

複写可

修士論文

授業データベース作成のための
遺伝的アルゴリズムを用いた部分板書画像の自動結合手法



平成 20 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

長縄 侑樹

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
第2章	個別学習用eラーニングシステム	8
2.1	本研究室のeラーニングシステム	8
2.2	ネットワークカメラ	12
2.3	従来システムの問題点	14
第3章	テンプレートマッチング法による部分板書画像結合手法	16
3.1	板書撮影視野の決定	16
3.2	テンプレートマッチング法	19
3.3	相違度算出法	23
3.3.1	相違度算出法概要	23
3.3.2	HSV表色系を用いたテンプレートマッチング	25
3.3.3	色チョークに対する汎用性の評価実験	29
3.4	環境変化とテンプレートの質	32
3.4.1	適応的なテンプレート生成の必要性	32
3.4.2	テンプレート評価実験	34
第4章	遺伝的アルゴリズムを用いたテンプレート自動生成法	37
4.1	テンプレート評価方法	37
4.1.1	評価関数	37
4.1.2	評価実験	39
4.2	遺伝的アルゴリズムを用いたテンプレート生成法	42
4.2.1	遺伝的アルゴリズムの適用	42
4.2.2	染色体のコーディング法	43
4.2.3	遺伝的演算	45
第5章	提案手法を用いた部分板書画像結合実験	46
5.1	実験環境	46
5.2	実験結果と考察	47
5.2.1	提案手法の実験結果と考察	47
5.2.2	比較手法の実験結果と考察	53
第6章	おわりに	55
	参考文献	56
	謝辞	58
	発表論文リスト	59

図一覧

- 図1 ブレンディッド・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 個別学習システムとデータベースの関係
- 図4 eラーニングシステム概要
- 図5 講義風景(固定視野カメラ)
- 図6 講師映像の撮影範囲(可動視野カメラ)
- 図7 板書画像(可動視野カメラ)
- 図8 キーワードに関連する補助説明表示システム
- 図9 可動視野カメラ(VB-C50i)
- 図10 固定視野カメラ(VB-C50Fi)
- 図11 教室の後部の壁に設置した IP カメラ(10 番教室)
- 図12 実際の教室において撮影された板書画像
- 図13 図 12 の板書画像より切り出された文字画像
- 図14 文字認識に適切な解像度の個別文字画像
- 図15 板書撮影視野(2 分割撮影)
- 図16 図 15 より切り出された文字画像
- 図17 板書撮影視野(8 分割撮影)
- 図18 図 17 より切り出された文字画像
- 図19 部分版書画像例
- 図20 結合板書画像例
- 図21 部分板書画像例
- 図22 テンプレートマッチングの原理
 - (a) フローチャート
 - (b) 探索範囲におけるテンプレートの走査
- 図23 提案テンプレートマッチング法概要
 - (a) テンプレートと探索範囲の候補領域(横結合)
 - (b) テンプレートと探索範囲の候補領域(縦結合)
 - (c) テンプレートマッチング処理
- 図24 相違度算出法概要
- 図25 結合画像例(RGB 表色系)
- 図26 HSV 色空間モデル
- 図27 色相上における各色チョークと黒板の領域
- 図28 白色の板書画像における輝度分布イメージ
- 図29 各色のチョークにより書かれた板書画像のグレースケール画像
- 図30 時間変化による輝度変化
 - (a) 左上黒板画像(8:50)
 - (b) 左上黒板画像(16:10)

- 図31 場所変化による輝度変化
- (a) 左上黒板左上部画像
 - (b) 左上黒板画像右下部画像
- 図32 作成テンプレート例
- (a) テンプレート候補領域
 - (b) 1文字を含むテンプレート
 - (c) 黒板領域のみのテンプレート
 - (d) 文字列を含むテンプレート
- 図33 マッチング失敗テンプレート例
- 図34 探索領域中の類似領域例
- 図35 マッチング成功テンプレート例
- 図36 (4.3)式における交点領域イメージ
- 図37 候補領域内の座標指定によるテンプレート作成イメージ
- 図38 評価実験における探索範囲決定イメージ
- 図39 評価値とマッチング精度の関係
- 図40 染色体のコーディング表
- (a) マッチング1回分の遺伝子イメージ
 - (b) マッチング1回分の染色体コーディングイメージ
 - (c) 7回分のマッチングを統合した場合の染色体コーディングイメージ
- 図41 評価値の推移
- 図42 テンプレート生成成功例
- (a) 相対的に複雑な文字を含むテンプレート
 - (b) 複雑な文字の一部分を含むテンプレート
 - (c) 曲線により構成される文字を含むテンプレート
- 図43 テンプレート生成失敗例
- (a) 交点部分のみを抽出したテンプレート
 - (b) 極めて小さいテンプレート
 - (c) 角の部分を抽出したテンプレート
- 図44 板書画像結合成功例
- (a) 全体画像
 - (b) 拡大画像
- 図45 板書画像結合成功例
- (a) 全体画像
 - (b) 拡大画像
- 図46 存在しない領域を含むテンプレート
- (a) テンプレート候補領域
 - (b) 探索候補領域
- 図47 比較手法における結合板書画像例

表一覧

表1 各手法における実験結果

表2 環境の変化による輝度変化のマッチング精度に対する影響

表3 テンプレート評価実験結果

表4 遺伝的アルゴリズムに使用したパラメータ

表5 t分布表

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、ネットワーク通信のブロードバンド化が急速に進むとともに、ネットワーク通信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性能の向上、および価格の下落が進んできている。これにより、実現可能となった応用技術のeラーニングが挙げられる。eラーニングとは、一般に「コンピュータやネットワークを利用した教育」と定義され、その有用性は実際の教育現場においても認められている。eラーニングには様々な形態があるが、その代表的な形態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), 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)に大別される。また、携帯電話に代表される携帯端末の技術革新と普及に伴い、モバイル・ラーニング(Mobile Learning)と呼ばれる形態も普及しはじめている。これらの学習ツールの技術革新および多様化により、市販製品やサービスにおいては同期型・非同期型の統合が進みつつある。

実際の教育現場では、従来の教室に一同を集めて教育を行う集合教育と、eラーニングを組み合わせる「ブレンディッド・ラーニング(Blended Learning)」が一般的となっている。図1に「ブレンディッド・ラーニング」の概念図を示す。「ブレンディッド・ラーニング」は学習者にとって最適な学習形態の提供を目標としており、現在のeラーニングのトレンドの一つとして、「ブレンディッド・ラーニング」の効果的な運用を実現する「学習者の視点に立ったシステム」の提供が挙げられる[1]-[3]。

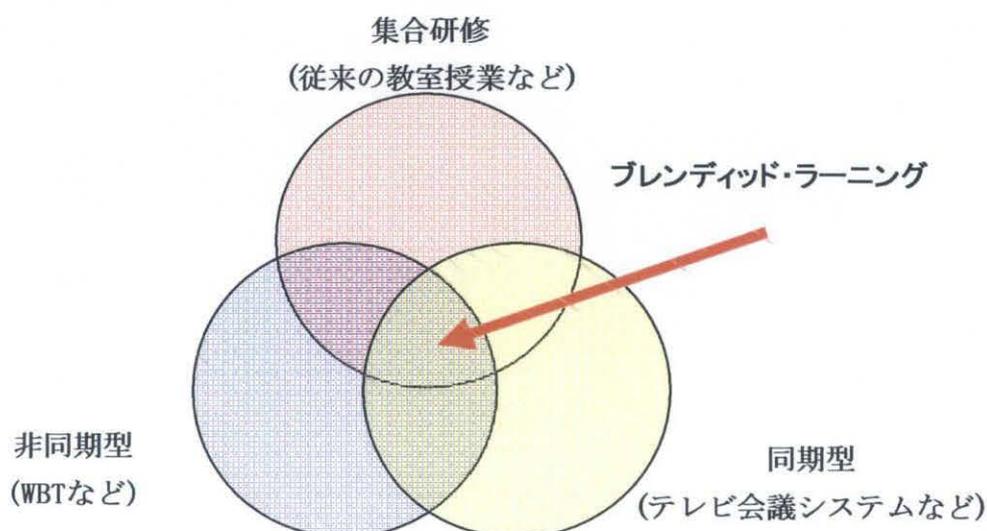


図1 ブレンディッド・ラーニングの概要

1.1.2 eラーニングの現状

前節にて述べたように、ネットワーク通信の高速化や情報処理端末の低価格化に伴い、高速インターネット環境の普及が進んでいる。これにより、教育現場においてもインターネットの整備が進められ、eラーニングを展開できる環境が整備されつつある。また、政府が大学におけるeラーニングの取り組みを後押しするなど、技術的な背景以外においても、eラーニング導入のための環境が整いつつある。

これまでに政府主導により、教育現場におけるeラーニングの導入を推進しようとする政策が提起されてきた。2001年3月には文部科学省の大学設置基準が改正され、従来のテレビ会議式の遠隔授業に加えてインターネット等活用授業も遠隔授業として位置づけられることとなり、単位取得が認められることとなった[4]。また、内閣府が2006年1月に公表した「IT新改革戦略」では、

「インターネットを用いた遠隔教育等を活用した特色ある取組を支援し、各大学間の競争的環境を醸成するなど、2010年度までにインターネット等を用いた遠隔教育を行う学部・研究科の割合を2倍以上にすることを旨とする」

としている[5]。

このような教育政策が進められる中、2004年4月1日より89の国立大学法人、及び大学利用共同機関法人が誕生した。これにより、大学の自主性・自律性が拡大し、各大学の個性を生かした運営が可能となった。特に本学においては、eラーニングにより新たな教育に対する取り組みという形でそれが表れている。本学では2005年度から英語教育にTOEIC英語自習システムを導入するとともに、共通教育の英語履修時に、本システムを用いた学習を義務づけている。また、2006年度から全教室において無線LANの使用が可能となり、教室にパソコン等の情報端末を持ち込めば、全学生がインターネットを利用することが可能となった。さらに、eラーニングプラットフォームである学習管理システム Moodle が導入され、従来の集合授業による学習に加えて、オンライン学習(資料の閲覧、レポート、フォーラムなど)を取り入れた新たな教育内容を組む授業科目が増えはじめている。

1.1.3 高等教育におけるeラーニングの役割

近年、大学は少子化や独立行政法人化などをはじめとした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加している。メディア教育開発センターの調査によると、4年制大学の約4割が何らかの形で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進んでいる。

以下に、高等教育機関で行われているeラーニング導入事例の特徴をまとめたものを示す。

- 授業の補助ツールとして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た後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利用して、国内のみならず海外に目を向けてシステム開発を行う例もある。
- eラーニングを学習プロセスに活用するだけでなく、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。

これらの特徴を見ると、eラーニングが高等教育において非常に幅広く活用されていることがわかる。これは、高等教育におけるeラーニングの役割に対し、複数の視点からニーズが存在するためである。これらの視点は二つに大別することができる。一つ目はIT活用による教育の質と提供手段の改善のための道具とする「教育の質と提供手段の視点」、二つ目は大学のIT化の推進や、社会人学生や留学生の獲得などの道具とする「大学経営の視点」である(図2)。これら二つの視点に含まれるニーズが大学の法人化、少子化といった大学を取り巻く環境の変化から発生したeラーニングに対するニーズである。

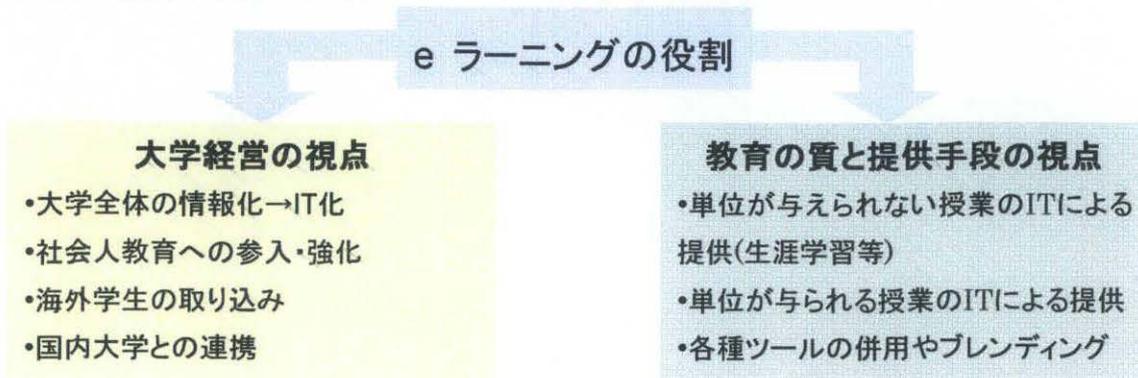


図 2 高等教育において期待されるeラーニングの役割

以上、大学の法人化、少子化といった大学を取り巻く環境変化から発生した eラーニングに対するニーズについて述べた。しかしながら、高等教育において eラーニングが求められている事例はこれだけではなく、大学の教育現場からも eラーニングに対するニーズが高まってきている。

一方、大学では、学生の入学時までのカリキュラムの相違や「ゆとり教育」による学生の既知内容の格差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ている。この結果、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が年々増加していることが、問題となっている[6]。このような問題を改善・解決するため、eラーニングに対する期待が大学の教育現場においても高まってきているのが実情である。これは、パソコンさえあれば学生が必要な時にどこでも学習出来ることや、学習者のレベルに応じた学習が可能であるという eラーニングの特長を活かしたシステムを用いて個別学習環境を配備することにより、これらの問題を改善・解決することが期待できるためである。

しかしながら、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。

① コンテンツの作成に膨大な費用が必要とされる

現代の学生はビデオゲームやテレビ等の映像に慣れており、これらの学生を満足させるコンテンツを作成する場合、莫大な費用が必要となる。例えば、ニューヨーク大学では支持の得られるコンテンツを作成するために、1コースにつき 2000 万円以上の費用が投じられた。しかしながら、日本では一つのコンテンツにかかることの出来る費用は 200~300 万円程度であり、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、一つの教材が完璧に全受講生に適応している事はあり得ない。全受講生に適応したコンテンツを作成するためには、複数のコンテンツをリンクさせ、互いに補間させる必要がある。この事実がコンテンツ作成をより困難としている。

③ 必要なコンテンツが存在しない

受講生数が多いほどコンテンツ作成に多くの費用を投じることができる。その反面、受講生の少ないコースは eラーニング化されず、受講したくても受講することが出来ない学生が存在する。

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。これらの問題点を改善することが、高等教育における eラーニングの導入に繋がると考えられる。具体的な改善策としては、従来の集団教育を変えること

なく、eラーニングのメリットを享受できるシステムの開発がそれにあたる。すなわち、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

講義自体の情報、教科書等の既存の教材をデジタル化する事で安価に作成する事が可能ではないかと考えられる。

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られる。

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になると考えられる。

1.2 本研究の目的

前節で述べたように、eラーニングの普及には質の高いコンテンツの充実が必要不可欠である。しかし、質の高いコンテンツを作成するには多くの時間と費用がかかる。そこで本研究室では、安価かつ容易に質の高いコンテンツを作成する機能を備えたeラーニングシステムを目標とし、大学における通常講義を対象とした非同期型の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[7-12]。

本システムの特徴は、3台のカメラを用いて講義の撮影を行い、得られた講義情報を利用することにより安価かつ効果的なコンテンツを生成・提供することが可能な点である。本システムにより提供されるコンテンツは以下の四つである。

1. 講師の動きに追従した臨場感のある講義映像
2. 黒板上の板書情報を提供するための板書画像
3. 2をテキスト化することにより得られた板書情報内に存在するキーワードとリンクした電子教科書中の本文テキスト
4. 3のテキスト中のキーワードの補足説明

これらのコンテンツは全て講義情報を利用しており、非常に安価かつ自動的に生成することが出来る。その結果、大学講義用eラーニングコンテンツの蓄積が可能となり、学生は必要に応じ利用することにより、1.1.3節で述べたような問題を改善・解決する手法の一つとなる事が期待される。また、将来構想として、キーワードを用いたリンクの対象を電子教科書だけでなく、複数のデータベースに拡張し、キーワードを軸としたリンクを構築することで互いを補間し更なる学習効率を学習者に提供することの出来る個別学習アシストシステムを考えている(図3)。

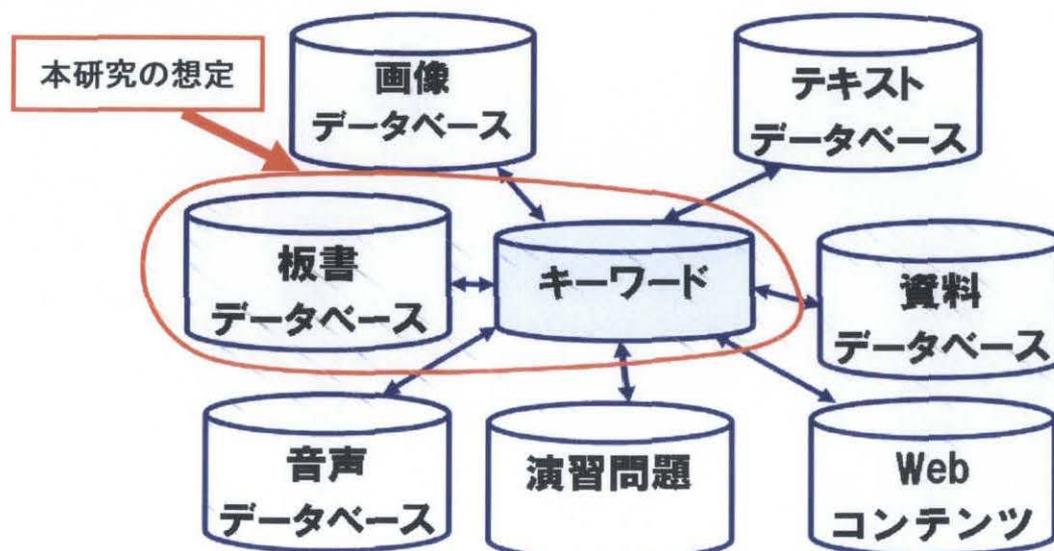


図3 個別学習システムとデータベースとの関係

本システムにおいて前頁3, 4のコンテンツを生成する際には, 板書画像に対し文字認識を適用し, 認識結果から講義中の重要単語であるキーワードを抽出することが必要不可欠である. しかしながら, 現在のシステムでは黒板を撮影対象とした場合, 板書画像の撮影に関してはほとんど検討されていない. そこで本研究では, キーワード抽出及び文字認識に適切な板書画像の取得を研究目的とする.

本論文の構成は, まず2章において本研究室でこれまで開発を行ってきたeラーニングシステムについて述べる. 3, 4章では板書画像の撮影方法及び生じた問題を解決するための提案手法について述べる. そして5章では提案手法の評価実験, 実験結果の考察を行う. 最後, 6章にてまとめとする.

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

前章で述べたように、本研究室ではeラーニングの利点に着目し、個別学習を行うことが出来る非同期型のeラーニングシステムの開発・提案を行ってきた(図4)[7-12].

本研究室で開発しているeラーニングシステムは可動視野カメラ2台, 固定視野カメラ1台の計3台のIPカメラにより構成される. 本システムでは3台のカメラを本学工学部の10番教室, 20番教室の後壁に設置し講義を撮影し, インターネットを経由して研究室内部のコンピュータにおいて配信処理を行っている. 可動視野カメラはそれぞれ黒板と講師を, 固定視野カメラは黒板と講師を含む講義風景を撮影する(図5). 可動視野カメラは, 利用者に臨場感のある映像を提供するため, 固定視野カメラを用いて撮影した講義風景映像から講師の位置を抽出し, 常に講師が中心となるような視野で撮影を行う(図6). また, もう1台の可動視野カメラは板書部分を撮影する(図7). 本システムはこれにより得られた板書画像中の文字を認識することにより, 板書中に含まれているキーワードを抽出する. この抽出されたキーワードと電子教科書をリンクさせることにより, キーワードに関連する電子教科書の説明部分の検索が可能となると共に, キーワードに関する説明を表示することも可能となる(図8).

本システムにおいて撮影された講義映像はリアルタイム配信されるだけでなく, 復習用の講義コンテンツとしてもオンデマンド配信されるため, データベースに蓄積される.

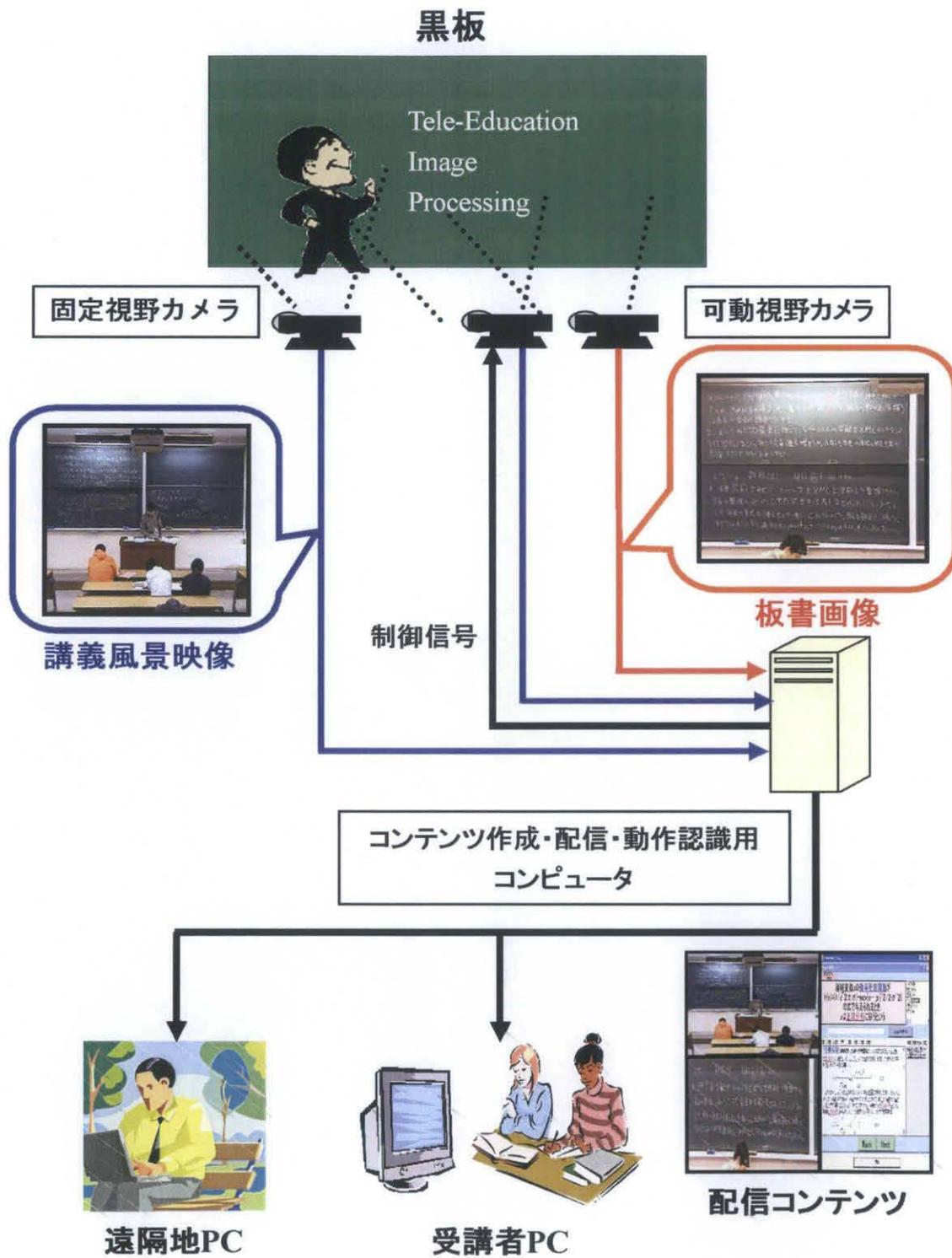


図4 eラーニングシステム概要

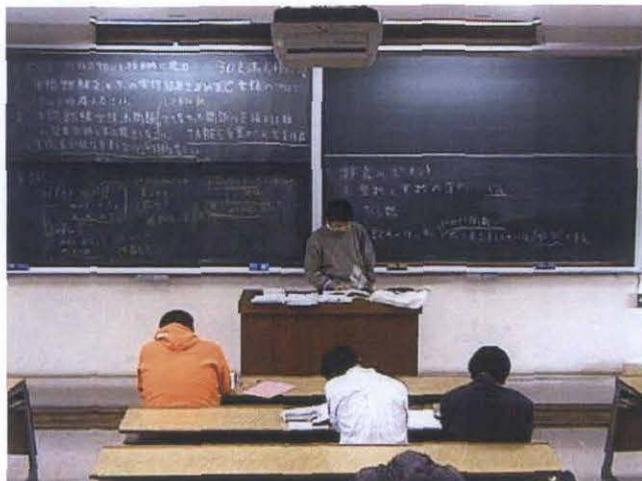


図 5 講義風景(固定視野カメラ)

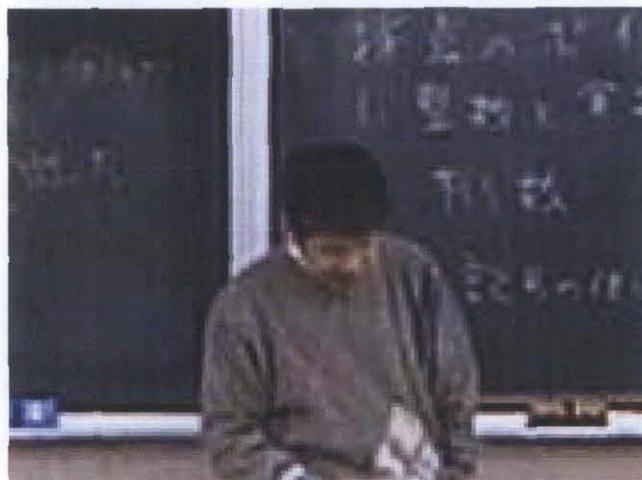


図 6 講師映像(可動視野カメラ)

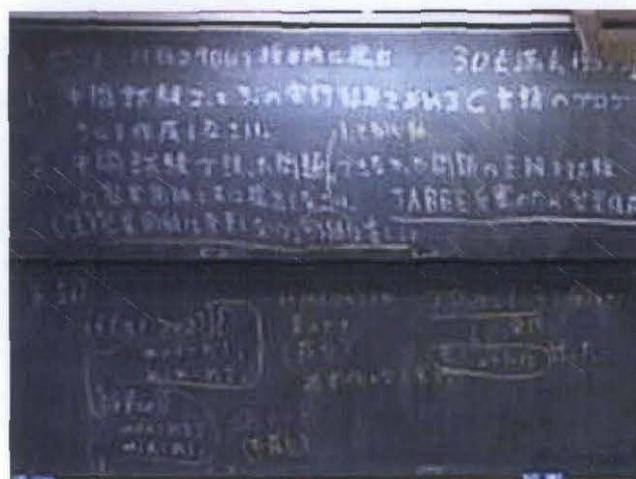


図 7 板書画像(可動視野カメラ)

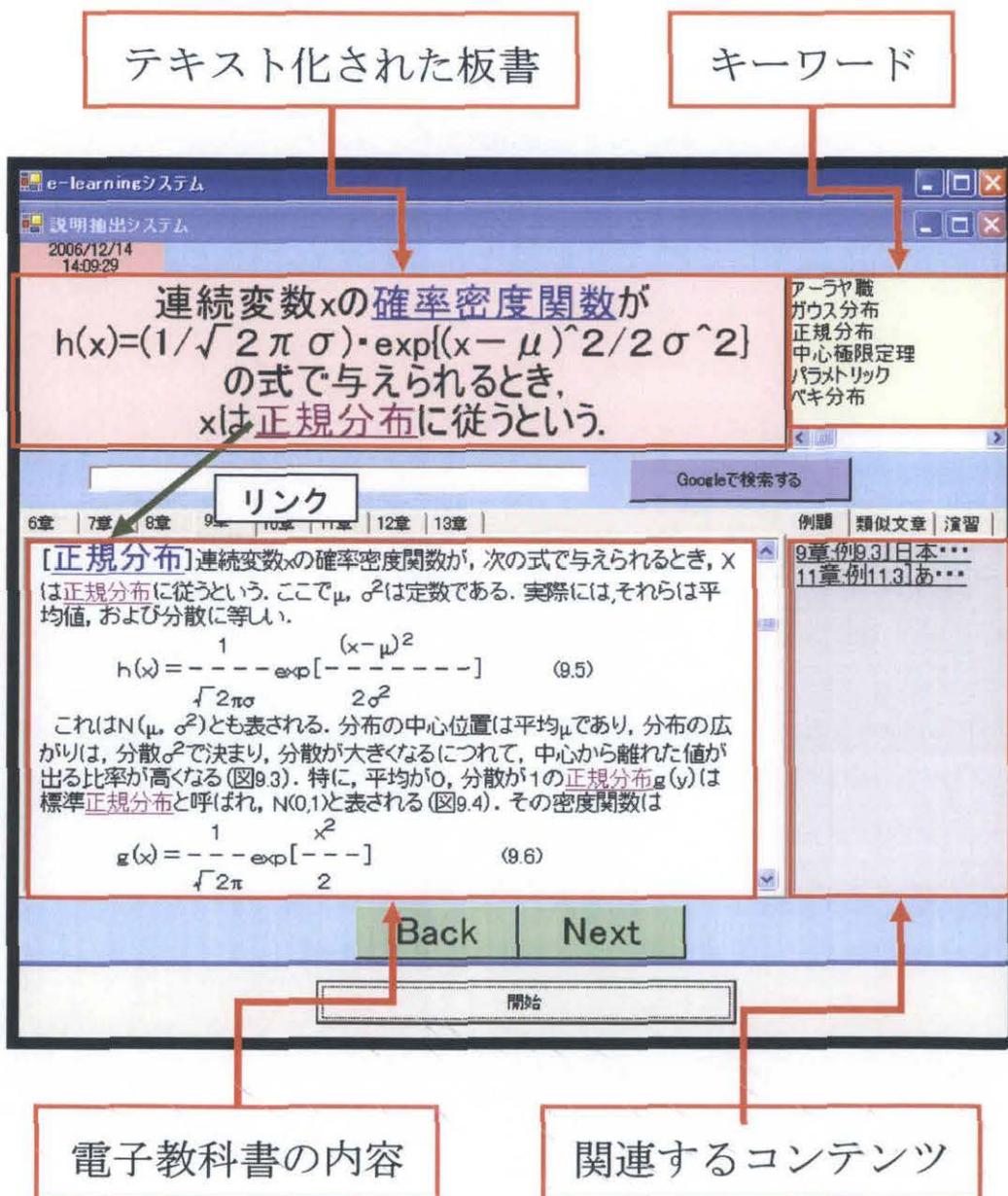


図 8 キーワードに関連する補助説明表示システム

2.2 ネットワークカメラ

本研究室のeラーニングシステムで使用しているネットワークカメラについて述べる。本システムでは臨場感ある講師映像と、複数存在する黒板の板書情報を同時に取得するため、可動視野カメラ2台を用いて、講師・板書画像を撮影する。この可動視野カメラにはパン・チルト・ズームがインターネット回線を経由して制御可能なネットワークカメラ(Canon社製 VB-C50i)(図9)を使用する。講義風景を撮影する固定視野カメラには、ズームのみ可能なネットワークカメラ(Canon社製 VB-C50Fi)(図10)を使用する。各カメラは、ネットワークに接続しているコンピュータから撮影・記録が可能であり、その画像サイズは640×480画素、320×240画素、160×120画素の3種が指定できると共に、動画の品質やフレームレートの指定も可能である。また、光学ズーム26倍(デジタルズーム16倍)が可能であるため、教室の後方にカメラを設置した場合においても、板書文字の内容を十分に読み取ることができる高品質画像を撮影することができる。カメラのそれぞれの配置を図11に示す。以下にVB-C50i, VB-C50Fiの性能を示す[13]。

可動視野カメラ : Canon社製 VB-C50i
固定視野カメラ : Canon社製 VB-C50Fi
撮影フレーム : 30fps
動画フォーマット : MS mpeg4
入力画像 : 640×480画素, RGB各8bitフルカラー
API : WebView Livescope SDK v2.0
カメラの設置場所 : 工学部10番教室(定員96名)の後部の壁



図 9 可動視野カメラ(VB-C50i)



図 10 固定視野カメラ(VB-C50Fi)



図 11 カメラの配置(10番教室)

2.3 従来システムの問題点

本研究室のeラーニングシステムには、板書画像から得たキーワードと電子教科書をリンクすることにより、学習者が学習している内容に合致したコンテンツを即座に提供できるという機能がある。また更に、このキーワードとのリンクを電子教科書以外のデータベースに拡張することにより、更に効率的な学習を行うことの出来る学習支援システムの構想も存在する。これらの実現には、板書画像に対する高精度な文字認識に基づくキーワード抽出が必要不可欠である。しかしながら、1.2節で述べたように従来システムでは、黒板全体を一度に撮影するのみで、高精度な文字認識を行うことを考慮した撮影は行われていない。実際にカメラを用いて黒板を撮影した場合、図12のような板書画像が得られる。図13は図12から文字領域のみを切り出した文字画像である。文字認識を高精度に行うためには、図14に示すような80×80画素程度の文字画像が有効であるとされているが、図13の文字画像の解像度は25×25画素程度であり、この解像度の画像を用いた文字認識では高い認識精度を期待することは出来ない[11]。加えて、現在の撮影方法により得られる黒板画像には、図12の上部の黒板に存在するものと同様の高輝度の領域が存在する場合がある。この状態の板書から文字情報を取得することは困難である。また、この領域は蛍光灯の光や日照が反射することにより生じるため、発生の防止も不可能である。

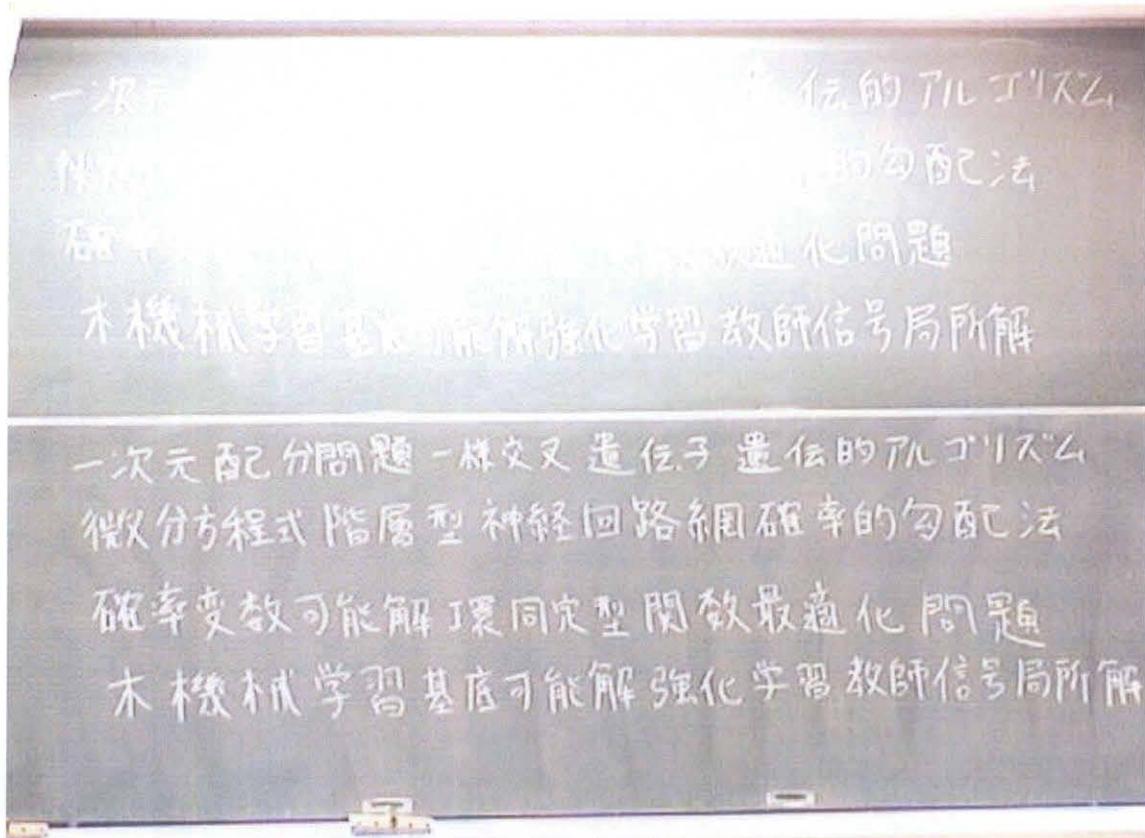


図 12 実際の教室において撮影された板書画像(640×480 画素)

関 数 教

図 13 図 12 の板書画像より切り出された文字画像(25×25 画素)

関 数 教

図 14 文字認識に適切な解像度の個別文字画像(80×80 画素)

第3章 テンプレートマッチング法による 部分板書画像結合手法

3.1 板書撮影視野の決定

前章で述べたように、今までの我々のシステムでは一度に黒板全体を撮影しており、解像度は 640×480 画素である。従って、黒板上の1文字の解像度は 25×25 画素程度になる。この解像度では、高精度に文字認識を行うことは難しい。文字認識に十分な解像度を持つ板書画像を得るためには、黒板全体を撮影するのではなく、カメラのズーム機能を用いて1文字が十分な解像度を持つまで拡大し、部分的に黒板を撮影した後に改めて1枚の板書画像を作成する必要がある。黒板を分割して撮影を行う場合、カメラから取得することが出来る画像の解像度(640×480 画素)の縦横比を考慮すると、図15のように黒板を二つの領域に分割して撮影を行う場合と、図17のように黒板を8領域に分割する場合の2通りが考えられる。いずれの場合も、前章で述べた光の反射を原因とする高輝度領域が存在している画像は得られず、この問題についてはどちらの撮影方法を選択しても良い。一方、文字画像の解像度については明確な差が生じた。黒板を2分割して撮影した場合、切り出した文字画像の解像度はおよそ 50×50 画素であり、8分割して撮影した場合の文字画像の解像度はおよそ 80×80 画素であった(図16, 図18)。2.3節で述べたように、効果的に文字認識を行なうことが出来る文字画像の解像度は 80×80 画素であることがわかっている。よって、本研究では黒板を八つの領域に分割して撮影を行うこととする。図19に実際に撮影した部分板書画像を示す。

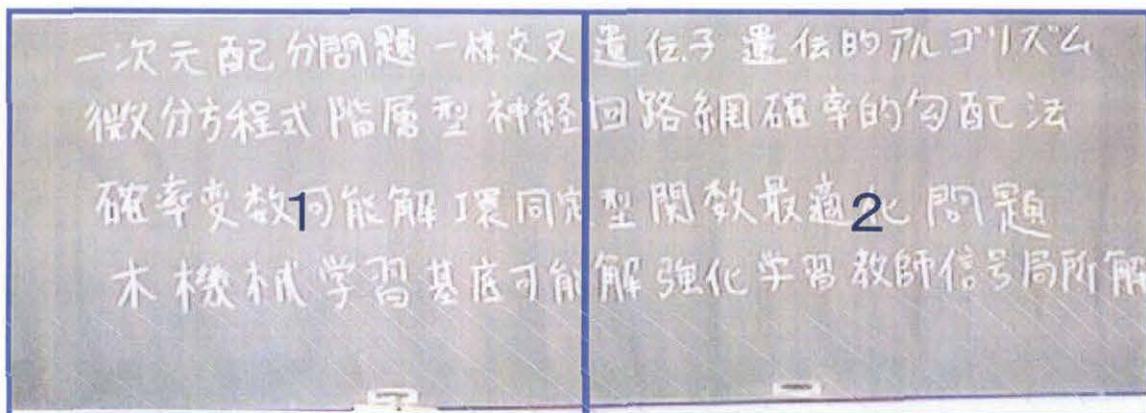


図 15 板書撮影視野(2分割)



図 16 図 15 より切り出された文字画像(50×50 画素程度)

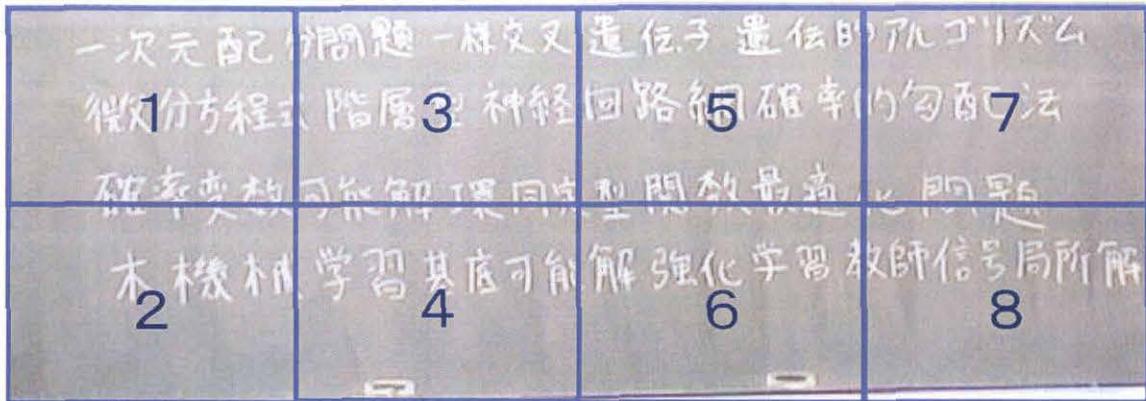


図 17 板書撮影視野(8分割撮影)



図 18 図 17 より切り出された文字画像(80×80 画素程度)

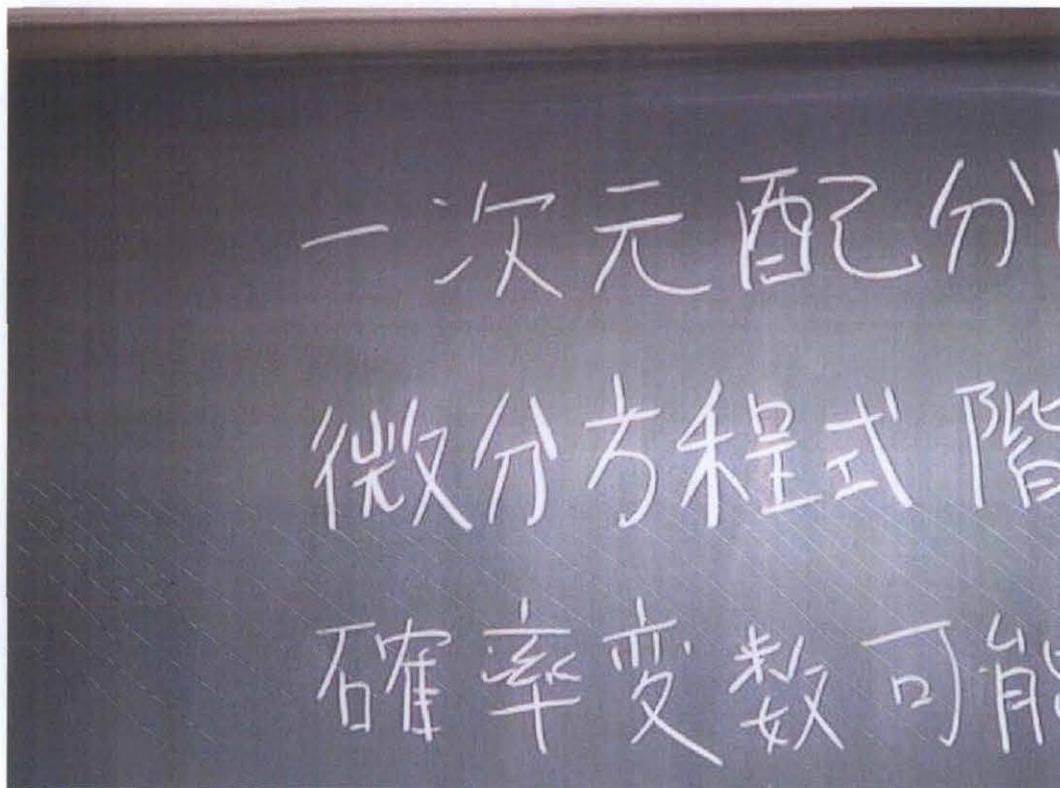


図 19 部分版書画像例(640×480 画素)

実際に板書を撮影し、得られた画像を並べた。結果、カメラのパン角度変更時の機械的な回転精度により、結合部分にずれが生じ、文字の重複や欠損が生じた(図20)。このような文字の損傷、重複は文字認識を困難にすると考えられる。代表的な文字認識手法の一つである加重方向指数ヒストグラム法を用いて実験を行った結果、撮影画像結合の際に生じた誤差により欠損、重複した文字は認識することができなかった[11]。この誤差は、可動視野カメラのハードウェア構造に由来するものであり、必然的に発生するため、これらの誤差の発生を完全に無くすことは困難である。そこで、これらのずれを解消するための結合手法について次節で検討する。

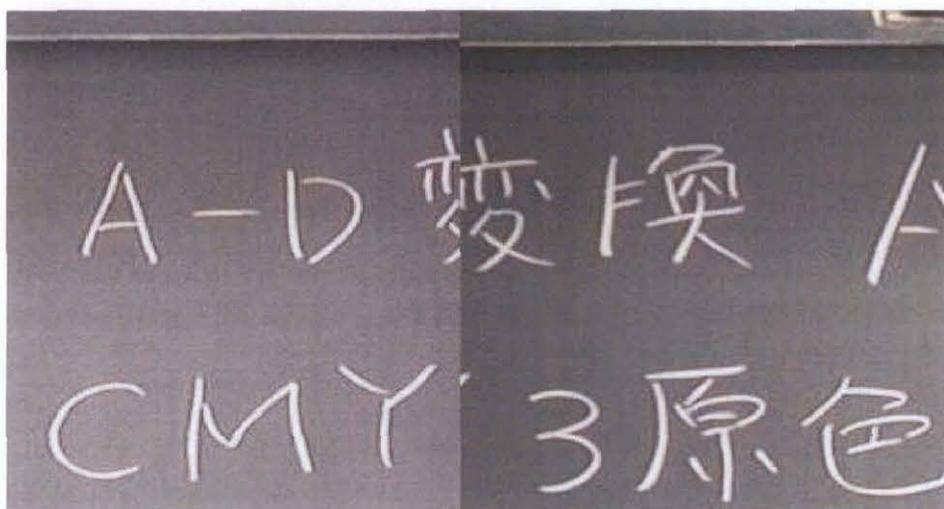


図 20 ずれの補正をしないで結合した板書画像の例

3.2 テンプレートマッチング法

カラー画像結合手法として様々なイメージモザイク手法が存在する。従来のイメージモザイク手法では、パノラマ画像、航空写真などのように風景、物体を撮影した画像が多い。従って、画像間に共通する特徴的な画素(物体の角、希少な濃度値の組み合わせを持つ画素)を対応点として結合の際に利用することが出来る[14][15]。しかしながら、本問題において対象とする画像(図21)は、物体を撮影した画像ではなく、黒板と板書の2種類の色しか存在しない上に、カメラのオートゲインコントロール及びレンズ歪曲によって画像の明るさに差が生じる[16]。このため、画像間に濃度差が生じ、通常イメージモザイク手法で用いられているように、特徴的な画素を用いて結合を行うことは難しいと考えられる。しかしながら、部分板書画像上には特徴的な形状をした結合画素を含む領域、すなわち文字領域が必ず存在する。そこで本研究では、特徴的な画素ではなく、特徴的な領域を利用して同一領域を抽出するテンプレートマッチング法を適用することにより、部分板書画像の結合を行う。

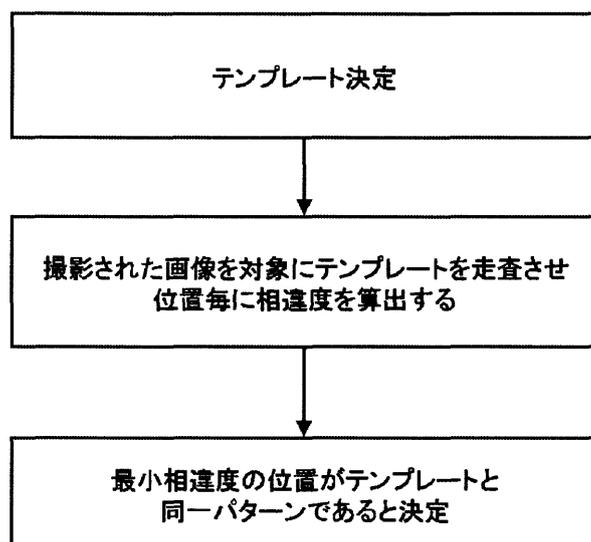


図 21 部分板書画像例

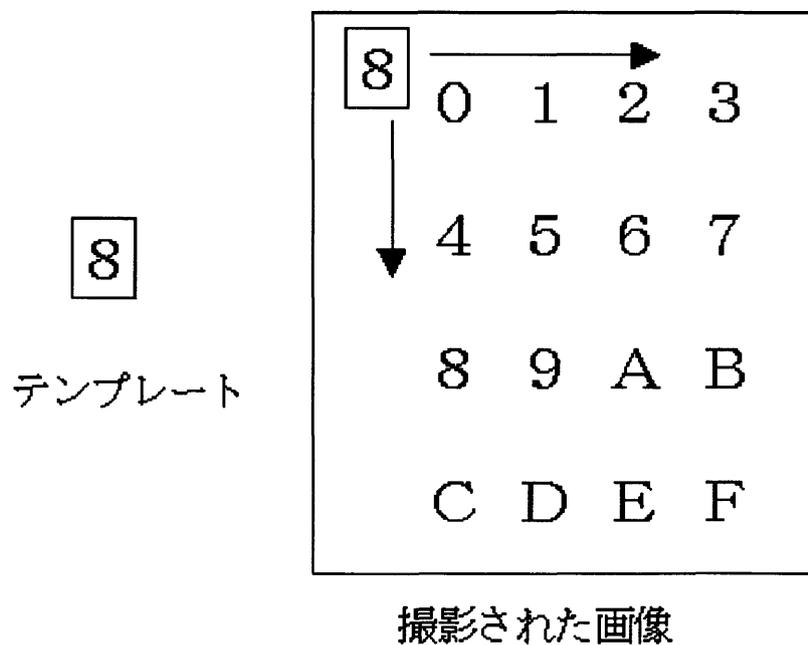
テンプレートマッチング法とは、安定して出現する画素数の少ない特徴的なパターン(以後テンプレート)をあらかじめ登録しておき、対象の画像内におけるテンプレートと同一パターンを持つ部分を検出する手法のことである(図22)。具体的にはテンプレートを決定し、テンプレートの中心を対象画像の局所領域内部のある点に置く。そしてテンプレートと対象画像の重なっている局所領域の相違度 R を計算し、最小になる位置を求めるという手法である[17]。テンプレートが対象画像内において特徴的であるほど、高精度にマッチングすることができるという性質を持つ。

本問題にテンプレートマッチングを適用するにあたり、隣接する部分黒板画像間においてテンプレートを生成する領域、すなわち共通する情報を持つ部分が存在する必要がある。そこで、黒板撮影時に撮影された隣接画像同士の一部が重複するように視野を決定し、各部分板書画像が隣接画像と共通の情報を持つように撮影した。この共通する情報を持つ領

域を候補領域とし，候補領域においてテンプレートの作成と探索範囲の決定を行う．図 23 に本研究におけるテンプレートマッチングの概要を示す．なお，図 23 中の C_t , C_s は各候補領域内におけるテンプレートの座標と探索範囲の左上座標，右下座標を示す． C_{in} は探索範囲内におけるテンプレートの左上座標，右下座標を示す．

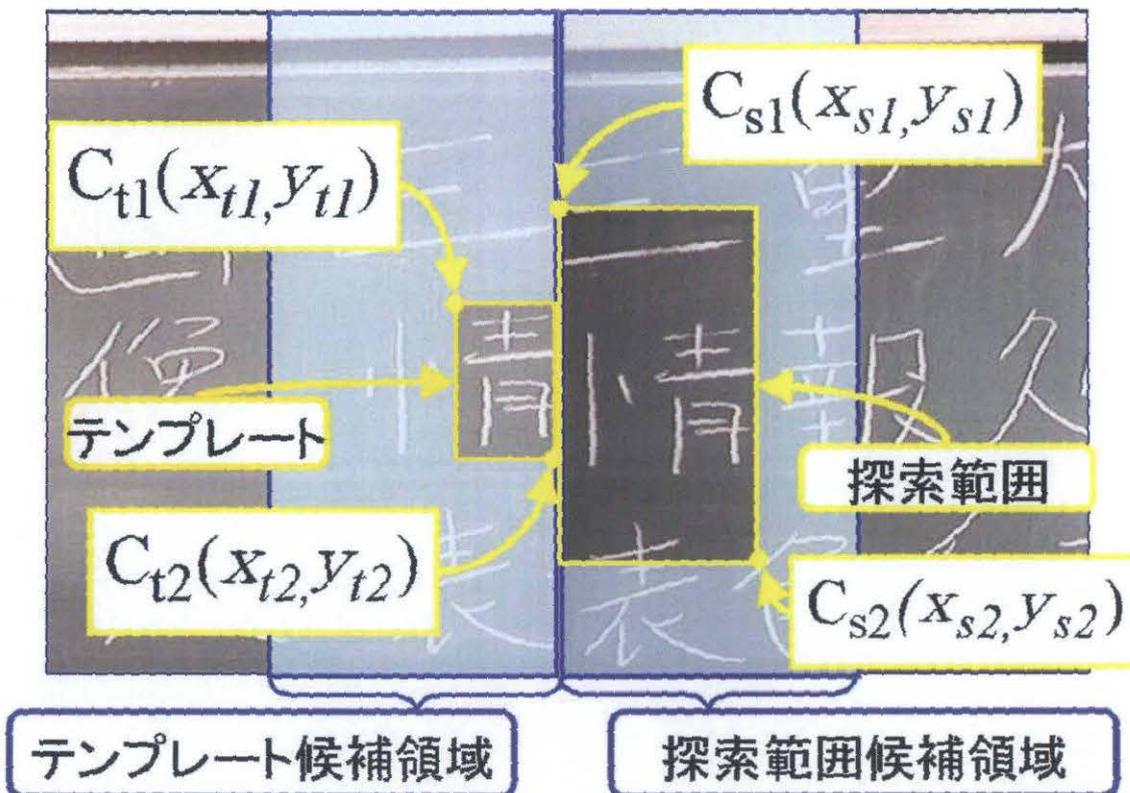


(a) フローチャート

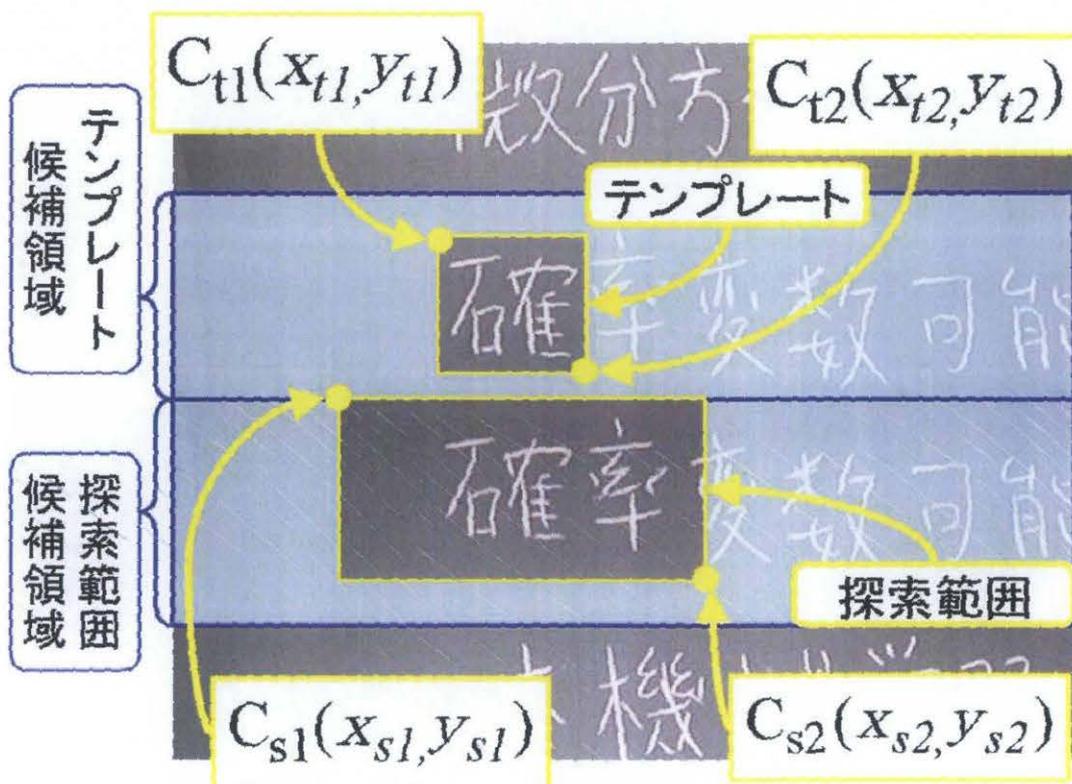


(b) 探索範囲におけるテンプレートの走査

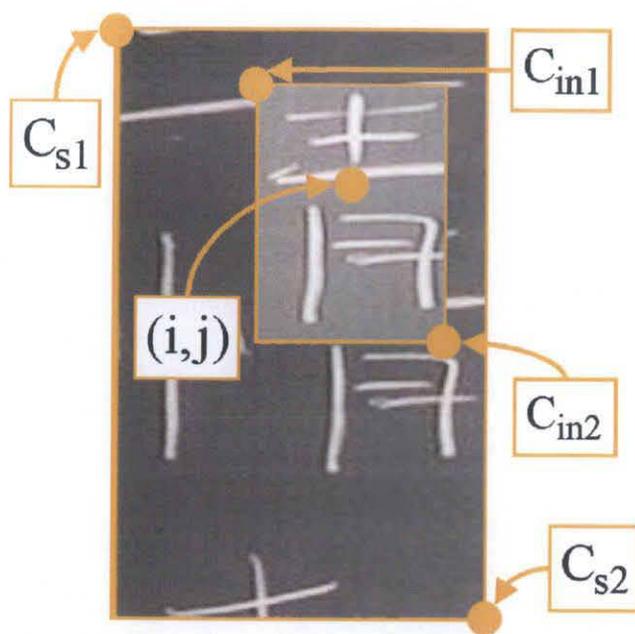
図 22 テンプレートマッチング法



(a) テンプレートと探索範囲の候補領域(横結合)



(b) テンプレートと探索範囲の候補領域(縦結合)



(c) テンプレートマッチング処理

図 23 提案テンプレートマッチング法概要

3.3 相違度算出法

3.3.1 相違度算出法概要

本論文では、テンプレートと同一のパターンを探索する際、テンプレートと重なる探索範囲中の局所領域の全ての画素において特徴ベクトルの比較を行うことにより相違度を算出し、最も相違度の小さくなる位置を求め、同一のパターンとみなす。そして、その位置を指標として画像を結合する(図24)。

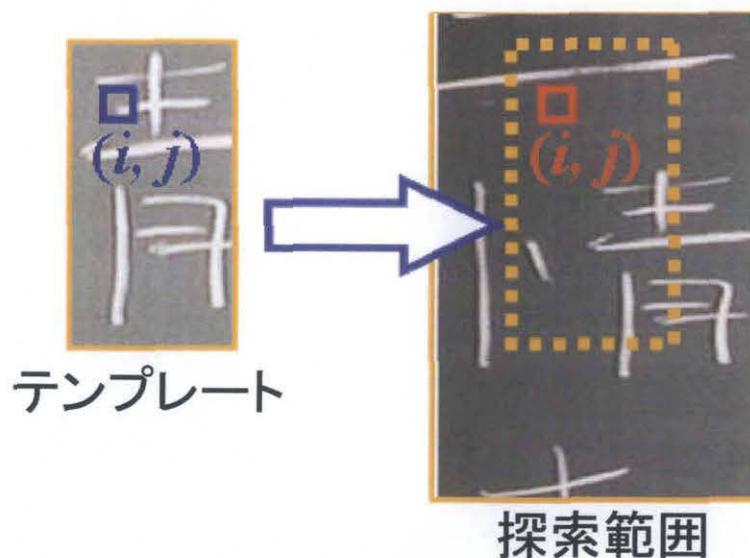


図 24 相違度算出法概要

まず、RGB 表色系の R, G, B 成分の濃度値(RGB 値)を特徴ベクトルとして相違度を算出し、部分板書の結合を行った結果、図 25 のような結合画像が得られた。図から、ずれが生じ正しく結合されていないことがわかる。このずれは 3.2 節で述べた画像間に存在する明るさの差を原因とする濃度の変化が原因と考えられる。このため、テンプレート上と探索範囲上では、黒板上の同一部分でも濃度値に差が生じ、本来最小の相違度となる座標と実際に最小となった座標が異なってしまったと考えられる。

以上より、RGB 表色系だけでなく、明るさの変化に影響を受けやすい特徴ベクトルを用いる場合、相違度が低下し、マッチングが正しく行われないと考えられる。そのため、テンプレートマッチング用いて画像結合するには、この明るさの差を補正する、あるいは明るさの変化に影響を受けにくい特徴ベクトルを用いて相違度を求める必要がある。そこで、本研究では HSV 表色系を特徴ベクトルとして相違度として用いるテンプレートマッチングを提案する。

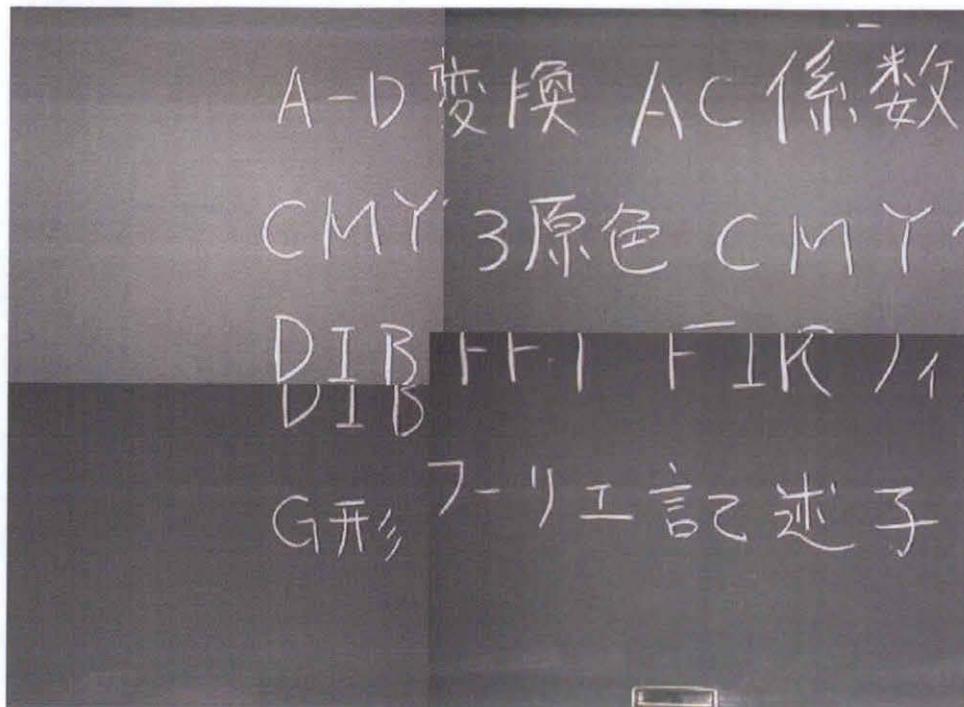


図 25 結合画像例(RGB 表色系)

3.3.2 HSV 表色系を用いたテンプレートマッチング

前節で述べたように、明るさの変化に伴って大きく値を変化させる特徴ベクトルを相違度として用いる場合、部分板書画像間に生じる濃度差を原因としてマッチング精度が低下する。そこで本研究では、明度を独立した値として持つ HSV 表色系を用いた相違度算出法を提案する。

HSV表色系とは色相(Hue)、彩度(Saturation)、明度(Value)の三つの成分を用いた色の表現空間である。光の三原色を利用したRGB表色系から変換が可能でコンピュータグラフィックスの世界においてよく用いられる[18]。

H(色相)：色の種類，360°の環で表現する。

S(彩度)：色の鮮やかさ，0～100%で表現する。彩度が0に近づくほど色味が失われ無彩色(灰色)へ近づく。

V(明度)：色の明るさ，0～100%の範囲で表現する。0なら黒を表し，100で白を表す。HSV色空間は図26に表すとおり立体柱モデルで表現することができる。

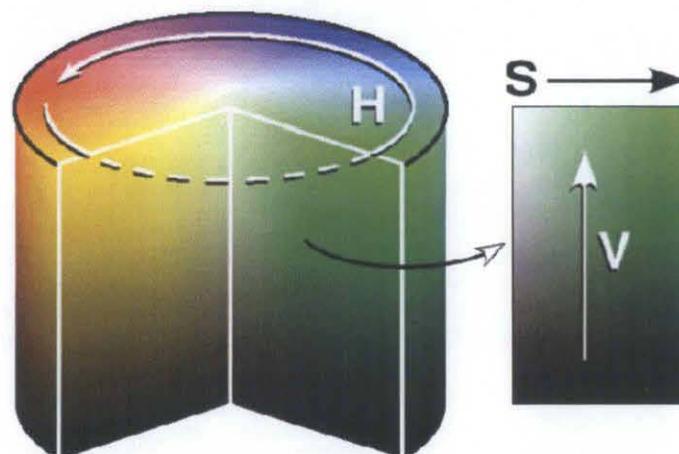


図 26 HSV色空間モデル

RGB色表色系からHSV表色系へ変換するには式(3.1)-(3.3)を用いる。

$$H = \begin{cases} 60 \times \frac{G-B}{MAX-MIN} + 0, & \text{if } MAX = R \\ 60 \times \frac{B-R}{MAX-MIN} + 120, & \text{if } MAX = G \\ 60 \times \frac{R-G}{MAX-MIN} + 240, & \text{if } MAX = B \end{cases} \quad (3.1)$$

$$S = MAX - MIN \quad (3.2)$$

$$V = MAX \quad (3.3)$$

HSV表色系からRGB表色系へ変換する場合は以下の式を用いる。

$$H_i = \frac{H}{60} \bmod 6 \quad (3.4)$$

$$f = \frac{H}{60} - H_i \quad (3.5)$$

$$p = V(1 - S) \quad (3.6)$$

$$q = V(i - fS) \quad (3.7)$$

$$t = V(1 - (1 - f)S) \quad (3.8)$$

$$\begin{aligned} H_i = 0 &\rightarrow R = V, G = t, B = p \\ 1 &\rightarrow R = q, G = V, B = p \\ 2 &\rightarrow R = p, G = V, B = t \\ 3 &\rightarrow R = p, G = q, B = V \\ 4 &\rightarrow R = t, G = p, B = V \\ 5 &\rightarrow R = V, G = p, B = q \end{aligned} \quad (3.9)$$

提案手法では、明度の変化に対して影響を受けにくいテンプレートマッチングを実現するため、(3.1)-(3.3)式を用いてRGB表色系からHSV表色系に濃度値を変換した後、明度V、色相H、彩度Sを特徴ベクトルとして相違度を算出する。HSV表色系は明度が独立しているため、明度の値を調節することにより、明るさの変化に影響を受けにくいマッチングが期待できる。特に、赤、黄、青のチョークの色及び黒板の色は、HS平面において図27のように表現でき、各色を明確に分割することが可能である。つまり、相違度を色相と彩度を利用することにより、明るさの変化の影響を小さくし、かつ黒板領域と文字領域を判別可能であると考えられる。

上述したように、HS平面上においては色相を用いることにより、色の違いから黒板領域と文字領域を区別できると考えられる。しかしながら、高精度なマッチングを行うためには、色相だけではなく、文字領域の特徴を強く示すと考えられる彩度、明度における相違度も可能な限り小さくなる必要がある。そこで、本論文では各パラメータを正規化した後、差分をとった値の和をそのまま相違度とする。(3.10)、(3.11)式に提案手法において用いる相違度Dを示す。

$$D(i, j) = \alpha \left| \frac{H_t - H_b}{360} \right| + \beta \left| \frac{S_t - S_b}{255} \right| + \gamma \left| \frac{V_t - V_b}{255} \right| \quad (\alpha, \beta, \gamma: \text{Weights}) \quad (3.10)$$

$$\text{Dissimilarity} = \sum_{i, j \in \text{Template}} D(i, j) \quad (3.11)$$

提案手法では、(3.10)式を用いてテンプレートの位置を順次移動させながら各位置における相違度を計算し、この値が最も小さくなる位置を指標として結合する。なお、各パラメータの添え字が t の場合、テンプレート中の画素のパラメータであることを示し、 b の場合、探索範囲内の画素のパラメータであることを示す。また、各パラメータの式中の α , β , γ はそれぞれ相違度内の色相 H , 彩度 S , 明度 V に対する重みである。これらの値は、マッチングの対象とする板書画像において、板書を書いたチョークの色に対応して変化させる。赤、黄、青色のチョークが持つ色は、色相上において明確に表現される領域が存在するため、 HS 平面のみに限定して相違度の算出を行う方が高い精度においてマッチングができると考えられる。そのため、これらのチョークにより書かれた板書画像に対しては γ の値を 0 とした。一方、白色のチョークは HS 平面上においては明確に白色を表現する領域が存在しないため、色相と彩度のみから黑板領域と文字領域を判別することは難しい。しかし、白色のチョークは図28に示すように、明度により表現した場合に特徴を示し、黑板領域と文字領域を判別することができる。そこで、白色のチョークにより書かれた板書画像をマッチングする場合には、各重みを調節した色相、彩度、明度を特徴ベクトルとして用いて相違度を算出する。これにより、画像間に生じる濃度差の影響を抑え、かつ文字領域の特徴を生かしたマッチングが期待できる。

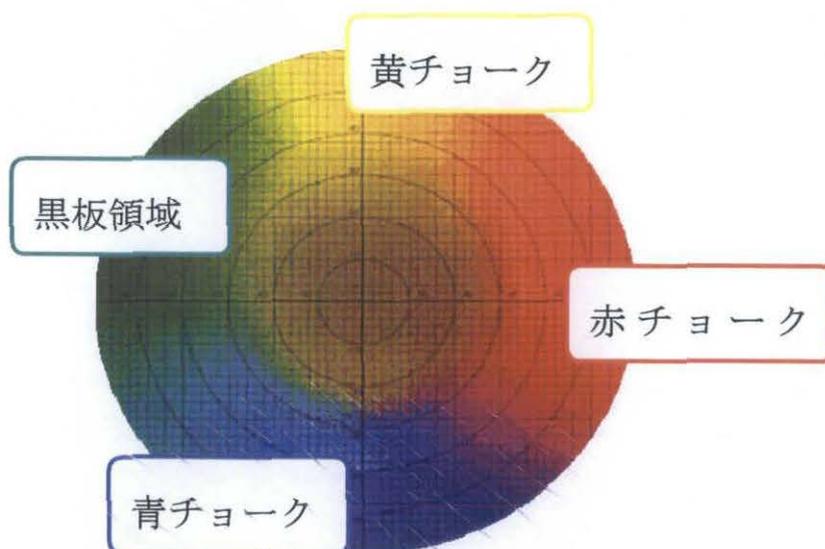


図 27 色相上における各色チョークと黑板の領域

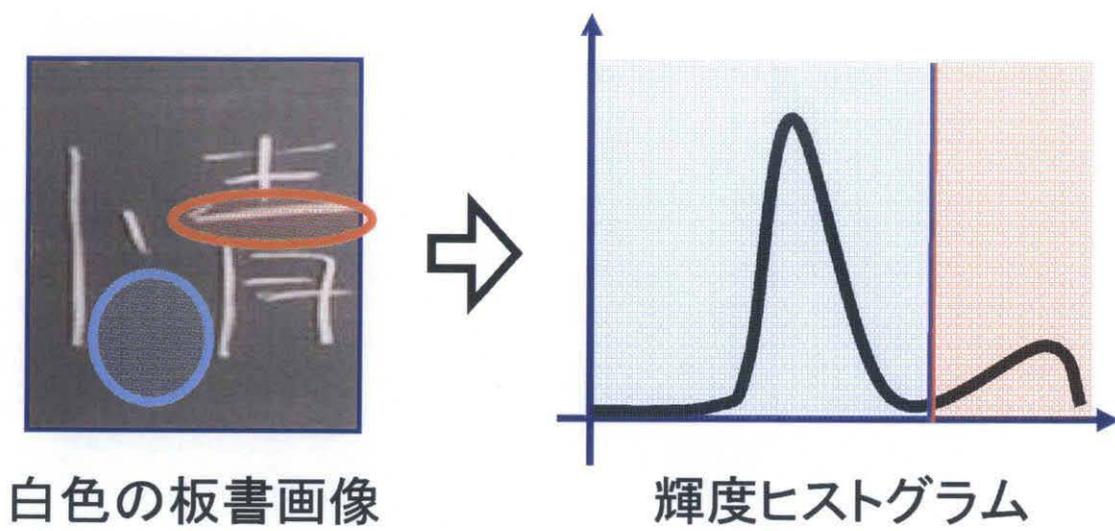


図 28 白色の板書画像における明度分布イメージ

3.3.3 色チョークに対する汎用性評価実験

前節にて示した提案手法の有用性を示すため、ここでは色チョークに対する汎用性の評価実験を行った。評価の比較対象として、RGB 表色系におけるカラーヒストグラムの差分を相違度とするマッチング方法(以下比較手法 1)と、グレースケールの濃度値の差分を相違度とするマッチング方法(以下比較手法 2)を用いた。実験環境及び使用機器を以下に示す。なお、この時の RGB 表色系からグレースケールへの変換には式(3.12)を用いた[17]。

$$Y = (0.298912 * R + 0.586611 * G + 0.114478 * B) \quad (3.12)$$

(実験環境)

CPU : Pentium D 2.80GHz
 メモリ : 512MB
 OS : Windows XP Professional
 使用言語 : Microsoft Visual Studio .NET 2003 C#

(使用機器)

板書撮影用カメラ : Canon社製 VB-C50i
 解像度 : 640×480画素, RGB各8bit/画素
 フレームレート : 0.1~30fps
 画像圧縮 : 静止画JPEG

(実験試料)

赤, 白, 黄, 青の色チョークのみを用いて書かれた板書を撮影した部分板書画像各色40枚
 縦結合テンプレート候補領域 : 640×120画素
 横結合テンプレート候補領域 : 120×480画素
 縦結合テンプレート : 580×60~580×80画素
 横結合テンプレート : 420×60~420×80画素

(評価方法)

手動にて作成した理想的な結合画像(理想画像)とマッチング手法により作成された画像(マッチング画像)との位置の差をマッチング評価値とする。

(実験概要)

板書に用いられるチョークの色に対する本手法の汎用性について評価するため、板書毎に白, 赤, 青, 黄各色のチョークを用いて筆記された 4 パターンの画像を作成した。この時、板書される文字などの特徴を揃えるため、一回の実験において筆記者は 1 人とした。実験では、撮影された板書画像に対し、各手法を用いてテンプレートマッチングを行った。そして、実験試料に記したテンプレートの中から、上述の評価方法による評価値が最も高かったテンプレートを用いたマッチ

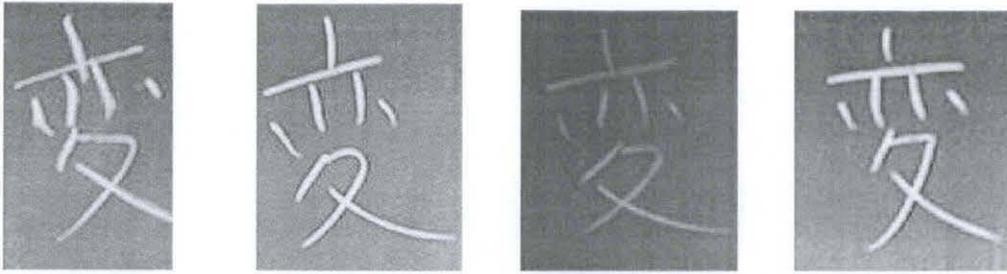
ングをその板書画像のマッチング結果とした。(3.10)式の α , β , γ は赤, 黄, 青色のチョークで書かれた板書画像に対しては $\alpha=1.0$, $\beta=1.0$, $\gamma=0$ とし, 白色のチョークについては $\alpha=1.0$, $\beta=1.0$, $\gamma=1.0$ として実験を行った。

表 1 に各色チョークで書かれた分割板書画像に対する各手法の評価値を示す。提案手法によるマッチングの評価値が最も高くなった。これは赤, 黄, 青色のチョークで書かれた板書画像については, 前節の図 27 に示したように, V 値を除いた HS 平面においてチョークと黒板が判別可能であるためと考えられる。また, 白色のチョークで書かれた板書画像については, HS 平面と明度を用いたことにより, 画像間の濃度差の影響が小さく, 文字領域と黒板領域との濃度差を特徴としてマッチングを行うことが出来たためと考えられる。実際, 文字部分の少ないテンプレートでは高い精度を得ることは出来なかった。

一方, 比較手法 1 の評価値が低い理由としては, 画像間に発生する明度の変化の影響により文字領域に比べて背景の黒板領域の似た色の部分を特徴としてマッチングが行われたためと考えられる。また, 比較手法 2 の結果が良くないのは, グレースケール化画像が明るさの変化を受けやすいことが原因として挙げられる。図 29 に各色のチョークにより書かれた文字画像のグレースケール化画像を示す。白色チョークにより書かれた文字と黒板領域とのコントラストに比べて, 各色チョークにより書かれた画像の文字と黒板領域とのコントラストが低くなっていることが図からわかる。また, 赤色チョークの場合においては, 他の色チョークの場合と比較してもコントラストが低くなっている。このように色チョークで書かれた板書画像をグレースケール化した場合, 文字領域と黒板領域の濃度値に顕著な差は生じていない。そのため, 背景の黒板領域において似た色の部分を特徴とし, さらにチョークの消し跡などのノイズの影響を受けやすくなったためマッチング精度が低下したと考えられる。なお, 現時点では本研究の目的として, マッチング時の誤差を 5 画素とすることを目標としている。これは板書文字を構成する線の太さが 5 画素程度であることから, 誤差を 5 画素以下とすれば文字認識が可能であるとの考えに基づいている。これらの結果より, 提案手法の有用性が示されたと考えることができるが, 高精度な文字認識のためには, ずれば小さければ小さい程良いと考えられるため, 次節では, 更なるマッチング精度の向上を目指し, マッチングに有効なテンプレートに関して検討する。

表 1 各手法におけるマッチング評価値

	青色	赤色	黄色	白色
提案手法(画素)	9.5	2.5	4.0	5.0
比較手法 1(画素)	31.0	25.5	21.0	32.5
比較手法 2(画素)	25.0	27.0	21.5	8.0



(a)青色チョーク (b)黄色チョーク (c)赤色チョーク (d)白色チョーク

図 29 各色のチョークにより書かれた板書画像のグレースケール画像

3.4 環境変化とテンプレートの質

3.4.1 適応的テンプレート生成の必要性

前節では固定的なテンプレートを用いた評価実験により、HSV 表色系における相違度算出法の有用性について述べた。この実験では複数のテンプレートにおいて評価を行い、最良の値を評価値としていたが、常に同一のテンプレートサイズが最良の値を示す訳ではなかった。経験的に、最良の値を示したテンプレートは、文字を多く含むテンプレートが多く、反対に黒板領域を多く含むテンプレートは最良の値となることは少なかった。後者が最良の値をとった場合は、文字領域が少ないテンプレートの中において相対的に文字領域を多く含んでいるテンプレートが多かった。このことから、板書をどれだけ含んでいるかがマッチング精度に大きく影響すると考えられる。

また、黒板の状態は書かれた板書の内容以外にも、時間の推移による日照、天気の変化、照明の有無、チョークの消し跡等、様々な要因により多様に変化する。これらの変化は文字領域の有無と同様、マッチング精度に大きく影響すると考えられる。特に図30, 31に示す時間の推移、撮影領域の変化による濃度値の変化は画像全体の濃度値を変化させ、マッチング精度に及ぼす影響は非常に大きいと考えられる。

以上の 2 点から、高精度に結合を行うためには、板書の状態毎にそれぞれ適切なテンプレートを作成する必要があると考えられる。

ここで、板書の内容を変化させず、時間の推移による環境の変化によりマッチング精度がどの程度影響を受けるか調査するべく、実験を行った。実験概要を以下に示す。実験環境、使用器具及び評価方法は 3.3.3 の実験と同じものを用いた。

(実験試料)

白色のチョークのみを用いて書かれた板書を撮影した部分板書画像120枚

縦結合テンプレート候補領域：640×120画素

横結合テンプレート候補領域：120×480画素

縦結合テンプレート : 580×60～580×80画素

横結合テンプレート : 420×60～420×80画素

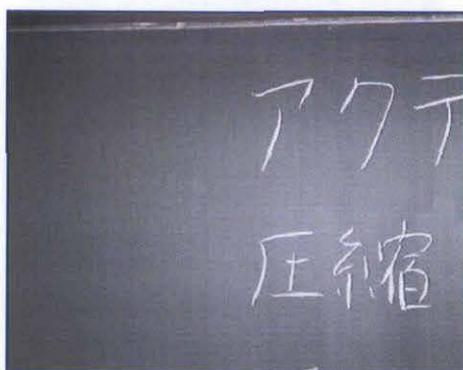
(実験内容)

1. 本学講義の1コマ目開始時(8:50)に黒板の撮影を行う。
2. 1. で得た画像に対し、3.3.3と同様に実験試料に記したテンプレートの中から、上記の評価方法による評価値が最も高かったものを用いた。
3. 2～5コマ目の開始時と5コマ目の終了時において黒板の撮影を行い、2. で決定したテンプレートを用いて各画像に対しマッチングを行う。
4. 2.の時点の評価値と3.で算出した評価値の差を計算する。

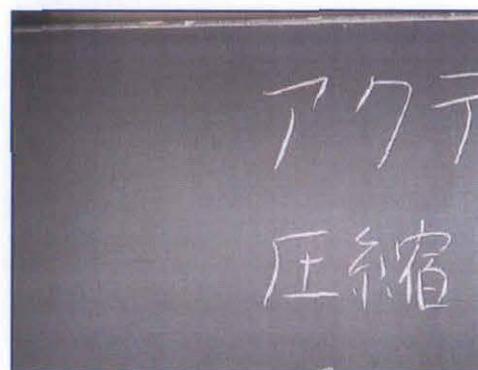
実験結果を表2に示す。本問題においては同一のテンプレートを用いた場合、マッチング精度が安定せず、常に高い精度でマッチングを行うことは困難であることがわかった。そこで次節においてどのようなテンプレートを用いれば高精度にマッチングを行うことが出来るか検討を行う。また、4章においては動的なテンプレート生成手法について検討する。

表 2 環境の変化による濃度変化のマッチング精度に対する影響

最大誤差(画素)	平均誤差(画素)
15.0	5.0

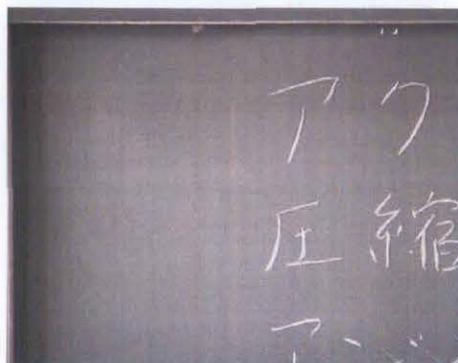


(a)左上黑板画像(8:50)

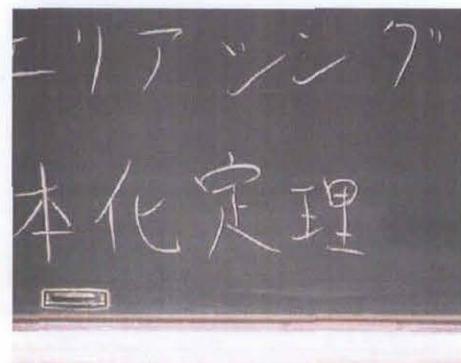


(b)左上黑板画像(16:10)

図 30 時間変化による濃度値の変化



(a)左上黑板左上部画像



(b)左上黑板右下部画像

図 31 場所変化による濃度値の変化

3.4.2 テンプレート評価実験

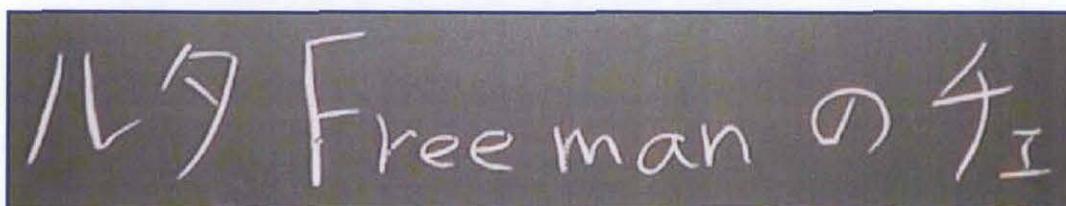
講義中，黒板は板書内容だけでなく，日照や照明，チョークの消し跡等，様々な規則性のない要因により，状態を変化させる．このような性質をもつ黒板を対象とする場合，高精度にテンプレートマッチングを行うためには，高精度でマッチングが可能な質の高いテンプレートがどのようなものか調査する必要がある．そこで，テンプレートの質を調査するため，実験を行った．実験概要を以下に示す．実験環境，評価方法，及び使用器具は3.3.3の実験と同じものを用いた．

(実験試料)

白色のチョークのみを用いて書かれた板書を撮影した部分板書画像40枚

(実験内容)

1. 一つの部分板書画像の候補領域において，図32のように(c)黒板領域のみを含む場合，(d)複数の文字を含む場合，(b)1文字を含む場合の3通りのテンプレートを作成する．
 2. 1. で作成したテンプレートを用いて，提案手法による画像の結合を行う
 3. 2. の結果において，手動にて作成した理想画像との誤差が5画素以内のマッチングを成功とし，成功率を算出する．
- ※ 1文字を含むテンプレートについては複数候補が存在するため，目視において最も複雑な文字と判断したものをテンプレートとして用いた．



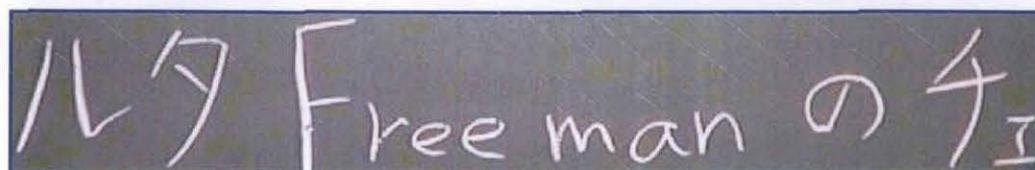
(a)テンプレート候補領域



(b)1文字を含むテンプレート



(c)黒板領域のみのテンプレート



(d)文字列を含むテンプレート

図 32 作成テンプレート例

実験結果を表3に示す。1文字を含むテンプレートが最も高い値を示した。1文字を含むテンプレートは黒板領域の割合が少なく、特徴となる文字領域の割合が最も多くなるためと考えられる。一方、黒板領域のみを含むテンプレートは非常に低い値となった。黒板領域のみでは文字が存在しないために特徴に欠け、画像間の濃度差の影響も受けやすいため、マッチング精度が低くなったと考えられる。文字列を含むテンプレートは80%以上を示した。この種のテンプレートを用いたマッチングでは、10画素以下6画素以上の誤差が生じる場合が多く見受けられた。これは、文字列を含むテンプレートは指標となる特徴は存在するが、文字同士の間隔が存在するため、1文字を含むテンプレートより黒板領域の割合が多くなるためであると考えられる。実験の結果、1文字を含むテンプレートよりも濃度差の影響を強く受けることとなり、マッチング精度が低下したと考えられる。

1文字を含むテンプレートを用いてマッチングした場合における、失敗例について考察する。マッチングに失敗した際に用いたテンプレートの例を図33に示す。図33(a)のテンプレートは文字の一部であり、サイズが小さく、文字画素が少なくなっている。結果、濃度差が影響しマッチング精度が低下したと考えられる。一方、図33(b)のテンプレートは文字が欠損することなく1文字全体を含んでいる。しかし、候補領域の中に同じ種類の文字が存在したため、その同じ種類の異なる文字とマッチングしてしまった。同様の例としてテンプレート図33(c)の例が挙げられる。このテンプレートのように非常に簡単な形であった場合、文字領域の割合が大きくても、図34のように探索範囲内に複数の似た領域が存在することになり、マッチングが失敗することがある。

以上より、質の良いテンプレートとは以下の条件を満たすテンプレートであると考えられる。

1. 黒板領域の割合が少なければ少ないほど、文字部分の割合が多ければ多いほど良い。
2. 候補領域中、出来るだけ複雑な文字領域を含むこと
3. 一定以上の大きさを持つこと。(1文字の大きさに近いほど良い)

図35にマッチングに成功した際のテンプレート例を示す。これを見ると、上記の条件を全て満たしていることがわかる。次章では、これらの条件から評価関数を定義し、自動的に質の良いテンプレートを生成するための手法について検討する。

表 3 テンプレート評価実験結果

テンプレート	文字列	黒板	1文字
誤差 5 画素以下	29/35 (82.8%)	1/35 (2.9%)	32/35 (91.3%)
誤差 6 画素以上	6/35 (17.2%)	34/35 (97.1%)	3/35 (8.7%)

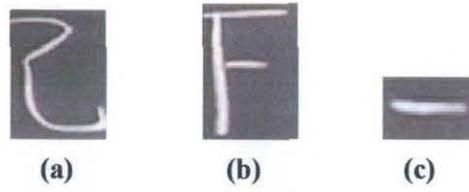


図 33 マッチング失敗テンプレート例

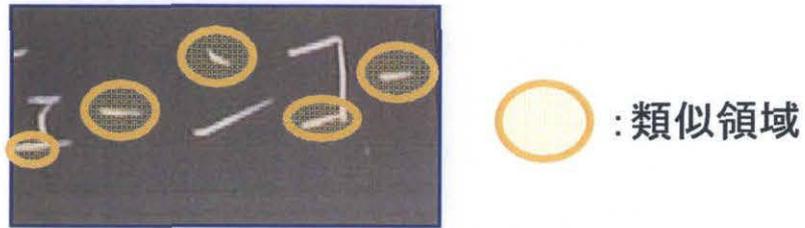


図 34 探索範囲中の類似領域例



図 35 マッチング成功テンプレート例

第4章 遺伝的アルゴリズムを用いたテンプレート自動生成法

4.1 テンプレート評価方法

4.1.1 評価関数

前章では、黒板の撮影方法、撮影した部分板書画像のテンプレートマッチングによる結合手法、テンプレートマッチングに用いるテンプレートの優劣について検討した。その結果、安定して高精度の部分板書画像の結合を行うためには、結合を行う際に毎回候補領域中から質の良いテンプレートを選択し、マッチングを行う必要があることがわかった。そこでこの章では、候補領域中から適切なテンプレートを自動的に生成する手法について述べる。

3.4.2節において、以下の三つの条件を満たすテンプレートが良いテンプレートであると考えた。

1. 黒板領域の割合が少なく、文字部分の割合が多いこと。
2. 候補領域中、出来るだけ複雑な文字領域を含むこと。
3. 一定以上の大きさを持つこと。

これらの条件から、次のようにテンプレートの評価基準を定めることとした。

1. テンプレート中の文字領域の割合が大きいほど良い。
2. 交点を多く持つ文字領域ほど良い。
3. 1文字の大きさに近いほど良い。

この評価基準を基にテンプレートを探査した場合、質の高いテンプレートが作成されることが期待される。また、同様に探索範囲も評価関数を用いて探索することにより、テンプレートと同一の情報を含み、かつテンプレートに近い大きさの探索範囲が作成されると考えられる。そこで本研究では、上記の評価基準を基にテンプレートと探索範囲を評価する関数を作成し、質の高いテンプレートと探索範囲の探索・作成を行う。

本研究で用いた評価関数((4.1)-(4.3)式)を以下に示す。

評価関数1

$$E_1 = \frac{N_{tmp-c} - N_{tmp-b}}{N_{tmp} \times N_{search}} \quad (4.1)$$

ここで(4.1)式中の N_{tmp-c} 、 N_{tmp-b} はテンプレート内の文字画素数とノイズ画素数を示し、

N_{tmp} はテンプレート内の白画素数(チョークの色の画素数)を, N_{search} はテンプレートのサイズを示す. この評価関数はサイズが小さく, より多くの文字画素を含むテンプレートに高い評価値を与える. なお, 文字領域とノイズの判別はラベリングを用いて算出した結合画素数の大きさを基準として判別した.

評価関数2

$$E_2 = \begin{vmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{vmatrix} \quad (4.2)$$

(4.2)式中の S_{xx} , S_{yy} はテンプレート内の文字画素の水平方向分散, 垂直方向分散を示し, S_{xy} はテンプレート内の文字画素の共分散を示す. この式ではテンプレート内において文字部分と考えられる画素の分散共分散行列を求め, その行列式すなわち一般化分散を求めることにより, 評価値とした. この評価関数はテンプレート内に文字画素が偏りなく分散している場合に高い値となる.

評価関数3

$$E_3 = \frac{A_{IP}}{A_{temp}} \quad (4.3)$$

(4.3)式中の A_{IP} は候補領域中の文字が持つ交点を中心として縦横共に n ピクセル以内の領域を交点領域としたとき, テンプレート内に存在する交点領域の面積を示す. A_{temp} はテンプレートの面積を示す. すなわち(4.3)式はテンプレート内に交点が多く存在するほど高い評価値を与える関数である. 図 36 に(4.3)式の交点領域のイメージ図を示す. なお, 交点は, 細線化した後に 3 分岐, 4 分岐の交点を検出することの出来る 3×3 の論理フィルタを用いることにより検出している[19].

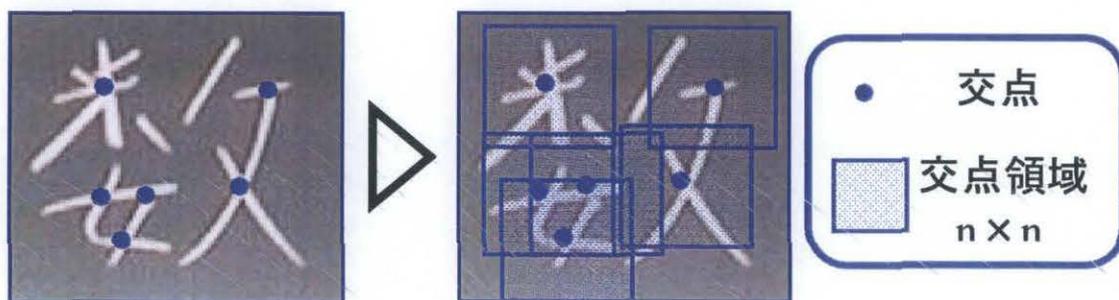


図 36 (4.3)式における交点領域のイメージ

以上, 三つの評価関数の線形和(4.4)式を用いてテンプレートの評価を行うこととした.

$$E = E_1 + E_2 + E_3 \quad (4.4)$$

4.1.2 評価実験

(4.1)-(4.3)式の有用性を確認するために、評価実験を行った。実験概要を以下に示す。実験環境及び、使用器具は3.3.3の実験と同じものを用いた。ここで、評価関数を用いる探索手法として、学習が不要であること、解空間の全域探索が可能であることから遺伝的アルゴリズムを用いた。探索に用いる解は候補領域内の座標とした。図37のように候補領域内の座標を指定することにより、テンプレートを作成可能である。図37中の $(x1,y1)$ 、 $(x2,y2)$ は候補領域におけるテンプレートの左上座標、右上座標を示す。同様に $(x3,y3)$ は候補領域における探索範囲の左上座標、右上座標を示す。なお、本実験は評価関数の有用性の確認を行う実験であるため、初期解を文字領域の近傍に設定した。また、作成したテンプレートの探索範囲は図38に示すように、探索範囲候補領域においてテンプレート候補領域において作成したテンプレートの周囲に40画素拡張した領域とする。探索範囲の座標が探索範囲候補領域を超えた場合、候補領域端の座標に置き換えることとする。

(実験試料)

白色のチョークのみを用いて書かれた板書を撮影した部分板書画像32枚

縦結合テンプレート候補領域：640×120画素

横結合テンプレート候補領域：120×480画素

(実験内容)

1. あらかじめ5000世代の試行を行い、その結果得たテンプレートの評価値を仮想的な飽和値とする。そして、飽和値の割合(99, 98, 97・・・%)から閾値を算出する。
2. 閾値ごとに、指定した閾値を超えるテンプレートを発見するまで探索を行う。
3. 発見されたテンプレートを用いてマッチングを行い、結合の成否を判定する。実験により得たマッチング画像と手動にて作成した理想画像の差が5画素以下の時、成功とする。

実験の結果、図39のような飽和値割合(評価値)とマッチング精度の関係が得られた。評価値の上昇に伴い、マッチング精度が上昇し、飽和値割合99%の時には86%(成功数28回のマッチング中24回)のマッチング精度を得ることができた(図39)。しかし、評価値が飽和値の90%を超える頃から、必ずしも評価値の上昇に伴い、精度が上昇しているわけではないことがわかる。飽和値に近づく程にこの傾向は顕著になり、飽和値まで算出した場合のマッチング精度よりも96~99%におけるマッチング精度の方が高くなっている。このことから、(4.1)-(4.3)式を用いたテンプレート評価法は更なる検討が必要ではあるが、有効な手法であると言える。しかしながら、この(4.1)-(4.3)を用いて探索を行う際、候補領域全域を全探索した場合、計算量は非常に大きくなる。そこで次節では、(4.1)-(4.3)式を用いて効率良く探索を行うための手法について検討を行う。

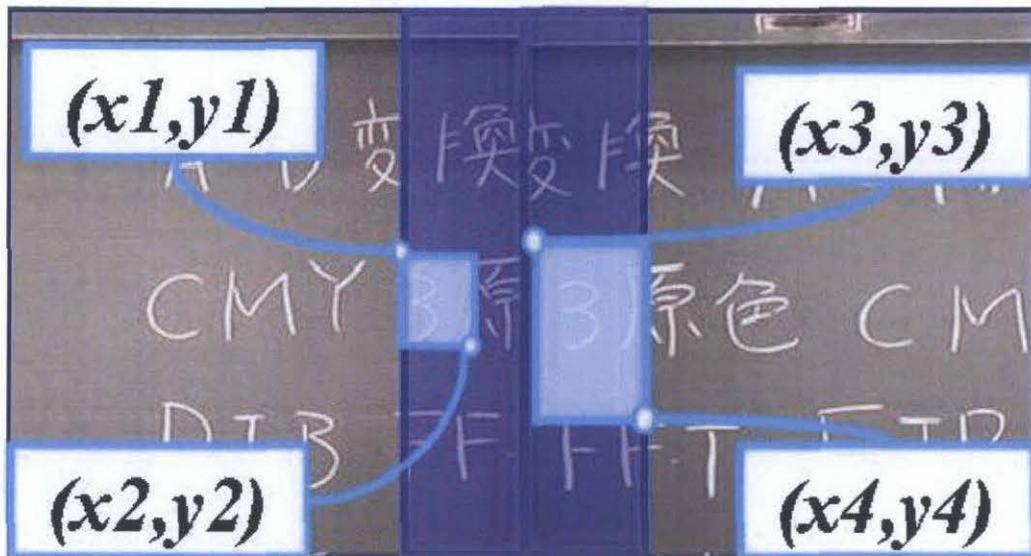


図 37 候補領域内の座標指定によるテンプレート作成・探索範囲決定イメージ

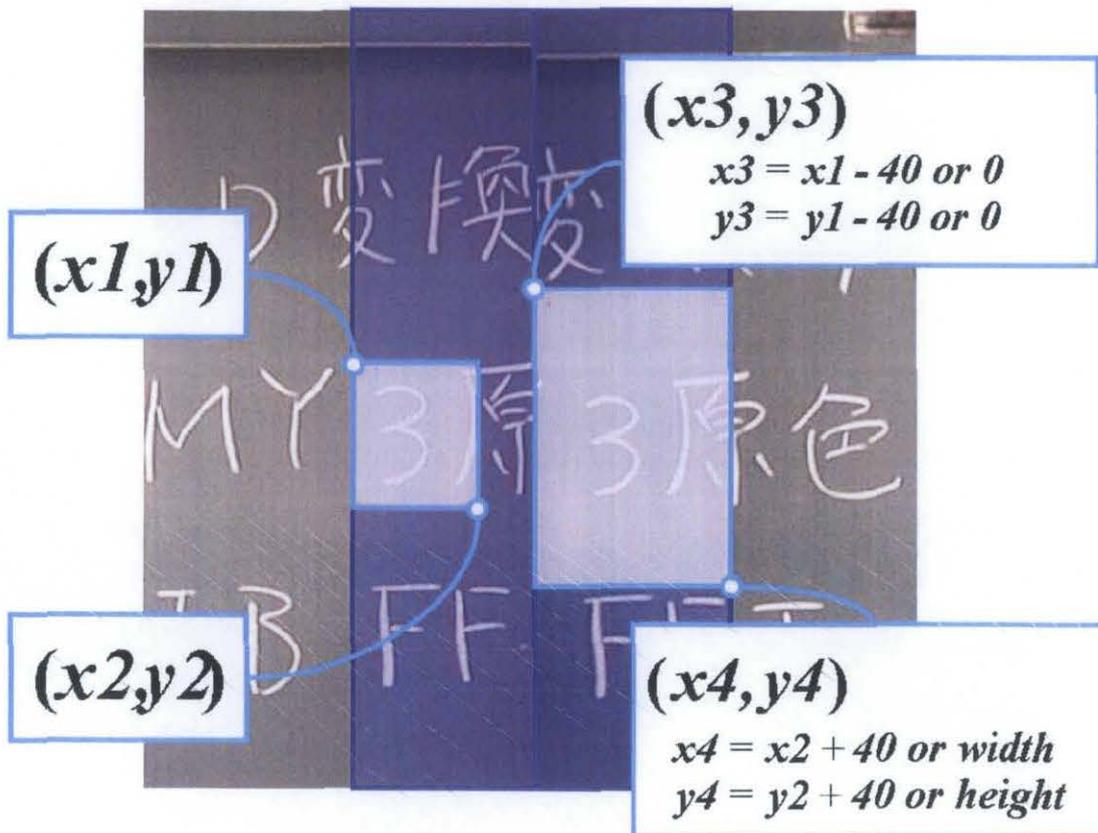


図 38 評価実験における探索範囲決定イメージ

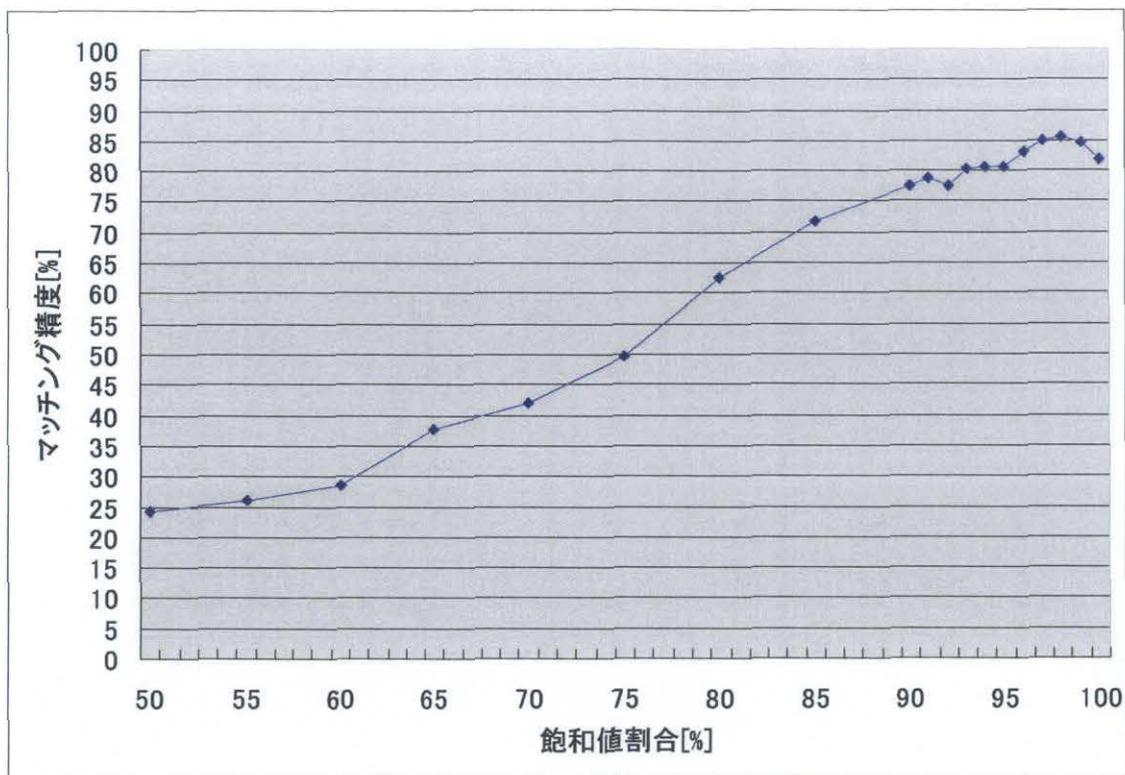


図 39 評価値とマッチング精度の関係

4.2 遺伝的アルゴリズムを用いたテンプレート生成法

4.2.1 遺伝的アルゴリズムの適用

前節では, (4.1)-(4.3)式を用いてテンプレートを抽出し, マッチングを行った. 結果, 86%のマッチング精度を得ることが出来た. しかしながら, 本問題における解空間は非常に広大であり, 最大で2の200乗を超える(縦結合テンプレート(探索範囲)候補領域×横結合テンプレート(探索範囲)候補領域の7乗). この解空間において, 全探索によりテンプレートを抽出することは現実的ではない. そこで, 効率的なテンプレートの探索法が必要となってくる. 本問題においては以下のような条件を満たす探索法が有効であると考えられる.

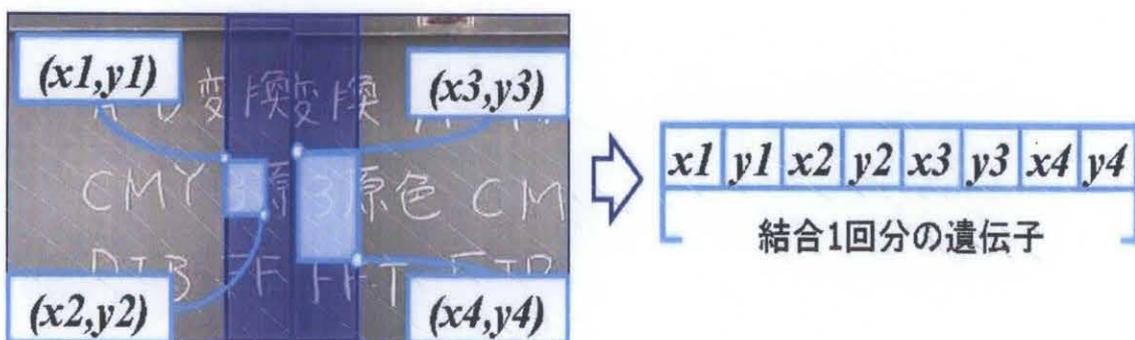
1. 非常に多様に黒板の濃度分布状態が変化するため, 黒板状態の変化パターンを学習することは困難である. そこで, 学習を前提としていない手法
2. 前節図39より, 飽和値よりも飽和値の96%から99%において最も高い値が示されていることから, 長い時間をかけて厳密に最高の評価値を解く手法ではなく, 短時間で最適解に近い解が得られる手法
3. 本問題は候補領域内において, テンプレートと探索範囲の座標を指定することによりテンプレートを作成している. すなわち, 本問題は組み合わせ最適化問題として考えることが出来るため, 組み合わせ最適化問題を解く手法
4. 複数存在するテンプレート候補の中から, 出来るだけ良いテンプレートを取得するため, 局所的ではなく, 全域的に探索を行うことの出来る手法

以上のことから本研究では, テンプレートの抽出法として, 遺伝的アルゴリズムを引き続き用いる.

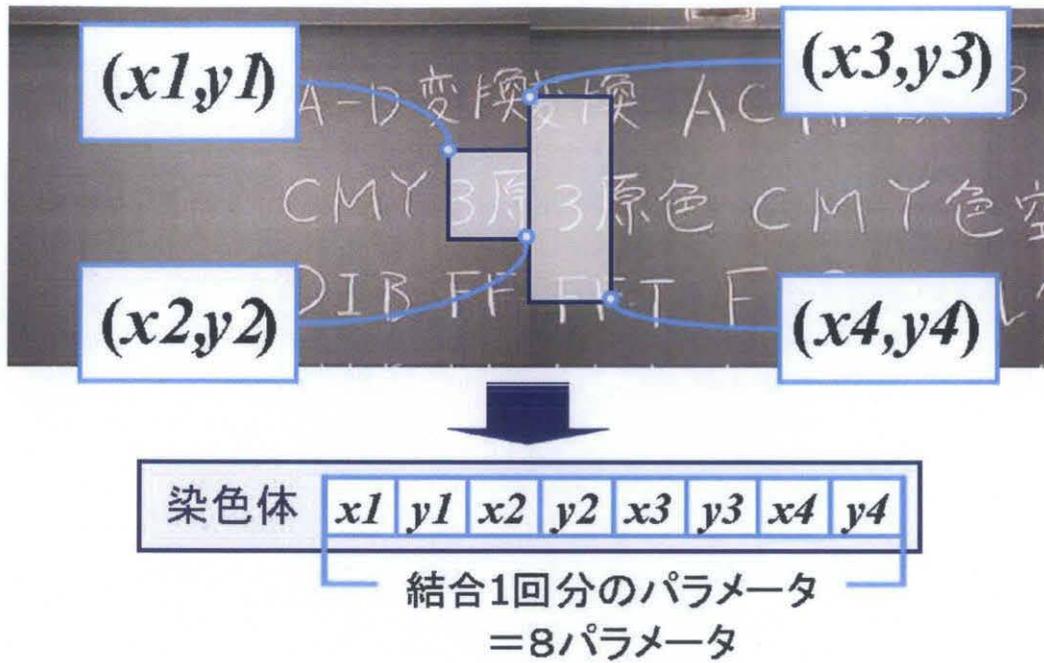
4.2.2 染色体のコーディング法

本問題における染色体のコーディング法について説明する。図40に染色体のコーディング法の概要を示す。本研究においては、遺伝的アルゴリズムの適用範囲を8枚の部分板書画像の結合とする。すなわち、8枚の部分板書画像の結合に必要なテンプレートと探索範囲を、遺伝的アルゴリズムと評価関数を用いることにより探索する。ここで、図37に示したように候補領域中の座標 $(x1,y1)$ と $(x2,y2)$ の差からテンプレートサイズ、位置を決定し、 $(x3,y3)$ と $(x4,y4)$ から探索範囲を決定する。これらを染色体に変換すると、マッチング1回分の遺伝子は図40(a)のように示される。ここで、本問題において、1枚の結合画像を作成するにあたり、マッチングは7回行う必要があることを考えると、コーディング法としてはマッチングの回数分探索を行う方法(図40(b))と、一度の探索によりマッチングに用いるテンプレートを全て探索する方法(図40(c))の2通りが考えられる。前者では、遺伝子数8(テンプレート座標 $(x1,y1)$ 、 $(x2,y2)$ + 探索範囲座標 $(x3,y3)$ 、 $(x4,y4)$)の染色体を用いてテンプレートの探索を7回繰り返した後に、画像の結合を行うこととなる。後者では、遺伝子数7(マッチング回数) \times 8 = 56変数の染色体を用いて、一度の探索から7枚のテンプレートを得て結合を行うこととなる。

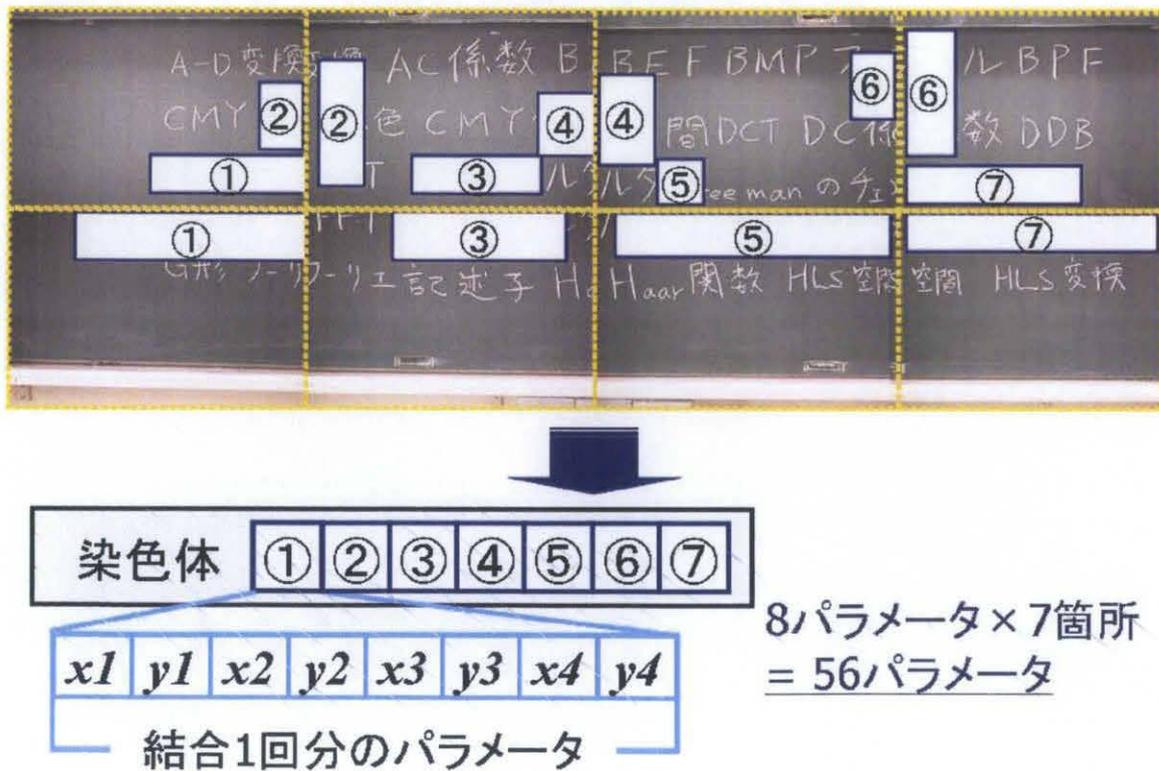
これらのコーディング法を比較するため、評価値の推移のみを見る程度の簡単な実験を行った。その結果、4.1.2節で示したマッチング精度が最大となる評価値には、図40(c)の7回分のマッチングを統合した場合の染色体として構築したコーディングによる探索の方が少ない世代数で到達する傾向が見られた。このことから、本論文では図40(c)に示す染色体のコーディング法を用いることとする。なお、染色体の評価関数としては(4.1)-(4.3)式を用いる。また、探索範囲の評価にはテンプレートよりも必ず大きい探索範囲となるような境界条件を付加する。



(a) マッチング1回分の遺伝子イメージ



(b) 1回分のマッチングの染色体イメージ



(c) 7回分のマッチングを統合した場合の染色体コーディングイメージ

図 40 染色体のコーディング法

4.2.3 遺伝的演算

本研究にて用いた遺伝的操作を表4に示す。これらのパラメータは簡単な予備実験の結果、最も有効だと考えられるものを用いた。図41は、提案手法における染色体の評価値の推移を示す。横軸は世代数、縦軸は、染色体の正規化された評価値を意味する。図41は10回試行の平均値を示す。図41では100世代付近から上昇が緩やかになり、300世代付近では評価値が収束に向かっていることがわかる。本研究において作成したシステムでは、一回の試行に約3分程度の時間を要した。また事前調査の結果、実際の講義において個人差はあるものの、撮影対象の黒板に再度板書されるには20分程度要することがわかっている。従って、現時点では時間のみに焦点を当てた場合、講義中に結合画像を作成可能であるといえる。

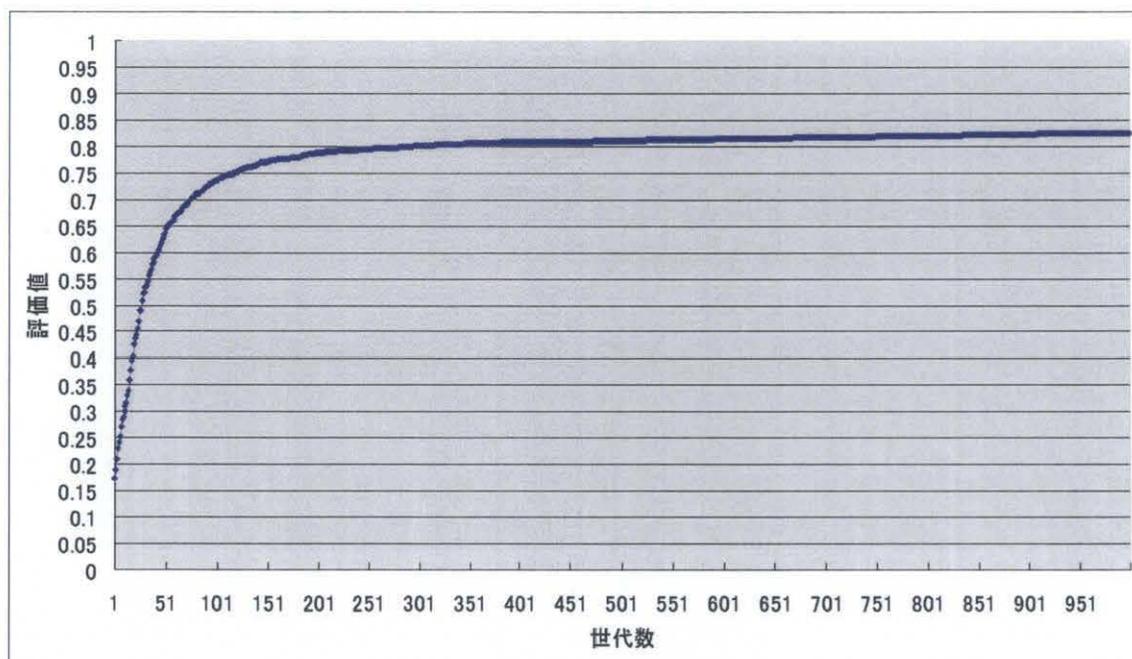


図 41 評価値の推移(1000 世代)

表 4 遺伝的アルゴリズムに使用したパラメータ

染色体数	30個
淘汰・増殖率	0.20
交叉手法	一様交叉
交叉率	0.50
突然変異率	0.15

第5章 提案手法を用いた部分板書画像結合実験

5.1 実験環境

3, 4章まで, 部分板書画像を結合するための手法について述べた. これを提案手法とし, 提案手法の有用性を確認するための実験を行う. 実験環境を以下に示す.

(実験環境)

OS : Windows XP Professional

CPU : Pentium D 2.80GHz

メモリ : 2048MB

(使用機器)

板書画像撮影用カメラ : Canon社製 VB-C50i

解像度 : 640×480画素, RGB各8bit/画素

フレームレート : 0.1~30fps

画像圧縮 : 静止画JPEG

(実験試料)

実際の教室において撮影した白色のチョークで書かれた部分板書画像56枚

縦結合時の候補領域サイズ : 640×120画素

横結合時の候補領域サイズ : 480×120画素

(比較手法)

イメージモザイク手法の一つである”Multi-histogram法による画像のモザイク”[14]を用いた.

(実験概要)

実験試料に対し, 提案手法, 比較手法それぞれを適用する. 各手法により得られた結合画像と理想画像との誤差が5画素以下のときに結合成功とし, 成功率を算出する. なお, 提案手法は表4の遺伝的演算を1000世代行い, それを10回試行して得られた成功回数の平均値から成功率を算出する.

5.2 実験結果と考察

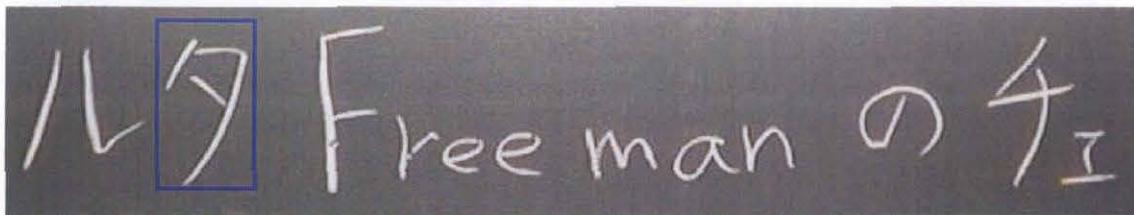
5.2.1 提案手法の実験結果と考察

実験の結果、得られたテンプレートの中から成功例を図42、失敗例を図43に示す。また、結合板書画像成功例を図44に、失敗例を図45に示す。高い評価値を得た場合、図42(a)のように候補領域内においては複雑であり、かつ1文字をちょうど含むテンプレートが生成された。あるいは図43(b)のように1文字丁度ではないが、複雑な文字を部分的に含むテンプレートが生成されていた。さらに、図43(c)のように、曲線を中心として構成されている文字を含むテンプレートも存在した。このようなテンプレートは希であり、高精度の結合が期待できると考えられる。実際、これらのテンプレートを用いたマッチングにおいては高い精度で画像を結合することができた(図44(a))。以上の結果より、さらに精度を向上させるためには、図42(c)のように曲線部分が多いテンプレートを抽出するため、文字を構成する線の形状を評価する評価関数の追加等も有効であると考えられる。

一方、図44(a)のような交点とその周辺部分のみを含むテンプレートも生成された。図44(a)のテンプレートは(4.3)式に評価が偏ったため、生成されたと考えられる。図44(a)のように交点のみをテンプレートとして抽出された場合、高い類似度を持つ領域が複数存在し、結果としてマッチング精度を大幅に低下させると考えられる。また、図44(b)のようなテンプレートも生成された。このようにテンプレートサイズが著しく小さい場合、文字部分の特徴である形状が存在しないため、一定の精度でマッチングすることは困難であると考えられる。しかし、このようなテンプレートは(4.1)、(4.2)式において高い評価値をとる。これは、(4.1)式においては評価値がテンプレートサイズに対する文字画素の面積の比率に強く影響を受けるため、(4.2)式においては画素が少量でも面積が小さい場合に評価値が高くなってしまふことが原因と考えられる。特に、後者は図44(c)のようなテンプレートを用いてマッチングを行った場合においても同様の結果となる。従って、これを改善するためには、一つの式で画素と面積を評価するのではなく、二つの式を用いてそれぞれを評価する、あるいは、分散に画素の概念を取り入れる等の評価関数の改善が必要ではないかと考えられる。これら図43に示したようなテンプレートを用いてマッチングを行った結果、図45(b)における「ずれ」が生じている部分のように結合に失敗した。これらのテンプレートの生成を防ぐためには、更なるテンプレートの質を考慮することができる評価関数が必要となると考えられる。また、図46のようにテンプレートの質に関わりなく、候補領域内における非共通部分からテンプレートが生成される場合も見受けられた。このようなテンプレートは例外的に低い評価を与えているが、隣接画像の候補領域において共通部分が少ない場合においても生じることがある。このようなテンプレートの生成を抑制するためには、探索以前の処理の時点で候補領域内における非共通部分の検出を行い、探索範囲から除外する等の処理を行う必要があると考えられる。

提案手法の隣接画像毎の結合成功率は81.4%であった。しかしながら、文字認識を安定し

て行うためにはより正確なマッチング精度が求められ、そのために更なる評価関数の検討が必要である。



(a) 相対的に複雑な文字を含むテンプレート



(b) 複雑な文字の一部分を含むテンプレート

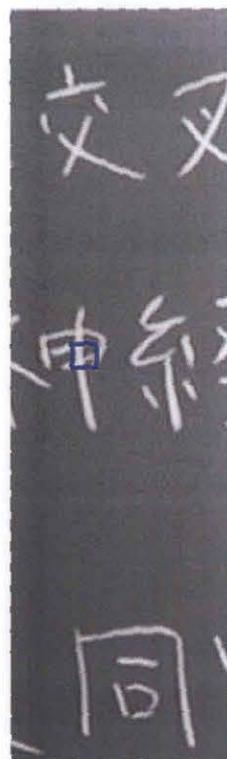


(c) 曲線により構成される文字を含むテンプレート

図 42 テンプレート生成成功例



(a) 交点部分のみを抽出したテンプレート

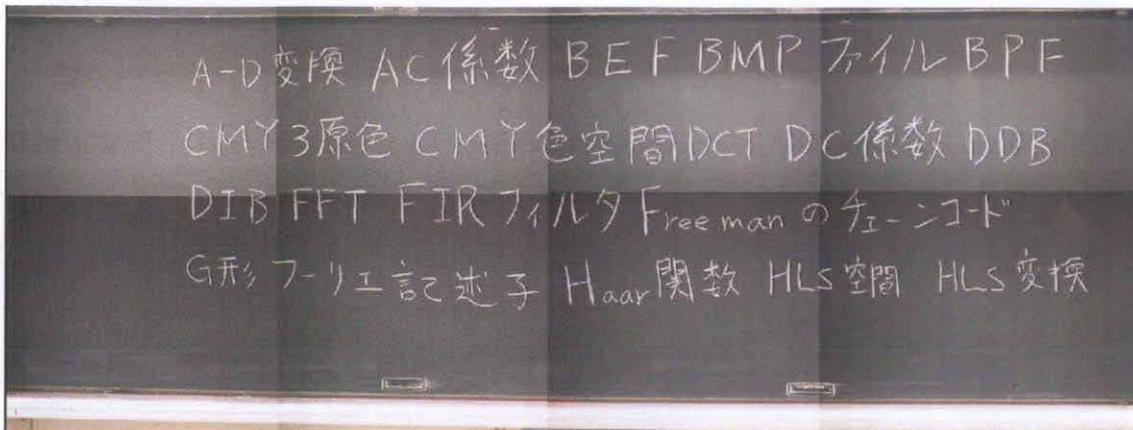


(b) 極めて小さいテンプレート

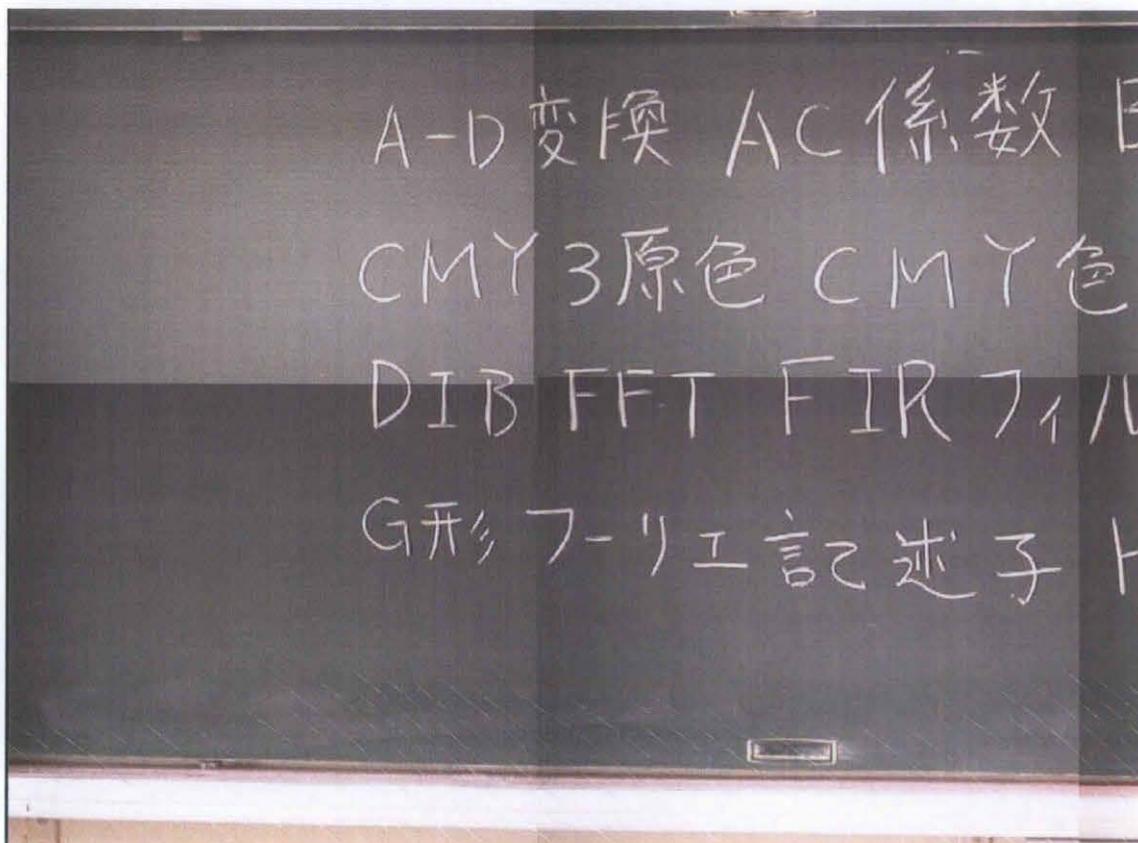


(c) 角の部分抽出したテンプレート

図 43 生成テンプレート生成失敗例

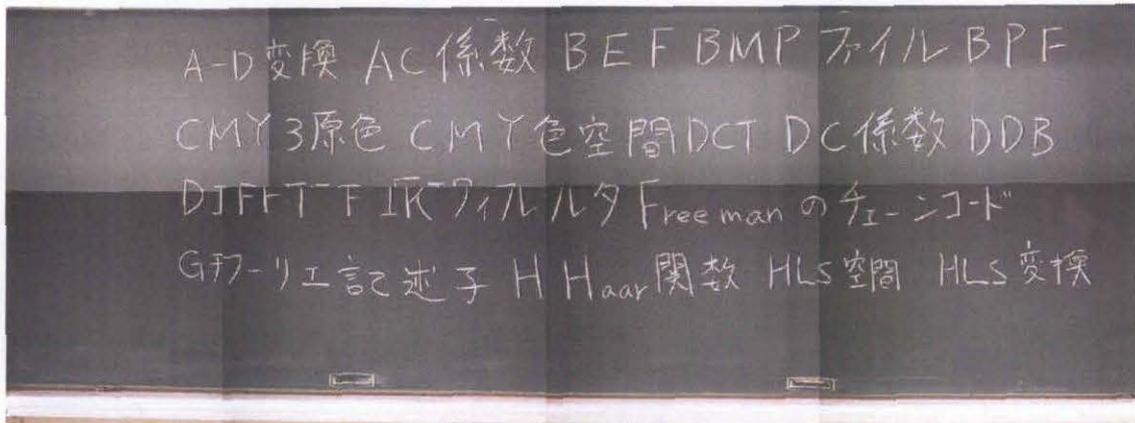


(a)全体画像

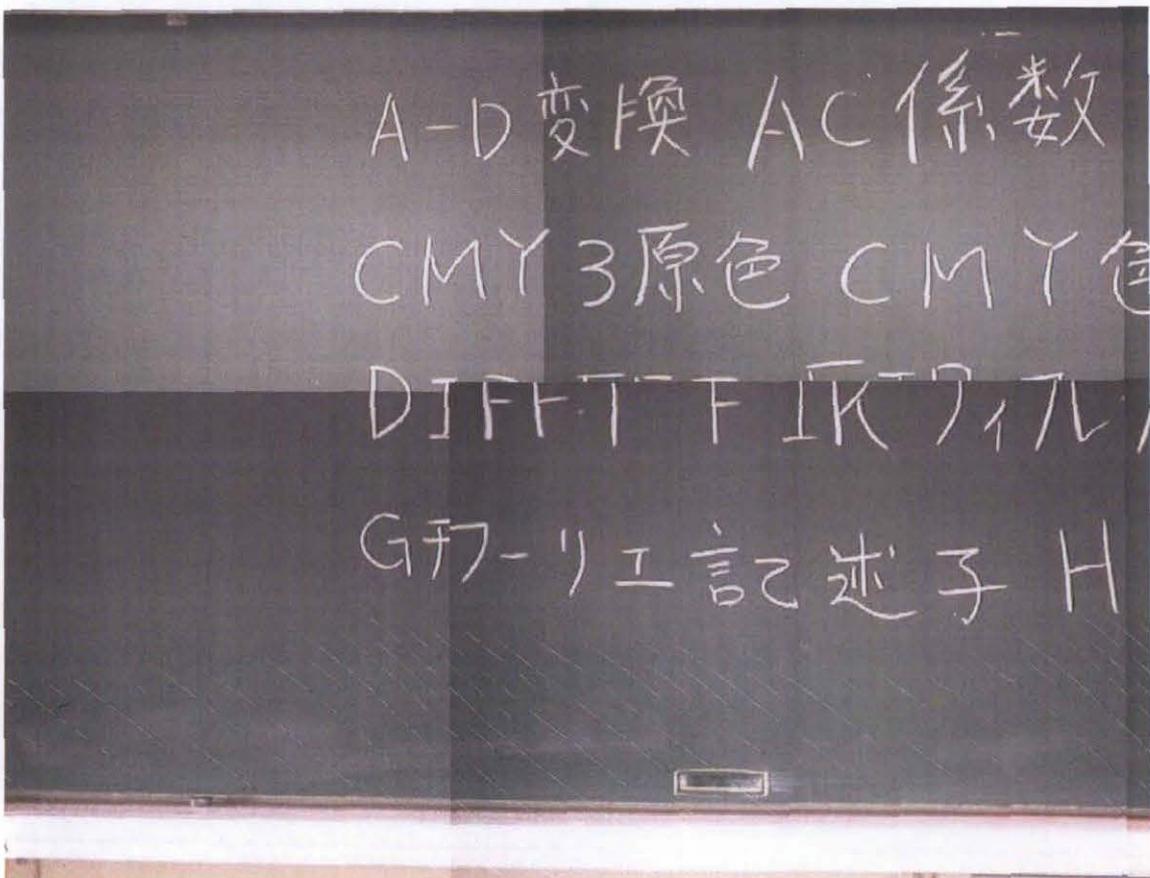


(b)拡大画像

図 44 部分板書画像結合成功例

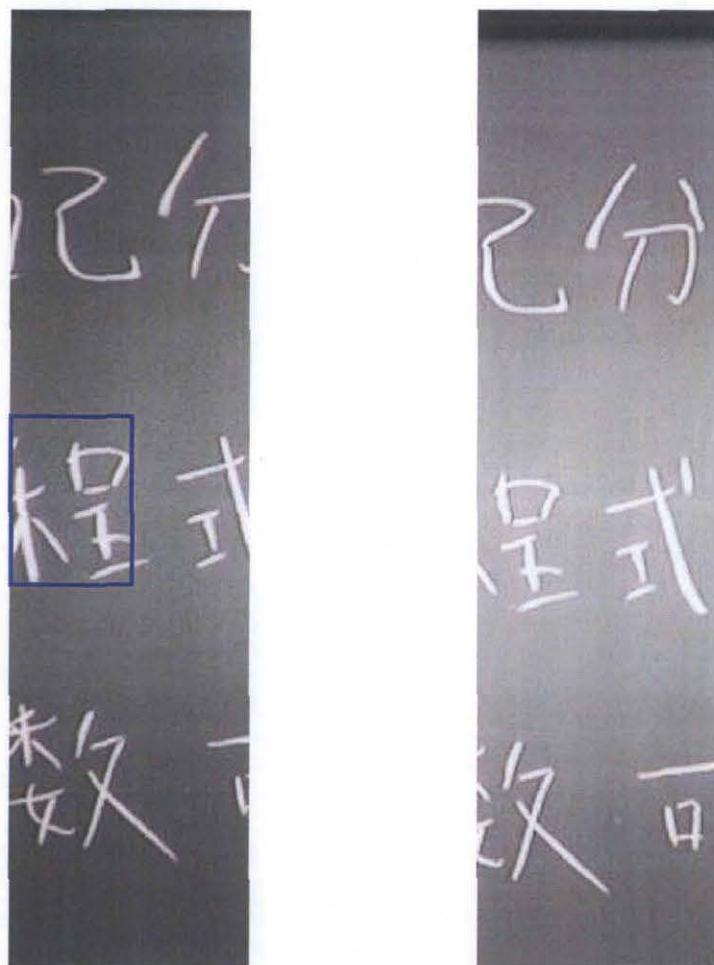


(a)全体画像



(b)拡大画像

図 45 部分板書画像結合失敗例



(a)テンプレート候補領域

(b)探索範囲候補領域

図 46 存在しない領域を含むテンプレート

5.2.2 比較手法の実験結果と考察

ここでは提案手法の有用性を議論するべく、他の部分画像結合手法との比較を行う。文献[14]で紹介されているイメージモザイクング手法を、提案手法との比較のため用いた。文献の手法では二つの画像間に共通し、かつ出現頻度の低い画素を特徴として用いることにより結合を行っている。5.1節の実験にて対象とした部分黒板画像と同一の画像に対し、文献のアルゴリズムを用いて結合を行った。図47に文献のアルゴリズムにより得られた結合画像を示す。図より、適切に部分板書画像を結合できなかったことがわかる。これは文献[14]のアルゴリズムでは出現頻度の低い濃度値を指標としているため、板書画像には特徴的な画素そのものが少ないこと、隣接画像間に濃度値の差が生じることから、指標とすべき画素を適切に選択出来なかったためと考えられる。本実験の結合成功率は55.7%であり、提案手法を用いた時の値よりも低くなった。

一方、提案手法の結合成功率は80%を超える結果となった。これは特徴的な画素そのものを指標とする比較手法と異なり、特徴的な領域を利用しているためであると考えられる。これにより、濃度値の差の影響は低くなり、反対に特徴的な部分は多く存在する。しかし、本実験における実験試料は少なく、統計的に十分に有意であるか断定できない。そこで、本問題における提案手法の優位性を確認するべく、前節に示した比較手法と提案手法において優位性に差がないと仮説を立て、検定を行った。検定方法にはt検定を用いた。対象データは実際の教室において撮影した56枚における画像結合回数49回の結果を用いた。データ数40, 60の場合のt分布表を表5に示す。検定の結果、検定統計量Tは2.439となり、データ数から有意水準 $\alpha=0.05$ の時の、 $\alpha=0.02$ の場合より上回っていることがわかる。この結果より、仮説は棄却され、本問題における提案手法の比較手法に対する優位性が示された[20]。

しかしながら、本手法も実際の講義中の板書に対する適用は行っておらず、図や表が存在した場合どのような挙動を行うかは未知であるため、実用化までにはまだ多くの問題点が残っていると言える。しかしながら、画像の結合手法は、高解像度画像が必要となるセキュリティ分野における人物の顔画像認識、統計的な生体の生育状況の把握(多数の微生物の変化)、生体組織変化の観察(眼底カメラ画像からの疾患の検出)では、強く要望されており、今後大きな研究分野に成長していくものと考えられる。

表 5 t 分布表

データ数/有意水準	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.709	3.551
60	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460

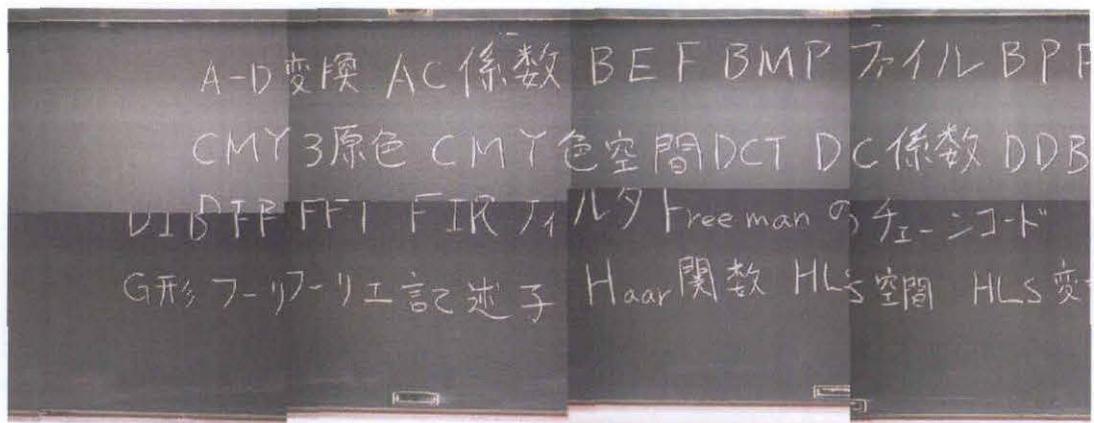


図 47 比較手法における部分板書画像結合例

第6章 おわりに

本論文では、板書データベース作成のため、高解像度に黒板の一部を撮影し、それらを遺伝的アルゴリズムとテンプレートマッチングを用いて結合することにより、高解像度な板書画像を取得する方法を提案し、その有用性について検討した。また、実際の教室にて撮影した板書画像に対し画像結合実験を行ったところ 80%を超える成功率が得られた。今後の課題として、以下の点が挙げられる。

(1)実際の講義での試行と、そのデータに基づく黒板の変化の調査

板書の変化パターンを解析した場合、更に有用な評価関数及び探索法の発見に役立つと考えられる。

(2)テンプレートの評価関数の改善

1文字を最適なテンプレートとして、評価関数を作成したが、1文字丁度を抽出する頻度は低く、実現には更なる評価関数が必要であると考えられる。

(3)マッチング時の相違度算出法、濃度差に対する検討

HSV表色系におけるマッチングは有用であるが、各パラメータがどのような影響を与えるかは不明であるため、それに対する検討が必要である。また、黒板上が非常に明るい場合には白色と黒板部分の濃度値の差が小さくなるため、特徴が弱くなる。その改善のため、濃度値の補正などが有効であると考えられる。

(4)板書画像の撮影方法の検討

板書が存在しない部分、あるいは少量しか含まない部分を撮影した場合、高い精度でテンプレートマッチングを行うことは難しい。この問題の改善のためには板書のある部分のみを撮影するか、黒板1枚あたりの撮影枚数を増やし、隣接する画像間の共通領域を拡大する必要がある。

参考文献

- [1] Josh Bersin著, 赤堀侃司監訳, “ブレンディッド・ラーニングの戦略 eラーニングを活用した人材育成”, 東京電機大学出版局, (2006-3)
- [2] 経済産業省商務情報政策局情報処理振興課編, “eラーニング白書2007/2008年版”, 東京電機大学出版局, (2007-8)
- [3] 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, (2004-11)
- [4] 文部科学省 <http://www.mext.go.jp/>
- [5] 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>
- [6] 和田公人, “失敗から学ぶeラーニング”, オーム社, (2004-5)
- [7] 山口達, 吉川大弘, 篠木剛, 鶴岡信治, “講師の動作認識に基づいた遠隔授業映像の自動撮影”, 電子情報通信学会技術研究報告, PRMU2000-181, pp. 149-156, (2001-1)
- [8] 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研究報告, PRMU2003-213, pp. 89-94, (2004-1)
- [9] 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”, Proceedings of the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp. 102-105, (2003-9)
- [10] 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 6th International Symposium on Advanced Intelligent Systems (ISIS2005), pp. 213-217, (2005-9)
- [11] Daisuke Yoshida, Shinji Tsuruoka, Hiroharu Kawanaka and Tsuyoshi Shinogi, “Keywords Recognition of Handwritten Character String on Whiteboard Using Word Dictionary for e-Learning”, Proceedings of 2006 International Conference on Hybrid Information Technology (ICHIT2006), pp. 140-145, (2006-11)
- [12] 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”, 計測自動制御学会中部支部平成18年度三重地区計測制御研究講演会講演論文集, pp.(P22-1)-(P22-4), (2006-11)
- [13] キヤノン株式会社 <http://cweb.canon.jp/webview/index.html>
- [14] 森明慧, 内田誠一, “Multi-histogram法による画像のモザイクング”, 画像の認識・理解シンポジウム(MIRU2005), pp.289-296, (2005-7)
- [15] 千葉直樹, 蚊野浩, 美濃導彦, 安田昌司, “画像特長に基づくイメージモザイクング”, 電子情報通信学会論文誌, pp.1581-1589, (1999-10)

- [16] 竹村裕夫, CCD カメラ技術入門, コロナ社, p.69, 1997.
- [17] 酒井幸市, デジタル画像処理の基礎と応用, 5章 パターン認識, pp.84-86, CQ出版社(2003)
- [18] ISP imaging-developers システム計画研究所, <http://image-d.isp.jp/index.html>
- [19] 奥富正敏編, デジタル画像処理, 10章 線画像のベクトル化, pp.185-190, CG-ARTS 協会(2006)
- [20] 前野昌弘, 三國彰, 図解でわかる統計解析, 7章 差の検定, pp.160-162, 日本実業出版社(2000)

謝辞

本研究の遂行及び修士論文の作成にあたり，丁寧なご指導とご助言を賜りました本学大学院工学研究科教授鶴岡信治先生，同准教授高瀬治彦先生，同助教川中普晴先生に深謝申し上げます。また，貴重な時間を割いて本論文を査読して頂きました本学大学院工学研究科准教授若林哲史先生に深く感謝致します。

そして，共に学び切磋琢磨し合ってきた情報処理研究室博士前期課程2年の前田祐樹氏，博士前期課程1年の平林邦章氏，服部正寛氏，川村敏仁氏に感謝致します。また，共に画像グループで学んできた情報処理研究室卒業研究生の青木優氏に感謝致します。

最後となりましたが，日頃お世話になりました情報処理研究室の皆様感謝の意を表します。

発表論文リスト

国内会議

- [1] 長縄侑樹, 鶴岡信治, 川中普晴, 篠木剛, “進化的計算を用いた映像eラーニングのためのテンプレート自動作成法に関する一考察“, 日本知能情報ファジィ学会ファジィシステムシンポジウム講演論文集, vol.23, pp.821-826, (2007-8)
- [2] 長縄侑樹, 川中普晴, 鶴岡信治, “映像eラーニングのための遺伝的アルゴリズムを用いた部分板書画像の自動結合方法“, 計測自動制御学会平成19年度SICE三重地区計測制御研究講演会講演論文集, pp.(A11-1)-(A11-6), (2007-12)
- [3] 長縄侑樹, 川中普晴, 鶴岡信治, “映像eラーニングのための遺伝的アルゴリズムを用いた部分板書画像の自動結合方法“, 平成20年度全日本教育工学協議会全国大会・三重大会, p.16, (2008-11)
- [4] 長縄侑樹, 川中普晴, 鶴岡信治, “映像eラーニングのための遺伝的アルゴリズムを用いた部分板書画像の自動結合方法“, 計測自動制御学会平成20年度SICE三重地区計測制御研究講演会講演論文集, pp.(B9-1)-(B9-4), (2008-12)

国際会議

- [5] Yuki NAGANAWA, Hiroharu KAWANAKA and Shinji TSURUOKA, ”A Study on Image Merge Method Using Genetic Algorithm for Image Based e-learning System”, Proc. of IEEE International Conference on Systems, Man, and Cybernetics (SMC2008), pp.1124-1129, (2008-10)