ニューラルネットワークによる位置ずれ・回転に 不変なパターン認識システム

正員 福 見 稔 (徳島大)
非会員 細 川 直 史 (徳島大)
正員 大 松 繁 (徳島大)

Pattern Recognition System Insensitive to Translation and Rotation by Neural Network

Minoru Fukumi, Member, Masafumi Hosokawa, Non-member, Sigeru Omatu, Member
(University of Tokushima)

In pattern recognition, we often deal with classification problems in which the retinal input patterns may be subject to translation, rotation, scale change, perspective change etc. Usual pattern recognition methods by statistical discrimination functions are not powerful to classify those patterns.

In this paper, a new approach for pattern recognition by a neural network is proposed based on the concept involving an invariance network and a descrambler network. The invariance network plays an important role in producing a set of outputs which are invariant to translation, and rotation, etc., of the retinal input pattern. The descrambler network is used to classify the scrambled data into the original patterns by using back-propagation algorithm. The structure of proposed neural networks is similar to that by Widrow, et al.

But the sigmoid functions are adopted as nonlinear elements in the neural networks while Widrow's MR II is based on signam functions. Hence, the back-propagation method can be used in the learning algorithm of descrambler network.

Finally, some numerical results are illustrated to show the effectiveness of the present algorithm for pattern recognition.

キーワード：ニューラルネットワーク，パターン認識，バックプロパゲーション

1. はじめに

近年，ニューラルネットワークが，パターン認識，特徴抽出などにおいて識別能力が高いことが示されている。ビジュアルパターンに対して，福島らのネオコグニトロン⁽¹⁾は，視覚システムと似た階層的構造をもった多層ネットワークであり，手書き文字認識に対して有効な手法を提供している。また，Widrowらは，基本構成要素としてシグナム関数を用いたモデルである線形適応ニューロン (ADALINE) を考案した⁽²⁾。

更に，これを用いたニューラルネットを提案し⁽²⁾⁽³⁾，これを基にした回転と位置ずれに対して不変な多層ADALINEネットワークを提案した。このネットワークは，インバリانسネットワークと，ディスクランブラネットワークよりなる。インバリانسネットワークは，位置ずれ，回転，スケール変換，視点変換に対して不変な出力を生成し，ディスクランブラネットワークは，ADALINEネットワークを学習させた後オリジナルパターンを再生する。ここでの学習アルゴリズムは，MR IIルールと呼ばれている。更に，

MR IIルールは学習アルゴリズムが統一的でなく試行錯誤的である。従ってWidrowらの手法は概念として優れたものであるが、パターン認識に対してあまり有効なものとなっていなかった。本論文では、新しいタイプのニューラルネットワークを用いた位置ずれ、回転、雑音などの影響を受けにくい文字認識手法を提案する。ここで提案されるニューラルネットワークは、非線形関数としてシグモイド関数を用い、最小二乗原理の代わりにバックプロパゲーションアルゴリズム⁽⁴⁾を採用したインバリانسネットワークとディスクランブラネットワークの部分からなっている。インバリانسネットワークは、入力パターンの位置ずれ、回転に対して不変な出力の集合を生成するのに重要な役割を果たす。ディスクランブラネットワークは、バックプロパゲーションアルゴリズム⁽⁴⁾に基づいてインバリانسネットワークの出力をオリジナルパターンに分類するのに使われる。このインバリانسネットワークの構造は、Widrowらによって提案されたものに似ているが⁽²⁾、WidrowらがADALINEに基づき、非線形関数としてシグナム関数を用いているのに対して、提案するインバリانسネットワークは、非線形関数としてシグモイド関数を採用している。これにより、ニューロンの出力値がアナログ値となり、こう配法に基づいたバックプロパゲーションアルゴリズムが適用可能なシステムとなっている。この構造を採用することによって、位置ずれや回転、または雑音や孤立点を含んだパターンの分類に対して、Widrowらの手法より効果的であることをシミュレーションにより定量的に検証する。

2. ADALINE と MR II ルールの要約

文献(2)のADALINEは、多層ニューラルネットの基本構成要素で、図1で表される。

WidrowらはADALINEを適応させることでネットワークの出力を目標の応答に漸近させるニューラルネットワークを提案した⁽²⁾⁽³⁾。

図2はWidrowらによる多層ADALINEネットワークを示している。このニューラルネットワークでは出力層を除いて個々のADALINEの目標値を設定することができないため、最小二乗誤差規範に基づいたアルゴリズム（以下、LMSアルゴリズムと呼ぶ）を直接適用できなかった。そこで、Widrowらは、最小二乗原理に基づいた多層ADALINEネットワークの学習則を提案した⁽²⁾⁽³⁾。これはMR IIルールと呼ばれており、その原理を以下に示す。まず、第一層のニューロンで、アナログ出力が最も0に近いユニットのバ

図1 ADALINEの構造
Fig. 1. Structure of ADALINE.

図2 多層ADALINEネットワーク
Fig. 2. Multi layered ADALINE network.

イナリー出力を反転させるような適応規則を与える。もし、この適応規則が出力のエラーの数を少なくするならば、ウェイトの変更を受け入れる。そうでなければ、ウェイトは適応以前の値へもどされる。次に、2番目にアナログ値が0に近いニューロンに対して同様な規則をあてはめ、これらのことを繰返し行い、できるだけネットワークをかく乱しないような適応規則を用いる。第一層のニューロンのすべてを一つずつ適応させた後、2個の組合せを、次に3個の組合せをそれぞれ、あらかじめ決めた個数まで適応を行う。次に、第二層に対してエラーを少なくするために同様な操作を繰返す。最近、Winterら⁽³⁾によって、できるだけ組合せの数を少なくするためのアルゴリズムが提案されているが、試行錯誤の域を出ていない直感的な手法である。このMR IIルールの有効性を見るために、XOR回路を図2のような多層構造を用いてシミュレーションを行う。しかし、それらのシミュレーション結果からADALINEのウェイトの初期値などにより、次の三つの場合（ケース1～3）が生じることがわかった。

(1) ケース1 数回の適応で出力誤差が消える。本論文でのシミュレーションで最も少ない適応回

図3 XORのシミュレーションによるビット誤り

Fig. 3. Bit error result of XOR by simulation.

数は3回であった。

(2) ケース2 一定の出力誤差が残る。本論文でのシミュレーションでは、ウェイトが収束した後の出力のハミング距離は2となった。

(3) ケース3 図3で示すように出力誤差が周期的に変化する。

従って、もしADALINEの出力を反転させる組合せをうまく選べばMR IIアルゴリズムは効果的に働くが、そうでなければケース2、ケース3のような現象が生じることがある。パターン認識において入力層のユニット数は多くなるため、多数のユニットをもったニューラルネットの出力を適応させるためには、多くの組合せを選ばなければならず、最適な組合せを選ぶことが困難な場合が多い。Widrowらのアルゴリズムがこのように複雑なものになったのは、ADALINEの出力が±1のような二値信号であり、出力関数が微分不可能なシグナム関数を使用している点にある。特に、MR IIルールではシグナム関数の性質からこう配法の原理が適用できず、最小かく乱原理を用いなければならない点はその複雑さの最大の原因である。この困難性を克服するために、ADALINEで用いたシグナム関数をシグモイド関数で置換したパターン認識システムを以下で提案する。特に、シグモイド関数の導入によりこの配法の適用が可能となるため、このパターン認識システムにおけるニューラルネットの学習は、バックプロパゲーションアルゴリズム⁽⁴⁾に基づいた構成となっている。なお、Widrowらはシグモイド関数の導入可能性についてふれているが⁽²⁾⁽⁵⁾、具体的なアルゴリズムの導出は行っていない。

3. バックプロパゲーションアルゴリズム

ニューラルエレメント（以下、NEと略記）が図4に示すようにモデル化される。

図2のADの代わりに、図4のNEを用い、これ

図4 ニューラルエレメント (NE)
Fig. 4. Neural element (NE).

図5 バックプロパゲーションによるXORの誤差

Fig. 5. Output error of XOR by back-propagation.

を基にニューラルネットワークを構成する。このネットワークの学習則としてバックプロパゲーションアルゴリズム⁽⁴⁾を用いる。

バックプロパゲーションアルゴリズムは、以下の関係式によって与えられる⁽⁴⁾。

$$\delta_j^M = (y_j - o_j^M) f_i(i_j^M)$$

$$\delta_j^k = f_i(i_j^k) \sum_m w(t)_{j,m}^{k,k+1} \delta_m^{k+1}$$

$$\Delta w(t)_{i,j}^{k-1,k} = \alpha \delta_j^k o_j^{k-1} + \eta \Delta w(t-1)_{i,j}^{k-1,k}$$

$$f_i(i_j^k) = 1 / (1 + \exp(-i_j^k))$$

ここで、 $0 \leq f_i(i_j^k) \leq 1$ 、 $w(t)_{i,j}^{k-1,k}$: 第 $k-1$ 層の i ユニットから第 k 層の j ユニットへの重み関数、 $\Delta w(t)_{i,j}^{k-1,k}$: $w(t)_{i,j}^{k-1,k}$ の増分、 $\Delta w(t-1)_{i,j}^{k-1,k}$: 過去のウェイトの増分、 o_j^k : 第 k 層の j ユニットからの出力、 i_j^k : 第 k 層の j ユニットでの入力と重み関数の積和、 δ_j^k : 一般化された誤差、 f_i : f_i の微分

ここで、 $o_j^k = f_i(i_j^k)$ より

$$f_i'(i_j^k) = o_j^k (1 - o_j^k)$$

である。図5はバックプロパゲーションアルゴリズムによるXOR問題の誤差の収束の結果を表している。

ここでは、2層のニューラルネットワークが使用さ

れている。図5において入力は、(0,0), (0,1), (1,0), (1,1)で対応する出力は、それぞれ、0,1,1,0である。しかし、収束速度を上げるために、0と1の代わりに0.1と0.9が使用された。また、 $\alpha=0.25$, $\eta=0.9$ とし、これらはRumelhartらによって推奨された値である⁽⁴⁾。この場合、約400回の繰返しで出力誤差が消えている。ゆえに、バックプロパゲーションアルゴリズムを使用することによって多層ニューラルネットワークに対する重み関数を適切に決定することができる。

4. 変換に対して不変なパターン認識システム

本章では、変換に対して不変なパターン認識アルゴリズムを新たに提案する。図6はWidrowらによる変換に対して不変なパターン認識システムを表す⁽²⁾。同図でインバリانسネットワークは、ADALINEやMAJを使って位置ずれ、回転、スケール変換、視点変換などに対して不変なスクランブルパターンを生成する。ディスクランブラネットワークは、MR II アルゴリズムに基づいたパターン認識の主要な部分である。

このMR IIルールを用いて幾つかのパターン認識のシミュレーションを行ったが、満足な結果は得られなかった。その主要な問題は以下に要約されている。

(1) MR IIアルゴリズムは、パターン認識の場

合、パターンの画素の増加と共に、符号を反転させるADの組合せ数が大きくなり、計算時間を増大させると共に適応手順が複雑になる。

(2) ADALINEやMAJ関数中で使われるシグナム関数はステップ関数であり、雑音に対して敏感である。そのため、ADALINEの二値出力も、それに対応して複雑に変動する。従ってスクランブラの出力である中間表現の画像は、雑音に対してロバストなものとなっていない。

この問題を克服し、バックプロパゲーションアルゴリズムを適用可能とするために、ニューロン中の非線形関数として、しきい値素子であるシグナム関数の代わりに、滑らかな特性をもつシグモイド関数を用いる。まず、インバリانسネットワーク中のADALINEの代わりに図7で示されるNEを採用する。ここで、非線形関数 f_2 は以下で示されるものを用いる。

$$f_2(x) = 2 / \{1 + \exp(-x)\} - 1$$

なお、 $f_1(x)$ と $f_2(x)$ の差異は、前者が出力を $[0, 1]$ としているのに対して、後者は多数決関数(MAJ素子)を効果的に用いるために $[-1, 1]$ に変更している点である。以下では、Widrowらのインバリانسネットワークの構成法と同様な位置ずれ、回転に対して不変なインバリانسネットワークを構成する。図6の中で、パターン画像は 8×8 ピクセルからなるとする。そのため、 8×8 個のNEをもったスラブと呼ばれる

図6 Widrowらによる変換に不変なパターン認識システム

Fig. 6. Transform invariant pattern recognition system by Widrow.

図7 インバリانسネットワーク中のNEの構成図

Fig. 7. Structure of NE in invariance network.

図8 スラブの構成図

Fig. 8. Structure of slab.

基本構成要素を作る。スラブの構成図が図8で示される。ここで示されるように、MAJ素子の中においても非線形関数として、シグモイド関数 f_2 を用いる。まず、乱数を使ってNEの 8×8 のウェイト行列 W_1 を決定する。スラブ中で左上隅のNEのウェイトの行列が W_1 、そして一つ下のNEのウェイトの行列を $T_{d1}(W_1)$ とする。ここで、 T_{d1} は一つ下にずらす変換を表す。このウェイトの集合は、全体的に1ピクセル下にずらされたことを除いて、最も上のウェイトと同じである。この場合、最下行はウェイト行列の第一行にもってくる。2行、3行、 $\dots$ 、7行ずらす操作を $T_{d2}, T_{d3}, \dots, T_{d7}$ で、それぞれ表現する。この変換を使って生成されたNEのウェイト行列は以下で表される⁽²⁾⁽³⁾。

$$\begin{bmatrix} W_1 \\ T_{d1}(W_1) \\ \vdots \\ T_{d7}(W_1) \end{bmatrix}$$

次に、これらのウェイトに対して右にずらす変換を $T_{r1}, T_{r2}, \dots, T_{r7}$ と記述する。ここで、 T_{ri} は i 列だけ右にずらす変換を表す。すなわち、一番目のスラブ中のウェイト行列は以下で表される⁽²⁾⁽³⁾。

$$\begin{bmatrix} W_1 & T_{r1}(W_1) & \dots & T_{r7}(W_1) \\ T_{d1}(W_1) & T_{r1}T_{d1}(W_1) & \dots & T_{r7}T_{d1}(W_1) \\ \vdots & \vdots & & \vdots \\ T_{d7}(W_1) & T_{r1}T_{d7}(W_1) & \dots & T_{r7}T_{d7}(W_1) \end{bmatrix}$$

そして異なる W_1 から W_{64} のNEのウェイト行列に対して同様な変換を行う。こうして、インバリانسネットワークから64ピクセルのスクランブルされたパターンを得る。

次に、Widrowらの手法⁽²⁾⁽³⁾と同様に、回転に対して不変なインバリانسネットワークを考える。まず、4個のスラブのセットを用意する。ここで、4個のスラブは 90° だけ4回転することを意味する。回転に対して不変のために、2番目のスラブのウェイトは 90° だけ時計回りに回転していることを除くと、1番目と同じである。これは、 $R_{90}(W_1)$ で表現される。左上隅のNEのウェイトの行列は、1番目のスラブで W_1 で始まり、2番目のスラブで 90° 時計回りに回転し、3番目のスラブで 180° 、4番目のスラブで 270° 回転する。従って、2番目のスラブは、

$$\begin{bmatrix} R_{90}(W_1) & T_{r1}R_{90}(W_1) \\ T_{d1}R_{90}(W_1) & T_{r1}T_{d1}R_{90}(W_1) \\ \vdots & \vdots \\ T_{d7}R_{90}(W_1) & T_{r1}T_{d7}R_{90}(W_1) \end{bmatrix} *$$

図9 スクランブルパターンの一画素を生成するための構成図

Fig. 9. Structure for producing a pixel in scrambled pattern.

$$* \begin{bmatrix} T_{r7}R_{90}(W_1) \\ T_{r7}T_{d1}R_{90}(W_1) \\ \vdots \\ T_{r7}T_{d7}R_{90}(W_1) \end{bmatrix}$$

4個のスラブからの出力はすべてMAJ素子を通して、左上隅のピクセルの値として格納される。同様に、 $W_1, W_2, \dots, W_{64}$ のウェイトを使って64ピクセルのスクランブルされた画像データを得ることができる。スクランブルされたパターンの一画素を生成するための構成図が図9で示される。

スクランブルされた画像は、バックプロパゲーションアルゴリズムに基づいたディスクランブラネットワークを経由し、標準パターンへ変換される。このようなパターン認識システムの全体が図10で示される。

これらの構成図からわかるように、Widrowらのシステムでは ± 1 の2値が使用されていたが、このシステム中ではアナログ値が使用されている。また、ディスクランブラネットワークでは、MR II アルゴリズムの代わりにバックプロパゲーションアルゴリズムが使用され、スクランブラネットワークからの出力を標準入力パターンに分類するように学習される。本論文で提案したアルゴリズムの有効性を、シミュレーションにより以下で定量的に検証する。

5. シミュレーション結果

標準入力パターンは 8×8 画素のE, I, L, Tで、図11で示される。

目標パターンは、標準入力パターンとする。これらのパターンによってネットを学習させた後で、位置ず

図 10 提案される変換不変パターン認識システム

Fig. 10. Proposed transform invariant pattern recognition system.

れなどの変換を受けたパターンを標準パターンへ分類する。比較のために三つの場合を考える。

(1) ケース 1 インバリアンスネットワークなしで、バックプロパゲーションアルゴリズムだけに基いた場合。

(2) ケース 2 インバリアンスネットワークとディスクランブラネットワークが存在し、インバリアンスネットワークの部分は ADALINE 素子からなり、ディスクランブラネットワークはバックプロパゲーションアルゴリズムを採用した場合。

(3) ケース 3 図 10 で示される、本論文で提案したパターン認識システムを用いる場合。

ディスクランブラネットワークは、64 入力 64 出力の 3 層ニューラルネットワークで、 $\alpha=0.25$, $\eta=0.9$, 学習回数は 10,000 回である。図 12 は入力パターンで、(a) は位置ずれ、回転の変換を加えたパターン、(b) は孤立点と雑音を含んだパターンである。図 13 はケース 1 の分類結果を示す。図 14 はケース 2 の分類結果を示す。図 15 はケース 3 の分類結果を示す。

これらの結果より、ケース 1 とケース 2 による分類結果には、幾つかのエラーが見られる。ケース 3 による分類結果からわかるように、提案されたパターン認識システムは、位置ずれ、回転、雑音、孤立点の影響をあまり受けにくいロバストなアルゴリズムであること

図 11 標準入力パターン
Fig. 11. Standard input pattern.

図 12 入力パターン
Fig. 12. Input pattern.

図 13 ケース 1 の分類結果
Fig. 13. Result of classification for case 1.

図 14 ケース 2 の分類結果
Fig. 14. Result of classification for case 2.

が観察される。

次に、標準パターン数を増加した場合にも本論文で提案した手法が有効であることを示すために、図 16 に示す 10 個の標準パターンに対するパターン認識について考察する。図 17 は入力パターンを示す。図 18 に示すように 10,796 回の学習後 10^{-3} 以下に出力誤差が減少できることがわかった。この場合の分類結果を図 19 に示す。図 19 からわかるように、図 17 の入力

図 15 ケース 3 の分類結果
Fig. 15. Result of classification for case 3.

図 16 10 個の入力パターン
Fig. 16. Input patterns of ten pieces.

図 17 パターン認識用入力パターン
Fig. 17. Input patterns for pattern recognition.

図 18 学習誤差曲線
Fig. 18. Learning error curve.

図 19 パターン認識結果
Fig. 19. Pattern recognition results.

パターンに対して完全なパターン認識結果が本論文の手法で得られることが示されている。

6. 結 論

位置ずれ、回転、雑音、孤立点の影響を受けにくい新しいニューラルネットワークを提案した。これらはバックプロパゲーションアルゴリズムのみの場合や、Widrow らによって提案されたパターン認識システムより有効であることがわかった。今後、スケール変換や視点変換がなされたパターン認識についても考察されるならば、実際の文字認識などへの応用も可能となるものと思われる。(平成元年3月8日受付)

文 献

- (1) K. Fukushima, S. Miyake & T. Ito: "Neocognitron: A

Neural Network Model for a Mechanism of Visual Pattern Recognition", *IEEE Trans. Systems, Man, & Cybernetics*, SMC-13, 826 (1983)

- (2) B. Widrow, R. G. Winter & R. A. Baxter: "Layered Neural Nets for Pattern Recognition", *IEEE Trans. Acoustic, Speech & Signal Processing*, ASSP-36, 1109 (1988)
(3) C. R. Winter & B. Widrow: "MADALINE RULE II: A Training Algorithm for Neural Networks", Second Annual International Conference on Neural Networks, i-401 (1988)
(4) D. E. Rumelhart, G. E. Hinton & R. J. Williams: "Learning Internal Representation by Error Propagation" In: Rumelhart, D. E. and J. L. McClelland (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations* (1986) Cambridge, MIT Press
(5) B. Widrow: "LMS Techniques in Neural Networks", Tutorial No. 13, International Joint Conference on Neural Networks, Washington, D. C. (1989)

福 見 稔 (正員)

昭和 36 年 9 月 21 日生。62 年 3 月徳島大学大学院工学研究科情報工学専攻修士課程修了。同年 4 月同情報工学科助手、現在に至る。適応信号処理、神経回路網による情報処理の研究に従事。IEEE, INNS, 電子情報通信学会などの会員。

細 川 直 史 (非会員)

昭和 39 年 8 月 25 日生。平成元年 3 月徳島大学大学院修士課程修了。在学中はニューラルネットワークの画像処理に関する研究に従事。現在、徳島県警勤務パターン認識の応用に関する研究に興味をもっている。

大 松 繁 (正員)

昭和 21 年 12 月 16 日生。44 年 3 月愛媛大学工学部電気工学科卒。46 年 3 月大阪府立大学大学院工学研究科修士課程修了。49 年 3 月同博士課程修了。49 年 6 月徳島大学工学部情報工学科助手、50 年 4 月同講師、55 年 3 月同助教授、63 年 6 月同知能情報工学科教授となり、現在に至る。分布定数系の推定制御、信号処理、画像処理、ニューラルネットワークの研究に従事。計測自動制御学会、IEEE 会員。