

eラーニングのための  
単語辞書を使用した板書手書き文字認識

平成18年度

三重大学大学院工学研究科  
博士前期課程 電気電子工学専攻

吉田大祐

修士論文

e ラーニングのための  
単語辞書を使用した板書手書き文字認識



平成 18 年度修了

三重大学大学院工学研究科

博士前期課程 電気電子工学専攻

吉田 大祐

# 目次

第1章 はじめに	1
1.1 研究の背景	1
1.1.1 eラーニングとは	1
1.1.2 eラーニングの現状	2
1.1.3 高等教育の状況とeラーニングの役割	3
1.2 本研究の目的	6
1.3 eラーニングシステムの特徴と計画	8
第2章 個別学習用eラーニングシステム	9
2.1 本研究室のeラーニングシステムの概要	9
2.2 従来システムの問題点	18
2.2.1 教室環境の問題点	18
2.2.2 システム上の問題点	20
第3章 加重方向指数ヒストグラム法を用いた文字認識	23
3.1 加重方向指数ヒストグラム法の概要	23
3.2 紙とホワイトボード上との文字認識結果の比較	30
第4章 単語辞書を用いたキーワード認識	35
4.1 単語辞書を用いない手法の問題点	35
4.2 単語辞書を用いたキーワード認識とは	36
4.3 単語辞書を用いたキーワード認識法の概要	38
第5章 キーワード認識の評価実験及び考察	45
5.1 実験環境	45
5.2 実験結果	46
5.3 考察	51
5.3.1 実験結果の考察	51
5.3.2 他の単語辞書を用いた研究との比較	52
第6章 おわりに	53
参考文献	54
謝辞	56
発表論文リスト	57

## 図一覧

- 図1 プレンディット・ラーニングの概要
- 図2 高等教育において期待される e ラーニングの役割
- 図3 本研究の位置付け
  - (a) 講義映像データベース
  - (b) 1つの講義シーン
  - (c) ヘッダの内容
- 図4 個別学習アシストシステムとデータベースとの関係
- 図5 e ラーニングシステムの構成
- 図6 ペン位置検出装置
- 図7 ホワイトボード上に筆記された板書
- 図8 計算機上に取り込まれた板書画像
- 図9 Canon VC-C3
- 図10 個別学習アシストシステムのブラウザ
- 図11 個別学習アシストシステムの処理の流れ
  - (a) 講義映像データベースの作成方法
  - (b) 講義映像データベースのブラウザでの利用方法
- 図12 キーワードファイル
- 図13 計算機上に取り込まれた板書画像
- 図14 リンクされたキーワードと補助説明の表示例
- 図15 本システムに期待される効果
- 図16 教室の後部の壁に設置した IP カメラ
- 図17 黒板を対象としたシステムの概要
- 図18 Script Converter による文字認識結果 1
- 図19 Script Converter による文字認識結果 2
- 図20 Script Converter による文字認識結果 3
- 図21 個別文字認識までの処理の流れ
- 図22 計算機上に取り込んだ板書画像
- 図23 作成された行画像
- 図24 垂直方向への黒画素の和を求めた射影ヒストグラム
- 図25 切り出された個別文字
- 図26 加重方向指数ヒストグラム法の処理過程
- 図27 入力画像 (8bit Bitmap 形式)
- 図28 8 近傍 (黒色: 注目画素 灰色: 近傍画素)
- 図29 4 近傍 (黒色: 注目画素 灰色: 近傍画素)
- 図30 前処理後の画像
- 図31 7×7 の小領域に分割された入力画像
- 図32 方向指数 (4 方向)

- 図33 空間的集約の中心点の位置
- 図34 空間的加重フィルタ（数値は重みを示す）
- 図35 加重方向指数ヒストグラム（ $4 \times 4 \times 4 = 64$  次元の特徴ベクトル）
- 図36 紙に筆記された文字列
- 図37 ホワイトボードに筆記された文字列
- 図38 紙の方が正読率の良かった例 1
- 図39 紙の方が正読率の良かった例 2
- 図40 ホワイトボードの方が正読率の良かった例 1
- 図41 ホワイトボードの方が正読率の良かった例 2
- 図42 ストロークの欠けた文字
- 図43 キーワードが認識できない例 1
- 図44 キーワードが認識できない例 2
- 図45 入力画像
- 図46 単語辞書
- 図47 人間の感覚での単語辞書によるキーワード認識
- 図48 入力文字と文字認識結果
- 図49 意図しないキーワードを認識する例
- 図50 キーワードが認識できない例
- 図51 1 位の文字候補の相違度の分布関数
- 図52 相違度差の統計結果 1
- 図53 相違度差の統計結果 2
- 図54 入力文字画像の例
- 図55 成功例 1
- 図56 成功例 2
- 図57 成功例 3
- 図58 失敗例 1
- 図59 失敗例 2
- 図60 失敗例 3

## 表一覧

表1 文字認識結果

表2 紙上・ホワイトボード上の認識精度の比較

表3 紙上・ホワイトボード上のストロークの違い

表4 文字認識結果

表5 キーワード認識実験結果

表6 “と”の文字認識結果

# 第1章 はじめに

## 1.1 研究の背景

### 1.1.1 e ラーニングとは

近年、光通信などのネットワーク通信のブロードバンド化が急速に進み、それと同時に、それに利用されるコンピュータや周辺機器等の情報処理端末の高精度化及び低価格化が進んできている。これにより、これらを利用した応用技術、その一つとして e ラーニングが注目されてきている。

e ラーニング、その言葉は広く認知されてきており、その定義に関しては様々な機関や研究者が行っているが、広い意味で「コンピュータやネットワークを利用した教育形態の総称」と定義される場合が多い。

e ラーニングには様々な形態がある[1]。代表的なものとして、インターネットを通じてオンラインで教材の配信やテストを行う「非同期型」(WBT(Web Based Training), オンデマンド型とも呼ばれる)や、衛星通信やテレビ会議システムを用いて講師が行う授業をリアルタイムに配信する「同期型」(リアルタイム型とも呼ばれる)、また近年では、従来からの教室に一同を集めて教育する集合教育と、これらの e ラーニングを組み合わせる「ブレンディット・ラーニング」の運用も進められてきている(図1)。「ブレンディット・ラーニング」における重要な点は、学習者の視点に立った個別学習を支援するシステムを提供する事である。

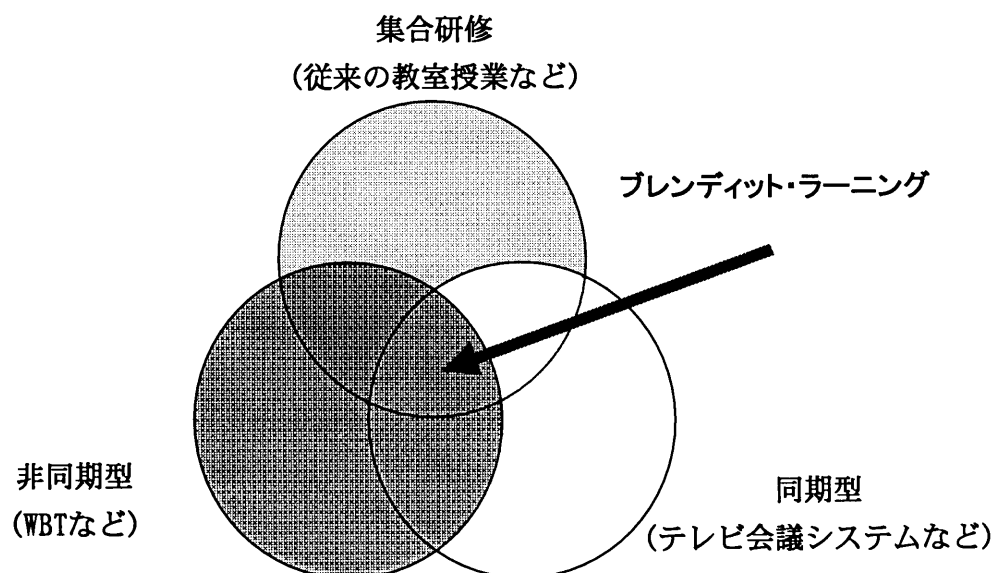


図 1 ブレンディット・ラーニングの概要

### 1.1.2 e ラーニングの現状

前節で述べたように、ネットワークの通信速度の高速化や利用料金や情報機器の低価格化、携帯電話の普及に伴い、日本においてもインターネットが急速に普及してきている。この事からも、家庭や学校等、どこでも e ラーニングが出来る環境が整いつつあり、「いつでも・どこでも・誰にでも」を実現させる「ユビキタス時代」が始まってきていると言える。

政府の e-Japan 戦略では、教育に関しては初等中等教育における条件整備と産業教育における IT 人材の育成に重点が置かれてきた。この追い風を受けて、企業内教育での e ラーニング実践が増えた。それに対し高等教育に関しては、英語 TOEIC 自習システム、JST の Web ラーニングプラザ、放送大学などの先端的な試みがいくつか始まった程度である。

それを受けて、2003 年 7 月に発表された e-Japan 戦略Ⅱでは新たに高等教育における遠隔教育が新たに推進される事になった。例えば「2005 年度までに IT を利用した遠隔教育を実施する大学学部・研究科を 2001 年度の 3 倍とすることを目指す」としている。また、「社会人等が時間や場所を選ばず、IT を活用して必要となる教育を効率的かつ、低廉な価格で受ける事ができる環境を整備する」とされている。これらは専門職大学院における実施を含んでいる。これらの戦略は、明らかに高等教育に重点が移ってきた事が読み取れる。

重点計画-2006 では、「世界に通用する高度 IT 人材の育成 次世代を見据えた人的基盤づくり」として普通教室における教育用 PC 設備の充実等（2010 年度までに生徒 3.6 人당に 1 台を目標とする）が挙げられている。

このように、大学教育で e ラーニングを展開できる環境が、インフラ面と受講対象者両面で整備されつつある（[2]～[4]）。これを受けて三重大学においても、2005 年度より英語教育に TOEIC を導入し、2006 年度から全教室で無線 LAN が使用できるようになった。また全学的に Moodle という文字情報を主体とした e ラーニングシステムを導入し、約 200 科目の登録がされ、今後 e ラーニングの普及期に入ると思われる。

### 1.1.3 高等教育の状況とeラーニングの役割

1.1.2節で述べたように、近年、eラーニング導入のための環境は徐々に整備されてきたが、企業に比べ高等教育においてeラーニングの導入はまだ進んでいないと言える。大学でのeラーニングの活用は以前から行われてきたが、研究や実験段階のもの等、規模的に小さいものが多かった。また、大学の情報環境の整備はかなり進んだが、それを活かした形でeラーニングが普及しているとは言い難い状況であった。

ところが近年、大学は少子化や独立行政法人化という非常に大きな環境変化に直面している。そのため、少子化による学生の獲得競争に生き残る方法の一つとして、先進的な学習環境の整備の一環でeラーニングの充実により他大学との差別化や、留学生の獲得などを意識してeラーニングに取り組む大学が増えている。メディア教育開発センターの調査によると、4年生大学の4割がインターネットを教育に利用しており、その比率は年々高まっている。また、社会人教育の必要性に対応して、通学せずに単位が取得できるeラーニングを活用した講座の開設が徐々に進んでいる。

ここで、高等教育でのeラーニング導入事例を見ると以下のような特徴がみられる。

- 授業の補助的集団としてeラーニングを活用するものから、eラーニングで学位を取得できるものまで幅広い。
- 一部の授業への導入事例もあれば、学部単位で導入を検証して現在は全学レベルで導入を始めた戦略的な事例もある。
- 正規の学生を対象としたものが中心であるが、一部は社会人向けの生涯学習や職業能力開発のためのものもある。
- eラーニングの特性を最大限に活用して、海外に目を向けて開発する例もある。
- eラーニングを学習プロセスに活かすだけでなく、教育機関として新たなビジネスモデルを模索している例もある。

このように、大学の分野ごとの教育内容・教育方法の多様性同様、eラーニング活用についても非常に幅広いといえる。

また、これらの事例から、高等教育においてeラーニングが期待される役割として、IT活用による教育の質と提供手段の改善を目的とした「教育の質と提供手段の視点」と、大学の情報化の推進やシステム統合、他大学との連携、社会人学生や留学生の獲得などの「大学経営の視点」に大別できる(図2)。

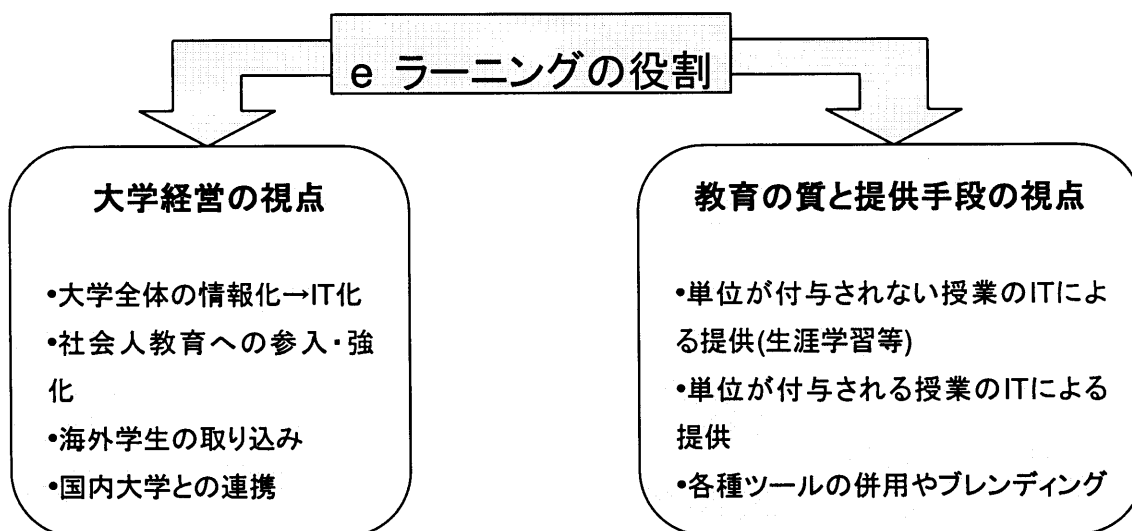


図2 高等教育において期待される e ラーニングの役割

それと同時に近年、大学教育における問題点として、学生の大学入学時までのカリキュラムの相違やゆとり教育の弊害等によって既知内容の格差が存在する事や学生の講義に対する興味の多様化等から、学生間の能力格差が拡大し、講義についていけず、脱落し結果的に留年してしまう学生、退学してしまう学生の増加が問題視されている。こういった問題を改善・解決するためにも、e ラーニングに対する期待が教育現場から高まってきている。e ラーニングの特長である、パソコンさえあれば学生が必要な時にどこでも学習が可能である事や学習者の学習レベルに応じた学習が可能である事を活かしたシステムによって個別学習環境を配備する事で、これらの問題を改善・解決する事が期待できるからである。

このように、高等教育での e ラーニングのニーズは高まっている反面、その導入は思うように進んでいない。その理由と従来の e ラーニングの問題点として以下の事項が挙げられる [5]。

#### ① コンテンツの作成に膨大な費用が必要とされる

このようなコンテンツに対して受講生の目は肥えており、例えば NHK の教育講座のようなテレビ番組やテレビゲームがライバルとなる。実際にそれらは莫大な費用と投入して作成している。参考として、ニューヨーク大学では1 コース 2000 万円以上かけて初めて支持の得られるコンテンツが得られた。それに対して、日本では一つのコンテンツに掛ける費用は 200～300 万円程度で、それと比較しても満足なコンテンツを得るための費用が不足していると言える。

#### ② コンテンツを作成する事が困難

受講生の基礎知識や興味の方向はそれぞれ違うため、一冊の本が完璧に受講生に適応している事はない。そのため、複数の教材と相互にリンクしたコンテンツを作成する

必要性があるため、コンテンツ作成をより困難にしている。

③ 必要なコンテンツが存在しない

①で挙げたように、多くのお金と専門家が投入した優れたコンテンツは受講生が多く集まるコースでしか開発されない。そのため、人気のないコースはeラーニング化されない。それゆえに、ある学生が習いたくても習えない、学習したくても先生がいないという状況が発生している。

④ 専門の管理者が必要

サーバ管理、ネットワーク管理、コンテンツ管理等様々な領域で高度な専門知識を持った管理者が必要とされる。

以上の点を踏まえて、従来のスタイルを変える事なく、eラーニングのメリットを享受できるシステム、具体的には以下の条件を満たしたシステムを開発する事が高等教育にeラーニングの導入を広めていく一つの方法と考えられる。

① 安価にコンテンツを作成可能

→授業を生中継したり、双方向性を持たせたり、その講義自体の情報を有効利用する事で安価に作成する事が可能ではないかと考える。

② 多種の教材と相互にリンクしたコンテンツ

→様々な受講生に適応すべく、一つの教材に拘るのではなく、複数の教材、関連する講義との連携をとる事が出来れば魅力的なシステムになると考える。

③ コンテンツの充実

→①、②の条件を満たす事で、安価に魅力的なコンテンツが作成可能となり様々な領域のコンテンツを作成可能と言える。

## 1.2 本研究の目的

このような状況から本研究室では、大学における e ラーニング用コンテンツの充実を目指し、大学の通常講義を対象とした e ラーニングによる個別学習システムの開発を行っている[6][7]。このシステムは、主に講義の復習を自主的に行う事や欠席した講義を別の時間に受講する事を目指しており、非同期型のシステムを対象としている。以上のようなシステムを自動的に作成する事で、安価に大学講義用 e ラーニングコンテンツの蓄積が可能になり、受講者となる学生は必要に応じ利用する事で、1.1.3 節で述べたような問題を改善・解決する一手法となる事が期待される。本システムの特徴は、以下のようになる。

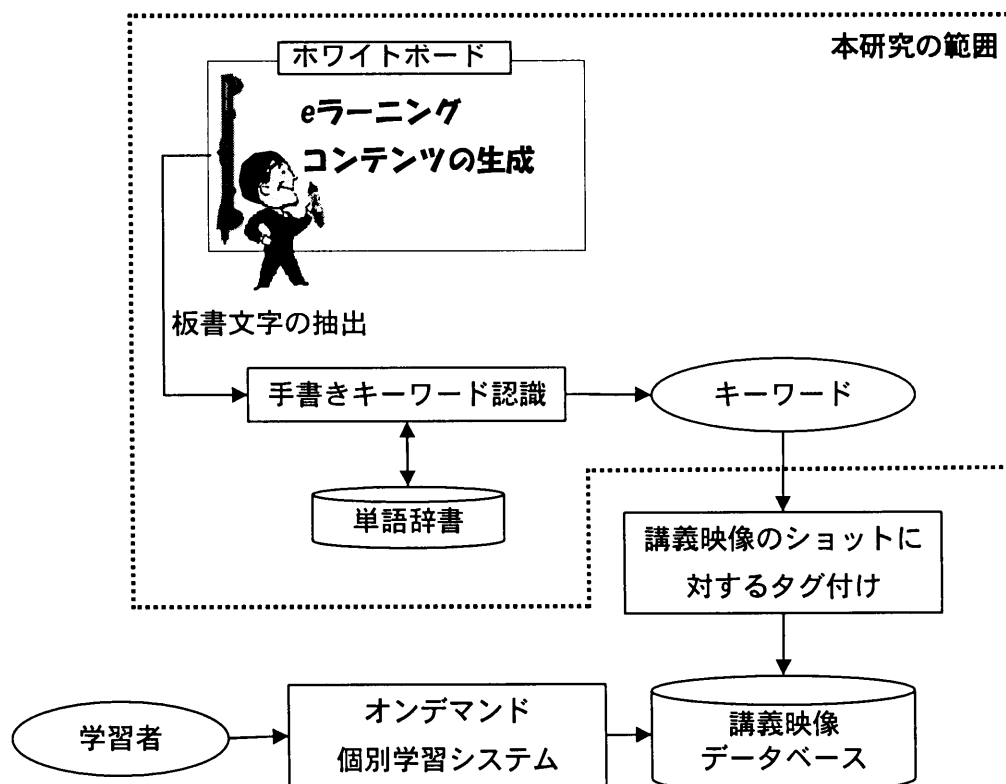
- (1) 講義映像と板書画像、テキスト化された板書、キーワードに関連する電子教科書の本文テキストがリンクされており、同時に見る事が可能である。
- (2) 板書画像を文字認識しており、受講者が必要に応じキーワードをクリックするだけで、電子教科書中の補足説明を同じウィンドウに表示する事ができる。

ここで、このシステムにおいて(2)の機能のために板書文字を文字認識し、キーワードを認識し関連事項とのインデキシングを行う事が重要な処理となるが、従来システムではこれまで文字認識をペン位置検出装置付属のソフトウェアで行っていた。しかし、認識率が e ラーニングに用いるには不十分であり、さらにソースが非公開であるためシステムの完全自動化が不可能で、改良が不可能であるという問題があった。また、ホワイトボードに筆記された文字は紙に筆記された文字に比べ、ペンの滑りが大きく、筆記に不慣れであるため、ストロークが紙に筆記された文字と異なり、不安定な文字になりやすい。そのため、文字認識精度が低く、従来システムで用いていたキーワード認識方法では認識結果とキーワードとの完全一致が必要であったため、教科書中のキーワードの説明部分とリンクさせるためには不十分で、ホワイトボード上の手書き文字を認識した文字列からキーワードを認識することができない場合が多いという事が問題となっていた。

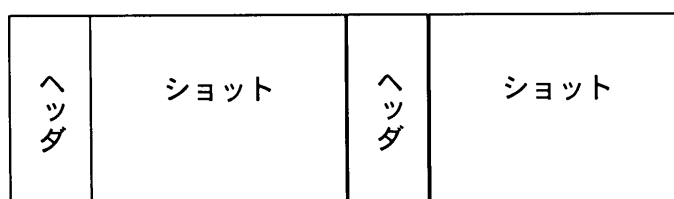
そこで本研究では、ソースが公開されている文字認識ソフトウェアで新たにシステムを構築し、高精度にキーワードを認識する事を目的に研究する。その内容を以下に示す。

- (1) 板書画像から文字領域を抽出し、板書文字を文字認識し、その結果から単語辞書を利用し、キーワードを認識する方法を提案する。
- (2) (1)の手法により、抽出したキーワードを e ラーニングに利用する方法を考察する。

本論文では、2 章でこれまでに研究開発してきた e ラーニングシステムの概要と問題点を述べ、3 章で新たに用いる文字認識手法を、4 章で単語辞書を利用したキーワード認識法について説明する。さらに 5 章でキーワード認識の評価実験と考察、他の手法との比較を行い、6 章でまとめを行う。



(a) 講義映像データベース



(b) 1つの講義シーン

講義名, キーワードのタグ,  
年月日, 録画時間等

(c) ヘッダの内容

図3 本研究の位置付け

### 1.3 eラーニングシステムの特徴

前節で述べたように、本eラーニングシステムの特徴は以下になる。

- (1) 授業映像と板書映像、テキスト化された板書、キーワードの補足説明がリンクされており、同時に見る事ができる。
- (2) 板書画像を文字認識しており、利用者が追加説明の欲しいキーワードをクリックするだけで、教科書の補足説明を同じウィンドウに表示する事ができる。
- (3) 講義の欠席者や講義内容の理解が不十分な生徒に対して、非同期型で学内での利用を現段階では対象としている。

将来構想としては、計算機やネットワークの更なる充実に従って同期型や学外での利用も提案している。また、図4のように、板書と教科書とのリンクのみではなく、様々なデータベースをキーワード検索の軸にした個別学習アシストシステムを構築したいと考えている。

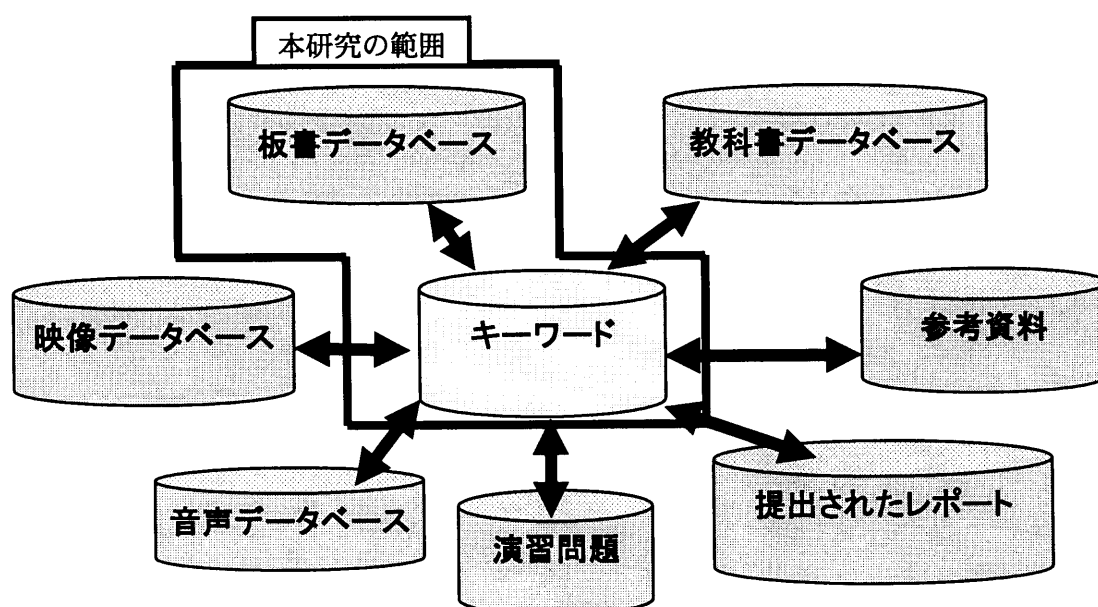


図4 個別学習アシストシステムとデータベースとの関係

## 第2章 個別学習用 e ラーニングシステム

### 2.1 本研究室の e ラーニングシステムの概要

これまでに本研究室では、ホワイトボードを用いた授業を対象とした、個別学習用 e ラーニングを提案してきた [6] [7]。その理由としては、ホワイトボードが教育施設の情報処理教室や会議室等によく使用されていて、さらに今後情報処理教室以外の教室にも拡大していくと思われるからである。

本研究室で開発している e ラーニングシステムの構成は、図 5 のようになる。

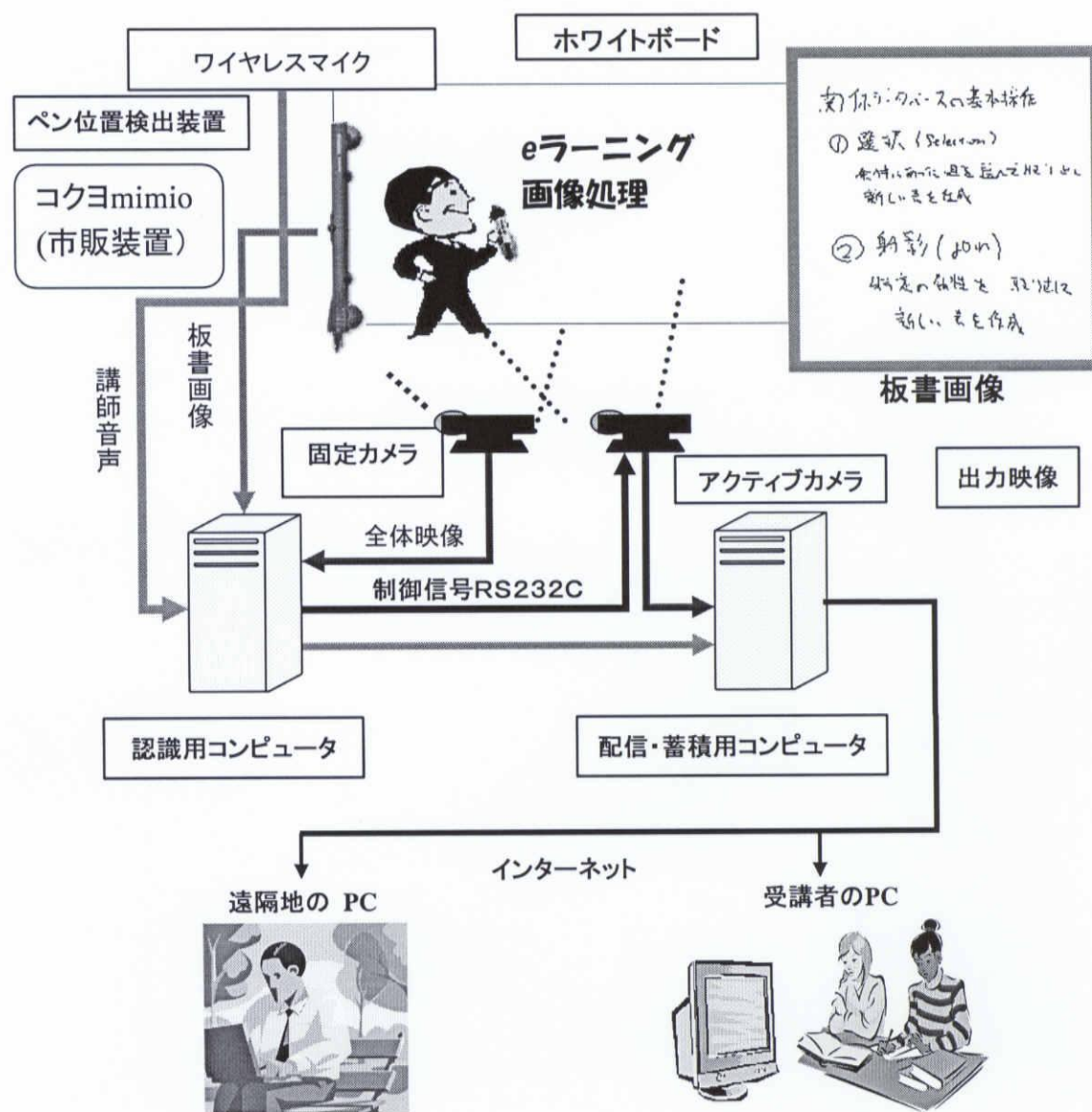


図5 e ラーニングシステムの構成

## 【使用機器】

### (1) ペン位置検出装置

ホワイトボード上に筆記された板書画像を取得するためにペン位置検出装置を用いる。使用するペン位置検出装置は米国のベンチャー企業 Virtual Ink 社が開発した、KOKUYO 社販売の mimio Xi である[8]。

mimioは図6のように受信装置とホワイトボード用のペンを格納するためのスタイラスから構成されている。スタイラスに収められたペンを用いてホワイトボード上に筆記することで、スタイラスのスイッチが入り赤外線と超音波を発する。受信装置には赤外線用の受信機が1機、超音波用の受信機が2箇所あり、スタイラスにより発せられた赤外線によりペンの色を、超音波によりペンの XY 座標を取得し、計算機へ情報を発信する。ペンの XY 座標はスタイラスとホワイトボードの左端に取り付けられた受信装置の2機の受信機による3点測量により行われる。以上によって、ホワイトボード上に筆記された文字と計算機上に取り込まれた板書画像の例を図7、8に示す。



図6 ペン位置検出装置

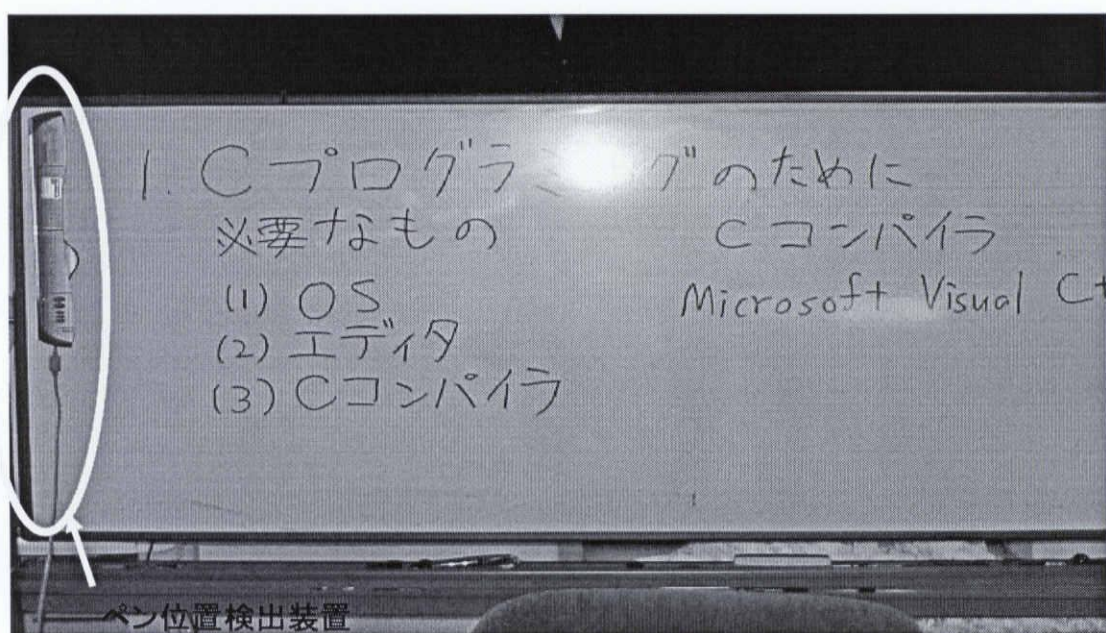


図7 ホワイトボード上に筆記された板書

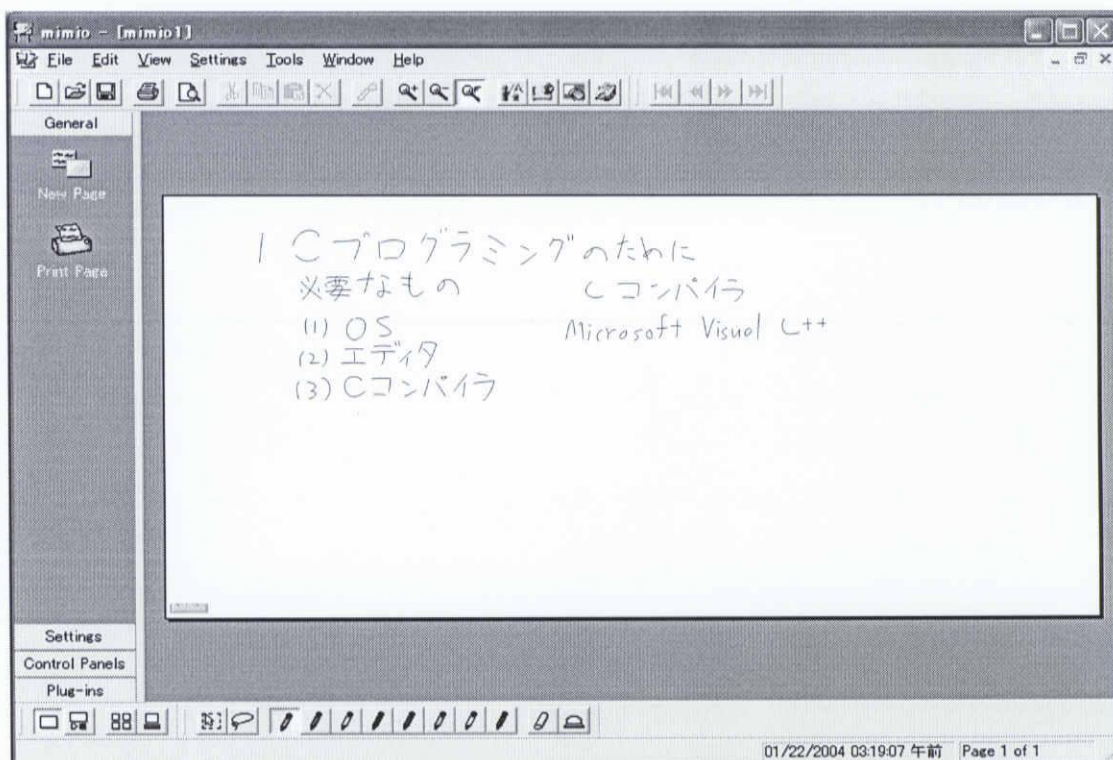


図8 計算機上に取り込まれた板書画像

ペン位置検出装置 mimio を用いる利点としては以下の事項が挙げられる。

- 電子ボードを購入するよりも低予算で導入可能。
- 装置を取り付けるだけなので、現状のホワイトボードがそのまま利用可能。
- スタイラスを利用する事で、講師に特別な技術を必要としないため、ストレスを感じさせず、容易に板書データを計算機上に取り込む事が可能。
- 板書を計算機上に取り込む事で、講師がホワイトボードの前に立っていても常にホワイトボードの板書を取得する事が可能。
- 板書した内容を計算機に蓄積する事ができ、文字認識ソフトを使用すれば、必要に応じてキーワードにより検索ができる。

## (2) コミュニケーションカメラ

講義映像撮影用カメラとして、2台のコミュニケーションカメラ（パン、チルト、ズームがパソコンからの RS232-C 信号により行えるアクティブカメラ）を用い、それぞれ固定視野カメラ及び、可動視野カメラとした。使用するのは Canon 社製の VC-C3[9] 2 台である（図 9）。



図9 Canon VC-C3

図 5 のように、2 台のカメラをそれぞれ固定カメラ、可動視野カメラ（パン、チルト、ズーム）として用いて、自動的に講師映像を撮影し[6]、ペン位置検出装置を用いて、ホワイトボードの板書画像を、ワイヤレスマイクを用いて講師の発話を計算機上に取り込み、それらから次節で述べるような e ラーニングコンテンツを作成し、XML タグを付加する事で講義データベースに蓄積する形態となっている。そのコンテンツを教室内で受講している受

講者や学内にいる受講者がPCを用いて利用する事を想定する。

受講生のPCに表示される表示ウィンドウは図10のようになる。

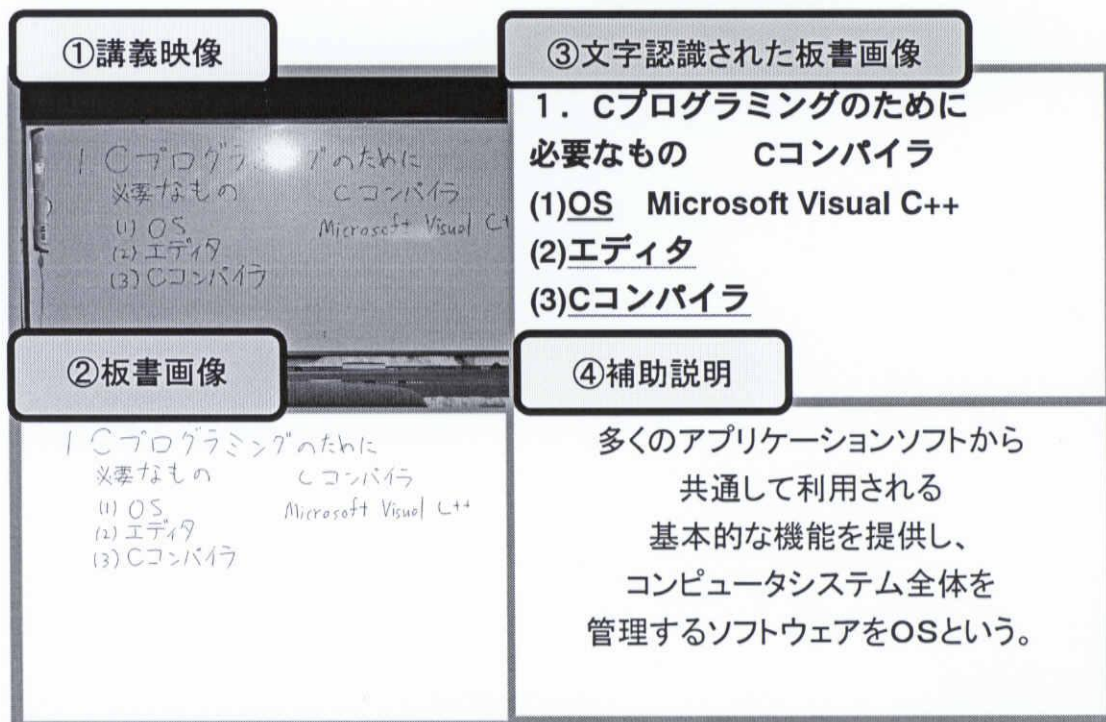


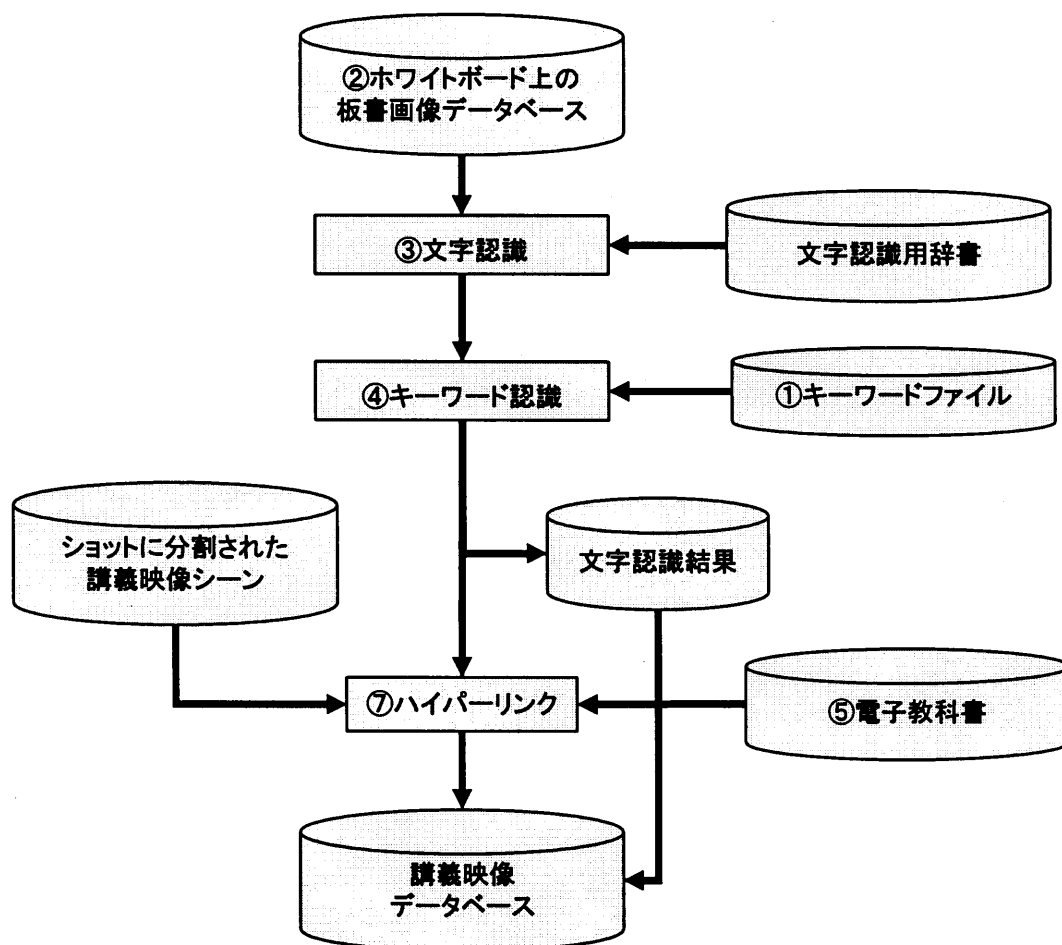
図10 個別学習アシストシステムのブラウザ

図10のように、受講生のPC上に表示されるコンテンツは以下の4つのサブウィンドウから構成される。

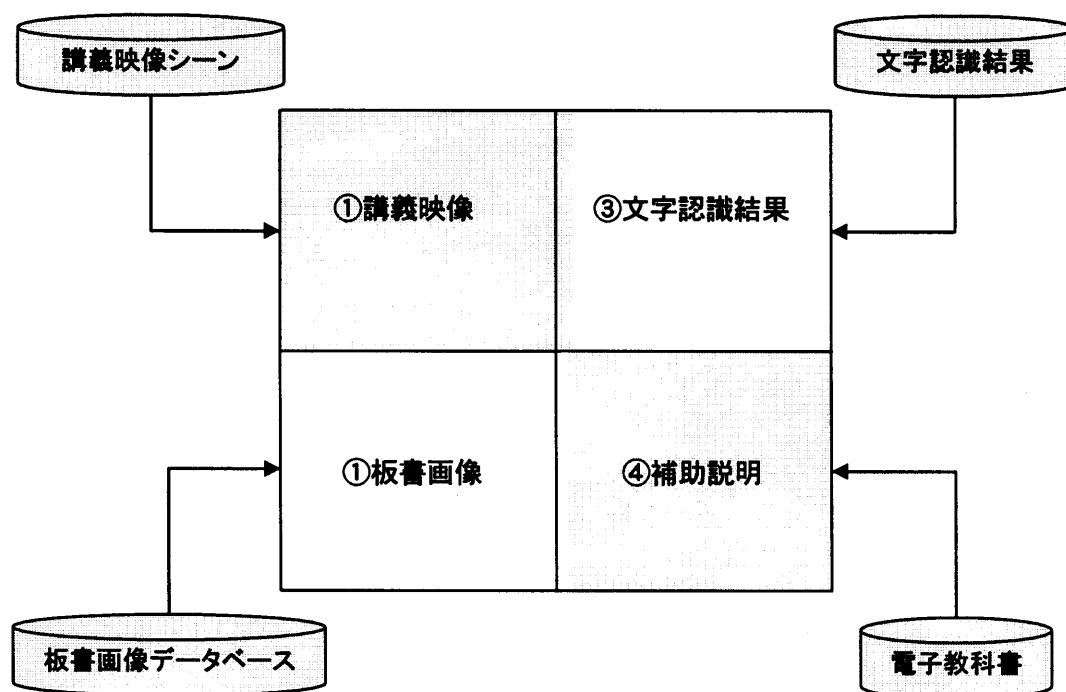
- ① 撮影用カメラにより、講師を中心とした講義映像を撮影した動画を表示するサブウィンドウ（講義映像は[6]の手法によって撮影される）
- ② 講師によって筆記された板書をペン位置検出装置によって取り込んだ画像を表示するサブウィンドウ
- ③ 計算機上に取り込まれた板書文字列をテキスト化し、同時に講義に重要なキーワードに対し関連事項へのリンクを表示するサブウィンドウ
- ④ リンクされたキーワードに対する関連事項（この場合教科書全文ファイルへのリンク）を表示するサブウィンドウ

関連事項へのリンクとして、教科書全文ファイルへのリンクを対象としたものが本研究室の従来システムとして開発されてきている[7][10]。また、前章で述べたように、教科書全文ファイル以外の、講師音声や講義映像等、様々なデータベースにキーワードを中心としてリンクする事を将来構想としている。

上記で述べた個別学習アシストシステム（以下、個別学習システム）の処理の流れを図11に示す。



(a) 講義映像データベースの作成方法



(b) 講義映像データベースのブラウザでの利用方法

図11 個別学習アシストシステムの処理の流れ

### ① キーワードファイル

講義で重要だと思われるキーワードをあらかじめキーワードファイルとして登録しておく。本研究ではその講義で用いる教科書の索引に列挙してある単語を使用した。

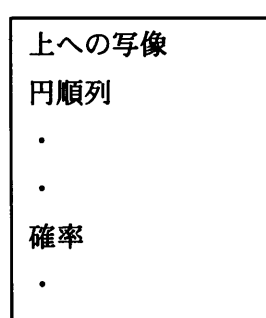


図12 キーワードファイル

### ② 板書画像

ペン位置検出装置を用いて、ホワイトボード上に筆記された文字列をリアルタイムで計算機上に取り込む。

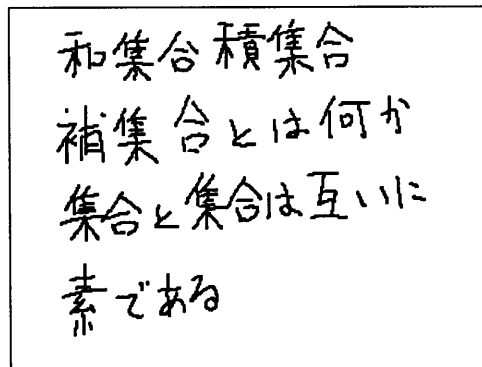


図13 計算機上に取り込まれた板書画像

### ③ 文字認識

②で得られた板書画像に文字認識処理を行う。本研究室の従来システム[7]では mimio 付属の mimio 用文字認識ソフト Script Converter を用いているが、文字認識精度が低いので、本研究では研究室で開発したソフトウェアを使用する。

### ④ キーワード認識

文字認識によって得られた板書文字列データとキーワードファイルを照合し、該当するキーワードが存在するかどうか探索する。本研究ではこの処理がメインテーマである。

### ⑤ 電子教科書

授業で用いる教科書や参考書等を事前に電子テキスト化しておく。本研究では、全文テキストがテキスト形式で著者の先生から使用許可が取れた教科書を使用した。

### ⑥ キーワード関連事項の抽出

⑤の電子教科書とキーワードファイルを照合し、キーワードに関連する部分を抽出する。従来システムではあらかじめ抽出したものを HTML ファイルとしてデータベース化していたが、現在は⑦のリンクを行う以前に適宜照合し抽出している[10]。

### ⑦ リンク

キーワードファイルに登録された単語に関して、④のキーワード認識で認識されたキーワードと⑥で抽出された教科書全文テキストとをハイパーリンクによりリンクさせる。

### ⑧ 個別学習アシストシステムの表示

リンクされたキーワードに対して受講者は必要に応じて、キーワードに対する補助説明を適宜利用することができる(図14)。

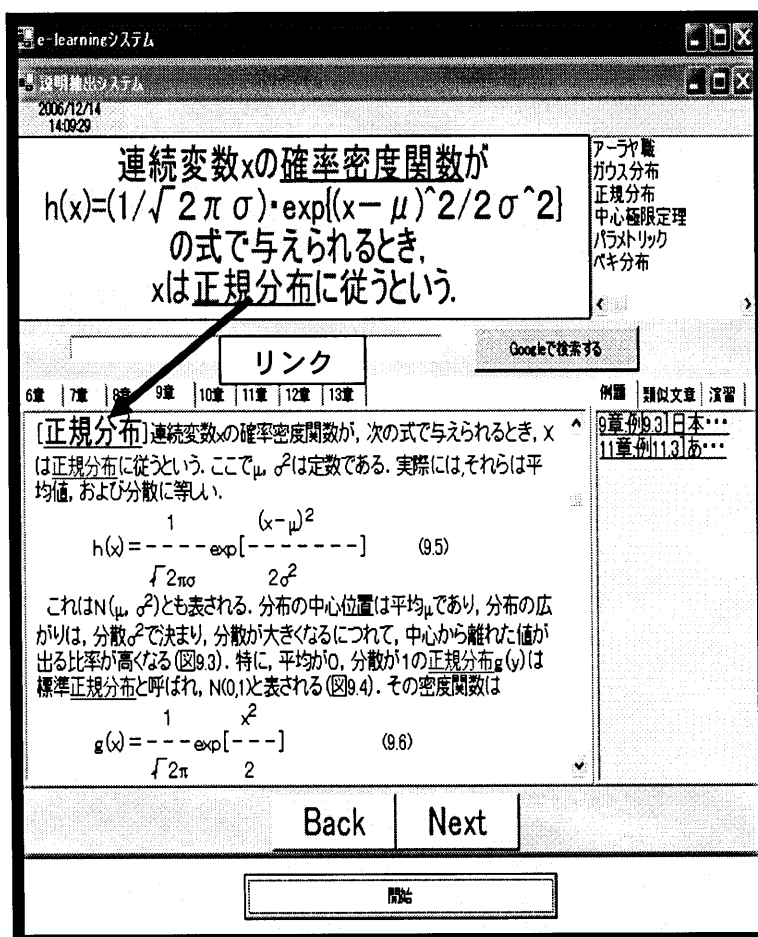


図14 リンクされたキーワードと補助説明の表示例

以上のようなシステムによって、受講者・講師それぞれにとって以下のような利点があると考えられる。

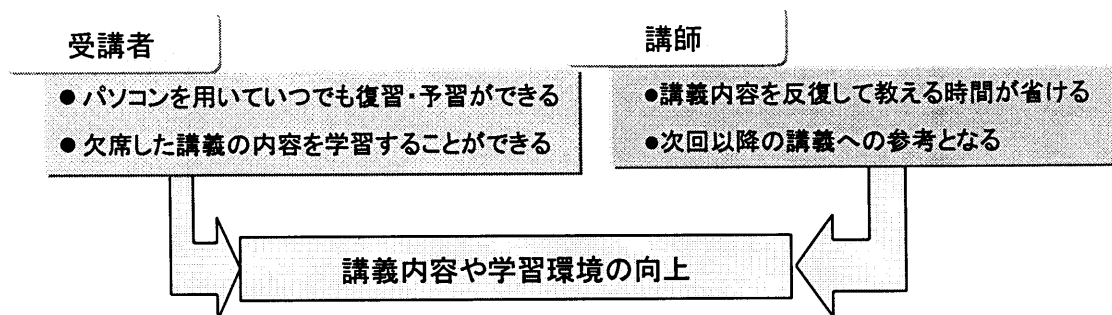


図15 本システムに期待される効果

## 2.2 従来システムの問題点

### 2.2.1 教室環境の問題点

従来システムは、ホワイトボードを用いた講義を想定している場合がほとんどである。しかし三重大学を含む大学の教室では、黒板を用いた講義がほとんどである。理由としては、大学の大教室ではホワイトボードの文字は読みにくく、またホワイトボードマーカーの寿命が短く、管理費がかかるからである。そのため、黒板での講義へのシステムの適応が一つと課題となっている。

この問題点を解決するために、本研究室では現在、実際の黒板での講義を対象としたシステムの開発を行っている[11]。

このシステムでは、従来システム同様、固定視野カメラ（Canon 社製 VB-C50Fi, ズーム可能）、可動視野カメラ（Canon 社製 VB-C50i, パン・チルト・ズーム可能）の2台のネットワークカメラを用いて講義映像を撮影している。また、ホワイトボードでの講義ではペン位置検出装置を用いる事で板書画像を取得していたが、このシステムでは、ペン位置検出装置を用いる事ができない。そこで、新たに可動視野カメラ（Canon 社製 VB-C10）を用いて板書画像を取得する予定である。

本研究では現在、ホワイトボードの板書を対象として研究を行っているが、後述するが、このシステムのための黒板の板書への対応も可能である。

利用するカメラを図16に、システムの概要図を図17に示す。



図16 教室の後部の壁に設置した IP カメラ

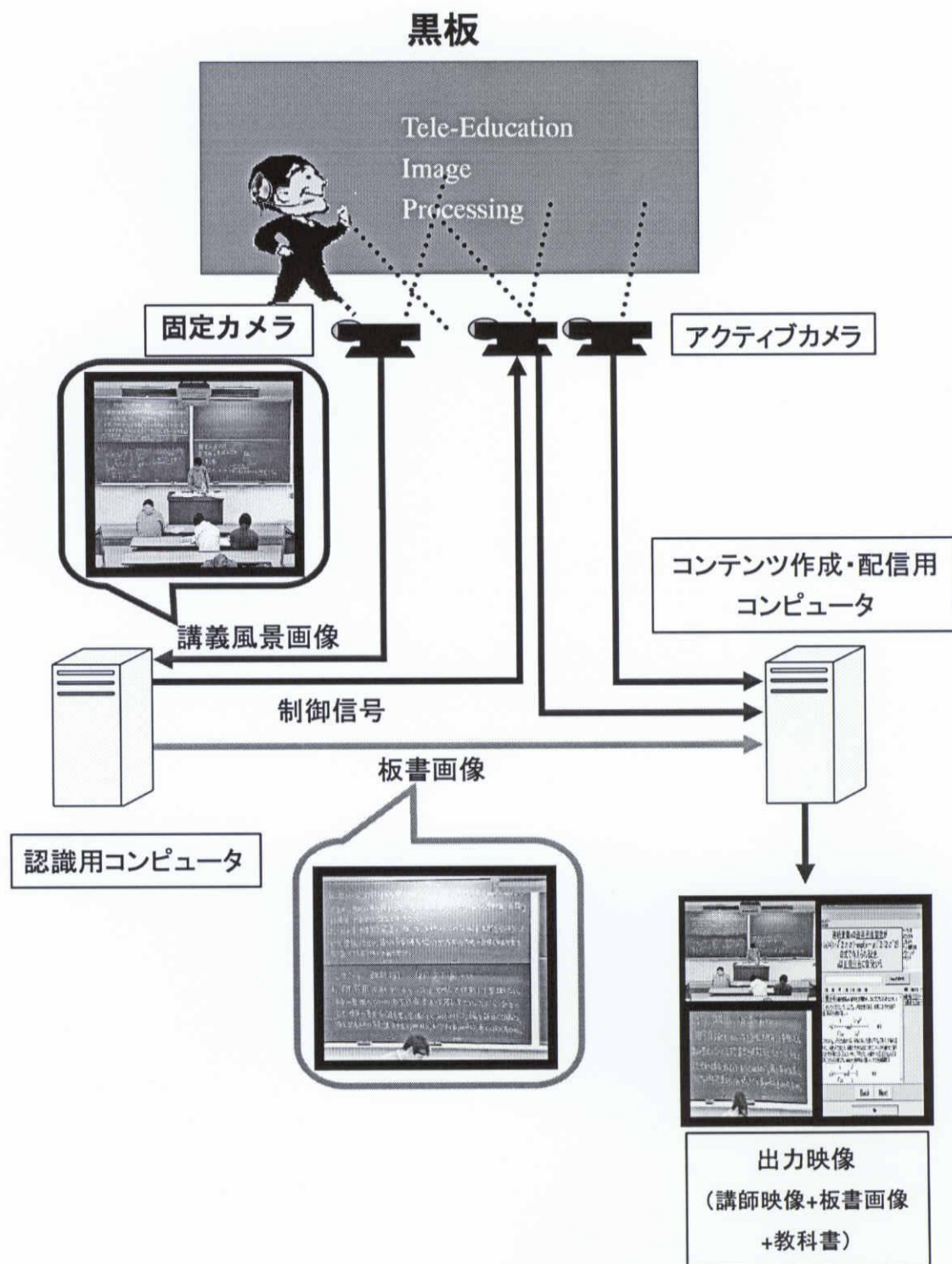


図17 黒板を対象としたシステムの概要

### 2.2.2 システム上の問題点

2.1 節で述べたように、従来システムでは文字認識にペン位置検出装置 (mimio) 用の文字認識ソフトウェア (Script Converter) を用いていた。そのため以下のような問題がある。

#### (1) システムの完全自動化が不可能

文字認識ソフトウェアのソースプログラムが公開されていないため、開発プログラムを文字認識ソフトと関係させる事が困難である。そのため、板書データが蓄積される度に文字認識ソフトウェアを起動させる人員が必要となるため、将来構想として考えている、同期型システムのためのシステムの完全自動化が不可能である。また非同期型システムの場合にも完全自動化は不可能で作業人員の分だけ余分にコストが必要となる。

#### (2) 認識精度の問題

Script Converter による文字認識では、活字のような字形で書かれた漢字の認識率は高いが、ひらがな・カタカナのような画数が少なく変動の大きい文字の認識率は低い。さらに、“偏”と“つくり”が分離されて記述されている漢字等の認識率が不十分である。これは、Script Converter はブロック体の英字用に開発されたアルゴリズムであるため、文字列から各文字を切り出す「個別文字の切り出し」の技術が不十分であり、一つの文字が複数の文字に分割されてしまうため、認識率が低下するものと考えられる。また、従来システムにおいてキーワード認識をする際には、キーワードとキーワードファイルとのキーワードが完全に一致する場合のみが認識されるため、キーワード認識率が低い。

Script Converter による文字認識例を以下に示す。

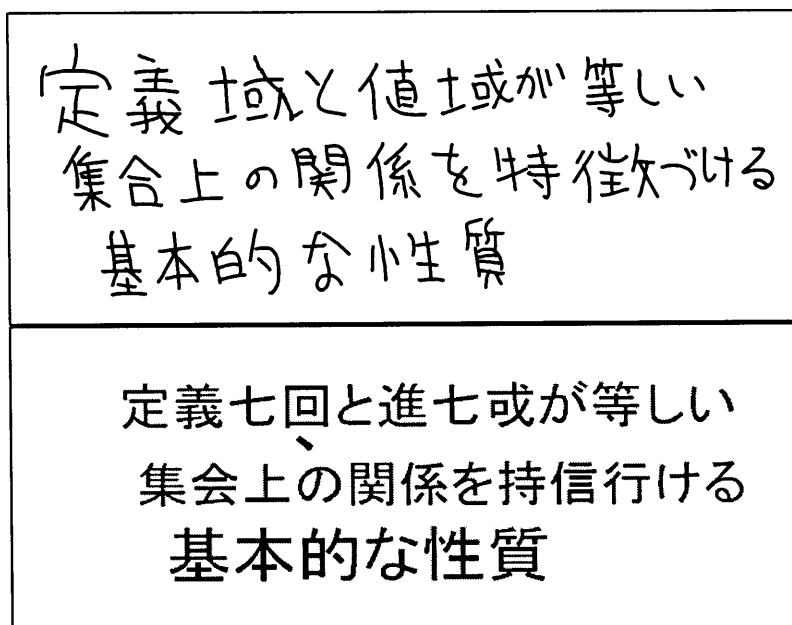


図18 Script Converter による文字認識結果 1

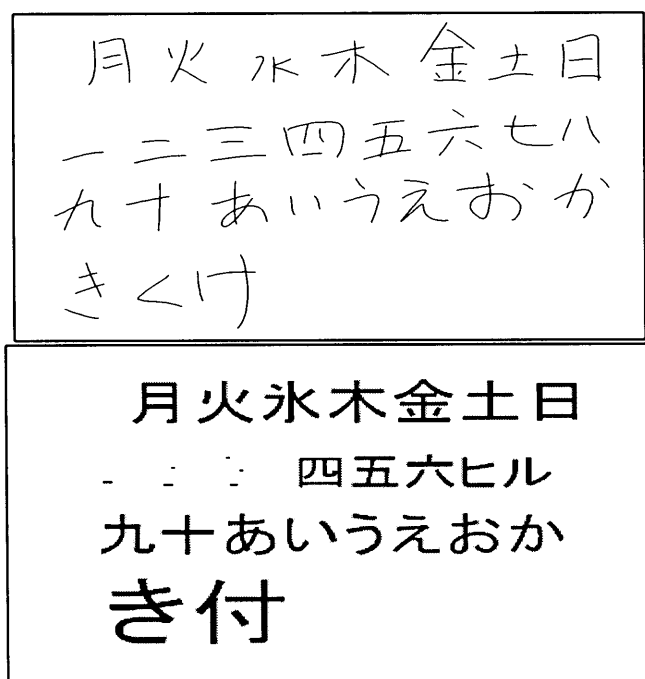


図19 Script Converter による文字認識結果 2

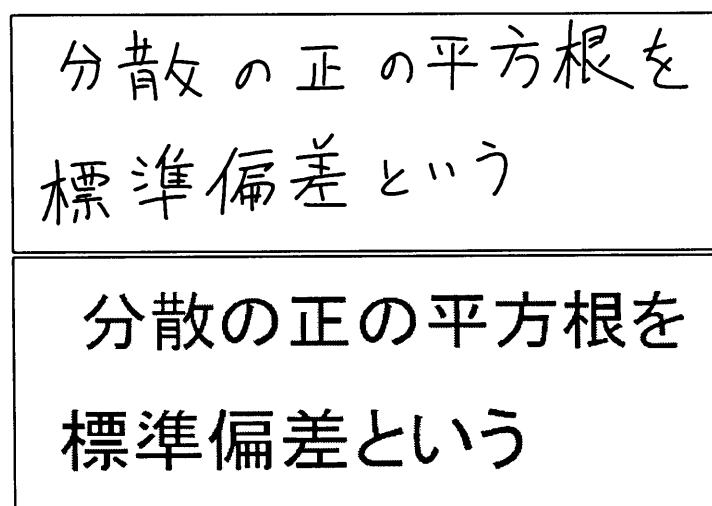


図20 Script Converter による文字認識結果 3

図 18 では個別文字切り出しのミスにより，‘域’が，‘七’と‘回’，‘七’と‘或’に分離して認識している．また“徴づ”が‘信’と‘行’に認識されている．図 19 では“くけ”が 1 文字として認識され‘付’となっている．また，“一，二，三”は漢数字でなく記号として認識され，‘七’，‘八’は漢数字をカタカナ‘ヒ’‘ル’と誤認識している．図 20 では活字体のような字形であり，全て正しく認識している．

### (3) ソースが非公開

(1)や(2)で挙げた事項にも関連するが、この Script Converter は市販のソフトウェアであるがゆえにソースが非公開で上記に挙げたような問題点を解決する改善手法を用いる事が出来ない。この問題が最も致命的である。

以上の事から、ソースが公開されている文字認識アルゴリズムを用いて、新たにシステムを構築する必要がある。

## 第3章 加重方向指数ヒストグラム法を用いた文字認識

### 3.1 加重方向指数ヒストグラム法の概要

前章で挙げた問題点の解決手法として，本研究では文字認識手法として，本研究室で研究開発された“加重方向指数ヒストグラム法[12][13]”を用いて個別文字認識を行う．以下に個別文字認識を行うまでの処理の流れを示す．

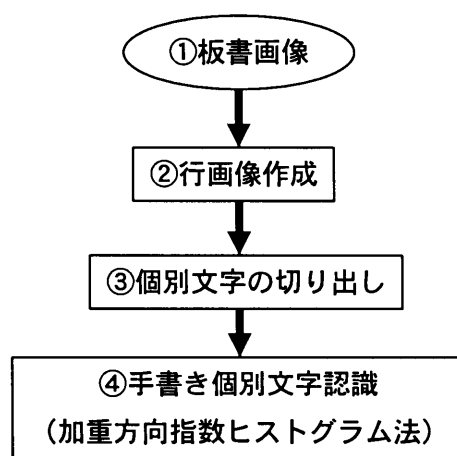


図21 個別文字認識までの処理の流れ

#### ① 板書画像の取り込み

従来法同様，ホワイトボードに筆記された板書を，ペン位置検出装置を用いてリアルタイムで Bitmap 形式（720×480pixel，32bits/pixel）計算機上に取り込む．

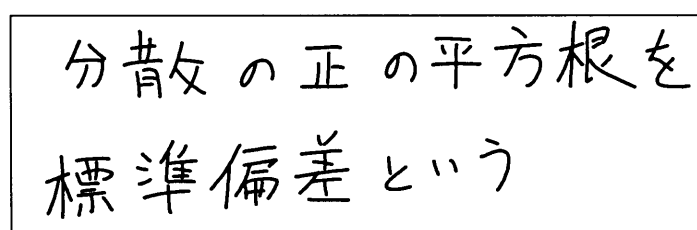


図22 計算機上に取り込んだ板書画像

## ② 行画像作成

取り込んだ板書画像から，文字領域を抽出し，各行ごとに画像（以下，行画像）を抽出する．

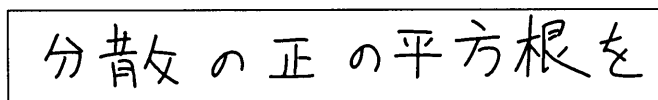


図23 作成された行画像

## ③ 個別文字の切り出し

作成された行画像に対して，図 24 のように垂直方向への黒画素の和をとった射影ヒストグラムを求め，そのヒストグラムの値が閾値以下の位置を個別文字切り出し位置とし，個別文字の切り出しを行う．

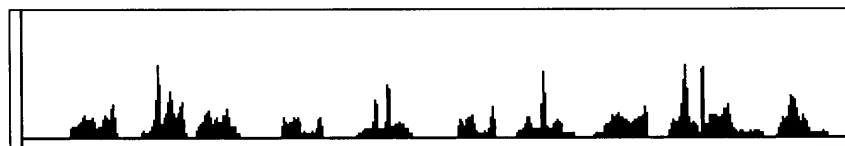


図24 垂直方向への黒画素の和を求めた射影ヒストグラム

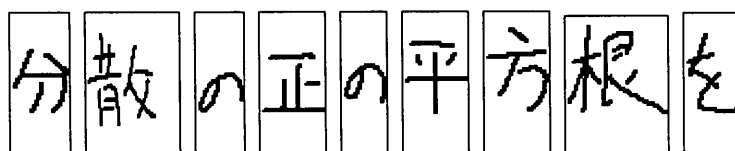


図25 切り出された個別文字

## ④ 手書き個別文字認識

加重方向指数ヒストグラム法を用いて文字認識を行う．次項にその詳細を示す．

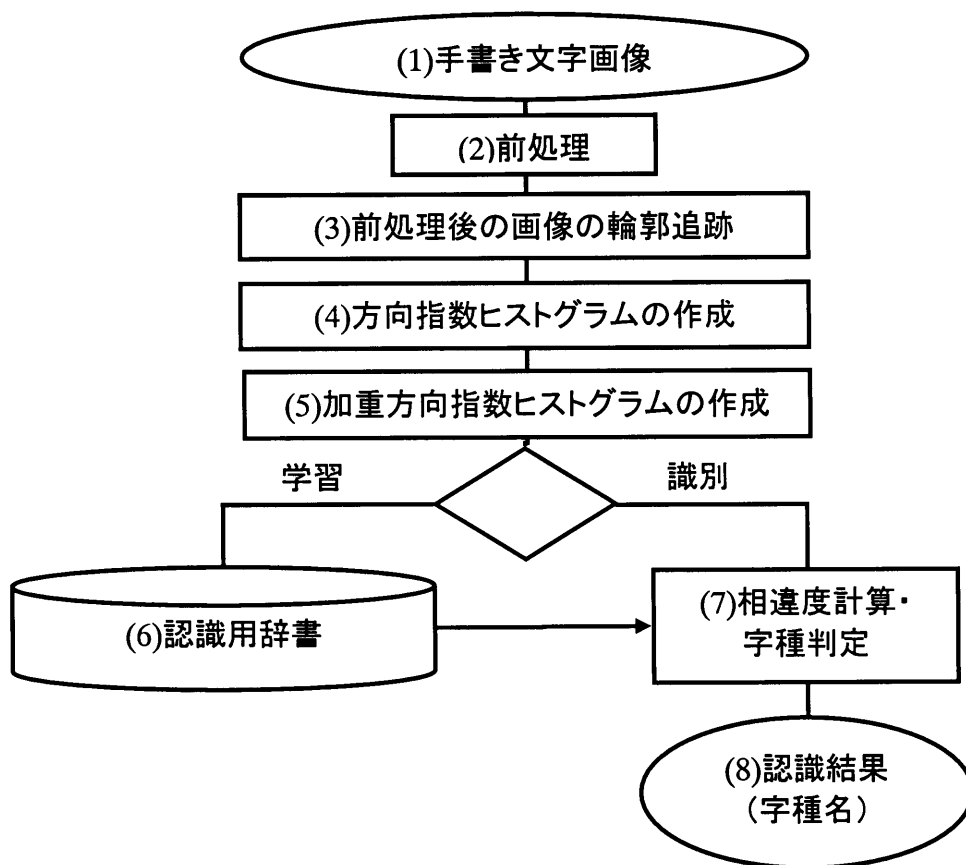


図26 加重方向指数ヒストグラム法の処理過程

(1) 手書き文字画像

入力として，図 27 のような 8bit の Bitmap 形式の個別文字画像を入力する．

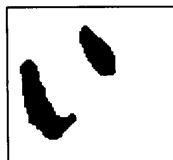


図27 入力画像 (8bit Bitmap 形式)

## (2) 前処理

前処理として、入力画像に対し、以下の処理を行う。

## (ア)2 値化

入力画像に対して 2 値化を行い、2 値化画像を作成する。

## (イ)ノイズ除去

注目画素及び 8 近傍画素（図 28）のうち、5 画素以上が黒画素ならば注目画素を黒画素に、4 画素以下ならば注目画素を白画素にする。

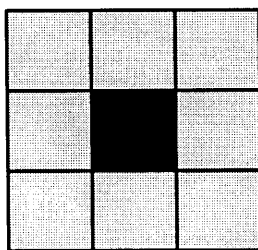


図28 8 近傍（黒色：注目画素 灰色：近傍画素）

## (ウ)正規化

ノイズ除去した画像に対して、文字の中心が画像の中心になるように 70×70 画素で画像の位置・大きさの線形正規化をする。

## (エ)平滑化

注目画素の 4 近傍（図 29）に黒画素が 3 画素以上存在する白画素を黒画素に変換し、8 近傍すべてが白画素である黒画素を白画素に変換する。

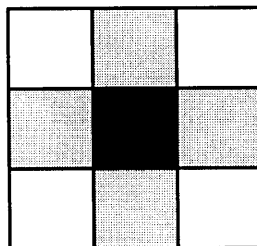


図29 4 近傍（黒色：注目画素 灰色：近傍画素）

以上の処理によって前処理後の画像は図 30 のようになる。



図30 前処理後の画像

### (3) 前処理後の画像の輪郭追跡

正規化後の文字パターンに対し、輪郭追跡を行い、各境界で輪郭線の方角を 4 方向に量子化した方向指数を求める。

### (4) 方向指数ヒストグラムの作成

図 31 のように文字パターンを  $7 \times 7$  の小領域に等分割して、それぞれの小領域において、4 方向の方向指数ヒストグラムを作成する。

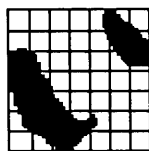
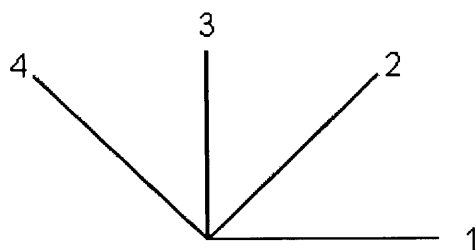
図31  $7 \times 7$  の小領域に分割された入力画像

図32 方向指数 (4 方向)

### (5) 加重方向指数ヒストグラムの作成

文字パターンが受けている変動を吸収するために、(4)で作成した方向指数ヒストグラムをぼかし、集約する。

空間的集約としては、図 33 に示すような図中灰色の小領域を中心として、図 34 のようなガウス形の重み係数を持つ空間フィルタによって方向指数ヒストグラムを  $7 \times 7$  個から  $4 \times 4$  個に集約した図 35 のような加重方向指数ヒストグラム (64 次元の特徴ベクトル) を得る。

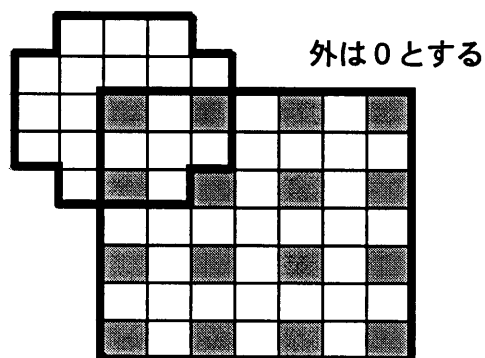
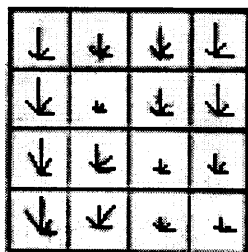


図33 空間的集約の中心点の位置

	0.009	0.017	0.009	
0.009	0.057	0.105	0.057	0.009
0.017	0.105	0.194	0.105	0.017
0.009	0.057	0.105	0.057	0.009
	0.009	0.017	0.009	

図34 空間的加重フィルタ (数値は重みを示す)

図35 加重方向指数ヒストグラム ( $4 \times 4 \times 4 = 64$  次元の特徴ベクトル)

## (6) 認識用辞書

この特徴ベクトルを各文字から抽出し、 $N$  個の 1 字種の学習文字から字種ごとに共分散行列を求め、この行列の固有値、固有ベクトル ( $\lambda_i, \Phi_i; i=1, 2, \dots, k (k < 64)$ ) を標準パターンとし、認識用辞書に登録する。

## (7) 相違度計算・字種判定

学習で得られた標準パターンを用いて相違度計算を行い、入力文字を識別する。識別関数としては、(1) 式の疑似ベイズ識別関数を用いる。

$$g'(x) = \sum_{i=1}^k \frac{\{\Phi_i'(x - \mu)\}^2}{\lambda_i} + \sum_{i=k+1}^n \frac{\{\Phi_i'(x - \mu)\}^2}{\lambda_{k+1}} + \ln\left(\prod_{i=1}^k \lambda_i \cdot \prod_{i=k+1}^n \lambda_{k+1}\right) \cdots \quad (1)$$

ただし、 $k < 64$ （特徴ベクトルの次元数）とする．実験では  $k=15$  とした．

#### (8) 認識結果

⑦の相違度計算の結果，最も小さな値となる字種（文字コード）を認識結果とする．

表1 文字認識結果

順位	文字候補	相違度
1	い	59
2	じ	87
3	ど	122
4	購	136
5	心	137
...	...	...

← 認識結果

### 3.2 紙とホワイトボード上との文字認識結果の比較

これまでに、加重方向指数ヒストグラム法を用いた文字認識は、JIS 第1水準手書き漢字データベース ETL9 で正読率 99.50%，旧郵政省が作成した手書き郵便番号数字データベースで正読率 99.17%を誇るなど、紙上の文字認識においては非常に有効な手法である事が確認されている。

ここで、本研究の分野に適用すると、ホワイトボードや黒板等、板書した手書き文字は、紙の上に筆記された文字と比べ、以下のような要因がある。

- 腕が肘で固定できず、ペンやチョークが滑りやすい
- 筆記者が板書に慣れていない
- ホワイトボードや黒板では、紙に比べペンやチョークが滑りやすい

以上の事から、紙の上に筆記された文字に比べ、不安定な文字となり、認識率が悪くなると考えられる。これは、従来法において、Script Converter の認識用辞書がホワイトボードに筆記した文字から学習したにも関わらず、キーワード認識率が実用に耐えない事からも分かる。以下に、同一筆記者によって紙・ホワイトボード上に筆記された同じ文を例として示す。

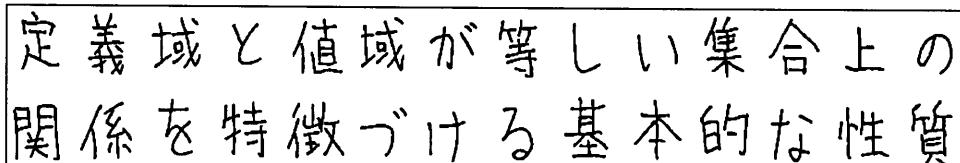


図36 紙に筆記された文字列

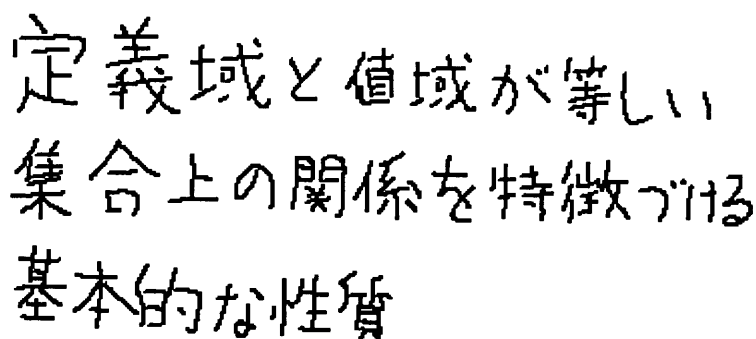


図37 ホワイトボードに筆記された文字列

図 36, 37 から, 同一筆記者でも, 紙上・ホワイトボード上では異なったストロークとなる事が分かる.

そこでまず予備実験として, 以下の実験環境で同一筆記者によって, 紙上とホワイトボード上に筆記された文字に対して, どれほどの違いが出るのかを実験した.

#### [実験環境]

##### (1) 実験用パソコン

OS : Windows XP SP2

CPU : Pentium(R)4 2.4GHz

メモリ : 512MByte

開発環境 : Microsoft Visual Studio .NET2003

##### (2) 筆記者数 6 人 (研究室の卒業研究生と大学院生)

##### (3) 収集文字

紙上, ホワイトボード上それぞれ同様の文字 104 字, 計 208 字/人

##### (4) イメージスキャナ

機種 : EPSON GT-7200

解像度 : 1200dpi

##### (5) 認識用辞書

JIS 第 1 水準手書き漢字データベースから学習した辞書 (3072 カテゴリ, 200 パターン, 64 次元)

#### [実験結果]

実験結果より, 以下に認識率の違い (表 2), 紙の方が正読率の良かった例 (図 41, 42), ホワイトボードの方が正読率の良かった例 (図 43, 44) をそれぞれ示す.

表2 紙上・ホワイトボード上の認識精度の比較

筆記者	正読率[%]	
	紙	ホワイトボード
A	80.8	91.3
B	86.5	76.0
C	95.2	83.7
D	71.2	78.8
E	79.8	72.1
F	88.5	81.7
平均	83.7	80.6

紙	ホワイトボード
確率を信頼係数値を信頼限界 区間を信頼区間という	石確率を信頼係数 値を信頼限界 区間を信頼区間という
確率を信頼係数 値を信頼限界 区間を信頼区間という	醜妻を信頼係数 値を信頼限界 区間を信頼区間という

図38 紙の方が正読率の良かった例1

紙	ホワイトボード
この確率分布を同時確率分布という この確率分布を同時確率分布という	この確率分布を 同時確率分布という
	この確率分布を 同時確率分布という

図39 紙の方が正読率の良かった例2

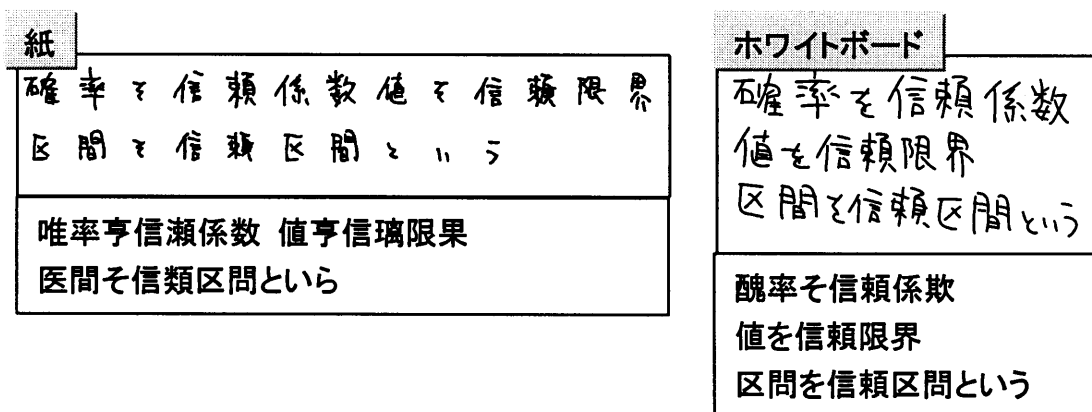


図40 ホワイトボードの方が正読率の良かった例1

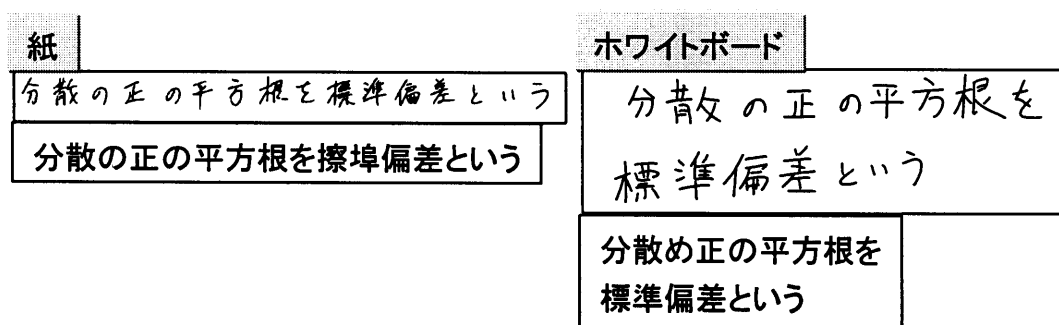


図41 ホワイトボードの方が正読率の良かった例2

表2から、紙上の正読率は最大95.2%、最小71.2%、平均で83.7%、ホワイトボード上の正読率は最大91.3%、最小72.1%、平均で80.6%となった。平均値で比較すると、紙上の方がホワイトボード上に比べ3.1%正読率が良いという結果になった。この結果だけみると、紙上・ホワイトボード上共にそれほど違いは無いと考えられるが、ここでホワイトボード上の正読率が紙上の正読率を上回った筆者A・Dに着目する。

筆者A・Dに対して以下の共通点がある。

- ホワイトボード上に筆記した経験が1年以上と他の筆者に比べ経験期間が長い
- ペン位置検出装置 mimio を扱った経験があり、スタイラスの操作に慣れている

図38～41からも分かるように、どの筆者においても紙上、ホワイトボード上ではそれぞれ異なったストロークとなる。ストロークが異なる原因の一つとして、ホワイトボードに筆記した経験が無い事で不安定な文字になりやすいと言えるが、上記の実験の結果から、筆記する経験を重ねる事で、ホワイトボードでの筆記に慣れ、ストロークの不安定さが吸収されるものと考えられる。同様に mimio のスタイラスでの筆記には慣れが必要で、操作になれていないと、スタイラスをホワイトボードに押しつける圧力が弱く、超音波が発信

されず、以下のようなストロークが途中で欠けてしまう場合も存在する。そのような文字を文字認識した場合も当然認識率は低下する。

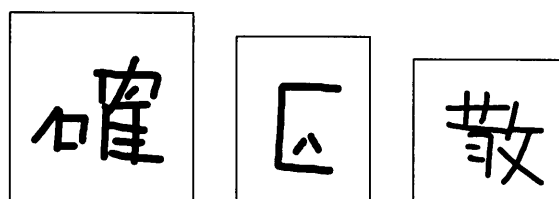


図42 ストロークの欠けた文字

また、筆記者にはそれぞれ固有の“書き癖”があるが、これを汎用辞書で吸収する事は困難である事がこれまでの研究[14]で明らかになっている。実験の結果のように、ストロークが異なる事で認識率に悪影響を及ぼしているケースがほとんどであると考えられるが、一部例外として、“書き癖”がストロークの違いにより、吸収される例も存在する事が分かった。表3にその例を示す。

表3 紙上・ホワイトボード上のストロークの違い

筆記場所	紙	ホワイトボード
入力文字例		
入力文字数	6	6
正解数	0	5
“そ”と認識	3	1

表3より、どちらも入力文字は“を”であるが、書き慣れた紙上では筆記者の“書き癖”が色濃く出ており、人間が見ても“を”なのか“そ”なのか判別する事が難しい。また、その影響が認識結果に反映されている。この例の様にホワイトボードに筆記することで、筆跡が変わり、癖が吸収される場合も存在する。即ち、ホワイトボードに筆記した場合の方が丁寧な文字に見える場合もごく一部であるが存在する。

結論として、本実験の結果のように、紙上の認識率よりもホワイトボード上の認識率が良くなる筆記者は例外的なケースと考えられる。さらに、本研究の目指すシステムは、ホワイトボードの筆記の経験の有無に影響されず汎用性の高いシステムを目指すのが望ましいと考える。いずれにせよ、文字認識率は100%には及ばず、単語辞書を用いない文字認識単体でのキーワード認識はeラーニングシステムには不十分な精度であると考えられる。

## 第4章 単語辞書を用いたキーワード認識

### 4.1 単語辞書を用いない手法の問題点

これまでに述べたように、従来システムで用いていた単語辞書を用いない基本的なキーワード認識では、文字認識部から得られる結果に対して、相違度の最も低い 1 位の文字候補を文字認識の結果としてキーワード認識を行う。そのため、文字認識結果とキーワードファイル中のキーワードとが完全に一致しないとキーワードを認識する事が出来ない。

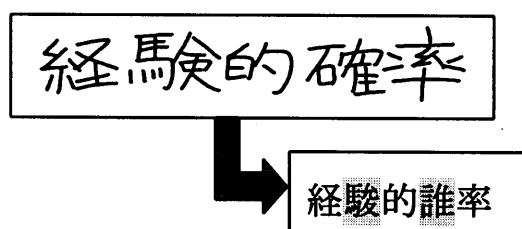


図43 キーワードが認識できない例 1

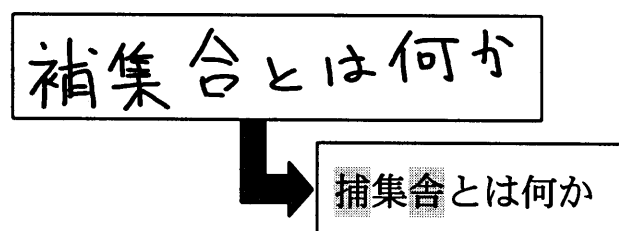


図44 キーワードが認識できない例 2

図 43, 44 を用いて説明する。それぞれ上が入力文字画像、下がその文字認識結果である。図 43 のケースでは“経験的確率”というキーワードを、図 44 のケースでは“補集合”というキーワードを認識したいが、それぞれ図中灰色の部分の文字が誤認識されているために、どちらのキーワードも認識する事が出来ない。そのため従来システムでも、キーワードを構成する文字のうち、1 文字でも認識結果が異なれば、キーワードを認識する事ができなかった。

また、前節で行った予備実験の結果より、本研究での文字認識手法においても、ホワイトボードに筆記された文字は紙上に筆記された文字に比べ認識率が低下する事からも、文字認識のみでのキーワード認識では十分なキーワード認識率が得られないと考えられる。

以上により、文字認識結果とキーワードが完全に一致せずともキーワードが認識できるようなロバスト性の高い手法が必要となる。

## 4.2 単語辞書を用いたキーワード認識とは

単語辞書を利用した文字認識とは、個別文字認識の後処理として行われ、品詞や単語等の言語情報を利用する事によって文字認識部の誤りを吸収・訂正し、精度を向上させようとするものである。応用分野として、主に郵便区分機などの住所認識[15][16][17]や設計図入力装置、ガス管・電線・水道管の設計図面入力装置、電子回路図面等の専門分野ごとのOCR(Optical Character Reader：光学式文字読取装置)に用いられている[18]など、「工(漢字)」や「エ(カタカナ)」,「テ(カタカナ)」と「〒(郵便番号)」,「ト(漢字)」と「ト(カタカナ)」,「二(カタカナ)」と「二(漢字)」,「パ(半濁点)」と「パ(濁点)」,「カ(カナカナ)」と「か(ひらがな)」,「力(漢字)」などのように類似した字種の多い日本語において非常に有効な誤り訂正手法である事が知られ、主に紙上での単語認識に用いられてきた。また、単語辞書を用いるための条件として文字認識部から複数の字種候補が与えられる事が挙げられ、この文字認識手法にも適合している。そこで、ホワイトボード上でのキーワード認識にも単語辞書を用いる事が有効ではないかと考える。

実際の郵便区分機の例で、単語辞書認識の概要を簡単に説明する。図45の入力文字に対し、文字認識結果が表4のようになったとする。また単語辞書として都道府県名が登録されたものを用いる。

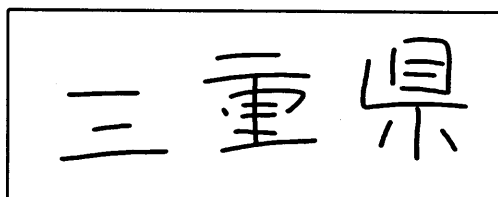


図45 入力画像

表4 文字認識結果

入力画像	類似度による字種の候補				
	1位	2位	3位	4位	5位
三	三	こ	重	夏	二
重	童	重	空	量	室
県	県	得	粟	呉	黒

単語番号	単語
1	愛知県
2	愛媛県
.	
.	
47	和歌山県

図46 単語辞書

この例の場合、認識対象となるのは住所である。ここで、個別文字認識のみで認識文字を判定すると、第1位の字種候補が認識文字となるので『三童県』となり誤認識となるが、ここで図47のような都道府県名が登録された単語辞書を参照し利用する字種候補を第5位の字種候補までに拡張すると、候補文字に『三重県』となる組み合わせが存在する。これにより、2文字目の誤認識文字『童』を訂正し『三重県』と正しい認識結果を得る。

これは、人間の感覚で説明すると、図47のように、個別文字だけではどの文字なのか判別に迷う場合でも、事前情報として「工学部」という単語を知っていれば、前後情報からこの文字が「工」とであると分かる。このように人間の感覚に近い理にかなった手法であるといえる。

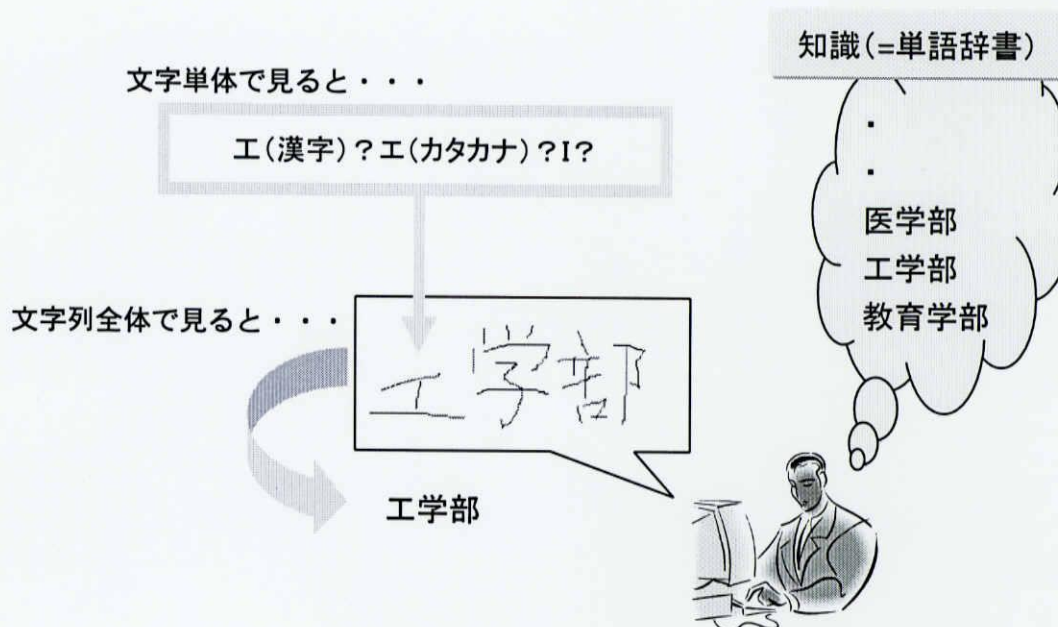


図47 人間の感覚での単語辞書によるキーワード認識

### 4.3 単語辞書を用いたキーワード認識法の概要

本研究では、単語辞書としてキーワードファイルを用いる。単語辞書を用いたキーワード認識の処理の流れを以下に示す。

step. 1 文字認識部から、10 位までの文字候補とその相違度を得る。

同		時		確		率		分		布	
文字候補	相違度										
同	82	時	103	確	103	幸	88	分	82	命	89
周	103	晴	113	確	117	率	98	命	94	希	91
月	116	靖	120	雄	123	章	104	弁	101	布	97
国	120	埼	122	礎	123	華	107	斉	105	帝	107
用	123	辞	125	椎	130	辛	108	合	106	侍	107
岡	128	涛	125	曜	133	卒	111	杏	107	奇	109
向	129	瞬	126	権	133	革	114	余	108	席	110
団	133	持	126	竣	137	宰	116	各	112	序	111
図	138	掠	128	腫	139	牽	119	斧	112	弄	114
回	142	碯	129	瞳	139	牢	120	帝	112	寄	114

(灰色：正解文字)

図48 入力文字と文字認識結果

step. 2 第N位の文字候補と1位の文字候補との相違度の差が30以上となれば、それ以降の認識結果をリジェクトとして破棄し、以降の文字候補を探索しない。

step. 3 単語辞書内のキーワードが文字候補中に存在するか探索する。その結果、以下の2つの場合、そのキーワードをキーワードとして認識する。

(ア) 構成文字中の75%の文字が存在する場合。

(イ) 単語辞書中の探索するキーワードを構成する文字のうち1文字以上が、探索を行う文字候補中の1位に存在し、かつ2位の文字候補との相違度の差が30以上であり、それに加え、構成文字中の60%の文字が存在する場合。

ここで(イ)の60%に設定した理由としては、1位と2位の文字候補の相違度差が大きいという事は、それ以外の文字候補に比べその1位の文字候補に対する信頼度が高

いという考えで、(ア)の場合よりも、各文字の信頼度が高いと考え、よりキーワードを認識しやすくするため、75%から60%に制約を緩めた。

図48の例の場合、単語辞書内に『確率』『確率分布』『同時確率分布』というキーワードが登録されていた場合、それぞれ認識される事になる。

step. 4 同じ文字候補中で複数のキーワードの組み合わせが存在する場合、(2)式で定義する評価関数を用いて評価値を算出し、評価値が最大となるキーワードとして認識する。

$$V(a_1, a_2, \dots, a_M) = M - c \sum_{i=1}^M g(a_i) \quad \cdots (2)$$

ここで、

$$\left\{ \begin{array}{l} V(a_1, a_2, \dots, a_M) : \text{単語『} a_1, a_2, \dots, a_M \text{』の評価値} \\ M : \text{単語の文字数} \quad c : \text{重み係数} \\ g(a_i) : a_i \text{の認識結果の順位, 正解が含まれない場合 } p(\text{定数}) \text{を与える.} \end{array} \right.$$

係数  $c$ ,  $p$  については、実験により正解が多くなる適切な値を決めたため、 $c$  には 0.3,  $p$  には 14 を与えた。

図48の例では、『確率』『確率分布』『同時確率分布』の3つのキーワードがstep.3までに認識されているが、ここで評価値の計算の結果、『確率』の評価値は1.1、『確率分布』の評価値は1.9、『同時確率分布』の評価値は3.3となり、『同時確率分布』がキーワードとして認識される。

また、これまでに相違度を用いない手法（上記 step.1 で 5 位までの文字候補を利用，step.3(ア)，step.4）でキーワード認識を行って来た[19][20]が以下のような課題を得てきた。

- 意図しないキーワードをキーワードとして誤認識してしまう

例を示す。図49のように入力文字に対して文字認識結果が与えられたとする。

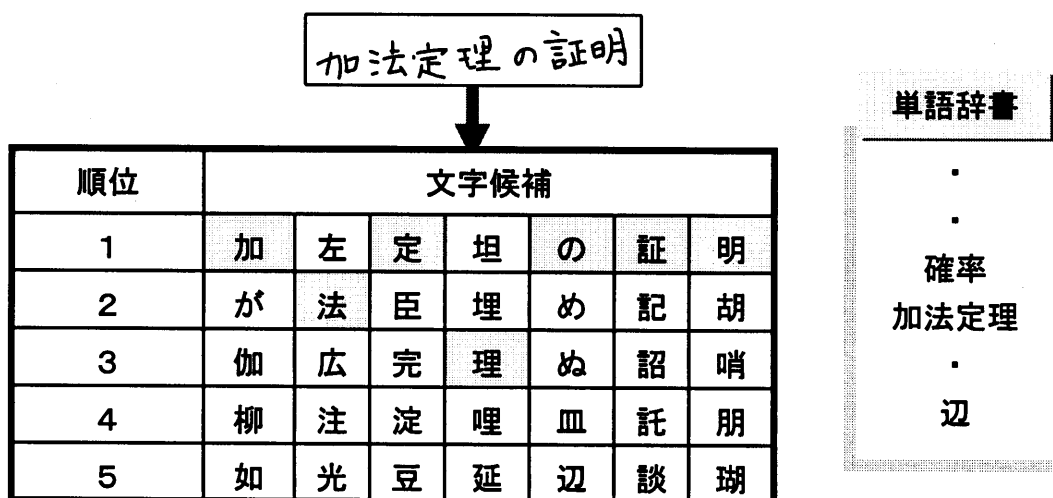


図49 意図しないキーワードを認識する例

この例では認識したいキーワードは“加法定理”であるが、加法定理は正解文字の組み合わせに存在するため認識される。しかし、入力文字“の”の文字候補中に“辺”という文字候補が存在し、単語辞書中にも“辺”というキーワードが存在する。結果として、“辺”もキーワードとして認識される。これは意図しないキーワードである。

このように相違度を用いない方法では、文字候補中に単語辞書内に該当するキーワードが存在すれば無条件で認識していたため、“辺”や“商”のような1文字単語のような文字数の少ないキーワードが誤って認識されてしまう危険性があり、利用する文字候補の数や単語辞書登録数が多いほどその危険性は高い。

この場合、システムでは、その板書では、そのキーワードに関する説明はされていないのにそのキーワードに関連する板書であるというインデキシングがされてしまう。

- 文字数の少ないキーワードが認識されにくい

例を示す。図 50 のように入力文字に対して文字認識結果が与えられたとする。

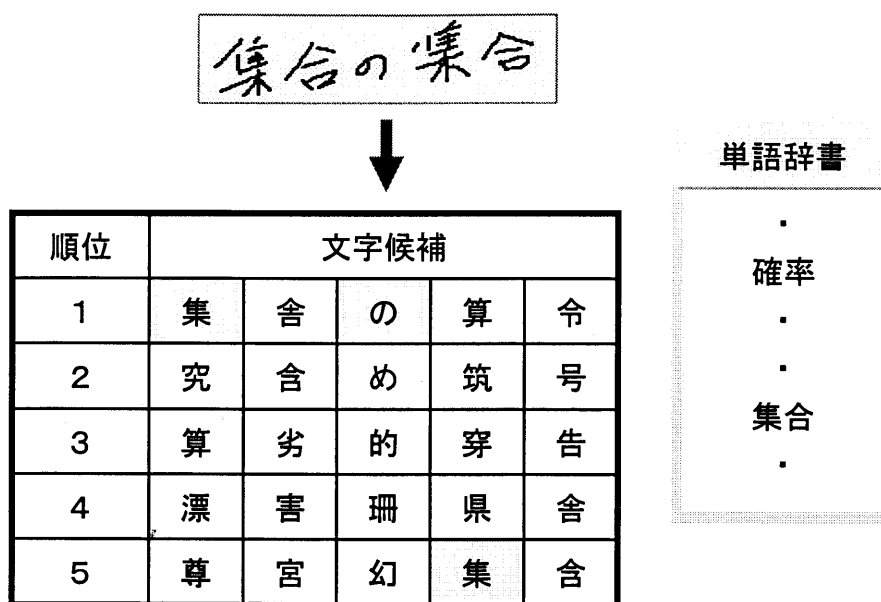


図50 キーワードが認識できない例

この例において認識したいキーワードは『集合』であるが、文字候補中にどちらの“合”という入力文字に対し、正解文字が存在しない。そのため、キーワードが認識できなかった。このように、文字数が2, 3文字とキーワードを構成する文字数が少ない単語では、1文字当たりの文字認識結果がより重要となり、1文字だけでも正解が含まれないと認識できなかった。

そこで、相違度をキーワード認識のアルゴリズムに加える事でより高精度に認識できると考え、次項に示す予備実験の結果を踏まえ、本手法で相違度を用いた手法を使用している。

# [予備実験]

## ● 実験内容

筆記者数 6 人のホワイトボード上に筆記された文字の単語辞書を使用しない文字認識結果に対して、統計的調査を行った。その内容と目的を以下に示す。

### ① 相違度の統計

内容：各入力文字の文字認識結果の 1 位の文字候補の相違度の値を、1 位が正解の場合の 501 パターン、不正解の場合の 122 パターンの計 623 パターンの統計をとる。

目的：相違度の値によって正解・不正解の判別が可能かどうか、及び統計的特徴が存在するかどうかを調べるため。

### ② 相違度差の統計

内容：1 位の文字候補と 2 位の文字候補との相違度の差の統計を取る（1 位が正解文字：501 パターン、不正解文字：122 パターン、計：623 パターン）。また、2 位以下に正解文字が存在する場合の 1 位の文字候補との相違度差の統計を取る（全 93 パターン）。

目的：相違度差によって、その文字を正解文字と判別する事が可能かどうか、もしくはそれに近い信頼性を設定し単語辞書利用に生かす事ができないかどうかを調査するため。

## ● 実験結果

### ① 相違度の統計

文字認識結果が 1 位の文字候補に対する相違度についての分布関数を図 51 に示す。

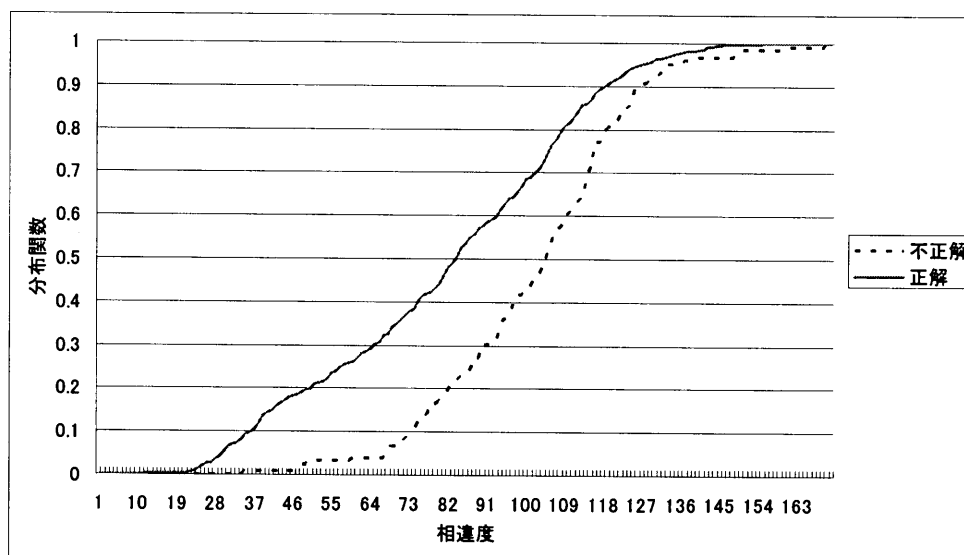


図51 1 位の文字候補の相違度の分布関数

図 51 より、正解文字は不正解文字に比べ、相違度の低い方の分布の割合が大きいと言えるが、不正解文字も割合が低いとは言え、相違度が小さい領域にも存在するため、ある相違度の値で正解・不正解を判別する事は難しいと考える。ただ、文字候補に何らかの信頼性を持たせる事は有効ではないかと考える。

## ② 相違度差の統計

統計結果より、不正解の 1 位の文字候補と 2 位の文字候補との相違度差の分布関数及び正解の 1 位の文字候補と 2 位の文字候補との相違度差の分布関数を図 52 に、2 位以下に正解文字が存在する場合、その文字候補の相違度と 1 位の文字候補との相違度差の分布関数を図 53 に示す。

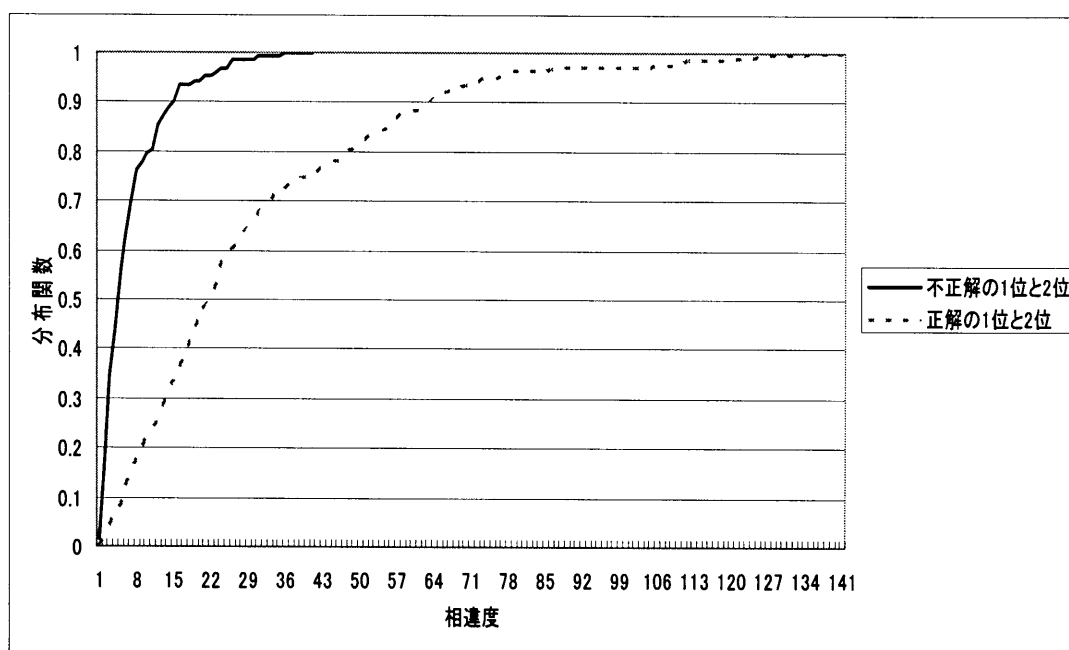


図52 相違度差の統計結果 1

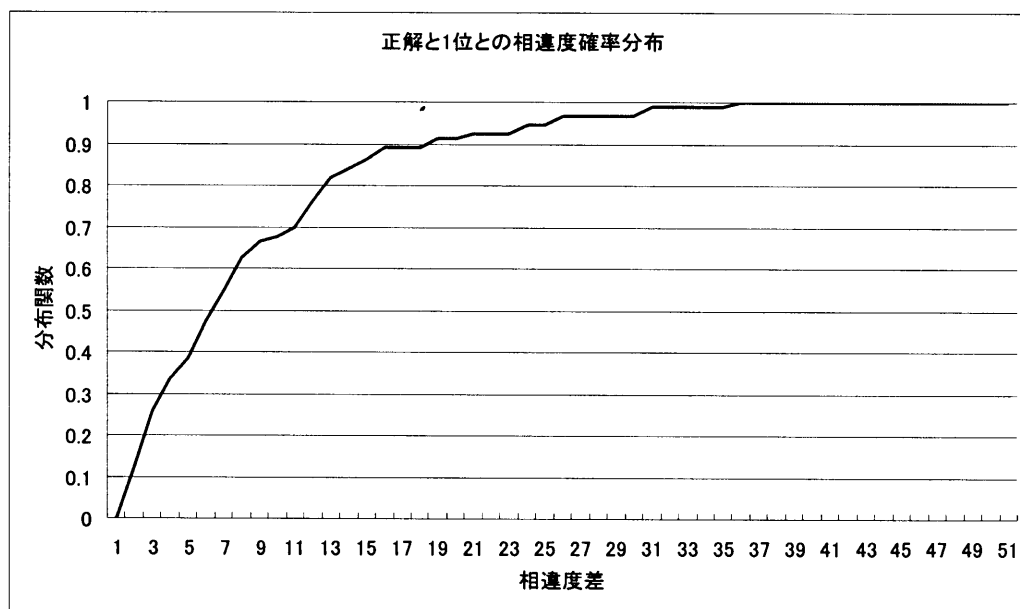


図53 相違度差の統計結果 2

図 52 より，1 位の文字候補が不正解だった場合，2 位との相違度の差が 35 以下の領域に全パターンが存在している．さらに，図 54 より，正解の文字候補が 1 位でなかった場合，その正解の文字候補と 1 位との文字候補との相違度差が 35 以下の領域に全パターンが存在している．この事より，1 位と 2 位の文字候補の相違度差が 30 以上ある場合，2 位以下に入力文字の正解文字候補が殆ど存在しない事から，相違度を用いた手法が，従来の順位のみを用いた手法よりもより信頼性の高い，高精度な単語認識ができるのではないかと考えられる．

以上の結果を踏まえて，本研究での手法では，利用する文字候補を 10 位まで使用し，step.2 及び step.3 (イ) の処理を加えた．

## 第5章 キーワード認識の評価実験及び考察

### 5.1 実験環境

キーワード認識法の評価実験として、以下の実験環境により、その有用性を確認した。

#### (1) 入力

以下に示すような、各文中に 1, 2 個程度を含むホワイトボードに筆記した文字列を入力として、6 人の筆記者、総キーワード数は 84 個で行った。

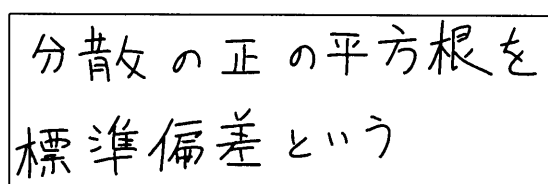


図54 入力文字画像の例

#### (2) 比較手法

##### ① 単語辞書を用いない従来のキーワード認識法

この手法では、文字認識部から与えられる相違度から最短経路を最適解とする（文字候補の第 1 位を利用する事に相当する）。その結果と、キーワードファイルのキーワードとを照合する。

##### ② 単語辞書を用いたキーワード認識法（4.3 節で述べた方法）

#### (3) 単語辞書

教科書“情報数学”（オーム社）[21]及び教科書“デジタル画像処理の基礎と応用”（CQ 出版社）[22]の 2 冊の教科書の索引単語、計 406 語を登録したものを用いた。

## 5.2 実験結果

実験結果を表5示す。また、キーワード認識の成功例（図55～57）、失敗例（図58～60）をそれぞれ示す。また、キーワード認識率及び文字認識率は以下の式で定義する。

$$\text{キーワード認識率}[\%] = \frac{\text{認識に成功したキーワード数}[\text{個}]}{\text{総キーワード数}[\text{個}]} \times 100 \dots (3)$$

$$\text{文字認識率}[\%] = \frac{\text{1位で認識した正解文字数}[\text{個}]}{\text{総文字数}[\text{個}]} \times 100 \dots (4)$$

表5 キーワード認識実験結果

筆記者	単語数 [個]	単語辞書を用いない キーワード認識		単語辞書を用いた キーワード認識		文字認識率 [%]
		正読した 個数[個]	認識率 [%]	正読した 個数[個]	認識率 [%]	
A	14	4	28.6	13	92.9	76.0
B		1	7.1	11	78.6	72.1
C		7	50	14	100.0	83.7
D		2	14.3	14	100.0	78.8
E		9	64.3	13	92.9	91.3
F		4	28.6	14	100.0	81.7
平均		4.5	32.1	13.2	94.0	80.6

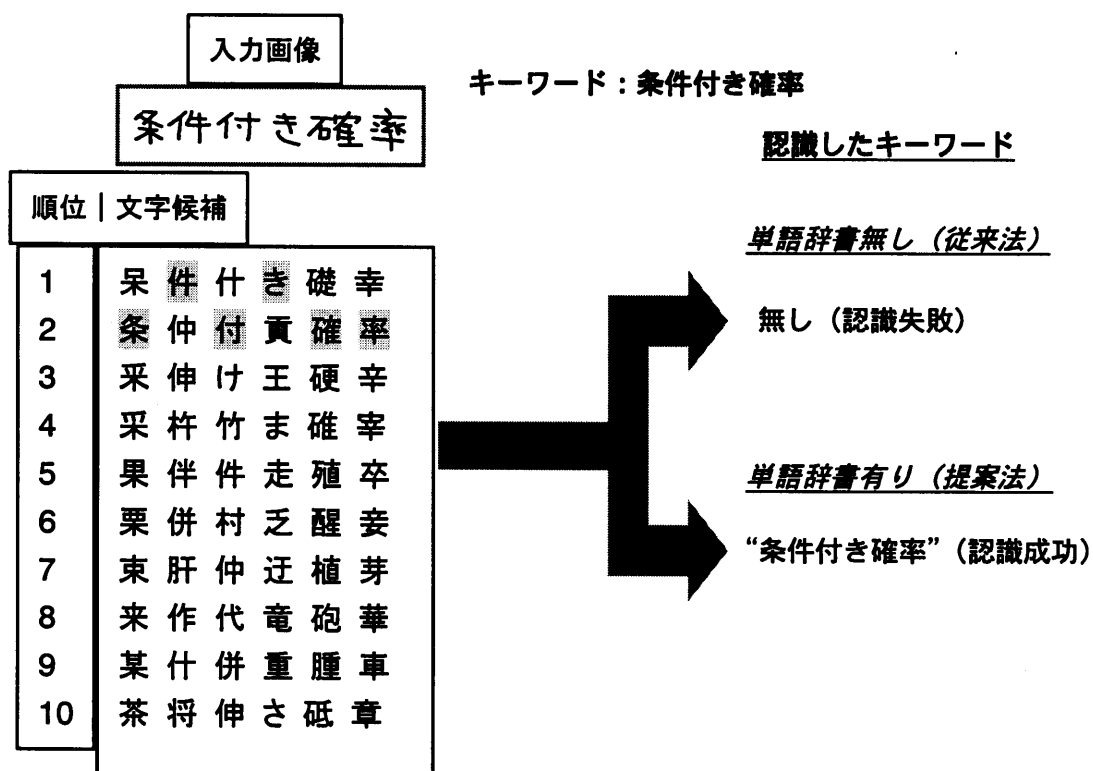


図55 成功例 1

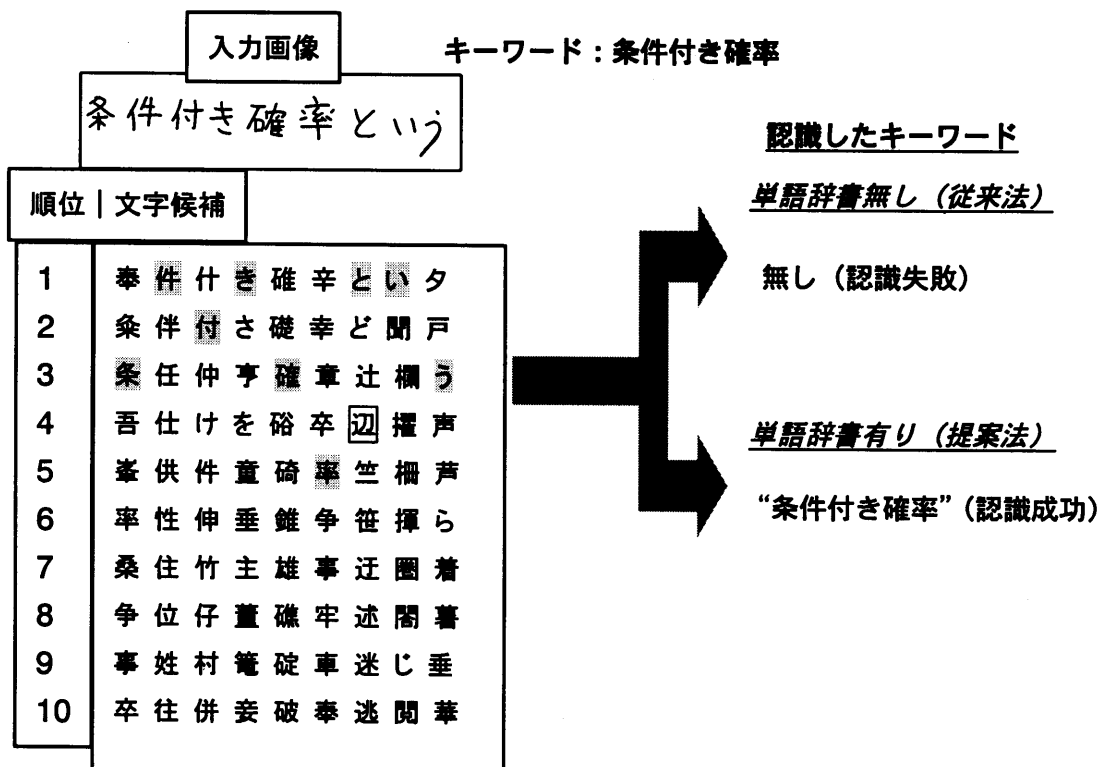


図56 成功例 2

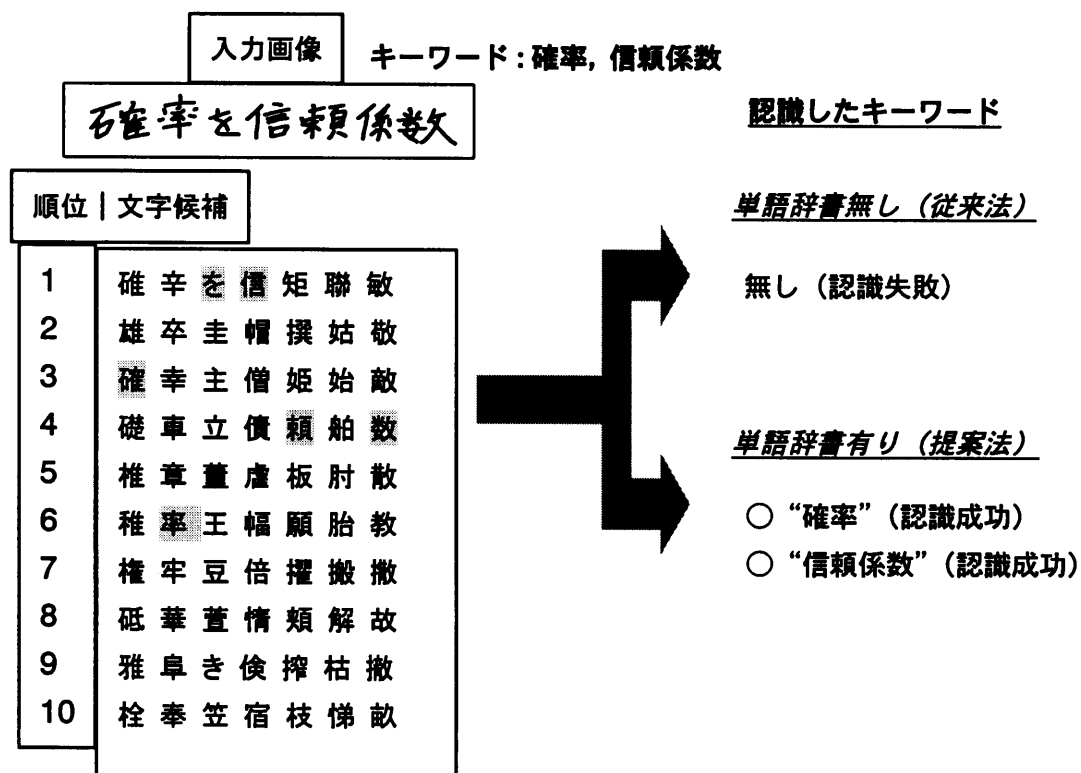


図57 成功例 3

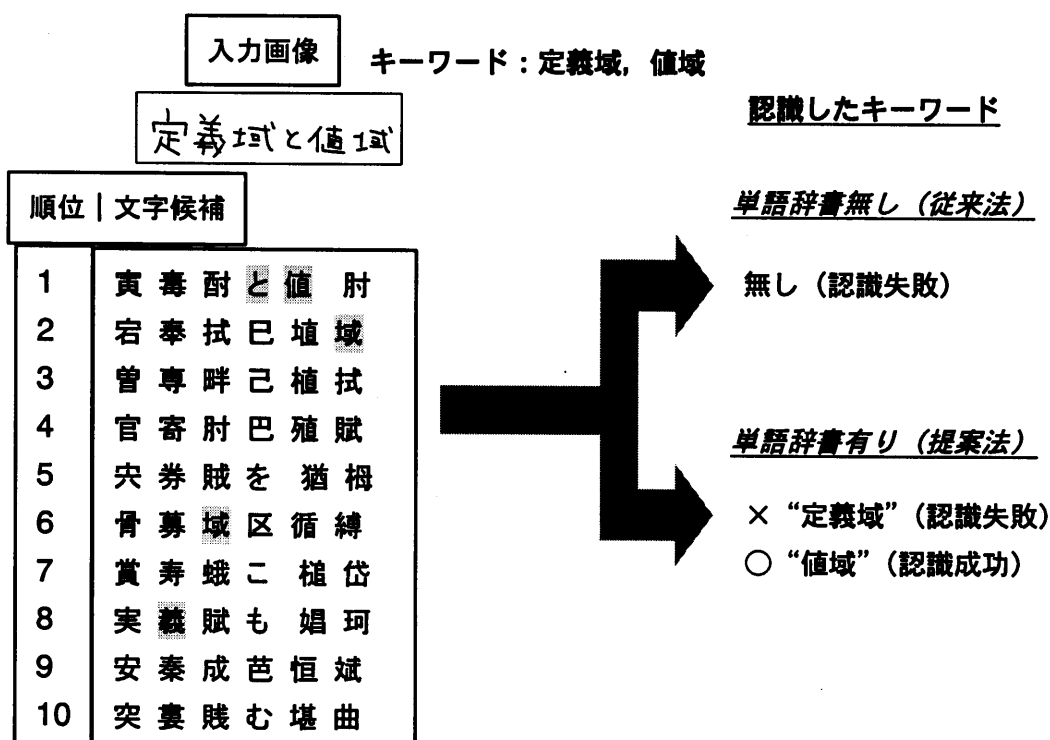


図58 失敗例 1

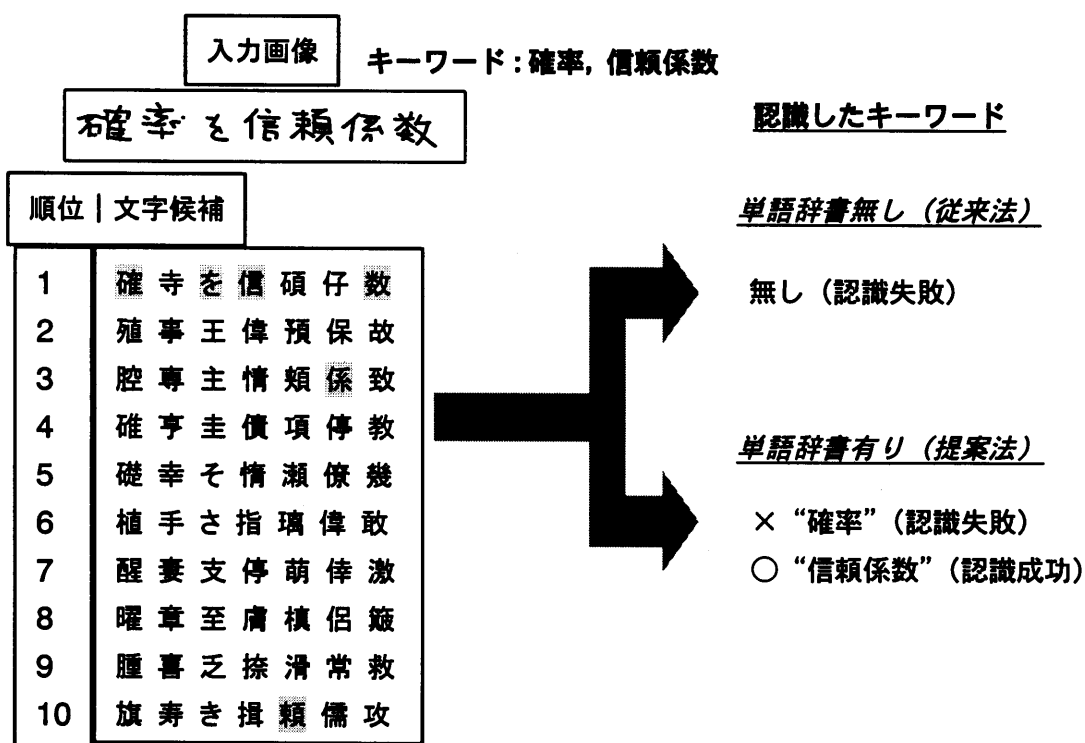


図59 失敗例 2

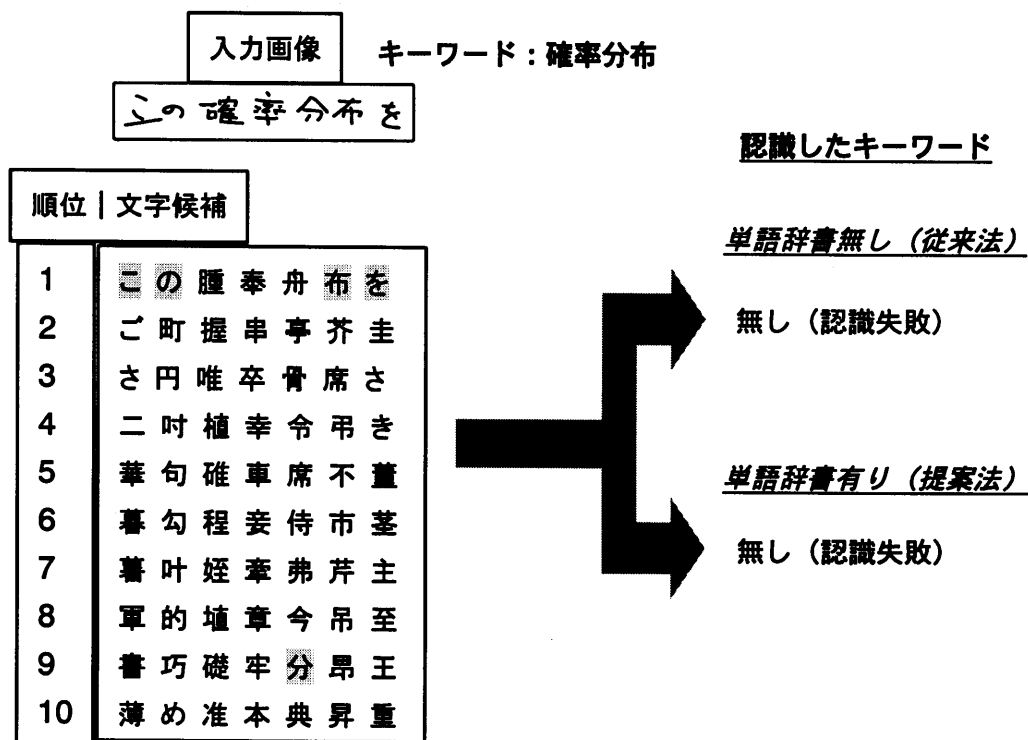


図60 失敗例 3

図 55 の成功例では、“条件付き確率”というキーワードが単語辞書を用いない手法では、認識できなかったが、単語辞書を用いた手法では認識する事に成功した。また、図 56 の成功例では、単語辞書を用いない手法・用いた手法ともに、“条件付き確率”というキーワードの認識に成功している。ここで、個別文字“と”の文字候補に注目すると、第 4 位の文字候補に“辺”が存在する。本実験で用いた単語辞書中には“辺”というキーワードが存在するため、相違度を用いない手法では“辺”がキーワードとして誤認識されていたが、この文字候補の相違度を見ると表 6 のようになる。ここで、本手法の step. 2 の処理により、1 位の文字候補“と”と 2 位の文字候補“ど”との相違度差が 30 以上であるため、2 位以下の文字候補はリジェクトとして破棄される。これにより“辺”を誤認識する事がなくなった。

表6 “と” の文字認識結果

文字候補	相違度
と	45
ど	100
辻	133
辺	139
竺	140

また、図 57 の例では、キーワード“信頼係数”に対し、入力文字の“係”も文字候補の 10 位以内に正解が存在しなかったが、4 文字中 3 文字（75％）の文字が存在するので認識に成功している。

失敗例を見ると、図 58 では“定義域”の“定”が、図 59 では“確率”の“確”が、図 60 では“確率分布”の“確”及び“率”のそれぞれが、文字候補中に存在しなかったために、認識に失敗している。このように、キーワードを構成する文字数が“確率”や“定義域”のように 2～3 文字と少ないキーワードでは、一つ一つの文字認識の結果の重要度が大きくなり、1 文字でも正解が存在しないと認識する事が出来ない。

## 5.3 考察

### 5.3.1 実験結果の考察

実験結果より、各筆記者に対して個人差はあるものの全員に対して単語辞書を用いないキーワード認識法よりも、単語辞書を用いたキーワード認識法で高いキーワード認識率、（平均で 32.1% から 94.0%）となることが確認できた。この結果より、紙上での単語認識同様にホワイトボード上での文字列に対しても単語辞書を用いた単語認識は有効であると考えられる。

失敗例の結果から、このような失敗例の場合、文字認識精度が上昇する事で解決する事が望ましいが、その他に対策として、利用する文字候補を 10 位までと限定するのではなく、10 位以降の候補も利用したり、相違度の差が 30 までの全ての文字候補を利用する事や、本手法の step.3 の 75% や 60% の各制約条件を下げる事が考えられる。この場合、意図しないキーワードの誤認識が増加する事が容易に予測できるが、本研究では現在、非同期型の e ラーニングシステムの開発を対象としているため、誤認識部分は後に訂正を行えば良い。よって、該当部分の認識は行わず、認識結果が疑わしいと知らせるため、文字色を変えるなどの処理を行う事が有効ではないかと考える。これは実際に OCR でも用いられている。また、本研究では用いなかった 1 位の文字候補の相違度によって信頼度を設ける手法も有効と考える。また、本研究の実験結果より、同じ文字でもやはり筆記者によって、文字のストロークや大きさが異なり、小さい文字やストロークが途中で切れている文字の認識率が悪い傾向があった。このように、各筆記者の字形についての個性“癖”を吸収させる文献[14]のような手法の導入も有効と考える。

上記をまとめると以下ようになる。

- 文字認識部の改良
- 利用する文字候補数の検討
- キーワード認識法の制約条件の検討
- 相違度による信頼度をキーワード認識に反映させる
- 筆記者に適応した字形に関する辞書を加える（汎用辞書+個別辞書）

次に、本実験では(3)式より評価値を算出する際に、各定数  $c, p$  に関して、 $c=0.3$ ,  $p=14$  としたが、 $c$  の値は大きいほど文字数の少ない単語が優位に評価され、小さい程文字数の多い単語が優位に評価されるといった特徴があり、文字数の多い単語の方が優位に評価されるべきとの考えからこの値を用いた。 $p$  の方は、ほとんどの文字において 14 位までに正解文字が存在する。という考えからこの値を用いた。しかし、これらの値は場合により変動するものと思われる。そのため、自動的に設定可能とするのが良いのか、また実験的に算出するのが良いのか、もしくは、評価値を算出したキーワード全てを評価値順に表示させる方が良いのか、今後検討する必要があると考えられる。

### 5.3.2 他の単語辞書や認識技術を用いた研究との比較

提案した手法と比較するために、紙上で用いられている単語辞書を用いた誤り訂正手法をいくつか紹介する。一つは木村らの文の前後関係を用いた手法[23]であるがホワイトボード上では、本実験の結果で示したように文字認識精度が低いため、適用する事が難しい。次に、三部らの形態素解析を行った後で、国語辞書を用い言語情報を利用する手法[24]があるが、この場合、辞書に登録してある単語数が非常に多いためにその処理時間が問題となる。一方で、本研究の場合、大学の講義を対象としているために、キーワード数を限定する事ができ、そのため少ない処理時間で処理を行う事が可能である。次に、鈴木らのキー文字を用いた日本の住所認識[16]においても単語辞書として県名・市名等の辞書が用いられているが、本研究の場合、日本語の住所認識のように“県”や“市”のようなキー文字が常に存在するとは限らず、キー文字を決定する事は困難であるためこの手法は本研究では向かないと考える。また、他の応用分野として宮本らはオフィス環境においてロボットに本棚の蔵書をカメラにより撮影し文字認識を行い、人間によって検索が行えるシステムの研究[25]を行っているが、認識率が低く、まだ実用段階に至っていない。

このように、紙上においては文字認識技術を利用したシステムが普及しているが、本研究のように板書を対象としたものや、カメラ画像を利用したものというのはまだまだこれからの研究領域である。

## 第6章 おわりに

本論文では、e ラーニングによる個別学習システムのための処理部分として、文字認識手法として“加重方向指数ヒストグラム法”を、キーワード認識手法として、“単語辞書を用いたキーワード認識法”を提案し検討した。また、実験により複数の筆記者に対してホワイトボード上での筆記文字でも、従来の紙上同様高いキーワード認識精度が得られる事を確認し、e ラーニングシステムに使用できる可能性を明確化した。

今後の課題として、以下の事項が挙げられる。

### (1) キーワード認識法及び文字認識部の改良

提案手法や文字認識のプログラムを改良する事で、認識率が上昇する余地はまだ残っていると考えられる。

### (2) より良い個別文字切り出し法の導入

本研究では、個別文字切り出しが完全に成功したものとし、射影ヒストグラムを求める手法により簡易的に個別文字の切り出しを行い、手動で、個別文字が切り出されている画像を選択したが、個別文字同士が接触するなど複雑な構造になるにつれ、さらに精度の高い切り出し法[26][27][28]の導入が必要と考える

### (3) 黒板に筆記された板書画像への対応

ホワイトボードを用いた講義に比べ、黒板を用いた講義がまだ多い事からも、黒板の板書への対応も重要な課題であると考えられる。

### (4) システムの完全自動化

文字切り出しや、行画像作成など、未だ一部手動による部分があるため、これらを自動化し、統一した使いやすいユーザーインターフェースの作成が必要である。

## 参考文献

- [1] 経済産業省商務情報政策局情報処理振興課, “e ラーニング白書 2005/2006 年版”, オーム社, 2006-7
- [2] 高度情報通信ネットワーク社会推進戦略本部 (IT 戦略本部)   
 <http://www.kantei.go.jp/jp/singi/it2/>, 内閣府, 2004
- [3] 先進学習基盤協議会 (ALIC) 編著, “e ラーニングが創る近未来教育 -最新 e ラーニング実践事例集-”, オーム社, 2002-9
- [4] 岡本敏雄, 小松秀圀, 香山瑞恵, “e ラーニングの理論と実際”, 丸善株式会社, 2004-11
- [5] 和田公人, “失敗から学ぶ e ラーニング”, オーム社, 2004-5
- [6] 矢田裕紀, 鶴岡信治, 吉川大弘, 篠木剛, “遠隔授業映像撮影のためのカメラ映像と板書画像を併用したカメラ視野の決定法”, 電子情報通信学会技術報告, PRMU2003 - 213, pp. 89 - 94, 2004
- [7] Kazuyuki Nishikimi, Yuki Yada, Shinji Tsuruoka, Tomohiro Yoshikawa and Tsuyoshi Shinogi, “Automatic Display of an Additional Explanation on a Keyword Written by a Lecture for e-Learning Using a pen Capture Tool on Whiteboard and Two Cameras”, Proc. of The 4<sup>th</sup> International Symposium on advanced Intelligent Systems (ISIS2003), pp.102-105, 2003
- [8] コクヨ mimio ホームページ <http://www.kokuyomimio.com/>
- [9] Canon 社ホームページ <http://canon.jp/>
- [10] 安井良, 鶴岡信治, 篠木剛, 川中普晴, “e ラーニングのためのキーワードに関連する教科書の説明部分の抽出”, 平成 18 年度三重地区計測制御研究講演会 pp.(P22-1)-(P22-4), 2006
- [11] Keiichi Shirasawa, Hiroharu Kawanaka, Shinji Tsuruoka, Tomohiro Yoshikawa and Tsuyoshi Shinogi, “Automatic Determination of an Active Camera View in an Image Based e-Learning System,” Proceedings of The 6<sup>th</sup> International Symposium on advanced Intelligent Systems (ISIS2005), pp.213-217, 2005-9
- [12] 栗田昌徳, 鶴岡信治, 三宅康二, 横井茂樹, “加重方向指数ヒストグラムと疑似マハラノビス距離を用いた手書き漢字・ひらがな認識”, 電子情報通信学会, 信学技報 PRL-82-79, pp105-112, 1983-01
- [13] 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”, 電子情報通信学会論文誌 D, vol. J70-D, No. 7, pp. 1390-1397, 1987-7
- [14] 鶴岡信治, 森田裕之, 木村文隆, 三宅康二, “筆記者に対して適応機能をもった自由手書き文字認識”, 電子情報通信学会論文誌 D, vol. J70-D, No.10, pp.1953-1960, 1987-10
- [15] 丸川勝美, 古賀昌史, 島好博, 藤澤浩道, “手書き漢字住所認識のためのエラー修正ア

- ルゴリズム”, 情報処理学会論文誌, vol.35, No.6, pp.1101-1110, 1994-6
- [16] 鈴木章, 宮原末治, “文字位置のずれを許容する枠なし筆記住所認識”, 電子情報通信学会論文誌 D-II Vol. J77-D-II No.1, pp.20-28, 1994-1
- [17] 嶺竜治, 古賀昌史, 酒匂裕, “N-gram 言語統計量を併用した島駆動型文字列認識方式”, 電子情報通信学会論文誌 D Vol. J89-D No.5, pp.1011-1018, 2006-5
- [18] 伊藤伸泰, 丸山宏, “OCR 入力された日本語文の誤り検出と自動訂正”, 情報処理学会論文誌, vol.33, No.5, pp.664-670, 1992-5
- [19] 吉田大祐, 鶴岡信治, 篠木剛, “e ラーニングのための単語辞書を用いたホワイトボード手書きキーワードの高精度認識”, 平成 17 年度電気関係学会東海支部連合大会論文集, 0-242, 2005-9
- [20] Daisuke Yoshida, Shinji Tsuruoka, Hiroharu Kawanaka and Tsuyoshi Shinogi, “Handwritten Keywords Recognition on Whiteboard Using Word Dictionary for e-Learning”, Proc. of The 6<sup>th</sup> International Symposium on advanced Intelligent Systems (ISIS 2005), pp.197-202, 2005-9
- [21] 鳥脇純一郎, 石井直宏, 鶴岡信治, “インターユニバーシティ情報数学”, オーム社, 2001
- [22] 酒井幸市, “デジタル画像処理の基礎と応用-基本概念から顔画像認識まで-”, CQ 出版社, 2003
- [23] Fumitaka Kimura, Shinji Tsuruoka, Yasuji Miyake and Malayappan Shridher, “A Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words”, IEICE TRANS.INF. & SYST. Vol.E77-D, No.7, pp.785-793, 1994-7
- [24] 三部裕史, 大森健児, “信頼性の低い文字認識結果に対する言語情報を用いた誤認識文字の訂正”, 情報処理学会論文誌, Vol.34, No.10, pp.2117-2124, 1993-10
- [25] 宮本圭, 上野敦志, 武田英明, “オフィス環境における文字情報の検出と利用に関する研究”, 人工知能学会知識ベースシステム研究会, SIG-KBS-9904, vol.47, 2000-3
- [26] Shinji Tsuruoka, Nobuyuki Watanabe, Nariyasu Minamide, Fumitaka Kimura, Yasuji Miyake, Malayappan Shridhar, “Base Line Correction for Handwritten Word Recognition”, Proc. of Third Int. Conf. On Document Analysis and Recognition (ICDAR'95), pp.902-905, 1995-8
- [27] Shinji Tsuruoka, Yusuke Adachi and Tomohiro Yoshikawa, “The Segmentation of a Text Line For a Handwritten Unconstrained Document Using Thinning Algorithm”, Proc. of 7<sup>th</sup> International Workshop on Frontiers in Handwriting Recognition, pp.505-510, 2000-9
- [28] T.Yamaguchi, T.Yoshikawa, T.Shinogi, S.Tsuruoka, M.Teramoto, “A Segmentation Method for Touching Japanese Handwritten Characters Based on Connection Condition of Lines”, Proc. of Int. Conf. On Document Analysis and Recognition (ICDAR'01), pp.937-841, 2001

## 謝辞

本研究の遂行及び修士論文の作成にあたり、日ごろ丁寧なご指導とご助言を頂きました本学工学部電気電子工学科の鶴岡信治教授、篠木剛助教授、川中普晴助手に深く感謝致します。また、貴重な時間を割いていただき、本論文を査読していただいた本学工学部電気電子工学科の北英彦助教授に深く感謝致します。また本学工学部電気電子工学科の研究生の Glenn Rayat 氏に感謝致します。

そして、苦楽を共にした情報処理研究室博士前期課程 2 年の伊藤聖太郎、大谷芳弘、岡山陽介、京谷忠雄、東海林正和、安井良、また共に画像グループで学んだ、情報処理研究室博士前期過程 1 年の伊藤哲也、大國聖治、柴田彰洋、Premachandra Halpage Chinthaka、卒業研究生の竹田洵哉、長縄侑樹に感謝致します。

最後となりましたが、日ごろお世話になった情報処理研究室の皆様に感謝致します。

## 発表論文リスト

### 国内会議

- (1) 吉田大祐, 鶴岡信治, 篠木剛, “e ラーニングのための単語辞書を用いたホワイトボード手書きキーワードの高精度認識”, 平成 17 年度電気関係学会東海支部連合大会論文集, 0-242, 2005-9
- (2) 吉田大祐, 川中普晴, 篠木剛, 鶴岡信治, “e ラーニングのための単語辞書を用いたホワイトボード上の手書き文字列中のキーワード認識”, 平成 17 年度 SICE 三重地区計測制御研究講演会, pp. (B8-1)-(B8-6), 2005-12
- (3) 吉田大祐, 鶴岡信治, 篠木剛, 川中普晴, “e ラーニングのための単語辞書を用いたキーワード認識～複数人の筆記文字の評価～”, 平成 18 年度電気関係学会東海支部連合大会論文集, 0-502, 2006-9

### 国際会議

- (1) Daisuke Yoshida, Shinji Tsuruoka, Hiroharu Kawanaka and Tsuyoshi Shinogi, “Handwritten Keywords Recognition on Whiteboard Using Word Dictionary for e-Learning”, Proc. of The 6<sup>th</sup> International Symposium on Advanced Intelligent Systems (ISIS 2005), pp.197-202, 2005-9
- (2) Daisuke Yoshida, Shinji Tsuruoka, Hiroharu Kawanaka and Tsuyoshi Shinogi, “Keywords Recognition of Handwritten Character String on Whiteboard Using Word Dictionary for e-Learning”, Proc. of 2006 International Conference on Hybrid Information Technology (ICHIT2006), pp.140-145, 2006-11

### 受賞等

- (1) Outstanding Paper Award  
2005 International Symposium on Advanced Intelligence Systems (Korea)  
Title: "Handwritten Keywords Recognition on Whiteboard Using Word Dictionary for e-Learning"  
Authors: D.Yoshida, S.Tsuruoka, H.Kawanaka, T.Shinogi