

3層BPモデルによる横書き文字列の認識

Recognition of Lateral Strings of Letters Using the 3-Layered BP Model

第2報・捺印数字列の認識

2nd Report: Recognition of Lateral Strings of Sealed Numerals

*今井勝次 **郷原一寿 **内川嘉樹
*Katsuji IMAI **Kazutoshi GOUHARA **Yoshiki UCHIKAWA

*トヨタ自動車(株) **名古屋大学工学部
*TOYOTA Motor Corporation **School of Engineering, Nagoya University

Abstract

We have proposed a new method for recognizing a lateral string of letters using the 3-layered BP with feedback connections. In this report, the ability of the proposed method for the recognition of a lateral string of sealed numerals is demonstrated. Experimental results show the ability of the proposed method to correctly recognize a lateral string of sealed numerals that have lacked parts, blurred parts or shaded parts, when the original images suffers bad conditions: i.e. when two numerals next to each other are connected and when lumped noises are included.

1 はじめに

筆者らは、フィードバック結合をもつ3層BPモデル^{[1][2][3]}を用いた印刷横書き文字列の認識手法を提案している。本手法によれば、原画像中の横書き文字列から個々の文字を切り出しながらか識別することができる。‘A’から‘E’までの5字種の印刷アルファベット大文字で構成される文字列認識実験を行った結果、原画像内に塊状ノイズが存在する場合や、隣合う2つの文字が接合している場合にも個々の文字を切り出して識別することが可能であることを示している^{[4][5]}。本報告では、この手法に対してさらに改良を加え、捺印数字列の認識に応用した事例を報告する。

本報告が対象とする捺印数字列は、電子部品(チップ抵抗器)に捺印されているものである。これらの捺印数字列に含まれる文字には次のような欠陥がある場合が多く、認識を困難にしている^{[4][5][6]}。

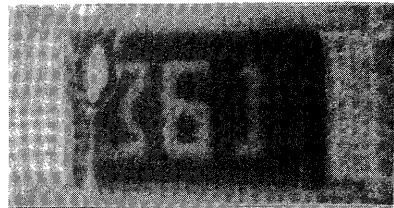
- ① 文字の一部が欠けており、一部が途切れているもの。
- ② 途切れてはいないが文字の一部が細くなったり、捺印が薄いもの。
- ③ 文字の一部、または全部が太くなっているもの。
- ④ 隣合う文字同士が接合している。

また、電子部品の性質上、次のような場合が多く含まれてくる。

⑤ 加工時の塵などにより、塊状のノイズが含まれているもの。

⑥ 部品表面の凹凸が不安定であるため、ハレーションが生じているもの。

これらの捺印数字列の認識に関しては実用に耐えうる手法がなく、これまで自動化は困難とされていた。図1は、認識目標とする電子部品番号の写真である。



(a) 塊状ノイズを含む例



(b) 隣合う文字が接合している例

図1 認識目標とする部品番号の例

(a) は⑤の場合に相当し、(b) は④の場合に相当する。

2 提案する捺印数字列認識手法

図2に、提案する手法を用いた捺印数字列認識の原理を示す。まず、個々の数字のピッチに比べて十分に狭いピッチのスリットを与える。このスリットを、始点の位置から、終点の位置まで移動させる。このことにより、原画像を横方向に走査する。原画像のある部分を走査しているとき、スリット内部に現れる画像をスリット画像と呼ぶことにする。ニューラルネットワークには、このスリット画像を逐次入力していく。

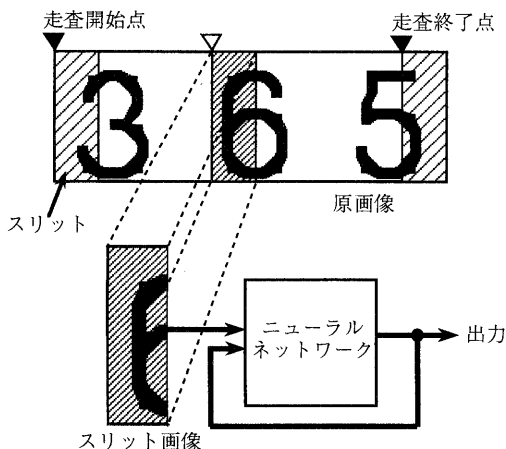


図2 提案する手法による捺印文字認識

図3に、本手法に用いるニューラルネットワークの構造と、入出力関係を示す。スリット画像を含んだ入力に対し、ニューラルネットワークは、走査ステップ信号 (SSS信号) と呼ぶ数字の切り出しに関する信号、所属カテゴリ信号 (CAS信号) と呼ぶ個々の数字の識別に関する信号、そしてセキュリティ信号

(SEC信号) と呼ぶ誤認識防止に関する信号、の3つを出力する。前回までのニューラルネットワークの構造^{[6][7]}と比較すると、SEC信号を加えた構造となっている。原画像の走査は、SSS信号、CAS信号が0の状態を初期状態として開始される。現在のスリット画像が識別対象とする数字の一部分であると判断される場合、SSS信号は、走査の進行と共にその値が増加した後、数字の切れ目で0に戻り、CAS信号は、その数字が続いている間中、1に近い値を取る。これに対し、スリット画像に識別対象とする文字が含まれていないと判断される場合は、2つの信号はともに0に近い値となる。これら2つの信号はフィードバックされ、スリット画像とともにニューラルネットワークの次の入力の一部となる。また、SEC信号は、個々の文字を識別する最初の段階と、最後の段階で変動し、どのSEC信号が変化したかを調べることにより、誤認識 (おもに捺印不良の文字に対する) を防止する。学習後に得られるニューラルネットワークの汎化能力を完全に把握することは容易なことではなく、ニューラルネットワークを用いた情報処理装置を構築する場合、その誤動作を防ぐために、このような監視機構を取り入れていくことが1つの方向であると考える。

SSS信号、CAS信号を観測していけば、個々の数字の切り出しと識別を、スリットによる原画像の走査とともに行うことができる。以前著者らが提案した手法では、'A' から 'E' までの5字種の印刷アルファベット大文字で構成される文字列認識実験の結果、93%程度の認識率を得たが、文字が読み飛ばされることが多いことを報告している^{[6][7]}。これはSSS信号の立ち下がりやCAS信号の立ち下りのタイミングにずれがある場合があることに原因があると考えられ、その方法の改良を行った。

SSS信号が0.2を超えた時点で、それぞれのCAS信号の積分を開始する。その後、SSS信号が

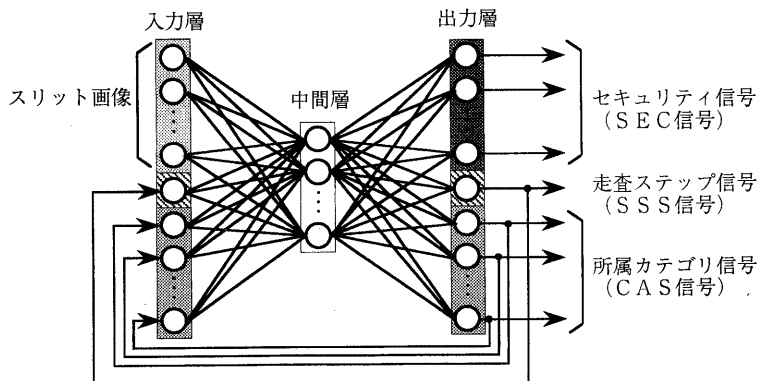


図3 ニューラルネットワーク

0.7を超える値を1回以上とった後、はじめて0.2を下回った時点で、それぞれのCAS信号の積分値を評価する。この間、それぞれのSEC信号が常に監視されており、0.5を超えた時点で、そのSEC信号の番号を順次記録しておく。積分値が最大であったCAS信号に割り当てた文字を認識結果とする。ただし、ここでSEC信号が0.5を超えた順位が異常である場合、認識に誤りがあるものとして判定する(認識除外判定)。また、途中でSSS信号の値が0.7を超えることなく0.2を下回った場合、CAS信号の積分を中止する。

3 捺印数字列認識実験

3.1 実験の概要

提案する手法の有効性を検証するため、電子部品に書かれている捺印数字列を対象とした認識実験を行った。これらの数字列は、'0'から'9'までの数字のうちの3個を任意に配列したものである。原画像は、数字列をCCDカメラによって取り込み、縦60画素、横128画素に2値化したものである。捺印数字列と導電部が同じ白色であって区別不可能であること、また捺印の濃度が一樣であるとは限らないため、2値化処理を行うことが必要であると判断した。これらの原画像は、次の3つのグループに分けられる。

- ① ノイズ、文字の欠陥が少なく、個々の文字が鮮明な原画像22例。
- ② ノイズ、文字の欠陥がある場合で、人間の目では判読可能であるが、比較的読み取りにくい原画像22例。
- ③ ノイズ、文字の欠陥がある場合で、人間の目でも判読不能(捺印不良)なものをいずれも含んでいる原画像11例。

3.2 ニューラルネットワークの構造

本実験では、SSS信号は1個、CAS信号はそれぞれの字種に対して1個(合計10個)、SEC信号は3個とした。ここで用いるニューラルネットワークは、フィードバック結合をもつ3層BPモデルとした。各層のユニット数を示す。中間層ユニット数は、120個、140個、160個、180個の各ユニット数からそれぞれ500回学習させ、誤差が最も減少した160個を採用した。

入力層 47 (= 36 + 1 + 10) 個
 中間層 160 個
 出力層 14 (= 1 + 10 + 3) 個

入力層と中間層の間、中間層と出力層の間に学習による可塑性結合、出力層と入力層の間の結合は固定結合である。

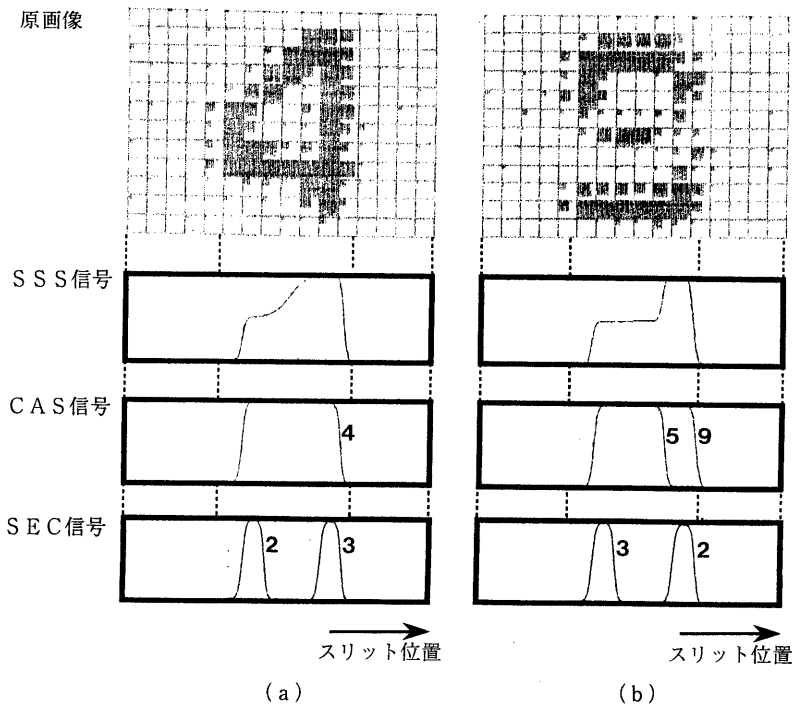


図4 入力画像と教師信号の例

3. 3 ニューラルネットワークの学習方法

入力画像と教師信号の組を乱数を用いて作成し、逐次ニューラルネットワークに提示した。提示する順序は、'0' → '1' → … → '9' → '0' → …の順に固定した。ここで、'0' から '9' の文字を一通り提示することを、1回の学習回数として定義する。誤差はさらに減少する傾向を示していたが、学習に多大な時間を要するため、1,500回で打ち切った。

入力画像と教師信号の例を図4に示す。CAS信号、SEC信号それぞれのグラフの右側の数字は、信号番号(CAS信号では各数字と対応、SEC信号は1から3の範囲)を表す。また、これらの作成に用いる基本パターンのフォントを図5に示す。入力画像は、基本パターンを変形させて画像中心部に置き、そこに塊状のノイズを設定したものである^[5]。今回は塊状のノイズの設定に関して、正のノイズのほかに負のノイズも与えることにより、文字のかすれや欠けに対応した。

3. 4 認識実験結果

図6は、'0' から '9' までの10字種の捺印数字で構成される数字列を認識した実験結果の例である。この図に関しても同様に、グラフ右側の数字が、信号番号を示している。(a) から (c) は、3文字すべての認識に成功した例である。また、(d) は3文字中2文字は正しく認識しているが、1文字が認識除外判定された例である。

まず、(a)を用いて文字列認識のプロセスを説明する。文字'4'がスリット内に現れはじめたスリット位置で、SSS信号、4番目のCAS信号、2番目のSEC信号がほぼ同時に立ち上がっている。その後すぐに、2番目のSEC信号の値は0.5を超えた後に立ち下がっていることがわかる。SSS信号はスリットによる原画像の走査に併せて徐々に増加し、0.7を超える値をとっていることがわかる。文字'4'の右端に近づいた時点で、今度は3番目のSEC信号が立ち上がり、0.5を超える値をとっている。この後、SSS信号は急降下して0.2を下回っている。また、これとほぼ同時に4番目のCAS信

号、3番目のSEC信号が立ち下がっていることがわかる。この時点で、それぞれのSEC信号の積分値を評価する。この場合は図から明かであるように、4番目のCAS信号の積分値が最大であった。この場合はSEC信号が立ち上がる順位が適当であったため、文字'4'と判定された。

ところで、(d)の2番目の文字であるが、これは捺印不良文字である。このような文字は認識除外することが望ましい。この例においても、SEC信号の立ち上がる順位が不適当であったため、認識除外判定が行われている。

結果をまとめる。3.1に示した原画像55種類に関して、認識実験を行った。

<①の画像に関して>

22種類の原画像には、判読可能な66文字が含まれていた。66文字中、64文字(97%)が正しく認識された。また、残りの2文字に関しては認識除外判定がなされた。誤認識は存在しなかった。

<②の画像に関して>

22種類の原画像には、判読可能な66文字が含まれていた。66文字中、54文字(82%)が正しく認識された。また、残りの12文字は認識除外判定がなされた。誤認識は存在しなかった。

<③の画像に関して>

11種類の原画像には、19文字の判読可能な文字が含まれていたが、それ以外は捺印不良であり、判読不能であった。19文字の判読可能な文字に関して、15文字(79%)に関しては正しく認識されたが、残りの6文字に認識除外判定がなされた。捺印不良の部分文字として切り出した場合が3回あったが、いずれも認識除外判定がなされ、誤認識は存在しなかった。

以上をまとめると、151文字の判読可能な文字に対して、133文字(88%)が正しく認識された。また、捺印不良がある場合にも認識除外判定がなされ、誤認識は1つもなかった。

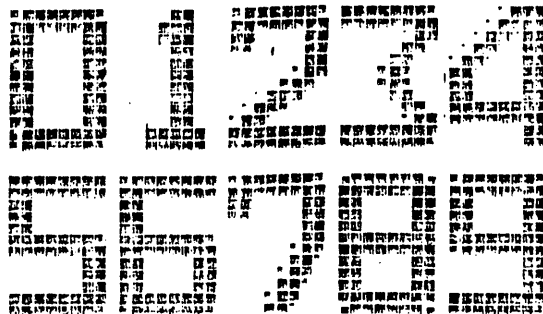


図5 入力画像作成に用いた基本パターン

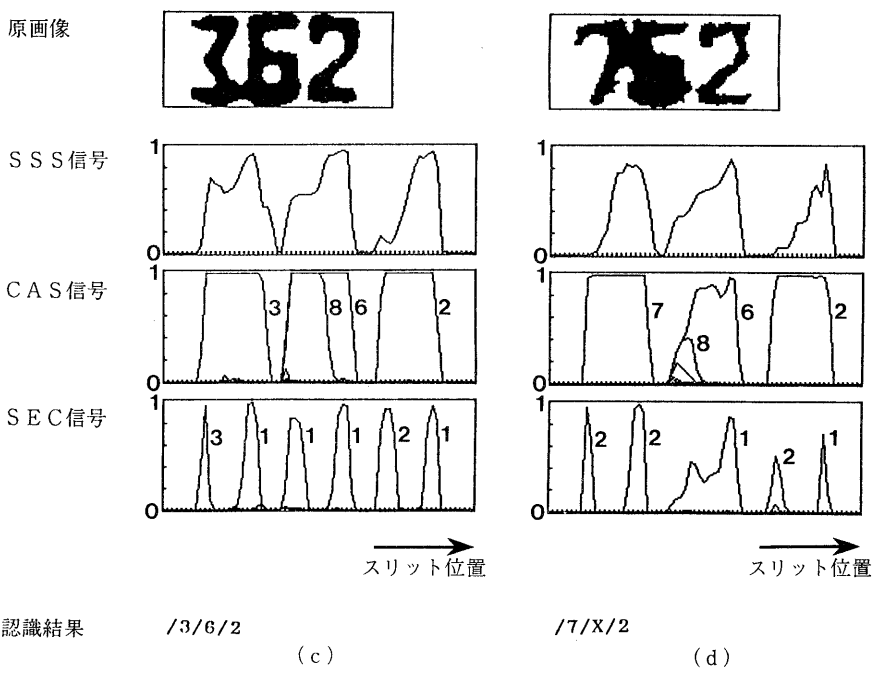
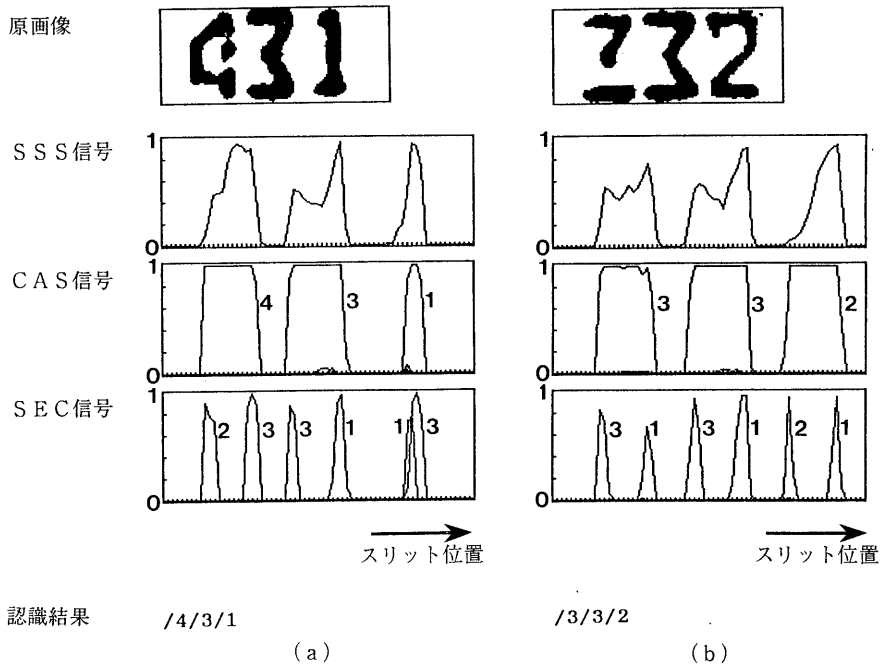


図6 実験結果の例

4. まとめ及び今後の課題

本報告では、筆者らの提案する手法にSEC信号を付加する改良と、ニューラルネットワークの出力として得られる信号の処理方法の改良を行い、この改良手法を用いてチップ抵抗の捺印数字列を認識する実験を行った事例について報告した。そして、判読可能な捺印文字に関しては88%程度の認識率を得ることができ、誤認識を防止できることを確認した。また、SEC信号の付加により、捺印不良の場合にも認識除外判定を正しく行うことができ、誤認識を防止できることを確認した。

今後の課題について述べる。第1に、今回の実験では、誤差がさらに減少していたが、計算に時間がかかるため、学習を途中で打ち切った。計算速度の大きな計算機を用いて、さらに多くの回数の学習を行ってみたい。第2に、SEC信号の設定に関するノウハウを蓄積することも課題であり、現在は試行錯誤の段階である。第3に、SSS信号、CAS信号、SEC信号の設定は手作業で行っており、手間がかかる。文字フォントに応じてこれらの信号を自動的に設定する方法が望まれ、これまでに蓄積してきたノウハウをまとめている。第4に、目的とする捺印数字には、部品上の印面における位置ずれがある。筆者らの提案する位置ネットワーク^{[7][8][9][10]}の応用を試みたいと考えている。第5に、本手法は印刷文字列にのみ有効であるが、手書き文字列の認識手法に発展させたいと考えている。

文 献

- [1]Rumelhart D.E., McClelland J.L., and PDP Research Group: "Parallel Distributed Processing", 1, Chap.8, MIT Press(1986).
- [2]Jordan M. I. : "Attractor dynamics and parallelism in a connectionists sequential machine", Proceedings of the Eighth Annual Conference of the Cognitive Science Society, pp. 531-546(1986).
- [3]Funahashi K.: "On the Approximate Realization of Continuous Mappings by Neural Networks", Neural Networks, 2, 3, pp.183-192(1989).
- [4]沼上泰彦他: "チップ抵抗機捺印文字の自動品質検査アルゴリズム"、情報処理学会第43回全国大会講演論文集、2、pp.319-320(1991).
- [5]今井勝次、郷原一寿、内川嘉樹: "フィードバック結合をもつ3層BPモデルを用いた印刷横書き文字列の認識"、信学論(D-II)、J74-D-II、11、pp.1556-1564(1991).
- [6]今井勝次、郷原一寿、内川嘉樹: "フィードバック結合をもつBPモデルによる電子部品番号の認識"、第22回画像工学コンファレンス論文集(1991).
- [7]Gouhara K., Imai K., Uchikawa Y. : "Position and Size Representations by Neural Networks", Proceeding of 5th IASTED, pp.148-153(1989).
- [8]今井勝次、郷原一寿、内川嘉樹: "ニューラルネットワークの学習による目標パターンの位置、大きさの検出"、信学論(D-II)、J74-D-II、6、pp.748-756(1991).
- [9]Imai K. , Gouhara K. , Uchikawa Y.: "Pattern Extraction and Recognition for Noisy Images Using the Three-Layered BP Model", IJCNN'91-Singapore, 1, pp.262-267(1991).
- [10]今井勝次、郷原一寿、内川嘉樹: "3層BPモデルを用いたノイズのある原画像に対する手書き文字認識"、信学論(D-II)、J75-D-II、2、印刷中(1992).