Improving Underwater Acoustic Echo Classification via Fusion of Multiple Information Processors

Su Yang

Department of Computer Science, Fudan University,
220 Handan Rd., Shanghai 200433, China

suyang@fudan.edu.cn

Abstract

Underwater signal classification is mostly based on frequency domain or time-frequency domain analysis so far. However, they are not enough to capture all useful information. To enhance the current solutions, we present a feature extraction method based on principal component analysis in reconstructed state space (RSS-PCA) for acoustic echo classification, with the emphasis on time domain. We utilized 3 kinds of features, RSS-PCA, spectra, and the combination of RSS-PCA and spectra, to classify the underwater signals backscattered from 5 categories of lake bottoms. The classification results show that: (1) The RSS-PCA feature and the spectra feature is alone effective for underwater echo classification. (2) The combination of both features performs much better than either feature alone. This confirms that the RSS-PCA feature is able to enhance the current solution and the fusion of the different features promises better performance. We attribute this to the following factor. Since reconstructed state space is a new domain to view signals, different natures of signals can then be revealed than those in frequency domain. As different feature extractors capture the profile of an object of interest from different viewpoints, fusion of multiple feature extractors usually leads to a more complete profile of the object.
Keyword: Acoustic Signal Classification, Feature Extraction, State Space Reconstruction, Principal Component Analysis (PCA).

I. Introduction

Underwater acoustic signals falls into two groups: acoustic echoes and acoustic emissions. The two groups are also referred to as active and passive sonar signals. Since the natures of the two categories of signals are quite different, the classification of either category of signals is an independent research area. This paper is dedicated to study acoustic echo classification. Underwater acoustic echo classification has been continuously receiving intensive investigations so far. Most feature extraction methods in the current literature are focused on frequency domain or time-frequency domain. Some representative methods are: wavelet packets in conjunction with linear predictive coding (LPC) [1], textural features derived from wavelet transformation [2], and partition of DFT data into subsets with constant variance [3]. In general, no single feature can capture all useful information. Different features working in collaboration with each other usually provide complementary information. As the underwater acoustic echo classification is an open problem so far, new feature extraction and classification methods are in urgent demand to augment the existing solutions. According to our investigations, time domain features also play an important role in underwater acoustic echo classification. However, few investigations in the current literature are devoted to time domain features. In this study, we apply principal component analysis [9] based on state space reconstruction [7,8] (RSS-PCA) as a feature extractor for active sonar signal classification. We firstly proposed this feature extractor in [6], aiming at passive sonar signal classification. We extend the usage of it to acoustic echo classification in this study and the emphasis is placed on investigating its performance with active sonar signals. By means of the state space reconstruction, the time sequence of interest can be embedded to a space with higher dimensionality. The benefit of applying state space reconstruction lies in that some regular patterns invisible in the time sequence can be revealed in the embedding space.

To investigate the contribution of RSS-PCA to active sonar signal classification, we conducted the following experiments. We classify the underwater echoes returned from 5 classes of objects, which are lake bottoms consisting of 5 different kinds of materials: rock, grit, scree, sand, and silt, using the RSS-PCA feature, a spectral feature, and the fusion of both. Here, the nearest neighbor classifier [11] is employed to perform the classification. The classification results justified two facts: (1) Either feature is alone effective for underwater acoustic echo classification. (2) The combination of both features promises a much better performance than either one alone. We attribute this to the following factors: (1) State space reconstruction provides a new insight into the different natures of active sonar signals in contrast to traditional means. (2) The fusion of different features could provide complementary information to each other and compensate for the limits of each other. Thus, fusion of features obtained in different domains is a way to enhance the existing solutions.

II. Feature Extractors

A. *Principal Component Analysis in Reconstructed State Space*

The state space reconstruction performed on a given time series $[s_k]$:k=1,2,…,N_T is described as follows. At first, we have to determine the values of two parameters, the delay time τ and the embedding dimension N. Then, we can construct $M=N_T-(N-1)\tau$ vectors $\{x_i|i=1,2,\dots,M\}$, where

$$x_i=(s_i,s_{i+\tau},s_{i+2\tau},\dots,s_{i+(N-1)\tau})^T. \tag{1}$$

There exist various ways in the literature to decide the parameter τ. Here, we employ the method presented in [4] because it is more suitable for processing underwater acoustic signals in accordance with our practice. Such method sets τ as the minimum duration during which the autocorrelation function of the time series $[s_k]$ approaches 0. The criterion to determine the embedding dimension N has been discussed in detail in our former report [10], that is, the embedding dimension resulting in the highest classification precision.

Let $X=[x_1,x_2,\dots,x_M]$ represent a matrix being composed of $\{x_i|i=1,2,\dots,M\}$. Such matrix forms a trajectory in the embedding space, which reveals the dynamical behaviors of the system behind the observed time series of interest.

Then, the pattern analysis in the embedding space is performed as follows. First, we construct a new base spanned by N vectors $\{u_1,u_2,\dots,u_N\}$. Let $\lambda_1\geq\lambda_2\geq\dots\geq\lambda_N$ denote the eigenvalues of XX^T.

The $\{u_1,u_2,\ldots,u_N\}$ are chosen to be the corresponding eigenvectors of XX^T. Then, x_i can be transformed to the new space spanned by $\{u_1,u_2,\ldots,u_N\}$ via

$$y_i=U^T x_i, \tag{2}$$

where $U=[u_1,u_2,\ldots,u_N]$ forms a matrix. Let $Y=[y_1,y_2,\ldots,y_M]$ and $X=[x_1,x_2,\ldots,x_M]$. On account of Eq. (2), it follows that

$$Y=U^T X. \tag{3}$$

It can then be derived that

$$XX^T=U\Lambda U^T, \tag{4}$$

where $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\ldots\lambda_N)$. According to Eq. (3) and (4), it is easy to deduce that

$$YY^T=U^T XX^T U=U^T U\Lambda U^T U=\Lambda. \tag{5}$$

It means that the correlations among different dimensions have been eliminated following the transformation defined in Eq. (3). Here, the eigenvalues $\lambda_1,\lambda_2,\ldots,\lambda_N$ are utilized as features because they indicate the characteristics of the reconstructed trajectory of interest along different dimensions in the space spanned by $\{u_1,u_2,\ldots,u_N\}$.

B. *Spectral feature*

The spectral feature is computed as follows. (1) Compute the power spectrum of every signal using Welch's averaged periodogram method, where a Hanning window is used and the computation is based on a 2-segment average without overlay. (2) The entire frequency axis is divided into some equal bins without overlay, where 10 bins are used. Then, the mean of the spectral values in every bin is used to form the pattern for every sample.

III. Nearest Neighbor Classifier

In this study, we employ the nearest neighbor classifier presented in [10] for signal classification. Following is the detailed implementation. Suppose that there are Q training samples and the pattern of the ith sample is denoted as $F_i=(F_{i1},F_{i2},\ldots,F_{iK})^T$: $i=1,2,\ldots,Q$. The class label of every training sample is prior known and denoted as $C(F_i)$: $i=1,2,\ldots,Q$. Let $f=(f_1,f_2,\ldots,f_K)^T$ represent the pattern of the sample to be classified. The distance between f and F_i is defined as

$$d(f,F_i)=\sqrt{\left(\frac{f_1-F_{i1}}{\sigma_1}\right)^2+\left(\frac{f_2-F_{i2}}{\sigma_2}\right)^2+\ldots+\left(\frac{f_K-F_{iK}}{\sigma_K}\right)^2} \tag{6}$$

where σ_j:$j=1,2,\ldots,K$ is the standard deviation of the jth attribute of the Q training samples, namely, $\{F_{ij}|i=1,2,\ldots,Q\}$. The function of $\{\sigma_j|j=1,2,\ldots,K\}$ is to compensate for the scale difference among different attributes. If

$$k=\arg\min_i\{d(f,F_i)\,|\,i=1,2,\ldots,Q\} \tag{7}$$

then, the class label of f is assigned to be $C(F_k)$. This means that f is classified into the class to which the kth training sample belongs.

IV. Classification experiments

The previously presented two features, the RSS-PCA feature and the spectral feature, are utilized for classifying underwater acoustic echoes with the nearest neighbor classifier presented in section 3. The data set used in the classification experiments contains 5 classes of underwater acoustic echoes returned from 5 categories of lake bottoms, which are composed of rock, grit, scree, sand, and silt, respectively. The number of the signal samples contained in every class is 144, 180, 179, 180, and 180, respectively. The length of each sample is 2000. Prior to the feature extraction, each sample was normalized to possess unit energy. We select at random half samples of every class to form the training set and the residual samples of every class are used for testing. We conduct 3 types of testing, classification with RSS-PCA, spectra, and the fusion of both features, respectively. For every group of testing, we run the classification test 10 times with randomly selected training and testing samples for every run. Then, we summarize the averaged classification accuracy over the 10 runs with regard to every class and the overall classification accuracy in Table 1. Here, in computing the RSS-PCA feature, the embedding dimension is set to be 30 (The reason will be explained later). It is obvious that the fusion of both RSS-PCA and spectral features outperforms either single one. The performance improvement achieved by using the fusion of the two features in contrast to the use of every signal feature is listed in Table 2. Here, the performance improvement is obvious. For every class, the combination of both features leads to better performance in contrast to every single one. According to Table 1, the spectral feature is far better than the RSS-PCA feature in classifying the echoes returned from rock, grit, and scree. With regard to the echoes returned from sand, the spectral feature also promises better performance. Only for the echoes returned from slit, the RSS-PCA feature outperforms the spectral feature, but not much. However, when combining the two features to classify the 5 classes of signals, the performance is improved obviously in classifying every class (See Table 2). We attribute the interesting phenomenon to the following cause. Since every single feature extractor views the signals from a specific viewpoint, it is only able to reveal a portion of the nature of the signals of interest. Different feature extraction methods reveal different aspects of the signals' nature. When different features are combined, the information regarding different aspects, which are in general complementary to each other, are organized. This benefits the classification. In this study, since reconstructed state space is a new domain to view signals, different natures of signals can be revealed than those in frequency domain.

Table 1: Classification rate using different features (%)

	Rock	Grit	Scree	Sand	Silt	Overall
Spectra&RSS-PCA	96.3889	95.2222	97.0000	99.6667	97.7778	97.2454
Spectra	92.2222	87.7778	92.0000	99.0000	90.0000	92.1991
RSS-PCA	77.0833	81.3333	87.8889	96.8889	92.2222	87.5000

Table 2: Performance improvement by using combination of features (%)

	Rock	Grit	Scree	Sand	Silt
Spectra	4.1667	7.4444	5.0000	0.6667	7.7778
PCA30	19.3056	13.8889	9.1111	2.7778	5.5556

According to our observation, the classification performance is subject to the embedding dimension. A too low or too high embedding dimension will result in performance degradation in terms of classification. So, the choice of embedding dimension plays an important role in the analysis based on state space reconstruction. It has been pointed out in [5] that the embedding dimension determination should be application-dependent. As for the practice in this study, the optimal embedding dimension is that resulting in the highest classification rate. To approach the optimal embedding dimension, we conduct the following tests. We compute the classification rates using the RSS-PCA features computed under different embedding dimensions. The classification results are presented in Table 3. Obviously, when the embedding dimension is 30, the highest classification rate is achieved. So, in previous classification experiments, we let the embedding dimension be 30.

Table 3: Classification rate using RSS-PCA under different embedding dimensions (%)

Dimension	Rock	Grit	Scree	Sand	Silt	Overall
5	72.0833	57.6667	79.5556	69.5556	74.2222	70.5556
10	77.3611	60.3333	86.0000	80.3333	81.1111	77.0139
15	74.8611	71.6667	84.7778	86.8889	84.8889	80.8565
20	76.6667	75.4444	86.2222	93.7778	92.0000	85.1620
25	74.4444	78.5556	88.4444	97.7778	92.0000	86.7361
30	77.0833	81.3333	87.8889	96.8889	92.2222	87.5000
35	76.3889	81.3333	88.0000	97.6667	88.0000	86.6898
40	76.1111	82.0000	87.8889	97.6667	85.8889	86.3194
45	77.0833	81.7778	87.7778	97.8889	86.1111	86.5046
50	76.1111	79.2222	87.3333	98.0000	86.0000	85.7176

V. Conclusion

The contributions of this study are as follows. (1) Principal component analysis in reconstructed state space is experimentally shown to be effective for underwater echo classification. (2) We justified through experiments that fusion of different features could provide complementary information to each single feature and thus promise better performance in terms of underwater acoustic echo classification. In this study, the fusion of RSS-PCA feature and the spectral feature outperforms either single feature in the sense of classification. Actually, RSS-PCA is not a powerful feature for underwater acoustic echo classification if used alone. When combined with spectra, however, it greatly improved the spectra based classification. In this sense, the RSS-PCA is a useful feature because it provides complementary information overcoming the limits of the spectral feature.

References

[1] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, G. J. Dobeck, “Underwater target classification using wavelet packets and neural networks”, IEEE Transactions on Neural Networks, 2000, Vol. 11, Issue 3, pp. 784-794.

[2] K. S. Thyagarajan, T. Nguyen, C. E. Persons, “Maximum likelihood approach to image texture and acoustic signal classification”, IEE Proceedings-Vision Image and Signal Processing, 1999, Vol. 146, Issue 1, pp. 34-39.

[3] R. J. Kenefic, “An algorithm to partition DFT data into sections of constant variance”, IEEE Transactions on Aerospace and Electronic Systems, 1998, Vol. 34, Issue 3, pp. 789-795.

[4] T. S. Parker, L. O. Chua, *Practical numerical algorithms for chaotic systems*. New York: Springer-Verlag, 1989.

[5] T. Schreiber, "Interdisciplinary application of nonlinear time series methods", Physics Reports, 1999, vol. 308, pp. 1-64.

[6] S. Yang, "High-dimensional shape analysis in state space: methodology for oceanic signal classification", Electronics Letters, 2004, vol. 40, pp. 513-515.

[7] N. H. Packard, J. P. Scutchfield, J. D. Farmer, R. S. Shaw, "Geometry from a time series", Physical Review Letters, 1980, vol. 45, pp. 712-716.

[8] F. Takens, "Detecting strange attractors in turbulence", in *Lecture Notes in Mathematics*, vol. 898, Springer, 1981, pp. 366-381.

[9] I. Jolliffe, *Principal component analysis*. New York: Springer-Verlag, 1986.

[10] S. Yang, "Nonlinear signal classification in the framework of high-dimensional shape analysis", IEEE Transactions on Circuits and Systems (II), 2005, Vol. 52, No. 8, pp. 512-516.

[11] A. V. Oppenheim, R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

Su Yang is with Department of Computer Science and Engineering at Fudan University, Shanghai, China. His research interests include Computer Vision, Signal Processing, and Pattern Recognition.