

## 学習機構を備えた工業用文字認識 エキスパートシステム

松崎吉衛 鈴木英明 岡本啓一  
日立製作所 生産技術研究所

エキスパートシステムとニューラルネットを用いた、文字等のパターン認識を行なう汎用的な工業用視覚システムについて述べる。このシステムは、認識対象の特性と適切な認識アルゴリズムの関係を示す知識を備えたエキスパートシステムにより認識方式を決定し、その方式に基づく最適な認識パラメタの学習と認識実行をニューラルネットにより行なう。認識モジュールは階層型であり、入力ユニットはエッジ抽出や特徴量計算等の画像処理機能を持つ。認識方式はこのネットワークの結合関係、および入力ユニットの機能タイプとして表現される。このシステムを製造現場で使用される文字読取りに用い実用性を検討した。

### Expert System with Learning Mechanism for Industrial Character Recognition

Kichie Matsuzaki Hideaki Suzuki Keiichi Okamoto  
Production Engineering Research Laboratory, Hitachi, Ltd.

292 yoshida-cho, Totsuka-ku, Yokohama 244, Japan

This paper describes an industrial pattern recognition system consisting of an expert system and a neural network. This system decides recognition method by the expert system using knowledge about object features and appropriate algorithm relations, learns optimum recognition parameters based on the decided method, and executes pattern recognition by the neural network. A recognition module has layered structure, and has input units with image-processing functions such as edge detection and feature extraction. The recognition method is presented by connection style and image-processing type to the recognition module. The system was evaluated by applications for character recognition system using factory floor.

## 1. はじめに

フレキシブルな生産設備を実現するための重要な要素として、文字や部品の種類等をパターン認識により判別する工業用視覚システムの適用が増大している。現在、このようなパターン認識は、認識対象の判別に適した特徴量を用いて行なわれるのが普通である。この方法は、適切な特徴量が見い出せた時は優れた認識機能を実現できるが、この特徴量は認識対象に依存するので、生産ラインで使用される多様な文字や部品に対応できる汎用的な装置の実現は難しい。このため認識対象毎に専門家による画像処理方式の検討やプログラム開発が必要となり、この工数削減が強く望まれている。

そこで、画像処理や視覚システムの開発の自動化を実現する方法の1つとして画像処理/視覚エキスパートシステム[1][2][3][4]が開発されている。これらは一般的な画像処理、および応用分野に関する知識を備え、処理手順等を生成するものである。ただし、各種の認識対象に適した処理のパラメタまで生成することは難しく、処理パラメタの決定はユーザにまかせていることが多い。

また、システム工数削減の他の方法として、学習機構を用いた自律的な認識システムの使用が考えられる。この方法にはニューラルネットを用いたもの[5][6]、統計的手法を用いたもの[7]等がある。文字認識に関しては、バックプロパゲーションアルゴリズムを用いて、かなり良い結果が得られているが、実用レベルの性能を得るための学習にはかなりの時間とノウハウを要するという問題がある。

本システムのねらいは、エキスパートシステムとニューラルネットの特徴を組合せ、認識対象に関する高次の知識を備えたエキスパートシステムにより認識方式を決定し、その方式に基づく認識パラメタの最適化と認識実行をニューラルネットで行なうことである。

ここで、エキスパートシステムの設計結果はニューラルネットの構造、すなわちニューロンの機能、ネットワークの接続関係、重み係数の初期値

等で表現される。このためには、各種パターン認識アルゴリズムがニューラルネットにより一般的に実行できる必要がある。これは、画像濃度値を入力データとし、シグモイド関数を用いた3層程度のニューラルネットでは困難と見られるので、本システムでは、入力ユニットに画像処理機能を持たせることで対処している。

## 2. システムの概要

本システムは図1に示す様に、設計モジュールと、認識モジュールから構成される。

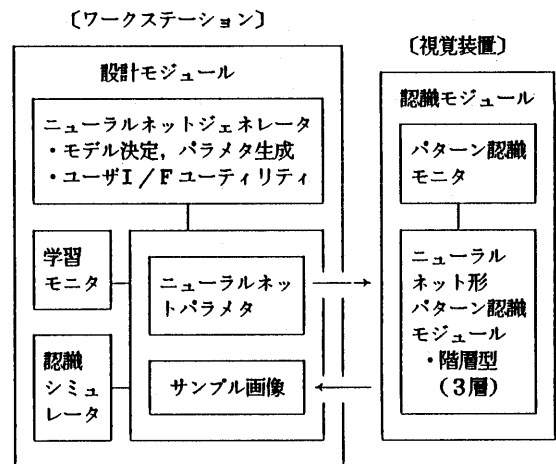


図1 システム構成

認識モジュールには、3層のニューラルネット形パターン認識モジュールがあり、工業用視覚装置で実行される。ここで、入力層ユニットの画像処理機能は画像処理回路、中間層と出力層ユニットの演算はDSPおよびマイクロコンピュータによって行なわれる。

設計モジュールは、ニューラルネットジェネレータ、学習モニタ、および認識シミュレータから構成される。ニューラルネットジェネレータはユーザに認識対象に関する質問を出し、その回答と内蔵するパターン認識に関する知識からニューラルネットパラメタを生成する。学習モニタはこのパラメタと、あらかじめ視覚装置からアップロード

しておいたサンプル画像を認識シミュレータのデータエリアにセットし、認識シミュレータを起動する。認識シミュレータは視覚装置に搭載されたニューラルネット形認識モジュールと同等のものであり、認識処理を行なう毎に学習モニタに結果を返す。学習モニタは認識結果に基づきニューラルネットパラメタを修正し、所定の認識性能が得られるまで学習を繰り返す。学習が終了したら設計モジュールは、そのニューラルネットパラメタを認識モジュールに転送し、認識実行ができるようになる。この設計モジュールは、エキスパートシステム構築シェルで記述されワークステーションで実行される。

るものであり、隠れユニットは、通常のように入力ユニットと全結合しているものである。明示ユニットは、入力ユニットと出力ユニットの結合関係が予測できる場合に生成される。なお、1つの入力ユニットとのみ結合するような中間層ユニットは原理的には不要であり、入力ユニットと出力ユニットを直接結合させてもよいが、同一の入力ユニットが明示ユニットだけでなく隠れユニットにも結合する場合があるので、ネットワークの階層構造を3層に規格化するために明示ユニットを設けた。

### 3. システム構成

#### 3.1 ニューラルネット形パターン認識モジュール

この認識モジュールは、図2に示すような3層の階層型ニューラルネットである。中間層と出力層のユニットは通常のシグモイド関数であるが、入力層のユニットは特定範囲の画像に対する画像処理演算を行なう。現在備えている画像処理演算には、濃度平均値、上下左右各方向の微分値、縦横比や縦／横構造ベクトル等数種類の特徴量計算がある。中間層のユニットには、明示ユニットと隠れユニットと呼ぶものがある。明示ユニットは、1つの入力ユニットとのみ閾値付きで結合してい

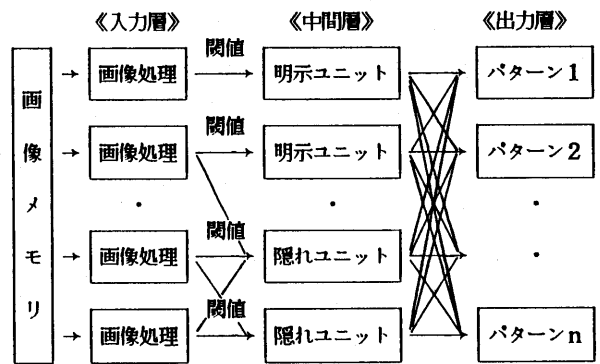


図2 ニューラルネット型パターン認識モジュールの構造

表1 ニューラルネットパラメタの種類と設定方法

分類	データ種類	設定方法
入力ユニット	個数、位置、形状、画像処理のタイプ	Nジェネレータがユーザ指定と知識ベースにより設定
明示ユニット	個数、入力ユニットとの結合係数、 閾値、出力ユニットとの結合係数	Nジェネレータがサンプル画像の解析により初期値を設定し、学習モニタが学習結果を設定 Nジェネレータがユーザ指定と知識ベースにより初期値を設定し、学習モニタが学習結果を設定
隠れユニット	個数、入力ユニットとの結合係数 出力ユニットとの結合係数	Nジェネレータがユーザ指定と知識ベースにより初期値を設定し、学習モニタが学習結果を設定 Nジェネレータがランダムに初期設定し、学習モニタが学習結果を設定
出力ユニット	個数	Nジェネレータがユーザ指定のパターン数だけ設定

### 3. 2 ニューラルネット表現モデル

ニューラルネット形パターン認識モジュールの構造はニューラルネットパラメタにより記述される。このニューラルネットパラメタにより、パターン認識モジュールは3層の範囲において任意の結合が定義される。

ニューラルネットパラメタの種類と設定方法は表1の通りである。

### 3. 3 パターン認識モニタ

ニューラルネットパラメタは1つのデータファイルとして保持されており、動的に変更が可能なので、認識対象パターンが変化した時は、このパラメタを変更することによりパターン認識モジュールの判別機能を変更できる。

パターン認識モニタは、パターン認識モジュールにあり、認識対象パターンに対して最適なニューラルネットパラメタ設定することにより判別機能を認識対象に応じて最適な状態に変更するものである。

この機能は、出現するパターンの候補が出現場所に対応して特定の種類に限定できる場合の多い工業用文字/パターン認識において特に有効である。例えば、ICパッケージに印字された型式名称の先頭はメーカーに固有の数種類の英字に限られることがある。このような場合、桁毎にその桁に対して指定された候補文字だけの判別を行なうニューラルネットパラメタを使用すると、英数字全部の判別をおこなうニューラルネットパラメタを使用するより認識率が高くなると考えられる。

### 3. 4 ニューラルネットジェネレータ

ニューラルネットジェネレータは、ユーザが認識対象の種類や特徴を指定すると、ユーザに質問を出しながら内蔵する知識を用いてニューラルネットパラメタの初期値を生成する。ユーザが指定するものとしては、対象とする実物パターンの他、特徴となりそうな部分、その範囲の特徴抽出に適した画像処理の種類等、人間にとっては容易に認識できるが、計算機にとっては難しい項目が含ま

れる。このユーザ指定項目については、ユーザの判断を容易にするためのユーティリティを備えている。この例としては、画像処理の種類を決定するために微分等の画像処理を行い結果を表示したり、多変量解析により重み係数の初期値を決めるもの等がある。

ここで蓄積した知識は、すでに開発されているパターン/文字認識アルゴリズムに基づいている。これまで検討した結果では、既存のパターン/文字認識アルゴリズムのかなりの部分はニューラルネットパラメタ、すなわちネットワークの結合関係を適切に設定することにより、画像処理機能を持つ入力ユニットを備えたニューラルネットで近似的に表現できる。これは、ニューラルネットモデルが一般的なパターン認識モジュールとして使えることを示していると考えられる。

なお、ニューラルネットジェネレータが生成するパラメタのうち、入力ユニットの個数、位置、形状、画像処理のタイプ、明示ユニットの個数、隠れユニットの個数、出力ユニットの個数は、そのまま認識モジュールに渡されるが、各ユニット間の結合係数は学習の初期値であり、最終的に認識モジュールに渡される学習結果とは異なる。

### 3. 5 学習モニタ

学習モニタは、ユーザからの学習に関する指示を受け付け、短時間に適切な学習結果が得られるように認識シミュレータを管理するものである。学習のさせ方にも認識パターンの特性に応じたノウハウがあり、学習モニタはこの知識を備えている。学習モニタの主要な機能は以下の通りである。

#### (1) ユーザインタフェース

通常の学習操作に必要な、温度、学習終了条件、学習係数等の学習パラメタを入力したり、結果をユーザに理解し易い形で出力する。

#### (2) 学習パターンの生成

実際に入力したサンプル画像から実際の認識環境(位置ずれ、明るさ変動、形状変形等)にあわ

せて、学習用画像の生成を行なう。これにより比較的少数のサンプル画像から、各種の変動に対応できる学習を行なうことができる。

### (3) 学習戦略立案

学習の収束時間を短くしたり、誤認識を極小にするような学習戦略が各種開発されている[8]。学習モニタはこのような学習戦略に関する知識を用い、誤りなくパターンの判別を行なうニューラルネットパラメタをできるだけ速く得るものである。

## 4. 工業用文字認識システムへの適用

本システムを工業用文字認識システムの開発に適用した結果を示す。工業用文字[9]は、工場で製品や部品を判別するために印字されるもので、フォントは決まっていることが多いが、複数のフォントが混在することはある。文書文字認識と比べた時の問題は、印字する相手が多様な材料や部品のため文字と背景のコントラストが悪く、文字領域抽出のため画像処理に工夫を要することである。また、よごれやかすれによるパターンの変形はランダムであり、パターンの特徴部にかかることもある。文字の種類は英数字が多いが、限定された漢字やマークも対象になる。

### 4.1 ニューラルネットパラメタ生成知識

ニューラルネットジェネレータがニューラルネットパラメタを生成する場合に使用する知識として、以下の様な文字の特性と認識アルゴリズムの関係[10]を用いている。各認識アルゴリズムは図3に示すニューラルネットの形に展開できる。いずれも明示ユニットだけを用いた機能的には2層のネットワークになる。ここにおいてニューラルネットを使う意味は、各種アルゴリズムを一般的なニューラルネットに対するパラメタとして扱えること、および図に示す基本構造から出発して必要に応じて隠れユニットを設けることにより、各アルゴリズムで認識できないパターンに対する機能拡張が可能なることである。

### (1) 定点サンプリングマッチング

〔文字の特性〕

- ・対象フォントが固定。
- ・位置ずれ、大きさ変動無し。
- ・文字領域と背景領域の区別が明確。
- ・文字の種類があまり多く無く類似文字無し。

〔認識方法〕

文字の判別に適した場所にウィンドウを設定し、ウィンドウ内の黒白パターンが辞書パターンに近い程出力ユニット値が高くなるようにする。ウィンドウとする部分はフォントからユーザが設定する。

〔ニューラルネット表現〕

出力ユニットの重み係数は一定。

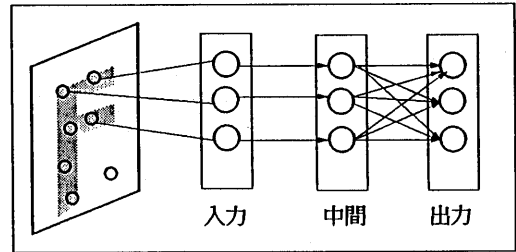


図3(1) 定点サンプリングの表現

### (2) 単純マトリックスマッチング

〔文字の特性〕

- ・対象フォントが固定。
- ・位置ずれ、大きさ変動無し。
- ・文字領域と背景領域の区別が明確。
- ・文字の種類が多くても可。

〔認識方法〕

文字枠内をメッシュに分割し、各メッシュの黒白パターンが辞書パターンに近い程出力ユニット値が高くなるようにする。

〔ニューラルネット表現〕

出力ユニットの重み係数は一定。

### (3) 重み付きマトリックスマッチング

〔文字の特性〕

- ・対象フォントが固定。
- ・位置ずれ、大きさ変動が若干ある。

- ・文字領域の濃淡にはバラツキがある。
- ・文字の種類が多くて可。

〔認識方法〕

文字枠内をメッシュに分割し、各メッシュの黒白パターンに重みを付け、辞書パターンに近い程出力ユニット値が高くなるようにする。

〔ニューラルネット表現〕

出力ユニットの重み係数は文字毎に黒白になる確率から求める。

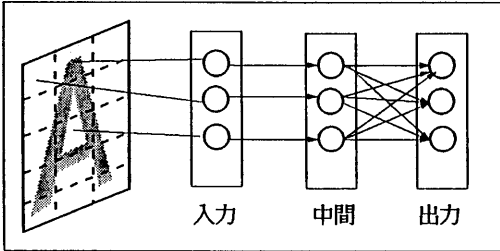


図3(2)(3) マトリックスマッチングの表現

(4) ストロークアナリシス

〔文字の特性〕

- ・対象フォントは直線部が多く固定。
- ・位置ずれ、大きさ変動が若干ある。

〔認識方法〕

文字線の直線部にあたる場所にウィンドウを設け、ウィンドウ内の文字線の有り無しパターンが辞書パターンに近い程出力ユニット値が高くなるようにする。ウィンドウとする部分はフォントからユーザが設定する。

〔ニューラルネット表現〕

入力ユニットは文字線検出のための画像処理機能をもつ。出力ユニットの重み係数は文字毎に文字線の存在する確率から求める。

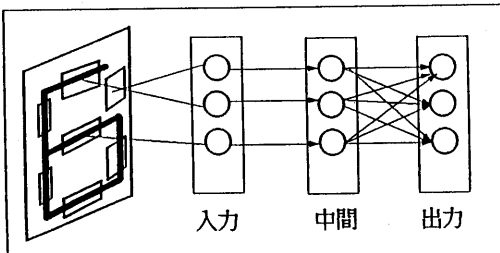


図3(4) ストロークアナリシスの表現

(5) 特徴マッチング

〔文字の特性〕

- ・対象フォントが固定でなくても可。
- ・位置ずれ、大きさ変動が若干ある。
- ・文字の種類が多くて可。

〔認識方法〕

従来の文字認識で有効とされている特徴量[11]を求める入力ユニットを設け、特徴量の値を組み合わせた値で文字の種類を評価する。

〔ニューラルネット表現〕

入力ユニットは特徴量抽出のための画像処理機能をもつ。入力ユニットの処理範囲は特徴量の種類に依存し、文字枠内全域となることもある。出力ユニットの重み係数は文字毎に文字線の存在する確率から求める。

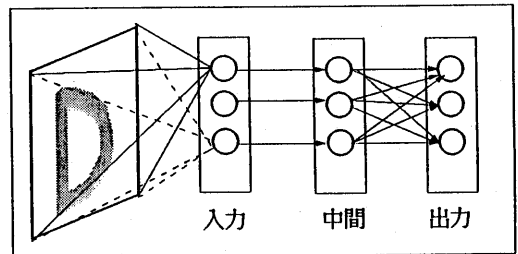


図3(5) 特徴マッチングの表現

4.2 ニューラルネットパラメタ生成におけるユーザインタラクション

(1) 多変量解析

隠れユニット、および出力ユニットの重み係数を学習で求める際、初期値の与え方がローカルミニマムの回避、学習効率の向上に大きな影響を与える。この初期値の決定には、多変量解析が有効な場合が多いとされている[12]。そこで、ニューラルネットジェネレータは、重回帰分析、判別分析、クラスタ分析のパッケージを備え、ユーザはこれを利用して重み係数の初期値を決めることができる。

(2) 入力ユニット部の特徴量分布

入力ユニットの場所を決定する際、入力ユニッ

トから得られる特徴量は、同じ文字に対しては値の分散が小さく、異なる文字間では分散が大きいほうが判別に有利と考えられる。ユーザは指定した入力ユニットに対して各文字の特徴量分布をグラフでみることができ、入力ユニットの場所を変動させながら適切な場所を見つけることができる。

また、入力ユニットで抽出した特徴量は明示ユニットで正規化されるので、この特徴量の分布は明示ユニット閾値の初期値を決める上での参考にもなる。

### (3) 候補文字指定

文字列の桁毎に使われる文字が決まっている場合、その候補文字についてだけ学習することにより、認識率を向上させることができる。ユーザが桁毎にこの候補文字を指定すると、それを記憶し学習の際利用できるようにする。

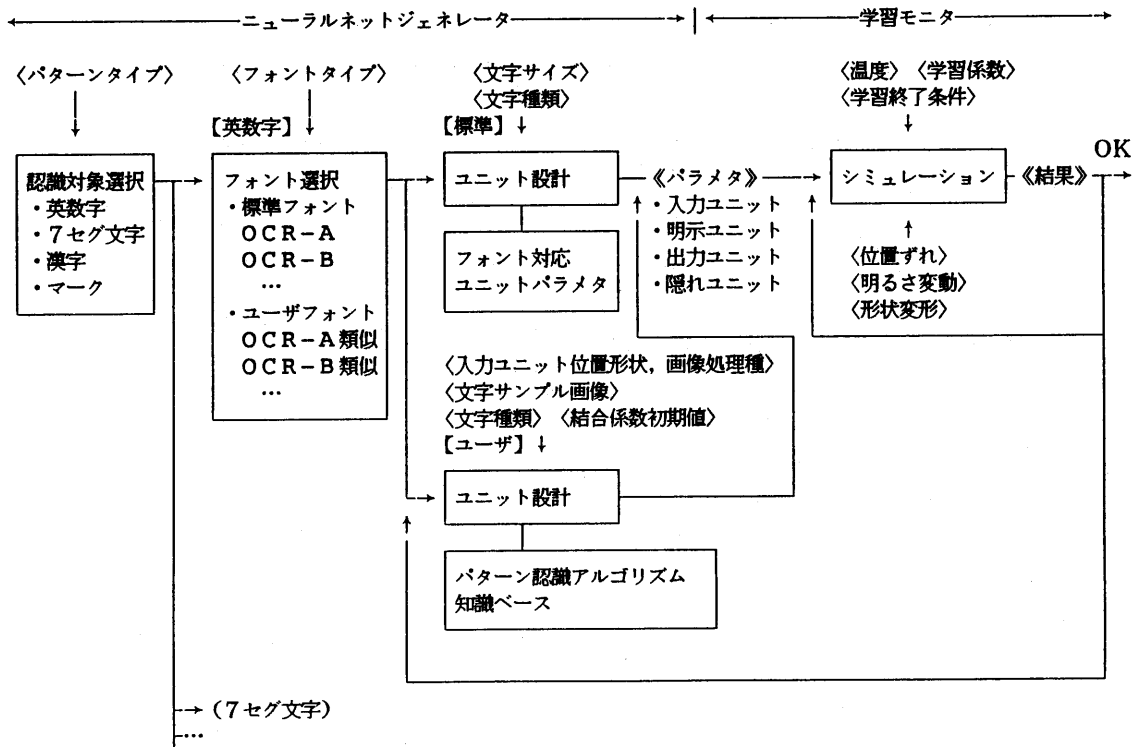
## 4. 3 ニューラルネットパラメタの生成フロー

設計モジュールを用いたニューラルネットパラメタの設計フローは図4に示すようになる。

### (1) ニューラルネットパラメタの初期設定

ここで用いた事例の文字フォントは、OCR-Aフォントをもとに若干の修正を加えたものである。そこで、はじめにフォント選択メニューが表示されるので、“OCR-A類似”のタイプを選択する。

これにより、ユーザ定義フォントとして設計が開始される。まず、直線的形状の“OCR-A類似”の文字に対してはストロークアナリシス法が適しているという設計知識に基づき、入力ユニットの場所となる文字線の検出に適した場所をユーザに設定させる。また、文字線検出は画像の1次元微分を用いることとし、文字線検出に適した微分方向をユーザに設定させる。この設定の際、文字判別に適した入力ユニットの場所すなわち微分処理の範囲を適切に決めるユーティリティとして、



ユーザが指定した入力ユニットの場所を中心として、一定の範囲でサイズと位置を変化させ、入力ユニット値の分散を計算表示する機能を使用する。また、この機能は明示ユニットの閾値の初期値を決めるのにも使用する。この操作により入力ユニットの個数、場所、機能が決まる。

つぎに、対象文字の種類を指定させる。この印字パターンの場合、桁毎に使われる文字が異なるので、桁毎に対象文字を指定する。これから各桁の対象文字を全て含んだものが出力ユニットの個数と種類になる。

この入力ユニットと出力ユニットの関係は、文字毎に入力ユニットの範囲における文字線の有無が予想できるので、中間層ユニットとしては、1つの入力ユニットとのみ結合し出力ユニットとの関係が予測できる明示ユニットが生成される。この明示ユニットと出力ユニットの結合係数の初期値は、多変量解析に基づく計算結果や、入力ユニットの場所の文字線の有無に基づくユーザによる直接的指定によって設定される。

## (2) ニューラルネットワークパラメタの学習

学習モニタは、ニューラルネットワークジェネレータにより設定されたニューラルネットワークパラメタに対し、まず明示ユニット閾値を固定して、出力ユニットへの重み係数をバックプロパゲーション法により学習させる。つぎに、この学習による認識率の変化が一定値以下になったところで、明示ユニット閾値も含めて学習を行なう。なお、学習においては、前に述べた学習パターンの生成や、学習戦略が利用される。

以上の学習によって必要な認識率が得られない時は、文字列全体についてではなく、桁毎に指定された候補文字についてのみ学習を行なう。本システムでは、パターン認識モニタが動的にニューラルネットワークパラメタを変更できるので、桁毎に対応するニューラルネットワークパラメタを用いることで認識率を上げることができる。

それでも必要な認識率が得られない場合は、入力ユニットや中間層ユニットの変更が必要なので、

一旦学習モニタからニューラルネットワークジェネレータに戻り、隠れユニットの追加等、新たなニューラルネットワークパラメタを生成する。これを用いて再度学習を行うプロセスを、良好な認識結果が得られるまで繰り返す。

## 4.4 考察

ニューラルネットワークの構造を認識対象に合わせるための効果を調べるため、学習回数と認識率の関係をいくつかのニューラルネットワークについて調べた。

図5において、(a)は始めの設計結果である微分入力ユニット13個とそれに接続した明示ユニットを用いたもの、(b)はこの入力ユニットに加え9個の入力ユニットとそれに接続した隠れユニットを用いたもの、(c)は $10 \times 13$ の130個の入力ユニットと、40個の隠れユニットを用いたものである。ここで、結合係数の初期値は明示ユニットに関連するものは設計者が判断した値を用い、隠れユニットに関連するものは乱数を用いた。

ここで、(a)は設計者が文字フォントからヒューリスティックに考えた認識アルゴリズムに相当し、(c)は典型的な文字認識用ニューラルネットワークである。(b)はヒューリスティックな認識アルゴリズムをベースにして学習的に特徴量を抽出するユニットを付加したものと考えられる。(a)より(b)の方が学習が速いのは、実際の文字画像が、印字機、下地材質、照明等の特性により、濃度むらや形の

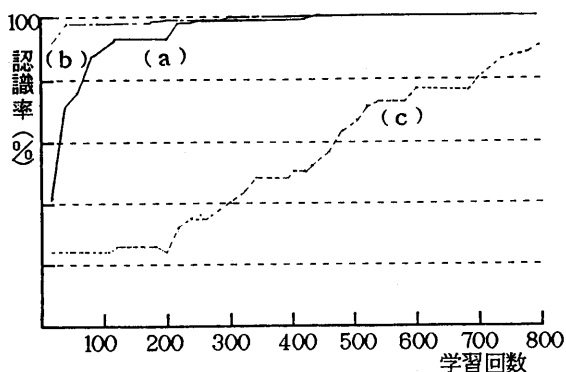


図5 ニューラルネットワーク構造による学習傾向の違い



ゆがみを持ち、フォントデータから考えた特徴領域以外にも特徴となるパターンがあるためと考えられる。また、(a)(b)が(c)より学習が速いのは、認識対象に関する先見的な知識によりニューラルネットを設計することの有効性を示していると考えられる。

#### 5. おわりに

文字等のパターン認識を行なう汎用的な工業用視覚システムの構築を目的として、これ迄に蓄積されたパターン認識に関する知識を用いて、認識対象に適したニューラルネットの構造を生成し、それに基づく学習によって認識パラメタを求め、認識を実行するシステムを開発した。

現在、ヒューリスティックなアルゴリズムによるパターン認識システムの開発には試行錯誤的調整を必要とし、また、ニューラルネットによる自律的パターン認識システムにおいても、学習方法に関するノウハウが重要といわれている。本システムは、エキスパートシステムとニューラルネットを組み合わせることで問題点を補うことをねらっている。

今後、さらに多くの事例に適用し、パターン認識アルゴリズム、およびニューラルネットの使用法に関するノウハウを知識ベースとして蓄積していくことが課題である。

#### 参考文献

- 1) 田村, 坂上: 画像解析エキスパートシステムのための3種の知識, 信学技報, PRL83-49, pp. 27-40 (1983)
- 2) 特集: 画像処理エキスパートシステム, 情報処理学会論文誌, Vol. 29, No. 2 (1988)
- 3) 北村, 佐藤, 田村: 産業用マシンビジョンのためのエキスパートシステム, 昭和63年度人工知能学会全国大会論文集, pp. 551-554 (1988)
- 4) 松崎, 泰, 岡本: オブジェクト指向モデルを用いた工業用ロボット視覚エキスパートシステム, 画像理解の高度化と高速化シンポジウ

ム講演論文集, pp. 43-46 (1989)

- 5) 麻生: ニューラルネットワーク情報処理, 産業図書(1988)
- 6) 山田, 上, 溝口, 天満: ニューラルネットを用いた文字認識, 信学技報, Vol. 58, No. 177, PRU88-58, pp. 79-86 (1988)
- 7) 大津: 適応学習型汎用図形認識—柔らかな視覚認知をめざして—, 昭和63年度第2回エキスパートビジョン研究会資料, (1989)
- 8) 香田, 阪上, 高木, 木: ニューラルネットによる英数字認識, 信学技報, Vol. 88, No. 492, PRU88-151, pp. 33-40 (1988)
- 9) 江尻: 工業用画像処理, 7章, 昭晃堂 (1988)
- 10) 橋本: 文字認識論, 第3章, オーム社 (1982)
- 11) 萩前, 泰, 日比, 軍司, 矢野: 多機能実装プリント板外観検査装置(2), 1989年度精密工学会春季大会論文集, pp. 533-534 (1989)
- 12) 二木: バックプロパゲーションアルゴリズムの初期値決定法, 信学技報, Vol. 89, No. 53, NC89-10, pp. 33-40 (1988)