

サッケード機構を組み込んだ神経回路モデル
手書き文字の位置ずれに対する性能評価

5E-7

早野克己 鷲沢輝芳
(株)テレマティーク国際研究所

1. はじめに

現在、手書き文字認識に対するアプローチとしてバックプロパゲーション法（以下単にBPと呼ぶ）を用いたニューラルネットワークが検討されている⁽¹⁾⁽²⁾⁽³⁾。しかし、手書き文字では、文字列を構成する文字が学習した文字に対して個々に平行移動している場合があるので、切り出しの良し悪しが認識率に与える影響は非常に大きい。これを解決するには、認識システムが最も良く認識できる領域で切り出しを行なうべきであろう。

以上のことから、我々はサッケード機構（視点移動）を組み込んだニューラルネットワークモデルを提案し、簡単なパターンに対するモデルの性能を評価した⁽⁴⁾。本報告では、このモデルのひらがなに対する性能を検討する。

2. モデルの構成

図1に本システムの構成図を示す。視点移動は以下のように行なわれる。視点の近傍内で、bit mapから切り出したおのおのの画像を学習済みのBP部へ入力し、得られた出力値からエントロピーEを求める。このEは(1)式から算出される。それぞれの位置で得られたEを比較し、最小値をとる方向へ視点は移動する。

$$E = \begin{cases} \frac{(V_{max} - V_{mean})(V_{max} - V_{min})}{V_{min} - V_{mean}} & \text{for } (V_{min} \neq V_{mean}) \\ 0 & \text{for } (V_{min} = V_{mean}) \end{cases} \quad (1)$$

ただし、Vmax : 最大出力値
Vmin : 最小出力値
Vmean : 平均出力値

3. シミュレーション実験

図1中のBP部は既に図2に示すパターンをBP法を用いて学習している。図2のパターンを大きな画像に貼り付け、その画像内で各視点の移動軌跡を示したものが図3で、(1)は「あ」、(2)は「い」の場合である。ただし、四角の表示のある点は視点移動しない点（極小値）を表わし、四角の大きさはエントロピー|E|を表わす。また、棒グラフはその時の出力層の興奮の状態を示している。

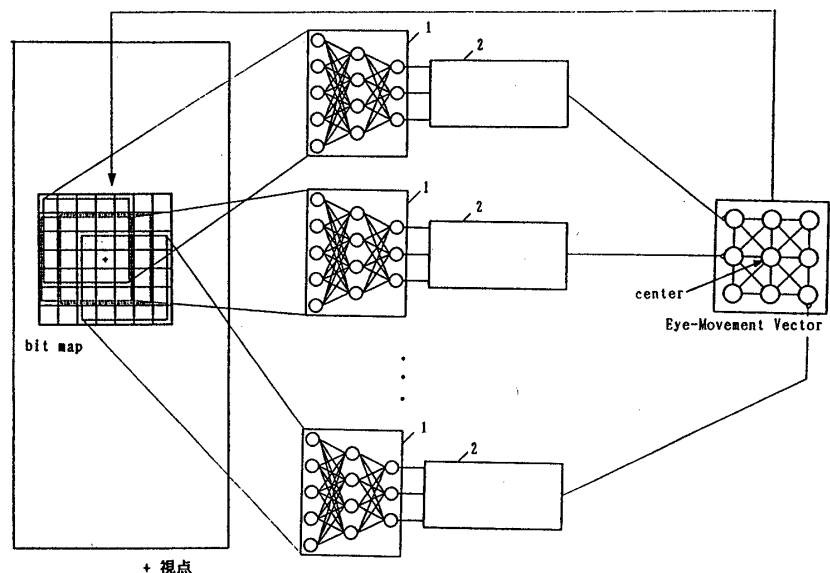


図1 システム構成図

- 1 フィードフォワード型ニューラルネットワーク (BP部)
- 2 エントロピーEを求める部分

あ	い	う	え	お	か	き	く	け	こ	さ	し	す	せ	そ	た	ち	つ	て	と
な	に	ぬ	ね	の	は	ひ	ふ	へ	ほ	ま	み	む	め	も	や	ゆ	よ	ら	り
る	れ	ろ	わ	を	ん														

図2 学習文字パターン

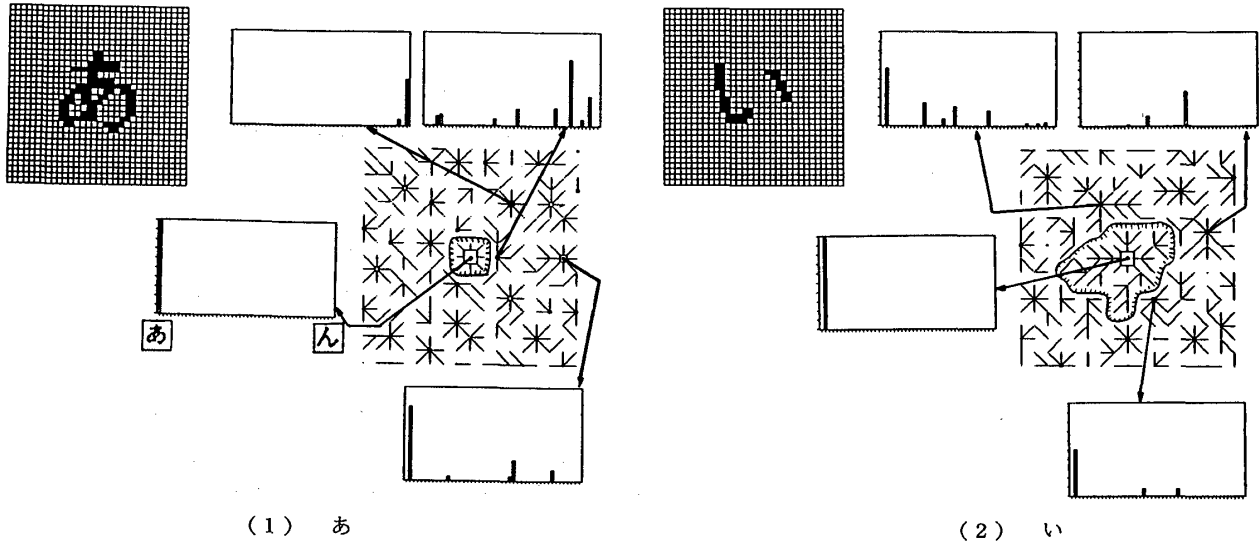



図3 視点移動の軌跡

4. まとめ

このシミュレーション結果から、図3において、で囲った領域内に視点が入っていれば、エントロピー最小点に到達し、正しい切り出しが可能となることが分かった。すなわち、一つの標準パターンを学習すれば、閉領域内に含まれる数だけの未学習パターンを認識できることを意味している。しかし、文字を切り出した時に視点が閉領域内に含まれていない場合には、極小点に捕まってしまう。この極小点から退避する方法として次のことが考えられる。

- 1) 視点移動を決定する際に、エントロピーを計算する範囲を広くとる。
- 2) エントロピーにしきい値を設け、それ以上の値で極小点となった場合その値を0とみなす。これによって多くの極小点が潰され、最小点に到達する範囲も広がることが期待される。

<参考文献>

1. D.E. RUMELHART, J.L. McCLELLAND, and the PDP Research Group, "Learning Internal Representations by Error Propagation" PARALLEL DISTRIBUTED PROCESSING Volume 1
2. 香田、阪上、高木、ノ木、「ニューラルネットによる英数字認識」 信学技法 (PRU)
3. 加藤、高木、「複数の特徴ベクトルを利用したニューラルネットによる手書き数字認識」 信学技法 (PRU)
4. 鷲澤、早野、「サッケード機構を組み込んだ神経回路モデル」 信学技法 (MBE)