Class 1 (red), +: Class 2 (green), ○: Class 3 (blue).

2) Experimental results: Fig. 8 shows the classification results with 100 reference vectors for each class. Fig. 8 (a) and (b) show the results of classification by the normal MkNN and by the proposed wMkNN, respectively.

Fig. 9 shows the classification results with 10 reference vectors for each class. Fig. 9 (a) shows the results of classification by the normal MkNN. Fig. 9 (b) shows the results of classification by the proposed wMkNN.

Table I shows the correct classification rates by each method. CCR by the normal MkNN is considerably decreased with the number of the reference vectors. However the proposed wMkNN can maintain the classification accuracy even the number of reference vectors becomes small.

TABLE I Correct classification rates

Number of reference vectors for each class	MkNN [%]	wMkNN [%]
100	94.5	88.1
10	77.5	85.1

B. Application to real IVUS data for coronary plaque

1) Experimental settings: The performances of the normal MkNN and the proposed wMkNN for tissue characterization of coronary plaque are compared. In this experiment, RF signals are observed from two different sections of the blood vessel. One is used as the training data, and the other is used as the test data for classification.

The prototype vectors are selected from the training data. The feature vectors used as the prototype vectors are the power spectra that are calculated for each of the corresponding RF signals reflected from each tissue, i.e., fibrous tissue, lipid tissue, and fibrofatty tissue.

Thirty prototype vectors are selected at random from each of three kinds of tissue for both cases of the normal MkNN and of the proposed wMkNN. In addition, the width of window for the short-time Fourier transformation is 64 points for both methods.

2) Experimental results: Fig. 10 shows the tissue characterization results for the training data. Fig. 10 (a) shows the tissue composition given by a medical doctor by examining the dyed tissue looking through a microscope. Fig. 10 (b) shows the results of tissue characterization by the normal MkNN. The number of prototype vectors is manually adjusted to get the best accuracy of characterization. The number of prototype vectors for fibrous, lipid and fibrofatty tissues are 850, 100, and 300, respectively.

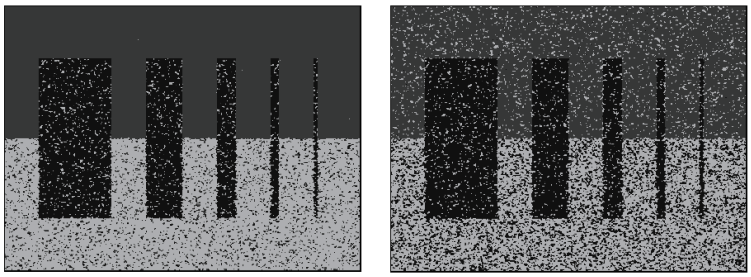
Fig. 10 (c) shows the results by the normal MkNN. Thirty prototype vectors are selected at random for each tissue. Fig. 10 (d) shows the results by the proposed wMkNN. The similar results to Fig. 10 (b) are obtained even if the prototype vectors are selected at random and even more the number of prototype vectors is smaller.

It is observed from those results that the accuracy of tissue characterization by the proposed wMkNN is kept good even if the number of prototype vectors is reduced and the prototype vectors are selected at random.

Fig. 11 shows the tissue characterization results for the test data. The characterization accuracy is a little bit behind though compared with that of Fig.10. However, the superiority of the proposed wMkNN over the normal MkNN still can be seen.

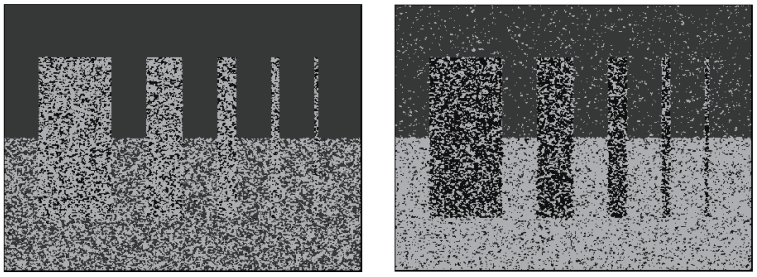
V. Conclusions

In this study, we have proposed an extended algorithm of a multiple k-nearest neighbor (wMkNN), which is applied to an intravascular ultrasound (IVUS)-based tissue characterization of coronary plaque. In the proposed algorithm, a weighted decision based on the distances between the input vector and the prototype (reference) vectors are employed instead of the majority decision in the labeling process of k-nearest neighbor (kNN).



(a) MkNN (k=5, k'=9) (b) wMkNN (k=5, k'=9, ε=0.05)

Fig. 8 Classification results. The number of reference vectors is 100 for each class. (a) Classification results by the normal MkNN [3]. (b) Classification results by the proposed wMkNN.



(a) MkNN (k=5, k'=9) (b) wMkNN (k=5, k'=9, ε=0.05)

Fig. 9 Classification results. The number of reference vectors is 10 for each class. (a) Classification results by the normal MkNN [3]. (b) Classification results by the proposed wMkNN.

The experimental results show that the accuracy of tissue characterization by the proposed method (wMkNN) is kept good even if the number of prototype vectors is reduced.

Future studies are to confirm further the effectiveness of the proposed method by applying to many other IVUS data, and to find out the best feature vectors for the best tissue characterization accuracy.

ACKNOWLEDGMENTS

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

[1] E. Falk, "Coronary plaque disruption." *Circulation*, vol. 92, pp. 657–671, 1995.

[2] J. B. Hodgson, S. P. Graham, A. D. Savakus, S. G. Dame, D. N. Stephens, P. S. Dhillon, H. S. D. Brands, and M. J. Eberle, "Clinical percutaneous imaging of coronary anatomy using an over-the-wire ultrasound catheter system," *The International Journal of Cardiovascular Imaging*, vol. 4, pp. 178–193, 1989.

[3] R. Kubota, E. Uchino, and N. Suetake, "Hierarchical k-nearest neighbor classification using feature and observation space information," *IEICE Electronics Express*, vol. 5, no. 3, pp. 114–119, 2008.

[4] E. Uchino, N. Suetake, R. Kubota, T. Koga, G. Hashimoto, T. Hiro, and M. Matuzaki, "An ROC performance validation of hierarchical k-nearest neighbor classifier applied to tissue characterization using IVUS-RF signal," *Proceedings of the 2009 RISP International Workshop on Nonlinear Circuits and Signal Processing*, pp. 333–336, 2009.

[5] T. Koga, S. Furukawa, N. Suetake, and E. Uchino, "Acceleration of the speed of tissue characterization algorithm for coronary plaque by employing GPGPU technique," *Proceedings of the 2011 International Conference on Multimedia Systems & Signal Processing*, pp. 52–57, 2011.

[6] J. C. Seabra, F. Ciompi, O. Pujol, J. Mauri, P. Radeva, and J. Sanches, "Rayleigh mixture model for plaque characterization in intravascular ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 1314–1324, 2011.

[7] A. K. Ghosh, P. Chaudhuri, and C. A. Murthy, "On visualization and aggregation of nearest neighbor classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 1592–1602, 2005.

[8] M. P. Moore, T. Spencer, D. M. Salter, P. P. Kearney, T. R. D. Shaw, I. R. Starkey, P. J. Fitzgerald, R. Erbel, A. Lange, N. W. McDiken, G. R. Sutherland, and K. A. A. Fox, "Characterisation of coronary atherosclerotic morphology by spectral analysis of radiofrequency signal: in vitro intravascular ultrasound study with histological and radiological validation," *Heart*, vol. 79, pp. 459–467, 1997.

[9] A. Nair, B. D. Kuban, E. M. Tuzcu, P. Achoenhagen, S. E. Nissen, and D. G. Vince, "Coronary plaque classification with intravascular ultrasound radiofrequency data analysis," *Clinical Investigation*, vol. 106, 2002.

[10] H. Seiji, K. Senya, and M. Sueharu, "Pattern recognition using average patterns of categorical k-nearest neighbors," *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 4, pp. 412–415, 2004.

[11] K. Tokunaga, E. Uchino, H. Tanaka, and N. Suetake, "Tissue classification of coronary plaque using intravascular ultrasound method by extended multiple k-nearest neighbor," *Proceedings of the International MultiConference of Engineers and Computer Scientists 2012*, pp. 39–42, IMECS 2012, 14-16 March, Hong Kong, 2012.

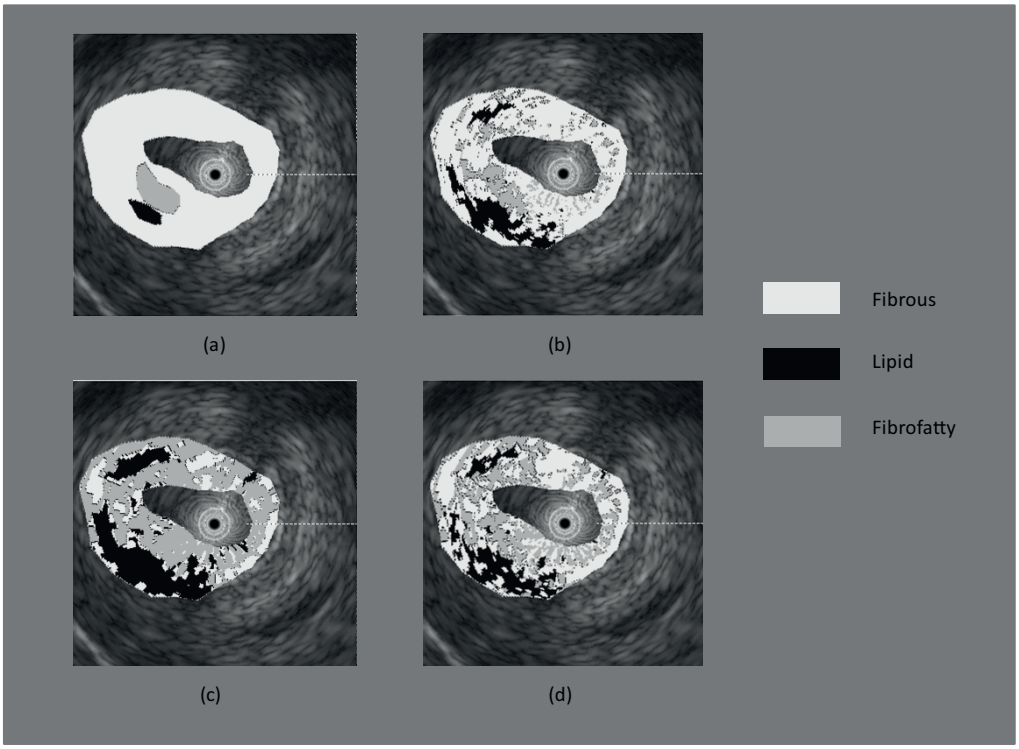


Fig. 10 Tissue characterization results by the normal MkNN and by the proposed wMkNN for the training data. (a) The tissue composition given by a medical doctor by examining the dyed tissue looking through a microscope. (b) The tissue characterization results by the normal MkNN. The numbers of prototype vectors for fibrous, lipid and fibrofatty tissues are 850, 100, and 300, respectively. (c) The tissue characterization results by the normal MkNN. The number of prototype vectors is 30 for each tissue. The prototype vectors are selected at random. (d) The tissue characterization results by the proposed wMkNN. The number of prototype vectors is 30 for each tissue. The prototype vectors are selected at random.

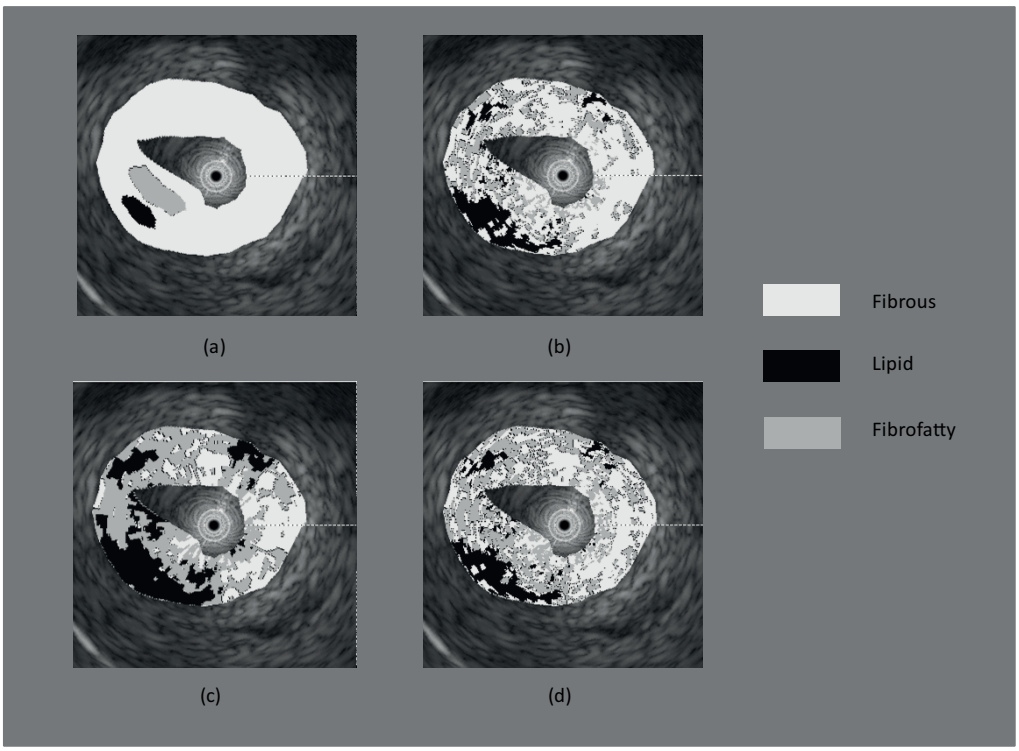


Fig. 11 Tissue characterization results by the normal MkNN and by the proposed wMkNN for the test data. (a) The tissue composition given by a medical doctor by examining the dyed tissue looking through a microscope. (b) The tissue characterization results by the normal MkNN. The numbers of prototype vectors for fibrous, lipid and fibrofatty tissues are 800, 130, and 450, respectively. (c) The tissue characterization results by the normal MkNN. The number of prototype vectors is 30 for each tissue. The prototype vectors are selected at random. (d) The tissue characterization results by the proposed wMkNN. The number of prototype vectors are 30 for each tissue. The prototype vectors is selected at random.