

逆想起ニューラルネットモデルを用いた フィードバック型認識方法

山 田 敬 嗣[†] 津 雲 淳[†]

フィードバック制御機能をもつ高精度なパターン認識を、逆想起ニューラルネットモデルとそれに基づくパラメータ調節法によって実現する方法を提案する。フィードフォワード処理としては、従来の特徴抽出処理の後に、その特徴ベクトルを入力して、逆想起ニューラルネットモデルのフォワード処理により識別を行う。識別結果より、認識結果の不確信度を計算し、その値がしきい値以上ならば、フィードバックにより特徴抽出での処理パラメータを調節する。まず、ニューラルネットの逆想起処理を起動して、不確信度を小さくするために必要な入力特徴ベクトルの変更量を求める。さらに、特徴ベクトルの変更量をもとに、特徴抽出のためのパラメータ値を変更し、再度特徴抽出、識別を行う。不確信度がしきい値以下になるまで、この処理を繰り返し、特徴抽出パラメータを調節する。これにより、個々のパターンごとに最適に特徴抽出パラメータを調節し、識別に有効な特徴を抽出できるため、認識精度を向上できる。この逆想起ニューラルネットモデルを用いたフィードバック型認識方法を、手書き英数字認識に適用し、同一誤認識率で、フィードバックを用いない場合に比べ、リジェクト率を半減できた。

Feedback Pattern Recognition Method Using Inverse Recall Neural Network Model

KEIJI YAMADA[†] and JUN TSUKUMO[†]

A feedback pattern recognition method based on an inverse recall neural network model is proposed. The inverse recall model is trained by the same method as that used for a typical multi-layer feedforward model. The model can produce an inverse mapping of the trained feedforward mappings to show the parts of an input pattern which are important for recognition and are not presented to the network. The model is applied to the feedback pattern recognition method which can extract features from input patterns and discriminates between them by the neural network model. The feedback recognition method adjusts feature extraction parameters so as to detect the important parts shown by the network model, to present them to the network model, and to generate more certain recognition results. This method was examined on handwritten alpha-numeric recognition and it was found that rejection ratio can be reduced by half at the same error ratio.

1. はじめに

パターン認識の入力、前処理、特徴抽出、識別等の各処理段階では、明示的にまたは暗黙に処理パラメータが用いられ、あらかじめ適切な値に設定されて利用されている。しかし、あらかじめ設定された1通りのパラメータ値を用いるだけでは、実際上の大きなパターンの変形を吸収して正しく認識を実行できないことがある。

予備的な手書き数字認識実験の結果でも、方向と曲率に基づく特徴を抽出し、多層のニューラルネットに

より非線形な識別関数を実現した場合には、線形識別関数を用いた場合よりも、誤認識率を3分の1に減少できた¹⁾。しかしながら、残る誤認識パターンの中には、人間には容易に認識可能であるものの、特徴抽出処理の段階で処理パラメータ値の設定が不適切であったために、識別に有効な特徴が抽出できず、誤認識となったものがある。

この問題を解決するためには、膨大にあるパラメータ値の組み合わせの中から、認識対象のおおのこのパターンに対して、最適な処理パラメータ値の組み合わせを探索しなければならず、試行錯誤的な山登り法などでは処理効率が悪い。試行錯誤的な探索を避け、効率よく探索するために、決定論的弛緩法などの最適値探索法が利用できる^{2), 3)}。決定論的弛緩法では、探索

[†] 日本電気(株)情報メディア研究所
Information Technology Research Laboratories,
NEC Corporation

規範値 R を探索空間の要素 p_i で偏微分した値を $\partial R/\partial p_i=0$ とする p_i の値を、反復法により求めるが、一般的に処理パラメータの関数としての認識結果は不連続で、微分可能ではない。そのため、決定論的な弛緩法を直接利用して、処理パラメータの調節を実現することができない。

本論文では、従来の特徴抽出と多層ニューラルネットによる識別処理を組み合わせたパターン認識方式において、より信頼できる認識結果が得られるように、特徴抽出の処理パラメータを適応的に調節できるフィードバック型の認識方法を提案する。まず、認識結果の信頼性の尺度である不確信度を定義し、識別部のフィードバックを実現する逆起ニューラルネットモデルを提案し、特徴抽出部へのフィードバックと特徴抽出パラメータの調節方法について述べる。最後に、手書き英数字認識実験結果をもとに、その性能を考察する。

2. フィードバック型認識

方向と曲率に関する特徴を用いた手書き数字認識実験の結果では、線形識別関数による認識率は 97.2% であったのに対して、多層ニューラルネットによる認識率は 99.1% であった¹⁾。また、濃淡の文字パターンを入力として、局所結合型の多層ニューラルネットを用いた場合、通常的全結合型の多層ニューラルネットに比べて認識率は向上した^{4),5)}。しかし、この場合には、最高でも認識率は 98.9% であり、方向曲率特徴を入力した場合よりも低かった。これより、従来の特徴抽出法により変動に強い特徴を抽出して多層ニューラルネットで識別すれば、高い認識率が得られることがわかる。しかし、誤認識パターンを調べると、あらかじめ設定された特徴抽出パラメータが不適切で識別に有効な特徴を抽出できていない場合がある。

すべての文字パターンの変形を吸収できるような特徴抽出法を考案することにより、この問題は解決できるが、これは非常に困難である。むしろ、図 1 の概念図に示すように、フィードバック制御を導入し、識別に有効な特徴が抽出できるように、個々の文字パターンに適応的に特徴抽出パラメータ値を調節すればよい。

図 2 を例に考えると、ある一定の触手長のベクトルにより疑似的な接線方向 ϕ と曲率 ϕ を求めて特徴とする場合に、小さな凹凸に影響されない程度に長い触手を利用する必要がある。しかし、図 2 左のパターンのように長い触手では識別に必要な凹部の特徴を検出

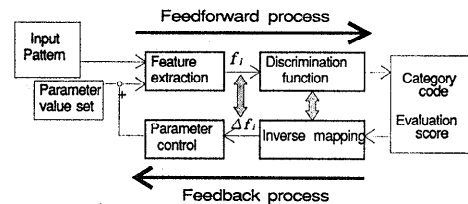


図 1 フィードバック型パターン認識の概念図
Fig. 1 Conceptual diagram of a feedback pattern recognition.

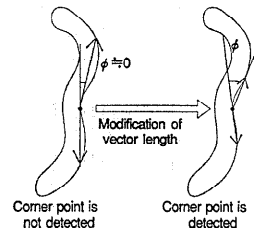


図 2 フィードバック型認識での触手長制御による凹部検出の例
Fig. 2 Concavity detection by adjusting vector length in a feedback pattern recognition method.

できず、誤認識を起こす。このような場合には、一時的に触手長を短くして凹部の特徴を検出しなければならない。同時に疑似曲率 ϕ から凹部を判定するしきい値も同時に制御することが望まれる。ここで提案するフィードバック型認識方法は、このようなしきい値や触手長などの処理パラメータを認識対象のパターンに合わせて自動的に調節することを目的とする。

このフィードバック型の認識法は、入力パターンを固定した上で認識結果を一種の評価値とし、 n 個の特徴抽出パラメータ値を探索対象とする、 n 次元空間での最適値探索とみなせる。これを実現するには、

1. 探索のための評価値を明確に定義すること
 2. 無意味な試行を避け、効率よく探索すること
- の 2 点が重要である。第 1 項の評価値をパターン認識に対して定義するために、認識結果の確信度または不確信度 R を定める。つまり、より信頼できる認識結果を得るようにパラメータ値を探索する。

第 2 項に関しては、図 1 に示すようなフィードバック方法で、最急勾配法を応用して、解決を試みる。入力パターンを P とし、特徴抽出パラメータの組を $\Theta = \{\theta_i\}$ として、抽出された特徴を $F = \{f_i(P, \Theta)\}$ で表す。さらに、識別関数の出力を $D(F)$ とし、認識結果の不確信度を $R(F)$ とする。 P と Θ の関数としての F と R がともに、 Θ の探索領域内で連続で偏微分

可能な場合には, θ_k の変更量 $\Delta\theta_k$ は,

$$\Delta\theta_k = -\alpha \frac{\partial R(F)}{\partial \theta_k} = -\alpha \sum_i \frac{\partial R(F)}{\partial f_i} \frac{\partial f_i}{\partial \theta_k} \quad (1)$$

で求められる。線形識別関数や多層ニューラルネットによる識別では, 微分可能な識別関数と不確信度を設計でき, 式(1)の $\partial R/\partial f_i$ (図1中の Δf_i) は, 容易に計算できる。しかし, 特徴抽出 $F(P, \theta)$ は, 一般には微分可能ではない。

$-\alpha \partial R/\partial f_i$ は, 不確信度 R を小さくするために変更すべき特徴量のステップで, $f_i' = f_i - \alpha \partial R/\partial f_i$ を得て, 再び $\{f_i'\}$ を識別関数に入力すれば, 式(1)の $\partial f_i/\partial \theta_k$ を計算することなく, より不確信度の小さい認識結果を得ることができる。しかし, このように特徴の変更量を求めることは, 識別関数の逆写像を求めることと同じで, 一種の不良設定問題である。よって, 変更された特徴量 f_i' は, 入力パターン P から実際に求め得る特徴とは限らず, 不当な誤認識結果を与えることがある。これを解決するために, あらかじめ $\partial f_i/\partial \theta_k$ をいくつかの場合に分けて近似的に求めておき, それを用いて式(1)を計算して, 特徴抽出パラメータを変更した後に, 再度特徴抽出, 識別を実行する。これにより得られた特徴量 f_i'' は, 入力パターンから実際に求め得る特徴であり, 特徴抽出パラメータを過度に変更しない限り, 前述のような不当な誤認識を起こさない。

この近似的に解法を実現できるのは, (a)特徴抽出パラメータ調節のための探索は, 大域的な探索を必要としないこと, (b)特徴量の値域は広く, 制約を加えにくい, 求める特徴が得られる特徴抽出パラメータは, 初期値に近くその探索範囲は限定できることによる。これらの条件は, 勾配法を用いて最適値探索を行う上での条件である。予備実験から, 微妙な変形のためにリジェクトされているパターンが多く存在することがわかっており, このようなリジェクトパターンは, 上の条件を満たすと見なすことができる。このようなパターンの頻度は, 提案方法の性能に多大な影響を与えるので, 認識実験の結果をもとに考察する。

以降に具体的なフィードバック型認識方法について述べる。

3. 逆想型ニューラルネットモデルによる フィードバック型認識

3.1 識別処理と不確信度

本節では, 図1のフィードバック型認識において,

識別関数とフィードバック時の逆写像関数を実現する逆想起ニューラルネットモデルについて述べる。ここでは, 多層ニューラルネットによる識別処理を述べた後に, 前節の2つの要件のうち, 第1項目の探索評価値を認識結果の信頼性を表す不確信度として定義する。

多層ニューラルネットにおける非線形関数として,

$$f(x) = \frac{1}{2}(1 + \tanh x) \quad (2)$$

を用いる。第 j 層の第 i ユニットの値を q_{ji} ($i=1 \sim N_j$) とし, N_j は第 j 層のユニット数を表す。このとき, 第 $j+1$ 層のユニット値は, 下式で計算する。

$$q_{(j+1)k} = f\left(\sum_{i=1}^{N_j} w_{jki} q_{ji} + b_{jk}\right) \quad (3)$$

入力層へは特徴抽出後の N_0 次元の特徴ベクトルを入力し, 上式に従って出力層 (第 M 層) までのユニット値を順次計算する。ただし, すでに重み値 w_{jki} とバイアス値 b_{jk} の学習は終了しているものとする。

出力層の第 k 番目のユニットの値は, 第 k カテゴリへの帰属の程度を表す。この場合, 出力値は0から1の値を取り, 1の場合には第 k カテゴリに属し, 0の場合には第 k カテゴリに属さないことを示す。これより, 唯一のユニットが値1をとり, 他のすべてのユニット値が0の場合には, 認識結果は明確である (不確信度規範1)。しかし, 最大ユニット値が低下したり, それ以外のユニット値が増加すると, 認識結果はより不明確になる (不確信度規範2)。これを評価するために, 不確信度 R を以下のように定義した。

$$R = \sum_{k=1}^{N_M} q_{Mk}(1 - q_{Mk}) + \beta \left(\sum_{k=1}^{N_M} q_{Mk} - 1 \right)^2 \quad (4)$$

前述の2つの不確信度規範を満足するために, β の値は0.5に定めた⁶⁾。不確信度が0.5を超える場合には, 無条件にリジェクトされる。初期的に, またはフィードバックにより不確信度 R が, しきい値 δ 以下になった場合には, 認識結果が受理される。

3.2 逆想起ニューラルネットモデル

認識処理の不確信度 R が, しきい値 δ と0.5の間にある場合, フィードバック処理が起動され, 特徴抽出パラメータの最適値探索が実行される。図1の逆写像を実行して, 不確信度を評価値とするフィードバック処理のために, すでに提案した逆想起ニューラルネットモデルを用いた⁷⁾。

これは, 学習済みの多層ニューラルネットを用い, 入力パターン (この場合は特徴ベクトル) から出力値を式(3)によって求め, さらに式(4)より不確信度

R を求める. R が一定しき値 δ より大きく, 入力特徴ベクトルが受理されない場合には, 不確信度を小さくするように入力特徴ベクトルの変更を行う. 3 層のニューラルネットを仮定し, 入力層を第 0 層とし, $f_i = q_{0i}$ として, 出力層を第 2 層とした場合には, 特徴量の変更量 Δf_i は, 下式で求められる.

$$\Delta f_i = -\alpha \frac{\partial R}{\partial f_i} = -\alpha \sum_{j=1}^{N_1} \left(\sum_{k=1}^{N_2} \frac{\partial R}{\partial q_{2k}} \frac{\partial q_{2k}}{\partial q_{1j}} \right) \frac{\partial q_{1j}}{\partial f_i} \quad (5)$$

$$\frac{\partial R}{\partial q_{2k}} = 1 - 2q_{2k} + 2\beta \left(\sum_{n=1}^{N_2} q_{2n} - 1 \right)$$

$$\frac{\partial q_{2k}}{\partial q_{1j}} = \frac{2}{u_0} 2q_{2k}(1 - q_{2k})w_{1kj}$$

$$\frac{\partial q_{1j}}{\partial f_i} = \frac{2}{u_0} q_{1j}(1 - q_{1j})w_{0ji}$$

この変更量を入力層の元の特徴ベクトルに足し込むことにより, 少し小さい不確信度を与える特徴ベクトル $f_i(t)$ を入力層に生成することができる.

$$f_i(t) = f_i(t-1) + \Delta f_i(t-1) \quad (6)$$

ここで, $f_i(t)$ は t 回目のフィードバック後の特徴ベクトルを表す. このニューラルネットモデルは, 識別関数の逆写像を実現するものであり, 逆想起ニューラルネットモデルと呼ぶ. 前述したように, このニューラルネットが与える $\Delta f_i(t-1)$ は, 識別関数に関わりの少ない部分空間 (補空間と呼ぶ) の要素をも含むため, (6)式から求めた $f_i(t)$ は, 必ずしも実際に求め得る特徴になるとは限らない. 例えば, 実際のあるカテゴリのデータでは値を持つはずのない要素 i で, $\Delta f_i(t-1)$ が比較的大きな正値を持つことがある. 線形識別においては, その直交補空間は識別に影響を与えないが, ニューラルネットの場合には, その補空間が影響を与えない保証はない. さらに, (6)式の計算を繰り返し実行した場合に, 識別結果に影響を与えるノイズが発生することが実験で確認されている⁷⁾. そこで, 以降に $\{\Delta f_i(t-1)\}$ のうち, 前述の補空間に関わる要素の影響を抑制して, それを元に実際に求められる $\{f_i(t)\}$ を得る方法を提案する.

3.3 特徴抽出パラメータの変更

逆想起により得た特徴ベクトルの変更量 Δf_i から, 式(1)を近似的に実現するように, 特徴抽出のパラメータ値の変更量を求める関数 $G(\{\Delta f_i\})$ を定める. 処理の簡易化のために, この関数の値はあらかじめ定めた一定の変更量の増減の符号を与えるものとする. これにより変更されたパラメータを用いて, 入力パターンから再度特徴抽出をし, 実際に求め得る特徴を得て, それを用いて識別を実行する.

基本的に $G(\{\Delta f_i\})$ の学習では, $\{\Delta \theta_m\}$ に影響を与えない $\{\Delta f_i\}$ の要素を統計的に求める. これにより, 安定な値を持つ要素のみならず, 実際の特徴値として現れにくい要素を検出し, 認識処理での特徴抽出パラメータの調節時に, その要素の影響を抑制する. 特徴抽出パラメータの変動に対して不安定な特徴の要素は各パターンによって異なるが, カテゴリごとに求めることにより, 不安定な特徴要素を限定できる. これは, 各カテゴリは, 現れ得ない特徴要素や安定して現れる特徴要素をもつこと, さらにこれらの特徴要素はカテゴリごとに異なることから仮定できる. それゆえ, ここではカテゴリごとに $G(\{\Delta f_i\})$ を生成する場合について説明する. 詳細な処理を以下に述べる.

(準備・関数 $G(\{\Delta f_i\})$ の学習)

1. 学習データセットの中の第 k カテゴリ C_k の第 j データに対して, 初期的な特徴抽出パラメータ $\{\theta_m\}$ を用いて特徴抽出し, $\{f_i(j, k)\}$ を得る. ここで, i は特徴ベクトルの要素の番号を表す.
2. 同一データに対して, 初期的なパラメータ群の内の第 m 番目の θ_m にのみ一定量 $\Delta \theta_m$ を加えた θ_m' を用いて, 特徴抽出をして $\{f_i'(j, k)\}$ を得る. また, 一定量 $\Delta \theta_m$ を減じた θ_m'' を用いて, 特徴抽出をして $\{f_i''(j, k)\}$ を得る.
3. 第 k カテゴリ C_k の全データから下式を計算する.

$$\Delta F_i(k, m) = \sum_{j \in C_k} (f_i'(j, k) - f_i(j, k)) \quad (7)$$

$$\Delta F_i'(k, m) = \sum_{j \in C_k} (f_i''(j, k) - f_i(j, k))$$

4. $\Delta F_i(k, m)$ の符号により, 特徴ベクトル内の要素を部分集合 $F_{++}(k, m)$ と $F_{+-}(k, m)$ に分類する.

$$F_{++}(k, m) = \{i | \Delta F_i(k, m) > 0\} \quad (8)$$

$$F_{+-}(k, m) = \{i | \Delta F_i(k, m) < 0\}$$

5. $\Delta F_i'(k, m)$ の符号により, 特徴ベクトル内の要素を部分集合 $F_{-+}(k, m)$ と $F_{--}(k, m)$ に分類する.

$$F_{-+}(k, m) = \{i | \Delta F_i'(k, m) > 0\} \quad (9)$$

$$F_{--}(k, m) = \{i | \Delta F_i'(k, m) < 0\}$$

6. 以上の処理をすべてのカテゴリ C_k とすべてのパラメータ θ_m に対して行い, 特徴ベクトルの要素の部分集合 $F_{++}(k, m)$, $F_{+-}(k, m)$, $F_{-+}(k, m)$, $F_{--}(k, m)$ を求める.
7. 第 m 番目の特徴抽出パラメータに関して, $F_{ss}(i, m)$, $ss \in \{++, +-, -+, --\}$ の集合と, 認識時に得られる特徴の変更量 $\{\Delta f_i\}$ を用いて

パラメータ変更による影響度を以下のように求める。

$$M(k, m), \{\Delta f_i\} = \sum_{i \in F_{++}(k, m)} \Delta f_i - \sum_{i \in F_{+-}(k, m)} \Delta f_i - \sum_{i \in F_{-+}(k, m)} \Delta f_i + \sum_{i \in F_{--}(k, m)} \Delta f_i \quad (10)$$

を用いて、 $G(k, m, \{\Delta f_i\})$ を求める。ただし、 $\Delta \theta_m$ は、パラメータごとにあらかじめ定めた一定値とする。

$$G(k, m, \{\Delta f_i\}) = \Delta \theta_m \quad \text{if } M(k, m, \{\Delta f_i\}) \geq 0 \\ = -\Delta \theta_m \quad \text{otherwise} \quad (11)$$

(認識処理)

1. 初期的なパラメータ $\{\theta_m\}$ を用いて、入力パターンから特徴抽出を行い、逆想起ニューラルネットに入力し、認識結果としてカテゴリ番号 k と式(4)に従って不確信度 R を求める。
2. 不確信度 R がしきい値 δ より小さい場合は、カテゴリ番号 k を認識結果として受理する。 R が 0.5 より大きい場合は、無条件にリジェクトする。もし、 R が δ より大きく、0.5 より小さい場合は、以下のフィードバック処理を行う。
3. ニューラルネットの逆想起処理を行い、式(5)に従って特徴量の変更量 $\{\Delta f_i\}$ を求める。
4. カテゴリ番号 k と $\{\Delta f_i\}$ から、第 m 番目の特徴抽出パラメータを下式に従って変更する。
$$\theta_m \leftarrow \theta_m + G(k, m, \{\Delta f_i\}) \quad (12)$$
5. すべての特徴抽出パラメータ θ_m に対して、4. の変更処理を行う。
6. 再度、入力パターンを用いて特徴抽出とニューラルネットによる識別を行い、不確信度 R がしきい値 δ 以下ならば、認識結果を受理し、認識処理を終了する。もし、しきい値より大きければ、3. から 5. のフィードバック処理を継続する。
7. 全パラメータの変更 (3. から 6.) を K 回行っても、認識結果が受理されない場合は、入力パターンを曖昧なパターンとしてリジェクトする。

4. 手書き英数字認識実験

電子技術総合研究所の手書き文字データベースのうちの ETL 6 データベースから英大文字と数字の計 36 カテゴリを選び出し、2 値化後に大面積の潰れなどを生じている文字パターンを除いた 48,621 データを生

成した。このデータベースのうち、奇数番目のデータである 24,311 データを学習に使い、残りの 24,310 データを認識実験に用いた。

特徴としては、図 3 に示す文字輪郭線の方角と曲率に基づく特徴を用いた。輪郭線上の 1 点 p_s から輪郭線沿いに長さ L の触手を両方向に延ばし、両端点間のベクトル $\overline{p_{s-L}p_{s+L}}$ の方向 ψ を点 p_s の疑似的な接線方向とする。また、2 本の触手 $\overline{p_{s-L}p_s}$, $\overline{p_s p_{s+L}}$ のなす角度 ϕ を疑似的な曲率とする。

これによって求められた方向を等間隔の 8 方向に分類し、曲率は 4 つのしきい値 ($\theta_1, \theta_2, \theta_3, \theta_4$) により、正曲率の屈曲点、正曲率の曲線分、直線分、負曲線の曲線分、負曲率の屈曲点の 5 種類に分類する。文字パターンを 4×4 の小領域に分割し、各小領域内での各方向、各曲率クラスに属する輪郭画素の数を計数し、特徴ベクトルの要素とする。それゆえ、 $4 \times 4 \times 8 \times 5$ の計 640 次元の特徴ベクトルを用いる。

フィードバック型認識のための特徴抽出パラメータと変更量は以下のとおり。

触手長 $L=5$

曲率分類しきい値 $(\theta_1, \theta_2, \theta_3, \theta_4) = (1.2, 0.2, -0.2, -1.2)$

触手長の変更量 $\Delta L=1$

曲率分類しきい値の変更量 $(\Delta \theta_1, \Delta \theta_2, \Delta \theta_3, \Delta \theta_4) = (0.12, 0.02, -0.02, -0.12)$

不確信度しきい値 $\delta \in [0.001, 0.5]$

フィードバック最大回数 $K=6$

フィードバックのない場合とフィードバックのある場合の、リジェクトしきい値を変化させて得た誤認識率対リジェクト率のグラフを図 4 に示す。図 4 中の単純山登り法とは、各特徴抽出パラメータを前述の一定値 $\Delta \theta_m$ ずつ増減させた場合と、変更しない場合の 3 つの場合を、全パラメータの組み合わせで試みる方法を指す。5 つのパラメータを調節する本実験では、3⁵ 通りの試行から得た不確信度を比較して、最小値を与

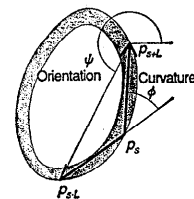


図 3 文字輪郭線の疑似接線方向と疑似曲率
Fig. 3 Pseudo tangent orientation and pseudo curvature of a character contour.

えるパラメータ値の組を、特徴抽出パラメータの変更値として選択する。同様に、不確信度がしきい値 δ 以下になるまで、すべてのパラメータ値を繰り返し変更する。この場合も、最大6回まで各パラメータを変更する。

フィードバックがない場合に比べ、提案したフィードバック法を用いた場合には、同一誤認識率の場合に、リジェクト率が減少している。例えば、誤認識率

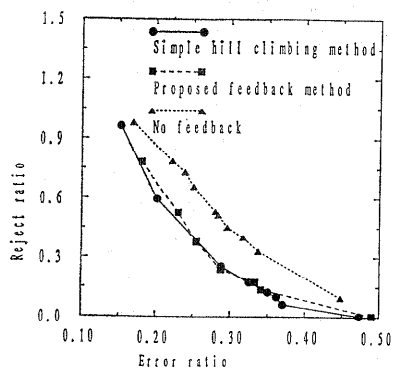


図4 フィードバックのある場合とない場合での誤認識率対リジェクト率と単純山登り法による誤認識率対リジェクト率

Fig. 4 Error ratio versus rejection ratio in the case when a feedback method is used, that in the case of a feedback method is not used, and that in the case of a simple hill climbing method.

約 0.34% の場合に、フィードバックがないときにはリジェクト率が 0.33% であるのに対して、提案したフィードバック法ではリジェクト率が約 0.14% である。同様に、誤認識率 0.25% の場合に、リジェクト率は、フィードバックがないときに 0.66% で、提案したフィードバック法では 0.38% である。全体的には、同一誤認識率でリジェクトを半減できている。これは、初期的なリジェクトパターンの中に、認識結果が曖昧であるものの、正解に近いパターンが多く含まれることを示している。このようなパターンは、2.3節で述べたフィードバック解法実現の条件(a), (b)を満たし、特徴抽出のパラメータを少々調節するだけで識別に有効な特徴を抽出できるので、リジェクトを正解にすることができる。

図5に、誤認識がフィードバックにより正しく認識されるようになったパターンの例を示す。ここでの誤認識とは、リジェクトパターンのうちで、誤ったカテゴリが第1候補となっているものを指す。'S', 'J', '8'などにみられるように微妙な特徴を捉えて、正しく認識されている。また、図6にフィードバックのために誤認識になったパターンを示す。'I'の場合のように、小さな特徴を捉えすぎて、誤認識となっているもの、'4'の場合のように、小さな特徴を無視するようにパラメータを調節したために、誤認識になっているものがある。これらは、局所的な位置によらず、全

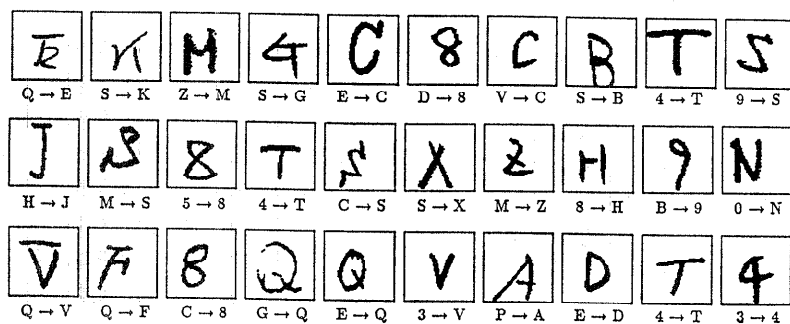


図5 フィードバックにより誤認識が正解になったパターンの例

Fig. 5 Pattern examples which are mis-recognized before feedback and correctly recognized by a feedback method.

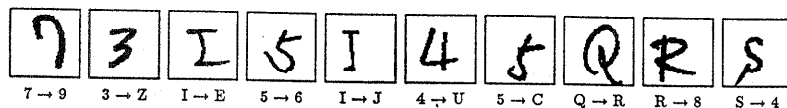


図6 フィードバックにより正解が誤認識になったパターンの例

Fig. 6 Pattern examples which are correctly recognized before feedback and mis-recognized by a feedback method.

体的にパラメータを調節しているため、局所位置ごとに特徴抽出パラメータを調節すれば、よりフィードバックの効果を高めることができると考えられる。

提案したフィードバック型認識法と単純山登り法とを比較すると、ほぼ同程度の誤認識率とリジェクト率を与える。ここでいう単純山登り法は、特徴抽出パラメータの初期値の周辺をすべて調べて、最良のパラメータ群を選択しながらパラメータを変更する方法であり、局所的な最適探索では、最良の結果を与える。しかし、本実験の場合では5次元のパラメータ空間中での探索であり、1回のパラメータ群の変更のために、3⁵回の特徴抽出と識別処理を実行する方法であり、効率は良くない。それに対して、提案したフィードバック法は、1回の識別と1回の逆想起処理と1回のパラメータ変更量の計算を実行するだけでよく、効率の良い最適な処理パラメータの探索を実現できる。

5. おわりに

本論文では、特徴抽出の後に多層ニューラルネットを識別関数に用いたパターン認識方法において、特徴抽出のための処理パラメータをフィードバックにより最適探索し、適応的に設定する方法を提案した。

まず、探索のための評価値として、多層ニューラルネットの出力値から求める認識結果の不確信度を定義し、それをより小さくするように入力特徴の変更量を求める多層ニューラルネットの逆想起モデルについて述べた。さらに、逆想起された特徴の変更量から、特徴抽出処理でのパラメータの変更量を求めて、特徴量の再抽出を行う。この処理を繰り返し行い、入力パターンに最適な特徴抽出パラメータを調節して、認識を実行する。この方法により、無駄な試行錯誤を避け、効率よく特徴抽出パラメータを調節できる。

この方法を用いて手書き英数字認識実験を行った結果、同一の誤認識率の場合にフィードバックがない場合に比べてリジェクト率を約半減できた。今後は、提案した処理パラメータの最適化法を応用して、文字パターン内で局所的に特徴抽出パラメータの最適化を試みるとともに、特徴抽出以外のスキミング等の処理段階のパラメータの最適化への利用を試みる。

参 考 文 献

- 1) Yamada, K., Tsukumo, J., Kami, H. and Temma, T.: Handwritten Numeral Recognition by Multi-layered Neural Network with Im-

proved Learning Algorithm, *Proc. of IJCNN* 89, pp. II259-II266 (1989).

- 2) Poggio, T., Torre, V. and Koch, C.: Computational Vision and Regularization Theory, *Nature*, Vol. 317, No. 6035, pp. 314-319 (1985).
- 3) 坂上勝彦: エネルギー最小化原理に基づく画像処理—弛緩法と正則化—, 電子情報通信学会技術研究会報告, IE 89-89 (1989).
- 4) 山田敬嗣, 上 博行: ニューロン回路—パターン認識, シミュレーション, Vol. 8, No. 1, pp. 17-20 (1989).
- 5) Le Cun, Y.: Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, Vol. 1, pp. 541-551 (1989).
- 6) Yamada, K.: Multi-font Alpha-numeric Recognition using Multi-layer Neural Networks with a Rejection Function, *Advances in Machine Vision, Strategies and Applications*, Archibald, C. and Petriu, E. (eds.), *World Scientific Series in Computer Science*, Vol. 32, pp. 211-227, World Scientific Publishing, Hong Kong (1992).
- 7) Yamada, K.: Learning of Category Boundaries Based on Inverse Recall by Multi-layer Neural Networks, *IJCNN-91-Seattle*, pp. II7-II12 (1991).

(平成4年12月8日受付)

(平成5年4月8日採録)



山田 敬嗣 (正会員)

昭和57年京都大学工学部情報卒業。昭和62年同大学院博士課程修了。同年日本電気(株)に勤務。現在、日本電気(株)情報メディア研究所主任。平成2～3年カリフォルニア大学サンディエゴ校客員研究員。パターン認識、ニューラルネットの研究に従事。昭和61年情報処理学会学術奨励賞、昭和62年画像電子学会論文賞受賞。電子情報通信学会など5学会各会員。



津雲 淳 (正会員)

昭和51年東京大学工学部計数工学科卒業。同年より日本電気(株)に勤務。現在、日本電気(株)情報メディア研究所研究課長。昭和60～61年Concordia大学(カナダ・モントリオール)客員研究員。文字認識、図形の重ね合わせ手法の研究に従事。電子情報通信学会、IEEE CS、AVIRG 各会員。