Classification of Environmentally Distorted Acoustic Signals in Shallow Water Using Neural Networks: Application to Simulated and Measured Signals

*Young-Nam Na, *Joung-Soo Park, *Duck-Hong Chang and **Chun-Duck Kim

Abstract

This study attempts to test the classifying performance of a neural network and thereby examine its applicability to the signals distorted in a shallow water environment. Linear frequency modulated (LFM) signals are simulated by using an acoustic model and also measured through sea experiment. The network is constructed to have three layers and trained on both data sets. To get normalized power spectra as feature vectors, the study considers the three transforms: short-time Fourier transform (STFT), wavelet transform (WT) and pseudo Wigner-Ville distribution (PWVD). After trained on the simulated signals over water depth, the network gives over 95 % performance with the signal to noise ratio (SNR) being up to -10 dB. Among the transforms, the PWVD presents the best performance particularly in a highly noisy condition. The network performs worse with the summer sound speed profile than with the winter profile. It is also expected to present much different performance by the variation of bottom property. When the network is trained on the simulated signals, it classifies over 90 % of the measured signals. On the contrary, when the network is trained on the measured signals, it gives a little better results than that trained on the simulated data. In conclusion, the simulated signals are successfully applied to training a neural network, and the trained network performs well in classifying the signals distorted by a surrounding environment and corrupted by noise.

I. Introduction

There has been an increasing interest in using neural networks to classify passive sonar data [1-5]. Several papers dealing with transient signal classification have been published in recent proceedings and journals. For example, Desai and Shazeer [6], Hemminger and Pao [7] found a back-propagation (BP) network to be useful for classifying targets based on transients. Kundu et al.[8] successfully used a hidden Markov model and a multilayer perceptron network to classify transient signals. Meanwhile, Na et al.[9,10] found that a BP network could be successfully applicable to tracking tonals of low frequency, and even to filtering out ambient noise from underwater noisy signals.

While these studies have been encouraging, there is considerable need to investigate further the use of neural network in the analysis of passive sonar data. In particular, while the classification of real ocean data is the ultimate goal of the sonar classifier, it is very difficult to access neural network capabilities when the nature of the source data is not well understood. Moreover, most of these studies have not considered or defined environmental conditions in which acoustic waves propagate. Consequently, it has been hard to estimate the applicability of neural networks to the signals distorted by the surrounding environment. The signals inevitably undergo distortions by the environment in which they propagate.

Few papers have been published to deal with the signals distorted by surrounding environment. Field et al.[11] tried to classify the signals distorted by the environment and found that the network was able to recognize the received signals with 90 % accuracy. However, the network was tested under limited environmental conditions of no noise and relatively short range (5 km) compared with the water depth (1.3 km). In these conditions, the traveling waves suffer one surface or/and bottom interaction, yielding small distortion and loss. Moreover, the study dealt only simulated signals and covered small portion of water column, the coverage being 125 m from the bottom.

This study assumes the environment to be shallow water, and considers environmental variations in sound speed profile (SSP) and bottom property. The linear frequency modulated (LFM) signals are simulated at four range steps (2, 5, 10 and 20 km) and then white Gaussian noise is added to test the network performance on noisy

* Agency for Defense Development, Chinhae
** Pukyong National University, Pusan
Manuscript Received : May 24, 1998.

signals. The LFM signals are chosen because they are relatively simple to be generated but enable to analyze the distorted signals with ease. They are regarded to be enough to accommodate the signal distortion caused by the environment. Time signals are simulated through an acoustic model to accommodate the environmental effects. To the simulated signals, noise is added to produce noisy signals. In addition, real measured signals are employed to test the network performance. The first goal of the study is directed to examining the applicability of the simulated signals to train a network. The second one is to examine the applicability of the trained network to the simulated and measured signals.

II. Theory

A. Spectrum Estimation for Non-Stationary Signals

Many temporal signals that are inherently non-stationary can be regarded as stationary if viewed over a short period of time. Since the frequency content of stationary signals is a very powerful method of description, the time history of the power spectrum is an important tool for characterizing signals whose statistical properties are changing slowly with time. This method of characterization has been especially important for speech processing and for sonar signal processing [12-15].

In this study, three methods are considered: short-time Fourier transform (STFT), pseudo Wigner-Ville distribution (PWVD), and wavelet transform (WT). The theory on PWVD is described in the accompanying paper [16]

Short-Time Fourier Transform (STFT)

The most direct approach to computing the time history of the power spectrum is to view the recorded data through a moving average window whose length corresponds to the time over which the data can be assumed to remain stationary. The Fourier transform of the windowed data is known as the STFT. The STFT of the given data is defined by [17]

$$X[n,\omega] \equiv \sum_{k=-\infty}^{\infty} X[k]W[n-k]e^{-j\omega k}, \tag{1}$$

where the window extends from n–L to n+L. The power spectral estimate is given by the short-time periodogram

$$\hat{P}_x[n,\omega] = \frac{1}{2L+1}|X[n,\omega]|^2 \tag{2}$$

The time history of the spectrum comprises the so called spectrogram (or lofargram in sonar signal processing).

Wavelet Transform (WT)

Instead of decomposing a signal into harmonic functions $e^{j\omega t}$ in the Fourier transform, in the WT, one can classify a signal by a series of orthogonal basis functions of finite length. Recently, much attention has been directed to this new transform, especially in signal and image processing [18]. An important property of the WT is its ability to capture major characteristics of transient or non-stationary signals, simply by changing the scales of local structure in the time-scale plane.

For a time signal $s(t)$, the WT may be expressed as

$$C(a,b) \equiv \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \overline{\Psi}\left(\frac{t-b}{a}\right) s(t)dt, \tag{3}$$

where Ψ is an analyzing wavelet, a and b denote time dilation and translation, respectively.

This study employs the wavelet such that its frequency axis is partitioned into successive octaves called scales. For computational efficiency, integral power of 2 is used for frequency partitioning so that the wavelet is scaled and translated as $\Psi(2^i-k)$, where i, k are integers. The analyzing wavelet $\Psi(2^i t-k)$ is obtained from $\Psi(t)$ by dilating $1/2^i$ and translating $k/2^i$. In this case, the WT becomes

$$C(2^i,b) \equiv \frac{1}{\sqrt{2^i}} \int_{-\infty}^{\infty} \overline{\Psi}\left(\frac{t-b}{2^i}\right) s(t)dt. \tag{4}$$

As a mother wavelet Ψ, a modulated Gaussian [18] is employed

$$\Psi(t) = e^{j\omega_0 t} e^{-t^2/2}. \tag{5}$$

The Fourier transform of Eq. (5) has the form

$$\hat{\Psi}(a,\omega) = ae^{-(\omega-\omega_0/a)^2 a^2/2}, \tag{6}$$

where ω_0/a is analysis frequency. The bandwidth of the WT, BW, is proportional to $1/a$, thus comprises a constant relative bandwidth since bandwidth divided by analysis frequency is constant.

B. Conjugate Gradients (CG)

The problem of minimizing continuous, differentiable functions of many variables is one that has been widely studied, and many of the conventional approaches are directly applicable to the training of neural networks. This study introduces the CG algorithm by Johanson et al.[19].

For a quadratic error function, an iterative equation may be derived in the following form through some manipulation [20]

$$w_{j+1} = w_j + \alpha_j d_j. \tag{7}$$

Here, w_j and d_j denote the point vector of j-th step, and vector which is mutually conjugate with respect to Hessian matrix **H**, respectively. This represents a succession of steps parallel to the conjugate directions, with step lengths controlled by the parameters α_j.

After some manipulations are made, α_j can be written in the form [20]

$$\alpha_j = \frac{g_{j+1}^T H d_j}{d_j^T H d_j}. \tag{8}$$

It can be shown that if the weights are incremented using Eq. (7) with the α_j given by Eq. (8), then the gradient vector g_j at the j-th step is orthogonal to all previous conjugate directions [19]. It therefore follows that after steps the components of the gradient along all directions have been made zero, and so one will have arrived at the minimum of the quadratic form.

The problem is how to construct a set of mutually conjugate directions. This can be achieved by selecting the first direction to be the negative gradient $d_1 = g_1$, and then choosing each successive direction to be linear combination of the current gradient and previous search direction.

$$d_{j+1} = -g_{j+1} + \beta_j d_j. \tag{9}$$

The coefficient β_j can be found by the same procedure as with α_j

$$\beta_j = \frac{g_{j+1}^T H d_j}{d_j^T H d_j}. \tag{10}$$

By examining Eqs.(7) and (9), one can notice that they are very similar to traditional BP algorithm with momentum. However, there are two differences: In traditional BP, the step size is fixed, while in the CG algorithm, it is carefully chosen to be the line minimum along the search direction. Also, in traditional BP the momentum is usually fixed. In the CG algorithm, the momentum is optimally adjusted for each new step. These are actually two quite significant differences, and they account for the typically huge difference in performance. In order to gain a more global view, an simulated annealing algorithm is employed in addition to the CG which tends to sink into any convenient minimum [19,20].

III. Signal Simulation and Measurement

A. Model Simulation

This study considers a LFM signal of center frequency 200 Hz and bandwidth 100 Hz so that the signal sweeps up or down in 150-250 Hz. The source signal is generated by the following equation [21]

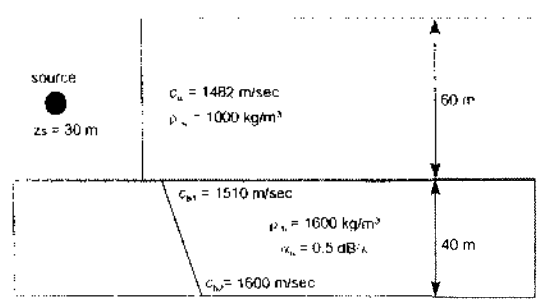
$$s(t) = \sin[2\pi(f_o t + mt^2/2)], \tag{11}$$

where f_o = center frequency (200 Hz), m = bandwidth (100 Hz). For the upsweep (class A) signal, the time goes from $T/2$ to $T/2$ and for the downsweep (class B) signal from $T/2$ to $-T/2$. The sampling frequency is 1024 Hz so that 1024 sequences are generated over a period of one second.

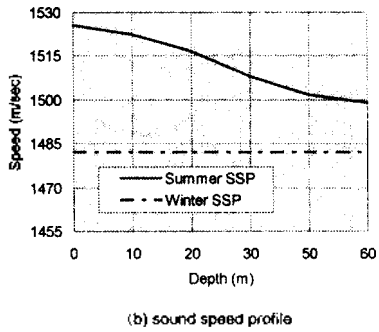
Before the simulation of signals, a modified Hamming window is applied to the signals generated by Eq. (11). The window is taken at the beginning and end of each 10 % sweeping period. This leads to reduce the energy leakage caused by the discontinuity of the finite record of data. Although the waveforms are different each other, both classes of the signals have identical power spectra [20,22].

To simulate the LFM signals distorted in the environment, the Fourier synthesis scheme [16,20] is employed. In the CW calculation with each frequency, an acoustic model based on the parabolic equation (PE) is used. After pressure fields are computed by the PE scheme, the time signals are obtained by convolving them with source spectra at four ranges: 2, 5, 10 and 20 km. To the generated time signals, white Gaussian noise is added to test the network performance with the signal to noise ratio (SNR) being 0 or -10 or -20 dB. And then the three techniques (STFT, WT, and PWVD) are applied to extract feature vectors, i.e., spectrograms.

The environment, with which time signal simulation is conducted, is a simple waveguide with pressure release surface and penetrable fluid bottom (Fig.1). The source depth is 30 m and two SSPs, being typical in winter and summer, are introduced. The sediment parameters are shown in the figure including sound speed (C_b), density (ρ_b), and attenuation (α_b). They are determined by referring to Miller and Wolf [23].



(a) geoacoustic data



(b) sound speed profile

Figure 1. Input data for time signal simulation.

B. Sea Experiment

In the experiment to obtain real sea data, one sound source and three receivers were used. The sound source projected four LFM signals centered on 200, 400, 600 and 800 Hz with each bandwidth of 100 Hz. The signals were swept up for one second and down for an another second. That is, they were repeated to produce the two classes of signals every two seconds.

Table 1 summarizes the operation conditions during the experiment. The sound source produced signals for more than 15 minutes at each depth. The source-receiver range was about 5.4 km. From now on, the drifting receiver is called as DRR and the bottom moored receiver as BMR.

Table1. Source and receiver operation conditions during the experiment.

Source Depth (m)	Signal Type	Source Duration (min)	Receiver	Receiver Depth (m)	Source-Receiver Range (km)
30	*LFM	19	*DRR 1,2 #BMR	18 60	5.4
10	*LFM	15	#BMR	60	5.4

(*) 4LFM signals centered on 200, 400, 600 and 800 Hz, each being swept up (class A) and down (class B) with bandwidth 100Hz
(+) drifting receivers connected each other by a 100 m-long rope
(#) bottom moored receiver

Two kinds of receivers, the sonobuoy AN/SSQ-57A (DRR1,2) and sonobuoy AN/SSQ-57B (BMR) were used. The latter was modified so that it could separate received signals into the north-south and east-west components. Two sonobuoys (DRR1 and 2) were connected each other by a 100 m-long rope and allowed to drift in water keeping water depth of about 18 m. However, they were again connected to the weight on the sea bottom via the rope so that they could drift in a limited area. The modified receiver (BMR) was installed on the sea bottom where the depth is around 60 m.

The bottom of the experiment area consists of sand-silt-clay. Its typical geoacoustic parameters are characterized by density 1600 kg/m^3, porosity 67.2 %, sound speed 1510 m/sec, and attenuation coefficient 0.5 dB/λ [24]. The winter SSP in Fig.1 comes from the measured data through a CTD equipment.

IV. Network Structure and Training

A. Structure

The optimum number of neurons (or units) can be found by trying experiment with the training data sets. Training starts with 9 units and continues up to 50 units in the hidden layer. After trained, the network is forced to classify the noise-free and noisy signals. The noise of SNR 0 or -10 dB is added to give noisy signals. Figure 2 shows network performance for verification by the variation of hidden unit for the class A signal with the SNR -10 dB. It can be shown that the overall performance is the best when the number of hidden units is selected to be 19. Only the STFT gives the best result with 30 units in the hidden layer. For increased number of hidden units, no practical improvement is achieved [20]. Hence, the network is built such that it has 19 units in the hidden layer.

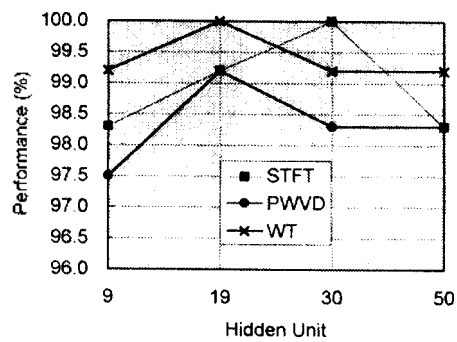


Figure 2. Network performance (%) by the variation of number of hidden units for the signals of class A and SNR-10dB.

The output layer has two units to classify if the received signal is the class A (upsweep) or B (downsweep).

The three transforms need different number of input units because they give spectrograms of different resolution. For the measured signals, the WT is not considered because it gives non-linear resolution in frequency axis where comparisons are needed among the four LFM signals of different frequency.

B. Training Data Set

The spectrograms at range 2 km are for training and others for test of the network performance. The model produces

time signals at each 0.5 m depth so that total of 120 spectrograms come out over water depth 60 m. The input parameters for the model simulation are described in Fig.1.

Figure 3 shows how a training (or test) data set is prepared from a spectrogram. At each receiver depth, one can obtain an input data set representing the characteristics of LFM signal of the class A or B. The numbers of time and frequency bin, NT and NF, vary with each feature extraction technique. The spectrum data are normalized relative to the maximum value in the spectrogram and converted to one dimensional data, x[k] (k = 1, 2,...,NT* NF). In practice it is nearly always advantageous to apply pre-processing transformations to the input data before it is presented to a network. One of the most common forms of pre-processing consists of a simple linear scaling of the input variables. This is often useful if different variables have typical values that differ significantly. By applying a linear transformation all the inputs are arranged to have similar values. Figure 4

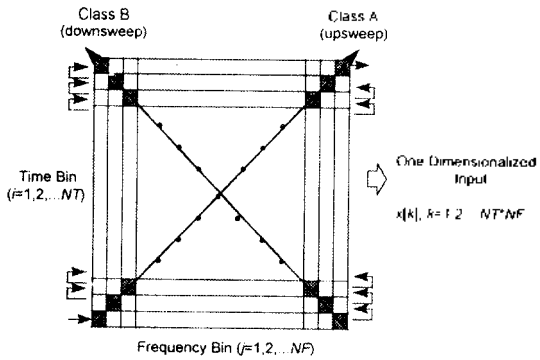
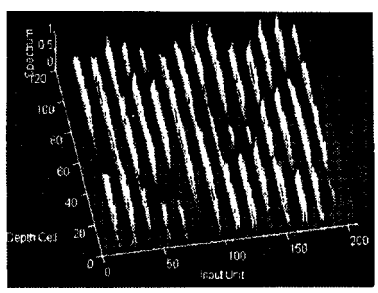
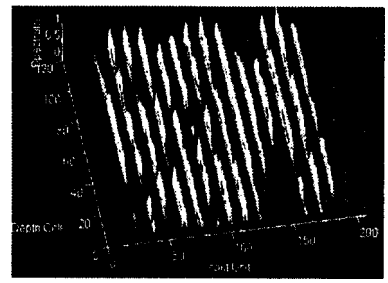


Figure 3. Preparing procedure of the training data set at each dept cell

presents examples of the training data sets. They are obtained by applying the PWVD over the simulated signals at range 2 km. As the training data represents more varieties of the target data, the network will be able to perform better on the real target data. In this sense, the two examples may be good training data sets because they give variable but almost independent spectra with time and input unit.



(a) class A



(b) class B

Figure 4. Power spectrum examples for the network training. They are obtained by applying the PWVD to the signals simulated at range 2 km.

C. Training

When the network is trained on the data sets as shown in Fig.4, it finishes learning typically within 3 steps where each step consists of 2000 iterations (Table 2). Here, the allowable error is 0.000005, and two temperatures of 0.20 and 0.10 are employed at each step. In calculating the error, a cross-entrophy error function [20] is used. The network completes its learning in four hours on a Pentium-chip based PC. This is for the STFT and it takes same order of time for other transforms.

Table 2. Learing process of the network on the simulated signals of which feature extraction technique is the STFT.

Step	Anneal Temperature	Error	Gradient Error	Anneal Error
1	0.20 0.10	0.000115 0.000024	0.000010	0.000010
2	0.20 0.10	0.000088 0.000011	0.000010	0.000010
3	0.20 0.10	0.000088 0.000004	0.000010	0.000010

V. Classification Performance

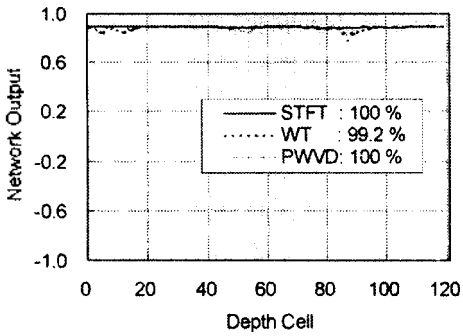
A. Simulated Signals

Variation of Feature Extraction Technique

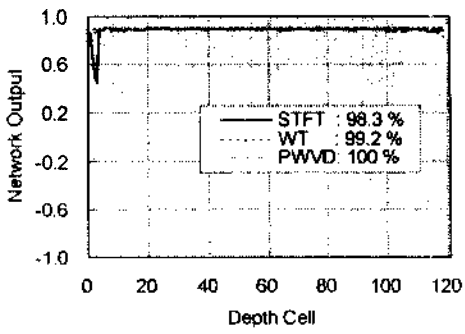
The network performance for the three techniques (STFT, WT and PWVD) is compared. The comparisons are attempted over the network performance with the winter SSP.

Figure 5 shows three examples that enable to compare the performance for each technique. The first picture (Fig.5a) comes from the class B signal at 5 km and with SNR 0 dB. When the classifying threshold is 0.8, the network performance reaches 100 % except for the WT where it is 99.2 % of 120 depth cells. The network performs almost perfectly in this case. In another case of range 10 km, SNR -

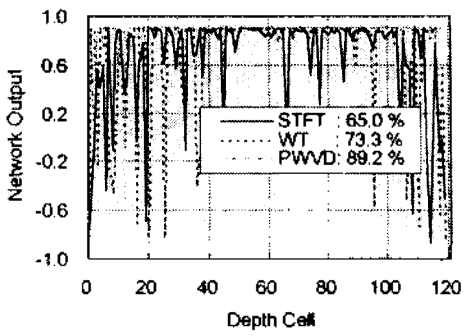
10 dB and class B (Fig.5b), the performance ranges from 98.3 (STFT) to 100 % (PWVD). At the worst case of range 20 km and SNR -20 dB, the last example (Fig.5c) shows that the network performance is 65.0 % for the STFT and 98.2 % for the PWVD. In particular, the performance is relatively poor near the surface and the bottom irrespective of the techniques. At ranges of 5 and 10 km, the simulated amplitudes are relatively small near the two interfaces [16]. Consequently, as strong noise is added to the signal, the performance is thought to be relatively poor near the two interfaces. Throughout the results considered, the PWVD gives the best performance and the STFT the worst.



(a) range 5 km, SNR 0 dB



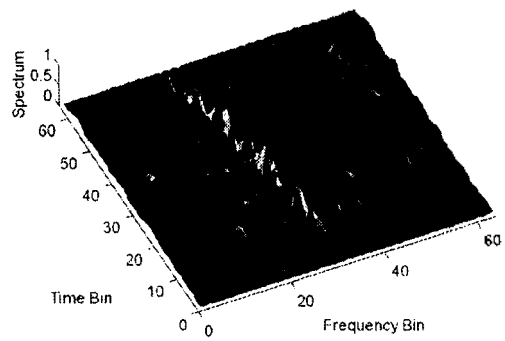
(b) range 10 km, SNR -10 dB



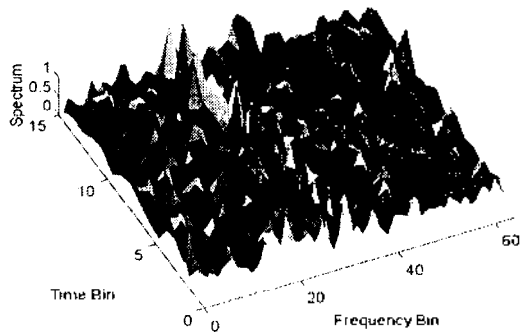
(c) range 20 km, SNR -20 dB

Figure 5. Examples of network outputs for the three techniques(STFT, WT and PWVD). They are for the class B signal.

Figure 6 presents two spectrogram examples by the PWVD and STFT. They are derived from the class B signals at range 20 km, receiver depth 30 m, and with SNR -20 dB. Most of noise is filtered out while it is not in the second one. The STFT was applied such that it may give 64 bins from 128 points representing the frequency range from 0 to 512 Hz. With a 64-point overlap, the STFT produces 15 time slices. As the number of points to transform increases, the frequency resolution decreases. Hence, a compromise should be made between the resolutions of time and frequency. On the contrary, the PWVD gives the resolution of 64 (time) x 64 (frequency). When the classifying threshold is chosen to be 0.8, the two spectrograms give right classification. The two spectrograms give network outputs of 0.90 and 0.81, respectively. In applying the PWVD, a sliding exponential window is adopted in the time-frequency domain, which reduces interference and avoids negative values. Inherently, the PWVD guarantees four times higher resolution than an ordinary power spectral distribution [25]. Instead, the computation time increases significantly in performing the autocorrelations and their Fourier transforms.



(a) PWVD, network output 0.90



(b) STFT, network output 0.81

Figure 6. Spectrogram examples of the class B signal at range 20km, SNR-20 dB and receiver depth 30m.

Figure 7 shows a spectrogram example corrupted by the strongest noise (SNR=-20 dB). It is derived by the STFT over the class B signal simulated at range 5 km and receiver depth 4 m. The signal varies within frequency bin 19-31 but it is not clear enough to identify the signal because of surrounding noise. The network gives output of -0.61, failing to classify the signal with this kind of spectrogram.

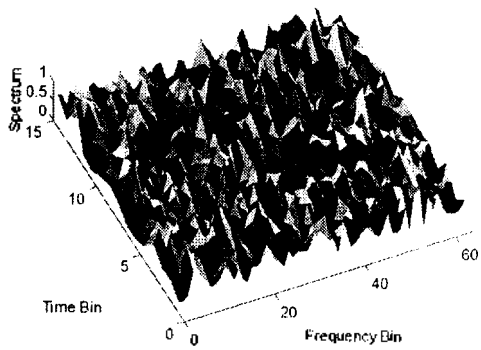


Figure 7. Spectrogram example on which the network fails to classify the signal. The network output is -0.61 where the technique STFT is employed. It is for the class B signal at range 5 km and with SNR-20 dB.

Table 3 summarizes the network performance for each technique. On an average, when the SNR is 0 or -10 dB, all the three techniques give over 95 % performance showing no difference between the techniques. When the SNR is -20 dB, however, the difference is clear: the PWVD promises the best performance giving as high as 90.3 % on the class B signal.

Table 3. Network performance(%) of each feature extraction technique where the classifying threshold is 0.8.

Range (km)	Technique	SNR		
		0 dB (α/β)	-10 dB (α/β)	-20 dB (α/β)
5	STFT	99.2/100	97.5/100	64.2/59.2
	WT	98.3/99.2	98.8/94.2	52.5/72.5
	PWVD	100/100	97.5/100	57.5/85.0
10	STFT	100/100	98.3/98.3	80.9/69.2
	WT	99.2/99.2	96.7/99.2	57.5/79.2
	PWVD	100/100	97.5/100	90.0/96.7
20	STFT	100/98.3	95.8/95.0	55.0/65.0
	WT	99.2/99.2	95.0/94.2	34.2/73.3
	PWVD	99.2/99.2	95.0/99.2	73.3/89.2
Avg.	STFT	99.7/97.8	97.2/97.8	66.7/64.5
	WT	98.9/99.2	95.8/95.8	48.1/75.0
	PWVD	99.7/99.7	96.7/99.7	73.6/90.3

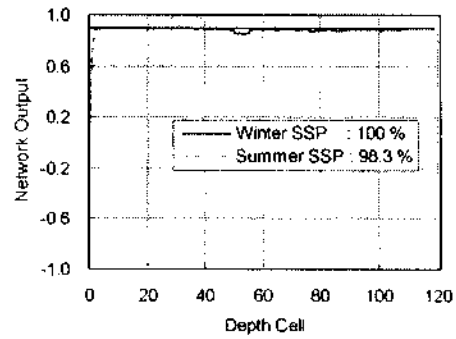
(α/β) : class A/class B

Variation of SSP

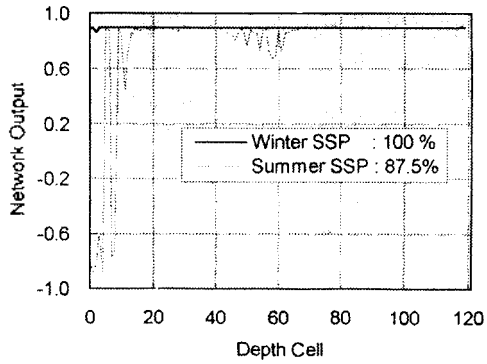
A SSP in water column decides the traveling patterns of acoustic waves. It depends on the SSP to which direction the acoustic wave refracts during the propagation. The refraction follows the Snell's law that is a function of sound speed. The two extreme profiles of the winter and summer are expected to give an effect of different degree on the signal distortion.

The summer SSP is a typical profile measured in the shallow water around Korea. As shown in Fig.1, it has great negative gradient (i.e., thermocline in terms of temperature) of nearly -0.9 sec^{-1} in depth 20-30 m. As a feature extraction technique, the PWVD is chosen since it proved to give the best performance in the case with the winter SSP.

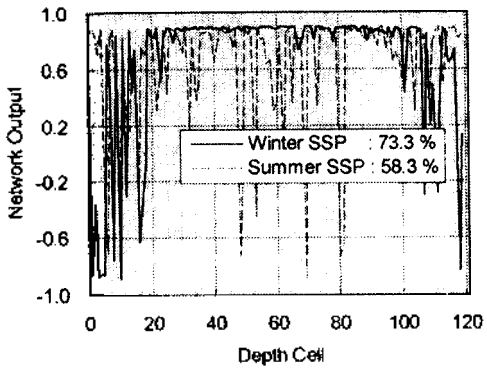
Figure 8 compares the network performance with the winter and the summer SSPs. At short range (5 km) and high SNR (0 dB), the network shows 98.3 % performance with the summer SSP while it does perfectly with the winter profile. When the range and the noise are increased up to 10 km and -10 dB, the winter SSP still gives performance of 100 % but the summer profile does only 87.5 % (Fig.8b). Most of misclassifications with the summer profile occur in the upper layer of the water column. As mentioned in time-frequency distributions of signal amplitude, this may be sufficiently expected because the summer profile is such that smaller amplitudes are formed in the upper layer. The above two examples are for the class B signal. The last picture gives the results for the class A signal. In this case, the range



(a) range 5 km, SNR 0 dB, class B



(b) range 10 km, SNR -10 dB, class B



(c) range 20 km, SNR -20 dB, class A

Figure 8. Examples of network outputs by the variation of SSP.

is 20 km and the SNR is -20 dB, comprising the worst situation for the network to identify the signal. The network performs up to 72.3 % with the winter SSP but only 58.3 % with the summer profile, giving 15 % difference. The network performance with the winter SSP shows poor performance near the surface and bottom, while that with the summer profile does in the upper layer of the water column.

Table 4 summarizes the network performance with the two SSPs. It can be shown that, in general, the network performance with the summer SSP is worse than that with the winter profile except the case of range 5 km and SNR -20 dB where the former is better than the latter. As the range or the noise increases, the network performs worse with the summer SSP. For example, when the SNR is -10 dB and the signal type is the class B, the network with the summer profile gives performance of 95.8, 87.5 and 64.2 % by the variation of range from 5 to 20 km. For the same condition, its performance with the winter profile is 100, 100 and 99.2 %, showing no considerable change. As the noise increases (as in the case of range 20 km and class B signal), the network with the summer profile performs much worse showing only 8.3 % classification at the worst situation.

Table 4. Network performance(%) by the variation of SSP. Classification is conducted on the test data via the PWVD where the classifying threshold is 0.8.

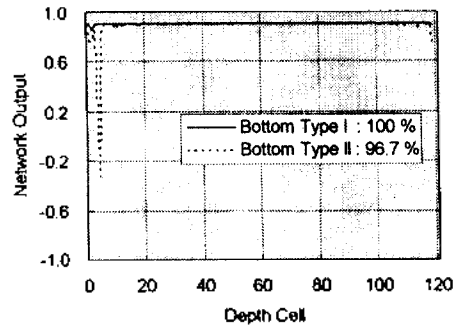
Range (km)	SSP	SNR		
		0 dB (α/β)	-10 dB (α/β)	-20 dB (α/β)
5	Winter	100/100	97.5/100	57.5/85.0
	Summer	100/98.3	96.7/95.8	92.5/78.3
10	Winter	100/100	97.5/100	90.0/96.7
	Summer	100/95.0	95.8/87.5	79.2/42.5
20	Winter	99.2/99.2	95.0/99.2	73.3/89.2
	Summer	100/98.3	83.3/64.2	58.3/8.3
Avg.	Winter	99.7/99.7	96.7/99.7	73.6/90.3
	Summer	98.6/95.0	91.9/82.5	76.7/43.0

(α/β) : class A / class B

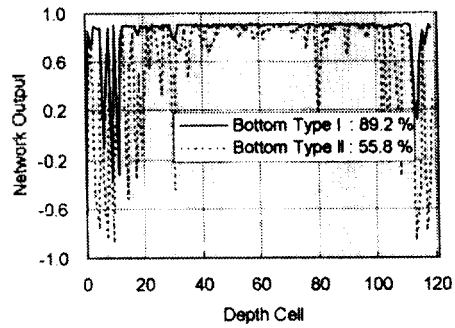
Variation of Bottom Property

To examine the effect of bottom property variations, this study considers an another type of bottom, which has geoacoustic properties of ρ_b=1500 kg/m^3, C_b=1500 m/sec and α_b=1.0 dB/λ. Compared with the type in Fig.1, the sound speed is decreased as much as 10 m/sec and the attenuation coefficient is doubled. This bottom is called as the type II bottom in comparison with the type I bottom in Fig.1. The SSP used in simulating the signal is the winter profile. The feature vectors, spectrograms, are derived via the PWVD.

Figure 9 demonstrates the bottom property effects on the network performance. For the class B signal at range 10 km and SNR -10 dB, the two results show performance of above 96 %. For the class B signal at range 20 km and SNR -20 dB, however, the network shows very poor performance (55.8 %) with the type II bottom compared to that with the type I (89.2 %). The two outputs show poor performance near the surface and the bottom, that with the type II doing much more considerably.



(a) range 10 km, SNR -10 dB, class B



(b) range 20 km, SNR -20 dB, class B

Figure 9. Examples of network outputs showing the effect of bottom property variation. Type I : C_{bI}=1510 m/s, α_b=0.5 dB/λ, and type II : C_{bI}=1500 m/s, α_b=1.0 dB/λ.

Table 5 summarizes the network performance with the two bottom properties. In all cases except for one (range 5 km,

SNR -20 dB and class A), the network performs better with the type I. With the type II bottom, however, the network performance is below 60 % for both classes of the signal at the worst condition (range 20 km and SNR -20 dB). On an average, the performance with the type II bottom is poor compared with the type I, and it is considerable for the class B signal with the SNR -20 dB.

Table 5. Network performance(%) by the variation of bottom property. Classification is conducted on the test data via the PWVD where the classifying threshold is 0.8.

Range (km)	Bottom Type	SNR 0 dB (α/β)	SNR -10 dB (α/β)	SNR -20 dB (α/β)
5	I	100/100	97.5/100	57.5/85.0
	II	99.2/100	95.8/97.5	85.0/82.5
10	I	100/100	97.5/100	90.0/96.7
	II	99.2/100	96.7/96.7	75.8/86.7
20	I	99.2/99.2	95.0/99.2	73.3/89.2
	II	97.5/97.5	95.0/89.2	59.2/55.8
Avg.	I	99.7/99.7	96.7/99.7	73.6/90.3
	II	98.6/99.2	95.8/94.5	73.3/75.0

(#) geophysical properties are shown in Fig. 1

(+) sound speed and attenuation coefficient are varied from the type I bottom

(α/β) : class A / class B

B. Measured Signals

Variation of Feature Extraction Technique

The classification is tried with the network trained on the feature vectors through the STFT and PWVD. The measured signal sets are four receiver cases (Table 1).

Figure 10 gives the four cases of the network performance with the transforms used. The first three pictures show that the two transforms guarantee almost the same performance.

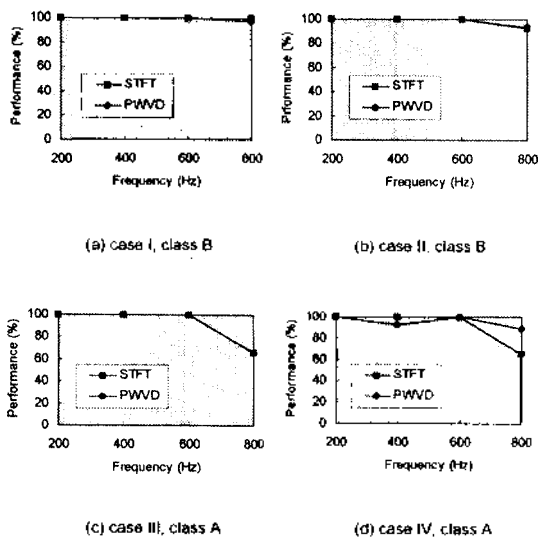


Figure 10. Four cases of the network performance with the two transforms, PWVD and STFT.

The last picture shows that the STFT outperforms at frequency 400 Hz but the PWVD at 800 Hz. In overall, on the measured signals, the two transforms give performance comparable to each other.

Table 6 gives the performance comparisons for the two feature vectors: spectra distributions via the STFT and PWVD. The network is trained on the simulated signals. One can see that the two transforms guarantee almost the same performance on an average. On the class A signal, the PWVD is superior to the STFT, and on the class B signal vice versa. Examining each performance, one can see that the network can classify nearly 90 % or more of the received signals except for the class A signals of center frequency 800 Hz in case II and III.

Table 6. Network performance(%) by the variation of the feature extraction technique. The classifying threshold is 0.8.

case	Receiver	Source Depth (m)	Center Freq. (Hz)	STFT (α/β)	PWVD (α/β)
I	BMR	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	93.8/100	100/97.9
II	DRR1	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	59.7/92.5	64.9/93.3
III	DRR2	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	60.4/91.8	65.7/90.3
IV	BMR	10	200	100/100	100/99.3
			400	92.5/100	100/100
			600	100/100	100/100
			800	89.2/100	100/97.9
Avg.				93.4/99.0	95.7/98.8

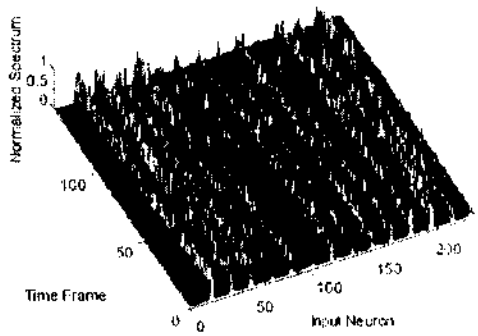
(α/β) : class A / class B

Variation of Training Data Set

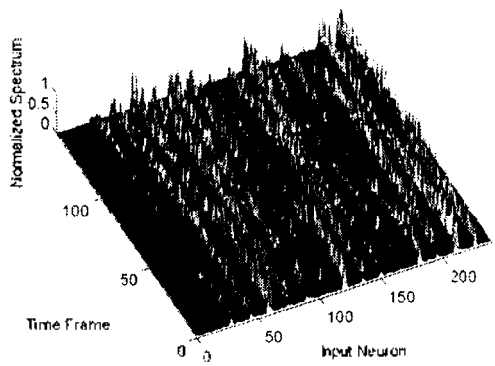
This section is directed to examining the applicability of simulated signals to training the network which is forced to perform on real measured data. That is, the network performance is compared to demonstrate the difference when the training data set is changed from the simulated signal to the measured one. It is often very expensive and difficult to obtain real sea data on which a neural network is trained. In this case, the analysis is needed to find the applicability of the simulated signals instead of the measured ones. Discussions are restricted to the results via the PWVD. As in the previous case, the training data set is obtained on one of the measured signal sets. As shown in the PWVD of the signals monitored at 1 m of the source [16,20], the signal centered on 600 Hz has the highest SNR. Hence, as the training data set, this study selects the PWVD from the

measured signal of center frequency 600 Hz in case I.

Figure 11 presents the training data sets via the PWVD on the measured signal (600 Hz in case I). Although not so dramatic as in the simulated signals, they also show some spectral variation with time frame and input unit, which is probably enough to accommodate variations or distortions of other test data sets. The two data sets themselves are also injected into the network and made to be classified by the network.



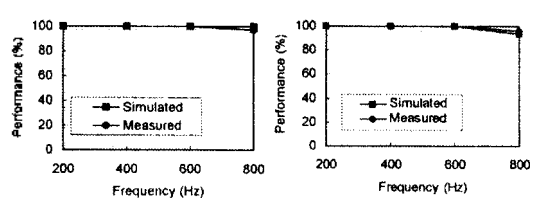
(a) case I 600 Hz PWVD, class A



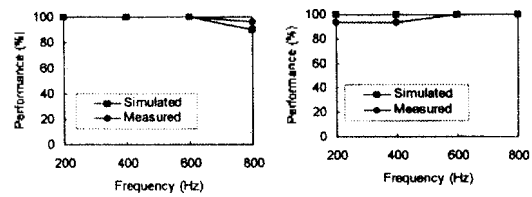
(b) case I, 600 Hz PWVD, class B

Figure 11. Power spectra of the real data on which the network is trained.

Figure 12 presents four selected cases of performance with the training data sets varied. The first three results show that little differences in performance exist between the two data



(a) case I, class A (b) case II, class B



(c) case III, class B (d) case IV, class A

Figure 12. Four cases of the network performance by the variation of training data set.

sets except for that at frequency 800 Hz. In the third picture, the measured data set gives a little better performance. In the last figure, however, the simulated one presents better results at the frequencies of 200 and 400 Hz.

Table 7 summarizes the network performance by the variation of training data set. On some cases, the network performs worse on the simulated training data (for example, 800 Hz in case II and III). However, the network shows better or comparable performance on the average for other cases, promising the applicability of the network trained on the simulated data. The network, trained on the measured signals, gives slightly better results than that on the simulated, the improvement being 1.40 % and 0.64 % on the average for the class A and B signals, respectively.

Table 7. Network performance(%) by the variation of training data set. The classifying threshold is 0.8.

case	Receiver	Source Depth (m)	Center Freq. (Hz)	Training Data Set	
				simulated (α/β)	Measured (α/β)
I	BMR	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	100/97.9	100/97.9
II	DRR1	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	64.9/93.3	85.1/95.5
III	DRR2	30	200	100/100	100/100
			400	100/100	100/100
			600	100/100	100/100
			800	65.7/90.3	84.3/96.3
IV	BMR	10	200	100/99.3	93.8/99.3
			400	100/100	93.2/100
			600	100/100	100/100
			800	100/100	100/100
Avg.				95.7/98.8	97.1/99.4

(α/β) : class A / class B

Variation of Center Frequency

It has been known that there exist the optimum frequencies at which wave transmission is best [22]. Optimum frequency is a general feature of ducted propagation in the ocean. It occurs as a result of competing propagation and attenuation

mechanisms at high and low frequencies. In the high-frequency regime there exist increasing volume and scattering loss with increasing frequency. At lower frequencies the situation is complicated. With increasing wavelength the efficiency of the duct to confine sound decreases (the cut-off phenomenon). Hence, propagation and attenuation mechanisms outside the duct (in the sea bottom) become important. In this study, the analyses are not tried any more on the propagation loss or the optimum frequency because they somewhat deviate from the topic. An analysis is attempted on the network performance in relation to only the SNRs of the four center frequencies.

Table 8 summarizes the performance by the variation of center frequency. The overall performance reaches over 90 % in all cases except for the case of frequency 800 Hz and class A where it is 87.2 %. Among the four center frequencies, the signals centered on 600 Hz are classified perfectly. When the network is trained on the measured data, the signal of frequency 600 Hz is chosen for the training. Thus, the network performance on the measured data of 600 Hz is in fact the verification results on the training data sets themselves. This perfect outputs may be also anticipated from the monitored signals [16,20] where the SNR is the highest on 600 Hz. The sound source was operated to have source levels of maximum 168 dB on 200 Hz and minimum 150 dB on 800 Hz [24], but the SNR is the highest on 600 Hz.

Table 8. Network performance(%) by the variation of center frequency. The classifying threshold is 0.8.

case	Training Data Set	Center Frequency				Avg. (α/β)
		200 Hz (α/β)	400 Hz (α/β)	600 Hz (α/β)	800 Hz (α/β)	
I	Simulated	100/100	100/100	100/100	100/97.9	100/99.5
		100/100	100/100	100/100	64.9/93.3	91.2/98.3
		100/100	100/100	100/100	65.7/90.3	91.4/97.6
		100/99.3	93.8/99.3	100/100	100/100	100/99.8
II	Measured	100/100	100/100	100/100	97.3/100	99.3/100
		100/100	100/100	100/100	85.1/95.5	96.3/98.9
		100/100	100/100	100/100	84.3/96.3	96.1/98.9
		93.8/99.3	93.2/100	100/100	100/100	96.8/99.8
Avg.		99.2/99.8	99.2/100	100/100	87.2/96.7	96.4/99.1

(α/β) : class A / class B

VI. Conclusions

This study attempts to test the classifying performance of a neural network and thereby examine its applicability to the signals distorted in a shallow water environment. Linear frequency modulated (LFM) signals are simulated by using an acoustic model and also measured through sea experiment.

After trained on the simulated signals over water depth, the network gives over 95 % performance with the SNR being up to -10 dB. Among the transforms, the PWVD presents the best performance particularly in a highly noisy condition. The network performs worse with the summer sound speed profile than with the winter profile. It is also expected to present much different performance with the variation of bottom properties.

When the network is trained on the simulated signals, it classifies over 90 % of the measured signals. On the contrary, when the network is trained on the measured signals which have the highest SNR, it gives a little better results than that trained on the simulated data.

In conclusion, the simulated signals are successfully applied to training a neural network, and the trained network performs well in classifying the signals distorted by a surrounding environment and corrupted by noise.

References

1. J. Z. Jacobson, M. Laplante and R. H. Mogford, "Artificial neural networks in naval warfare," *Operations and Analysis Establishment*, Report No. 2-6803, Ottawa, Ont., Canada, 1988.
2. C. Liou, "Design of neural networks to classify sonar targets," in International Neural Society 34, 1988.
3. R. H. Baran and J. P. Coughlin, "A neural network for target classification using passive sonar," *Proc. Intl. Soc. Optical Eng.*, vol. 1471, pp.164-176, 1991.
4. J. Briddle and D. Smith, "Line tracking using a artificial neural network and a dual process model Baysian mathematical frameworks," in Colloquim on the Role of Image Processing in Defense and Military Electronics, 1990.
5. D. W. Cottle and D. J. Hamilton, "All Neural Network Sonar Discrimination System," *IEEE* Conf. Neural Networks for Ocean Engeering, pp.13-19, 1991.
6. M. Desai and D. J. Shazeer, "Acoustic Transient Analysis Using Wavelet Decomposition," *IEEE* Conf. Neural Networks for Ocean Engeering, pp.29-40, 1991.
7. T. L. Hemminger and Y. H. Pao, "Detection and Classification of Underwater Acoustic Transients Using Neural Networks," *IEEE J.*, vol.5(5), pp.712-718, 1994.
8. A. Kundu and G. C. Chen, "Transient Sonar Signal Classification Using Hidden Markov Models and Neural Nets," *IEEE J. Ocean Eng.*, vol.19(1), pp.87-99, 1994.
9. Y. N. Na, T. B. Shim, J. W. Han and C. D. Kim, "A Study on the Classification of Underwater Acoustic Signal Using an Artificial Neural Network," *Acoust. Soc. Kor.*,

vol. 14(2E), pp. 57-64, 1996.

10. Y. N. Na, J. S. Park, J. Y. Choi and C. D. Kim, "Filtering Random Noise from Deterministic Underwater Signals via Application of an Artificial Neural Networks," *Acoust. Soc. Kor.*, vol. 15(3E), pp. 4-12, 1996.
11. R. L. Field, E. J. Yoerger, and P. K. Simpson, "Performance of Neural Networks in Classifying Environmentally Distorted Transient Signals," *Oceans' 90 Proc.*, pp.13-17, 1991.
12. B. Boashash, Advances in Spectrum Analysis and Array rocessing: Time-frequency signal analysis, Prentice Hall Inc., Eaglewood Cliffs, New Jersey, 1991.
13. L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol.77(7), pp.941-981, 1989.
14. S. H. Nawab and T. F. Quantieri, Advanced Topics in Signal Processing: Short-time Fourier transform, Prentice Hall Inc. Eaglewood Cliffs, New Jersey, 1988.
15. I. Daubechies, Advances in Spectrum Analysis and Array Processing:The wavelet transform-a method for time-frequency localization, Prentice Hall Inc., Eaglewood Cliffs, New Jersey, 1991.
16. Y. N. Na, T. B. Shim, M. S. Jurng and C. D. Kim, "Simulation and Experiment of Underwater Acoustic Signals Distorted in Shallow Water Environment, " submitted to *J. Acoust. Soc. Kor.*, 1998.
17. E. Parzen, Multiple Time Series Modeling, Tech. Rep. 12, Stanford University, Stanford California, 1968.
18. S. G. Mallet, "Multi-frequency Channel Decompositions of Images and Wavelet Models," *IEEE J. Acoustics, Speech, and Signal Processing*, vol.37(12), pp.2091-2110, 1989.
19. E. M. Johanson, F. U. Dowla, and D. M. Goodman, "Back-propagation learning for multilayer feedforward neural networks using the conjugate gradient method," *Intl. J. Neural Systems*, vol.2(4), pp.291-301, 1992.
20. Y. N. Na, Classification of Environmentally Distorted Acoustic Signals in Shallow Water Using Neural Networks, Ph. D. Thesis in Pukyong National University, Pusan, pp. 1-202, 1998.0
21. R. O. Nielson, Sonar Signal Processing, Artech House Inc., Boston, 1991.
22. F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, "Fundamentals of Ocean Acoustics," in Computational Ocean Acoustics, AIP Press, Woodbury, NY, 1994.
23. J. F. Miller and S. N. Wolf, "Modal Acoustic Transmission Loss (MOATL) : A transmission-Loss Computer Program Using a Normal-Mode Model of the Acoustic Field in the Ocean," Naval Research Lab.,Rep. No. 8429, pp.1-126, 1980.
24. Y.N.Na,T.B.Shim,S.I.Kim and C.D.Kim," Study on Propagation Characteristics of LFM Signals in Shallow Water," *Jour.Acoust.Soc.Kor.*, vol.15(1s),pp.331-338,1996.
25. J. J. Jeon and Y. S. Shin, Pseudo Wigner-Ville Distribution, Computer Program and Its Application to Time-Frequency Domain Problems, Naval Postgraduate School Rep. NPS-ME-93-002, 1993.

▲ Young-Nam Na

Senior Research Scientist, Agency for Defence Development(Vol. 15, No. 3E, 1996)

▲ Joung-Soo Park

Senior Research Scientist, Agency for Defence Development(Vol. 15, No. 3E, 1996)

▲ Dung-Hong Chang

Senior Research Scientist, Agency for Defence Development(Vol. 15, No. 3E, 1996)

▲ Chun-Duch Kim

Professor, Department of Electrical Engineering Pukyong National University(Vol. 15, No. 3E, 1996)