

オンライン手書き数式認識システムにおける行列要素の位置関係に基づく行列認識機構

豊住 健一[†] 鈴木 隆広^{†*} 森 健策^{††} 末永 康仁^{††}

A Real-Time Recognition of Handwritten Mathematical Expression Including Matrix Structure Based on Position Relations among Elements of a Matrix

Kenichi TOYOZUMI[†], Takahiro SUZUKI^{†*}, Kensaku MORI^{††},
and Yasuhito SUENAGA^{††}

あらまし 本論文では、オンライン手書き数式認識システムにおける行列構造の認識手法について述べる。このシステムは手書き入力された数式の文字と構造を逐次認識し、認識結果を $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ のソースコードとして出力する。これまで数式の2次元構造として、添字、根号、総和、分数などを認識対象としていたが、行列構造には対応していなかった。行列構造は明示的な記号が現れずに数式要素が垂直方向に位置するため、専用の認識機構が必要であると考えられる。本論文では、行列の幾何学的特徴や各要素の位置関係を考慮し、実時間の利用に適した行列認識機構の追加を行った。拡張したシステムに行列を入力する実験を行い、本手法の有効性を確認した。
キーワード 手書き入力, 実時間認識, 数式構造認識, 行列構造認識, 行列要素の位置関係

1. ま え が き

計算機を利用して技術文書を作成する際に、図や表と同様に数式の作成にも時間を要する場合が多い。キーボードなどを用いて容易に入力できる通常のテキストとは異なり、数式は垂直方向へも文字が展開する構造をもっているため、計算機上で入力する場合には専用の機構を要する。文書処理の一つとして計算機への数式入力を実現することを目的とした数式認識の研究は長い歴史をもっており [1], [2], スキャナや表示一体型タブレット（以下タブレットと略記）等の入力装置技術の成熟に伴って特に注目が集まってきている。数式認識は、文字認識と文字の位置関係を解析する数式構造認識で構成され、文字認識と同様にオフライン手法とオンライン手法が存在する。

技術文書等の電子化を目的としたオフライン数式認識の研究は、スキャナによって入力される活字体の数式画像を対象とする。数式の認識は画像から切り出された各文字の空間的な位置関係を、添字や分数などの数式構造の意味関係へと解釈することによって行われる。これまで、2次元文法の構文解析を用いる手法や、記号の位置関係の統計的な性質と構文の知識を用いた手法などが提案されている [1]。これらの手法は、数式表現の複雑さにより計算コストが高くなる傾向があった。しかしながら、岡本ら [3] は数式を垂直・水平方向に再帰的に投影する手法によって効率の良い数式認識機構を実現し、数式認識技術の実用化の可能性が高まっている。

一方、数式を実時間で対話的に計算機へ入力する機構として、記号や構造をメニューから選択する MathType (Design Science Inc., CA, USA) などの数式エディタや $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ などのマークアップ言語がある。しかし、数式は元来人が紙に筆記具を用いて手書きするものであり、タブレット等による数式の手書き入力が可能であれば、計算機へのより直感的な数式入力機構を実現可能である。

このような手書き数式を対象としたオンライン数式

[†] 名古屋大学大学院工学研究科計算理工学専攻, 名古屋市
Dept. of Computational Science and Engineering, Graduate
School of Engineering, Nagoya University, Nagoya-shi, 464-
8603 Japan

^{††} 名古屋大学大学院情報科学研究科メディア科学専攻, 名古屋市
Dept. of Media Science, Graduate School of Information Science,
Nagoya University, Nagoya-shi, 464-8603 Japan

* 現在, キヤノン株式会社画像技術研究所

認識の研究では、手書き文字とともに数式構造を逐次認識することが必要となる。数式認識における手書き文字認識は、文字の切出しにおいて文字同士の位置関係を数式構造としての可能性とともに考慮する必要があり、PDAなどで用いられる手書き文字認識とは異なっている。また、手書き数式は活字体の数式と比べて文字の大きさや位置の変動が大きいという性質や、逐次入力方式では数式全体の情報は利用できないという性質などによって、オンライン数式認識においてオフラインの手法をそのまま適用することは難しい。これまでに開発されているオンラインの手書き数式認識システムは、文字の入力順に制約を設けたり [4], [5], 対象とする数式の範囲を限定して実現されている [4]~[6]。また、入力の実時間性を生かした対話的な修正操作も検討されているが [7]~[9]、一方で現在のところ認識処理のすべてを実時間で行うことが困難なシステムもある [9], [10]。

これらの数式認識の研究では、認識対象を文字の垂直関係^(注1)に対して数式の意味関係を明示的に定義可能な総和や分数などの構造に限定している場合が多い。この条件を満たす数式はすべて階層的な木構造で表現可能であり、比較的扱いやすいためである。しかし、行列構造に現れるような行要素の垂直関係においては、多少の空白が存在するのみで明示的な記号は存在しない。このような数式の中でも特有の性質をもつ行列構造は、多くの研究で認識対象とされておらず [1]、特に我々の知る限りでは行列表現を対象とするオンライン手書き数式認識の研究は存在しない。一方、オフライン数式認識の研究では、比較的大きな括弧を手掛りとして空間的な領域として行列を検出した後に、括弧内に含まれる各要素間の空白を利用して各要素を切り出す手法や [11]、省略表記を含む複雑な行列構造を要素間の位置関係についての連立方程式を解くことによって認識する手法の検討などがある [12]。しかし、オンラインシステムでは行列が逐次構築され、入力中には行列の全体構造を把握することができないため、これらの手法はそのままでは適用できない。

筆者らの研究グループでも実時間で動作する手書き数式認識システムの開発を行ってきたが [13], [14]、行列構造は対象とされてこなかった。そこで、本論文では従来システムにおける認識機能の拡張として行列構造に対する認識機構の追加について述べる。以下、2. で従来システムの数式認識手法について述べる。3. で行列構造認識機構の実現、4. で本システムに対する評

価実験について述べ、5. で考察を加える。

2. 従来の手書き数式認識手法

この章では本論文の基礎をなすこれまで我々が開発を進めてきたオンライン手書き数式認識システム [13] について説明する。システムの概観を図 1 に示す。本手法の基本的な方針は、数式構造として想定する構造 (図 2) すべてについて場合分けを行い、経験的知識を用いてそれぞれの構造を解釈することによって、実時間での逐次認識に適したシステムを実現することである。詳細は文献 [13] を参照されたい。以下の節で図 3 に示す各処理についてそれぞれ述べる。この章で説明する数式認識手法は行列構造を認識対象としておらず、3. で行列認識機構を追加することが本論文の主旨である。

2.1 入力

本システムはタブレット等を用いてシステムの入力

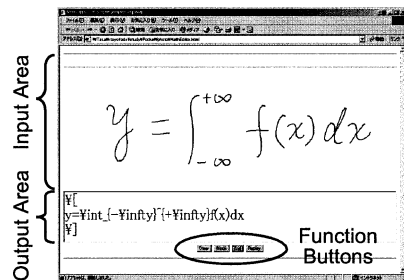


図 1 本システムのウィンドウ画面
Fig. 1 Window picture of this system.

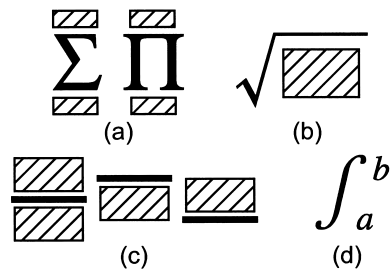


図 2 認識可能な構造 : (a) 総和, 総積 (b) 根号 (c) 分数, 上線, 下線 (d) 添字

Fig. 2 Recognizable structure : (a) summation, production (b) radical sign (c) fraction, upperline, underline (d) adscription.

(注1): 各数式要素の存在する x 座標の範囲に共通部分のある位置関係。例えば、一般に総和記号の上下にはその x 座標の範囲に含まれるような初期条件などを表す数式要素が垂直関係で現れる。

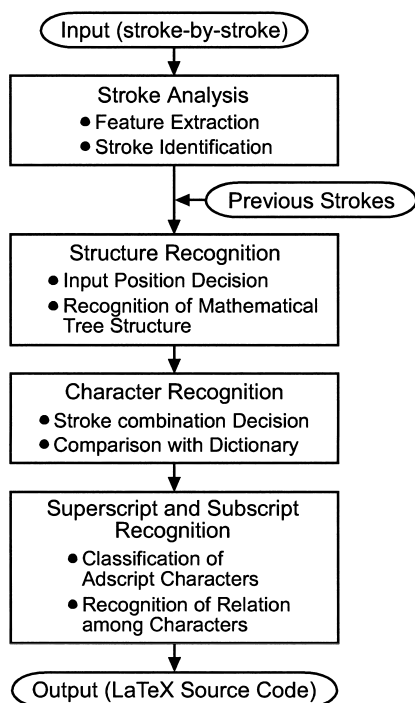


図 3 本システムの処理の流れ
Fig. 3 Flow chart of this system.

領域に 1 ストロークずつ逐次構築される手書きの数式を入力とする。このストロークを入力順に解析して数式を認識し、認識結果を \LaTeX のソースコードとして実時間で提示する。また、本システムの認識機構は文字の入力順序に依存しないため、入力の修正を目的とした任意文字の削除機能を利用した後などにも、任意位置への文字入力が可能である。

2.2 数式構造認識

入力された各ストロークは単独で、若しくは既に入力されているストロークと結合して文字を形成し、文字はその他の文字との位置関係によって数式の構造を構築する。ここでは、図 2 に示す記号がもつ入力ブロック^(注2)を基準としてストロークの入力位置を解析する。これらの記号の入力ブロックが入れ子構造を形成することによって、階層的な木構造をなす一般的な数式が表現される。

2.3 文字認識

入力ストロークを文字として認識するために、まずストロークの特徴をあらかじめ抽出し、仮文字を割り当てておく。本システムでは文字認識には、ストロークの等分割ベクトル列や 8 方向チェーンコードヒス

トグラムなどの特徴量を結合し、一つの特徴ベクトルとして利用している。複数画文字の認識では、各ストロークは仮文字の情報を用いながら 2.2 で同じ入力ブロックとして認識された既入力文字と結合される。この結合は文字パターンの辞書を参照しながら最も可能性の高い組合せとして決定され、それぞれが文字として認識される。入力と辞書との類似度計算には相関係数を用いている。

2.4 添字構造認識

文字認識処理が終了した時点で、それぞれの入力ブロックには文字の添字構造のみが現れ、多項式を形成している。多項式は上付き添字、下付き添字、同一ラインに分類される各文字の位置関係の連なりで構成される。二つの文字の位置関係は文献 [3], [15], [16] と同様に、各文字の外接方形の高さ、幅、中心の位置関係に関する指標を用いたしきい値処理によって決定する。ただし、本手法は右上若しくは右下に付く添字のみを認識対象としており、左上・下付きの添字には対応していない。

本システムでは数式にストロークが追加されるごとに各ブロック内の多項式の添字構造を毎回解析することで、文字の入力順に依存しない多項式の認識機構を実現する。

2.5 出力

本システムの認識結果は、2 次元構造をもつ数式を 1 次元の文字列^(注3)で表現する \LaTeX のソースコードとして各ストロークの入力ごとに実時間で逐次出力される。CPU が PentiumIII 600 MHz の PC での 1 ストローク当りの平均処理時間は約 40 ms である。利用者は \LaTeX を用いて文書を作成する際に、この認識結果を利用することができる。また、この文字列を解析することでその他の数式表現形式に変換することも可能である。認識精度は、現在のところ文字認識率が 80%、構造認識率は 92% であり、実時間での利用には適したシステムとなっている。

3. 行列認識機構の追加

行列構造は、従来システムで認識対象としてきた数式構造とは異なり、各要素間に 2.2 で述べたような位置関係を明示的に表す記号は現れない。行列の各要素は左右を大きな括弧に囲まれて格子状に並んでおり、

(注2): 図 2 の (a)~(c) の構造における斜線を施された方形領域。

(注3): \LaTeX において数式構造はキーワードを用いた文字列として表現される。例えば、分数 $\frac{1}{2}$ は $\text{\frac{1}{2}}$ と記述される。

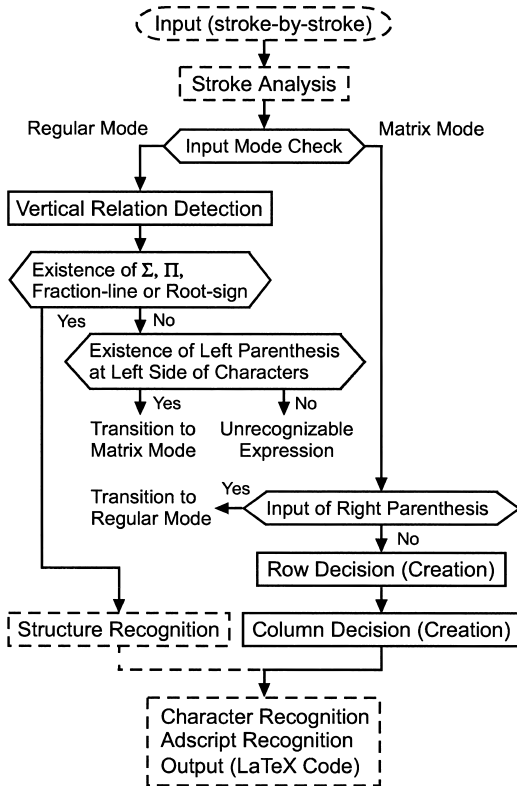


図 4 行列認識機構の処理の流れ：点線部は従来システム
の処理部を利用

Fig. 4 Flow chart of the proposed method for matrix
recognition: Dotted line indicates the proced-
ures of former system.

特に行要素は互いに垂直関係に位置している。このよ
うな位置関係は一般的な数式では現れないため、行列
構造の認識のためには専用の認識機構を開発する必要
がある。今回は行列構造認識の第 1 段階として、

- (1) 格子状に配置される各要素はすべて存在する
- (2) 省略表記は存在しない
- (3) 各要素間には自然な空白が存在する
- (4) 各要素内には添字構造のみが存在する

の条件を満たす最大 N 行 N 列までの行列を認識対
象とする。文字の筆記条件としては、斜体や続け字な
どがなく上記の認識対象構造の条件を満たす程度の
丁寧な入力を想定している。行列認識の目標認識率
は、手書き入力によるパターンの変動などを考慮し
て 95%^(注4)とする。以下で行列認識機構のそれぞれの
処理について詳述する。また、本手法の処理の流れを
図 4 に示す。

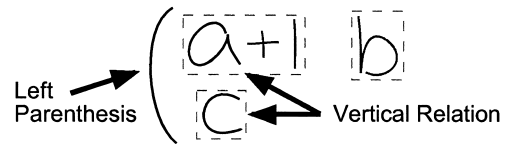


図 5 行列構造の検出
Fig. 5 Detection of the matrix structure.

3.1 行列構造の検出

行列構造を認識するためには、まず数式の逐次入力
において行列構造が入力されていることを検出する必
要がある。従来システムでは行要素間に現れるような
垂直関係が独立に現れた場合は、解釈不能な構造と認
識してエラーを返していた。そこで、このような数式
要素の垂直関係と、一般的な行列構造の幾何学的特徴
を考慮して行列を検出する(図 5)。処理手順は以下の
ようになる。

[step 1] 入力された文字が既入力文字と垂直関係を
なし、かつ、既入力文字が 2.2 で述べた入力ブロック
を介して他の要素と垂直関係を形成可能な記号^(注5)
でない場合は [step 2] へ。それ以外は従来の認識処理を
行う。

[step 2] 垂直関係をなす各要素の左隣の文字が左括弧
であれば行列の入力開始とみなし、以降行列構造専用
の処理を行う(行列モード)。それ以外は [step 3] へ。

[step 3] 左括弧でなければ、左隣の文字に対する添
字構造として認識。左隣に文字が存在しない場合は、
従来と同様にエラーを返す。

3.2 行列構造の認識

行列構造が検出された後の行列モードでは、行列の
格子状の構造を認識するために各格子点に入力ブロッ
クを仮定する。オンラインの入力方式では行列の全体
構造をあらかじめ決定することはできないので、行列
構造内の入力ブロックをストロークの入力位置に基づ
いて動的に生成する必要がある。また、既に存在する
入力ブロックにストロークが入力された場合には、従
来の処理部を利用してブロック内の多項式構造を認識
する。行列構造の認識処理は図 6 に示す幾何学的特徴
を利用しており、以下に処理手順を示す。

[step 1] 既入力の行について、各行に含まれる要素
の外接方形の y 座標中心を平均し、中心線 (Center

(注 4): 図 7(5) に示す行列の入力において誤りが一つ含まれる程度。

(注 5): 本手法では、総和、総積、根号、分数線(上線、下線)の 4 種。

line) を定義．[step 2] へ．

[step 2] 入力ストロークが終了条件を満たさないならば [step 3] へ．ここで、終了条件とは入力ストロークが右括弧かつ、その右括弧の高さ h_r と行列構造検出の基準とした左括弧の高さ h_l との比 $\frac{h_r}{h_l}$ が $0.8 < \frac{h_r}{h_l} < 1.2$ であることとする．終了条件を満たす場合は行列モードを終了し、以降は従来の認識手法による通常モードとなる．

[step 3] ストロークの入力行として、ストロークの外接方形の y 座標中心と行の中心線の y 座標が最も近い行を選択．選択された行が最下行ならば [step 4] へ．それ以外の行ならば [step 5] へ．

[step 4] 入力ストロークの y 座標最小値が最下行の先頭要素の y 座標最大値よりも大きい場合には新たな行を生成する．[step 5] へ．

[step 5] 入力ストロークが入力行の最右要素よりも、しきい値 θ 以上離れている場合には新たな列を生成する．それ以外は [step 6] へ．

[step 6] 入力行におけるストロークの入力列として、ストロークの外接方形の x 座標中心と列の x 座標中心が最も近い既存の列を選択．その行、列に対応する入力ブロックにストロークを追加し、添字構造の認識を行う．

本手法による既存の行列の要素ブロックへの入力にはストロークの入力順に依存しておらず、新たな行や列の生成方法も行優先と列優先の入力方式両方に対応している．ただし、既存の行や列の間への新たな入力ブロックの生成には対応していない．

また、[step 5] の列生成において入力ストロークと入力行の最右要素との距離を比較するしきい値 θ は、

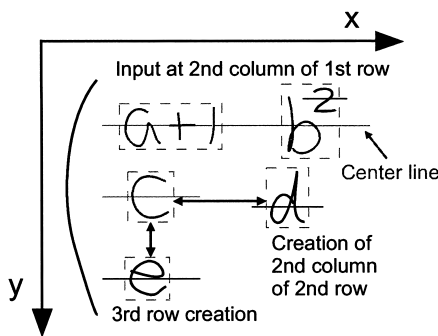


図 6 ブロック生成と入力位置判定

Fig. 6 Criterion of input block creation and input position analysis.

入力ストロークの幅と高さを w_{in}, h_{in} 、最右ブロックの先頭要素の幅と高さを w_{base}, h_{base} としたとき

$$\theta = \frac{1}{2} \max\{w_{in}, h_{in}, w_{base}, h_{base}\} + \alpha$$

と定義される．ここで、しきい値決定の基準として最右ブロックの先頭要素の大きさを利用するのは、添字など小さい文字が入力された場合にも安定したしきい値を求めるためである．また、「1」などの横幅が小さいストロークについては通常の要素と比較して列方向に隣り合う要素との距離が大きくなる性質をもっている．そこで、入力ストローク若しくは最右ブロックの先頭要素が、横幅と比較して縦幅が 3 倍以上である縦に長いストロークの場合、補正項 α としてそれぞれの縦幅の $\frac{1}{8}$ をしきい値に加える．

4. 評価実験

追加した行列認識機構の有効性を確認するために、作成したシステムを使用して評価実験を行った．システムは Java アプレットとして実装されており、入力領域の制限から認識可能な行列のサイズ $N = 10$ としている．評価実験に用いた計算機は CPU が Pentium III 750 MHz の PC であり、入力には液晶ディスプレイ付きタブレット WACOM PL-500 を用いた．実験では図 7 に示す 6 種類の行列を被験者 13 人がそれぞれ 5 回ずつ入力した．手書き入力は人間が目視によって正

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \quad \begin{pmatrix} 0 & 111 & 11 \\ 1 & 21 & 111 \\ 11 & 101 & 11 \end{pmatrix}$$

(1) (2)

$$\begin{pmatrix} a+1 & b+2 \\ c+3 & c+4 \end{pmatrix} \quad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

(3) (4)

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(5)

$$\begin{pmatrix} m_0^2 + n_0^2 & m_0 m_1 + n_0 n_1 \\ m_0 n_0 + m_1 n_1 & m_1^2 + n_1^2 \end{pmatrix}$$

(6)

図 7 実験に用いた行列

Fig. 7 Matrices inputted for the experiment.

しく行列構造を理解できるように、比較的丁寧に行われている。実際に入力されたパターンの例を図 8 に示す。

入力された手書きの行列に対して、本手法における行列構造の認識率と各処理部に費やす時間の平均を求めた。表 1 に認識率に関する実験結果を示す。行列構造の検出率 (Detection Rates) は入力された行列の数に対する 3.1 で述べた行列構造の検出が正しく行われた数の割合である。また、正しく検出された行列の数に対する行列構造が正しく認識された数の割合を行列構造の認識率 (Recognition Rates) とした。この実験結果から、比較的単純な構造の行列の認識については手書き入力によるパターンの変動や個人差を考慮すれば、目標精度はほぼ達成されたと考えられる。

平均処理時間に関しては、行列の検出、行の生成、列の生成の各処理にそれぞれ 23.5 ms, 15.4 ms, 17.5 ms を要した。なお、添字構造のみで構成されている行列要素ブロックへのストロークの追加には 24.1 ms を必要とした。また、行列のサイズによる処理時間の変化はほとんど見られなかった。手書き数式認識システムを実時間で利用するためには、利用者があるストロークを書き終えてから次のストロークを書き始めるまでに認識処理を終了する必要がある。我々の予備実験においては、一般的に利用者が数式を手書き入力する際、

$$\begin{pmatrix} 0 & ||| & || \\ 1 & 21 & ||| \\ || & 10 & || \end{pmatrix} \quad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

図 8 入力パターン例 (正しく認識された例)
Fig. 8 Examples of inputted patterns (recognized correctly).

表 1 行列構造認識の結果
Table 1 Result of matrix structure recognition.

Recognition Targets	Detection Rates	Recognition Rates
Matrix (1)	97% (63/65)	97% (61/63)
Matrix (2)	92% (60/65)	98% (59/60)
Matrix (3)	83% (54/65)	81% (44/54)
Matrix (4)	98% (64/65)	91% (58/64)
Matrix (5)	97% (63/65)	83% (52/63)
Matrix (6)	68% (44/65)	59% (26/44)
Average/Total	89% (348/390)	86% (300/348)

あるストロークを書き終えてから次のストロークを書き始めるまでの時間はほとんどの場合 100 ms 以上であったため、本手法は実時間で利用する目的に十分適している。

5. 考 察

処理時間に関する実験結果より、提案した行列認識機構は実時間で快適に動作することが確認された。また、人間による手書き入力に対する実時間処理の基準 (処理時間 100 ms 以下) からすると、認識結果の出力などの時間を考慮しても認識精度の向上や認識構造の拡張のために更に複雑な処理を付加可能である。したがって、本システムは多くの構造に適用可能で高精度な手書き数式認識システムの基盤になり得ると考えられる。

行列構造においては、各要素間に構造を明らかに示す記号は存在しないため、人が行列構造を理解する場合にも各要素間に存在する空白が重要な役割を担っていると考えられる。本論文における行列認識機構はこの空白 (各要素の距離) を利用して構築されており、比較的単純な構造の行列や丁寧に入力された行列に対しては妥協し得る程度^(注6)の精度で効率的な逐次認識が可能である。これは、特有の構造をもつ行列を含めて、認識対象とする数式構造について場合分けを行い、それぞれの構造に対する専用の機構を用いて認識を行うことの有効性を示している。しかし、行列構造の検出失敗と検出後の行列構造の誤認識は合わせて 23% (90/390) 起こっており、実用に耐え得る精度が達成されているとはいえない。誤認識の例を図 9 に示す。

図 9 の上段の例は行列構造検出時の誤りを示している。これは、逐次認識処理が近傍の数式要素との位置関係に依存しやすいため、行列の第 1 行の要素と第 2 行の要素の位置関係が垂直方向に離れすぎた場合に、行列構造における垂直関係とは認識されず、‘(’に上付きと下付きの添字が付随する構造であると解釈されたためである。しかし、元来数式において左括弧 ‘(’ の右隣に添字が付くことはあり得ない。したがって、今後数式の意味理解などを含めて、行列の検出方法を改善する必要がある。

図 9 の中段と下段は、列方向の分割の失敗例である。中段の例では、計 4 箇所で列方向の分割に失敗し

(注 6): 実時間で利用するシステムであるため、対話的な修正機能の存在を仮定することができる。

$$\begin{pmatrix} a+1 & b+2 \\ c \end{pmatrix}$$

$$\begin{pmatrix} a+1b+2 \\ c \end{pmatrix}$$

→ (L^AT_EX source code)
 $\backslash c^{a+1b+2}$

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

→ (L^AT_EX source code)
 $\backslash begin{array}{cccccccc}$
 $1&1&1&0&1&0&0&0 \\ 0&1&1&1&0&1&0&0 \\ 1&1&0&1&0&0&1&0 \\ 1&0&1&1&0&0&0&1$
 $\backslash end{array}$

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 01 & 1 & 1 & 0 & 1 & 0 & 0 & \\ 11 & 0 & 1 & 0 & 0 & 1 & 0 & \\ 10 & 1 & 10 & 0 & 0 & 1 & & \end{pmatrix}$$

$$\begin{pmatrix} m_0^2 + n_0^2 & m_0 m_1 + n_0 n_1 \\ m_0 n_0 + m_1 n_1 & m_1^2 + n_1^2 \end{pmatrix}$$

→ (L^AT_EX source code)
 $\backslash begin{array}{cccc}$
 $m_0^2&+&n_0^2&m_0m_1+n_0n_1 \\ m_0n_0+m_1n_1&&m_1^2&+&n_1^2$
 $\backslash end{array}$

$$\begin{pmatrix} m_0^2 & + & n_0^2 & m_0 m_1 + n_0 n_1 \\ m_0 n_0 + m_1 n_1 & m_1^2 + & n_1^2 & \end{pmatrix}$$

図 9 行列構造の誤認識例：枠内に認識結果に基づいて整式された行列を表示

Fig. 9 Examples of matrices recognized incorrectly : Typesetting images of recognition results are shown in the boxes.

ているが、局所的な情報のみで判断する場合、分割しない方が自然であると考えられる。これは、局所情報のみに基づいた認識手法の限界を示しており、大域的な情報を併用する必要がある。また、下段の例

では演算子 (+) がその他の文字に比べ小さく入力されたため、結果的に隣り合う要素間との距離が大きくなり列方向に過剰に分割されたものである。この例は、人間が見れば自然な構造の行列として入力されているため、文字認識の結果や大域的な要素の位置関係を併用することによって、認識精度の向上が可能であろう。

ここで示した例における誤認識は、今回実装した行列構造の認識手法において列の決定が行の判定後に各行において独立に行われており、異なる行の列要素の位置関係は考慮されていないことが主要な原因となっている。また、一般の行列表記では要素が省略される場合も多く、複数の行や列にまたがる大きな記号も存在する。したがって、各行と各列が存在する座標の範囲を正確に把握して、各ストロークの入力位置に対応した行や列を厳密に判定する必要がある。特に、手書き数式では利用者の書き癖などによる活字体の数式では見られないような行や列の斜行(図 9 中段参照)への対応が重要になるだろう。

6. むすび

本論文では、実時間で動作する手書き入力による数式入力システムを拡張し、従来の添字、根号、総和、分数などによる数式構造に加え、要素間に特異な垂直関係が現れる行列構造の認識を可能にした。また、評価実験によって本論文で述べた行列認識手法の実時間における有効性を確認した。今後の課題としては、更に複雑な行列構造の認識機構の実現、意味理解まで含めた数式認識機構の検討などが挙げられる。

謝辞 日ごろ、御指導ごべんたつ頂く中京大学情報科学部鳥脇純一郎教授、名古屋大学大学院目加田慶人助教授、並びに熱心に御討論頂いた末永・鳥脇両研究室の諸氏に感謝する。また本研究の一部は文部科学省科学研究費補助金、21 世紀 COE プログラム「社会情報基盤のための音声・映像の知的統合」によった。

文 献

- [1] D. Blostein and A. Grabavec, "Recognition of mathematical notation," in Handbook of Character Recognition and Document Image Analysis, ed. H. Bunke and P. Wang, pp.557-582, World Scientific, Singapore, 1997.
- [2] K.F. Chan and D.Y. Yeung, "Mathematical expression recognition: A survey," International Journal on Document Analysis and Recognition, vol.3, no.1, pp.3-15, Aug. 2000.
- [3] 岡本政行, 東 裕之, "記号のレイアウトに注目した数式構造認識," 信学論 (D-II), vol.J78-D-II, no.3, pp.474-

482, March 1995.

- [4] A. Kosmala and G. Rigoll, "On-line handwritten formula recognition using statistical methods," Proc. Int. Conf. on Pattern Recognition, pp.1306-1308, Aug. 1998.
- [5] R. Fukuda, S. I. F. Tamari, X. Ming, and M. Suzuki, "A technique of mathematical expression structure analysis for the handwritten input system," Proc. Fifth Int. Conf. on Document Analysis and Recognition, pp.131-134, Sept. 1999.
- [6] K.F. Chan and D.Y. Yeung, "*PenCalc* : A novel application of on-line mathematical expression recognition technology," Proc. Sixth Int. Conf. on Document Analysis and Recognition, pp.774-778, Sept. 2001.
- [7] 村瀬敦史, 佐藤 俊, 中川正樹, "オンライン手書き数式認識システム「METHA」の試作," 情処学研報, HI48-4, pp.23-32, May 1993.
- [8] A. Kosmala, G. Rigoll, and A. Brakensiek, "On-line handwritten formula recognition with integrated correction recognition and execution," Proc. Int. Conf. on Pattern Recognition, vol.2, pp.590-593, Sept. 2000.
- [9] S. Smithies, K. Novins, and J. Arvo, "A handwriting-based equation editor," Proc. Graphics Interface '99, pp.84-91, June 1999.
- [10] A. Kosmala, G. Rigoll, S. Lavirotte, and L. Pottier, "On-line handwritten formula recognition using hidden Markov models and context dependent graph grammars," Proc. Fifth Int. Conf. on Document Analysis and Recognition, pp.107-110, Sept. 1999.
- [11] M. Okamoto, H. Imai, and K. Takagi, "Performance evaluation of a robust method for mathematical expression recognition," Proc. Sixth Int. Conf. on Document Analysis and Recognition, pp.121-128, Sept. 2001.
- [12] 金掘利洋, 鈴木昌和, "可変ブロックパターンによる矩形領域分割を用いた行列構造の認識," 信学技報, PRMU2000-201, March 2001.
- [13] 青島史郎, 鈴木隆広, "実時間手書きストローク解析による数式入力システム," 信学論 (D-II), vol.J83-D-II, no.5, pp.1232-1245, May 2000.
- [14] T. Suzuki, S. Aoshima, K. Mori, and Y. Suenaga, "A new system for the real-time recognition of handwritten mathematical formulas," Proc. Int. Conf. on Pattern Recognition, vol.4, pp.515-518, Sept. 2000.
- [15] Z.X. Wang and C. Faure, "Structural analysis of handwritten mathematical expressions," Proc. Int. Conf. on Pattern Recognition, pp.32-34, 1988.
- [16] 江藤裕子, 笹井真樹, 鈴木昌和, "仮想リンクネットワークを用いた数式構文認識," 信学技報, PRMU2000-202, March 2001.

(平成 14 年 5 月 15 日受付, 10 月 3 日再受付)



豊住 健一 (学生員)

2000 名大・工・情報卒。現在同大大学院工学研究科博士課程後期課程在学中。オンライン手書き数式認識, ユーザインタフェースに関する研究に従事。2002 年度電気関係学会東海支部連合大会奨励賞受賞。情報処理学会会員。



鈴木 隆広 (正員)

1999 名大・工・情報卒。2001 同大大学院工学研究科博士課程前期課程了。現在キヤノン(株)勤務。在学中, ユーザインタフェースに関する研究に従事。



森 健策 (正員)

1992 名大・工・電子卒。1996 同大大学院博士課程後期課程了。1994~1997 まで日本学術振興会特別研究員。1997 より名大大学院工学研究科助手, 2000 同大講師。2001 同大難処理人工物研究センター助教授。2001 より 2002 まで米国スタンフォード大学客員助教授。2003 より名古屋大学大学院情報科学研究科助教授, 現在に至る。主に 3 次元画像処理とコンピュータグラフィックスの医用画像への応用に関する研究に従事。日本医用画像工学会奨励賞, 日本エム・イー学会論文賞・坂本賞, 丹羽記念賞, 本会情報システムソサイエティソサイエティ論文賞, 日本気管支学会優秀演題賞各賞受賞。工博。日本エム・イー学会, コンピュータ支援画像診断学会, 日本気管支学会各会員。IEEE Member。



末永 康仁 (正員:フェロー)

1968 名大・工・電子卒。1973 同大大学院博士課程了。同年日本電信電話公社(現 NTT)電気通信研究所に入所。以来, NTT 電気通信研究所において, 一貫して画像処理, 画像認識, 画像生成の研究に従事。1985 年 3 月より 1 年間, 米国マサチューセッツ工科大学メディア研究所客員研究員。NTT ヒューマンインタフェース研究所研究グループリーダー, 主幹研究員を経て, 1997 年 2 月より, 名大大学院工学研究科計算理工学専攻教授。2003 年 4 月より名古屋大学大学院情報科学研究科教授。文部科学省 21 世紀 COE プログラム拠点リーダー。情報処理学会会員。工博。