Signal Classification Based on Frequency Analysis Using Multilayer Neural Network with Limited Data and Computation

Kazuyuki HARA
Graduate School of Nat. Sci. and Tech.,

Kenji NAKAYAMA
Faculty of Engineering, Kanazawa Univ.

2-40-20 Kodatsuno, Kanazawa, 920 JAPAN
e-mail : nakayama@ec.t.kanazawa-u.ac.jp

ABSTRACT

Signal classification performance using multilayer neural network (MLNN) and the conventional signal processing methods are theoretically compared under the limited observation period and computational load. The signals with N samples are classified based on frequency components. The comparison is carried out based on degree of freedom the signal detection regions in an N-dimensional signal space. As a result, the MLNN has higher degree of freedom, and can provide more flexible performance for classifying the signals than the conventional methods. This analysis is further investigated throught computer simulations. Multi-frequency signals and the real application, a dial tone receiver, are taken into account. As a result, the MLNN can provide much higher accuracy than the conventional signal processing methods.

1. Introduction

Recently, neural networks NNs) [1] have been applied to signal processing field. Signal detection [2]-[6], a demodulator for shift keying signals [7]-[9] and multi-frequency signal classification [10],[11] are included. In these applications, the NN methods can provide better performance than conventional signal processing methods. However theoretical comparisons between them have not been well discussed.

The purpose of this paper is to compare performance of signal detection or classification by a multilayer (ML)NN and conventional signal processing methods. The comparison is discussed based on their pattern classification mechanism. Classification of the signals carrying information as frequencies are taken into account. Because, frequency analysis is one of the important processings in communication and signal processing fields. Multi-frequency signal classification, such as a dial tone detector, is a concrete example of the above classification.

The following environment is taken into account in this paper. First, signal classes are defined by frequency combination. Thus, multi-frequency signals include several frequency components, which are located alternately among the classes. Amplitude and phase for each frequency component may have arbitrary values. Moreover, we impose the following conditions : An observation period and computational load are strictly limited. We encounter such limitations in some practical applications. Furthermore, from the viewpoint of frequency analysis, the above limitations are severe conditions. Hence, accuracy of signal detection and classification by the above methods can be compared.

2. Signal Detection and Classification by Conventional Methods

2.1. Frequency Analysis Methods

Frequency analysis methods, including filters and Fourier transform, extract frequency components by calculating the inner products of the input signal $x(n)$ and filter coefficients $h(n)$ or Fourier kernel $e^{-j\omega nT}$. The frequency components can be found by detecting large outputs in the power among the signal groups. In the filter method, the output is given by

$$y_p(n) = \sum_{m=1}^{N} x(m)h_p(n-m), \quad p = 1 \sim P \qquad (1)$$

where $h_p(n)$ is the filter coefficients for the pth signal group.

In the Fourier transform method, the frequency components are also extracted by

$$F(j\omega_p) = \sum_{n=1}^{N} x(n)e^{-j\omega_p nT} \qquad (2)$$

where f_p is the frequency included in the pth signal group. If power of $y_p(n)$ or amplitude of $F(j\omega_p)$ take the maximum value, then the input signal is classified into the pth group.

The frequency components can be detected by suppressing other group frequencies, too. The processing is the same as Eq.(1). However, the transfer function has zeros on the frequencies included in the pth signal group.

$$H_{FIR2}(z) = h_0 \prod_{p=1}^{P}(1 - 2\cos\omega_p T z^{-1} + z^{-2}) \qquad (3)$$

In this method, if power of $y_p(n)$ takes the minimum value, then the input signal is classified into the pth group.

There are two kinds of filter structures, that is a finite impulse response (FIR) filter and an infinite impulse response (IIR) filter. Among them, we employ an FIR filter, because it can simulate an IIR filter. The frequency extracting and suppressing filter methods are denoted FIR1 and FIR2, respectively in this paper. Comparison between them will be described in Sec.5.

0-7803-2768-3/95/$4.00

2.2. Pattern Matching Methods

Euclidean distance and Maharanobis' generalized distance (MGD) are employed for the pattern matching method. These methods calculate the distance between the templates and the input signals. The Euclidean distance from the input signal $x(n)$ to the mth template $x_{pm}(n)$ in the pth group is given by

$$d(pm) = \{\frac{1}{N}\sum_{n=1}^{N}(x_{pm}(n) - x(n))^2\}^{\frac{1}{2}} \quad (4)$$

If

$$d(p'm) = \min_{p}\{d(pm)\} \quad (5)$$

then $x(n)$ is classified into the p'th group.

For Maharanobis' generalized distance, the distance from the input signal $x(n)$ to the centroid of pth group templates is defined by

$$d_p^2 = (x - \mu_p)^T C_p^{-1}(x - \mu_p) \quad (6)$$

Here, C_p^{-1} is the inverse matrix of the covariance matrix of the pth templates, μ_p is the centroid vector of the pth signal group. Thus, the variance of the signal group distribution is taken into account and this process reform the signal group distribution into a sphere. The same decision rule is used as in the Euclidean distance methods.

Each template forms a sub-group and the performance is highly depends on how well the templates cover the region in which all signals locate. So, selection of the templates is a critical problem.

The k-mean clustering is the most popular clustering algorithm by MacQueen[13] and Anderberg[14]. This clustering method selects the best template signal to classify the signals based on the Euclidean distance. This method estimates the distribution of the signal in an N-dimen-sional space as a sphere. That means the signals are not correlated each other. So, this estimation has some limitation in the signal classification of interest.

The maximum-likelihood Gaussian classifier [15] is another method for classification. This classifier uses the maximum-likelihood estimation to decide the fixed coefficients which maximize the accuracy. The inner product of the input signals and the coefficients are calculated. The inner product will be a large value for one group and small value for the others. This method assumes Gaussian distribution. However, distribution of multi-frequency signals does not always follow this assumption.

2.3. Spectrum Estimation Method

Maximum entropy method (MEM) [12], which is a spectrum estimation method, estimates an AR model of the given data. From the Wienner-Khinchin's law, the spectrum is equal to the Fourier transform of the autocorrelation of the signal of interest. The power spectrum $P(\omega)$ is given by

$$P(\omega) = \sum_{i=-M}^{M} \gamma_i e^{-j\omega k} \quad (7)$$

Here, γ_i is the autocorrelation sequence of the signal $x(n)$ with i lag. The power spectrum $P(\omega)$ is modeled by

$$P(\omega) \approx \frac{a_0}{\left|1 + \sum_{k=1}^{M} a_k e^{-j\omega k}\right|^2} \quad (8)$$

a_0 and $\{a_k\}$ are unknown coefficients of the prediction-error filter. M is the order of the filter. From Eqs.(7) and (8), Mth-order filter coefficients a_0 and $\{a_k\}$ are estimated so as to maximize the entropy. The frequency component is detected by finding the frequencies at which the spectrum $P(\omega)$ has peak values. From the detected frequencies, the input signal can be classified.

3. Signal Detection and Classification by Multilayer Neural Networks

We use a single hidden layer whose activation function is monotonically increased and squashed function (sigmoid function). If we can assume that the hidden unit outputs can approach to the saturation regions as the learning converges, then region boundaries are formed by hyper-planes, whose equations are formulated by the input-hidden layer connection weights.

Let the input signal is an N-dimensional vector. One hyper-plane divides the N-dimensional input space into two regions. When M hidden units are used, there are M hyper-planes, and the input space can be separated into 2^M sub-regions at maximum. In other words, the hidden layer outputs are encoded with M bits. Hence, we have 2^M bit patterns.

These sub-regions are further combined at the output layer through the hidden-output layer connection weights. In this step, set of the hidden layer outputs, which can be used simultaneously, should include linearly separable patterns, because only a single layer remains. Although this limitation is imposed, a variety of the signal detection regions can be formed at the output layer. Thus, the MLNN method has much higher degree of freedom of forming the signal detection regions.

4. Comparison of Signal Detection and Classification

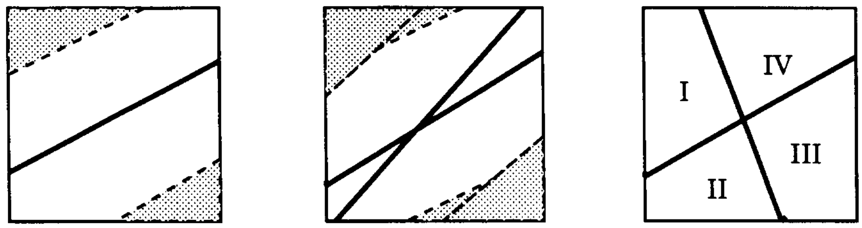
4.1. Conventional Signal Processing Methods

When a large number of samples are used to represent the input signals, highly accurate signal classification is possible by the conventional methods. Letting the signal detection accuracy be 100%, all signal vectors locate in the detection regions. For example, in the case of FIR1 in Sec.III, if the signals can be detected by using only one output sample, the following inequality may be held.

$$\left|\sum_{m=1}^{N} x(m)h_{p'}(n-m)\right| \gg \max_{p \neq p'}\left|\sum_{m=1}^{N} x(m)h_p(n-m)\right| \quad (9)$$

Supposing an appropriate threshold α, this condition can be replaced by

$$\left|\sum_{m=1}^{N} x(m)h_{p'}(n-m)\right| > \alpha \quad (10)$$



(a) FIR filter with one output sample (b) FIR filter with two output samples (c) MLNN with two hidden units

Fig. 1: Signal detection regions of FIR filter and MLNN.

This means that all the signals belong to the p'th group are concentrated in the above region in the N-dimensional space.

An example of the signal detection region given by Eq.(10) for two-dimensional signals is shown in Fig.1(a). The shade parts are the signal detection regions.

When the number of the input signal samples is limited, many outputs samples are needed in the power computation. In this case, the signal detection region is formed by

$$\sum_{k=1}^{K} \left| \sum_{m=1}^{N} x(m) h_{pk}(n-m) \right|^2 > \alpha' \tag{11}$$

This region is wider than that given by Eq.(10). Because all the inner products included in Eq.(11) are not required to satisfy Eq.(10). It is enough to satisfy Eq.(11) that some of them have large magnitude. An example of the extended region given by Eq.(11) is illustrated as shade parts in Fig.1(b).

In case of the filter methods, however, sub-regions, formed by hyper-planes, cannot be freely combined like the MLNN as described in Sec.III. The regions are uniquely determined by the hyper-planes. This is disadvantage of the conventional signal processing methods.

From this discussion, we can estimate the distribution of the N-dimensional multi-frequency signals. When the signal have a large number of samples, they will distribute in some limited region formed by a single hyper-plane. However, by limiting the number of the input signal samples, they spread out of the limited region. Because, the signals are not exactly orthogonal in this case. The filter outputs become smaller. Thus, a wide region is required for detecting the signals. The signal distributed regions may be very complicated. This means many hyper-planes will be required in the filter method. If computational load is limited to be the same as the MLNN method, performance of the filter methods will be inferior to that of the MLNN method. Discussions based on simulation will be given in Sec.V.

The same discussion as in the filter method can be valid in the pattern matching methods. A template corresponds to a set of filter coefficients. In this case, however, performance is highly dependent on selection of a set of templates.

4.2. Multilayer Neural Network Method

Signal detection regions formed by the MLNN method is investigated. When two hidden units are used, the input space can divided into four regions as shown in Fig.1(c). One hidden layer output corresponds to one of four reions. These regions are further combined at the output layer. In the MLNN with a single hidden layer, the hidden layer outputs must be linearly separable in this case. After taking this condition into account, the MLNN still has much higher degree of freedom for forming the regions. In Fig.1(c), when the number of signal groups is two, any combinations of the regions I $\sim$ IV , except for a combination of (I,III) and (II,IV) which is linearly nonseparable, are possible.

As discussed above, the input signals with a small number of samples distribute widely and complicatedly. In this case, the MLNN can be expected to provide high performance in the signal detection and classification.

On the other hand, the MLNN is trained by supervised learning. Then how well form the region is strongly dependent on a set of the training signals. Therefore, the training signals should be carefully selected, and the number of them should be enough.

5. Simulation Results and Comparisons

5.1. Multilayer Neural Network

A two-layer neural network is taken into account. N samples of the mth signal $x_{pm}(n)$, belongs to the pth group, are applied to the input layer. The nth input unit receives x_{pm}(n). The connection weight from the nth input to the jth hidden unit is denoted w_{nj}. The input potential and output of the jth hidden unit are given by

$$net_j = \sum_{n=1}^{N} w_{nj} x_{pm}(n) + \theta_j \tag{12}$$

$$y_j = f_H(net_j) \tag{13}$$

Letting the connection weight from the jth hidden unit to the kth output unit be w_{jk}, the input potential and output of the kth output unit are given by

$$net_k = \sum_{j=1}^{J} w_{jk} y_j + \theta_k \tag{14}$$

$$y_k = f_O(net_k) \tag{15}$$

The activation function of the hidden and output layers are the sigmoid function.

The number of output units is equal to that of the signal groups, that is $k = 1 \sim P$ in Eqs.(14) and (15). The neural network is trained so that a single output unit responds to one of the signal groups.

5.2. Multi-frequency Signal

The following multi-frequency signals are used in the simulation.

$$x_{pm}(n) = \sum_{r=1}^{R} A_{mr} \sin(\omega_{pr} nT + \phi_{mr}) \quad (16)$$

Here, $\omega_{pr} = 2\pi f_{pr}, m = 1 \sim M, n = 1 \sim N$. T is a sampling period. In the pth group, the signals have the same frequencies.

$$F_p = [f_{p1}, f_{p2}, \ldots, f_{pR}] \quad (17)$$

Amplitude A_{mr} and phase ϕ_{mr} are different for each frequency in the sample group. They are generated as random numbers, uniformly distributed in following ranges.

$$0 < A_{mr} \leq 1, \quad 0 \leq \phi_{mr} < 2\pi \quad (18)$$

5.3. Training and Classification of MLNN

Error back propagation(BP) algorithm [1] is used for training. The signals are divided into the training signals and the test signals. Noise free and noisy signals are used for both training and testing. Additive noise, uniformly distributed in [-0.5,0.5], is employed. After training, test signals are applied to the MLNN to investigate its generalization.

The number of frequencies included in each group is R=3, and that of the signal groups is P=2. The number of samples is N=10 or N=20. Group frequencies in group 1(#1) are 1, 2 and 3 Hz, and in group 2(#2) are 1.5, 2.5 and 3.5 Hz, respectively. A sampling frequency is 10Hz. 200 training signals and 1800 test signals in each group are used. Here, number of the training signals are determined by experience with which the generalization performance can be guaranteed.

5.4. Signal Classification

The number of parameters are listed in Table 1. One inner product or one sample power calculation is counted as one computation. After the learning converges, the hidden unit outputs approach to '1' or '0'. So, the sigmoid function can be replaced by a threshold function in the test signal classification. Therefore, the calculation of the sigmoid function is omitted from the computation. The parameters for each method is as follows; MLNN : the number of hidden units, Euclidean distance and MGD : the number of templates, Fourier transform and MEM : the number of frequency components, FIR1 and FIR2 : the number of output samples.

Figure 2 shows an example of an impulse response of 1000 lengths used in FIR1. The band width is 0.02Hz. The output signal of FIR1 is calculated by Eq.(1) in the steady state. Thus, calculating a single output requires N computations. This means only N samples of the 1000 length coefficients are used.

FIR2 does not require the output power. A single output signal in the steady state can be used to detect the frequency component. Order of the transfer function is 9th and 19th-order for N=10 and 20, respectively.

Accuracy of signal classification in percentage for the test signals by all the methods previously described

Table 1: Number of parameter of NN and conventional methods.

	limited		not limited	
Methods	N=10	N=20	N=10	N=20
MLNN	3	3	40	40
FIR1	2	2	10	10
FIR2	1	1	1	1
Fourier	1	1	3	3
Euclid	2	2	200	200
MGD	2	2	200	200
MEM	1	2	3	3

N : number of samples

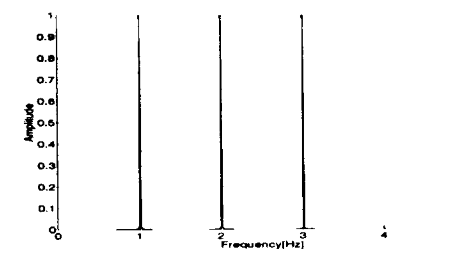


Fig. 2: Amplitude repsonse of FIR filter which designed to extract group 1 signals.

are listed in Table 2. In this table, "computation limited" means computational load required in the conventional methods is limited to be the same as in the MLNN method. In this case, the MLNN method performs well. The accuracy for N=10, noise free and noisy signal are 97.6% and 85.4%, respectively. FIR2 shows better accuracy than the MLNN with noise free signals, however it decreases the accuracy for noisy signals to 50.3%. Because FIR2 cannot suppress the noise spectrum. The accuracy of the rest methods are lower than the MLNN.

In the conventional methods, the accuracy can be improved by increasing computational load. In Table 2, "computation not limited" means the conventional methods can use computations so as to provide the highest accuracy. The valley shape activation function [8] is used instead of the sigmoid function in the hidden layer of the MLNN in this case. For 10 samples, MLNN and FIR1 show the best accuracy among all methods. For 20 samples, MLNN shows good accuracy which is the same as other methods. This means the MLNN can provide good performance in the multi-frequency signal classification in all cases.

Figure 3 shows the accuracy of the signal classification by the MLNN and the conventional methods. The horizontal axis means the number of hidden units, which equivalently indicates computational complexity. The noisy signals are used. From these figures, the MLNN performs better than conventional methods for any computational requirements.

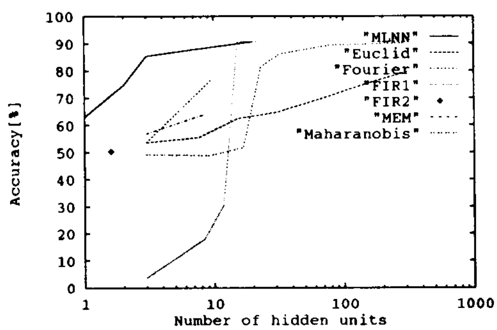
Robustness of the MLNN for noise level changes is further investigated. The MLNN, which is trained using noisy 20 signal samples with ±0.5 additive noise, provides 91.7% accuracy for ±0.2 additive noise, and 91.3% accuracy for the noise free signals. The number of the training signals are increased from 200 to 400 in

Table 2: Accuracy of signal classification in percentage.

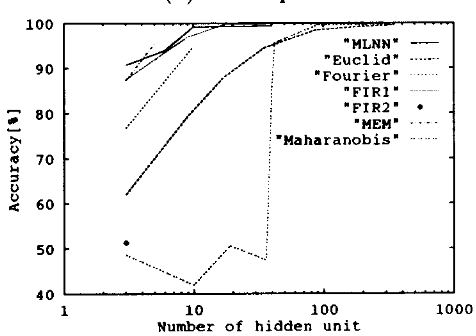
Methods	Computation limited				Computation not limited			
	N=10		N=20		N=10		N=20	
	NFS	NS	NFS	NS	NFS	NS	NFS	NS
MLNN	97.6	85.4	97.4	90.6	100	90.6	100	99.3
FIR1	4.7	3.7	100	87.5	100	90.5	100	99.8
FIR2	100	50.3	100	51.3	-	-	-	-
Fourier	56.1	53.6	77.9	76.7	70.6	65.7	100	94.8
Euclid	49.6	52.1	59.4	62.0	86.0	79.5	100	99.5
MGD	48.6	48.8	50.8	48.6	100	90.2	100	99.7
MEM	60.8	56.8	87.7	87.3	62.9	63.7	97.3	95.4

N : Number of samples, NFS : Noise Free Signal, NS Noisy Signal

this case. Thus, the robustness can be confirmed.



(a) 10 samples



(b) 20 samples

Fig. 3: Accuracy of the MLNN and the conventional methods for (a)10 samples and (b)20 samples noisy signal.

5.5. Signal Detection Regions by MLNN

The signal detection regions formed by the MLNN are investigated based on the hidden layer outputs and the connections from the hidden layer to the output layer. 10 samples noisy signals are used.

Figure 4 shows input-output distributions of the hidden units. In this figure, (a,b), (c,d) and (e,f) correspond to the 1st, 2nd and 3rd hidden units, and (a,c,d), (b,d,f) correspond to #1, #2 signal groups, respectively. From these figures, there are two types of distributions, concentrated and spread. For example, Fig.4(b) shows the 1st hidden unit inputs are concentrated for #2 signals. Hence, the 1st hidden unit is used for detecting #2 signals. The same situation occurs in the 2nd and 3rd hidden unit as shown in Fig.4(d) and (e), respectively. From these distributions, however, we can not estimate the combination of the hidden unit outputs.

The connection weights from the hidden layer to the 1st and the 2nd output units, which respond to the #1 and #2 signal group, respectively, are listed in Table 3. The connection weights from 1st, 2nd, 3rd hidden units and the bias unit to the 1st output unit are -18.95, 18.27, 11.63 and -2.0, respectively. The connection weights to the 2nd output unit have the opposite polarity to those of the 1st one.

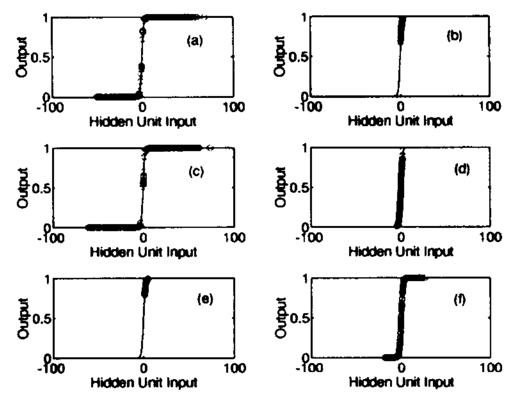


Fig. 4: Input-output distribution of hidden units. (a), (c) and (e) are for #1, and (b), (d) and (f) are for #2 signals, respectively.

Table 3: Connection weights from hidden layer to output layer.

Hidden	Output	
	1st	2nd
1st	-18.95	18.95
2nd	18.27	-18.27
3rd	11.63	-11.63
Bias	-2.0	2.0

Based on these connections, and in order to activate the 2nd output unit, the following pattern is permitted, that is (Hidden units : 1st, 2nd, 3rd) = (H, L, L), (H, H, L), (H, L, H), where H and L mean 'High level' and 'Low level', respectively. On the other hand, there are 4 patterns to activate the 1st output unit, (Hidden units : 1st, 2nd, 3rd) = (L, H, H), (L, H, L), (H, H, H) and (L, L, H).

Table 4 shows the combination of the hidden unit outputs for #1 and #2 signals. For #1, there are two combinations, (Hidden units : 1st, 2nd, 3rd) = (H, H, H), (L, L, H). For #2, three combinations, that is (Hidden units : 1st, 2nd, 3rd) = (H, H, L), (H, L, H), (H, L, L) are formed. As expected from the theoretical discussion, the MLNN effectively classified the multi-frequency signals using higher degree of freedom of forming the signal detection region.

Table 4: Hidden unit output distribution.

Group 1 (#1)				Group 2 (#2)			
1st	2nd	3rd	Number	1st	2nd	3rd	Number
H	H	H	112	H	H	H	0
H	H	L	0	H	H	L	52
H	L	H	0	H	L	H	109
H	L	L	0	H	L	L	39
L	H	H	0	L	H	H	0
L	H	L	0	L	H	L	0
L	L	H	88	L	L	H	0
L	L	L	0	L	L	L	0

Number : number of hidden units
1st *sim* 3rd : hidden unit output

6. Dial Tone Recognition

Dial tone recognition is used in the push button telephone to generate the signals correspond to the numerical and function buttons. This is an example of the multi-frequency signal classification. Two groups of high and low frequencies are used. Table 5 shows the combination of frequencies.

The amplitude range and phase are distributed in the same ranges given by Eq.(18). The sampling frequency is 4Hz. The number of samples is 10 or 20.

Table 5: Relation between combination of frequencies and dial tones #1 ~ #16.

	1.209	1.366	1.477	1.633
0.697	#1	#2	#3	#4
0.770	#5	#6	#7	#8
0.852	#9	#10	#11	#12
0.941	#13	#14	#15	#16

6.1. Classification by MLNN

Table 6 shows classification accuracy for the test signals. 50 hidden units, whose activation function is a sigmoid function, are used. In the both cases, the accuracy is very high. From the results, This kinds of complex problem can be solved by the MLNN with small computation.

Table 6: Accuracy of NN for dial tone recognition using MLNN method.

Signal Sample	Accuracy[%]
10	90.6
20	95.7

6.2. Classification by FIR1

As the conventional methods, FIR1 is used to classify the dial tone signals. Eight kinds of FIR filters are designed to extract each frequency component. A single FIR filter extracts only one frequency component. The power of the FIR1 outputs are calculated, and are added to extract one of 16 combinations.

Table 7 shows accuracy of dial tone recognition. "output samples" means the number of the output samples used in the power calculation. 200 output samples is sufficient to provide the highest accuracy in the FIR1 method. The accuracies of the FIR1 are lower than those of the MLNN.

Table 7: Accuracy of dial tone recognition using FIR1 method.

Signal Sample	Output Sample	Accuracy[%]
10	14	23.3
	200	41.2
20	10	79.4
	200	83.6

7. Conclusion

Classification performance of the signal which carring information as frequencies, using the MLNN and the conventional signal processing methods has been compared under the limited observation period and computational load. The comparison has been carried out based on degree of freedom to form the signal detection regions in the N-dimensional signal space. The MLNN method has higher degree of freedom, and can provide more flexible performance for detecting the frequency components than the conventional methods. The result has been supported through the simulations of multi-frequency signal classification and the dial tone receiver.

References

[1] D.E.Rumelhart and J.L.McCelland et al. Parallel Distributed Processing, MIT Press, 1993

[2] J.Zhan and F.Li, "A self-organization neural network and detecting signals", Proc. WCNN'93, Portland, Oregon, vol.IV, pp.748-751,1993

[3] R. Vanderbeek and A. Garper, "A back-propagation network for analog signal separation in high environments", Proc. IJCNN'92, pp.I-664 - 669, Boltimore MD, June,1992.

[4] K.A.Al-Mashouq and I.S. Reed, "The use of neural nets to combine equalization with decoding for severe intersymbol interference channels", IEEE Trans. on Neural Network, Vol.5, No.6, pp.982–988, Nov., 1994.

[5] D.M.Hummels, W.Ahmed, and M.T.Musavi, "Adaptive detection of small sinusoidal signals in non-gaussian noise using an RBF neural network", IEEE Trans. on Neural Network, Vol.6, No.1, pp.214–219, Jan., 1995.

[6] Z.H.Michalopoulou, L.W.Nolte, and D.Alexandrou, "Performance evaluation of multilayer perceptrons in signal detection and classification", IEEE Trans. on Neural Network, Vol.6, No.2, pp.381–386, March, 1995.

[7] G.Veciana and A.Zakhor, "Neural Net-Based Continuous Phase Modulation Receivers", IEEE Trans. Communications, vol.40, No.8, 1992.

[8] K.Nakayama and K.Imai, "A neural demodulator for amplitude shift keying signal", Proc. ICANN'94, Sorrent, vol.2, 1017-1020, 1994.

[9] D.R.Hush and B.G.Horne, Progress in supervised neural network, IEEE Signal Processing Magazine, pp.8-39, 1993.

[10] K.Hara and K.Nakayama, "High resolution of multi-frequencies using multilayer networks trained by back-propagation algorithm ", Proc. WCNN'93, Portland Oregon, vol.IV, pp.675–678,1993.

[11] K.Hara and K.Nakayama, "Comparison of activation functions in multilayer neural network for pattern classification ", Proc. ICNN'94, Orland Florida, pp.2997-3002, 1994

[12] Jae S.Lim and Alan V. Oppenheim, Advance topics in Signal Processing, Prentice–Hal, Inc., 1992.

[13] J.B.MacQueen, "Methods for classification and analysis of multivariate observations", Proc. Symp. Math. Statist. and Probability, 5th, Berkeley, 1, pp.281–297, 1967.

[14] M.R.Anderson, Cluster analysis for applications, Academic Press, pp.162–163, 1973.

[15] R.O.Duda and P.E.Hart, Pattern classification and scene analysis. Jhon Wiley and Sons, 1973.