Estimation and Extraction of Unstable Frequency Lines of Acoustic Signal Using Neural Network

*Seok-Wun Ha, **Soo-Bok Hwang, ***Jae-Chang Kim

Abstract

In passive sonar, underwater moving objects are identified by the acoustic sounds they transmit. The spectrum of these sounds show features about the mechanism of the sound source, these features are discrete frequencies on the spectrum and frequency lines on the spectrogram. Variability in the underwater environment produce discontinuous broken or unstable fluctuating frequency lines.

In this paper, we propose an efficient algorithm that estimate continuities of the discontinuous frequency lines and extract presence of the unstable frequency lines using neural networks and represent the proposed algorithm shows good performance in estimation and extraction the unstable frequency lines through the experiments.

I. Introduction

Passive sonar system is designed to exploit the radiated noise of underwater acoustic targets for the purpose of detecting and classifying the underwater acoustic targets. The main features of a radiated noise are tonals and frequency lines.1)-2)

Tonal is a discrete frequency component and frequency line is the line formed in spectrogram by a tonal of constant frequency in each frame.3) Frequency lines are of particular importance for target detection and classification. The frequency lines of spectrogram are divided into stable and unstable lines, and may produce a discontinuous (or broken) line by variability in the underwater environmental conditions.

Therefore, the estimation and extraction of frequency lines is a crucial issue in the so-called classification process of passive sonar system. But, the stable frequency lines make this task easy but non efficient since poor significance in specific acoustic targets which are fast and small mobiles.

In this paper, an efficient neural network of the estimation of discontinuous frequency lines and a neural network for separate extraction of unstable frequency line has nonlinear characteristics which vary with time are presented.4)

* Dept. of Computer Science, Gyeongsang National University,
** Agency for Defence Development,
*** Dept. of Electronics Engineering, Pusan National University.

The input to a neural network of the estimation of frequency lines is the modified spectrogram which is an output of TPM(Two Pass Mean)5) and frequency line matching method applied to the original spectrogram. TPM represents the process of whitening the noise spectrum and frequency line matching method eliminates the false detection of noises with some high level values after TPM executed.

Structure of the neural network of the estimation of discontinuous frequency lines is applied to a 4 by 7 pixel section of the spectrogram map. It is assumed that each neuron corresponding with the pixel of frequency bin receives simultaneously the excitation of frequency line coming from three directions, namely, upper left, upper center, and upper right.

Thus, activities of the three directions are then summed to determine the activation level of the neuron. The output of each neuron determines the estimation of frequency line of the corresponding pixel of frequency bin. The synthetic test signals is constructed by adding Gaussian noise with zero mean and unit variance to the various discontinuous frequency lines, it is confirmed that the proposed method shows good performance.

The output of a neural network of the estimation of frequency lines is input to the neural network for separate extraction of unstable frequency lines. This network consists of a MLP(Multi-Layer Perceptron) trained using the back propagation algorithm.4)

Each data frame representing the frequency lines of the spectrogram utilizes a two-dimensional array with a 11 by 13 pixel section of the time / frequency bin map. Thus the utilized MLP needs 143 input nodes. The hidden layer consists of 23 nodes and output layer 2 nodes. The two nodes of output layer shows the existence or no-existence of unstable frequency lines. A MLP is trained with examplars of specific line patterns which are composed of various unstable frequency lines and non-specific patterns which are composed of only background noise with no frequency line and stable frequency line.

By the simulation using the test signals that have unstable frequency lines generated from various characteristic conditions of underwater acoustic targets and application to underwater vessels, it is confirmed that the proposed neural network for the passive sonar system shows good performance in extracting the unstable frequency lines.

II. Overall Concept

A block schematic diagram of the system for estimation and extraction of the unstable frequency lines is shown at Figure 1.

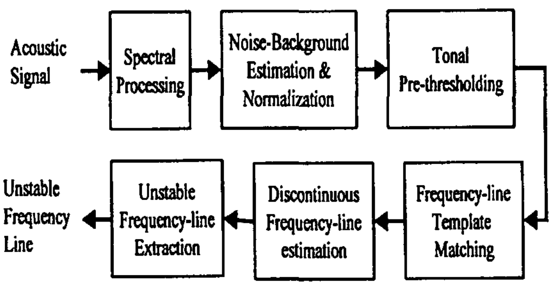

Figure 1. Block diagram of the overall system.

Overall procedure for extraction of unstable frequency lines are composed of four interactions process as follows.

① Spectrum normalization and tonal pre-detection : Analogue recordings of underwater sound are digitized into binary form and processed by FFT, background noise normalization, and pre-detection thresholding techniques to produce a spectrogram.

② Frequency line template matching : Frequency line matching method is applied to a 3 by 5 pixel section of the spectrogram to eliminate the false detection of noises.

③ Discontinuous frequency line estimation : Algorithm of neural network is applied to a 4 by 7 pixel section of the frequency line detected map to estimate continuities of discontinuous broken frequency lines.

④ Unstable frequency line extraction : A neural network, MLP, is trained with various frequency line patterns of the size of 11 by 13 to extract the existence of the unstable frequency lines.

III. Spectrum Analysis and Tonal Detection

1. Spectrum Noramlization

Spectrum analysis is employed in signal detection systems to determine the presence and frequency of signals. It is performed by taking the sampled waveform in the time domain and converting it to the frequency domain by means of a frequency domain by means of a FFT(fast Fourier transform). The magnitude spectrum is then obtained from the FFT output by computing the envelope of each frequency component, that is tonal bin. Large peaks in the magnitude spectrum indicate the presence of high noise values and signals. In order to facilitate the search for the signals, the noise background must be whitened, usually at unit height.

The process of whitening the noise spectrum is called normalization. And it is mathematically defined by

$$X(k)=F(k)/Z(k) \tag{1}$$

where F(k) is the magnitude in frequency bin k, and Z(k) is the noise mean estimate in the bin k.

Typically, there are four noise mean estimation techniques- TPM, OTA, S3PM, SAXA. The TPM is currently in widespread use and requires relatively few computations, so we use the TPM in our experiments. The following steps describe the technique:

Step 1 : A local first-pass mean is calculated for each bin by

$$\overline{F(k)} = \sum_{i\in\Omega_k} \frac{F(k)}{K} \tag{2}$$

where Ω(k) is a local window denotes the set of bin numbers that can be used to estimate Z(k), then k is the bin number of interest is centered in Ω(k) and (2M+1) is the length of the local window.

$$\Omega(k)= \{k-M,\ldots,k,\ldots, k+M\} \text{ for } k\geq 0 \tag{3}$$

Step 2 : Each bin value is then compared against its respective local first-pass mean and is replaced, i.e.,

$$E(k)=\begin{cases} F(k) & if \quad X(k)\langle\alpha\overline{F(k)} \\ \overline{F(k)} & if \quad X(k)\langle\alpha\overline{F(k)} \end{cases} \tag{4}$$

where $\alpha(=2.0)$ is the shearing threshold ratio.

Step 3 : The noise mean estimate is then calculated for each bin using the smoothed data, this is the second-pass mean, i.e.,

$$Z(k) = \sum_{i \in \Omega_k} \frac{E(k)}{K} \tag{5}$$

In this approach, high bin values are replaced by the first-pass mean prior to obtaining the noise mean estimate.

Figure 2 shows the results, here (a) is the raw magnitude spectrum, (b) is the estimated background noise, and (c) is the normalized spectrogram. Figure 3 shows the normalized spectrogram of an underwater moving object.

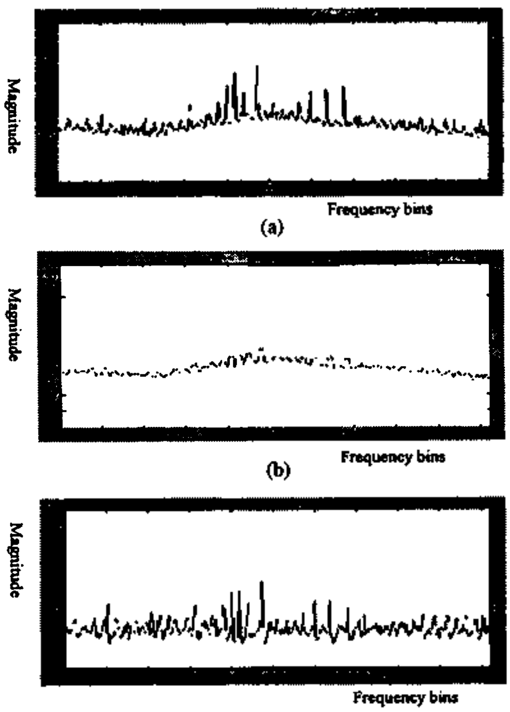

Figure 2. (a) the raw magnitude spectrum, (b) the background noise, and (c) the normalized spectrum.

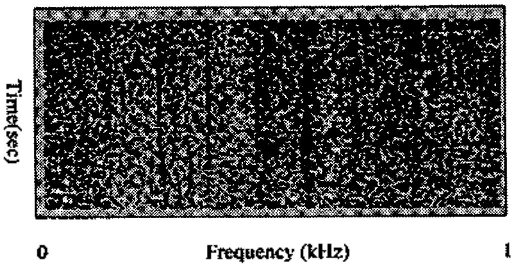

Figure 3. The normalized spectrogram of an underwater moving object.

2. Tonal pre-detection

The detection threshold is adaptively applied in the normalized spectrum of each frame and the discrete frequency components that exceed the detection threshold are assumed in the signal frequency components. The frequency line is the line formed in spectrogram by a signal frequency component in each frame. Frequency lines are divided into a stable and unstable lines.

In case of constant signal frequency, the frequency line indicates the stable line. Unstable frequency line has nonlinear characteristics which vary with time. In this case, a weak signal frequency below detection threshold enables the frequency line to detect the discontinuous or unstable frequency line. Therefore, first of all, it is necessary to enhance the continuity of frequency line.

In this paper, the adaptive pre-detection threshold, is a α-fractional of the detection threshold, is adopted to detect a weak signal. The adaptive pre-detection threshold T_{pd} is applied to the normalized spectrum X(k). In case of exceeding the pre-detection threshold $T_{pd} = \alpha T_d$ (α =0.8), its value indicates 1, otherwise 0. The output of spectrum Y(k) represents binary value

$$Y(k) = \begin{cases} 1 & if \quad X(k) < T_{pd} \\ 0 & if \quad X(k) \geq T_{pd} \end{cases} \tag{6}$$

3. Frequency line template matching

The frequency line template matching method is adopted to eliminate the false detection of noises with some high level values after TPM executed.

For frequency line template matching, we obtain a 3 by 5 local window of the time / frequency of binary spectrogram map with one present frame and two past frames.

The conditions of frequency line template matching is as follows. First, a convolution sum of pre-detected binary map Y(i,j) and template T(i,j) is equal to 3, and a convolution sum of normalized map X(i,j) and template T(i,j) is greater than the detection threshold T_d a s in equation (7).

We determine that the continuous frequency line exist as the condition is completed.

$$\sum_{i=t-2}^{t_0} \sum_{j=k-2}^{k+2} Y(i,j)T(i,j) = 3 \quad and$$
$$\frac{1}{3} \sum_{i=t-2}^{t_0} \sum_{j=k-2}^{k+2} X(i,j)T(i,j) > T_d \tag{7}$$

Figure 4(a) represents the templates of continuous frequency lines, (b) the generated spectrogram with unstable frequency lines, (c) the pre-detection thresholded time/frequency map, and (d) the frequency line template matched time/frequency map.

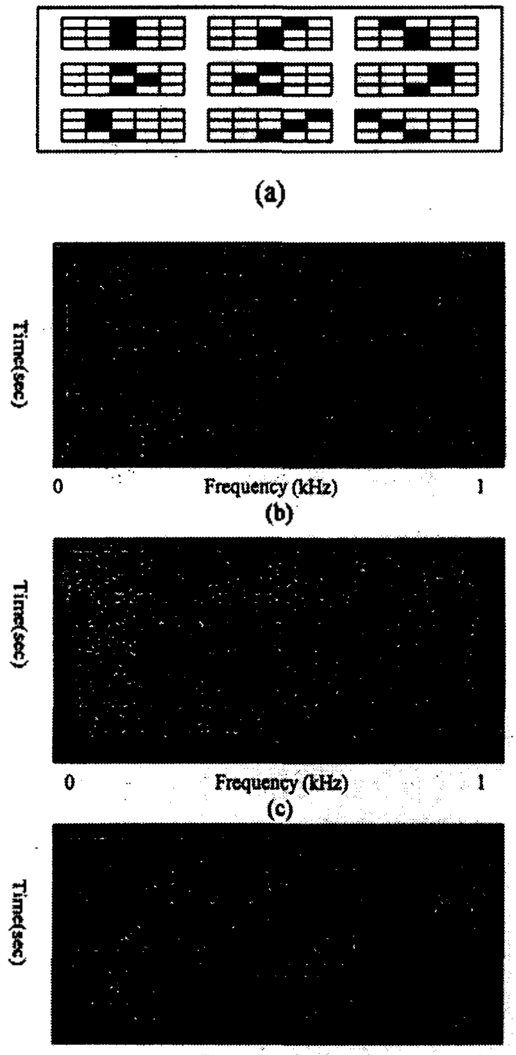

Figure 4. (a) The templates of continuous frequency lines, (b) a spectrogram generated for simulation, (c) the pre-detection thresholded map, and (d) the matched map.

IV. Estimation and extraction of Unstable Frequency Lines

1. Estimation of the discontinuous frequency lines

The frequency line detected using frequency line matching method also contains the discontinuous components. Especially, in case that the frequency lines vary abruptly it is difficult to understand its characteristics owing to the discontinuity of frequency line. Therefore, the continuity estimation of discontinuous frequency line is executed by estimating the signal magnitude with reference to activation value of frequency bin from frame passed.

The neural algorithm for estimation of the unstable frequency lines shown in Figure 5.

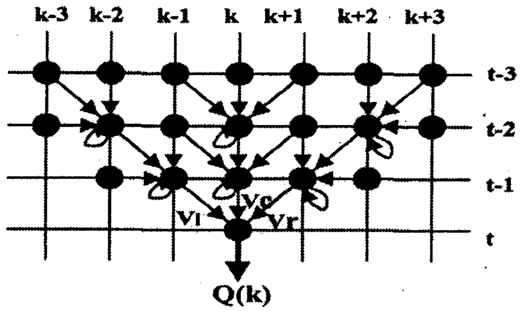

Figure 5. The neural algorithm for estimation of discontinuous frequency lines.

In the Figure 5, each section is divided into a 4 by 7 pixel representing the time / frequency bin data in the spectrogram and the frequency bin of the corresponding pixel is estimated from the mutual relations of neurons in this section.

It is assumed that each neuron corresponding with the pixel of frequency bin receives simultaneously the excitation of frequency line coming from three directions, namely, upper left, upper center, and upper right.

Thus, all of the three directions are then summed to determine the activation level of the neuron. The output of each neuron determines the estimation of frequency line of the corresponding pixel of frequency bin. The procedure is as follows.

Step 1: Calculate the excitation coming from each direction. First, the excitation of frequency line coming from upper left direction, Vl(k) is

$$V_l(k) = \beta Z_{t-1}(k-1) + \gamma\{V_{l1} + Z_{t-2}(k-1) + Z_{t-1}(k-2)\}$$
$$V_{l1} = \beta Z_{t-2}(k-2) + \gamma\{Z_{t-3}(k-3) + Z_{t-3}(k-2) + Z_{t-2}(k-3)\} \tag{8}$$

And the excitation of frequency line coming from upper center direction Vc(k) is calculated by

$$V_c(k) = \beta Z_{t-1}(k) + \gamma\{V_{c1} + Z_{t-2}(k-1) + Z_{t-2}(k+1)\}$$
$$V_{c1} = \beta Z_{t-2}(k) + \gamma\{Z_{t-3}(k) + Z_{t-3}(k-1) + Z_{t-3}(k+1)\} \tag{9}$$

Also, the excitation of frequency line coming from upper right direction Vr(k) is calculated by

$$V_r(k) = \beta Z_{t-1}(k+1) + \gamma\{V_{r1} + Z_{t-2}(k+1) + Z_{t-1}(k+2)\}$$
$$V_{r1} = \beta Z_{t-2}(k+2) + \gamma\{Z_{t-3}(k+2) + Z_{t-3}(k+3) + Z_{t-2}(k+3)\} \quad (10)$$

Step 2: All of the three directions are then summed to determine the activation level of the neuron.

$$V(k) = V_r(k) + V_c(k) + V_l(k) \quad (11)$$

Step 3: The output of each neuron determines the estimation of frequency line of the corresponding pixel of frequency bin.

$$Q(k) = \lambda V(k) \quad (12)$$

where β=1/2, γ=1/3, λ=0.7.

Figure 6 represents the time/frequency map estimated the unstable frequency lines

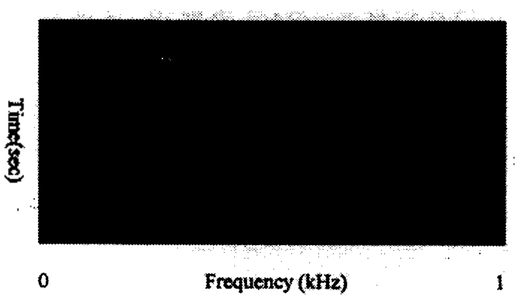

Figure 6. The result of estimation of disontinuous frequency lines from the raw spectrogram generated for simulation.

2. Extraction of unstable frequency lines

The estimation and tracking of frequency lines is crucial issue in the so-called passive sonar classification. The stability of frequency line makes this task easy but non efficient since with poor significance. This is why, most of the time, operator's attention is focused on the unstable frequency lines.

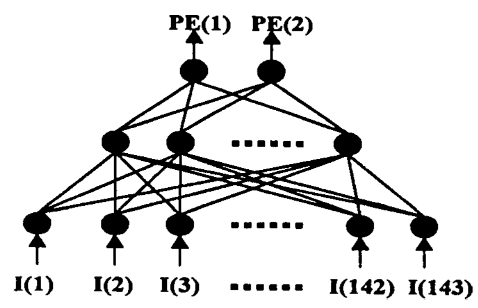

Figure 7. The MLP neural network for extraction of unstable frequency lines.

In this paper, to extract separately the unstable frequency lines the MLP neural network is utilized as shown in Figure 7.

In Figure 7, number of input neuron has a 11 by 13 pixel section, namely 143 nodes. The output layer consists of 2 nodes and obtains the opposite result of existence of unstable frequency lines or no-existence of unstable frequency lines. In the output layer, [PE(1), PF(2)] produces a [1 0] when the unstable frequency line is existed and a [0 1] when it is not existed. The hidden layer is determined by the number of training pattern and consists of 23 nodes.

The processing method of MLP neural network is as follows.

Step 1: Train the MLP using the backpropagation training algorithm.

Step 2: Train the result of existence of unstable frequency lines as [1 0] and the result of no-existence of unstable frequency lines as [0 1] with reference to the training pattern shown in Figure 8.

Step 3: Input the data section centered in each pixel in the spectrogram to the MLP network and extract whether a specific frequency lines exist or not.

Figure 8 shows the exemplars of the training pattern representing the result of existence of unstable frequency lines and no-existence of unstable frequency lines, the pattern of stable frequency lines and the pattern of unstable frequency lines.

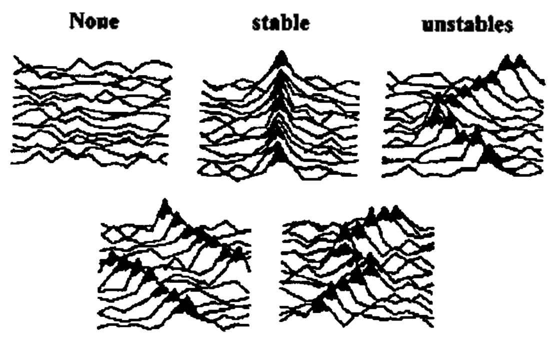

Figure 8. The exemplars of the training pattern representing the result of 'existence' or 'no-existence' of unstable frequency lines.

Figure 9 shows, (a) the time / frequency map which the specific and non-specific frequency lines exist simultaneously, and (b) the resulting output of non-specific frequency lines obtained using the neural network for extraction of specific frequency lines.

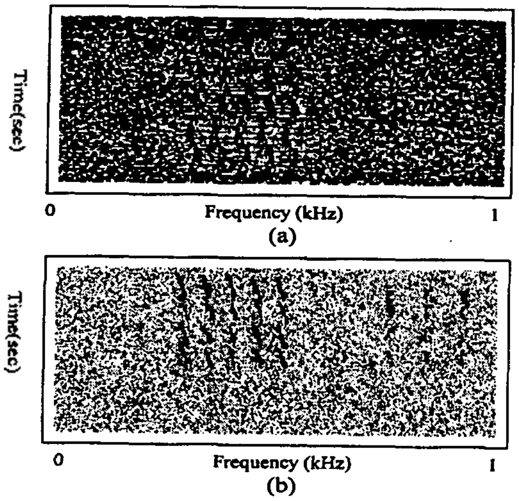

Figure 9. (a) The modified spectrogram of a real sound of an object with discontinuous unstable frequency lines, and (b) the resulting output extracted the unstable frequency lines by the proposed neural algorithm.

V. Conclusion

In this paper, a neural network algorithm for extraction of unstable frequency lines of an acoustic underwater object is proposed.

First, to enhance the detection of weak signals and eliminate the false detection of noises with some high level values, the adaptive pre-detection threshold and the frequency line matching method is adopted. And then, the two stage's hierarchical neural network for estimation of discontinuous frequency line and extraction of unstable frequency line is proposed.

By the simulation using the test signals that have unstable frequency lines generated from various characteristic conditions of underwater acoustic objects. From the expremental results of the modified spectrogram of an underwater object, it is confirmed that the proposed neural algorithm shows a good performance in extracting the unstable frequency lines.

References

1. R. J. Urick, *Principles of Underwater Sound,* McGraw-Hill, New York, pp. 328-353, 1983.
2. R. O. Nielson, *Sonar Signal Processing,* Artech House, Boston, pp. 143-185, 1991.
3. L. Cohen, "Time-Frequency Distributions - A Review," *Proc. IEEE,* vol. 77, pp. 941-981, 1989.
4. C. P. Sheppard, C. R. Gent, and SD-Scicon Uk Limited, "A Neural Network Based Sonar Classification System," *Underwater Systems Design,* pp. 372-375, Nov./Dec. 1991.
5. W. A. Struzinski and E. D. Lowe,"A Performance Comparison of Four Noise Background Normalization Schemes Proposed for Signal Detection Systems," *J Acoustical Society of America,* vol. 76, no. 6, pp. 1738-1742, Dec. 1984.

▲Seok-Wun Ha

Seok-Wun Ha was born in Pusan, Korea on Sep. 25, 1956. He received B.S., M.S., and Ph.D. all in electronics engineering from Pusan National University, Pusan, Korea in 1981, 1986, and 1995 respectively. Since 1995, he has been with the department of computer science at Gyeongsang National University as a assistant professor. He current research interests in digital signal processing, image processing, and neural networks.

▲Soo-Bok Hwang

The Journal of The Acoustical Society of Korea, Vol. 15, No. 6, 1996.

▲Jae-Chang Kim

The Journal of The Acoustical Society of Korea, Vol. 15, No. 6, 1996.