

## 数式の構造認識と意味理解

岡田 稔 中部大学工学部情報工学科教授

### 1 はじめに

書籍、学術論文などの文書画像を対象とした文書理解 (Document Analysis) は実用志向のパターン認識研究の一分野として注目されている。文書理解の研究では、その構成要素 (文字、文章、図、画像、表など) への分解、構成要素自体の認識理解が行われてきたが、特に科学技術論文などでは、数式や化学式は文書を構成する重要な要素であるにも関わらず、その認識の研究は立ち遅れていると言わざるを得ない。

本研究では、文書に含まれる数式の画像を入力とし、その数式を認識・理解し、符号化された数式情報を出力するためのシステムを構築することを目的とする。ここでいう数式の認識とは、数式の構成要素 (文字・記号) のいわゆる文字認識を含む、それらの文字形式 (書体、書式、大きさなど) とそれらの間の位置関係、などのレイアウト構造を符号化することをいい、特に構造認識という。また、数式の理解とは、数式の文脈的意味、数式を構成する要素の数学的意味までを対象として符号化することをいい、特に意味理解という。具体的には、数式の構造認識によって清書が可能となり、数式の意味理解によってその数式の計算が可能となる。たとえば、 $z = \sin x \cos y$  という数式では、構造認識プロセスによってその数学的意味は不明であるが、 $z, =, s, i, n, x, c, o, s, y$  の文字列からなることとそれらの配置が判り、清書が可能となる。次に意味理解プロセスによって  $x$  は4つの変数  $s, i, n, x$  の積ではなく、変数  $x$  の正弦関数であることがわかり、計算が可能となる。構造認識による清書では、タイプセッティング言語として地位を確立している  $\text{T}_\text{E}_\text{X}$ 、 $\text{L}^\text{A}_\text{T}_\text{E}_\text{X}$  ソースリスト形式の出力、意味理解による計算では同様に Mathematica スリスト形式の出力を想定している。本研究では、第一段階として高校の数学教科書レベルの単独の数式を想定しているが、第二段階として技術論文などに見られる専門的な数式群も対象とする。

### 2 国内外の関連研究と本研究の位置付け

数式認識に関する文献上に現れた最初の研究論文は 1968 年の R. H. Anderson によるもの [1] であると考えられる。その後、散発的に研究が行われ、最近では国際会議などでしばしば数式認識に関する論文が発表されるようになってきた。しかし、多くは手法の提案と簡単な実験報告だけであり、共通のデータベース等での評価がないために優劣を比較することは困難である。また、初期の研究論文は、2次元的な文法の構文解析や、グラフ書き換えなどを用いたものなど、実効速度の面で実用化には適さないものであった。

Zhao, Sugiura らは数理知識に基づく、積み木構成的な数式作成システムを提案した。これらは数式を如何に簡単にコンピュータに入力するかを目的としたオンライン入力支援システムであり、論文、書籍等からのオフラインによる数式符号化とデータベース化を目的とした本研究とは性格が異なるものである。信州大学の岡本は実効速度と認識精度の双方から実用化可能なアルゴリズムを発表した。九州大学の鈴木は岡本の研究を継承し、数式構文解析が文字の誤認識や類文字の影響を受けない方法を示した。また、ドイツのエッセン大学では、数学論文をイメージデータとしてデータベース化し、再利用可能なかたちで保存するプロジェクトが開始されている。

最近になって、コンピュータで数式を入力するためのオンライン支援システムとして構成的組み立て法による杉浦ら、ペン入力型による末永らの研究がある。

このように高度な理解を伴った数式の符号化、及び符号化数式データベースの構築を目的とした研究は報告されていない。また、数式の持つ曖昧性や経験・感性による記述と解釈の違いに関する検討も全くなされていない。

このように、数式の単なる清書を目的としたいわゆるレイアウト認識はいくつか報告があるが、本研究は書体等をも含めた詳細認識による高精度タイプセッティング、数式の意味理解による再計算も目的としたものである。しかし、本研究では、印刷済み (あるいは手書き済み) の書籍・論文を対象とするオフライン認識・理解が対象であり、その目的と方法・戦略は大きく異なる。その意味で本研究はすでに発表済みの論文によって学界から評価を得ている。

なお、数式認識の研究は、on-line / off-line の区別、印刷 / 手書きの区別により、以下の4つに大別できる。

1. off-line 手書き  
手書き原稿の清書を主目的とする。
2. off-line 印刷  
過去の論文・書籍の符号化を主目的とする。
3. on-line 手書き  
数式の簡易入力インタフェースを主目的とする。
4. on-line 印刷  
demand的な論文・書籍の符号化を主目的とする。

本研究は 2 に属し、この目的を持つ類似研究は僅少である。

### 3 システムの考え方と基本構成

#### 3.1 数式の構造認識と意味理解の考え方

文書理解は書籍、雑誌、などの符号化に重要なものとされている。数式は数学公式集、自然科学系の学术论文に多く含まれているが、その認識研究は文書理解の要素技術としては不十分である。本研究の目的はスキャナーなどによって画像として入力された数式をオフラインでその構造の認識・理解を行うことである。

本研究では、数式の構成要素 (文字・記号) のいわゆる文字認識を含む、それらの文字形式 (書体、書式、大きさなど) とそれらの間の位置関係、などのレイアウト構造を符号化することをいい、特に構造認識という。また、数式の理解とは、数式の文脈的意味、数式を構成する要素の数学的意味までを対象として符号化することをいい、特に意味理解という。構造認識によってその構造を表す木、すなわち構造ツリー (layout tree) を生成し、意味理解によって数式の意味を表す木、すなわち意味ツリー (semantic tree) を生成する(図1)。

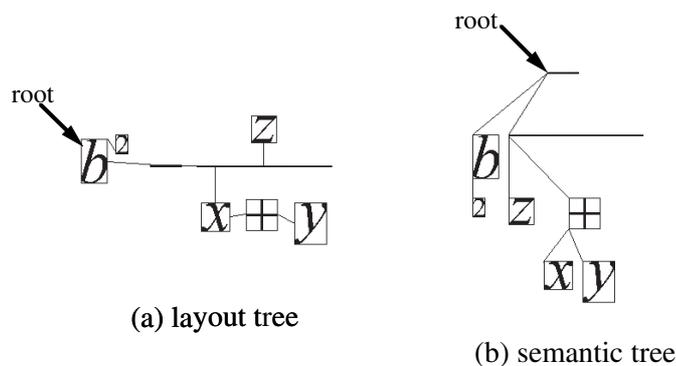


図1：構造ツリーと意味ツリー

#### 3.2 本システムの基本構成

本研究では、図2に示すような6要素からなる基本構成により、数式構造の認識と理解を行うシステムを構成した。各別にまとめると以下ようになる。

- 画像処理部: 画像の入力、ノイズ除去、切り出しなどの認識前処理。
- 文字認識部: 文字・記号の認識を行う。併せて文字タイプの認識も行う。
- 構造認識部: 文字認識の結果に基づき、構造ツリーを作成。
- ルールベース部: 数学ルール、経験ルール、感性ルールから構成される。
- 意味理解部: 構造ツリーとルールベースにより意味理解し、意味ツリーを作成。
- 変換部: 構造ツリーからTeX ソース。意味ツリーからMathematica ソースを出力。

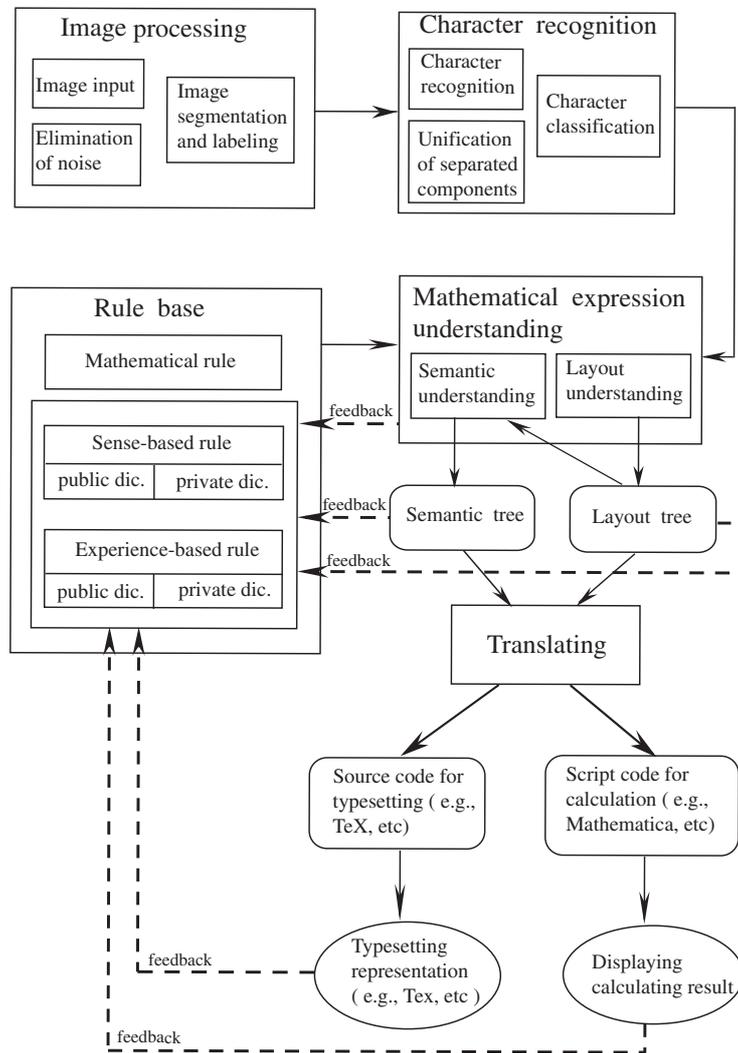


図2：システム構成の概要

#### 4 実験

構築したシステムにおいて、基礎実験を行った。対象とする数式を以下に示す。

$$x + \frac{y + z^2 - l}{h - d} \frac{f}{w \times s} + y^4 \quad (1)$$

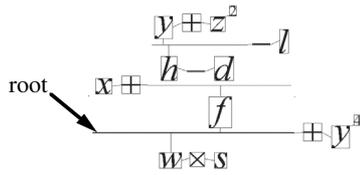
$$z = \frac{\arccos x}{x^2 + y^2} + \int \log x dx \quad (2)$$

$$\frac{x^a}{\arctan x} + \sum_{n=1}^k (n + \log n) \quad (3)$$

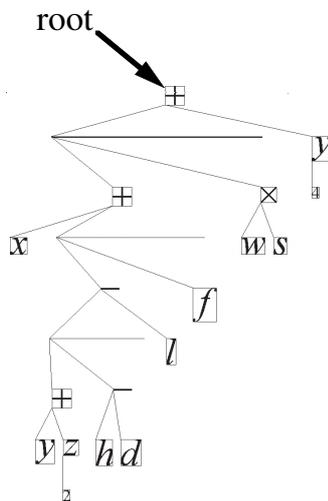
図3は式1の処理結果である。同図において(a)は構造認識の結果であり、レイアウトツリー(構造ツリー)を示している。また、(b)は意味理の結果であり、セマンティックツリー(意味ツリー)を示している。また、図4、図5はそれぞれ、式2、式3の処理結果である。このほか、高等学校の数学教科書から任意抽出した数式について実験した結果、成功率はほぼ90%であった。失敗例は、本研究では対象としなかった行列、ベクトルなどであり、今後の対応は検討課題である。

$$\frac{\frac{y+z^2}{x+\frac{h-d}{f}}-l}{w \times s} + y^4$$

(a)原画像



(b)生成された構造ツリー

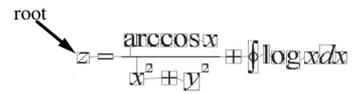


(c)生成された意味ツリー

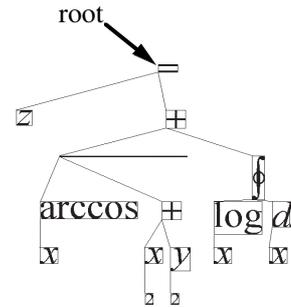
図3：実験結果1

$$z = \frac{\arccos x}{x^2 + y^2} + \int \log x dx$$

(a)原画像



(b)生成された構造ツリー



(c)生成された意味ツリー

図4：実験結果2

## 5 数式の曖昧性

数式を表記する規則、すなわち数式の記述文法は完備なものではなく、日本語、英語と同様な自然言語であると考えられる。それによって適切に書かれていないひとつの数義性、曖昧性を有する。例として、

$$z = \sin x \cos y \tag{4}$$

の数学的意味は普遍的に一義的ではない。数式4の右辺:  $\sin x \cos y$  は、一般的には  $(\sin x) (\cos y)$  と理解されるが、 $\sin(x \cos y)$  という解釈も可能である。なぜなら他方、数式(2):  $\sin xy$  は、 $(\sin x) (y)$  ではなく、一般的には  $\sin(xy)$  と理解される。数式(1),(2)の一般的解釈においては、項の積と、関数名とその独立変数の結合及びそれらの優先順位が、一義的には定められないという一例である。他の例として  $z = f(x+y)$  を示す。この数式では、 $f$  がスカラー変数なのかあるいは関数なのかは単独の数式のみでは判断できず、数式の前後関係あるいはテキストとしての本文などによる文脈・コンテキストによってしか推定できない。これらは、数式を言語として捉えたとき、その文法が曖昧であることを示している。

また、数式の構成法は AMS (American Mathematical Society) において定められた清書標準 (明文化された文法規約は存在しないが、現実的に  $AMST_{E}X$  としての標準がある) があるものの、公刊書籍、学術論文などでは必ずしも従われていない(例えば、 $\sin x$  とすべきところを  $\sin x$ ,  $\sin x$  とするなど)。また、数式の(書き手の)構成法・清書法と(読み手の)解釈法には、国、分野 (純粋数学、電気工学、機械工学など) によってもその違いがあることが事前調査で判明している。本研究では数式の曖昧性、多様性についても詳細に検討し、ルールベースの構築に資する。

$$\frac{x^a}{\arctan x} + \sum_{n=1}^k (n + \log n)$$

(a)原画像

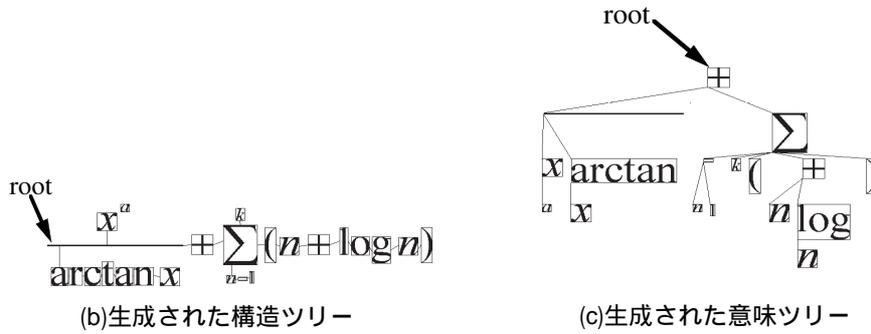


図5：実験結果3

表1：数式構成要素の異なる解釈

高等学校での数学教科	具体例
集合論	$A, B, C$ (集合を表す)
確率・統計	$A, B, C$ (事象を表す)
三角関数	$A, B, C$ (角度を表す)
行列	$A, B, C$ (行列を表す)
平面幾何	$m, n$ (直線を示す)
二次関数	$m, n$ (定数を表す)
指数関数	$m, n$ (指数部を表す)

### 5.1 レイアウト依存と非依存

数式の持つ曖昧性は、レイアウト依存とレイアウト非依存の曖昧性に大別できる。数式を構成する要素の位置関係によって生じる曖昧性をレイアウト依存の曖昧性、位置関係ではなく、構成文字の組み合わせによるものをレイアウト非依存の曖昧性という。レイアウト依存の曖昧性は、手書き数式では頻繁に出現し、印刷数式においてもタイプセットのミスなどにより出現する。

表2：レイアウト依存の曖昧性の例

曖昧性の種類	例	出現数
平方根記号の範囲	$\sqrt{a + bx}$	31
平方根記号の範囲	$\sqrt{a + b/x}$	9
平方根記号の基数	$x^{a+b} \sqrt{y}$	23
類似した関数名	$\sinh x$	792

表3：レイアウト非依存の精曖昧性の例

曖昧性の種類	例	出現数
算術演算 (乗算)	$\sin x \cos y$	700
算術演算 (割算)	$\sin x / \cos y$	44
算術演算 (割算)	$\sin(a_p)/(y_p)$	1
割算	$x/yz$	31
階乗	$n!m!$	
階乗	$2^{n+1}n!$	
$\Sigma$	$\sum_{k=0}^n k \sin x + b$	

## 5.2 曖昧性の調査

本研究のひとつの目的は、オフライン印刷数式の理解である。このため、既存の印刷数式における曖昧性を調査することは極めて重要である。本研究ではその対象として岩波数学公式集I, II, IIIに基づき、これらの曖昧性がどの程度あるのかを調査した。表2には、レイアウト依存の曖昧性の出現数、表3には、レイアウト非依存の曖昧性の出現数、をそれぞれ示す。

## 6 まとめ

本研究では、数式の構造認識と意味理解を行うシステムを構築することそのものを目的としている。一方で、本システムで作成された数式の符号化データは、利用価値が高い。特に数学公式集などの符号化データベースは、オンデマンドで使用するためのデータベースコンテンツとして計り知れない実用性、応用性、付加価値がある。また、付帯価値として、公式の検算が可能となり、データベースの信頼性も向上するであろう。公式集などは過去の著名なものから複写転載がくり返され、数十年経ってから誤植が発覚したこともあると聞く。目視による校正に代わる自動的数式検算も可能である。一方、化学式も数式と並んで科学技術文書のコンテンツとして重要であり、その構造の認識・理解も意義深いものである。本研究では数式理解の詳細検討と共に化学式の認識・理解の基礎検討も行ったが、将来的には詳細検討と実装、システム構築を図りたい。

## 7 謝辞

本研究は以下の学生諸君、名古屋大学大学院工学研究科・Yun CHEN、清水智巨、中部大学工学部工業物理学科・富永年彦、山内公英、前島良彦、木下祥一、(敬称略)の協力無しには遂行できえなかった。ここに記して謝辞に代える。

## 参考文献

- [1] R. H. Anderson, "Syntax-Directed Recognition of Hand-printed Two-dimensional Mathematics," *Interactive Systems for Experimental Applied Mathematics*, pp. 436-459, Academic Press (1968).
- [2] M. Okamoto and H. M. Twaakyondo, "Mathematical Expression Recognition by Projection Characteristics," *IEICE Trans.*, Vol. J78-D-II, No. 2, pp. 366-370 (1995).
- [3] H. M. Twaakyondo and M. Okamoto, "Structure Analysis and Recognition of Mathematical Expression," *Proc. of ICDAR-3; the Third Int'l Conf. on Document Analysis and Recognition*, pp. 430-437, 1995.
- [4] M. Okamoto and H. Higashi, "Mathematical Expression Recognition by the Layout of Symbol," *IEICE Trans.*, Vol. J78-D-II, No. 3, pp. 474-482 (1995).
- [5] Y. A. Dimitriadis and J. L. Coronado, "Towards an Art based Mathematical Editor, That Uses On-line Handwritten Symbol Recognition," *Pattern Recognition*, Vol. 28, No. 6, pp. 807-822 (1995).
- [6] R. J. Fateman and T. Tokuyasu, "Optical Character Recognition and Parsing of Typeset Mathematics," *Visual Communication and Image Representation*, Vol. 7, No. 1, pp. 2-15 (Mar. 1996).
- [7] E. G. Miller and P. A. Viola, "Ambiguity and Constraint in Mathematical Expression Recognition," *Proc. of the Fifteenth National Conf. on Artificial Intelligence*, pp. 784-791 (1998).
- [8] H. J. Winkler and M. Lang, "Symbol Segmentation and Recognition for Understanding Handwritten Mathematical Expressions," *Progress in Handwriting Recognition* pp. 407-412, Academic Press (1996).
- [9] H. J. Lee and J. S. Wang, "Design of a Mathematical Expression Recognition System," *Proc. of ICDAR-3; the Third Int'l Conf. on Document Analysis and Recognition*, pp. 1084-1087 (1995).
- [10] H. J. Lee and M. C. Wang, "Understanding Mathematical Expressions using Procedure-oriented Transformation," *Pattern Recognition*, Vol. 27, No. 3, pp. 447-457 (1994).
- [11] F. Kimura and M. Shridhar, "Handwritten Numerical Recognition based on Multiple Algorithms," *Pattern Recognition*, Vol. 24, No. 10, pp. 969-983 (1991).
- [12] A. Beláid and J. P. Haton, "A syntactic approach for handwritten mathematical formula recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 1, pp. 105-111 (1984).
- [13] Y. Nakayama, "A prototype pen-input mathematical formula editor," *Proc. of EDMEDIA 93-World Conference on Educational Multimedia and Hypermedia*, pp. 400-407 (1993).
- [14] F. Kimura, T. Wakabayashi, S. Tsuruoka and Y. Miyake, "Improvement of Handwritten Japanese Character Recognition Using Weighted Direction Code Histogram," *Pattern Recognition*, Vol. 30, No. 8, pp. 1329-1337 (1997).

[ 15 ] H. Tota, " The Newest Mathematics Textbook in High School (6 Volumes) ", *Keirinkan Press* (1998).

## 8 本研究の成果による公開出版物

- [ 1 ] 清水智巨, 陳うん, 岡田稔: " 数式表記ルールに着目した数式構造理解 ", 平成11年電子情報通信学会総合大会, (1999-3)
- [ 2 ] 清水智巨, 陳うん, 岡田稔: " 曖昧さに注目した数式構造理解 ", 電子情報通信学会技術研究報告, パターン認識とメディア理解研究会, PRMU99-49, pp. 1-8 (1999-7)
- [ 3 ] Y. Chen, T. Shimizu and M. Okada: " A Fundamental Study on Mathematical Expression Understanding Based on Rule Base ", *Technical Report of IEICE, SIG-Pattern Recognition and Media understanding*, PRMU99-50, pp. 9-16 (Jul. 1999)
- [ 4 ] 清水智巨, 陳うん, 岡田稔: " 曖昧さを考慮したオフライン数式構造理解 ", 平成11年度電気関係学会東海支部連合大会, No. 767 (1999-9)
- [ 5 ] 陳うん, 清水智巨, 岡田稔: " ルールベースに基づく数式レイアウトと意味理解の検討 ", 平成11年度電気関係学会東海支部連合大会, No. 768 (1999-9)
- [ 6 ] Y. Chen, T. Shimizu and M. Okada: " Fundamental Study on Structural Understanding of Mathematical Expressions ", *Proc. of SMC '99 - 1999 IEEE Systems, Man, and Cybernetics Conference*, Vol. II, pp. 910-914 (Oct.1999, Tokyo International Forum/Tokyo/Japan)
- [ 7 ] 富永年彦, 清水智巨, 陳ウん, 岡田稔: " 化学式の構造認識・理解のための基礎検討 ", 第60回情報処理学会全国大会, 3ZB-03 (2000-3)
- [ 8 ] 清水智巨, 陳ウん, 岡田稔: " 文書画像における数式構造理解の一手法 ", 第60回情報処理学会全国大会, 3ZB-01 (2000-3)
- [ 9 ] Y. Chen, T. Shimizu and M. Okada: " Structure and Semantic Understanding for Mathematical Expressions ", *The 60th Annual Conf. of IPS. Japan*, 3X-05 (Mar. 2000)
- [ 10 ] Y. Chen, T. Shimizu, M. Okada, and K. Yamauchi: " Ambiguous problem investigation in off-line mathematical expression understanding ", *Proc. of SMC ' 2000 - 2000 IEEE Systems, Man, and Cybernetics Conference*, Vol. 4, pp. 2917-2922 (Oct. 2000, Nashville/Tennessee/USA)
- [ 11 ] Y. Chen, T. Shimizu and M. Okada: " Mathematical Expression Structure Understanding ", *Proc. of VSMM ' 99 - 5th International Conference on Virtual Systems and MultiMedia 1999*, (Sept. 1999, Dundee/UK)
- [ 12 ] 前島良彦, 陳うん, 岡田稔: " 化学式の構造認識・理解の一手法 ", 情報処理学会第62回全国大会, 慶応義塾大学, 1M-3 (2001)
- [ 13 ] 木下祥一, 陳うん, 岡田稔: " 数式認識・理解システムにおける文字認識部の検討 ", 情報処理学会第62回全国大会, 慶応義塾大学, 1M-4 (2001)
- [ 14 ] Y. Chen, T. Shimizu, and M. Okada: " Structural Analysis and Semantic Understanding for Off-Line Mathematical Expressions ", *IJPRAI - Int ' l J. of Pattern Recognition and Artificial Intelligence*, Vol. 15, No. 8, (Aug, 2001)

## 9 本研究に関連した受賞

- 1. 清水智巨: " 第60回情報処理学会全国大会学生奨励賞 "( 上記8- [ 8 ] による ), (社) 情報処理学会 (2000-3)
- 2. Yun CHEN: " 電子情報通信学会東海支部学生研究奨励賞 ", (社) 電子情報通信学会東海支部学生会 (2001-6)

< 発 表 資 料 >

題 名	掲載誌・学会名等	発表年月
数式表記ルールに着目した数式構造理解	平成11年電子情報通信学会総合大会	1999年3月
曖昧さに注目した数式構造理解	電子情報通信学会技術研究報告 パターン認識とメディア理解研究会 PRMU99-49,PP.1-8	1999年7月
A Fundamental Study on Mathematical Expression Understanding Based Rule Base	Technical Report of IEICE, SIG-Pattern Recognition and Media understanding, PRMU99-50, PP. 9-16	1999年7月
曖昧さを考慮したオフライン数式構造理解	平成11年度電気関係学会東海支部連合大会、No.767	1999年9月
ルールベースに基づく数式レイアウトと意味理解の検討	平成11年度電気関係学会東海支部連合大会、No768	1999年9月
Fundamental Study on structural Understand of Mathematical Expressions	Proc. of SMC 99-1999 IEEE Systems, Man, and Cybernetics Conference, Vol. II pp. 910-914	1999年10月
化学式の構造認識・理解のための基礎検討	第60回情報処理学会全国大会,3ZB-03	2000年3月
文書画像における数式構造理解の一手法	第60回情報処理学会全国大会,3ZB-01	2000年3月
Structure and Semantic Understanding for Mathematical Expressions	The 60th Annual Conf. of IPS. Japan, 3X-05	2000年3月
Ambiguous problem investigation in off-line mathematical expression understanding	Proc. of SMC' 2000-2000 IEEE Systems, Man, and Cybernetics Conference, Vol. 4, pp. 2917-2922	2000年10月
Mathematical Expression Structure Understanding	Proc. of VSMM '99-5th International Conference on Virtual Systems and MultiMedia 1999	1999年9月
化学式の構造認識・理解の一手法	情報処理学会第62回全国大会、慶応義塾大学、1M-3	2001年3月
数式認識・理解システムにおける文字認識部の検討	情報処理学会第62回全国大会、慶応義塾大学、1M-4	2001年3月
Structural Analysis and Understanding for off-Line Mathematical Expressions	IJPRAI-Int'l J. of Pattern Recognition and Artificial Intelligence	2001年8月