

# イメージスキャナ入力による数式認識システムの開発

4G-2

須田 哲也, 成嶋 弘, 峯崎 俊哉

東海大学理学部

## 1. はじめに

文字認識は、情報処理技術の一分野であり、人間と機械の快適なインタフェースを目指してさまざまな研究がなされている。文字認識を入力の方から大別すると、ペンなどで書いている文字をリアルタイムに読みとる方式と、書かれた文字や印刷された文字をスキャナなどで読みとる方式に分けることができる。前者の場合、筆順も考慮できることから認識率はかなり高く、さまざまな製品に応用されているが、後者の場合はまだ認識率も低く、実用化という面ではやや後れている。

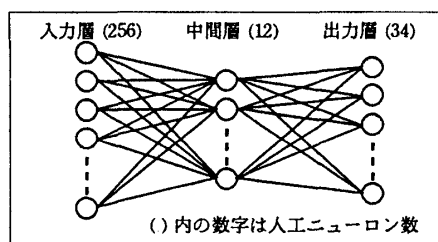
本稿では、後者の方式を用い、数式認識を目標とした「スキャナ入力による数式認識システム」について述べる。本システムは、「数式の TeX 書式出力」という特徴を持っている。

## 2. ニューラルネットワークによる文字認識

一般的な文字認識の場合、入力文字のサイズごとに基本パターンを用意し、最短距離法などでパターン照合を行い最も似ているものを選択する。しかし、この方法では常に比較するための基本パターンを用意しなければならない。また、全パターンと照合することから認識時間は長くなってしまふ。それに対しニューラルネットワークでは、システムの学習時のみ基本パターンが必要で、認識時に比較することはない。認識時間については、並列処理上では飛躍的に短くなるが、逐次処理上ではネットワーク構成により短くも長くもなる。本システムでは、ニューラルネットワークによる文字認識を採用した。

## 3. システムの概要

本システムのニューラルネットワーク構成は、図1のようになっている。



【図1 ネットワーク構成】

入力層の人工ニューロン数は、 $16 \times 16$  のドットマトリックスに対応しており、文字の切り出しで得た任意の大きさ（最大  $100 \times 100$  ドット）の入力パターンは、分数など一部を除き、すべて以下のように正規化される。

入力パターンの  $(x_{i-1}, y_{i-1}) \sim (x_i, y_i)$  のブロックに対して、ブロック内が黒  $\geq$  白の時は黒、その他の時は白に変換する。但し、

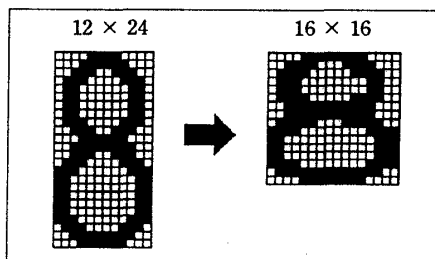
$$x_i = (\text{入力パターンの } x \text{ 座標幅} / 16 \times i \text{ の四捨五入})$$

$$y_j = (\text{入力パターンの } y \text{ 座標幅} / 16 \times j \text{ の四捨五入})$$

( $i=0, \dots, 16, j=0, \dots, 16$ )

黒とは点のある所、白はその逆の意味。

$12 \times 24$  ドットの例を図2に示す。



【図2 パターン正規化の例】

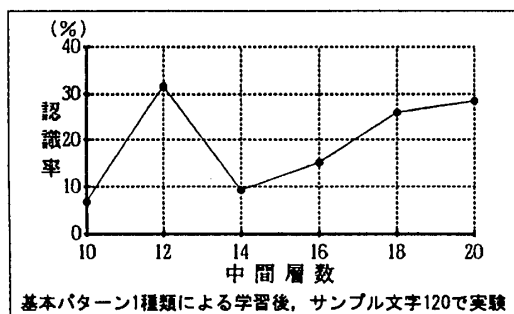
A Formula Recognition System by Scanner  
 Tetsuya Suda, Hiroshi Narushima, Toshiya Minezaki  
 Faculty of Science, Tokai University  
 1117 Hiratsuka, Kanagawa 259-12, Japan

出力層は、図3に示す34文字の認識文字に対応しており、34とした。

0	1	2	3	4	5	6	7	8	9				
a	b	c	d	e	f	g	h	i	j	k	x	y	z
+	=	÷	/	(	)	{	}	Σ	∫				

【図3 認識文字】

中間層の数は、実験結果(図4)より12とした。



【図4 中間層数と認識率】

#### 4. 数式のTeX書式出力

従来の文字認識ソフトでは、数式なども一般的な文書と同じように文字として扱われてきた。しかし、本システムでは、数式の認識結果を図5のようにTeXの書式に合わせて出力するため、そのままTeXの文書に利用することができる。

$f(a) = \frac{3}{2b+1}$
$f(a) = \frac{3}{2b+1}$

【図5 数式認識結果の例】

具体的には、まず、文書イメージ全体のx軸方向、y軸方向の点の分布を調べ、文字単位のブロックになるまで分割していく。そして、分割されたブロックに対して、÷を $\frac{\text{div}}$ 、総和記号Σを $\sum$ 、積分記号∫を $\int$ として出力し、その他の文字はそのまま出力する。また、分数や- (マイナス) などのような横に細長いものは、文字の大きさや位置関係によって認識し、分数は $\frac{\text{num}}{\text{den}}$ の形で出力

する。

TeXでの数式処理が苦手な人にとっては、便利なツールになるものと考えられる。

#### 5. おわりに

本システムは、文字認識の中でも特に、印刷または書かれた数式の処理に重点をおき、そしてTeX書式の出力を行うことが大きな特徴である。システム自体は開発中であり性能的にはまだ満足のいかないものであるが、分数などの簡単な数式は認識できることから、その可能性は十分にあると考えられる。

ところで、本システムではニューラルネットワークの利用とはいっても、逐次処理による単なるシミュレーションに止まっている。最近、ニューロチップなる並列処理が可能なものも登場し、今後これらのチップを搭載したコンピュータが出回れば、処理速度が飛躍的に向上し、ニューラルネットワークも利用価値が上がるものと予想される。

入力パターンの正規化による文字変形対策、ネットワークの学習時間短縮、認識率の向上、認識文字数の増加など、残された課題は多いが、今後はこれらの課題を解消し、実用性の向上を目指していきたいと考えている。

#### 参考文献

- [1] J.E.Dayhoff 著, 桂井浩訳: “ニューラルネットワークアーキテクチャ入門”, 森北出版, 1992.
- [2] 船橋誠壽: “ニューロコンピューティング入門”, オーム社, 1991.
- [3] 安居院猛, 中嶋正之, 木見尻秀子: “C言語による画像処理”, 昭晃堂, 1990.
- [4] 平野広美: “Cでつくるニューラルネットワーク”, パーソナルメディア, 1991.
- [5] 須田哲也: “イメージスキャナ入力による数式認識システムの開発”, 東海大学大学院 理学研究科 数学専攻 1995年度修士論文.