

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6694638号
(P6694638)

(45) 発行日 令和2年5月20日(2020.5.20)

(24) 登録日 令和2年4月22日(2020.4.22)

(51) Int. Cl. F I
G06K 9/62 (2006.01) G06K 9/62 G
G06K 9/00 (2006.01) G06K 9/00 P

請求項の数 6 (全 15 頁)

<p>(21) 出願番号 特願2016-570658 (P2016-570658)</p> <p>(86) (22) 出願日 平成28年1月19日 (2016.1.19)</p> <p>(86) 国際出願番号 PCT/JP2016/051457</p> <p>(87) 国際公開番号 W02016/117564</p> <p>(87) 国際公開日 平成28年7月28日 (2016.7.28)</p> <p>審査請求日 平成30年12月4日 (2018.12.4)</p> <p>(31) 優先権主張番号 特願2015-9207 (P2015-9207)</p> <p>(32) 優先日 平成27年1月21日 (2015.1.21)</p> <p>(33) 優先権主張国・地域又は機関 日本国 (JP)</p> <p>前置審査</p>	<p>(73) 特許権者 504132881 国立大学法人東京農工大学 東京都府中市晴見町3-8-1</p> <p>(74) 代理人 100090398 弁理士 大淵 美千栄</p> <p>(74) 代理人 100090387 弁理士 布施 行夫</p> <p>(72) 発明者 中川 正樹 東京都府中市晴見町3-8-1 国立大学 法人東京農工大学内</p> <p>(72) 発明者 ファン ミン カン 東京都府中市晴見町3-8-1 国立大学 法人東京農工大学内</p> <p style="text-align: right;">最終頁に続く</p>
---	--

(54) 【発明の名称】 プログラム、情報記憶媒体及び認識装置

(57) 【特許請求の範囲】

【請求項1】

手書き入力されたストローク列から文字列を含む数式を認識するためのプログラムであって、

入力されたストローク又はストローク列が、文字列と数式とを区分するための特定ジェスチャに該当するか否かを判定し、前記特定ジェスチャに該当すると判定した場合に、前記特定ジェスチャで指定される範囲に関する情報を登録する登録部と、

入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを数式に対応するストロークとして設定する区分部と、

文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、1つの前記特定ジェスチャと当該特定ジェスチャで指定される範囲に含まれるストロークとを1つのシンボルとして扱って、当該シンボルとして扱う当該特定ジェスチャの座標情報と、数式に対応するストロークとして設定されたストロークの情報とから、数式認識エンジンにより数式を認識する認識部としてコンピュータを機能させ、前記区分部は、

入力されたストローク又はストローク列が前記特定ジェスチャに該当すると判定された場合に、当該特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、入力されたストロークが前記特定ジェスチャに該当しないと判定された場合に、当該入力されたストロークが前記特定ジェスチャで指定される範囲に含

まれるか否かを判定し、前記特定ジェスチャで指定される範囲に含まれる場合に、当該入力されたストロークを文字列に対応するストロークとして設定することを特徴とするプログラム。

【請求項 2】

請求項 1 において、

入力されたストローク列を表示部に表示させる制御を行う表示制御部として更にコンピュータを機能させ、

前記表示制御部は、

前記特定ジェスチャに該当すると判定されたストローク又はストローク列を、前記特定ジェスチャに対応する形状を表す画像として表示させる制御を行うことを特徴とするプログラム。

10

【請求項 3】

請求項 2 において、

前記表示制御部は、

前記特定ジェスチャで指定される範囲に含まれるストロークを、前記特定ジェスチャで指定される範囲に含まれないストロークとは異なる色で表示させる制御を行うことを特徴とするプログラム。

【請求項 4】

請求項 1 乃至 3 のいずれか 1 項において、

前記登録部は、

入力されたストローク又はストローク列が略矩形を形成する場合に、当該ストローク又はストローク列が前記特定ジェスチャに該当すると判定することを特徴とするプログラム。

20

【請求項 5】

コンピュータ読み取り可能な情報記憶媒体であって、請求項 1 乃至 4 のいずれか 1 項のプログラムを記憶したことを特徴とする情報記憶媒体。

【請求項 6】

手書き入力されたストローク列から文字列を含む数式を認識する認識装置であって、

入力されたストローク又はストローク列が、文字列と数式とを区分するための特定ジェスチャに該当するか否かを判定し、前記特定ジェスチャに該当すると判定した場合に、前記特定ジェスチャで指定される範囲に関する情報を登録する登録部と、

30

入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを数式に対応するストロークとして設定する区分部と、

文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、1つの前記特定ジェスチャと当該特定ジェスチャで指定される範囲に含まれるストロークとを1つのシンボルとして扱って、当該シンボルとして扱う当該特定ジェスチャの座標情報と、数式に対応するストロークとして設定されたストロークの情報とから、数式認識エンジンにより数式を認識する認識部とを含み、

前記区分部は、

40

入力されたストローク又はストローク列が前記特定ジェスチャに該当すると判定された場合に、当該特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、入力されたストロークが前記特定ジェスチャに該当しないと判定された場合に、当該入力されたストロークが前記特定ジェスチャで指定される範囲に含まれるか否かを判定し、前記特定ジェスチャで指定される範囲に含まれる場合に、当該入力されたストロークを文字列に対応するストロークとして設定することを特徴とする認識装置。

【発明の詳細な説明】

【技術分野】

【0001】

50

本発明は、プログラム、情報記憶媒体及び認識装置に関する。

【背景技術】

【0002】

従来、数式を入力する手法として、L a T e Xなどの専門言語を用いる方法や、数式入力エディタを用いる方法等が用いられてきたが、前者は専門用語を習得する必要があり、後者はメニュー選択と記号選択を繰り返す必要があり、ともに一般人には使い勝手が悪かった。これに対して、タブレットなどに手書きされた数式をコンピュータにより認識させる方法は古くから研究されてきた。そして、長年の研究により認識率は向上しており、また、近年におけるタブレット型PCの急速な普及から、一般人に浸透する可能性も増している。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2009-80615号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

数式は、シンボルと、四則演算子、括弧、分数記号やルート、積分記号、べき乗などを表す記号の位置や大きさの関係からなる。手書き数式認識システムでは、シンボルとして英数字記号を仮定するのが普通であるが、実際には、シンボルとして英数字記号以外の文字列が書かれる場合も多く、従来の手書き数式認識システムでは、数式に含まれる文字列が認識されないという問題があった。こうした文字列を含む数式は、教科書のなかで良く現れるし、黒板に板書されることも多い。

20

【0005】

本発明は、以上のような課題に鑑みてなされたものであり、その目的とするところは、文字列を含む数式を認識することが可能なプログラム、情報記憶媒体及び認識装置を提供することにある。

【課題を解決するための手段】

【0006】

(1)本発明は、手書き入力されたストローク列から文字列を含む数式を認識するためのプログラムであって、入力されたストローク又はストローク列が、文字列と数式とを区分するための特定ジェスチャに該当するか否かを判定し、前記特定ジェスチャに該当すると判定した場合に、前記特定ジェスチャで指定される範囲に関する情報を登録する登録部と、入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列及び数式的一方に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを文字列及び数式の他方に対応するストロークとして設定する区分部と、文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、数式に対応するストロークとして設定されたストロークから数式認識エンジンにより数式を認識する認識部としてコンピュータを機能させるためのプログラムに関する。また、本発明は、コンピュータ読み取り可能な情報記憶媒体であって、上記各部としてコンピュータを機能させるためのプログラムを記憶した情報記憶媒体に係る。また、本発明は、上記各部を含む認識装置に係る。

30

40

【0007】

本発明によれば、入力されたストロークが特定ジェスチャに該当するか否かを判定し、特定ジェスチャで指定される範囲に含まれるストロークを文字列及び数式的一方に対応するストロークとして設定して、特定ジェスチャで指定される範囲に含まれないストロークを文字列及び数式の他方に対応するストロークとして設定し、文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、数式に対応するストロークとして設定されたストロークから数式認識エンジンにより数式を認識することで、数式に含まれる文字列と数式とを明確に区分してそれぞれを認識することがで

50

きる。なお、特定ジェスチャを用いずに、数式と文字列を自動的に分離（区分）する方法も考えられる。しかしながら、自動的に数式と文字列を区分することは困難であり、誤分割は避けられないため、本発明では、文字列と数式とを区分するための特定ジェスチャを導入している。

【0008】

(2) また本発明に係るプログラム、情報記憶媒体及び認識装置では、前記区分部は、入力されたストローク又はストローク列が前記特定ジェスチャに該当すると判定された場合に、当該特定ジェスチャで指定される範囲に含まれるストロークを文字列及び数式の一方に対応するストロークとして設定し、入力されたストロークが前記特定ジェスチャに該当しないと判定された場合に、当該入力されたストロークが前記特定ジェスチャで指定される範囲に含まれるか否かを判定し、前記特定ジェスチャで指定される範囲に含まれる場合に、当該入力されたストロークを文字列及び数式の一方に対応するストロークとして設定してもよい。

10

【0009】

本発明によれば、文字列（又は数式）を筆記してから特定ジェスチャを筆記する場合であっても、特定ジェスチャを筆記してから文字列（又は数式）を筆記する場合であっても、文字列と数式とを区分することができ、ユーザの利便性を向上することができる。

【0010】

(3) また本発明に係るプログラム、情報記憶媒体及び認識装置では、前記区分部は、入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを数式に対応するストロークとして設定し、前記認識部は、1つの前記特定ジェスチャで指定される範囲に含まれるストロークを1つのシンボルとして扱って、数式認識エンジンにより数式を認識してもよい。

20

【0011】

本発明によれば、1つの特定ジェスチャで指定される範囲に含まれるストローク（文字列に対応するストロークとして設定されたストローク）を1つのシンボルとして扱って数式認識エンジンにより数式を認識することで、文字列を認識できない数式認識エンジンを用いて文字列を含む数式を認識することができる。

【0012】

(4) また本発明に係るプログラム及び情報記憶媒体では、入力されたストローク列を表示部に表示させる制御を行う表示制御部として更にコンピュータを機能させ、前記表示制御部は、前記特定ジェスチャに該当すると判定されたストローク又はストローク列を、前記特定ジェスチャに対応する形状を表す画像として表示させる制御を行ってもよい。

30

【0013】

また本発明に係る認識装置では、入力されたストローク列を表示部に表示させる制御を行う表示制御部を更に含み、前記表示制御部は、前記特定ジェスチャに該当すると判定されたストローク又はストローク列を、前記特定ジェスチャに対応する形状を表す画像として表示させる制御を行ってもよい。

【0014】

本発明によれば、特定ジェスチャに該当すると判定されたストロークを、特定ジェスチャに対応する形状を表す画像として表示させることで、特定ジェスチャが正しく判定されたことをユーザに認識させることができ、ユーザの利便性を向上することができる。

40

【0015】

(5) また本発明に係るプログラム、情報記憶媒体及び認識装置では、前記表示制御部は、前記特定ジェスチャで指定される範囲に含まれるストロークを、前記特定ジェスチャで指定される範囲に含まれないストロークとは異なる色で表示させる制御を行ってもよい。

【0016】

本発明によれば、特定ジェスチャで指定される範囲に含まれるストロークを、特定ジェ

50

スチャで指定される範囲に含まれないストロークとは異なる色で表示させることで、特定ジェスチャにより文字列と数式とが正しく区分されたことをユーザに認識させることができ、ユーザの利便性を向上することができる。

【0017】

(6) また本発明に係るプログラム、情報記憶媒体及び認識装置では、前記登録部は、入力されたストローク又はストローク列が略矩形を形成する場合に、当該ストローク又はストローク列が前記特定ジェスチャに該当すると判定してもよい。

【図面の簡単な説明】

【0018】

【図1】 図1は、本実施形態の認識装置の機能ブロック図の一例である。

10

【図2A】 図2Aは、数式と文字列及び特定ジェスチャが混在した筆記の一例を示す図である。

【図2B】 図2Bは、特定ジェスチャテーブルの一例を示す図である。

【図3A】 図3Aは、特定ジェスチャが筆記される順番について説明するための図である。

。

【図3B】 図3Bは、特定ジェスチャが筆記される順番について説明するための図である。

。

【図3C】 図3Cは、特定ジェスチャが筆記される順番について説明するための図である。

。

【図3D】 図3Dは、特定ジェスチャが筆記される順番について説明するための図である

20

。

【図4A】 図4Aは、矩形の特定ジェスチャの筆記の一例を示す図である。

【図4B】 図4Bは、矩形の特定ジェスチャの筆記の一例を示す図である。

【図5】 図5は、逐次方式における処理の一例を示すフローチャートである。

【図6A】 図6Aは、逐次方式における具体的な表示例を示す図である。

【図6B】 図6Bは、逐次方式における具体的な表示例を示す図である。

【図6C】 図6Cは、逐次方式における具体的な表示例を示す図である。

【図6D】 図6Dは、逐次方式における具体的な表示例を示す図である。

【図6E】 図6Eは、逐次方式における具体的な表示例を示す図である。

【図6F】 図6Fは、逐次方式における具体的な表示例を示す図である。

30

【図7】 図7は、一括方式における処理の一例を示すフローチャートである。

【図8A】 図8Aは、一括方式における具体的な表示例を示す図である。

【図8B】 図8Bは、一括方式における具体的な表示例を示す図である。

【図8C】 図8Cは、一括方式における具体的な表示例を示す図である。

【発明を実施するための形態】

【0019】

以下、本実施形態について説明する。なお、以下に説明する本実施形態は、特許請求の範囲に記載された本発明の内容を不当に限定するものではない。また本実施形態で説明される構成の全てが、本発明の必須構成要件であるとは限らない。

【0020】

40

1. 構成

図1に本実施形態の認識装置の機能ブロック図の一例を示す。なお本実施形態の認識装置は図1の構成要素(各部)の一部を省略した構成としてもよい。

【0021】

文字入力部160は、ユーザが筆記媒体(ペン、指先等)で手書き文字を入力するためのものであり、その機能は、タブレット、タッチパネル等の筆記面などにより実現できる。文字入力部160は、筆記媒体が筆記面に触れてから離れるまでの筆記媒体の位置を表す座標データを一定時間間隔で検出し、検出された座標データ列(座标点系列)をストローク(筆画)のデータとして処理部100に出力する。なお、ストロークの終点から次のストロークの始点までのベクトルをオフストローク(運筆ベクトル)と呼び、ストローク

50

とオフストロークの連続する系列をストローク列と呼ぶ。

【0022】

記憶部170は、処理部100の各部としてコンピュータを機能させるためのプログラムや各種データを記憶するとともに、処理部100のワーク領域として機能し、その機能はハードディスク、RAMなどにより実現できる。

【0023】

表示部190は、処理部100で生成された画像を出力するものであり、その機能は、文字入力部160としても機能するタッチパネル、LCD或いはCRTなどのディスプレイにより実現できる。

【0024】

処理部100（プロセッサ）は、文字入力部160からの座標データやプログラムなどに基づいて、認識処理、表示制御などの処理を行う。この処理部100は記憶部170内の主記憶部をワーク領域として各種処理を行う。処理部100の機能は各種プロセッサ（CPU、DSP等）、ASIC（ゲートアレイ等）などのハードウェアや、プログラムにより実現できる。処理部100は、登録部110、区分部112、認識部114、表示制御部120を含む。

【0025】

登録部110は、入力されたストローク又はストローク列が、文字列と数式とを区分するための特定ジェスチャに該当するか否かを判定し、前記特定ジェスチャに該当すると判定した場合に、前記特定ジェスチャで指定される範囲（領域）に関する情報を登録する処理を行う。登録された情報は記憶部170に記憶される。ここで、登録部110は、入力されたストローク又はストローク列が略矩形を形成する場合に、当該ストローク又はストローク列が前記特定ジェスチャに該当すると判定してもよい。

【0026】

区分部112は、入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列及び数式の一方に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを文字列及び数式の他方に対応するストロークとして設定する。すなわち、区分部112は、入力されたストロークのうち、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定し、前記特定ジェスチャで指定される範囲に含まれないストロークを数式に対応するストロークとして設定してもよいし、その逆でもよい。

【0027】

認識部114は、区分部112により文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、区分部112により数式に対応するストロークとして設定されたストロークから数式認識エンジンにより数式を認識する処理を行う。ここで、区分部112が、前記特定ジェスチャで指定される範囲に含まれるストロークを文字列に対応するストロークとして設定する場合、認識部114は、1つの前記特定ジェスチャ及び当該特定ジェスチャで指定される範囲に含まれるストロークを1つのシンボルとして扱って、数式認識エンジンにより数式を認識してもよい。例えば、図2Aに示す例において、文字列「area」、「height」、「width」を、それぞれ「x」、「y」、「z」のような1つのシンボルとして扱って、数式認識エンジンにかけて認識させてもよい。

【0028】

表示制御部120は、入力されたストローク列を表示部190に表示させる制御を行う。ここで、表示制御部120は、前記特定ジェスチャに該当すると判定されたストローク又はストローク列を、前記特定ジェスチャに対応する形状を表す画像（例えば、矩形を表す画像）として表示させる制御を行ってもよい。また、表示制御部120は、前記特定ジェスチャで指定される範囲に含まれるストロークを、前記特定ジェスチャで指定される範囲に含まれないストロークとは異なる色で表示させる制御を行ってもよい。また、表示制御部120は、認識部114により認識された文字列を含む数式（認識結果）を表示部1

10

20

30

40

50

90に表示させる制御を行う。

【0029】

2. 本実施形態の手法

次に本実施形態の手法について図面を用いて説明する。

【0030】

手書き入力された文字列を含む数式をコンピュータに認識させる場合、理想的には、入力されたストローク列（手書きパターン）から自動的に数式と文字列を分離（区分）し、それぞれを認識することが好ましい。しかしながら、そもそも自動的に数式と文字列を区分することは困難であり、不完全な区分のうえに数式と文字列をそれぞれ認識させようとすると認識率が著しく低下する恐れがある。

10

【0031】

そこで、本実施形態の手法では、文字列と数式とを区分するためのジェスチャを導入する。これを特定ジェスチャと呼ぶ。ここで、特定ジェスチャとは、認識対象となる文字や数式以外の記号を筆記すると、当該ジェスチャで指定される範囲に含まれるストロークを特定して、それに応じた処理を行う操作を意味する。

【0032】

本実施形態の手法では、特定ジェスチャとして、文字列を囲む矩形を採用する。すなわち、手書き入力されたストローク（又はストローク列）が略矩形を形成する場合に、当該ストロークが特定ジェスチャに該当すると判定し、当該特定ジェスチャで指定される範囲に含まれるストロークから文字列を認識し、当該特定ジェスチャで指定される範囲に含まれないストロークから数式を認識する。特定ジェスチャの形状として矩形を採用するのは、矩形が普通の数式には現れず、また、範囲を指定するために矩形を筆記することが人間にとって自然だからである。なお、数学のある種の領域では、正方形や長方形がシンボルとして使われることがあるが、その場合は、矩形の代わりに別の形状のシンボルを特定ジェスチャとして採用すればよい。すなわち、特定ジェスチャは、形状がユニークで（他のシンボルや記号と区別が付き易く）、それが指示する範囲が明確であればどのようなものでもよい。例えば、特定ジェスチャの形状は、矩形の対角線の片方或いは両方を右から左に筆記する（左から右に筆記すると「X」と区別がつかなくなるため）ものでもよいし、矩形の対角線の両方を続けて筆記するもの（「」を逆さにした形状）でもよい。また、ここでは、数式に含まれる文字列を区分するために特定ジェスチャを用いるが、文字列に含まれる数式を区分するために特定ジェスチャ（例えば、数式を囲む矩形等）を用いてもよい。

20

30

【0033】

図2Aは、数式と文字列及び特定ジェスチャが混在した筆記の一例を示す図である。登録部110は、ストロークが入力される毎に、そのストロークが特定ジェスチャであるか否かを監視し、特定ジェスチャであると判定した場合に、当該特定ジェスチャの筆順（特定ジェスチャに該当する筆画の順番）と座標（矩形の4角の座標、或いは矩形の対向する2角の座標）を、特定ジェスチャテーブルに登録する。図2Bは、図2Aの筆記例に対応する特定ジェスチャテーブルを示す図である。例えば、図2Bに示す特定ジェスチャテーブルでは、特定ジェスチャBG₁について、4画の文字列「area」の後の5画目に筆記されたことと、その左上角及び右下角の座標（特定ジェスチャで指定される範囲に関する情報の一例）が登録され、特定ジェスチャBG₂について、1番目の「x」の後の13画目に筆記されたことと、その左上角及び右下角の座標が登録され、特定ジェスチャBG₃について、2番目の「x」の後の18画目に筆記されたことと、その左上角及び右下角の座標が登録されている。区分部112は、図2Bに示すような特定ジェスチャテーブルを参照して、いずれかの特定ジェスチャに囲まれたストロークを文字列の一部（文字列に対応するストローク）として区分し、それ以外のストロークを数式の一部（数式に対応するストローク）として区分する。

40

【0034】

図2Aに示す例では、特定ジェスチャを1画で筆記した場合を示しているが、特定ジェ

50

スチャを複数画で筆記することを許容してもよい。この場合には、処理の都合上、特定ジェスチャを筆記する最大画数（例えば、4画）を規定し、また、特定ジェスチャを筆記し始めてから筆記し終わるまでは、数式或いは文字に対応するストロークを筆記しないことを条件とする。この程度の条件であれば、ユーザにとって大きな制約とはならない。また、特定ジェスチャを複数画で筆記することを許容する場合には、特定ジェスチャを構成する最初のストロークの筆順と最後のストロークの筆順の2つを特定ジェスチャテーブルに登録する。なお、図2Bの特定ジェスチャテーブルにおける「調整後筆順」については後述する。

【0035】

本実施形態の手法では、特定ジェスチャがどのタイミングで筆記されても、文字列と数式とを区分する処理が実行されるようにしている。特定ジェスチャを筆記する順番を指定すると、ユーザにとって大きな制約となり、使い勝手が悪くなるからである。図3に示すように、特定ジェスチャが筆記される順番には4つの場合が考えられる。図3Aは、最初に特定ジェスチャBGを筆記して、その次に、特定ジェスチャBGの中に文字列を筆記する場合を示し、図3Bは、最初に文字列を筆記して、その次に、当該文字列を囲む特定ジェスチャBGを筆記する場合を示し、図3Cは、最初に文字列を筆記して、その後他の筆画を筆記した後に、当該文字列を囲む特定ジェスチャBGを筆記する場合を示し、図3Dは、既に筆記された文字列を囲む特定ジェスチャBGの中に文字列を書き足す場合を示す。

【0036】

図4に、矩形の特定ジェスチャの筆記の一例を示す。ユーザは特定ジェスチャBGを筆記する場合、矩形の左上角から書き始めて、時計周り（図4A参照）、或いは反時計回り（図4B参照）に書き進めて左上角に戻る。区分部112は、入力されたストロークを構成する筆点の少なくとも一部（例えば、1/3以上）が特定ジェスチャに含まれる場合に、当該ストロークが当該特定ジェスチャに囲まれる（当該特定ジェスチャで指定される範囲に含まれる）と判定する。これは、ユーザが筆記した特定ジェスチャが他のストロークと交差する場合がある（図3参照）ことを考慮したものである。

【0037】

本実施形態の手法では、特定ジェスチャを高精度に認識するために、ストロークの時系列情報を使うオンライン認識手法と、ストロークの画像としての特徴を使うオフライン認識手法を併用して特定ジェスチャを認識する。特定ジェスチャを認識する際の認識対象は、数式（英文字などのシンボル、四則演算子、分数記号、ルート、積分記号などの演算子、括弧）と特定ジェスチャ（囲み矩形）となる。

【0038】

オンライン認識手法では、MRF（Markov Random Field）モデルを用いる。まず、入力パターンを標準サイズに正規化し、Ramer-Douglas-Peucker法によって、端点や角などの特徴点を抽出する。そして、MRFモデルの特徴点と伸縮マッチングを行う。スコアの評価には、閾値を用いて、閾値より小さい場合はリジェクトする。

【0039】

オフライン認識手法では、MQDF（Modified Quadratic Discriminant Function）を用いる。ストロークは、時系列順に筆点列をつなぐことで、容易に画像に変換することができる。ストロークを画像に変換すると時系列情報を失うものの、画像に変換することで筆順違いや重ね書き等による影響を排除することができる。この画像に対して、Sobelフィルタにより勾配特徴を抽出して8方向に射影する。元画像が64×64ピクセルとして、これを同一サイズの8×8の区画に分割する。これに対して18×18ピクセルのガウスフィルタを適用してぼかす。このようにすると、中心に近い方向特徴は大きく寄与し、周辺の寄与は小さくなり、隣の区画からも多少の寄与を受け、これにより位置ずれに強くなる。その結果、8×8区画ごとに8方向の特徴が抽出され、全体で512次元の特徴が得られる。更に、フィッシャーの判別分析により次元を圧縮して、256次元の特徴を得る。識別にはMQDFを使う。更に、閾値を設定し、入力パターンの矩形への距離が最短

10

20

30

40

50

でも、閾値より大きい場合はリジェクトする。

【0040】

最後に、オンライン認識手法とオフライン認識手法の結果を統合する。ここで、オンライン認識手法は確率値の対数值（負値であるが、大きいほど良い）を出し、オフライン認識手法は距離（小さいほど良い）を出すため、これらの尺度を合わせる必要がある。そこで、オンライン認識のスコアに - 1 を乗じることで正值にして（従って、小さいほどよい）、重み付の和をとることで統合スコアを算出する。統合スコア $score_{combination}$ は、次式により算出される。

【0041】

【数1】

$$score_{combination} = (-1) \times w_1 \times score_{online} + w_2 \times score_{offline} \quad (1)$$

【0042】

ここで、 $score_{online}$ 、 $score_{offline}$ は、それぞれオンライン認識手法のスコア、オフライン認識手法のスコアである。また、 w_1 、 w_2 は、統合の重みであり、これらは訓練パターンを用いて最適化される。より詳細には、通常、 $score_{online}$ は、- 10 程度から 0 までの値をとり、 $score_{offline}$ は、0 以上 1024 未満の値をとる。すなわち、 $-score_{online}$ は $score_{offline}$ の 100 分の 1 程度なので、現実には、 $0 < w_1 < 0.1$ 、 $w_2 = 1 - w_1$ の条件を満たすときに最適な値が得られる。登録部 110 は、認識対象のなかで、 $score_{combination}$ が最も小さいものを特定ジェスチャとして認識する。

【0043】

3. 処理

次に、本実施形態の認識装置の処理の流れについて説明する。認識装置には、タブレット型 PC のように、時系列の筆点を実時間で表示して認識のフィードバックを筆画（ストローク）毎に行える対話型装置と、時系列の筆点をメモリを備えたタブレットや電子ペンで記録し、そのデータの読み込みと処理は PC 等で一括してなされる非対話型装置がある。

【0044】

3 - 1. 逐次方式の処理

対話型装置で採用される逐次方式では、ユーザは、文字列を認識する言語（例えば、英語か日本語）を指定してシステム（プログラム）を立ち上げ、タッチパネル等の入力部に数式や文字列を筆記し始める。システムは、特定ジェスチャの認識、文字列及び数式の区分をストロークが筆記される毎に行い、筆記終了後に、文字列及び数列の認識と認識結果の表示を行う。図 5 は、逐次方式における処理の一例を示すフローチャートである。

【0045】

まず、処理部 100 は、文字入力部 160 に入力されたストロークを取得する（ステップ S10）。このとき、表示制御部 120 は、取得されたストロークを表示部 190 に表示させる制御を行う。次に、登録部 110 は、取得されたストロークから遡って特定ジェスチャの最大画数（例えば、4 画）までの最近のストロークが特定ジェスチャに該当（特定ジェスチャを構成）するか否かを判断する（ステップ S12）。例えば、登録部 110 は、最近の 1 画分のストローク、2 画分のストローク列、3 画分のストローク列、4 画（最大画数）分のストローク列のそれぞれについて統合スコア $score_{combination}$ を算出し、認識対象のなかで、特定ジェスチャが一番小さい $score_{combination}$ を与えるとき、特定ジェスチャと認識する。但し、オンライン認識手法とオフライン認識手法のいずれかでリジェクトされたストロークについては、 $score_{combination}$ を算出せず、いずれのストロークもリジェクトされた場合には、特定ジェスチャに該当しないと判断する。

【0046】

特定ジェスチャに該当する場合（ステップ S12 の Y）には、登録部 110 は、特定ジェスチャに該当すると判断されたストロークの情報から、特定ジェスチャの情報（筆順と、対向する 2 角の座標）を特定ジェスチャテーブルに登録する（ステップ S14）。ここ

10

20

30

40

50

で、登録部110は、特定ジェスチャに該当すると判断されたストロークが、以前に数式に対応するストロークとして設定されていた場合には、その設定を取り消す。また、表示制御部120は、特定ジェスチャに該当すると判断されたストロークの画像を、他のストロークとは異なる色（例えば、赤色）の矩形（特定ジェスチャに対応する形状）を表す画像として表示させる制御を行う。これにより、特定ジェスチャが正しく認識されたことをユーザにフィードバックすることができる。

【0047】

次に、区分部112は、特定ジェスチャテーブルを参照して、登録された特定ジェスチャで指定される範囲に含まれるストロークを、文字列に対応するストロークとして設定する（ステップS16）。このとき、表示制御部120は、特定ジェスチャで指定される範囲に含まれるストロークの色を変化（例えば、黒色から青色に変化）させる制御を行う。これにより、特定ジェスチャによって文字列が正しく区分されたことをユーザにフィードバックすることができる。なお、特定ジェスチャで指定される範囲に含まれるストロークが存在しない場合には、ステップS16の処理をスキップする。

【0048】

特定ジェスチャに該当しない場合（ステップS12のN）には、区分部112は、特定ジェスチャテーブルを参照して、取得されたストロークが、いずれかの特定ジェスチャで指定される範囲に含まれるか否かを判断する（ステップS18）。取得されたストロークがいずれかの特定ジェスチャで指定される範囲に含まれる場合（ステップS18のY）には、区分部112は、取得されたストロークを、文字列に対応するストロークとして設定し（ステップS20）、表示制御部120は、取得されたストロークの色を変化（例えば、黒色から青色に変化）させる制御を行う。一方、取得されたストロークが、特定ジェスチャで指定される範囲に含まれない場合（ステップS18のN）には、区分部112は、取得されたストロークを、数式に対応するストロークとして設定する（ステップS22）。

【0049】

次に、処理部100は、ストロークの入力があるか否かを判断し（ステップS24）、入力がある場合（ステップS24のY）には、ステップS10の処理に移行し、ユーザによる筆記が終了するまで、ステップS10以降の処理を繰り返す。

【0050】

ストロークの入力がないと判断した場合（ステップS24のN）には、認識部114は、ステップS16、S20で文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、ステップS22で数式に対応するストロークとして設定されたストロークから数式認識エンジンにより数式を認識する（ステップS26）。そして、表示制御部120は、認識結果（文字列を含む数式）を表示部190に表示させる制御を行う。

【0051】

ここで、数式認識エンジンは、特定ジェスチャに対応するストロークも文字列に対応するストロークも理解しないが、特定ジェスチャで囲まれた文字列の情報がないと（すなわち、数式に対応するストロークの情報からのみでは）、数式の構造を解析することができない。そこで、本実施形態の手法では、1つの特定ジェスチャと当該特定ジェスチャに囲まれる文字列を1つのシンボル（数式における1つの記号）として扱って、当該シンボルとして扱う特定ジェスチャの座標情報と数式に対応するストロークの情報から数式を認識する。そのために、特定ジェスチャの筆順を調整する。具体的には、特定ジェスチャを構成する最初のストロークと当該特定ジェスチャで囲まれる文字列に対応する最初のストロークの筆順のうちより筆順の早い方を、当該特定ジェスチャの調整後筆順として特定ジェスチャテーブルに登録する。図2に示す例では、特定ジェスチャBG₁、BG₂、BG₃の調整後筆順として、それぞれ、「1」、「9」、「12」が登録されている。この場合、認識部114（数式認識エンジン）は、特定ジェスチャテーブルを参照することで、特定ジェスチャBG₁とこれに含まれる文字列「area」を1画目に筆記された1つのシ

10

20

30

40

50

ンボルとして扱い、特定ジェスチャ $B G_2$ とこれに含まれる文字列「height」を9画目に筆記された1つのシンボルとして扱い、特定ジェスチャ $B G_3$ とこれに含まれる文字列「width」を12画目に筆記された1つのシンボルとして扱って、これらシンボルと、数式に対応するストロークとの位置や大きさの関係から、数式の構造解析を行うことができる。

【0052】

逐次方式では、ユーザが書き間違いをすることがあるため、Undo機能を備えることが望ましい。この場合、ユーザがUndoコマンドを入力したときに、Undoされた最新のストロークが特定ジェスチャである場合には、当該特定ジェスチャとその登録情報を削除し、当該特定ジェスチャで指定される範囲に含まれるストローク（文字列に対応するストロークとして設定されたストローク）を数式に対応するストロークとして設定し直す。一方、Undoされた最新のストロークが文字列或いは数式に対応するストロークである場合には、単に当該ストロークを削除する。

【0053】

図6は、逐次方式における具体的な表示例を示す図である。図6に示す例では、ユーザは、まず、文字列「value」と数式を筆記し、文字列「value」を囲む特定ジェスチャ $B G_1$ を筆記する（図6A）。すると、登録部110は、特定ジェスチャ $B G_1$ を認識し、表示制御部120は、特定ジェスチャ $B G_1$ を、赤色で特定ジェスチャ $B G_1$ と同等の大きさの矩形を表す画像 $B I_1$ （特定ジェスチャ画像）に置換し、特定ジェスチャ $B G_1$ で囲まれた文字列「value」の色を青色に変化させる（図6B）。次に、ユーザは、新たな特定ジェスチャ $B G_2$ を筆記する（図6C）。すると、登録部110は、特定ジェスチャ $B G_2$ を認識し、表示制御部120は、特定ジェスチャ $B G_2$ を特定ジェスチャ画像 $B I_2$ に置換する（図6D）。次に、ユーザが、特定ジェスチャ画像 $B I_2$ の中に文字列「result」を筆記すると、表示制御部120は、この文字列「result」を、ストロークを取得する毎に青色で表示させる（図6E）。最後に、ユーザが認識の実行を指示する操作を行うと、認識部114は、文字列と数式の認識を行い、表示制御部120は、認識結果を表示させる（図6F）。なお、ここでは特定ジェスチャ画像 $B I_1$ 、 $B I_2$ を除去して認識結果を表示している。

【0054】

3-2. 一括方式の処理

非対話型装置で採用される一括方式では、ユーザは、システムを立ち上げ、タブレット等のインターフェースに数式や文字列を筆記する。そして、筆記後に、ストローク列のデータをPC等（認識装置）に読み込ませ、その際に、文字列を認識する言語を指定する。一括方式では、特定ジェスチャの認識、文字列及び数式の区分、文字列及び数列の認識と認識結果の表示の3つの処理は、全てのストローク列のデータがPC等に送られてから実行される。このとき、入力されたストロークを順次処理する際に、逐次方式と同様の処理を行うこともできる。また、一括方式では、特定ジェスチャの認識と、文字列及び数式の区分とを実時間で言う必要がないことから、全ての特定ジェスチャを認識した後に、それ以外のストロークについて文字列と数式とに区分し、最後にそれらを認識する方法を採用してもよい。図7は、一括方式における後者の方法の処理の一例を示すフローチャートである。

【0055】

まず、処理部100は、文字入力部160に入力されたストローク列を取得する（ステップS30）。次に、処理部100は、変数 i に1をセットする（ステップS32）。次に、登録部110は、 i 番目（ i 画目）のストロークから遡って特定ジェスチャの最大画数（例えば、4画）までの最近のストロークが特定ジェスチャに該当するか否かを判断する（ステップS34）。

【0056】

特定ジェスチャに該当する場合（ステップS34のY）には、登録部110は、特定ジェスチャに該当すると判断されたストロークの情報から、特定ジェスチャの情報（筆順と

10

20

30

40

50

、対向する2角の座標)を特定ジェスチャテーブルに登録する(ステップS36)。特定ジェスチャに該当しない場合(ステップS34のN)には、i番目のストロークを、非ジェスチャストローク(特定ジェスチャに該当しないストローク)として設定する(ステップS38)。

【0057】

次に、処理部100は、変数iの値を1だけ増加させ(ステップS40)、変数iの値がn(nは、取得されたストローク列に含まれるストロークの総数)以下であるか否かを判断する(ステップS42)。変数iの値がn以下である場合(ステップS42のY)には、ステップS34に移行し、変数iの値がnを超えるまで、ステップS34以降の処理を繰り返す。

10

【0058】

変数iの値がnを超えた場合(ステップS42のN)には、処理部100は、変数iに1をセットする(ステップS44)。次に、区分部112は、特定ジェスチャテーブルを参照して、ステップS38で設定されたi番目の非ジェスチャストロークが、いずれかの特定ジェスチャで指定される範囲に含まれるか否かを判断する(ステップS46)。i番目の非ジェスチャストロークがいずれかの特定ジェスチャで指定される範囲に含まれる場合(ステップS46のY)には、区分部112は、i番目の非ジェスチャストロークを、文字列に対応するストロークとして設定する(ステップS48)。一方、i番目の非ジェスチャストロークが特定ジェスチャで指定される範囲に含まれない場合(ステップS46のN)には、区分部112は、i番目の非ジェスチャストロークを、数式に対応するスト

20

【0059】

次に、処理部100は、変数iの値を1だけ増加させ(ステップS52)、変数iの値がm(mは、非ジェスチャストロークの総数)以下であるか否かを判断する(ステップS54)。変数iの値がm以下である場合(ステップS54のY)には、ステップS46に移行し、変数iの値がmを超えるまで、ステップS46以降の処理を繰り返す。

【0060】

変数iの値がmを超えた場合(ステップS54のN)には、認識部114は、ステップS48で文字列に対応するストロークとして設定されたストロークから文字列認識エンジンにより文字列を認識し、特定ジェスチャテーブル(シンボルとして扱う特定ジェスチャの座標情報)とステップS50で数式に対応するストロークとして設定されたストロークから数式認識エンジンにより数式を認識する(ステップS56)。そして、表示制御部120は、認識結果(文字列を含む数式)を表示部190に表示させる制御を行う。

30

【0061】

図8は、一括方式における具体的な表示例を示す図である。一括方式では、ユーザにより筆記され認識装置に取り込まれた全てのストロークが表示され(図8A)、ユーザが認識の実行を指示する操作を行うと、特定ジェスチャの認識、文字列及び数式の区分、文字列及び数列の認識が実行され、認識結果が表示される(図8C)。なお、特定ジェスチャの認識と、文字列及び数式の区分を実行した後、文字列及び数列の認識を実行する前に、図8Bに示すように、特定ジェスチャの認識結果と数式及び文字列の区分の結果を表示してもよい。図8Bに示す例では、特定ジェスチャBG₁、BG₂、BG₃を、それぞれ特定ジェスチャ画像BI₁、BI₂、BI₃に置換し、特定ジェスチャBG₁、BG₂、BG₃で囲まれた文字列「area」、「height」、「width」の色をそれぞれ青色に変更している。

40

【0062】

本実施形態によれば、文字列と数式とを区分するための特定ジェスチャを導入して、特定ジェスチャで囲まれるストロークを文字列に区分し、特定ジェスチャで囲まれないストロークを数式に区分し、文字列認識エンジンと数式認識エンジンによりそれぞれを認識することで、数式に含まれる文字列と数式とを明確に区分してそれぞれを認識することができる。また、数式を認識する際に、特定ジェスチャと当該特定ジェスチャで囲まれるスト

50

ロークを1つのシンボルとして扱うことで、数式認識エンジンを用いて、文字列を含む数式の構造解析を行うことができる。

【0063】

また、本実施形態によれば、逐次方式において、特定ジェスチャが筆記されたときに当該特定ジェスチャで囲まれるストロークを文字列に区分する処理を行い(図5のステップS16)、また、筆記されたストロークが特定ジェスチャで囲まれる場合に当該ストロークを文字列に区分する処理を行う(図5のステップS20)ことで、特定ジェスチャがどのタイミングで筆記されても、文字列と数式とを区分することができ、ユーザの利便性を向上することができる。すなわち、本実施形態によれば、特定ジェスチャを筆記してから文字列を筆記する場合(図3A)や、文字列を筆記してから特定ジェスチャを筆記する場合(図3B、図3C)や、特定ジェスチャの中に文字列を書き足す場合(図3D)であっても、文字列と数式とを区分する処理を行うことができる。

10

【0064】

なお、本発明は、上述の実施の形態に限定されるものではなく、種々の変更が可能である。本発明は、実施の形態で説明した構成と実質的に同一の構成(例えば、機能、方法及び結果が同一の構成、あるいは目的及び効果が同一の構成)を含む。また、本発明は、実施の形態で説明した構成の本質的でない部分を置き換えた構成を含む。また、本発明は、実施の形態で説明した構成と同一の作用効果を奏する構成又は同一の目的を達成することができる構成を含む。また、本発明は、実施の形態で説明した構成に公知技術を付加した構成を含む。

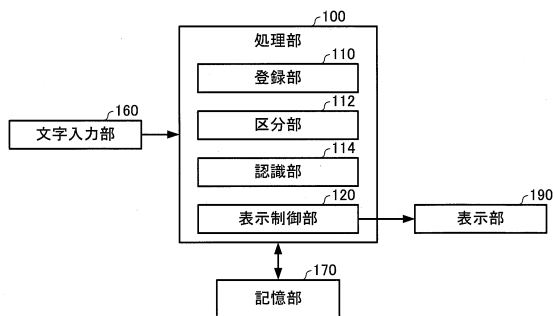
20

【符号の説明】

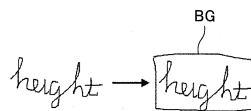
【0065】

100 処理部、110 登録部、112 区分部、114 認識部、120 表示制御部、160 文字入力部、170 記憶部、190 表示部

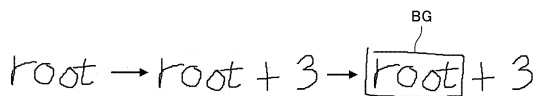
【図1】



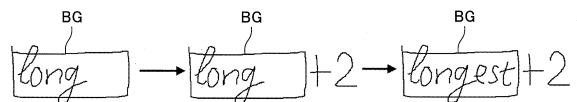
【図3B】



【図3C】



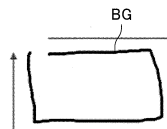
【図3D】



【図2A】

$$Area = \frac{1}{2} \times height \times width$$

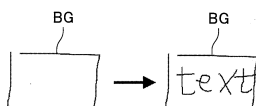
【図4A】



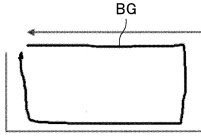
【図2B】

特定ジェスチャ	筆順	調整後筆順	Top-left corner (x,y)	Bottom-right corner (x,y)
BG ₁	5	1	(185,250)	(431,357)
BG ₂	13	9	(663,258)	(886,387)
BG ₃	18	12	(932,245)	(1116,371)

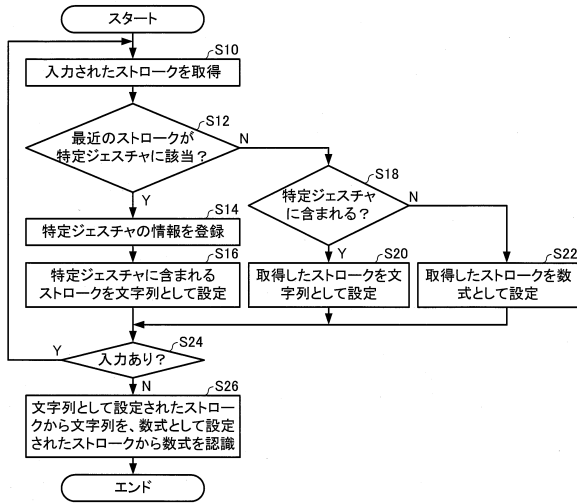
【図3A】



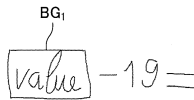
【図4B】



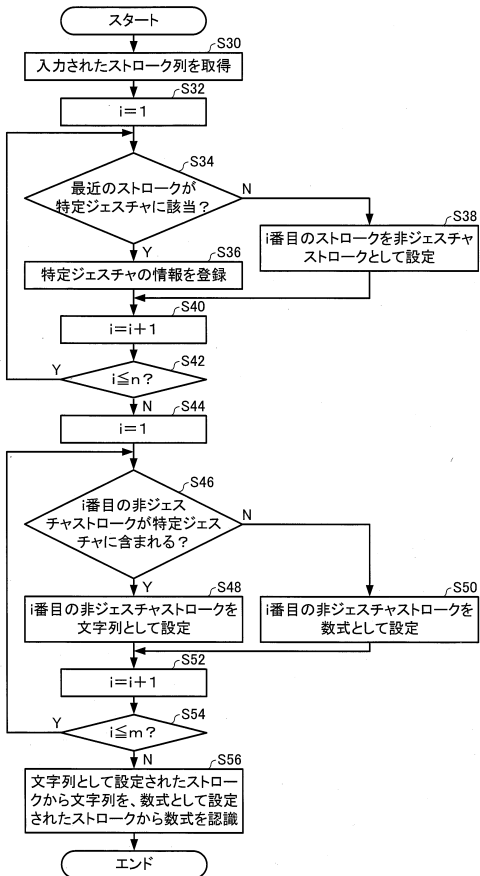
【図5】



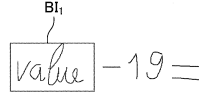
【図6A】



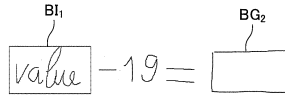
【図7】



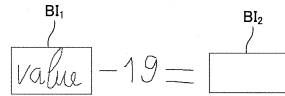
【図6B】



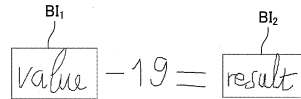
【図6C】



【図6D】



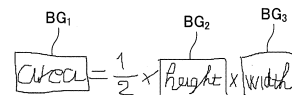
【図6E】



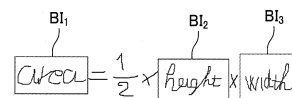
【図6F】

value - 19 = result

【図8A】



【図8B】



【図8C】

area = 1/2 * height * width

フロントページの続き

(72)発明者 レ ドウック アイン

東京都府中市晴見町3 - 8 - 1 国立大学法人東京農工大学内

審査官 佐田 宏史

(56)参考文献 米国特許出願公開第2006/0001667(US, A1)

特開2011-053986(JP, A)

特開平11-203412(JP, A)

特開2000-080615(JP, A)

特開2014-127188(JP, A)

村瀬 敦史、外3名, “手書き入力による数式認識システム”, 情報処理学会研究報告, 日本, 社団法人情報処理学会, 1991年 5月 9日, Vol.91, No.34, pp.1-8

持田 桂介、外4名, “文字列と数式が混在したオンライン手書きパターンの認識手法”, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2005年 3月10日, Vol.104, No.741, pp.43-48

偉 叢、外3名, “ストロークの相互関係を用いたオンライン手書き数式認識”, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 1998年12月17日, Vol.98, No.489, pp.9-16

中川 正樹、佐藤 俊, “表示一体型タブレット上でのペンの囲みに対する対象の包含を判定する高速アルゴリズムの実現と評価”, 電子情報通信学会論文誌, 日本, 社団法人電子情報通信学会, 1994年 8月25日, Vol.J77-D-II, No.8, pp.1630-1639

(58)調査した分野(Int.Cl., DB名)

G06K 9/00, 9/62