Neural Network Based Classification of Time-Varying Signals Distorted by Shallow Water Environment

(천해환경에 의해 변형된 시변신호의 신경망을 통한 식별)

Young-Nam Na, Taebo Shim, Duck-Hong Chang

(Agency for Defense Development)

Chun-Duck Kim

(Pukyong National University)

Abstract

In this study, we tried to test the classification performance of a neural network and thereby to examine its applicability to the signals distorted by a shallow water environment. We conducted an acoustic experiment in a shallow sea near Pohang, Korea in which water depth is about 60 m. The signals, on which the network has been tested, is linear frequency modulated ones centered on one of the frequencies, 200, 400, 600 and 800 Hz, each being swept up or down with bandwidth 100 Hz. We considered two transforms, STFT (short-time Fourier transform) and PWVD (pseudo Wigner-Ville distribution), from which power spectra were derived. The training signals were simulated using an acoustic model based on the Fourier synthesis scheme. When the network has been trained on the measured signals of center frequency 600 Hz, it gave a little better results than that trained on the simulated. With the center frequencies varied, the overall performance reached over 90 % except one case of center frequency 800 Hz. With the feature extraction techniques (STFT and PWVD) varied, the network showed performance comparable to each other. In conclusion, the signals which have been simulated with water depth were successfully applied to training a neural network, and the trained network performed well in classifying the signals distorted by a surrounding environment and corrupted by noise.

I. Introduction

Based on the presence or absence of features, the sonar operator attempts to determine the identity of the target. This classification task has important military consequences because passive sonar enables covert detection of unfriendly vessels. However, the task is difficult to perform requiring lengthy and often intensive training. As sensor technology develops and target becomes increasingly sophisticated, this task is becoming more and more difficult due to increasing volume and complexity of the data available for processing. These have led to the urgent need for increased computer assistance.

As the demand increases to address even more complex problems, such as pattern recognition, the limitations of conventional approaches are becoming increasingly pronounced. It is in this area that neural networks promise a significant breakthrough [1].

This paper is directed to test the performance of a neural network based classifier on the distorted signals in a shallow water. The network is trained on the signals simulated at 2 km using a time-domain acoustic model. The experiment has been conducted with a sound source emitting four LFM signals, and three receivers transmitting signals to land site via radio frequency.

The network performance is delivered in relation to four LFM signals, two feature extraction techniques, and training data sets.

II. Spectrum Estimation for Non-Stationary Signals

2.1. Short-Time Fourier Transform (STFT)

As feature vectors for the network, we employ spectrograms, i.e., spectrum distribution with time and frequency. The spectrograms have been especially important for speech processing and for signal processing [2-5]. To obtain spectrograms, we consider two transforms: short-time Fourier transform (STFT) and pseudo Wigner-Ville distribution (PWVD).

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 as [6]

$$X[n,\omega] = \sum_{k=-\infty}^{\infty} x[k]w[n-k]e^{-j\omega k}, \quad (1)$$

where $w[k]$ extends from $n-L$ to $n+L$. The power spectral estimate is given by the short time periodogram

$$\hat{P}_x[n,\omega] = \frac{1}{2L+1}\left|x[n,\omega]^2\right|. \quad (2)$$

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

2.2. Pseudo Wigner-Ville Distribution (PWVD)

This is a kind of time-frequency distributions and is known to be suitable for analyzing transient or other non-stationary phenomena. It has been widely used in optics [7] and speech processing [8].

The Wigner-Ville distribution (WVD) is defined as [9]

$$W(t,\omega) = \int_{-\infty}^{\infty} s^*(t-\frac{\tau}{2})s(t+\frac{\tau}{2})e^{-j\omega\tau}, \quad (3)$$

and discrete form as [10]

$$W(t,\omega) = 2\sum_{\tau=-\infty}^{\infty} s^*(t-\frac{\tau}{2})s(t+\frac{\tau}{2})e^{-j2\omega\tau}. \quad (4)$$

For a sampled signal $s[n]$ $(n=0,1,2,\ldots,N-1)$, Eq. (4) changes into

$$W[l,k] = \frac{1}{N}\sum_{n=0}^{N-1} s[l+n]s^*[l-n]e^{-j4\pi nk/N},$$

$$k=0,1,2,\ldots N-1 \quad (5)$$

where $s[m]=0$ for $m<0$ and $m>N-1$.

Basically, Eq.(5) has the form of the FFT and we can utilize a FFT algorithm. However, it has a $N/2$ period so that even when the sampled data satisfies the Nyquist criterion, there would be still aliasing component in the WVD [10]. A simple way to avoid the aliasing is to introduce the analytic signal beforehand [11].

Since the WVD has a $N/2$ period, we can rewrite Eq.(5) as follows

$$W(m\Delta t, k\Delta\omega) = 2\Delta t \sum_{n=0}^{2N-1} s[(m+n)\Delta t]s^*[(m-n)\Delta t]$$

$$e^{-j2\pi nk/2N}, \qquad (6)$$

where $\Delta\omega = \pi/(2N\Delta t)$ and Δt is the sampling interval. In Eq.(6), the frequency resolution $\Delta\omega$ is 1/4 of the ordinary FFT, implying that the WDF guarantees four times of frequency resolution.

To suppress the interference arising from cross terms, we apply a sliding window in time-frequency domain. The WDF, with the window included, is usually called the PWVD or smoothed WVD. We obtain the PWVD by convolving the WVD and Gaussian window function.

III. Sea Experiment

Figure 1 shows the locations of the sound source and receivers. We used one sound source and three receivers. The sound source projected four LFM signals centered on 200, 400, 500 and 800 Hz with bandwidth 100 Hz. The signals were swept up for one second and down for an another

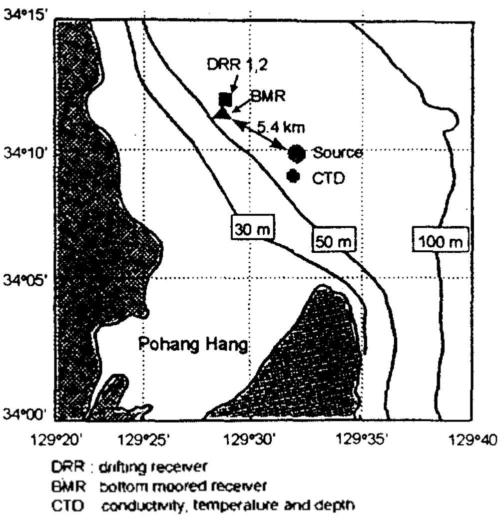


Fig.1. Station map of the acoustic experiment.

second. That is, they were repeated to produce the class A (swept up) and B (swept down) signals every two seconds. To check the quality of the projected signals, we installed a hydrophone at 1 m away from the source and monitored the signals. The source was operated on the water depths of 10 and 30 m.

We also deployed a CTD (conductivity, temperature and depth) equipment to get oceanographic data for resolving water conditions. A few days before the experiment, there was north-westerly strong enough to mix the whole water column.

We used two kinds of receivers, the sonobuoy AN/SSQ-57A (DRR1,2) and sonobuoy AN/SSQ-57B (BMR). The latter was modified so that it could separate received signals into the north-south and east-west components. Two sonobuoys (DRR1,2) 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 just in a limited area. The modified receiver (BMR) was installed on the sea bottom in which 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/s, and attenuation coefficient 0.5 dB/λ [12].

The profile shows typical pattern of very well mixed water, remaining almost same velocity from the surface to the bottom. This pattern was caused by the strong north-westerly a few days ago.

Figure 2 presents an example of the PWVD of the signals monitored at 1 m away from the source. In this case, the signal was swept down (class B). It was to check the quality of the projected signal with time. In the figure, the frequency and time axes span 256 bins, representing 1024 Hz and 1 second, respectively. Each LFM signal has bandwidth 100 Hz. Among

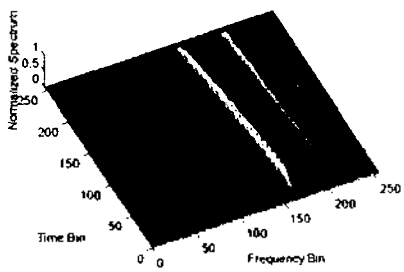


Fig.2. Spectrogram example via the PWVD on the class B signal monitored at 1 m away from the source.

the four LFM signals, the one of center frequency 600 Hz, is the most intensive (i.e., SNR is the highest). Even in a 1-second period, we can see that there exists intensity variation with time, particularly in the fourth LFM signal, center frequency being 800 Hz.

An example of the received signal at 5.4 km away from the source (Fig.3) shows that there are

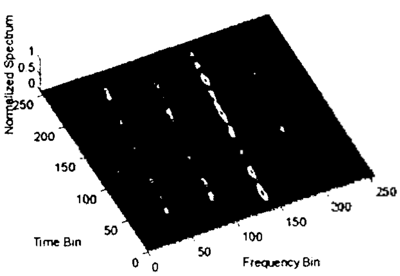


Fig.3. Spectrogram example via the PWVD on the measured signal (class A).

obvious four LFM signals. In this example, the signal was swept up (class A). As shown in the monitored signal, the third LFM signal has the strongest intensity. The frequency and time axes correspond to 1024 Hz and 1 second, respectively. It is via the PWVD on the signals at depth 60 m.

IV. Neural Network Training

In performing the STFT, we use 256 points of data with Hamming window and 64 points overlap. This transformation gives 128 frequency bins and 15 time frames. Among 128 frequency bins, 15 frequency bins are shared for each LFM signal. That is, each LFM signal (class A or B) spans for 15 frequency bins. Hence, the network needs 225 input neurons (15 bins x 15 frames) for each signal. Meanwhile, the PWVD presents 128x128 spectrum data in time-frequency domain where each LFM signal spans for 15 frequency bins. We select time frame in every 8-step interval so that the network needs 240 input neurons for each LFM signal.

We choose the network of three layers; input, hidden and output. The number of neuron in the hidden layer is set to 19.

To prepare the training data set for the network, we simulate time signals at range 2 km. We take the sound speeds in water to be constant value of 1482 m/s and other input parameters to be typical values of sand-silt-clay bottom. The LFM signal is swept up or down with bandwidth 100 Hz and center frequency 800 Hz. Time signals are simulated such that they give 2048 points per one second.

The way how to train the network is shown in Fig.4. At each receiver depth, we obtain one input data set representing the characteristics of the class A or B signal. The STFT and PWVD require 225 and 240 input neurons, respectively. The spectrum data are normalized relative to the maximum value and converted to one dimensional data, x[k] (k=1,2,...,NT*NF), where NT and NF are the numbers of time frame and frequency bin, respectively.

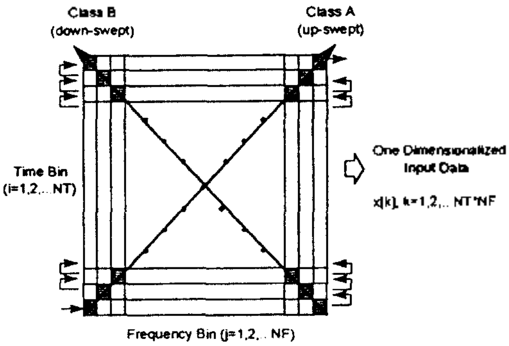
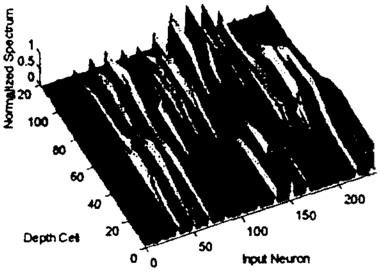
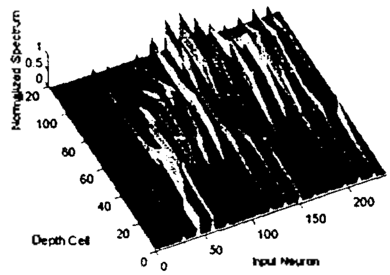


Fig.4. Preparing procedure of the training data.

Figure 5 shows the training data sets. They are obtained by applying the PWVD over the simulated signals at range 2 km. The water depth is 60 m and the model gives time signals at each 0.5 m so that 120 spectrograms come out. The two training data sets give obviously different patterns with depth cell. As the training data represent more varieties of the target data, the network would be able to perform better on the test data. In this sense, the two spectra examples may be good training data sets because they have



(a) class A (swept up)



(b) class B (swept down)

Fig. 5. Power spectra examples for the network training.

variable but almost independent spectra with time and input neuron. In training the network, we introduce the minimum-seeking algorithm, annealing plus conjugate gradients [13]. It combines the global search strategy of simulated annealing with the powerful conjugate gradient algorithm. Typically, the network converges to the minimum within three steps where the allowable error is 0.0005.

V. Network Performance on Measured Signals

5.1. Variation of Feature Extraction Technique

Table 1 gives the performance comparisons from the two feature vectors: spectra distributions by the STFT and PWVD. The network has been trained on the simulated signals. The table shows that the two transforms guarantee almost 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, we 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 1. Network performance (%) with the feature extraction techniques varied.

Case	Receiver (SD, m)	Center Freq. (Hz)	STFT	PWVD
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	98.5/100	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	100/100
Avg.			93.4/99.0	95.7/98.8

(*) class A / class B

As a whole, the network performance by the two transforms (STFT and PWVD) present similar trends and are very comparable to each other on an average.

5.2. Variation of Training Data Set

We examine the network performance when the training data sets are changed from the simulated signals to the measured. We restrict our discussion to the results via the PWVD. As can be seen in the PWVD of the signals monitored at 1 m of the source (Fig.2), the signal centered on 600 Hz has the highest SNR. Hence, we select the PWVD from the measured signal of center frequency 600 Hz in case I as the training data set.

Table 2 summarizes the network performance with training data sets varied from the simulated data to the measured. In 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 % for the class A and B signals, respectively.

Table 2. Network Performance with the training data sets varied.

Case	Receiver (SD, m)	Center Freq. (Hz)	Training Data Set	
			Simulated	Measured
I	BMR (30)	200	*100/100	100/100
		400	100/100	100/100
		600	100/100	100/100
		800	100/97.9	97.3/100
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

(*) class A / class B

5.3. Variation of Center Frequency

We present the network performance when center frequencies are varied from 200 to 800 Hz. Table 3 summarizes the performance with center frequencies varied on the simulated and measured signals. The overall performance reaches over 90 % in all cases except the case of center 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 we have trained the network on the measured data, we chose the signals of center frequency 600 Hz. Thus, the network performance on the measured data of 600 Hz is actually verified results on the training sets. This perfect outputs may be also anticipated from the monitored signals in Fig.2 where the SNR is the highest on 600 Hz and the lowest on 800 Hz. The sound source was operated to have source levels of maximum 168 dB on 200 Hz and minimum 150 dB on 800 Hz [14], but the SNR is the highest on 600 Hz.

Table 3. Network performance (%) with the center frequencies varied.

Case	Center Freq. (Hz)			
	200	400	600	800
I (S)	*100/100	100/100	100/100	100/97.9
II (S)	100/100	100/100	100/100	64.9/93.3
III (S)	100/100	100/100	100/100	65.7/90.3
IV (S)	100/99.3	100/100	100/100	100/100
I (M)	100/100	100/100	100/100	97.3/100
II (M)	100/100	100/100	100/100	85.1/95.5
III (M)	100/100	100/100	100/100	84.3/96.3
IV (M)	93.8/99.3	93.2/199	100/100	100/100
Avg.	99.2/99.8	99.2/100	100/100	87.2/96.7

(*) class A / class B

S : simulated, M : measured

Anyway, even though the network is trained on the signal on a particular frequency (800 Hz on the simulated data and 600 Hz on the measured), it performs well over signals on other frequencies.

5.4. Variation of Receiver Depth

We conducted the acoustic experiment at two receiver depths where the receivers were three. One of them was installed on the sea bottom (BMR i.e., case I) and the other two (DRR1,2, i.e., case II, III) were allowed to drift around the installed location keeping their depths to be 18 m.

In Table 3, we can see that the network performs better in case I than in case II and III (particularly on the class A signals of 800 Hz) irrespective of the training data types (simulated or measured). However, the average performance is over 90 % in all receivers.

5.5. Variation of Source Depth

We changed source depths during the experiment between 10 and 30 m, and in each depth we operated the sound source for more than 15 minutes.

As can be examined in Table 3, the network performs over 90 % in all cases except for the two cases (800 Hz in case II and III). When we compare the case I and IV, where the signals were received through same receiver but source depth was changed, we can see that the network shows no significant change in performance on the average. That is, when the network is trained on the simulated signals, the performance is 100 and 99.5 % for source depth 30 m (case I), but it is 100 and 99.8 % for 10 m (case IV). In the network trained on the measured signals (case I vs. IV), we can obtain a little better performance for source depth 30 m which reaches up to 99.3/100 % (class A/class B) compared with 96.8/99.8 % for source depth 10 m.

VI. Conclusion

We tried to test the classification performance of a neural network and thereby to examine its applicability to the measured signals in a shallow water environment. The training signals were simulated using a time-domain acoustic model.

Once the network has been trained, it classified over 90 % of the measured signals on the average. In conclusion, the signals, which have been simulated through an acoustic model, were successfully applied to training a network and the trained network performed good enough in classifying the measured signals.

References

[1] C. P. Sheppard, C. R. Gent, "A Neural Based Sonar Classification System," *Underwater Systems Design*, Nov./Dec., pp.28-31, 1991.

[2] B. Boashash, Advances in Spectrum Analysis and Array rocessing: Time-frequency signal analysis, Prentice Hall Inc., Eaglewood Cliffs, New Jersey, 1991.

[3] L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol.77(7), pp.941-981, 1989.

[4] S. H. Nawab and T. F. Quantieri, Advanced Topics in Signal Processing: Short-time Fourier transform, Prentice Hall Inc. Eaglewood Cliffs, New Jersey, 1988.

[5] 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.

[6] E. Parzen, Multiple Time Series Modeling, Tech. Rep. 12, Stanford University, Stanford California, 1968.

[7] M. J. Bastians, "The Wigner Distribution Function Applied to Optic Signals and Systems," *Optics Comm.*, vol.25(1), pp.26-30, 1978.

[8] E. Velez and R. Absher, "Transient Analysis of Speech Signals Using the Wigner Time-Frequency Representation," *IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, vol.4, pp.2242-2245, 1989.

[9] E. Wigner, "On the Quantum Correction for Thermodynamic Equilibrium," *Physic Review*, vol.40, pp.740-759, 1932.

[10] T. Classen and W. Mecklenbrauker, "The Wigner Distribution - A Tool for time-Frequency Signal Analysis," *Philips J. Res.*, vol.35, part I:217-250, part II:276-300, part III:372-389, 1980.

[11] J. Ville, "Theorie et Applications de la Notion de Signal Analytique," *Cables et Transmission*, vol.29(1), pp.61-74, 1948.

[12] 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.

[13] 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.

[14] 나영남, 심태보, 김성인, 김천덕, "천해에서 선형 주파수 변조신호의 전파특성에 관한 연구," 한국음향학회 학술발표회 논문집, vol.15(1s), pp.331-338, 1996.