

ディレイ素子を含むニューラルネットワーク におけるゲイン学習の効果

西 正明* · 降矢順治** · 林川基治***

An Effect of Gain Training on Neural Networks Including Delay Elements

Masaaki NISHI* , Junji FURUYA** , and Motoharu HAYASHIKAWA***

(Received September 28, 2001)

1. まえがき

ニューラルネットワークは脳の処理様式を模した処理システムであり、我々はこれまで、内部にディレイ素子を組み込んで過去の出力を学習と処理に利用することのできる階層型ニューラルネットワーク (BPD) を提案してきた¹⁾。本論文では、この BPD に FIR シナプスを付加したニューラルネットワークを用いる²⁾。このようなニューラルネットワークのユニットでは前の時刻の出力を入力データとして再入力すると同時に、前の時刻の入力を再び入力データとして入力することになり、様々な適用問題において識別能力や汎化能力が高まる²⁾。学習アルゴリズムには、バックプロパゲーション学習アルゴリズムを用いる。学習では通常結合荷重だけを更新して学習を行うが、本論文では各ニューロンに組み込まれる応答関数のゲインも学習の対象にして学習を行い、結合荷重だけを学習する場合と比較し検討する。適用問題として、AM 波、FSK 波および矩形波におけるノイズ除去問題を扱って検討する。

2. ニューラルネットワークの構成と学習アルゴリズム

ディレイ素子をユニットの出力に付加して前の時刻の出力を入力データとして再入力するようにしたユニットを BPD ユニットと呼ぶことにする。ディレイ素子をユニットの入力側に多段に付加して前の時刻の入力を入力データとして再入力できるような FIR シナプスを組み込んだユニットを FIR ユニットと呼ぶことにする。図 1 に FIR シナプスを示す。BPD ユニットと FIR ユニットの組み合わせたユニットを用いて、3層構成の出力フィードバックにディレイ素子を含む FIR ニューラルネットワークを図 2 に示す。学習アルゴリズムは、従来から用いられているバックプロパゲーション学習アルゴリズムで最急降下法によって得られ、以下に簡単に述べる。

* 信州大学教育学部

** 城西国際大学経営情報学部

*** 山口大学教育学部

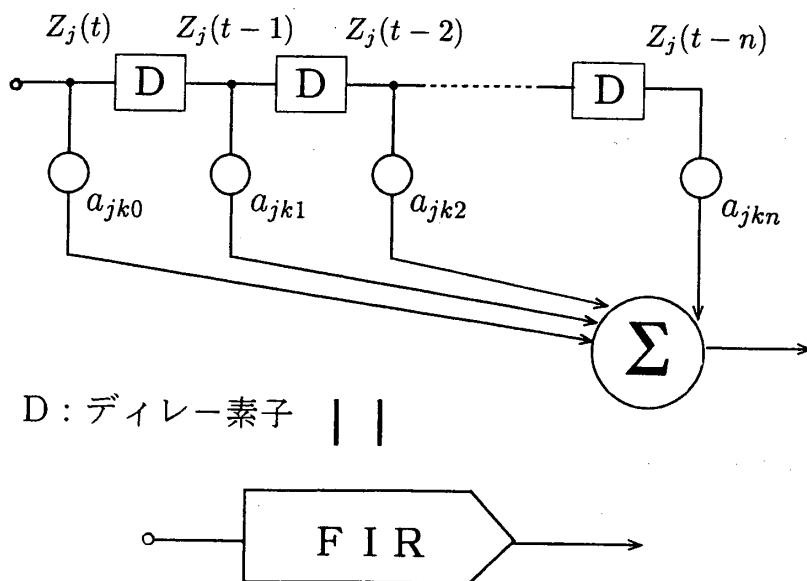


図1 FIR シナプス

第 l 層の k 番目のニューロンの内部状態 $X_k^l(t)$ と出力 $Z_k^l(t)$ を次式で求める。

$$X_k^l(t) = a_{0k} Z_0^{l-1}(t) + \sum_{j=1}^{N_j} \sum_{p=0}^L a_{jkp} Z_j^{l-1}(t-p) + \sum_{n=1}^N b_{kn} Z_k^l(t-n) \quad (1)$$

$$Z_k^l(t) = f(X_k^l(t)) = \frac{1}{1 + \exp(-2X_k^l(t)/u_0)} = \frac{1}{1 + \exp(-g_k^l X_k^l(t))} \quad (2)$$

式 (1) はニューロンの内部状態 $X_k^l(t)$ であり、第 1 項はオフセット部分で $Z_0^{l-1}(t)$ をオフセット値 1 とし結合荷重 a_{0k} で調整する。第 2 項は FIR シナプス部分の算出項、第 3 項は出力からのフィードバック部分の算出項である。式 (2) はニューロンの出力 $Z_k^l(t)$ であり、 u_0 は応答関数の傾きを規定するパラメータである。本論文では g_k^l をゲインとして扱う。結合荷重の修正量は、教師信号を $d_k(t)$ とし平均自乗誤差を式 (3) とすれば、式 (4) で求められる。ここで、 γ はオフセットの学習係数で、 α は FIR シナプスの学習係数、 β は出力からのフィードバックでの学習係数である。式 (4) を求める過程で微分式に間接的な項が生じるが、これを省略しても結果にほとんど影響が無いことが確認されているので³⁾、省略して簡単化して求めている。ゲインの修正量も同様にして式 (5) で求められる。 η はゲインに対する学習係数であり、ここでも間接的な微分項を省略して簡単化して求めることにした。

$$E_k(t) = \frac{1}{2} \sum_k (d_k(t) - Z_k^l(t))^2 \quad (3)$$

$$\Delta a_{0k} = -\gamma \frac{\partial E_k(t)}{\partial a_{0k}}, \quad \Delta a_{jkp} = -\alpha \frac{\partial E_k(t)}{\partial a_{jkp}}, \quad \Delta b_{kn} = -\beta \frac{\partial E_k(t)}{\partial b_{kn}} \quad (4)$$

$$\Delta g_k^l = -\eta \frac{\partial E_k(t)}{\partial g_k^l} \quad (5)$$

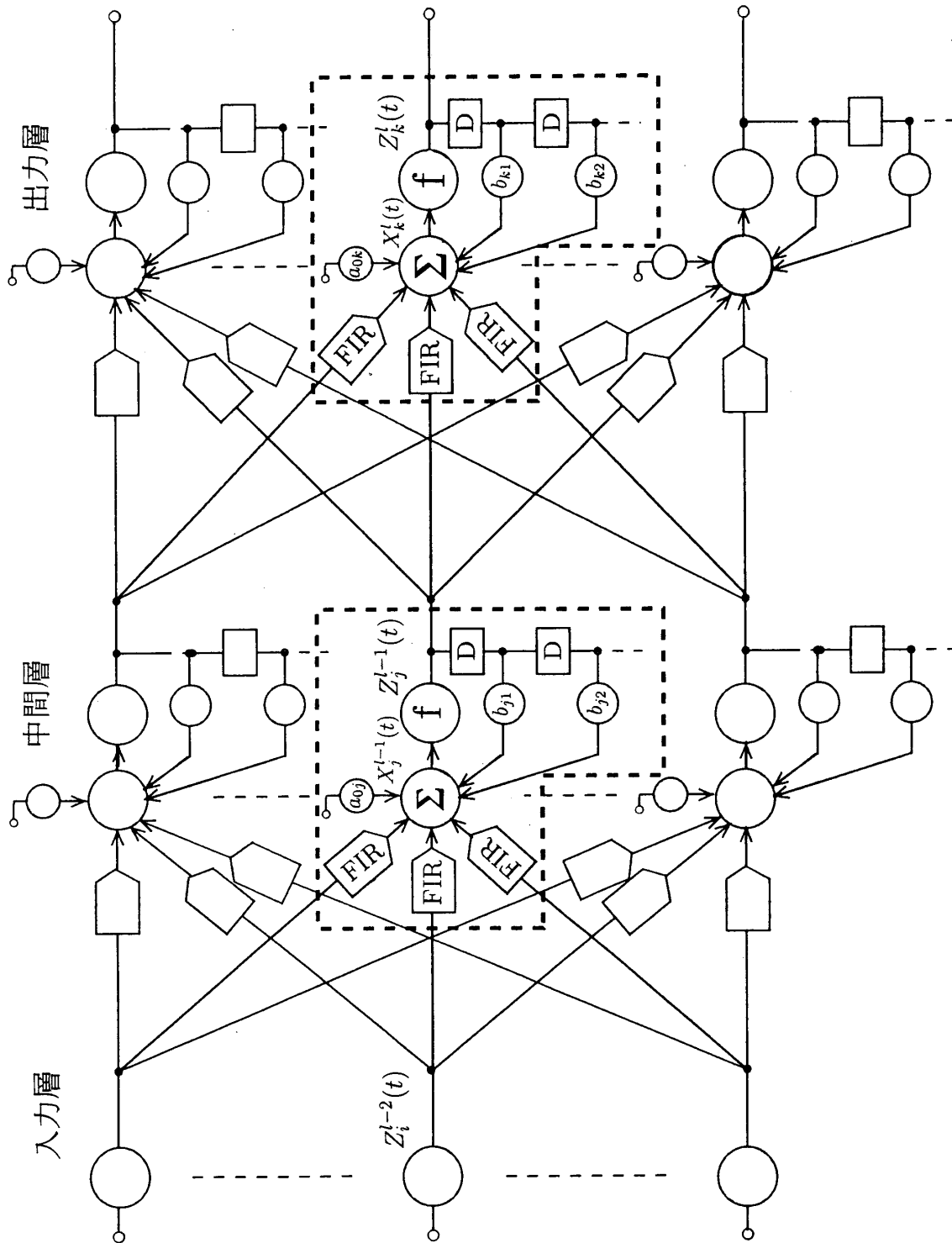


図 2 出力フィードバックにディレイ素子を含む FIR ニューラルネットワーク

3. シミュレーション結果

本論文では、出力からのフィードバックにディレー素子を多段に含む FIR ニューラルネットワークを扱う。入力層、中間層、出力層の各ユニット数を i, j, k とし、FIR シナプスの中間層と出力層の各ディレー段数を p, q とする。中間層と出力層での出力からのフィードバックディレー段数をそれぞれ m, n とする。これらの記号を用いて、本論文で扱うニューラルネットワークの構成を $i - (p)j(m) - (q)k(n)$ と表すことにする。

シミュレーションは中間層ユニット数 j を 5, 10, 15 の 3 通り、ディレー段数 p, q および m, n を 5, 10, 15 の 3 通りとして、その全ての組み合わせ 243 通りについて行う。これらのニューラルネットワークを AM 波、FSK 波、矩形波のノイズ除去問題に適用して、誤差の最も小さいネットワーク構成をそれぞれ選定した。学習原波形を図 3、 $\pm 10\%$ のノイズを付加した学習処理波形を図 4 にそれぞれ示す。学習では図 4 のノイズを付加した波形を用いる。認識処理には学習に用いた波形とは位相の違えた別の波形を用いる。認識原波形と $\pm 10\%$ のノイズを付加した認識処理波形を図 5 と図 6 にそれぞれ示す。図 6 のノイズを付加した波形を認識に用い、結果を図 5 の認識原波形と比較する。シミュレーションは以下のパラメータで行った。学習は誤差が 0.006 以下になるまで行い、5000 回で学習を打ち切ることとした。

- ・ゲインを学習しない場合

- シグモイド関数の傾き (出力層 g_k , 中間層 g_j): 0.75

- FIR シナプスの結合荷重修正用学習定数 (α): 0.4

- オフセットの結合荷重修正用学習定数 (γ): 0.11

- 出力フィードバックの結合荷重修正用学習定数 (β): 0.4

- ・ゲインを学習する場合

- シグモイド関数の傾き初期値 (出力層 g_k , 中間層 g_j): 0.8

- FIR シナプスの結合荷重修正用学習定数 (α): 0.6

- オフセットの結合荷重修正用学習定数 (γ): 0.1

- 出力フィードバックの結合荷重修正用学習定数 (β): 0.6

- ゲイン修正用学習定数 (η): 0.2

シミュレーションの結果、ゲインの学習をしない場合には、AM 波で $1 - (15)15(10) - (10)1(15)$ 、FSK 波で $1 - (15)15(5) - (15)1(10)$ 、矩形波で $1 - (5)15(15) - (15)1(5)$ の構成の場合がもっともノイズ除去特性がよかった。認識出力波形を図 7 に示す。同様に、ゲインの学習をする場合には、AM 波で $1 - (15)15(15) - (15)1(5)$ 、FSK 波で $1 - (15)15(15) - (15)1(10)$ 、矩形波で $1 - (10)15(5) - (15)1(5)$ の構成の場合がもっともノイズ除去特性がよく、これらの認識出力波形を図 8 に示す。図 7 と図 8 を比較すると、AM 波の場合と FSK 波の場合では、図 7 よりも図 8 に示すゲインを学習した場合の方が周期の短い細かいノイズをよく除去して認識原波形に近い波形を出力していることがわかる。矩形波の場合には差は大きくは無いが、同様の傾向が認められる。すなわち、図 7 よりも図 8 に示したゲインを学習する場合の方が周期の短い細かいノイズをよりよく除去して出力する傾向が認められる。

学習経過における誤差の推移はパラメータによって一定ではないが、ゲイン学習の有無と

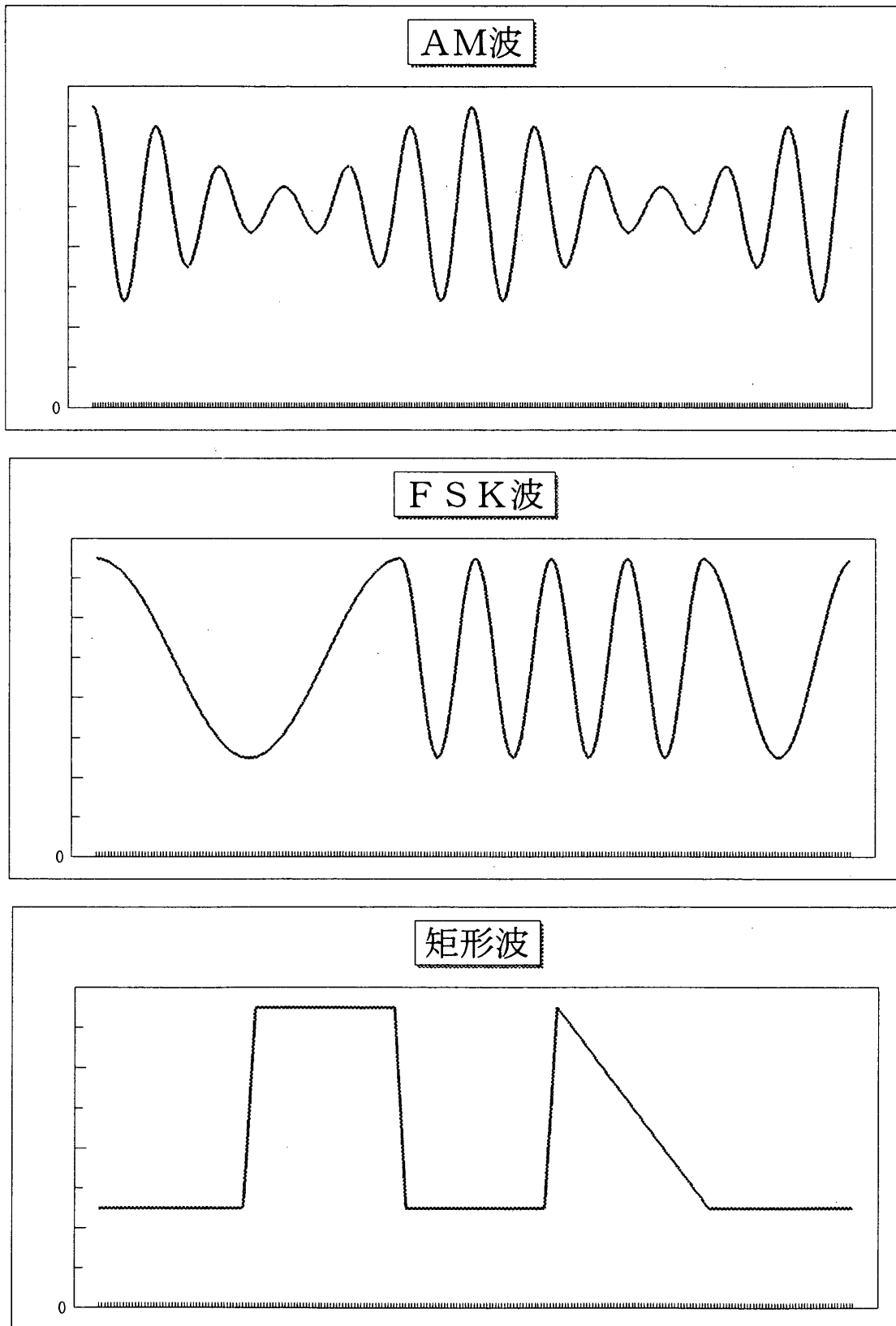


図3 学習原波形

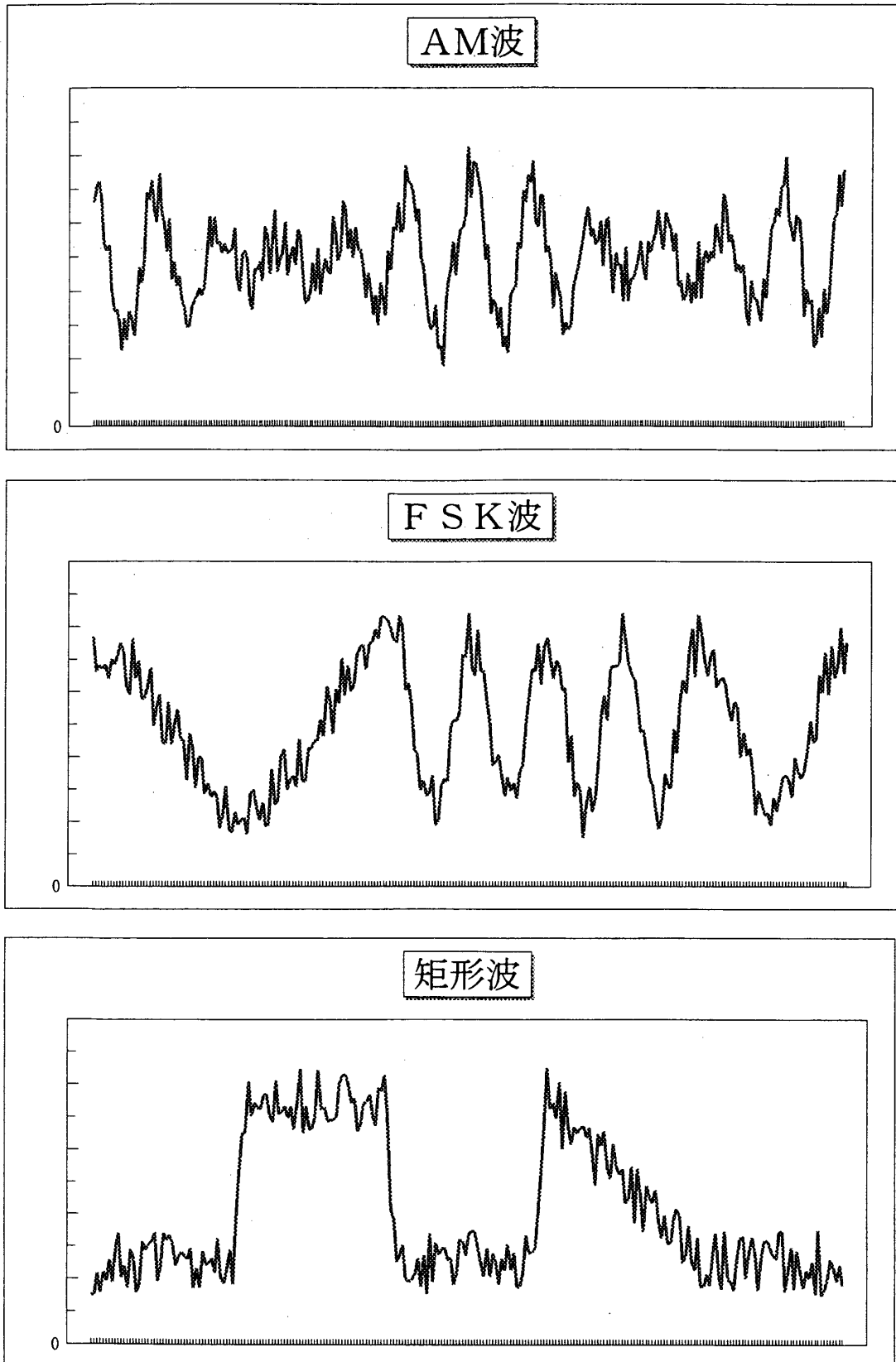


図 4 ノイズ付加学習波形

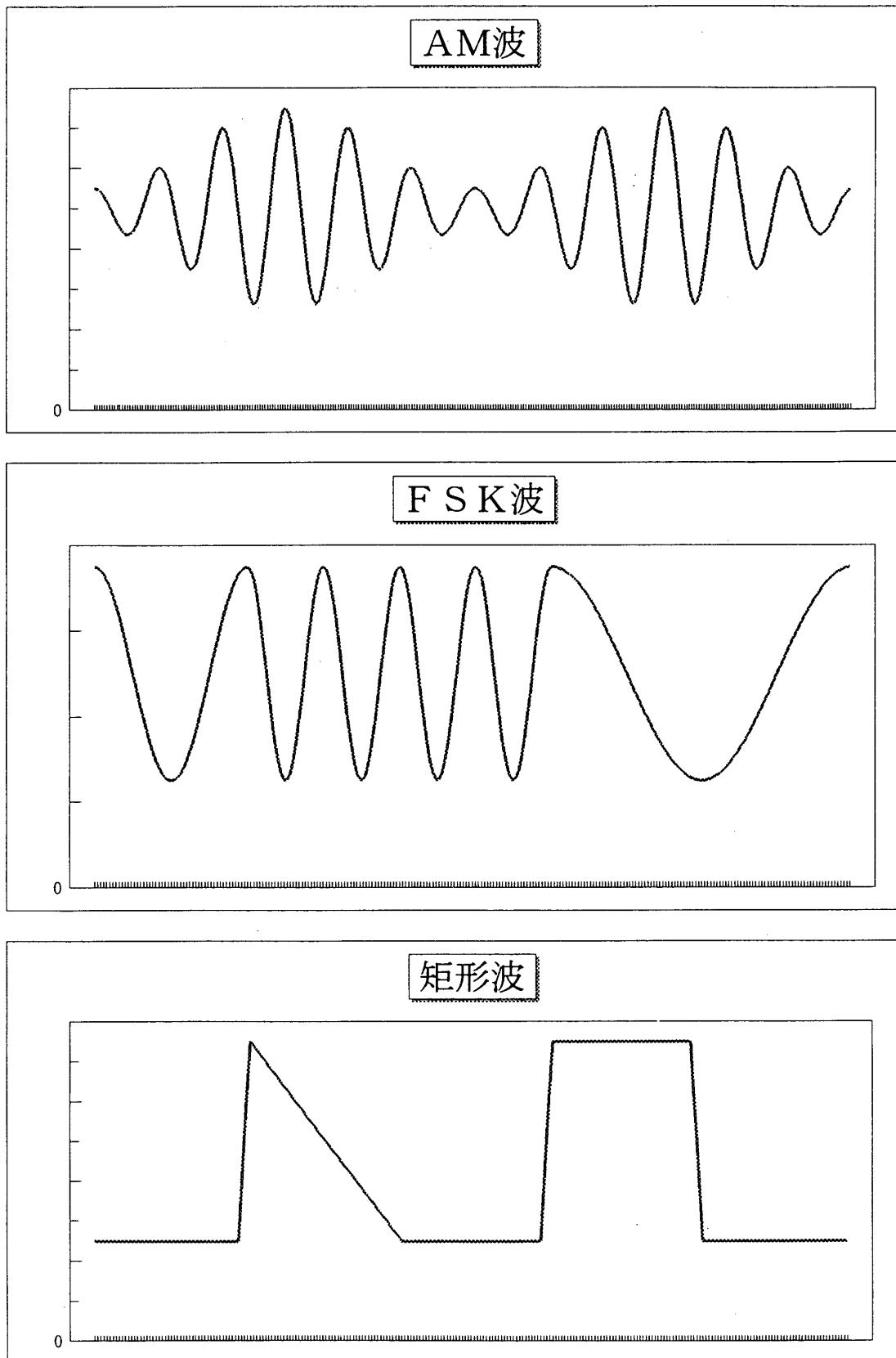


図5 認識原波形

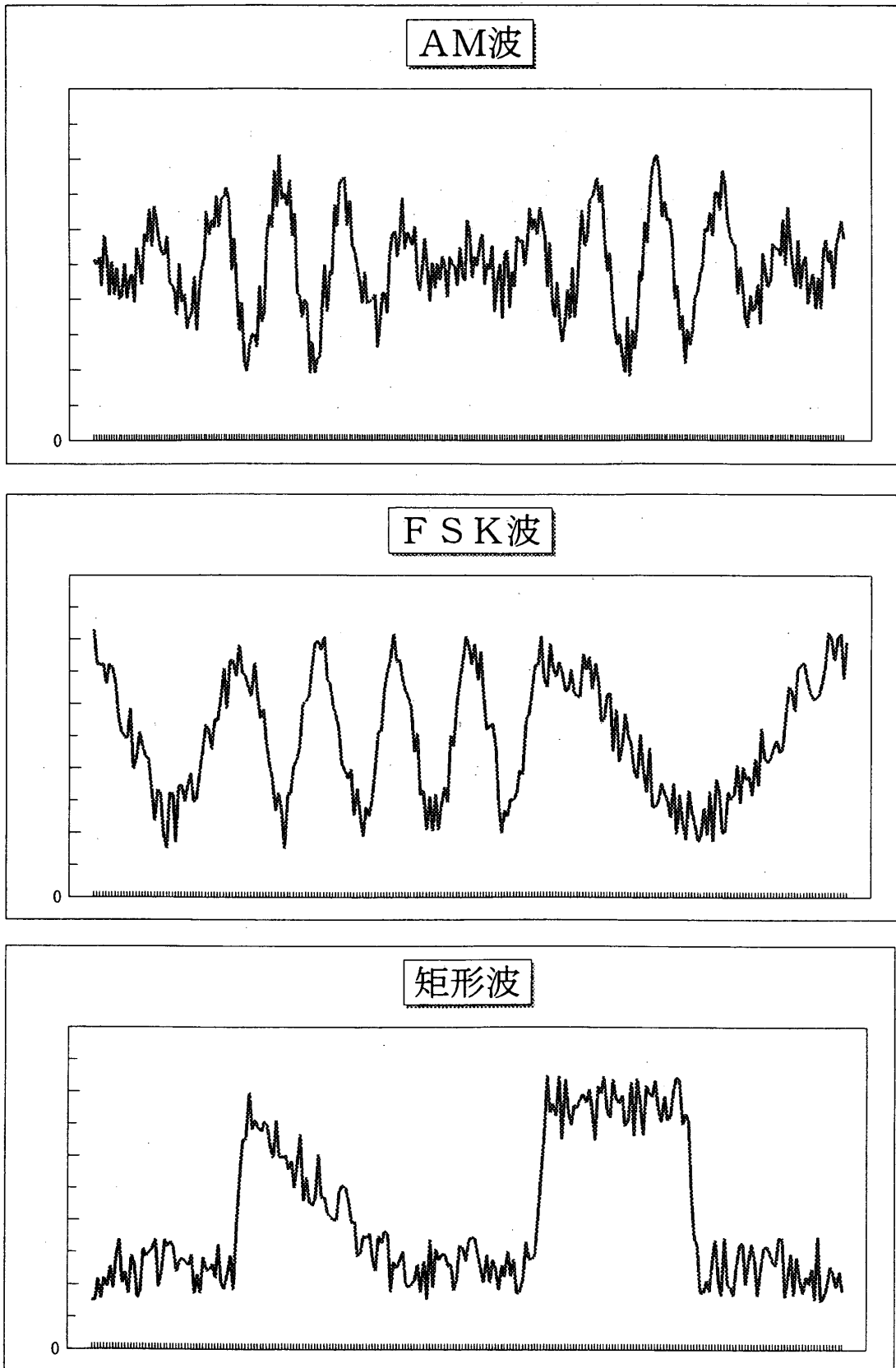


図 6 ノイズ付加認識波形

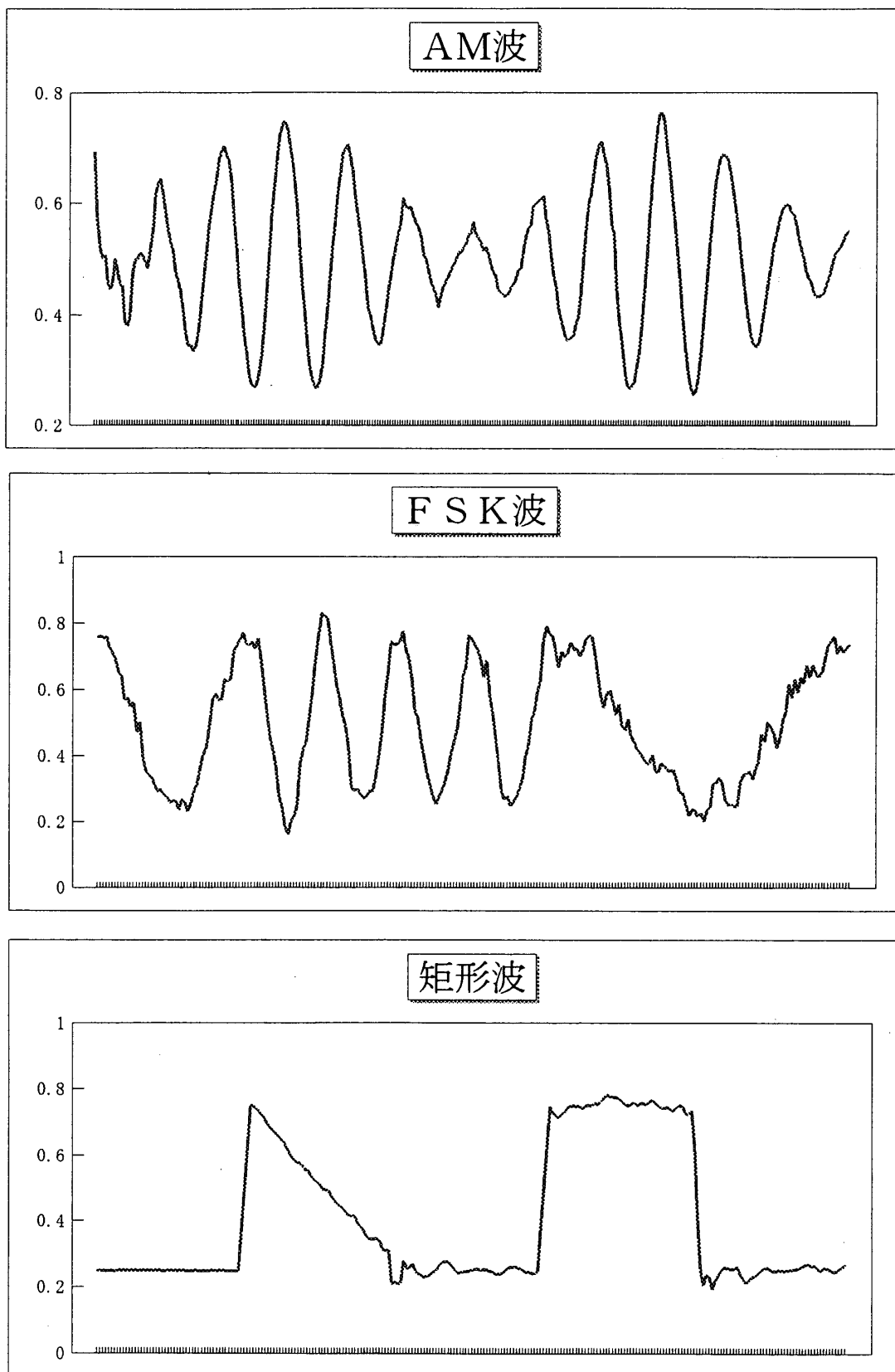


図7 ゲイン学習なしの場合の認識出力波形

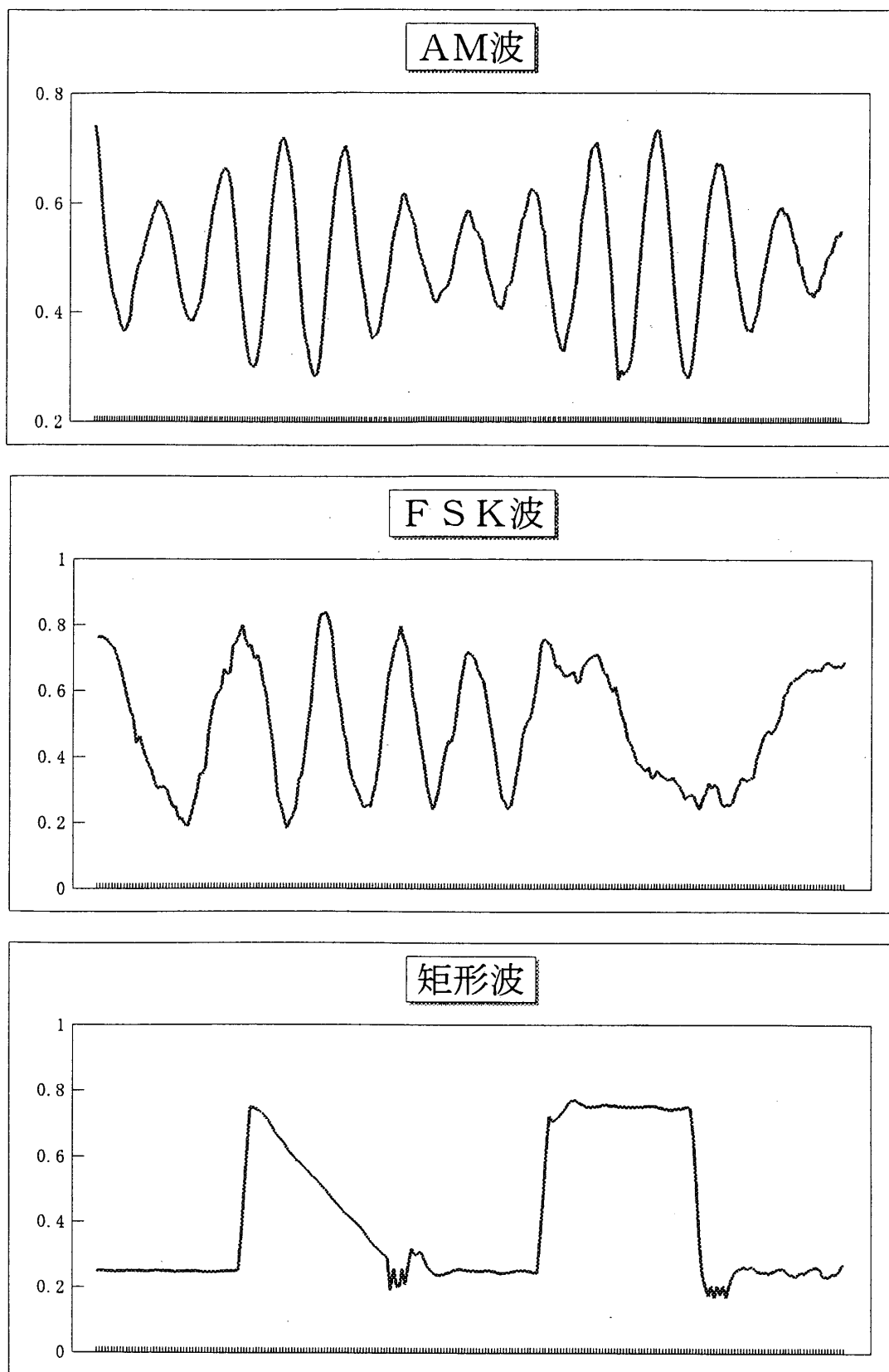


図 8 ゲイン学習ありの場合の認識出力波形

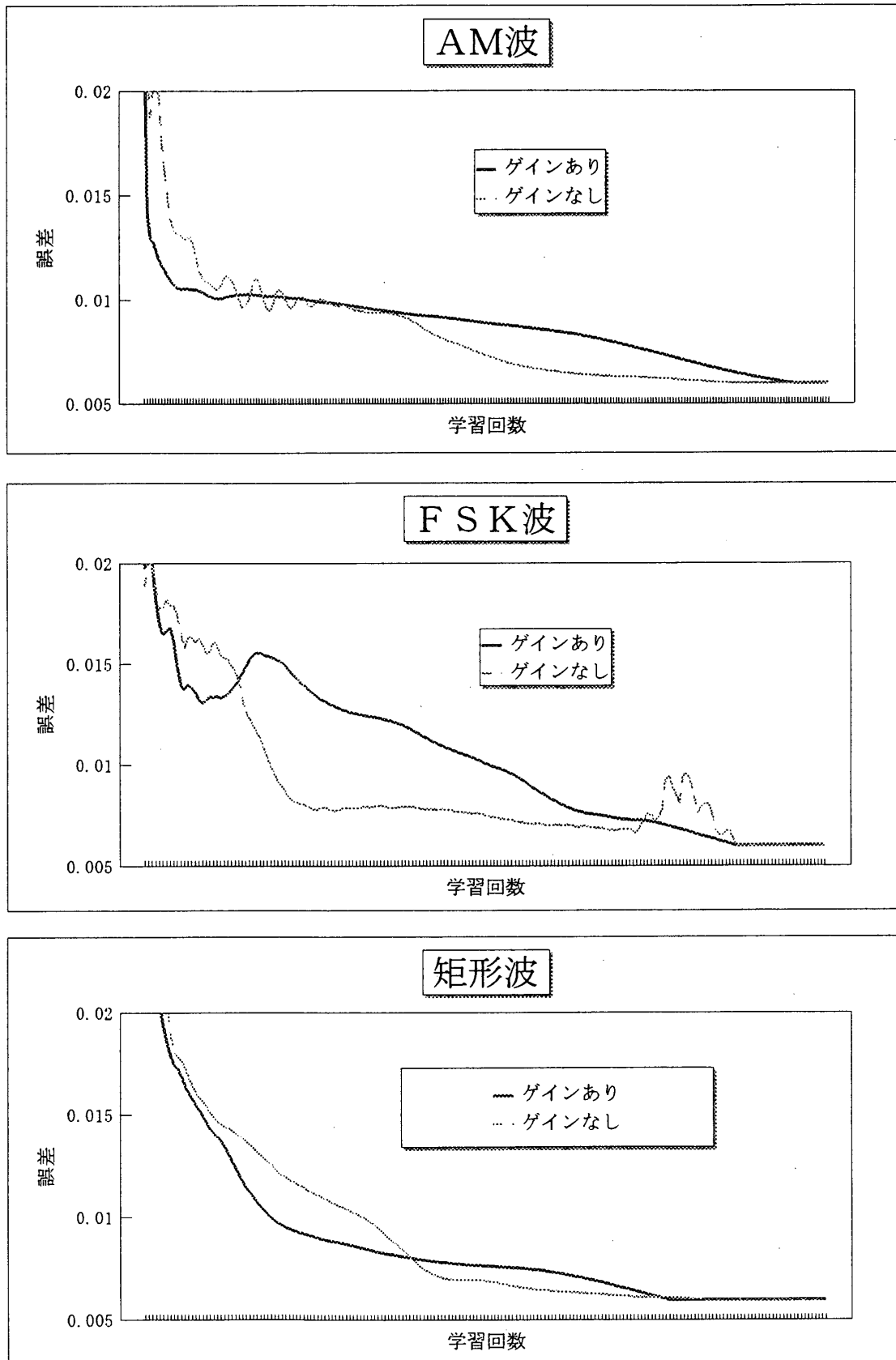


図9 ゲイン学習の有無と学習誤差の推移

学習経過における誤差の推移の一例を図9に示す。図9から、ゲイン学習なしの場合は誤差値が大小に大きく振動する傾向があることが認められる。ゲイン学習ありの場合は学習の初期では誤差値がゲイン学習なしの場合よりも急激に小さくなるが、学習が進むにつれて誤差の減少は小さくなり、最終的には多くの学習回数を必要とする傾向があることが認められる。

以上より、出力のフィードバックにディレー素子を含むFIRニューラルネットワークにおけるゲイン学習は、学習の高速化にはほとんど効果は認められないが、学習後のニューラルネットワークのノイズ除去特性の向上には効果があることがわかった。

4. むすび

出力からのフィードバックに多段のディレー素子を付加して過去の出力値を入力データに用い、同時にFIRシナプスによって過去の入力値も入力データに用いる構成のニューラルネットワークを、AM波、FSK波および矩形波におけるノイズ除去問題に適用した。このニューラルネットワークの学習において、各ユニット内の応答関数のゲイン学習の有無による学習特性とノイズ除去特性を比較検討した。その結果、ゲイン学習は学習の高速化にはほとんど効果はないが、ノイズ除去特性の向上には効果があることがわかった。このことは、今後の情報教育に応用していくうえで有益な知見である。

今回はノイズ除去特性で比較評価したが、今後は音声認識や符号化並びに暗号化など他の時系列処理問題にも適用して検討していきたい。

参考文献

- 1) M.Nishi, J.Furuya, and T.Nakamura, "Backpropagation Networks Including Time Delay elements(BPD)," Proc. of IJCNN'93, IEEE, pp.1681-1684, 1993.
- 2) 西正明, 降矢順治, 林川基治, "ディレー素子を含むニューラルネットワークによるノイズ除去特性," 山口大学教育学部附属実践総合センター研究紀要, 巻12号, pp.195-208, 2001.
- 3) 西正明, 降矢順治, 中村維男, "ディレー素子内包型バックプロパゲーションニューラルネットワーク(BPD)の一構成," 信学論(D-II), Vol.J78-D-II, No.10, pp.1522-1530, 1995.