

## 高次情報を利用した数式文字認識の誤り訂正法の一検討

瀧口 祐介<sup>†</sup> 岡田 稔<sup>††</sup> 三宅 康二<sup>†††</sup>

<sup>†</sup>, <sup>††</sup> 早稲田大学大学院情報生産システム研究科 〒808-0135 北九州市若松区ひびきの 2-7

<sup>†††</sup> 中部大学工学部 〒487-8501 愛知県春日井市松本町 1200

E-mail: <sup>†</sup>tar@asagi.waseda.jp, <sup>††</sup>mokada@waseda.jp, <sup>†††</sup>miyake@cs.chubu.ac.jp

あらまし 我々の研究室では印刷数式の知的二次利用を目的とした、オフライン数式認識理解システム(以下、本システム)の開発を進めている。本システムには、数式構成記号の文字認識を行う低次認識と、数式の二次元的な構造の認識と数学的な意味の理解を行う高次認識の二つの認識過程が含まれる。本論文では、その高次認識過程における文字認識の誤り訂正法を提案する。提案手法では文字認識と数式の構造認識の結果に基づいてコストツリーを生成し、数式に現れるキーワードと隣接する文字の共起確率の2種類の情報を利用することで、文字の誤認識を訂正する。提案手法の有効性を確かめるために幾つかの数式画像を用いて実験を行った結果、記号単位の文字認識率が79.8%から90.2%に、また数式単位の認識率が5.8%から41.1%に改善されることを確認した。

キーワード 数式認識, 高次情報, 文字認識誤り訂正, コストツリー, 共起確率

## A Study on Character Recognition Error Correction Method for Mathematical Formulae using Higher Level Information

Yusuke TAKIGUCHI<sup>†</sup>, Minoru OKADA<sup>††</sup>, and Yasuji MIYAKE<sup>†††</sup>

<sup>†</sup>, <sup>††</sup> Graduate School of Information, Production and Systems, Waseda University

7 Hibikino 2, Wakamatsu-ku, Kitakyushu, 808-0135 Japan

<sup>†††</sup> College of Engineering, Chubu University

1200 Matsumoto-cho, Kasugai, Aichi, 487-8501 Japan

E-mail: <sup>†</sup>tar@asagi.waseda.jp, <sup>††</sup>mokada@waseda.jp, <sup>†††</sup>miyake@cs.chubu.ac.jp

**Abstract** In this paper we propose a method for correcting character recognition errors at the higher level recognition step of the mathematical formulae recognition and understanding system. The system consists of two-level recognition steps: the low level recognition including character recognition, and the higher level recognition including layout recognition. We apply the layout information recognized in the latter step, and character recognition errors are corrected by using two sources of information. One is by some keywords such as mathematical function names, and the other is based on a cost tree and co-occurrence probabilities between symbols. Availability of the proposed method is indicated by some experimental results, and the character recognition rate raised from 79.8% to 90.2% and the formula recognition rate raised from 5.8% to 41.1% are confirmed.

**Key words** mathematical formulae recognition, higher level information, character recognition error correction, cost tree, co-occurrence probability

### 1. はじめに

近年、紙媒体に印刷された書籍や論文(紙文書)のデジタルアーカイブ化が行われるようになった。紙文書をデジタルアーカイブ化する場合、一般にイメージスキャナが用いられ、画像としてコンピュータに取り込まれる。取り込まれた画像はそのままでも閲覧できるが、光学式文字読取装置(OCR: Optical

Character Recognizer)を用いることで、画像中に書かれている文字を読み取ることが可能になる。このような方法で文書に含まれる文字情報までデジタルアーカイブ化すれば、高精度清書や再編集といった閲覧以外の二次利用が可能となる。本研究ではこのような二次利用を特に知的二次利用と呼び、重要視している。前述した事項については、数式を数多く含む科学技術系の文書(科学技術文書)の場合でも同様である。しかしながら、一

般に数式は二次元的な構造を持つため、水平もしくは垂直に一次元的に並んだ文字を認識対象として設計された OCR を用いてこれらを正確に認識することは困難である。

数式の構造認識を行う手法はその入力形態 (オンライン、オフライン) によって、二種類に大別される。オンライン入力とは、マウスやペンタブレット等のヒューマンインタフェースを利用した形態である。この場合、入力の際に得られるストローク (筆順やその軌跡) 情報を利用して文字や数式構造の認識を行う手法が一般的である。これに対してオフライン入力とは、イメージスキャナ等によって画像を入力する形態である。そのためオンライン入力とは異なり、認識にストローク情報を利用することは不可能である。本研究は印刷された書籍や論文を対象とした、オフライン入力を対象とする。オフライン入力による数式を認識対象とした研究は既に幾つか報告されている。岡本らは、大きな記号や空白に注目する基本構造処理と、添字式や根号式を処理する個別構造処理を組み合わせて、印刷数式の構造認識を行う手法 [1] を提案している。しかし彼らは数式の構造認識における文字の誤認識に関する問題については言及していない。現実には、画像中に含まれる全ての文字・記号を正しく認識することは、非常に困難である。この事実から、文字の誤認識に対する訂正処理を備えた数式認識の手法が、幾つか提案されている。Lee らは、ルールベースによる誤り訂正機構を備えたシステム [2] を提案しているが幾つかの特定の文字を訂正対象としているため、不特定に現れる誤認識への対応は難しい。また、鈴木らは認識対象とする文字を、アルファベット類、演算子、インテグラル、総和演算子類の 4 種類に分け、その分類別に得られた特徴量と文字認識の結果得られた正読文字の候補を利用することで、正しい文字と構造を同時に決定する手法 [3] を提案している。しかしながら、この手法では数式構造の候補を削減するために文字認識で得られた幾つかの文字を利用しているため、それらの文字の訂正を行うことが不可能である。

本論文では、筆者らの研究室で開発されたオフライン数式認識理解システム [4], [5] における、文字認識誤りの訂正手法を提案する。提案手法では文字認識処理によって得られた文字・数学記号の候補と、構造認識処理によって文字認識結果とは独立に得られた数式の構造情報に注目する。更に構造中で隣接した文字・記号同士の接続方向を考慮した共起確率と数式に含まれる特定のキーワード情報を利用することで、文字認識の誤りを訂正する。本システムでは数式構造の認識を文字認識の結果を利用せずに行っているため、本提案手法を利用することで、数式画像中に現れる全ての文字・記号の誤認識に対応することが可能である。また幾つかの数式画像を用いた実験により、提案手法の有効性を示す。なお本稿では、数式画像中に現れる全ての文字・記号を単に文字と呼び、文字認識過程で得られる正読文字の候補を候補文字と呼ぶ。

## 2. オフライン数式認識理解システム

数式認識理解システムの目的は、計算機による数式の高次符号化およびその知的二次利用である。本システムでは、入力された数式構造を表す構造ツリーと関数名や演算の優先順位など

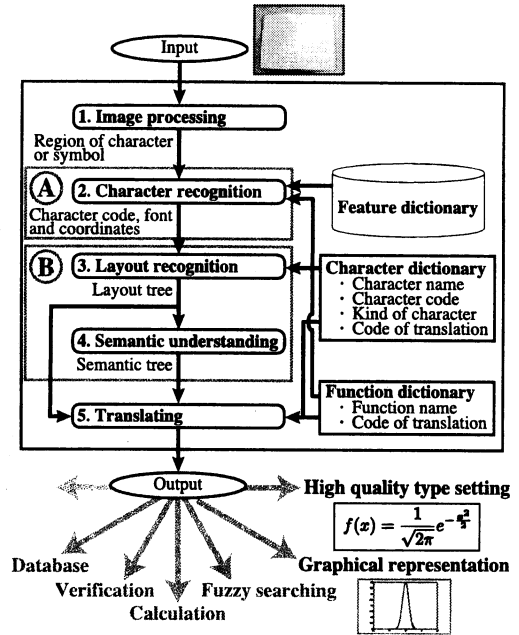


図1 本システムの概要  
Figure 1 Outline of the system.

の数学的な意味を表す意味ツリーの二つを特に高次符号と呼び、これらを生成することをそれぞれ構造認識と意味理解と呼んでいる。本システムの概要を図1に示す。本システムは5つの処理部、1) 画像処理部、2) 文字認識部、3) 構造認識部、4) 意味理解部、5) 出力変換部 (それぞれの数字と図1中の番号は対応) と3つの辞書、特徴量辞書、文字辞書、関数辞書によって構成されている。文字・記号の認識を行う低次認識処理 (図1(A)) と、数式の構造認識・意味理解を行う高次認識処理 (図1(B)) の二つの部分に大別しており、本研究では前者で得られる情報を低次情報、後者で得られる情報を高次情報としている。上掲した処理を含んだ本システム全体の処理の概要を以下に述べる。認識対象として入力された画像に対して、システムはまず、画像処理部でノイズ除去、二値化、ラベル付けといった認識の前処理を行い、画像に含まれる全ての文字領域を切り出す。次に文字認識部では切り出された文字領域に含まれる文字の認識を行い、文字コード、識別順位、相違度、書体、字体 (Typeface)、文字領域を出力する。なお、文字認識部については3.節で詳しく述べる。また構造認識部では、切り出された文字同士の二次元的な接続関係を認識することで、入力画像に含まれた数式の二次元的な構造を表す高次符号である構造ツリーを生成する。さらに生成された構造ツリーに基づいて、意味理解部では、経験的に省略された積記号の挿入、数学関数名や操作名 ('sin', 'lim' など) の認識、連続する数字の統合といった処理を行い、その後、演算子の優先順位に従って構造ツリーを変換することで、数式の数学的な意味を表す高次符号である意味ツリーを生成す

表1 本システムにおける認識対象字種  
Table 1 Recognized characters in our system.

対象とする文字・数学記号		字種数	学習文字数
アルファベット	Roman 体	52	52
	Italic 体	52	86
数学記号		53	53
ギリシャ文字		41	42
数字		10	10
合計		208	243

る。最後に、出力変換部では生成された二種類の高次符号（構造ツリー、意味ツリー）に基づき、様々なソフトウェアのソースコードを出力する。現在、本システムでは $\text{\LaTeX}$ ソースコードとgnuplotスクリプト（前者は数式の高精度清書、後者は数学関数の可視化を目的とする）の出力に対応している。

### 3. 文字認識部

文字認識部では、画像処理部（図1の1.）で切り出された全ての文字領域に含まれる文字を個別に認識する。数式は、アルファベット、数学記号、ギリシャ文字および数字で構成され、それらの書体（Font）や字体（Typeface）と二次元的な配置で数式としての意味を表す。このため本システムの高次処理の一つである意味理解の精度を向上させるために、本処理部では特に字種と書体を含めて認識する。現在、本システムの文字認識部で学習している字種は、アルファベット（大文字と小文字のRoman体及びItalic体） $26 \times 2 \times 2$ 字種、数字10字種、数学記号53字種、ギリシャ文字41字種の全208字種である。認識対象字種と学習文字数を表1にまとめる。文字認識では、三宅らによって開発された加重方向指数ヒストグラム[6]（ $7 \times 7 \times 8 = 392$ 次元）を特徴量として用い、特徴空間における入力パターンと学習パターンの市街区距離によってクラス判別を行う。文字認識では入力された文字領域一つに対して類似度が上位 $n$ 個の候補文字の、文字コード、識別順位、相速度（パターン間の市街区距離）、書体、字体、文字領域（候補文字の外接長方形）を得る。本処理部では入力された全ての文字領域に対して上掲した認識結果を得て、構造認識部に出力する。

### 4. コストツリーと共起確率行列を用いた文字の誤認識の訂正

#### 4.1 コストツリー

本提案法は、高次認識処理の一つである構造認識部の認識結果を利用する。そのため本提案法の処理は、構造認識部の最後で行われる。まず入力画像中の各文字に対応する候補文字の上位5個と構造認識部で得られた構造ツリーに基づき、コストツリーを生成する。コストツリーとは数式における構造情報と候補文字同士の繋がりを示したものである。ツリーのノードは入力画像における文字領域を示し、その内部には複数の候補文字を持つ。特に異なるノードにおける候補文字間の枝に重みを付けることで、その候補文字同士の繋がりの強さを表現する。図2にコストツリーの例を示す。ここで、図2の $N_i$ はツリーの $i$

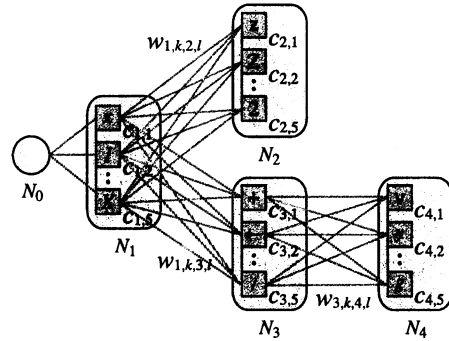


図2 コストツリーの例（ $x^2 + y$ ）

Figure 2 A cost tree for  $x^2 + y$ .

番目のノードを、 $c_{i,k}$ は $i$ 番目のノードに属する識別順位 $k$ 位の候補文字を、 $w_{i,k,j,l}$ は候補文字 $c_{i,k}$ と $c_{j,l}$ の間の結合の重みをそれぞれ表す。ただし、 $N_0$ はツリーの仮想的なルートノードである。次に、これらを利用して候補文字 $c_{i,k}$ と $c_{j,l}$ の間のコストを与える関数 $C(i,k,j,l)$ を式(1)の様に定義する。

$$C(i,k,j,l) = w_{i,k,j,l} D(c_{i,k}, c_{j,l}) \quad (1)$$

$$D(c_{i,k}, c_{j,l}) = \text{dis}(c_{i,k}) + \text{dis}(c_{j,l}) \quad (2)$$

$$w_{i,k,j,l} = 1 - P_{dir}(c_{i,k}, c_{j,l}) \quad (3)$$

ここで、 $D(c_{i,k}, c_{j,l})$ は相速度を与える関数であり、 $\text{dis}(\ast)$ は文字認識部の結果として出力された、候補文字 $\ast$ の相速度を表す。また $w_{i,k,j,l}$ と $P_{dir}(c_{i,k}, c_{j,l})$ はそれぞれ、候補文字 $c_{i,k}$ と $c_{j,l}$ の間の重みとそれらの共起確率である。なおこれらについては4.3節で詳しく述べる。ここで $N_0$ をルートとして、全てのノードを一度のみ通るツリーのコストの総和が最小となるものが、正しい文字認識の結果であることが期待される。この基本概念による誤り訂正処理の流れは、以下の通りである。

(1) コストツリーの水平方向の接続に注目して数式キーワード(4.2節)を探索し、発見した場合はそのキーワードを構成する候補文字の相速度を減小させる。

(2) 文字同士の接続方向を考慮した共起確率を利用してコストツリー内の候補文字間の重み $w_{i,k,j,l}$ を決定する。

(3) コストツリー内の全てのノードを一度だけ通るツリーの中から、コストの総和が最小のものを得る。

また、提案手法による誤り訂正が失敗する要因を以下に挙げる。

- 共起確率の偏り
- 正読文字が候補文字の中に存在しない
- 数式構造の誤認識

#### 4.2 数式におけるキーワード

数式には‘sin’, ‘cos’, ‘tan’等の関数名や、‘lim’, ‘log’等の操作名といったキーワードが多く存在する。本研究では特にこれらを数式キーワードと呼び、文字の誤認識の訂正を行う際の重要な情報のひとつとして扱う。表2に現在のシステムで考慮している数式キーワードの一覧を示す。我々はこれらの数式キーワードを利用して、‘Sin’, ‘CoS’, ‘lim’, ‘log’等の様な入力さ

表2 本システムで考慮している数学キーワードの一覧

Table 2 List of the mathematical keywords considered in our system.

acos	arccos	arcsin	arctan	arg	asin
atan	constant	cos	cosec	cosh.	cot
coth	esc	curl	deg	det	div
lim	exp	gcd	grad	hom	inf
ker	lg	lim	liminf	limsup	ln
log	max	min	rot	sec	sin
sinh	sup	tan	tanh		

□: A symbol of attention

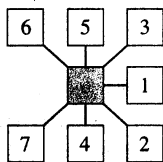


図3 文字の接続方向  
Figure 3 Directions of adjacent symbols.

表3 接続方向別の出現数 ([7]より)

Table 3 Total numbers in the directions from [7].

番号	接続方向	出現数
1	水平	101,133
2	右下	14,446
3	右上	8,440
4	下	4,084
5	上	1,390
6	左上	10
7	左下	2
total		129,505

れた数式画像に含まれる数式キーワード中に生じる文字の誤認識の訂正を行う。具体的には、コストツリー内の水平方向に接続されたノードに注目し、それらのノードに含まれる候補文字の組合せによってキーワードの構成が可能かを調べる。そしてキーワードの構成が可能であれば、キーワードを構成する全ての候補文字の相違度を軽減させる。

#### 4.3 文字の接続方向別の共起確率行列

コストツリーにおける重み  $w_{i,k,j,l}$  を定義するために、隣接文字の接続方向別の共起確率行列を利用する。図3で示す様に、提案法では数式中の文字の接続方向を水平、右下、右上、下、上、左上、左下の計7つと仮定している。共起確率行列  $P_{dir} = \{P_{dir}(c_{i,k}, c_{j,l})\}$  は、式(4)で定義される共起確率を要素に持つ、非対称正方行列である。

$$P_{dir}(c_{i,k}, c_{j,l}) = \frac{\text{num}_{to}(dir, c_{i,k}, c_{j,l})}{\text{num}_{from}(dir, c_{i,k})} \quad (4)$$

ここで  $dir \in \{1, 2, \dots, 7\}$  はコストツリーにおける候補文字  $c_{i,k}$  から  $c_{j,l}$  へのそれぞれの接続方向を表し、 $\text{num}_{from}()$  と  $\text{num}_{to}()$  は、それぞれ接続元の候補文字の出現数と接続元と接続先の候補文字が同時に生じた場合の出現数を表す。

## 5. 実験結果

### 5.1 共起確率行列の生成

鈴木らによって提供されている数式画像データベース (InfyCDB-1) [7] を利用して、数式における7種類の方向別共起確率行列を生成した。InfyCDB-1とは数式画像の正解付きデータベースであり、含まれる文字の総字種数は1,561である。このため、InfyCDB-1によって生成された共起確率行列は1,561行1,561列である。本研究で調査したデータベース中

表4 文字認識率 (入力文字総数: 373文字)

Table 4 Character recognition rate.

識別順位	1	2	3	4	5	6
順位別認識率 [%]	79.6	10.2	5.6	1.3	0.8	0.3
累積認識率 [%]	79.6	89.8	95.4	96.8	97.6	97.9

$$\log \eta = x_1 \log \eta_1 + x_2 \log \eta_2$$

(a) 入力画像 (1,124 × 195 [pixel])

$$1 \ \text{\texttt{0g}} \ \text{\texttt{eta}} = x_{\{1\}} \ 1 \ \text{\texttt{0g}} \ \text{\texttt{eta}}_{\{1\}} + x_{\{2\}} \ 1 \ \text{\texttt{0g}} \ \text{\texttt{eta}}_{\{2\}}$$

$$10g\eta = x_1 10g\eta_1 + x_2 10g\eta_2$$

(b) 認識結果 (誤り訂正無し)

$$\backslash \log \ \text{\texttt{eta}} = x_{\{1\}} \ \backslash \log \ \text{\texttt{eta}}_{\{1\}} + x_{\{2\}} \ \backslash \log \ \text{\texttt{eta}}_{\{2\}}$$

$$\log \eta = x_1 \log \eta_1 + x_2 \log \eta_2$$

(c) 認識結果 (誤り訂正有り)

図4 実験結果 A (成功)

Figure 4 Experimental result A (successful).

の20,097個の数式についての文字同士の接続方向別の出現数を表3に示す。

### 5.2 誤り訂正

約50年前に印刷された物理学の書籍 [8] から、イメージスキャナを用いて解像度600[dpi]で取り込んだ数式画像を用いて実験を行った。なお、実験には5.1節で生成した共起確率行列に経験的なバイアスを加えたものを用いた。本実験に先だって行った文字認識の実験結果を表4に示す。表中の識別順位とは、文字認識処理の結果得られた候補文字の識別順位を表す。表4より識別順位6位における累積認識率の増加率が0.5%以下であったため、本実験では識別順位の上位5個までを利用した。本実験は書籍 [8] から取り込んだ17枚の数式画像、318個の文字について行った。実験の結果、提案手法によって文字認識率が約79.8%から90.2%に、数式単位の認識率が5.8%から41.1%に改善したことを確認した。

成功3例(実験結果A, B, C)と失敗2例(実験結果D, E)について以下に述べる。まず図4に、数式キーワードを構成する文字に誤認識が生じた場合の実験結果Aを示す。ここで図4(a)は幾つかの数式キーワード 'log' を含んでいる。また図4(b), (c)はそれぞれ上から、システムによって生成された $\LaTeX$ ソースコードと、そのソースコードによって生成されたプレビュー画像である。(b)は誤り訂正を行わなかった場合で、(c)は提案手

(a) 入力画像 (711 × 233 [pixel])

$$mV = \sum_a m_a V_a$$

(b) 認識結果 (誤り訂正無し)

$$mV = \sum_a m_a V_a$$

(c) 認識結果 (誤り訂正有り)

図5 実験結果 B (成功)

Figure 5 Experimental result B (successful).

法による誤り訂正を行った場合である。図4(b)において入力画像中の全ての数式キーワード‘log’が‘10g’の様に誤認識されているが、これに対して図4(c)では提案手法によって全ての誤認識が訂正されている。図5は数式中の1文字が誤認識された場合の実験結果Bである。図5(a), (b), (c)は図4の場合と同様にそれぞれ、入力画像と提案手法の有無による結果を示している。図5(a)は総和演算子を含んだ数式の入力画像である。図5(b)では総和演算子‘ $\Sigma$ ’がギリシャ文字‘ $\Sigma$ ’となる文字の誤認識が生じているが、(c)では提案手法によって正しく訂正されている。また図6は、図4や5の様な誤認識が同時に生じた場合の実験結果Cである。図6(a)に示す入力画像は幾つかの分数線と数式キーワード‘sin’等を含んでいる。図6(b)より、 $x$ の指数部の2、分数線、数字の5や数式キーワード‘sin’がそれぞれ、*italic*体のアルファベット‘l’, ‘i’, Roman体のアルファベット‘S’および‘Sin’と誤認識されていることが確認できる。これに対して図6(c)では、全ての誤認識が提案手法によって訂正され、入力画像に含まれる数式が正しく表示されている。

誤認識が正しく訂正されなかった実験結果Dを図7に示す。図7(a)に示す原画像では二つの分数線と数式キーワード‘cos’を含んだ数式が書かれている。次に誤り訂正を行わない結果である図7(b)を見ると、入力画像の右辺側に位置する分数線が‘←’に、またキーワード‘cos’が‘cOS’と誤認識されているのが確認できる。これに対して図7(c)では、数式キーワードの誤認識のみが正しく訂正されており、分数線の誤認識は誤つ

(a) 入力画像 (1,721 × 204 [pixel])

$$-\frac{\pi x^2}{8} + \frac{\pi^2 x}{8} = \sin x + \frac{\sin 3x}{3^3} + \frac{\sin 5x}{5^3} + \frac{\sin 7x}{7^3}$$

(b) 認識結果 (誤り訂正無し)

$$-\frac{\pi x^2}{8} + \frac{\pi^2 x}{8} = \sin x + \frac{\sin 3x}{3^3} + \frac{\sin 5x}{5^3} + \frac{\sin 7x}{7^3}$$

(c) 認識結果 (誤り訂正有り)

図6 実験結果 C (成功)

Figure 6 Experimental result C (successful).

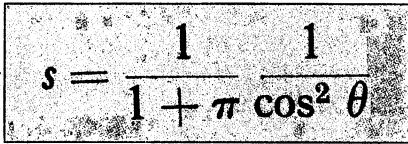
て訂正されている。この原因は、正読文字が誤認識を生じた分数線における文字の認識結果の上位5位以内に存在していなかった(実際には第12位)ためである。この様に候補文字中に正読となる文字が存在していない場合、本提案手法で誤認識を訂正することは不可能である。次に、正しく認識されていた文字を誤って訂正した実験結果Eを図8に示す。図8(a)に示す原画像では総和演算子や数式キーワード‘constant’等が含まれている。ここで図8(b)には、数式中の総和演算子‘ $\Sigma$ ’と変数‘k’および‘constant’がそれぞれ、ギリシャ文字‘ $\Sigma$ ’および‘h’および‘COntant’と誤認識されている。これに対して図8(c)では、提案手法によってこれらの誤認識は正しく訂正できた。しかしながら、図8(b)で正しく認識されていた‘g’が、(c)では‘!’と誤訂正されている。この訂正誤りは連続した3項‘P!’と‘P<sub>g</sub>=’に関する共起確率の逆転、すなわち、

$$D('P', '!')(1 - P_1('P', '!')) + D('!', '=')(1 - P_1('!', '=')) < D('P', 'g')(1 - P_1('P', 'g')) + D('g', '=')(1 - P_1('g', '='))$$

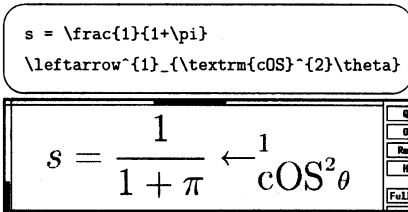
となっていることに起因している。

## 6. まとめ

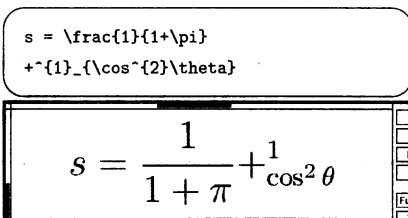
本稿では、オフライン数式認識理解システムの高次認識処理における文字認識の誤り訂正法について述べた。提案法では、構造認識処理で得られた数式の構造情報に基づき、数学関数名等の数式キーワード情報と接続方向別に学習した文字の共起確



(a) 入力画像 (749 × 201 [pixel])



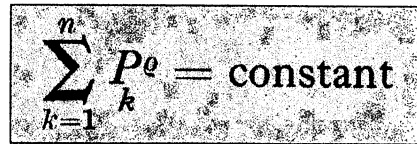
(b) 認識結果 (誤り訂正無し)



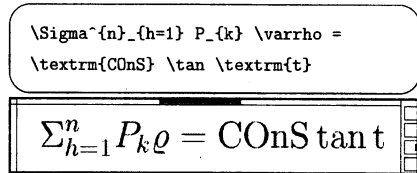
(c) 認識結果 (誤り訂正有り)

図7 実験結果 D (一部失敗)

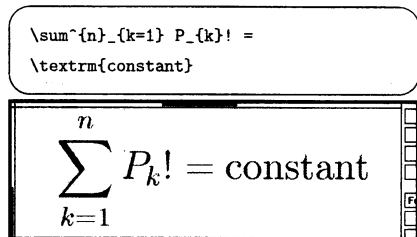
Figure 7 Experimental result D (partially failure).



(a) 入力画像 (720 × 270 [pixel])



(b) 認識結果 (誤り訂正無し)



(c) 認識結果 (誤り訂正有り)

図8 実験結果 E (一部失敗)

Figure 8 Experimental result E (partially failure).

率を要素として持つ共起確率行列を利用した。まず、文字認識結果として得られた候補文字の上位 5 位と数式の構造情報からコストツリーを生成し、その後、コストツリー中の全てのノードを一度だけ通るツリーにおけるコストの総和を計算し、その中で最小のコストとなるものを正しい認識結果であると考えた。幾つかの数式画像を用いた実験の結果、提案法によって記号単位の文字認識率が 79.8% から 90.2% に、数式単位の認識率が 5.8% から 41.1% に改善されたことを確認した。また、総和演算子 'Σ' をギリシャ文字 'Σ' に、分数线を '/' にそれぞれ誤認識した結果や、'lOg', 'Sin', 'cOS', 'COntant' といった数式キーワード中に現れた誤認識等に対して正しく訂正できたことを確認した。しかしながら、共起確率の偏りや候補文字中に正読文字が存在していないことに起因する誤訂正が幾つか確認された。このため、前者に対しては文字認識精度の向上を行い、後者に対しては数式に関する一般的な知識を利用して共起確率を修正することで対応する。なお、現在のシステムにおける文字認識処理では認識対象 1 字種あたり 1 サンプル強しか学習していないため、学習サンプル数を増加させることで、認識精度の改善が見込まれる。また提案手法は数式の構造情報が正しく認識されていることに基づいて設計されているため、文字認識の結果に依存しない高精度な構造認識手法を開発する必要がある。これらの改善が今後の課題である。

## 文 献

- [1] 岡本正行, 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 (1995)
- [2] H. J. Lee and J. S. Wang: “Design of a Mathematical Expression Understanding System”, *Pattern Recognition Letters*, Vol. 18, No. 3, pp. 289-298 (1997)
- [3] Y. Eto and M. Suzuki: “Mathematical Formula Recognition using Virtual Link Network”, *Proc. of ICDAR 2001 - 6th Int'l Conf. on Document Analysis and Recognition*, pp. 762-767 (2001)
- [4] Y. Chen and M. Okada: “Structural Analysis and Semantic Understanding for Offline Mathematical Expressions”, *Int'l J. of Pattern Recognition and Artificial Intelligence*, Vol. 15, No. 6, pp. 967-987 (2001)
- [5] Y. Takiguchi, M. Okada and Y. Miyake: “A Fundamental Study of Output Translation from Layout Recognition and Semantic Understanding System for Mathematical Formulae”, *Proc. of ICDAR 2005 - 8th Int'l Conf. on Document Analysis and Recognition*, pp. 745-749 (2005)
- [6] 鶴岡信治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二: “加重方向指数ヒストグラム法による手書き漢字・ひらがな認識”, 信学論 D, Vol. 70-D, No. 7, pp. 1390-1397 (1987).
- [7] M. Suzuki, S. Uchida and A. Nomura: “A Ground-truthed Mathematical Character and Symbol Image Database”, *Proc. of ICDAR 2005 - 8th Int'l Conf. on Document Analysis and Recognition*, pp. 675-679 (2005)
- [8] D. H. Menzel: “Fundamental Formulas of Physics”, Prentice-Hall, Inc., (1955)