

NAIST-IS-DD0661011

博士論文

形状に基づくアピアランスモデルを用いた
画像理解とその応用に関する研究

草地 良規

2007年 9月 28日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(工学) 授与の要件として提出した博士論文である。

草地 良規

審査委員：

横矢 直和 教授 (主指導教員)

千原 國宏 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

形状に基づくアピアランスモデルを用いた 画像理解とその応用に関する研究*

草地 良規

内容梗概

画像中に撮影された文字，物体，顔などの対象を理解する画像認識技術は，人間の視覚機能の代替および補強技術として幅広い応用が期待できるため，汎用的かつ実用的な技術の確立が望まれている．対象を識別する際に利用可能な情報としては，形状，色，テクスチャなどがあるが，本研究では，照明光や撮像系などの撮影環境に左右されにくい形状を基にしたアピアランスベースのモデルを用いて，実環境において対象を識別する技術を確立する．提案技術を景観中の文字認識による画像検索効率化，画像を検索キーとした直感的インタフェース実現に応用し，その有用性を実証する．まず第1章では，画像認識技術の従来研究と技術課題を概観し，本研究の目的と意義を明確化する．第2章では，形状を用いたアピアランスベースの手法として，利用するパターン特徴量，そのモデル化方法，画像内からの対象検出方法について提案する．第3章では，提案法を景観中の活字文字認識に適用し，画像内容に基づく自動インデクス付与による画像検索に応用する手法について述べ，その実験結果について報告する．また第4章では，提案法を剛物体認識に適用し，カメラ付き携帯電話を用いて対象を撮影することにより対象のコンテンツ情報を入手するインタフェースに応用する手法について述べ，その有効性を確認する．最後に第5章で本研究を総括する．

キーワード

画像理解，アピアランスモデル，形状，文字，剛物体，画像検索，情報取得

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 博士論文, NAIST-IS-DD0661011, 2007年9月28日.

Image Understanding with Shape Based Appearance Model and Its Applications *

Yoshinori Kusachi

Abstract

Image understanding technology, which recognizes characters or objects or faces, has been studied for many years in order to substitute or enhance the human vision using shapes, colors and textures that are useful for recognition. In this thesis, a novel shape based appearance model is proposed for character and rigid object recognition in real environments. This model is robust against the varying imaging conditions such as illuminations and imaging devices. The proposed method is actually applied to image search and information retrieval systems. Chapter 1 gives a perspective of the present study in the area of image understanding. Chapter 2 proposes a new appearance based model, which consists of features, modeling and pattern detecting algorithms. Chapter 3 describes a new method for image search system using index of characters generated by the appearance based model. Proposed in Chapter 4 is a new information retrieval system on cellular phones using a photo as a searching key. Chapter 5 summarizes the present work.

Keywords:

image understanding, appearance model, shape, character, rigid object, image search, information retrieval

*Doctoral Dissertation, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0661011, September 28, 2007.

目次

1. 序論	1
1.1 画像認識技術概観	1
1.1.1 文字認識技術の従来研究とその技術課題	1
1.1.2 剛物体認識技術の従来研究とその技術課題	2
1.1.3 画像認識技術の応用における課題	4
1.2 研究の目的と意義	6
1.2.1 形状を基にしたアピアランスベースの認識法	6
1.2.2 景観中文字の認識と画像検索への応用	7
1.2.3 剛物体の認識と画像撮影による情報取得手段への応用	8
1.3 本論文の構成	11
2. 形状を基にしたアピアランスモデル	12
2.1 はじめに	12
2.2 問題設定と課題解決方針	12
2.3 利用する特徴ベクトル	13
2.3.1 縦横微分画像	14
2.3.2 方向のみ微分画像	15
2.3.3 加重方向指数ヒストグラム特徴	16
2.4 パターン学習	17
2.4.1 特徴抽出とカテゴリ階層構造	17
2.4.2 幾何学変形によるパターン生成	18
2.4.3 モデル生成	19
2.5 部分空間法の拡張	20
2.5.1 不定次元の導入	20
2.5.2 不定次元を有する特徴ベクトル群の共分散行列	21
2.5.3 特徴ベクトルのノルム正規化への対応	22
2.6 パターン認識	23
2.7 むすびに	27

3. 景観中文字の認識と画像検索への応用	29
3.1 はじめに	29
3.2 問題設定と課題解決方針	29
3.3 画像検索	30
3.4 実験結果	32
3.4.1 漢字画像の識別実験	35
3.4.2 対象画像からのインデクス抽出実験	40
3.4.3 検索実験	42
3.5 むすびに	44
4. 剛物体の認識と画像撮影による情報取得手段への応用	50
4.1 はじめに	50
4.2 問題設定と方針	50
4.3 シミュレーション実験	52
4.4 平面デザインの認識実験	55
4.4.1 ちらしの認識	55
4.4.2 絵葉書の認識	59
4.4.3 ロゴの認識	60
4.5 立体物の認識実験	63
4.5.1 ペットボトルの認識	63
4.5.2 フィギュアの認識	64
4.6 むすびに	76
5. 結論	77
5.1 本論文のまとめ	77
5.2 今後の検討課題	78
謝辞	81
参考文献	82

目 次

1	従来法では識別困難な景観画像中の文字例	9
2	画像撮影による情報取得手段のサービスイメージ図	10
3	画像撮影による情報取得手段への応用において扱う課題	10
4	提案法の概要	14
5	縦横微分画像例	15
6	階層構造の例	18
7	部分パターンの例	20
8	ノルム正規化における雑音の影響	24
9	正規化係数を複数考慮した分布	25
10	ピーク検出の例	26
11	提案アルゴリズムおよび画像検索アプリケーションの概要	30
12	パラメータ算出過程	32
13	「銀」の文字例	34
14	非階層化モデルによる累積分類率	36
15	階層化モデルによる累積分類率	37
16	1位として識別された漢字画像例	37
17	識別不能例（フォント）	38
18	識別不能例（変形）	38
19	識別不能例（つぶれ）	39
20	識別不能例（背景パターン）	39
21	シェーディングを伴う文字列の検索成功例：検索語は「平和通り」	45
22	立体文字列の検索成功例：検索語は上から「東郷神社」,「飯田ビル」	46
23	複雑背景上の文字列の検索成功例：検索語は上から「台東区役所」,「健康食品」	47
24	傾斜文字列の検索成功例：検索語は「豊島区」	48
25	文字列の一部の誤検出による検索失敗例：検索語は「出入口」	48
26	背景パターンからの誤検出による検索失敗例：検索語は「出入口」	49

27	システム概要図	51
28	オブジェクトの学習画像	52
29	マスク画像	53
30	テストパターン	54
31	撮影ポイント	56
32	実験結果	58
33	撮影画像例（ちらし）	59
34	撮影画像例（絵葉書）	60
35	撮影画像例（ロゴ）	62
36	撮影画像例（ペットボトル）	64
37	認識対象9種類	66
38	学習画像例	67
39	背景パターン	67
40	無背景（Obj2）	68
41	デスク（Obj2）	69
42	カレンダー（Obj2）	70
43	本棚（Obj2）	71
44	廊下（Obj2）	72
45	識別結果	73
46	1枚あたりの処理時間	73

表目次

1	アルゴリズムの選択肢	28
2	非階層化モデル生成のパラメータ	33
3	階層化モデル生成のパラメータ	33
4	文字サンプル数	35
5	漢字画像識別実験のパラメータ	35
6	各地域の文字再現率	41

7	各地域の文字適合率	41
8	インデクス抽出に関するパラメータ	42
9	各地域の画像検索再現 / 適合率	44
10	検索に関するパラメータ	44
11	姿勢識別結果	54
12	適用アルゴリズム	57
13	モデル作成パラメータ (ちらし)	57
14	モデル作成パラメータ (絵葉書)	61
15	モデル作成パラメータ (ロゴ)	62
16	モデル作成パラメータ (ペットボトル)	63
17	モデル作成パラメータ (フィギュア)	65
18	対象毎識別率	74
19	角度毎識別率	75
20	環境毎識別率	75

1. 序論

1.1 画像認識技術概観

今日、様々なセンサデータを画像で表現している。カメラ、ビデオカメラ、レンジセンサ、スキャナ、レントゲン、などが代表的なセンサである。これらのセンサはネットワークの広帯域化に伴って多く配備され、また、ハードディスクの大容量化に伴って画像データも飛躍的に増大している。人間はそのデータを見るとさまざまなことを理解できるが、逐一目視することは多大な労力を要するため、計算機が自動でその内容を理解し、人間の作業を効率化する技術、即ち人間の視覚機能の代替/補強技術、が求められている。

計算機に自動で画像の意味内容を把握させることは、情報科学の研究において最も困難な課題の一つである。困難とする要因は、環境、対象、混入雑音のバリエーションが大きく、これらの完全なモデルを作成することが困難であることが挙げられる。この困難を回避するために、各分野、各課題で制約条件を設定し実用化を実現するのが一般的である。例えばスキャナによる入力を前提としたOCR[梅田 94b]、固定カメラ設置を前提としたナンバープレート認識[高橋 04]、撮影環境を制御した顔認証[土橋 02]などが良い例である。

本論文では、その代表的な1分野と考えられる、カメラおよびビデオカメラにより撮影された画像中の活字文字、平面パターン、3次元剛物体を認識する課題を扱う。以下では、文字認識/剛物体認識技術の従来研究の流れを説明し、現状の技術課題を整理する。次に本研究の位置づけについてまとめる。

1.1.1 文字認識技術の従来研究とその技術課題

文字認識は、1960年代後半に郵便番号の読み取り(手書き数字認識)が実用化され、1970年代になると手書きの英数字にカタカナまで読めるOCR(Optical Character Recognition:光学式文字認識)が出現した。1980年代には字種の拡大、字体の自由化、書式の自由化、を課題として活発に印刷漢字認識、手書き漢字認識の研究が行われた[津雲 88, 梅田 94a]。印刷漢字認識に対してはあらかじめ用意したパターンと入力データをマッチングするパターン整合法がよく用いられ、大きな成

果を挙げた [目黒 84]。手書き漢字認識に対しては、文字のストロークを抽出してモデルと比較する構造解析法がよく用いられた [山本 82, 山田 84, 高橋 84, 田中 86]。しかし、構造解析方法は雑音に弱く安定したストロークを抽出できないという課題があり、パターン整合法が主流となった。

パターン整合法を用いた OCR では、文書構造解析、文字位置抽出、2 値化、正規化、特徴抽出、パターンマッチング、後処理のモジュールから構成される。この内、文字認識処理を行うのは 2 値化～パターンマッチングであり、その課題は、字種が多く計算時間がかかること、あらかじめパターンを全部用意することが困難であること、であった。前者は階層的な分類 [山本 73] により解決された。この方法では、全字種 (3000 字種程度) に対して計算コストの低い処理で 10～100 字種に候補を絞り、絞られた候補に対して詳細な識別処理を行う。後者の課題は、図形特徴 [岡 82, 岡 83]、統計的特徴 [鶴岡 87] の大きく 2 つの特徴抽出方法が研究開発され、印刷漢字認識に関しては実用化課題が解決された。

印刷漢字認識の研究より得られた知見は多いが、本論文に関係するものとしては、

- 統計処理を用いたパターン整合法の有用性
- 方向性特徴の有用性
- 階層構造の有用性

である。特に、統計処理の基本的な考え方は、アピアランススペースの方法として顔認識、物体認識に多大な影響を与えた [Tur91, 村瀬 94]。現在では、印刷漢字認識においてはデジタルカメラによる入力を前提とした低解像度文字への対応や撮影環境の条件緩和を課題として活発に研究されている [森 01, 黄瀬 04]。

1.1.2 剛物体認識技術の従来研究とその技術課題

物体の画像認識は、1960 年代にその萌芽を見ることができる。しかし、当初より課題として困難であることは認識されており、制約条件として撮影環境の限定が導入された「積木の世界」にて研究が行われていた。画像より線を抽出し、そ

の構造，位置関係によって対象を識別するアプローチが一般的であった [Clo71] . しかし，実世界に適用すると，前提となる撮影環境の限定ができず，安定動作は困難であった．実画像より線を安定的に抽出することは現在も解決されていない課題である．

次に，2次元的に扱うことのできる画像を構造化する研究が行われた．画像より領域を抽出し，形状，色，模様などの統計量，およびその位置関係によって対象を識別するアプローチが一般的であった [Ten77] . しかし，実世界に適用すると，前提となる撮影環境の限定ができず，安定動作は困難であった．実画像より領域を安定的に抽出することは現在も解決されていない課題である．また，対象毎に人手でモデル構造を与える知識ベースの方法 [Oht85, Dra89] も試みられたが，認識対象を増やすことが困難であった．

次に，3次元の情報復元を用いたモデルベースの物体認識の研究が盛んに行われた [Pop94] . モデルベースの物体認識においては，認識対象の3次元形状モデルをあらかじめ用意しておき，画像とモデルの照合を行うことでモデルパラメータと対象を識別するアプローチが取られた．モデルは簡単な幾何形状を基本要素とし，その位置関係，接続関係をパラメータにより構成されるのが一般的であった．これらの方法の課題としては，パラメータが多くなると識別に時間がかかる，3次元の幾何モデルは情報獲得に手間がかかる，精緻な幾何モデルは用意困難であり誤差を伴う，が挙げられる．

上記の限界を超えるために，文字認識の統計処理技術を応用したアピアランスベースの方法が提案された．アピアランスベースの方法では，画像情報をベクトルとみなして複数の学習画像情報に対して統計処理を行いモデルを生成する．この画像認識のアプローチは大きく分けて全体的なパターンを用いる方法 [Tur91, 村瀬 94, 村瀬 95, Kus04, 坂上 03b, 坂上 03a, 安藤 06] と局所的なパターンを用いる方法 [Vio01, Cha99, 山口 02, 大場 97, Sch98] の2方式が研究された．

全体的なパターンを用いた物体識別法としては，画像を1ベクトルと見なしたパターンの分布を主成分で近似する固有空間法 [Tur91] や固有空間上でパターンを非線形関数により補間するパラメトリック固有空間法 [村瀬 94] が提案されている．しかし，学習画像より背景領域を除いて学習する方法 [村瀬 95] がある

が、隠蔽や全反射には対応できていないという課題を有していた。また、雑音領域の特定と部分的な対象領域による識別を同時に行うことで雑音に頑健な識別方法 [坂上 03b, 坂上 03a] があるが、繰り返し演算を伴うため計算量が多い、ぼけやシェーディングなどの全体的なパターン変化には対応できないなどの課題を有していた。

一方、局所的なパターンを用いる方法により、隠蔽の課題を解決することが提唱された [Cha99, 山口 02, 大場 97, Sch98]。大場らは物体の局所的なパターンの変動を主成分分析により吸収し、その位置関係を利用してモデル化を行い、識別を行っている [大場 97]。特徴欠損/雑音付加には頑健になるが、全体的なパターンと比べてモデルを構成する情報が少なく、対象が多くなる場合は、対象を区別するために必要な情報が十分に記述されない可能性がある点が課題であった。

その後、SVM [Vap95]、ADABOOST などのパターン認識技術を用いた物体認識法が提案された。Viola らは単純な特徴による弱識別器を組み合わせることで精度の高い物体認識が可能であることを示した [Vio01]。これらは、対象、非対象の実サンプルを用いた統計学習により、さまざまな切り出し領域から高速に対象を抽出するアプローチである [Che04, Zhu05] が、対象と非対象の判定境界は実学習サンプルの分布に影響を受けるため、サンプルを十分に収集し吟味する必要がある。学習サンプルが十分に集められる分野には有効な手法であるが、そのような分野は限られる。

1.1.3 画像認識技術の応用における課題

画像認識を実環境に適用困難とする要因は、計測データ（パターン）の変動要素（パラメータ）が多いため、一般的にパラメータの組み合わせ空間は指数関数的に大きくなり、汎用的な解法（全探索など）は計算コストが高い、ことである。また、精緻なモデル化はモデル化誤差に伴う弊害が大きくなることが統計的に示されており [Aka73]、安易なパラメータ導入は控える必要がある。さらに、パラメータを特定できない場合（例えば服の形状変化など）も多く、汎用的な解法を適用できない。

そのため、応用の際には計算時間の制約を考慮して、少ないパラメータでパ

ターンの変動をモデル化し，利用シーンから合理的な制約条件を導入して判別することが現実的なアプローチである．

以下に筆者の研究を例にあげる．[草地 98] では，人物領域を抽出し，歩行動作によるパターン変動を 9 個のパラメータで記述した．また，パターンによって人物領域の変化の独立性が高いこと，脚部の歩行動作を関数近似できること，時系列データではパラメータが滑らかに変化すること，の 3 つの制約により，現実的な時間で歩行動作を解析し，キャラクタの歩行映像を合成した．[草地 00] では，ワーク形状の構造変化・隙間の生成/消滅・障害物による変動を 10 程度のパラメータで記述した．また，時系列データではパラメータが滑らかに変化するという制約を用いて，実用時間内で溶接点を認識し，質の高い溶接作業が実現できることを示した．

画像認識で強く開発が要望されているアプリケーションを 2 種類に大別し，それぞれに対して制約条件の要件を述べる．

- 大量画像データベースからのインデクシングによる受動検索

テロップ認識 [Lie96, 新井 00]，Web コンテンツの解析 [Lop00] 等が代表例である．検索の需要は撮影後に生じるため，処理の高速性の要求度は高くない．インデクスがシンボル（キーワード）化されているため，キーワード入力による検索に制約を設けることができる．この制約条件は，例えば文字認識によるインデクシングにおいては，文字の配置関係など例外はあるがほぼ一般的と考えられる規則を用いる必要がある．

- 画像をキーにした対象物の情報能動検索

実世界へのリンク付与 [Har01] やカメラ画像による英単語の翻訳 [メデ 06] 等が代表例である．検索の需要は撮影前に生じるため，処理の高速化が必要である反面，撮影条件に制約を設けることができる．この制約条件は直感的でわかりやすく，ユーザの適応が容易である必要がある．

1.2 研究の目的と意義

本研究では活字文字および剛物体（平面パターンを含む）を対象とし，認識対象数が多く，様々な撮影環境に頑健なアピアランスベースの認識技術を確立することを第一の目的とする．また，提案技術を用いて景観画像内に存在する文字を自動でインデクスとして付与し，画像検索を効率化することを第二の目的とする．さらに，実世界の物体を多数認識する提案技術を利用して，対象のデザイン性を損なわずに，ユーザが直感的に携帯から対象のコンテンツ情報にアクセスするインタフェースを実現することを第三の目的とする．

第一の目的の意義は，対象および環境の自由度の高い画像認識技術が望まれるが，考慮すべきパラメータが多く，課題解決は困難であり，実際の需要が高いと考えられる環境の自由度を重視して画像認識技術の課題を解決し，技術成果を早期に社会に還元することにある．第二，第三の目的の意義は，その達成により，画像認識に期待される代表的なアプリケーションにおける提案法の実用性を実証することである．

1.2.1 形状を基にしたアピアランスベースの認識法

活字文字および剛物体（平面パターンを含む）を対象とし，屋内外の照明光，カメラ撮像系，撮影者，視点変動に対処する課題を扱うために，形状を基にしたアピアランスベースの認識方法を提案する．

屋内外の照明光，カメラ撮像系，撮影者，視点変動，背景パターンなどによって剛物体，文字の見え方は変化する．また，フォントによって文字の見え方は変化する．これらの変化を克服して剛物体，文字を認識する必要があるが，以下3つの観点から要件を述べる．

- 特徴量

認識に用いる特徴量の基本要素としては色，形，テクスチャがある．文字の場合は色やテクスチャは認識にとって有用な情報となり得ない．また，数少ない剛物体であれば色やテクスチャで判別できる可能性があるが，多数の剛物体に適用する場合，色やテクスチャでは特定性が乏しく認識が困難

となる。数多くの対象を扱う特徴量の検討においては、対象画像空間と特徴量空間の写像関係が重要となるが、中でもアピアランスベースの特徴量は、その関係を制御し易いため、数多くの物体を扱うことが容易であるという利点がある。そのため、形状を基本としたアピアランスベースの特徴量を用いる必要がある。その際、1章で述べたように、方向性特徴が有用であることがわかっている。

- パターン認識

背景パターンが変動するため、できるだけ背景パターンを除いた信号パターンとモデルパターンを比較する必要がある。そのため、画像中の位置、大きさを変えながら小領域を切り出し、モデルパターンと照合する必要がある。

- パターン収集/学習

多くの対象パターンを収集し、各対象をモデル化する必要がある。収集方法には大別して以下の2つの方法がある。

- － 実画像よりパターンを収集する方法。
- － 定型の基本パターンに変形等を加え、複数のサンプルを生成する方法（生成型学習 [村瀬 05]）。

高い自由度を有する実環境下で前者を用いるのは困難であり、生成型学習が必要である。また、モデル化方法としては、各カテゴリの特徴空間中の範囲を定義する方法と、各カテゴリ間の境界領域を定義する方法がある。パターン認識を行う際に切り出した小領域には非対象の場合があり、各カテゴリの特徴空間中の範囲を定義する方法が望ましい。

1.2.2 景観中文字の認識と画像検索への応用

景観画像内に存在する活字文字を自動でインデクスとして付与し、画像検索を効率化する課題を扱う。デジタルカメラの普及に伴い、ユーザはより多くの画像を撮影し、保存するようになった。しかし、その検索に必要なインデクスは、機

器によって自動的に付与される時刻等の情報だけでは不十分である．画像内容に基づく自動インデクス付与は画像検索に有用である．例えば，史跡の説明看板，展示会等での説明資料をインデクシングすることで備忘録として利用できる．また，街中の画像を撮影してインデクシングすることでキーワードによって街の地図画像を参照することができる．

対象の撮影条件（特にコントラスト条件）を設けることで，文字列領域を切り出し，文字位置を推定して文字を認識する研究が多くある [Ohy94, Yan01, Sat99, Har01, Nag00, Lu06] が，撮影条件の緩和が課題となっている．また，文字，非文字の実サンプルを用いた統計学習により，さまざまな切り出し領域を高速に文字/非文字を判定する研究がある [Che04, Zhu05] が，その判定境界は実サンプルの分布に影響を受けるため，サンプルを十分に収集し吟味する必要がある．文字領域の抽出，文字位置の特定の成功後は，識別に濃淡特徴を用いることで照明条件の変化を克服することができる [Yos00, Zah02, Wak01, 高橋 04] ．

これらの処理性能定量化のために，文字領域抽出，文字認識，単語認識，文章認識，と課題を整理し，同一のデータベースに対して評価する試みがなされており，文字領域抽出においては様々な撮影条件の下，その有効性が示されつつある [Luc03]．しかし，景観画像中の文字に多く含まれる，図 1(a) シェーディングを伴う文字，(b) 立体文字，(c) 複雑背景上の文字，(d) 切り出しの困難な傾斜文字，を扱い，単語認識，文章認識まで成功させることは困難であった．

本研究では，上記 (a) ~ (d) を扱うことのできる画像インデクシングおよび検索法を提案する．インデクシングには背景中の偽文字候補が多数出現してしまうという課題と，影や隠蔽により抽出できない文字があるという課題がある．提案法では，検索文字列の空間的な配置関係を利用することにより，多くの偽文字候補の出現や文字候補の検出もれを許容できる．

1.2.3 剛物体の認識と画像撮影による情報取得手段への応用

携帯電話の普及に伴い，携帯から情報にアクセスするインタフェースを改善したいという需要が高まっている．そのため，一般的となったカメラ付携帯電話によって，対象を撮影することにより対象のコンテンツ情報をインターネットより

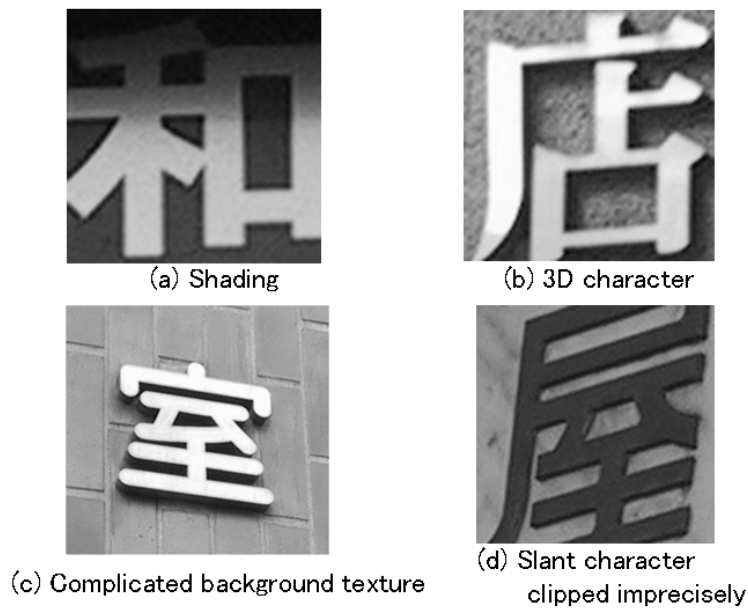


図 1 従来法では識別困難な景観画像中の文字例

入手する技術が注目されている．主な方式としてはQRコード[デン]や電子透かし[中村04]等があげられるが，対象に加工が必要である点，デザイン性を損なう点が課題であった．画像認識技術を用いて，実世界の物体を多数認識することでその課題を解決できる．そのサービスイメージを図2に示す．

本研究では，2章で提案する技術を用いて，カメラ付き携帯電話で対象を撮影する場合に課題となる，ボケ，照明全反射，照明光の多様性，影，撮影角度，撮影距離，複雑背景，隠蔽(図3)を扱うことのできる画像撮影による情報取得手段を提案する．情報取得手段には処理高速化の課題があるが，提案法では，ユーザが対象を中心に全体が画像内におさまるように撮影することを前提とすることで，実用化可能な処理性能を実現する．



図 2 画像撮影による情報取得手段のサービスイメージ図



図 3 画像撮影による情報取得手段への応用において扱う課題

1.3 本論文の構成

以降の本論文の構成は、以下のとおりである。第2章では、照明光や撮像系などの撮影環境に左右されにくい形状を用いたアピアランススペースの基本手法を述べ、利用するパターン特徴量、そのモデル化方法、画像内からの対象検出方法について提案する。第3章では、第2章で述べる方法を景観中の活字文字に適用し、画像内容に基づく自動インデクス付与による画像検索に応用する手法について述べ、その実験結果について報告する。また、第4章では、第2章で述べる方法を一般の2次元パターンおよび剛物体に適用し、カメラ付き携帯電話を用いて対象を撮影することにより情報を入手する情報取得手段への応用について述べ、その有効性を確認する。最後に第5章で本研究を総括する。

2. 形状を基にしたアピアランスモデル

2.1 はじめに

本章では、活字文字および剛物体（平面パターンを含む）を対象とし、認識対象数が多く、様々な撮影環境に頑健なアピアランスベースの認識技術を確立する課題を論じる。アピアランスベースの方法はこれまで数多く研究されているが、屋内外の照明光、カメラ撮像系、撮影者、視点変動、背景に起因するパターン変化を同時に克服する手法はまだ確立されていない。

以下では、2.2 節において、本章で扱う問題の設定とその解決方法について述べ、2.3 節において対象の性質に応じた3つの特徴量を提案する。2.4 節においてパターンの学習方法を説明し、2.5 節において従来法である部分空間法の拡張である、不定次元を含んだ部分空間法を提案する。2.6 節において、そのモデルを利用し画像中から対象を抽出する方法を述べる。最後に、2.7 節において本章をまとめる。

2.2 問題設定と課題解決方針

複数ユーザが多様な環境下で撮影する際に発生する

- (1) 撮影位置，方向
- (2) 照明変動，ぼけ，影，照明反射
- (3) 背景
- (4) 独自のフォント，厚み
- (5) 隠蔽

によるパターン変化を克服することを問題として設定する。(4) は文字特有の課題，(5) は主に剛物体における課題である。

提案法では、生成型学習 [村瀬 05] の考えに基づいた部分空間法 [Oja83] により (1) に、濃淡画像の低周波と高周波を除去した特徴ベクトルを利用することで (2)

に対処する。また、パターン全体を用いるのではなく、パターンの一部である部分パターンを利用して観測窓中の背景領域を減少させ、識別の際にはあらかじめ用意したモデルを用いてサイズや位置を変えながらスキヤニングを行うことで(3)に対処する。また、スキヤニングにおいては、粗探索の候補結果から、各カテゴリで探索範囲を限定することにより高速化を実現する。提案法は、正面および複数視点から撮影した画像を用いて、人工的に幾何変形を行った多数のサンプルを学習してモデルを作成するため、対象外との境界領域は実サンプルの影響を受けにくく、対象依存性は少ない。

また(2)の特徴をフォントによる形状、立体文字における厚み部位による微小変動を吸収できるように改良した特徴量を導入することで(4)に対処して識別可能な文字フォントを拡大して汎用性を高める。

さらに、部分空間法を拡張し、特徴欠損/雑音付加の伴うパターンを学習可能とすることで(5)に対処する。一般には特徴欠損/雑音付加の伴うパターンは指数関数的に多くなるが、特徴欠損/雑音付加の伴う画像を生成せずに、学習画像のみを用いてモデル化を行うことができる。

提案法の流れを図4に示す。提案法は学習系と識別系から構成される。学習系においては、対象の画像(学習画像)を用意し、特徴ベクトルを抽出して学習パターンとし、主成分分析によりモデルを作成する。識別系においては、画像を走査しながら一定の大きさで領域(観測画像)を切り出し、特徴ベクトル(観測パターン)を抽出してモデルと比較することで識別結果を得る(部分空間法)。

2.3 利用する特徴ベクトル

本節では、対象の特性に応じた特徴ベクトルについて述べる。一般の2次元平面に対しては縦横微分画像、背景変動を伴う2次元平面および剛物体に対しては方向のみ微分画像、文字に対しては、加重方向指数ヒストグラム特徴を改良した特徴ベクトルを用いる。

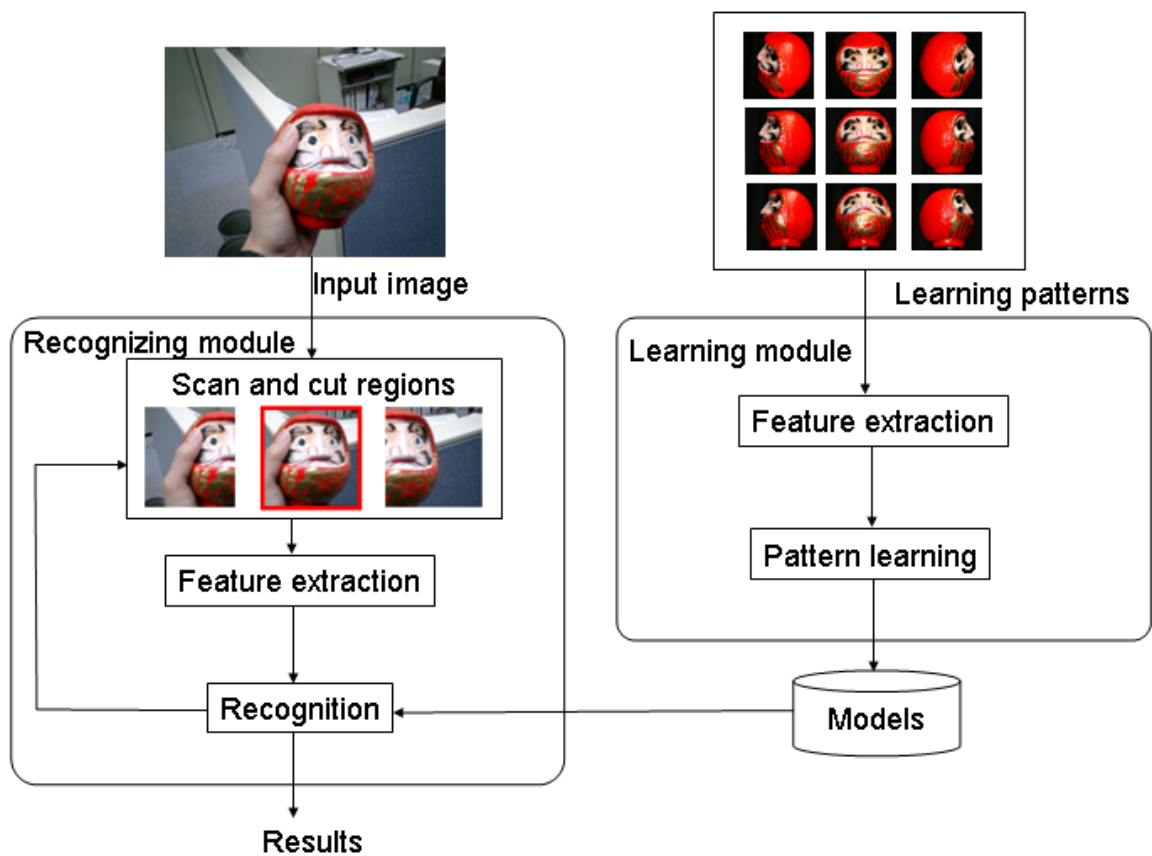


図 4 提案法の概要

2.3.1 縦横微分画像

観測パターンにおいて照明変動，シェーディングによるパターン変化は低周波に，ぼけによるパターン変化は高周波に偏在する傾向がある．また，走査と部分空間法では吸収しきれない

- 微小な撮影位置と方向の変動
- 微小な位置ずれ
- 画像圧縮

によるパターン変化は，高周波に偏在する傾向がある．

上記の考えから，濃淡画像の低周波と高周波を除去した特徴ベクトルを利用し，雑音，観測誤差の影響を低減する [高橋 04]．具体的には，学習画像をぼかした後，縦方向のエッジ抽出の結果画像と横方向のエッジ抽出の結果画像を接続することにより作成する（縦横微分画像，図 5）．例えば画像のぼかしにはガウシアンフィルタ，エッジ抽出にはソーベルオペレータを適用することで作成することができる．縦横微分画像は，エッジの強度とともに，エッジの角度情報を有する特徴ベクトルである．また，色情報を保持しておらず，カメラ付携帯電話の機種依存性は少ない．

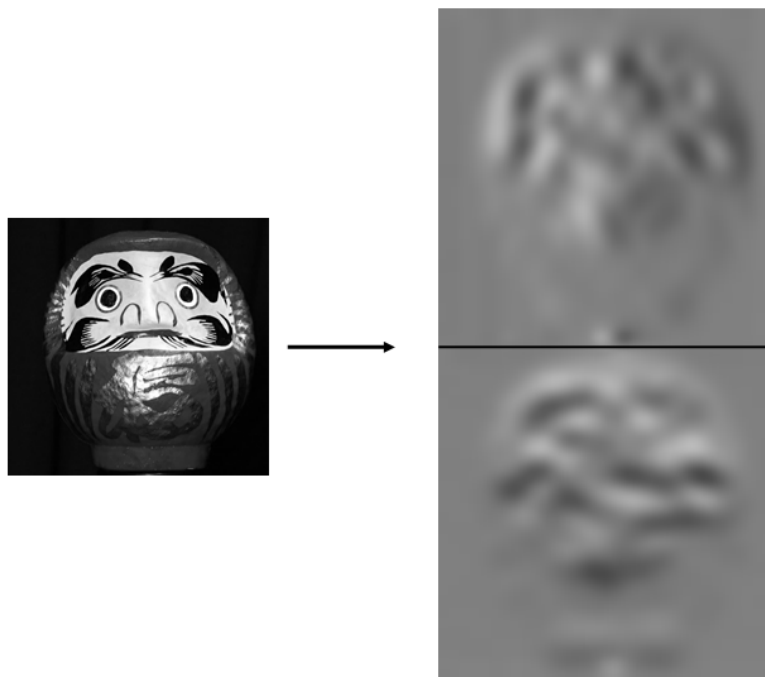


図 5 縦横微分画像例

2.3.2 方向のみ微分画像

2.3.1 で述べた縦横微分画像は，一般的な 2 次元平面の場合には適用可能であるが，背景色が異なる 2 次元平面や 3 次元物体に適用することは困難である． 3

次元物体の形状は輪郭により構成される場合が多いが、背景が異なると輪郭部のエッジは大きさ、方向共に変化してしまう。大きさについては背景の濃淡値によって様々な値を取り得てしまうため、不変量にはならない。方向については、変化しない、もしくは180度回転を伴うため、学習サンプルに複数の背景を有するものを収集することにより、背景の濃淡値に影響を受けない不変量となり得る。即ち、次元 d の縦横微分値を $(D_x(d), D_y(d))$ とすると、

$$(N_x(d), N_y(d)) = \frac{(D_x(d), D_y(d))}{\sqrt{D_x(d)^2 + D_y(d)^2}} \quad (1)$$

として方向のみ微分画像と定義する。ただし Q を定数として、

$$\frac{\sqrt{D_x(d)^2 + D_y(d)^2}}{\sqrt{\sum_d (D_x(d)^2 + D_y(d)^2)}} \leq Q \quad (2)$$

が成立する場合は、 $(N_x(d), N_y(d)) = 0$ として微小な雑音の影響を排除する。

2.3.3 加重方向指数ヒストグラム特徴

活字文字においては、フォントや厚みによる形状の微小変動が伴うため、加重方向指数ヒストグラム特徴を改良した特徴ベクトルを利用する。特徴抽出においては、文字を正面より撮影した原パターン（サイズ $W \times W$ ）を用意し、加重方向指数ヒストグラム特徴（WDCH）を求めてノルムの大きさを正規化する [Kim97]。WDCH は、OCR 用に開発され、2 値画像を対象としていたが、グレースケール画像に容易に拡張可能である。以下にそのアルゴリズムの概要を示す。ただし、 M, N は正定数である。

- < 1. エッジ抽出 > 原パターンから、微分オペレータを用いてエッジの大きさおよび方向を求め、エッジの方向を M 方向に量子化する。
- < 2. ヒストグラム作成 > 原パターンを $N \times N$ のグリッド（領域）に分割し、各グリッドの各方向で、エッジの大きさを加算する。
- < 3. ノルム正規化 > $N \times N \times M$ の特徴ベクトルと考え、ノルムの大きさを K に正規化する。

正規化された WDCH は、微分値をベースとしているため、明るさ変動の影響を受けにくい。また、グリッド内の微分値の和を用いることにより、フォントによる形状の微小変動を吸収することができる。同時に、立体文字における厚み部位によって付加されるエッジは、文字パターンの近辺に位置し、かつ方向が一致する傾向にあり、吸収することができる。

2.4 パターン学習

パターン学習においては原パターンを収集し、以下の4ステップを通してモデルが作成される。

- (1) 特徴抽出
- (2) カテゴリの階層構造の作成
- (3) 幾何学変形によるパターン生成
- (4) モデル作成

以下、順に文字を例として説明する。

2.4.1 特徴抽出とカテゴリ階層構造

特徴抽出においては、文字を正面より撮影した原パターン（サイズ $W \times W$ ）を用意し、2.3 節で述べた3つの形状特徴の内の1つを求めてノルムの大きさを正規化する [Kim97]。

カテゴリの階層構造の作成においては、上記特徴ベクトルの類似性から各カテゴリをクラスタリングし、図6に示すように階層構造（本図では2階層）を作成する。各ノードは複数のカテゴリが含まれる。最下層のノードでは、単体のカテゴリのみが含まれる。

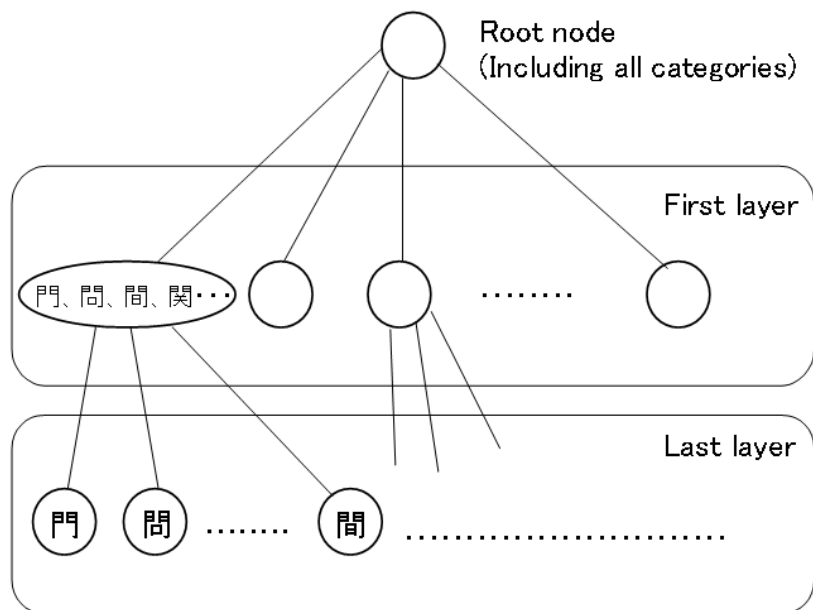


図 6 階層構造の例

2.4.2 幾何学変形によるパターン生成

幾何学変形によるパターン生成では、各カテゴリに対し、視点の変動に伴う変形パターンをすべて生成する。文字の場合は、複数フォントの文字変形パターンを生成する。例えば、原パターンを回転、垂直スキュー、水平スキュー、縦横比、拡大縮小の5パラメータのアフィン変換により幾何変形する。または、4パラメータの透視投影変換により幾何変形する。対象がカメラから遠くに存在する場合はアフィン変換、近くにある場合は透視投影変換を選択する。

生成されたパターンの大きさは、原パターンよりも大きくなる場合があるが、図7のように原パターンの窓サイズ内の部分パターンのみを用いて特徴を抽出し、その特徴ベクトルをモデル生成に用いる。このような部分パターンは、幾何変形パターンに比べて相対的に背景パターンを要因とするノイズを小さくすることができる。例えば、文字の場合について説明する。WDCHにおける第 q 次元の背

景パターンを要因とするノイズ量は、

$$H_q = \left| \frac{g_q}{\sqrt{G}} - \frac{g_q + b_q}{\sqrt{G + B}} \right| \quad (3)$$

で定式化される。ただし、 g_q は正規化前の第 q 次元の信号、 G は観測窓内の信号の 2 乗和、 b_q は第 q 次元に混入する背景パターンのエッジ量、 B は観測窓内の背景パターンによるエッジの 2 乗和、である。この式から、背景パターンについて以下の 3 つの場合に認識が困難となることがわかる。

- コントラストが大きく、広範囲に存在する。
- コントラストが大きく、局所的に存在する。
- コントラストは小さいが、広範囲に存在する。

部分パターンを利用することにより、文字領域に対して周辺の背景領域の割合が減るため、 b_q および B が小さくなり、 H_q が小さくなる効果が見込める。ただし、部分パターンを利用することにより、似たカテゴリのパターンが近づくという問題が発生する。この問題については、応用の際に回避する手段が必要である。

2.4.3 モデル生成

モデルは部分空間法 (付録) をベースとする。モデル生成においては、まず、各階層において、特徴を圧縮する。部分パターンを含む全特徴ベクトルを主成分分析し、上位の固有値を有する固有ベクトルを用いて圧縮する。この圧縮特徴ベクトルを $f(c, r, \vec{p})$ と表す。ただし、 c はカテゴリ、 r は圧縮率、 \vec{p} は変形パラメータである。

次に、各ノードでモデル作成を行う。各ノードのカテゴリ集合を C とすると、 $f(c', r, \vec{p}), c' \in C$ のベクトル集合を主成分分析し、部分空間 $E_d(C, r)$ を求める。ただし、 $1 \leq d \leq D$ であり、部分空間次元数 D は寄与率により求めるか、システムにより定められる正整数である。

各階層での圧縮率は、下層に向かうに従い低く設定することで、2.6 節で詳細を述べる粗密探索を実現する。上層では、精度は低いが高速度な識別を行って候補を絞り込み、下層では、低速ではあるが高精度な識別を実行する。



Original pattern Transformed pattern Partial pattern

図 7 部分パターンの例

2.5 部分空間法の拡張

2.5.1 不定次元の導入

文字を対象とする場合、隠蔽の発生は多くないが、剛物体を対象とする場合、把持などによりパターンの欠損が生じる場合があり、部分空間法はこのような特徴欠損/雑音付加によって誤識別が発生するという課題がある。

本論文では、特徴ベクトルの値として、任意の値の入力を前提とする「不定値」を“*”として表記する。また、その次元を「不定次元」と呼ぶ。すると、不定次元を含む G 次元の特徴ベクトルは、以下のように表記できる。

$$(v_1, v_2, \dots, *, \dots, v_G) \quad (4)$$

不定次元を用いることで、特徴欠損/雑音付加次元を伴う特徴ベクトルを表現することが可能となる。例えば、1枚の画像中の不定次元数を d 、各画素の値が

0 ~ 255 の濃淡値と考えると，不定次元を含む画像は，おのこの 256^d の画像集合と考えることができる．

2.5.2 不定次元を有する特徴ベクトル群の共分散行列

前節で述べた不定次元を有する特徴ベクトル群から，共分散行列を求めることができる．主成分分析を行い，従来の部分空間法に適用することができる．

G 次元の不定次元を含む特徴ベクトルが Q 個あるとし，

$$\vec{v}_{*q} = (v_{*q1}, v_{*q2}, \dots, v_{*qg}, \dots, v_{*qG}) \quad (5)$$

と定義する ($1 \leq g \leq G, 1 \leq q \leq Q$)．また，各画像の不定次元数を d_q ，不定値 * の最大値を w_{max} ，最小値を w_{min} ，刻み数を $w_{step} - 1$ ，不定次元値の順列を $w_r = w_{min} + \frac{r(w_{max} - w_{min})}{w_{step} - 1}$ ， $r = 0, \dots, w_{step} - 1$ ，不定次元での値 w_r の生起確率を $p(w_r)$ とする．ここで，(16) 式で定義した共分散行列の各行列要素 $c(i, j)$ に着目すると，

$$c(i, j) = \frac{m(i, j) - \frac{s(i)s(j)}{N}}{N - 1} \quad (6)$$

と式変形できる．ただし， \vec{v}_n の i 次元要素を v_{ni} として， $m(i, j) = \sum_{n=1}^N v_{ni}v_{nj}$ ， $s(i) = \sum_{n=1}^N v_{ni}$ である．

不定次元を含む特徴ベクトル群に対して， $N, s(i), m(i, j)$ の項は，以下の式で求めることが可能である．

$$N = \sum_{q=1}^Q w_{step}^{d_q} \quad (7)$$

$$s(i) = \sum_{q=1}^Q s_q(i) \quad (8)$$

$$m(i, j) = \sum_{q=1}^Q m_q(i, j) \quad (9)$$

ただし，

$$s_q(i) = \begin{cases} v_{qi} w_{step}^{d_q} & (v_{qi} \neq *) \\ \sum_{r=0}^{w_{step}-1} w_r p(w_r) w_{step}^{d_q} & (v_{qi} = *) \end{cases}$$

$$m_q(i, j) = \begin{cases} v_{qi} v_{qj} w_{step}^{d_q} & (v_{qi} \neq *, v_{qj} \neq *) \\ \sum_{r=0}^{w_{step}-1} v_{qi} w_r p(w_r) w_{step}^{d_q} & (v_{qi} \neq *, v_{qj} = *) \\ \sum_{r=0}^{w_{step}-1} v_{qj} w_r p(w_r) w_{step}^{d_q} & (v_{qi} = *, v_{qj} \neq *) \\ \sum_{r=0}^{w_{step}-1} \sum_{r'=0}^{w_{step}-1} w_r p(w_r) w_{r'} p(w_{r'}) w_{step}^{d_q} & (v_{qi} = *, v_{qj} = *) \end{cases}$$

2.5.3 特徴ベクトルのノルム正規化への対応

本節では，特徴ベクトルのノルムを正規化する場合に拡張する．画像認識においては，特徴ベクトルのノルム正規化は環境光による明るさ変動を吸収するのに効果的であり，広く用いられる手段である．(10) 式のように学習パターンのノルムを正規化してから部分空間法を適用する．ただし， S を正規化係数とする．

$$\vec{v}'_n = \frac{S}{|\vec{v}_n|} \vec{v}_n \quad (10)$$

しかし，観測パターンに雑音が含まれる場合，正規化が正しく行われないう課題が残る．以下で特徴ベクトルが原画像である場合を例に説明する．前節で述べた提案方法においても，モデル生成の際に，学習パターンのノルムを正規化し，不定次元を付与した特徴ベクトル（図 8-a）と，実際に雑音が混入された対象画像のノルムを正規化した特徴ベクトル（図 8-b）は，包含関係にないため有用な学習が行われていない．そこで，図 8-c のように，モデル生成の際に，特徴ベクトルのノルムを小さく（ $S' < S$ ）正規化し，不定次元を付与する．この特徴

ベクトル (図 8-c) は, 図 8-b の特徴ベクトルと包含関係にあるため, 現実に即した学習が可能である. 具体的には, 図 9 に示すように, 特徴ベクトルの正規化係数を複数考慮した分布を求め, その共分散行列を主成分分析する. 求めたい共分散行列の各要素を $c'(i, j)$ とすると, 式 (6) を拡張し,

$$c'(i, j) = \frac{m'(i, j) - \frac{s'(i)s'(j)}{N}}{N - 1} \quad (11)$$

であり, 共分散行列 Σ'_c を求めることができる. ただし, 式 (7) ~ (9) を拡張し,

$$N = \sum_{q=1}^Q S_{step} w_{step}^{d_q} \quad (12)$$

$$s'(i) = \sum_{q=1}^Q \sum_{t=0}^{S_{step}-1} s_q(i) S_t \quad (13)$$

$$m'(i, j) = \sum_{q=1}^Q \sum_{t=0}^{S_{step}-1} m_q(i, j) S_t^2 \quad (14)$$

である. また, 最小パワー倍率を S_{min} , 最大パワー倍率を S_{max} , 刻み数を S_{step} , 倍率順列を $S_t = S_{min} + \frac{t(S_{max}-S_{min})}{S_{step}-1}$, $t = 0 \sim S_{step} - 1$ とする.

2.6 パターン認識

対象の大きさおよび位置の変動に対処するために, 複数解像度画像全面に位置を動かしながら, 大きさ $W \times W$ の小領域を切り出し, パターン識別を行う. パターン識別では, モデルの階層構造において複数のルートを進みながら, 粗密探索を実行する. 以下にそのアルゴリズムを記載する.

- < 1. 特徴抽出 > 各解像度画像に対し位置を変化させつつ, 領域を切り出して特徴を抽出する. すべての切り出し領域の特徴をあらかじめ算出しておく.
- < 2. 初期化 > 木構造のルートノードを出発点とする.

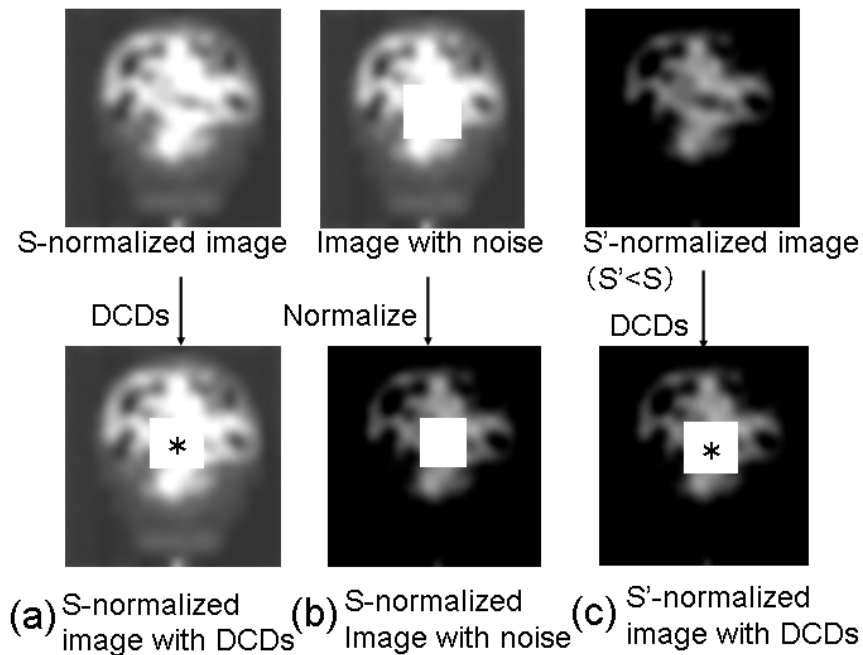


図 8 ノルム正規化における雑音の影響

- < 3. 候補ノードの設定 > すべての切り出し領域に、候補ノードとして第一階層のノードをセットする。各切り出し領域に対して、4 ~ 6 を繰り返す。ただし、切り出し領域にエッジ成分が無い場合には候補ノードをなしとする。
- < 4. 圧縮 > 下層の圧縮率を用いて切り出し領域の特徴を圧縮する。これを $I'(r)$ と表す。
- < 5. 投影距離計算 > 式 15 に従い、候補ノード A の部分空間を用いて投影距離 $L(A)$ を求める。

$$L(A) = |I'(r)|^2 - \sum_{d=1}^D \{I'(r) \cdot E_d(A, r)\}^2 \quad (15)$$

- < 6. スクリーニング > 上記の距離値から、各候補ノードの順位を計算する。その距離閾値 Th_d および順位閾値 Th_r から、候補ノードを更新する。

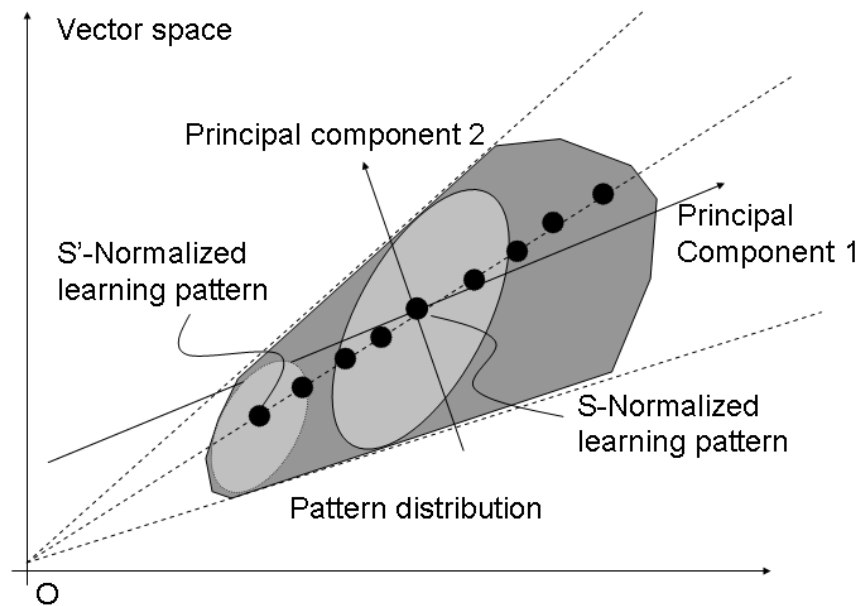


図 9 正規化係数を複数考慮した分布

- < 7. ピーク検出 > すべての切り出し領域の各候補ノードについて，図10に示すように3次元（縦，横，解像度）の空間的な連結性を算出し，連結成分を求める．各連結成分内の距離値の最小ピークを有する候補ノードのみを残す．その他の候補ノードは削除する．空間的な連結性は3次元ではなく2次元（縦，横）等で行ってもよい．このようにピーク検出を行うことにより，電線やビルなどの長いエッジから発生する偽候補の数を抑え，かつ高速化を実現することが出来る．
- < 8. 候補ノードの更新 > 候補ノードの下層に接続されたノードを新しい候補ノードとして登録する．最下層に辿り着いていれば9へ．その他であれば4に戻る．
- < 9. 局所領域でのスクリーニング > 同一解像度の候補ノードの集合各々に対して以下の処理を行なう．まず空間を $P_b \times P_b$ ピクセルのブロックに分割し，

各ブロック内に含まれる候補ノードを距離値によってソートして上位から一定個数 Th_{br} 以内だけ採用する．その他の候補ノードは削除する．その後，ブロック分割の位置を水平または垂直に半ブロックずらして同じ処理を行なう．

< 10. 同一候補ノードの統合 > 同一カテゴリである2つの候補ノードを取り出して3次元の中心座標間の差が閾値 Th_c 以下である場合に，距離値の小さい方に統合する処理を，統合する候補ノードの対が存在しなくなるまで繰り返す．

< 11. 終了 > 残った候補ノードを，インデクスとして出力して終了．インデクスのフォーマットは（カテゴリ名，座標，解像度，距離値）である．

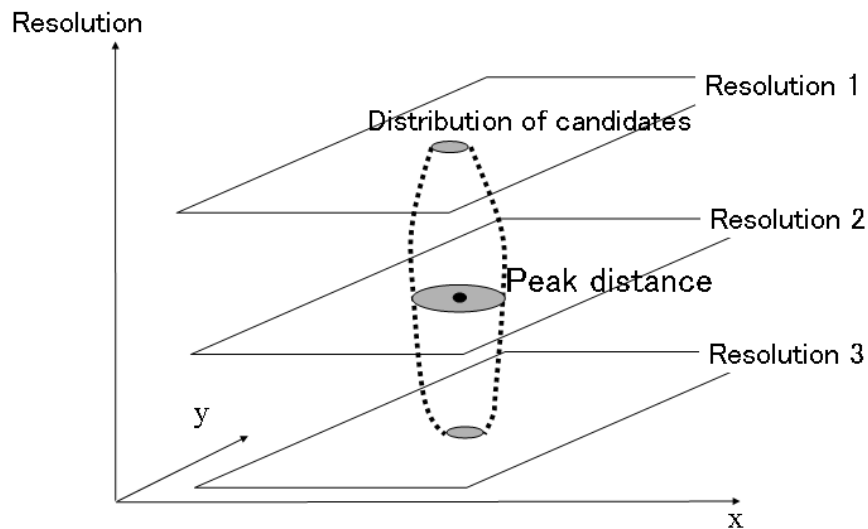


図 10 ピーク検出の例

2.7 むすびに

本論文では，文字，剛物体の画像認識において，複数ユーザが多様な環境下で撮影する際に発生する課題である，

- (1) 撮影位置と方向の変動
- (2) 照明変動，ぼけ，シェーディング
- (3) 隠蔽，照明光の全反射，影，背景変動
- (4) 独自のフォント，厚み
- (5) 隠蔽

によるパターン変化に対処する方法を提案した．

生成型学習 [村瀬 05] の考えに基づいた部分空間法 [Oja83] により (1) に，濃淡画像の低周波と高周波を除去した，対象の性質に応じた形状に基づく 3 つの特徴量を利用することで (2)，(4) に対処する．また，パターン全体を用いるのではなく，パターンの一部である部分パターンを利用して観測窓中の背景領域を減少させ，識別の際にはあらかじめ用意したモデルを用いてサイズや位置を変えながらスキミングを行うことで (3) に対処した．また，スキミングにおいては，粗探索の候補結果から，各カテゴリで探索範囲を限定することにより高速化が可能である．さらに，部分空間法を拡張し，特徴欠損/雑音付加の伴うパターンを学習可能とすることで (5) に対処した．提案法では，特徴欠損/雑音付加の伴うパターンを生成せずに，モデルを求めることができることを示した．

上記のアルゴリズムは，高速性と精度のトレードオフを操作する処理が多数含まれる．各ステップは課題に応じて取捨選択をしたり，簡略化することも可能であり，その方が実用上望ましい場合もあることに留意して応用に利用する．その選択肢について表 1 にまとめる．

表 1 アルゴリズムの選択肢

パターン学習	特徴ベクトル	縦横微分画像 方向のみ微分画像 加重方向指数ヒストグラム特徴
	幾何学変形によるパターン生成	アフィン変換 透視投影変換
	特徴抽出と階層構造	階層構造/非階層構造
	モデル生成	不定次元の利用/不利用
パターン認識	ピーク検出	連結成分の算出/非算出
	局所領域でのスクリーニング	実行/非実行
	同一候補ノードの統合	実行/非実行

3. 景観中文字の認識と画像検索への応用

3.1 はじめに

本章では、大量画像データベースからのインデクシングによる受動検索の一例として、景観画像内に存在する文字を自動でインデクスとして付与し、画像検索を効率化する課題を扱う。デジタルカメラの普及に伴い、ユーザはより多くの画像を撮影し、保存するようになった。しかし、その検索に必要なインデクスは、時刻等の機器によって付与される情報だけでは不十分である。画像内容に基づく自動インデクス付与は画像検索に有用である。

以下では、3.2 節において、本章で扱う問題の設定と 2 章において述べた提案法を用いたインデクシング方法について述べ、3.3 節において画像検索の方法を提案する。3.4 節においてインデクシングおよび画像検索の実験結果を述べ提案方法の有効性について示す。最後に、3.5 節において本章をまとめる。

3.2 問題設定と課題解決方針

本章では、ユーザが撮影した個人的な画像 DB に対して文字のインデクスを付与することを前提とする。特徴量は 2 章にて述べた加重方向ヒストグラム特徴を利用する。また、景観画像中の文字列は平面上にあるという制約を設け、文字はある程度遠くに撮影されていると仮定し、学習パターンは正面画像に対してアフィン変換を加えて生成型学習を行う。文字においては隠蔽の課題は発生しづらく、似た文字パターンも多いため、モデル化方法としては通常の部分空間法を用いる。対象の文字は活字体に限定し、手書き文字は対象から除外する。

図 11 は提案手法および画像検索アプリケーションの処理手順を示した図である。画像 DB 内のすべての画像に対し、文字候補群を検出してインデクスを生成する。本手法では、大きさの異なる文字に対応するために、複数解像度画像を生成し、位置をずらしながら定められた大きさの画像を切り出し、粗密探索を行う。その後、結果をインデクスとして利用する。画像検索では、キーワードが入力されると、インデクスから該当する文字候補のみを抽出して幾何学的な位置関係の

規則性を判定し，規則的と判定された画像を結果として出力する．

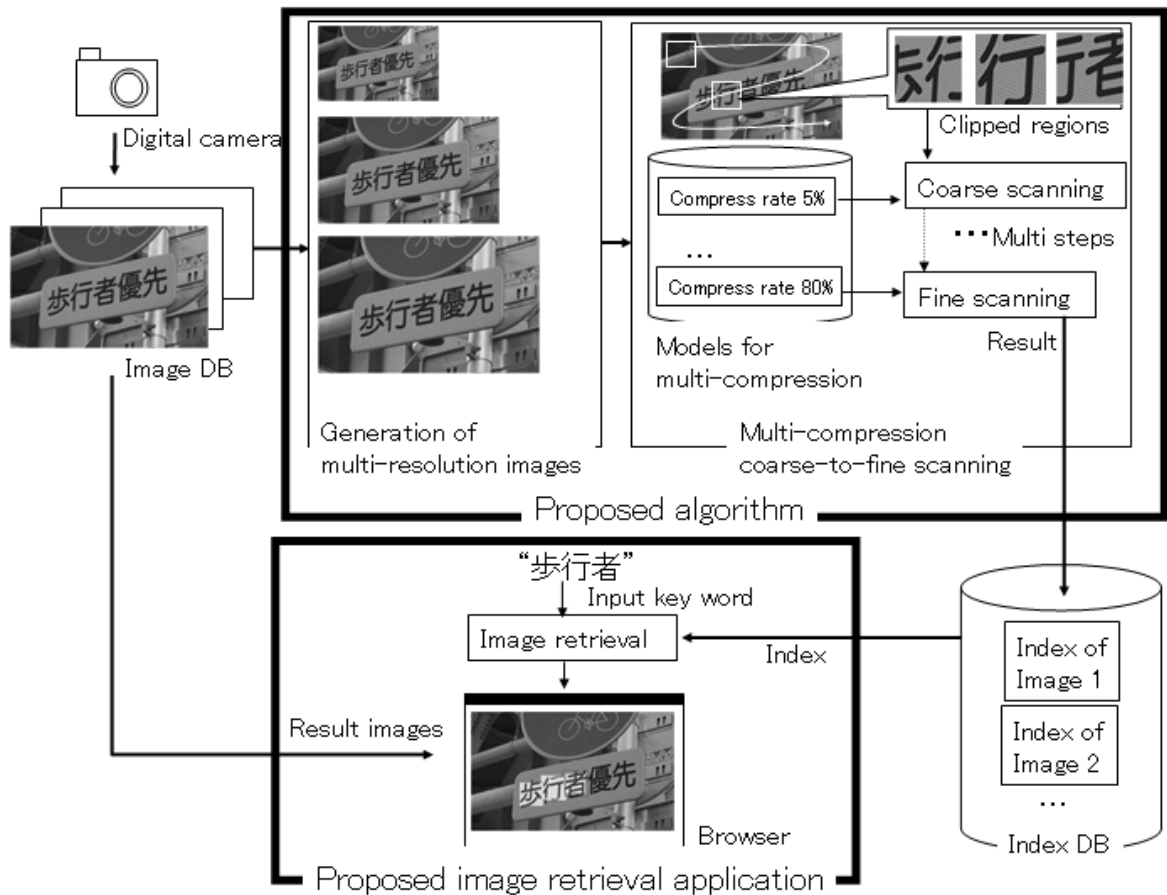


図 11 提案アルゴリズムおよび画像検索アプリケーションの概要

3.3 画像検索

画像検索では，キーワードが文字列として入力されると，2章で述べた提案手法で得られるインデクスの中からパターンが空間的に規則的に配置された個所を探索し，その位置と画像を検索結果として出力する．

パターンの空間的な配置の規則として，ここでは

1. ピッチが一定であること

2. パターンの大きさが揃うこと
3. ピッチが個別のパターンの大きさに対して相対的に一定の範囲内に収まること
4. パターンの並ぶ順序が入力された文字列の順序と一致し、かつパターンの並ぶ方向と水平方向もしくは垂直方向との角度の差が一定の範囲内であること
5. パターンは最初の文字から最後の文字まで全て画像の中に収まっていること

の5つを用いる。探索アルゴリズムに求められる条件は、a) 偽の文字候補を多く含むインデクスから高速に文字列を探索できること、b) インデクス中に正しい文字候補が一部存在しなくても探索できること、の2点である。2章で述べた提案手法で得られるインデクスには、背景において特徴量が文字に近い部分パターンが多く偽の文字候補として現れやすく、また劣化した文字パターンに対しては認識誤りが生じやすいため、上記 a), b) の条件が必要とされる。

提案アルゴリズムでは、まず入力文字列を構成する順方向の任意の2個の文字の組み合わせ(入力中の文字位置を各々 Q_1, Q_2 とする)がインデクスの中で存在する個所を全て探す。インデクスの情報から、元画像($L_x \times L_y$)における座標(X_1, Y_1), (X_2, Y_2)、およびそれらの解像度から大きさ W_1, W_2 、文字送りを表す2次元のベクトル(D_x, D_y)を算出する。また、ピッチ一定と見なし、開始位置の座標(X_s, Y_s)および終了位置の座標(X_e, Y_e)を算出する。次に、(X_s, Y_s, D_x, D_y)の4パラメータで構成される投票空間に投票を行なう。投票の際には、その組み合わせが上記(2)~(5)の規則を満たすか否かを以下の式でチェックし、満たす場合のみ投票する。

$$2. \quad Th_1 < \frac{Max(W_1, W_2)}{Min(W_1, W_2)}$$

$$3. \quad Th_2 < \frac{(X_2 - X_1)}{(Q_2 - Q_1)Max(W_1, W_2)}$$

$$4. \quad Th_3 < |\tan^{-1} \frac{Y_2 - Y_1}{X_2 - X_1}|$$

$$5. \quad Th_4 < Min(\frac{L_x - X_e}{L_x}, \frac{X_s}{L_x})$$

これらは横書きの文字列の探索の場合の式であり、縦書きの文字列の探索の場合にはこれらに準じて規定される。また補助的な規則として、その組み合わせが持つ認識結果としての距離値の合計値が一定の閾値以上の場合には信頼性が低いとみなして投票しない。

図12は投票に用いる4つのパラメータの算出過程を説明する模式図であり、「会」と「社」の組み合わせから、この2つが空間的に4文字の文字列「通信会社」の3番目と4番目を構成することを想定した際の仮想的な開始点の座標と文字送りを表すベクトルが示されている。そして最後に、投票空間の中からスコアが閾値以上の個所の有無を探索する。

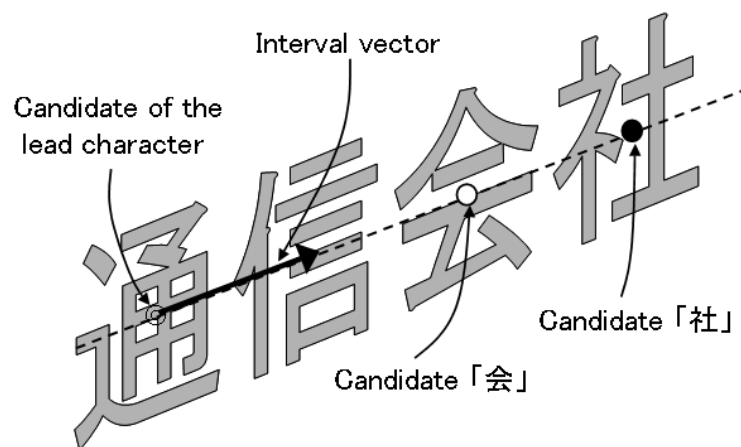


図 12 パラメータ算出過程

3.4 実験結果

[学習条件]

実験のために，3201 カテゴリ，14 フォントの活字パターンを収集し，非階層化モデルおよび階層化モデルの2種類を生成した．14個のフォントを用いて正面からの文字画像を生成し，表2に示すパラメータを用いて非階層化モデルを，表3に示すパラメータを用いて階層化モデルを生成した．ただし，階層化モデルの第二階層は，非階層化モデルを用いた．

表2 非階層化モデル生成のパラメータ

カテゴリ数	3201
フォントの数	14
画像サイズ (W)	40
グリッド (N)	8
方向数 (M)	4
特徴の次元数	256
ノルムの大きさ (K)	1.0
圧縮率 (r)	0.7
一カテゴリ当たり部分パターン数	6300
縦横比	1:2 ~ 2:1
大きさ倍率	0.5 ~ 1.5
スキュー角度	20 ~ -20 度
回転	10 ~ -10 度
木構造の階層数	1

表3 階層化モデル生成のパラメータ

木構造の階層数	2
第一階層のノード数	1200
第一階層の圧縮率 (r)	0.2

[評価対象]

提案アルゴリズムを評価するため、6地域（吉祥寺，銀座，原宿，浅草，巣鴨，池袋）の景観画像（1600x1200）を各150枚，計900枚を収集し，短辺が90ピクセル以上のサイズとなる全1760サンプルの漢字画像（正方形）を手動で切り出した（表4）。ただし，幾何変形した漢字画像については，図1（d）のように短辺を基準として部分画像として切り出している。

図13に例として「銀」に対してのサンプルを示す。



図13 「銀」の文字例

表 4 文字サンプル数

地域名	サンプル数
吉祥寺	315
銀座	322
原宿	251
浅草	234
巣鴨	294
池袋	344
評価サンプル計	1445

[評価環境]

OS は Windows2000 , CPU は PentiumIV の 3GHz , モデルメモリとして約 100 メガバイト , 作業メモリとして約 180 メガバイトを利用した .

3.4.1 漢字画像の識別実験

切り出した漢字画像に対して , 階層化モデルと非階層化モデルを用いて累積識別率を測定した . 非階層化モデルを用いた累積識別率を図 14 に , 階層化モデルを用いた累積識別率を図 15 に示す . ただし , 吉祥寺から切り出した漢字画像を用いてパラメータを表 5 のように調整した .

表 5 漢字画像識別実験のパラメータ

第一階層の部分空間次元数 (D)	10
第一階層の距離閾値 (Th_d)	0.2
第一階層の順位閾値 (Th_r)	100
第二階層のの部分空間次元数 (D)	40

困難な課題設定であるにも関わらず , 階層化モデルを用いて全地域において 15

位識別率 86%以上を確認した．図 16 に，1 位として識別された漢字画像の例を示す．

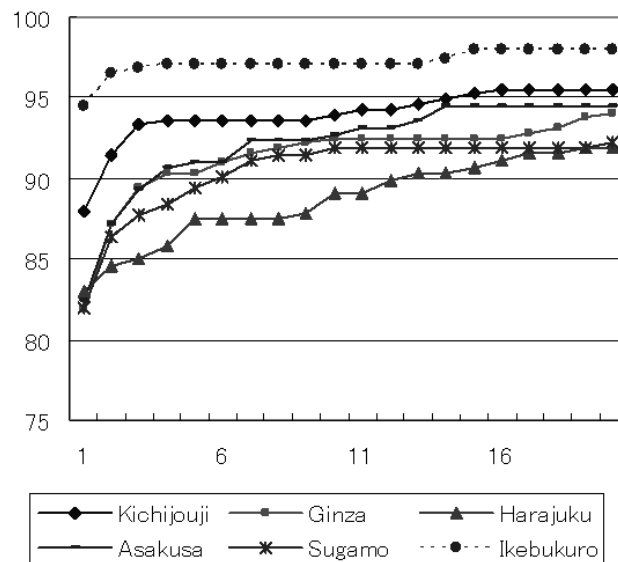


図 14 非階層化モデルによる累積分類率

OCRのように高い識別率を実現できない理由は，大きく4つの要因があり，これらが同時に発生する場合もある．

- 想定外のフォント（例：図 17）．
- 想定外の変形（例：図 18）．
- 立体, ボケによるパターンのつぶれ（例：図 19）．
- 背景パターンによるノイズ（例：図 20）．

背景パターンによるノイズに関しては，2.6 節で述べた走査では，より小さい切り出しによって文字候補として抽出できる可能性がある．ただし，小さい文字については部分パターンが切り出しの対象とならないため，上記のように識別不能となる．

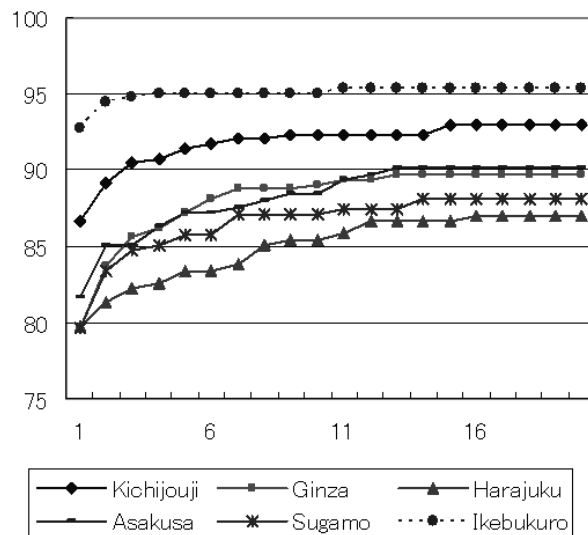


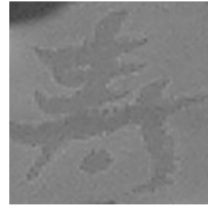
図 15 階層化モデルによる累積分類率



図 16 1位として識別された漢字画像例



「生」



「寿」



「蔵」

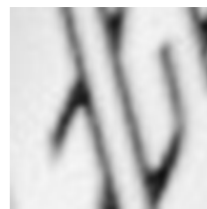


「雑」

図 17 識別不能例 (フォント)



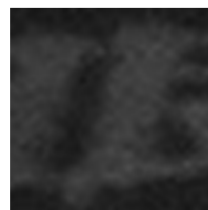
「転」



「北」



「駅」



「眼」

図 18 識別不能例 (変形)



「吉」



「齒」



「朝」



「障」

図 19 識別不能例 (つぶれ)



「託」



「券」



「事」



「損」

図 20 識別不能例 (背景パターン)

地域により傾向があり，上記サンプルが少なかった池袋では高い識別率を実現している．階層化モデルについては，非階層化モデルを用いた結果に比べて3～4%程度識別率が落ちるものの，階層化モデル：非階層化モデルの計算コストは1:18.5であり，速度/計算コストの良好なトレードオフと考える．

3.4.2 対象画像からのインデクス抽出実験

900枚のデータベース内の各画像に対してインデクス抽出実験を行った．短辺が90ピクセル以上の漢字とひらがな・カタカナの再現率，適合率を算出した．その結果を表6，表7に示す．ただし，再現率が切り出した漢字画像の1位識別率と同程度になるように吉祥寺のデータを用いて各パラメータを表8のとおり調整した．各画像に対し1000程度の候補文字であった．その他のパラメータは表5を用いている．処理時間は，1枚につき平均1分程度であった．

すべての地域において約80%程度の再現率を確保しており，3.3節で記載した検索技術を用いて，キーワードを用いた画像検索に耐え得る水準と考える．再現率と切り出し漢字画像の1位認識率には相関関係（相関係数0.72）があり，かつ吉祥寺のデータと比較して性能低下は小さく，インデクス抽出に関するパラメータの対象依存性は少ないと考える．再現率が切り出し文字の累積識別率よりも低くなる理由は，9.局所領域でのスクリーニングで，距離値の悪い候補が削除されてしまうのが主な要因である．また，ひらがな・カタカナの再現率が低い理由は，漢字を用いてチューニングをしたためである．

適合率が低い理由は，提案法がインデクス抽出の段階では文字候補としてできるだけ多くの可能性を残しておき，最終的に文字配列の幾何学的な規則性というヒューリスティックを用いて文字候補を絞り込むためである．ひらがな・カタカナの適合率が漢字よりも高い理由は，ひらがな・カタカナの種数が漢字の種数よりも少なく，相対的にひらがな・カタカナの文字候補が少ないためである．

表 6 各地域の文字再現率

地域	漢字 (%)	ひらがな・カタカナ (%)
吉祥寺	88.2	87.4
銀座	85.0	81.0
原宿	81.0	79.7
浅草	87.6	86.4
巣鴨	87.2	79.0
池袋	95.0	87.5
平均	87.2	82.7

表 7 各地域の文字適合率

地域	漢字 (%)	ひらがな・カタカナ (%)
吉祥寺	0.22	0.48
銀座	0.19	0.50
原宿	0.15	0.41
浅草	0.16	0.54
巣鴨	0.17	0.47
池袋	0.27	0.54
平均	0.19	0.49

表 8 インデクス抽出に関するパラメータ

複数解像度画像数	9
解像度画像の横サイズ	189 ~ 760
第二階層の距離閾値 (Th_d)	0.2
第二階層の順位閾値 (Th_r)	10
P_b	20
Th_{br}	5
Th_c	1.0

処理時間は、主に二つの処理によりばらつきが生じる（最大2分から最小30秒程度）。一つは、切り出し領域にエッジ成分が無い場合に、識別処理を行わずに、文字候補なしとする処理である。空や壁などのエッジのない領域が多く存在する画像に対しては処理が速くなる。もう一つは、粗探索後に連結成分を算出し距離値のピークのみを文字候補とする処理である。文字、看板やビルの端などの領域は粗探索の結果、同一の文字候補が集中しやすいため、そのような領域の多い画像に対しては密探索処理が速くなる。

3.4.3 検索実験

900枚のデータベース内の短辺が90ピクセル以上の文字を有する各画像に対して主キーワード（ひらがな、カタカナを含む）を定義し、そのキーワードを用いて検索実験を行った。各画像のインデクスは3.4.2節で生成したものをを用いて実験した結果を表9に示す。 P を正しいキーワードを含む検索結果数、 P' をデータベース中の正しいキーワードを含む画像数、 T を検索結果数、とすると、再現率は $100 \times P/P'$ 、適合率は $100 \times P/T$ である。ただし、本実験においては吉祥寺のデータを用いた予備的な検索実験により、再現率が80%を超えるように各パラメータを表10のとおり調整した。

平均再現率 78.9%、平均適合率 54.9%、検索時間は PentiumIV プロセッサで平

均0.1秒/語以下であり、個人用の比較的小規模な画像データベースであれば、実用可能な精度・計算コストと考える。表9の結果を見ると、画像検索の再現率と切り出し漢字画像の1位認識率には相関関係（相関係数0.67）があり、かつ吉祥寺のデータと比較して性能低下は小さく、検索に関するパラメータの対象依存性は少ないと考える。

図21にシェーディングを伴う文字列、図22に立体文字列、図23に複雑背景上の文字列、図24に傾斜文字列の検索成功例を示す。ただし、検索語の部分のみをハイライト表示している。照明条件、フォント、背景などのパターン変動が大きい条件であるため、文字候補の検出もれが発生しているが、正しく検索できていることがわかる。図22の「神」に関しては、学習したフォントとは類似していないため、検出に失敗している。また、図23の「品」に関しては、文字の中央部（「口」の内部）や上部（口の左右）に背景パターンが混入するため、部分パターンによるノイズ低減では対応できず、検出に失敗している。

画像検索の失敗要因は、大きく以下の3つに分類できる。

- 文字認識に失敗する。
- 文字列の一部を誤検出する。
- 背景のパターンを誤検出する。

文字列の一部を誤検出する例として、図25にキーワード「出入口」による検索結果例を示す。「出」は検出されていないが、提案法では文字候補の一部欠落を許容するため、「入口」という文字列の一部を検出してしまふ。また、背景のパターンを誤検出する例として、図26にキーワード「出入口」による検索結果例を示す。「出」、「口」ともに単純なパターンであり、提案法は切り出した局所画像のみから文字認識を行うため、背景のテクスチャ上に偽文字候補が出現してしまふ。また「入」は検出されていないが、提案法では文字候補の一部欠落を許容するため、検出されてしまふ。

表 9 各地域の画像検索再現 / 適合率

地域	再現率 (%)	適合率 (%)
吉祥寺	82.5	57.6
銀座	72.3	47.8
原宿	75.9	53.0
浅草	81.2	59.0
巣鴨	79.2	57.9
池袋	85.7	56.9
平均	78.9	54.9

表 10 検索に関するパラメータ

Th_1	2.0
Th_2	5.0
Th_3	40 (度)
Th_4	0.1

3.5 むすびに

本章では、景観画像中の文字候補群を用いた画像インデクシングおよび検索法を提案した。提案法では、従来法では検出困難なシェーディングを伴う文字、立体文字、複雑背景上の文字、傾斜文字、を文字候補として検出する能力がある。また、検索の際に検索文字列の空間的な配置関係を利用し、多くの偽文字候補の出現や文字候補の検出もれを許容して、実用的な画像検索を実現できる。

漢字の単文字認識実験を行い、認識率が 86.0%以上となることを確認した。また、景観画像のインデクシング実験を行い、漢字・ひらがな・カタカナの再現率が 80.0%以上となることを確認した。作成されたインデクスを用いて、キーワード

による画像検索精度の測定実験を行い，平均再現率 78.9%，平均適合率 54.9%を確認した．



図 21 シェーディングを伴う文字列の検索成功例：検索語は「平和通り」.

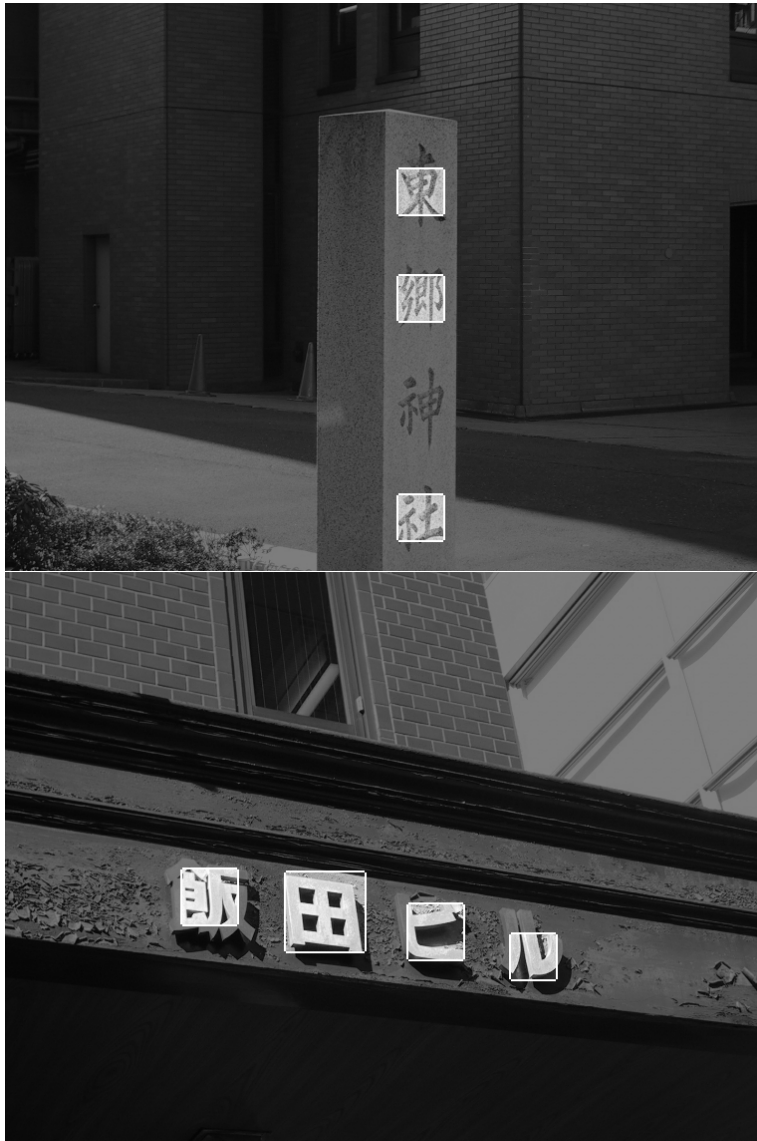


図 22 立体文字列の検索成功例：検索語は上から「東郷神社」、「飯田ビル」.

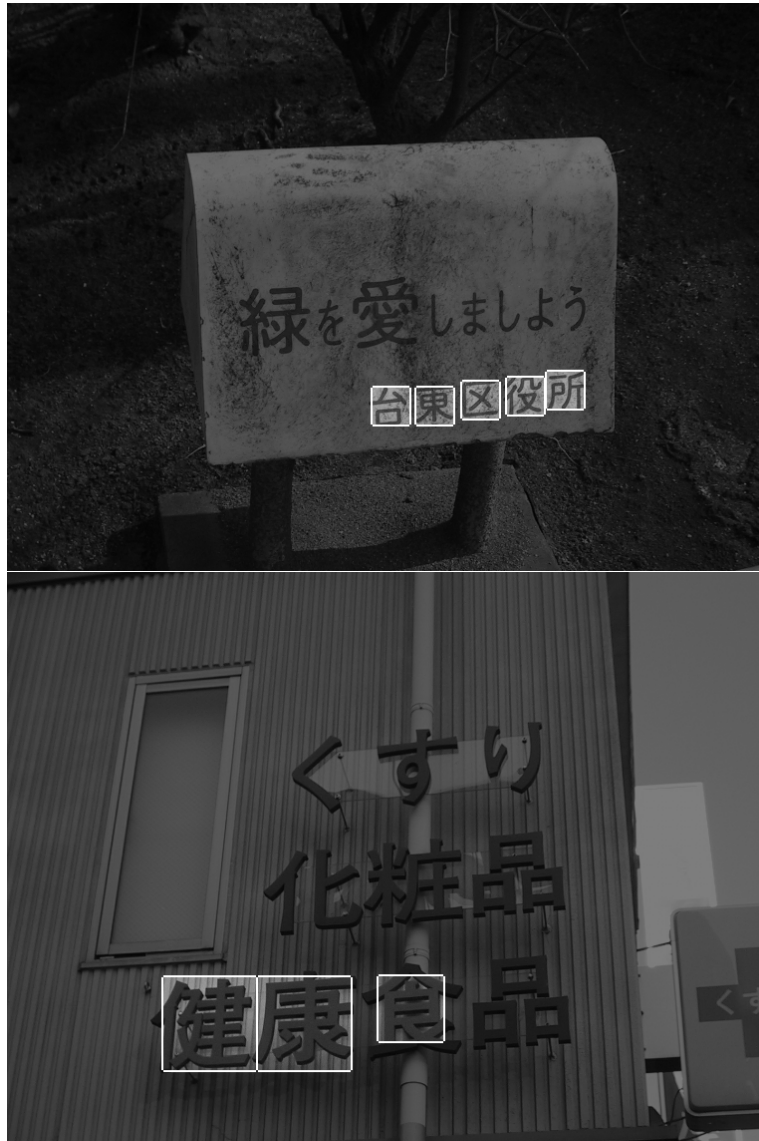


図 23 複雑背景上の文字列の検索成功例：検索語は上から「台東区役所」、「健康食品」。



図 24 傾斜文字列の検索成功例：検索語は「豊島区」

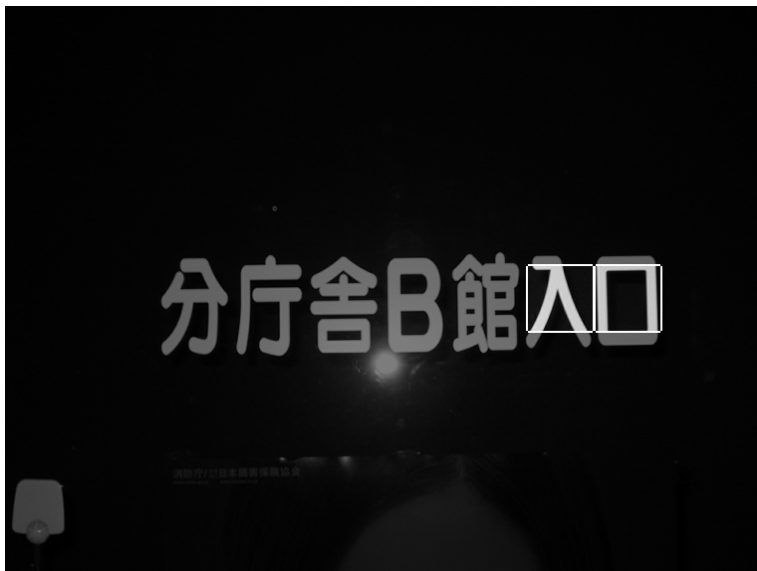


図 25 文字列の一部の誤検出による検索失敗例：検索語は「出入口」



図 26 背景パターンからの誤検出による検索失敗例：検索語は「出入口」

4. 剛物体の認識と画像撮影による情報取得手段への応用

4.1 はじめに

本章では，画像をキーにした対象物の情報能動検索の一例として，対象のデザイン性を損なわずに，ユーザが直感的に携帯から情報にアクセスするインタフェースを実現する課題について扱う．携帯電話の普及に伴い，携帯から情報にアクセスするインタフェースを改善したいという需要が高まっている．そのため，一般的となったカメラ付携帯電話によって，対象を撮影することにより情報を入手する技術が注目されている．

以下では，4.2節において，本章で扱う問題の設定と2章にて述べた提案法を用いた対象識別方法について述べる．4.3節において提案法を用いたシミュレーション実験，4.4節において2次元平面への適用実験，4.5節において3次元物体への適用実験を述べて，提案方法の有効性について示す．最後に，4.6節において本章をまとめる．

4.2 問題設定と方針

本章では，ユーザが対象を中心に全体が画像内におさまるように撮影することを前提とする．ただし，ピンボケ，若干の位置ずれ，大きさの分散は許容する．ユーザは対象の情報を得る目的で撮影するため，若干の制約条件は許容できる．特徴量は2.3節にて述べた縦横微分画像と方向のみ微分画像を利用する．2次元平面パターンに対しては，カメラの近くに撮影されていると仮定し，学習パターンは正面画像に対して透視投影変換を加えて生成型学習を行う．モデル化方法としては不定次元を用いた部分空間法を用いて，部分的な隠蔽に対処する．照合方法としては，粗密探索を用いる．



図 27 システム概要図

図 27 に示すカメラ付携帯電話による情報提供システムはカメラ付携帯電話と認識サーバから構成され、ユーザは以下の手順により動画像などのコンテンツを取得できる。

1. カメラ付携帯電話にてユーザが対象を撮影。
2. 画像を剛物体認識サーバに送信。
3. 剛物体認識サーバにて画像をモデルと照合・認識し、登録済URLを返信。
4. ユーザがコンテンツを閲覧。

本システムを利用することで、情報取得手段のインタフェースを改善することができる。

4.3 シミュレーション実験

本節では、把持した3次元物体を撮影し、その姿勢を推定する課題を想定し、提案手法の妥当性を検証する。

[学習条件]

図28のようにオブジェクトを垂直軸方向-90度から90度まで、水平軸方向40度から-20度まで回転させ、各10度間隔で133枚撮影し、学習画像とした。

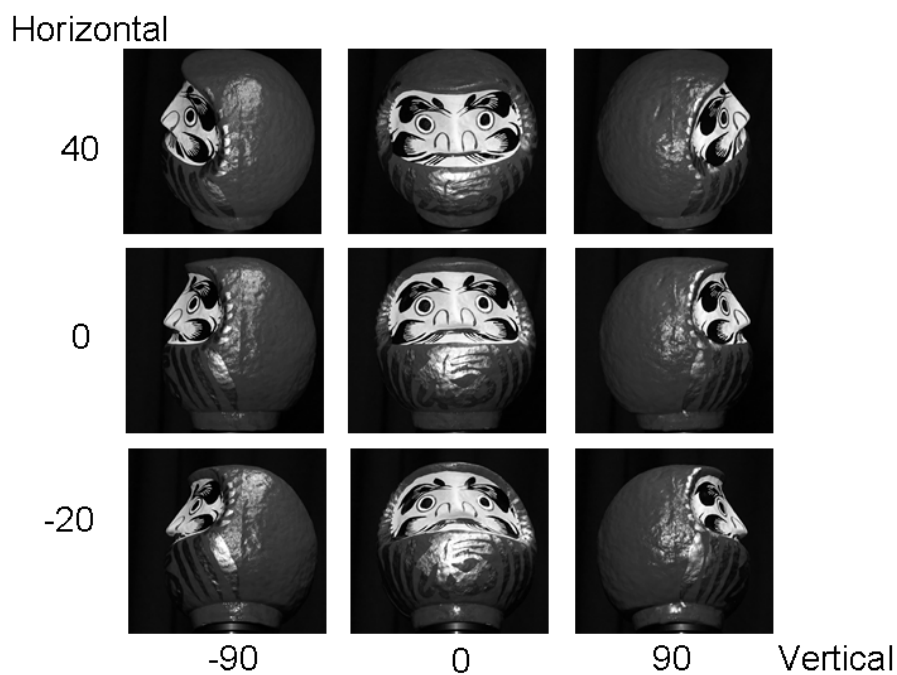


図28 オブジェクトの学習画像

学習画像を16分割し、 $({}_{16}C_2) = 120$ パターンのマスク画像(図29)を生成した。マスク画像の黒い領域を不定次元とし、各学習画像に対して120枚の不定次元を含む縦横微分画像を生成した。微分オペレータにはソーベルオペレータを利用した。また、不定次元に対するソーベルオペレータの結果は不定次元とした。学習パターンは、 $32 \times 32 \times 2 = 2048$ 次元、不定次元は各 $8 \times 8 \times 2 \times 2 = 256$ 次元である。また、 $w_{max} = 255$ 、 $w_{min} = 0$ 、 $w_{step} = 256$ 、 $p(w_r) = \frac{1}{256}$ (一様分布と

仮定), $S = 100$, $S_{min} = 0.5$, $S_{max} = 2.0$, $S_{step} = 16$ とした. 各姿勢において部分空間を求め, 識別のための部分空間次元には, 寄与率 95%を用いた.

[評価対象]

テストパターンとして, 図 30 のように, 領域選択パターンとして上下左右の 2 領域を選択した 4 パターン, テクスチャパターンとして白色格子と白色面の 2 パターンの, 合計 8 パターンを用意した. 姿勢識別においては, 学習パターン 1 つに対して部分空間を作成することはできないため, 比較従来法としてパラメトリック固有空間法 (PEM) を用い, 固有空間内での NN 法により姿勢を推定した. 特徴ベクトルは, 原画像および縦横微分画像を利用し, 寄与率は 99.9%とした. ただし, 評価には学習パターンに雑音を付与した画像を用いたため, 姿勢変化に伴う補間はない. また, 各主成分の次元の値が 0 である次元を排除して正規化を行った.

[評価結果]

テストパターンを全 133 枚に適用して識別実験を行った結果を表 11 に示す.

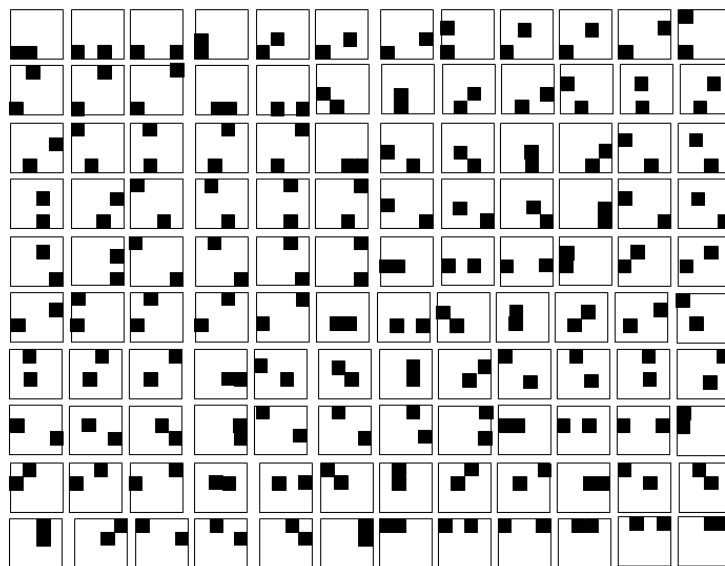


図 29 マスク画像

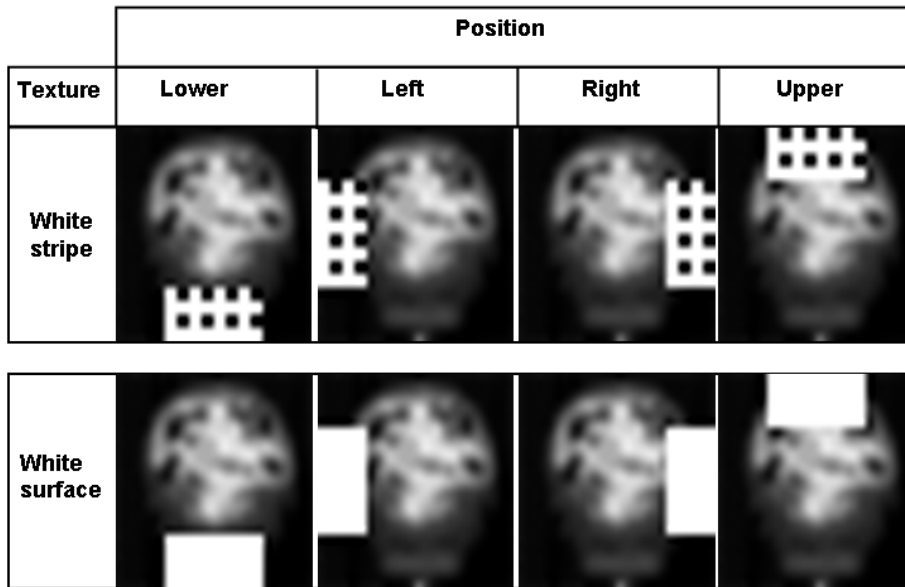


図 30 テストパターン

表 11 姿勢識別結果

	下	左	右	上
白色 (PEM 原画像)	60.2	73.7	49.6	59.4
白色格子 (PEM 原画像)	94.0	97.7	77.4	81.2
白色 (PEM 縦横微分画像)	100.	100.	100.	97.7
白色格子 (PEM 縦横微分画像)	100.	100.	100.	89.5
白色 (提案法)	100.	100.	100.	99.3
白色格子 (提案法)	100.	100.	99.3	98.5

従来法 (PEM 原画像) においては、特に白色パターンにおいて識別精度が低い。正規化による変動の影響が大きくなるためと考えられる。従来法 (PEM 縦横微分画像) においては、精度がよく改善されており、縦横微分画像は局所的な

低周波雑音にも効果があることがわかる。白色格子パターンの“上”で他に比べて識別率が低くなる要因は，“上”領域が固有空間に変換する際に重要視され，かつ雑音の量が白色パターンよりも大きいためと考える。縦横微分画像は，局所的な中周波雑音には対応できていないことがわかる。

提案法は，全パターンにて高い識別率が得られ，妥当性が実験的に示された。白色格子パターンの“右”および“上”，白色面パターンの“上”において識別率が若干低下する要因は，不定次元を定めた結果，部分空間同士が重なるためと考える。

本結果より提案法である，縦横微分画像および不定次元を用いた部分空間法の妥当性が実証されたと考える。

4.4 平面デザインの認識実験

本節では，提案法を適用して，対象のデザイン性を損なわずに，ユーザが直感的に携帯から情報にアクセスするインタフェースを実現し，平面デザインに対して評価実験を行った結果を報告する。

[評価環境]

OSはWindows2000，CPUはPentiumIVの3GHz，メモリは2ギガバイトのマシンを利用した。

4.4.1 ちらしの認識

2005年8月13日から20日まで開催された株式会社レッド・エンタテインメント企画・監修，株式会社セガ主催の「サクラ大戦 スーパー歌謡ショウ - 新・青い鳥 -」にて配布されたちらしを利用し，俳優の動画メッセージや壁紙を配信する実験を行った。

[学習条件]

撮影ポイントは，各俳優および各キャラクターとロゴ，ちらし全面，キャラクター全写真，ポスター，ポスターロゴの23種類（図31）とした。

[評価条件]

ちらしはコーティングされており、照明反射が頻発する。また、ポスターは会場内にライティングして1枚配備した。対応端末はFOMAの900i, 901i, 902i, 700i, 701iシリーズとした。撮影画像は128×96とし、圧縮率を一番高く設定した。また、ユーザには対象を画面一杯に中央に撮影して下さいというガイダンス提示を行っている。ユーザ数は430人、撮影画像数は4859枚であった。



図 31 撮影ポイント

収集した画像サンプルに対して、特徴ベクトルを2種類（原画像，縦横微分画像），識別アルゴリズムを2種類（部分空間法，不定次元を用いた部分空間法）を組み合わせた表12に示す4つのアルゴリズムを適用した。モデル作成の際に，ユーザの撮影角度と撮影距離のばらつきに対応するために，平面画像を透視投影変換した複数の学習画像を人工的に生成して学習画像とした。各方法のモデル作成のパラメータを表13に示す。

各学習画像に対してガウシアンフィルタをかけ原画像とし，原画像に対してソーベルオペレータを適用して縦横微分画像とした。また，不定次元を用いた部

表 12 適用アルゴリズム

	原画像	縦横微分画像
部分空間法	手法 A	手法 C
不定次元を用いた部分空間法	手法 B	手法 D

表 13 モデル作成パラメータ (ちらし)

手法	手法 A	手法 B	手法 C	手法 D
次元数	40 × 40	40 × 40	40 × 40 × 2	40 × 40 × 2
不定次元数	0	10 × 10	0	10 × 10 × 2
水平軸回転	-10 ~ 10 度	同左	同左	同左
垂直軸回転	-10 ~ 10 度	同左	同左	同左
投影軸回転	-10 ~ 10 度	同左	同左	同左
拡大縮小	56% ~ 100%	同左	同左	同左
拡大縮小ステップ数	5	同左	同左	同左

分空間法を適用する際には、マスク画像は学習画像を 4 × 4 領域に分割して、1 領域を不定次元とした 16 パターンを用意した。

[評価結果]

撮影ずれに対応するために、画像内で中心から 10% の位置ずれを許容する条件下で走査した実験結果を図 32 に示す。図 33 に示す認識成功例のように、ユーザの撮影環境は不定であり、照明反射、陰影、照明光の色の違い、ぼけのある撮影環境下であった。提案法 (部分空間 50 次元) で識別時間は平均約 0.5 秒、アンテナの状況によるが、撮影からコンテンツ閲覧まで 5 ~ 10 秒程度であった。

計算コストは、探索点数 × 次元数 × 部分空間数であり、原画像が 1600 次元であるのに対し、縦横微分画像は 3200 次元であるため、同計算コストで性能比較を行うと、原画像の部分空間次元を 1/2 として比較する必要がある。図 32 から、以下のことがわかる。

- 同じ計算量で、原画像よりも縦横微分画像の方が性能が良い。

- 同じ計算量で，部分空間法よりも不定次元を用いた部分空間法の方が性能が良い．
- 不定次元を用いた部分空間法の効果は，雑音混入量が大きいと考える原画像の方が顕著である．
- 不定次元による効果は部分空間次元の 25 次元以降に発現しており，幾何変形による学習パターンの分散よりも不定次元による分散が小さいと考えられる．

以上より，提案手法が従来手法よりも識別性能が高く，ユーザ，照明環境，および撮影機材による対象のパターン変化に頑健であることがわかる．具体的には撮影位置と方向の変動，照明変動，ぼけ，シェーディング，照明光の全反射によるパターン変化に頑健であることがわかる．

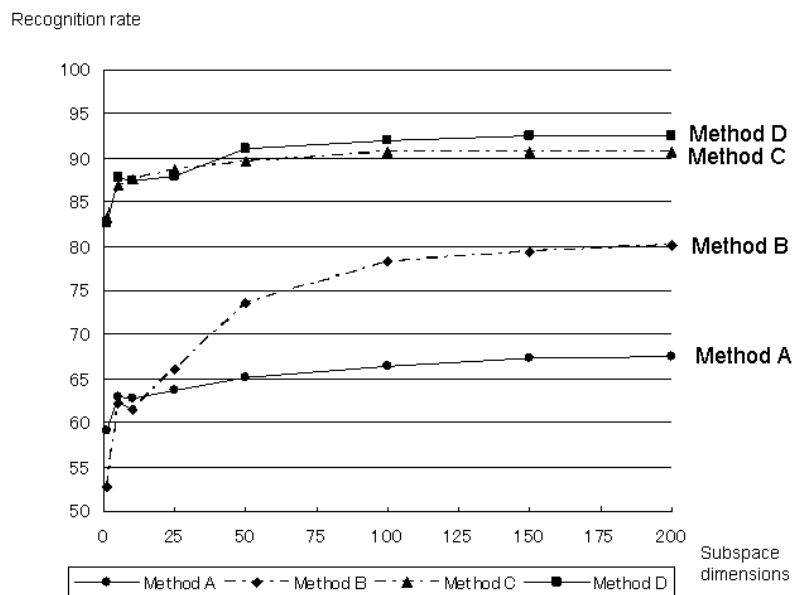


図 32 実験結果



図 33 撮影画像例（ちらし）

4.4.2 絵葉書の認識

[学習条件]

シャガールによって描かれた 30 種類の絵葉書をスキャナで読み込み，表 14 のパラメータに従って透視投影変換画像を人工的に生成し，学習画像とした．マスク画像は学習パターンを 4×4 領域に分割して，1 領域を不定次元とした 16 パターンを用意し，1 絵葉書につき 1 個のモデルを作成した．

[評価条件]

3 人のユーザが異なる場所 (1 機種) で絵葉書を把持して 30 回試行した (90 サンプル)．図 34 に撮影された画像例を示す．照明変動や指による隠蔽が発生していることがわかる．画像内で 10% 程度の位置ずれを許容する条件下で運用した．

[評価結果]

画像および結果の送付時間は4秒程度，識別時間は2秒程度であった．識別結果は，1位識別率92.9%，累積3位識別率97.1%であった．本実験により，提案法が実用上発生する照明変動/隠蔽にロバストであることがわかる．



図 34 撮影画像例（絵葉書）

4.4.3 ロゴの認識

[学習条件]

109種類のロゴを，表15のパラメータに従って透視投影変換画像を人工的に生成し，学習画像とした．マスク画像は学習パターンを 4×4 領域に分割して，1領域を不定次元とした16パターンを用意し，1ロゴに対し1個のモデルを作成した．

[評価条件]

雑誌等にロゴが配置されていることを想定し，背景に文字列を加えたロゴを撮影対象とした．4人のユーザが異なる8環境下で109回試行した（872サンプル）．

表 14 モデル作成パラメータ (絵葉書)

	絵葉書
次元数	50 × 74 × 2
不定次元数	13 × 19 × 2
水平軸回転	-10 ~ 10 度
垂直軸回転	-10 ~ 10 度
投影軸回転	-10 ~ 10 度
拡大縮小	56% ~ 100%
拡大縮小ステップ数	5

図 35 に撮影された画像の例を示す．大きな背景変動が発生していることがわかる．画像内で 10% 程度の位置ずれを許容する条件下で運用した．

[評価結果]

画像および結果の送付時間は 4 秒程度，識別時間は 0.5 秒程度であった．識別結果は，1 位識別率 86.7%，累積 3 位識別率 88.5% であった．本実験により，提案法が背景変動にロバストであることがわかる．



図 35 撮影画像例（ロゴ）

表 15 モデル作成パラメータ（ロゴ）

	ロゴ
次元数	40 × 40 × 2
不定次元数	10 × 10 × 2
水平軸回転	-10 ~ 10 度
垂直軸回転	-10 ~ 10 度
投影軸回転	-10 ~ 10 度
拡大縮小	56% ~ 100%
拡大縮小ステップ数	5

4.5 立体物の認識実験

本節では、提案法を適用して、対象のデザイン性を損なわずに、ユーザが直感的に携帯から情報にアクセスするインタフェースを実現し、立体物に対して評価実験を行った結果を報告する。

[評価環境]

OSはWindows2000，CPUはPentiumIVの3GHz，メモリは2ギガバイトのマシンを利用した。

4.5.1 ペットボトルの認識

[学習条件]

15種類のペットボトルを、表16のパラメータに従って撮影し、学習画像とした。また、投影軸回転、大きさ変動に関しては人口的に画像を生成し、学習画像に加えた。マスク画像は学習パターンを 4×4 領域に分割して、1領域を不定次元とした16パターンを用意し、1ペットボトルに対し1個のモデルを作成した。

表 16 モデル作成パラメータ (ペットボトル)

	ペットボトル
次元数	$23 \times 72 \times 2$
不定次元数	$6 \times 18 \times 2$
水平軸回転	-20 ~ 30 度
垂直軸回転	-30 ~ 30 度
投影軸回転	-10 ~ 10 度
拡大縮小	56% ~ 100%
拡大縮小ステップ数	5

[評価条件]

ペットボトルを把持し、4人のユーザが異なる4環境下(1機種)で15回試行した(60サンプル)。図36に撮影された画像の例を示す。局所的な照明光全反射や

指による隠蔽が発生していることがわかる．画像内で 10%程度の位置ずれを許容する条件下で運用した．

[評価結果]

画像および結果の送付時間は 4 秒程度，識別時間は 0.2 秒程度であった．識別結果は，1 位識別率 91.7%，累積 3 位識別率 95.0%であった．本実験により，提案法が局所的な照明光全反射や指による隠蔽にロバストであることがわかる．



図 36 撮影画像例（ペットボトル）

4.5.2 フィギュアの認識

[学習条件]

9 種類のフィギュア（図 37）を，ビデオを用いて対象を回転させて撮影した．その映像から静止画を抽出して 20 枚の学習画像とした（図 38）．また，背景変動への対処として，黒，白，青の 3 種類用意し撮影した（図 39）．また，表 17 に

従い，投影軸回転，大きさ変動に関しては各学習画像から人口的に画像を生成し，学習画像に加えた．マスク画像は学習パターンを 4×4 領域に分割して，1 領域を不定次元とした 16 パターンを用意し，1 オブジェクトに対し 1 個/背景のモデル（全 27 個）を作成した．

表 17 モデル作成パラメータ（フィギュア）

	フィギュア
次元数	$20 \times 30 \times 2$
不定次元数	$5 \times 8 \times 2$
水平軸回転	-45 ~ 45 度
垂直軸回転	-20 ~ 30 度
投影軸回転	-10 ~ 10 度
拡大縮小	50% ~ 100%
拡大縮小ステップ数	20

[評価条件]

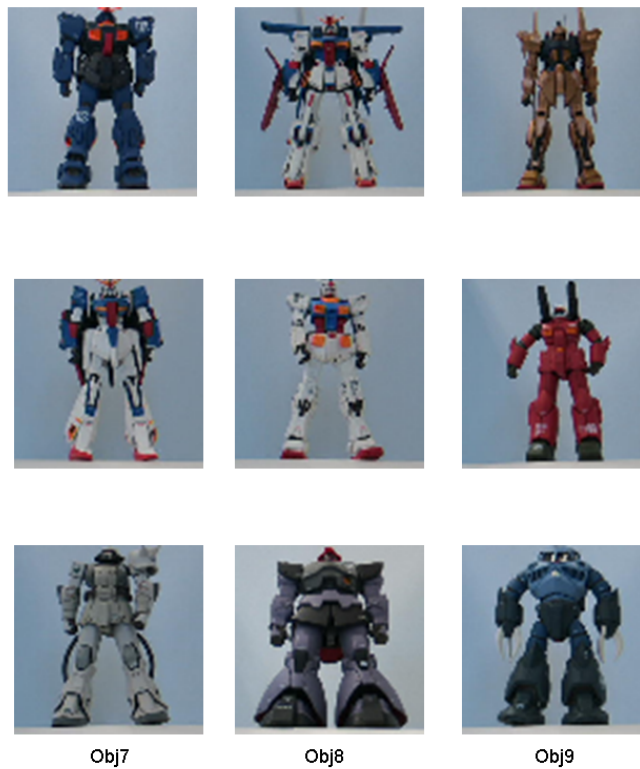
1 人のユーザが異なる以下の 5 環境で 7 方向（D1 D7）より撮影したサンプル（315 サンプル）で評価した．

- 無背景（図 40）
- デスク（図 41）
- カレンダー（図 42）
- 本棚（図 43）
- 廊下（図 44）

画像内で 20% 程度の位置ずれ，0.5 倍程度の大きさ変動を許容する条件下で運用した．

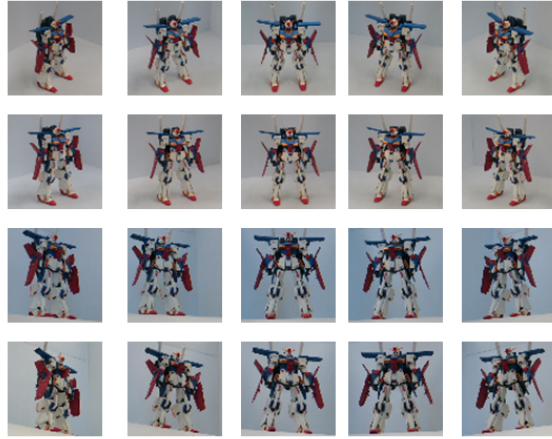
[評価結果]

評価の結果を図 45，図 46 に示す。



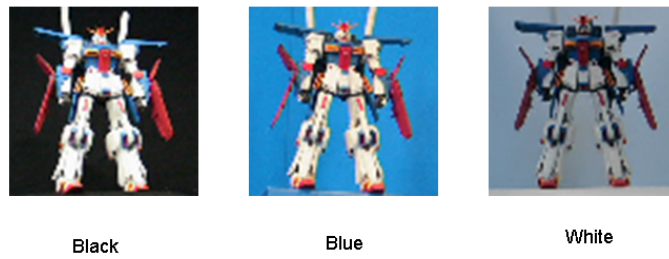
©創通・サンライズ

図 37 認識対象 9 種類



©創通・サンライズ

図 38 学習画像例



©創通・サンライズ

図 39 背景パターン



D1



D2



D3



D4



D5



D6



D7

©創通・サンライズ

図 40 無背景 (Obj2)



D1



D2



D3



D4



D5



D6



D7

©創通・サンライズ

図 41 デスク (Obj2)



D1



D2



D3



D4



D5



D6



D7

©創通・サンライズ

図 42 カレンダー (Obj2)



D1



D2



D3



D4



D5



D6



D7

©創通・サンライズ

図 43 本棚 (Obj2)



D1



D2



D3



D4



D5



D6



D7

©創通・サンライズ

図 44 廊下 (Obj2)

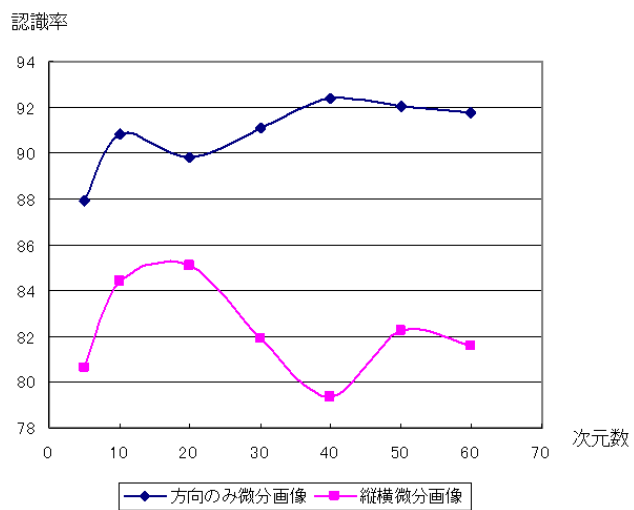


図 45 識別結果

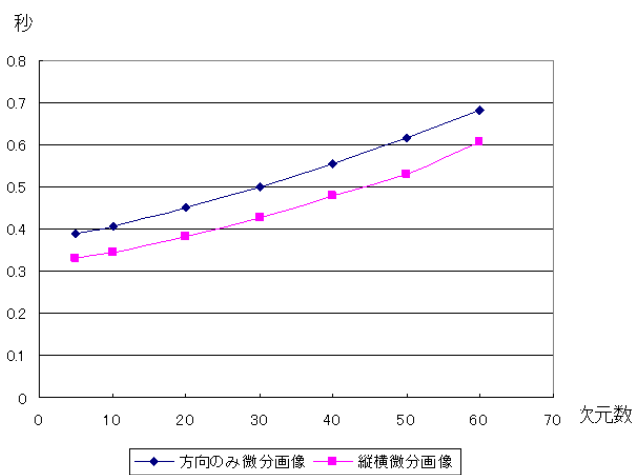


図 46 1枚あたりの処理時間

図 45 の結果について考察する．縦横微分画像を用いた認識結果は 20 次元でピークとなり，1 位識別率 86.7%であった．方向のみ微分画像を用いた認識結果は 40 次元でピークとなり，1 位識別率 92.7%であり，3 次元物体を対象とする場合，縦横微分画像を用いるよりも全次元において方向のみ縦横微分画像を用いた方が良いことがわかる．本実験により，ピークの次元数が異なる理由は，縦横微分画像においては，20 次元以上のデータはエッジの大きさに影響を受け，識別の情報としては役に立たない情報であり，方向のみ微分画像では有用な情報を含んでいたためである．

図 46 の結果について考察する．方向のみ微分画像においては切り出し画像に対して 2 値化処理を行う分だけ処理時間がかかっていることがわかる．また，次元数に対して処理時間はほぼ線形に増加しており，部分空間法における計算コストの理論と一致していることがわかる．

次に方向のみ微分画像の 40 次元の識別結果について詳細に考察する．表 18 は対象別，表 19 は角度別，表 20 は環境別の識別率結果である．

表 18 対象毎識別率

対象	識別率
Obj1	94.3%
Obj2	97.1%
Obj3	71.1%
Obj4	100.0%
Obj5	100.0%
Obj6	71.1%
Obj7	94.3%
Obj8	97.1%
Obj9	94.3%

表 19 角度毎識別率

角度	識別率
D1	93.3%
D2	93.3%
D3	88.9%
D4	91.1%
D5	97.8%
D6	91.1%
D7	91.1%

表 20 環境毎識別率

環境	識別率
無背景	100.0%
デスク	95.2%
カレンダー	82.5%
本棚	88.9%
廊下	93.7%

表 18 より，obj3 と obj6 の識別精度が低下している．輪郭形状が類似しているためと考えられる（図 37）．類似カテゴリに対しては，検索結果を複数候補を提示することでユーザ利便性を確保できると考える．方向のみ微分特徴においては，原画から特徴空間への写像に課題があり，十分な対象記述能力を確保できなかったためと考えられる．類似カテゴリの問題は文字認識でも研究され，専用の識別平面を用意することで解決可能なことが知られている．

表 19 より，D3 の識別性能がやや低下しているが，全体として大きな傾向はない．提案法が角度によって性能劣化が生じないことがわかる．表 20 より，背景が

カレンダーの場合に、識別性能が低下している。図 39 で示したように、背景色が一様なパターンにより学習したため、背景がカレンダーの場合に 2 値化傾向が異なったためと考えられる。

全体として類似カテゴリの課題を除くと各環境における識別率は 80% を超えており提案法による実用化は十分可能であると考えられる。

4.6 むすびに

本章では、画像をキーにした対象物の情報能動検索の一例として、対象のデザイン性を損なわずに、ユーザが直感的に携帯から情報にアクセスするインタフェースを実現する際のシステム構成について述べた。また、把持した 3 次元剛物体を撮影し、その姿勢を推定する課題を想定したシミュレーション実験を行い、提案法である縦横微分画像および不定次元を用いた部分空間法の妥当性を実証した。

実際にカメラ付携帯電話を用いた情報提供システムに提案法を適用し、実用上発生するユーザ、照明環境、周辺環境、および撮影機材による対象のパターン変化に頑健であることを示した。具体的には撮影位置と方向の変動、照明変動、ぼけ、シェーディング、照明光の全反射、隠蔽、背景変動、によるパターン変化に頑健であることを示した。同時に、平面パターンに対しては縦横微分画像、3 次元剛物体に対しては方向のみ微分画像が有用であることを評価実験により示した。また、従来の部分空間法に比べて不定次元を用いた部分空間法の優位性を示した。さらに、平面パターン、立体物ともに複数の実験を行い、そのすべてにおいて 85% 以上の識別率を確認し、提案法の実用性、および汎用性の高さを実証した。

5. 結論

5.1 本論文のまとめ

本論文では、形状を用いたアピアランスベースの利用するパターン特徴量，そのモデル化方法，画像内からの対象検出方法について提案した．提案法を景観中の文字に適用し，画像内容に基づく自動インデクス付与による画像検索に応用した．また，提案法を一般の2次元パターンを含む剛物体に適用し，カメラ付き携帯電話を用いて対象を撮影することにより情報を入手する情報取得手段に応用し，その有効性を確認した．

一般にさまざまなパラメータを伴う広範な対象を認識する技術を評価する場合，完備な評価データベースを構築することはできず，対象数などの実験条件によって評価結果には偏りが生じる．そのため，検証項目に即した評価パターンの収集と実際の利用想定環境での評価パターンの収集により，複数の視点から本技術を検証した．

第2章では，形状に基づくアピアランスベースの認識方法を提案した．提案法は，生成型学習 [村瀬 05] の考えに基づいた部分空間法 [Oja83] により撮影位置，方向によるパターン変動に，濃淡画像の低周波と高周波を除去した特徴量を利用することで照明変動，ぼけ，影，反射などのパターン変動に対処する．また，パターン全体を用いるのではなく，パターンの一部である部分パターンを利用して観測窓中の背景領域を減少させ，識別の際にはあらかじめ用意したモデルを用いてサイズや位置を変えながらスキヤニングを行うことで背景によるパターン変動に対処する．また，スキヤニングにおいては，粗探索の候補結果から，各カテゴリで探索範囲を限定することにより高速化を実現し，多数のカテゴリにも適用できる．

また，特徴量としては，フォントによる形状，立体文字における厚み部位による微小変動を吸収できるように改良し，文字パターンへの汎用性を高めた．さらに，部分空間法を拡張し，特徴欠損/雑音付加の伴うパターンを学習可能とすることで剛物体には発生しやすい隠蔽の課題に対処した．一般には特徴欠損/雑音付加の伴うパターンは指数関数的に多くなるが，特徴欠損/雑音付加の伴う画像

を生成せずに，学習画像のみを用いてモデル化を行うことができることを示した．

第3章では，景観画像中の文字候補群を用いた画像インデクシングおよび検索法を提案した．提案法では，従来法では検出困難なシェーディングを伴う文字，立体文字，複雑背景上の文字，傾斜文字，を文字候補として検出する能力がある．また，検索の際に検索文字列の空間的な配置関係を利用し，多くの偽文字候補の出現や文字候補の検出もれを許容して，実用的な画像検索を実現できる．漢字の単文字認識実験を行い，認識率が86.0%以上となることを確認した．また，景観画像のインデクシング実験を行い，漢字・ひらがな・カタカナの再現率が80.0%以上となることを確認した．作成されたインデクスを用いて，キーワードによる画像検索精度の測定実験を行い，平均再現率78.9%，平均適合率54.9%を確認した．

第4章では，画像をキーにした対象物の情報能動検索の一例として，対象のデザイン性を損なわずに，ユーザが直感的に携帯から情報にアクセスするインタフェースを実現する際のシステム構成について述べた．また，把持した3次元剛物体を撮影し，その姿勢を推定する課題を想定したシミュレーション実験を行い，提案法である縦横微分画像および不定次元を用いた部分空間法の妥当性を実証した．

実際にカメラ付携帯電話を用いた情報提供システムに提案法を適用し，実用上発生するユーザ，照明環境，周辺環境，および撮影機材による対象のパターン変化に頑健であることを示した．具体的には撮影位置と方向の変動，照明変動，ぼけ，シェーディング，照明光の全反射，隠蔽，背景変動，によるパターン変化に頑健であることを示した．同時に，平面パターンに対しては縦横微分画像，3次元剛物体に対しては方向のみ微分画像が有用であることを評価実験により示した．また，従来の部分空間法に比べて不定次元を用いた部分空間法の優位性を示した．さらに，平面パターン，立体物ともに複数の実験を行い，そのすべてにおいて85%以上の識別率を確認し，提案法の実用性，および汎用性の高さを実証した．

5.2 今後の検討課題

本研究により，標題技術が確立され実用化の目処がたったが，高性能化や適用範囲の拡大等に研究開発の余地がある．本研究成果の今後の発展のため，以下に

課題を列挙する。

- 適用範囲の検討

モデルとしてリッチなパターン特徴量を採用しているため対象依存性が少なく、さまざまな変化を伴う物体認識に容易に適用可能である。

- － 文字に関しては、韓国語、中国語などにも適用することにより、より広い画像検索を実現できる可能性がある。
- － 剛物体の角度パラメータの算出について、シミュレーションにより有効性を実証したが実物体には未適用である。適用実験を行うことで、適用範囲を広げることが可能である。例えば、角度パラメータの算出に特化した方法 [安藤 06] と組み合わせることで高速化、高性能化が期待できる。
- － 学習サンプルを収集可能であれば、原理的には、関節物体や柔軟物体への適用は可能である。実際に適用を試み、提案アルゴリズムがどの程度のパラメータ数に対応できるかを検討することで、本技術の適用領域が広がる可能性がある。

- 高速化、高精度化

文字認識においては、活字文字という特性を利用し、文字抽出の高速化を試みる [Lu07, Lu06] で応用面での高速化が期待できる。また、各文字の認識率を高める [And06] ことで、再現率、適合率を高めると同時に、OCRで利用される言語処理を導入することにより、各画像に書かれた文字列を特定する精度まで高性能化が期待できる。剛物体認識においては、類似カテゴリへの対処、2値化傾向の多様性への対処を解決することにより、高性能化が期待できる。

- 実用化

提案アルゴリズムをアプリケーションに応用する際には、設計の段階において、対象の登録数、原画像の次元数、計算時間、要求精度を決定する必要

がある．対象の登録数，原画像の次元数，計算時間，精度の間にはトレードオフの関係があり，その指針が必要となる．景観中文字の認識の知見から対象数の上限は3000程度まで，剛物体の認識の知見から原画像の次元数は40x40程度が良好なトレードオフを実現する，と考える．ただし，実験結果は評価パターンの偏りにより影響を受けるため，新しい対象に適用する場合は事前の実験が必須である．

また，アルゴリズム/利用特徴等を選択し，かつ様々なパラメータを調整する必要がある．調整者には高度な知識を要求することは出来ないため，ユーザインタフェースを整備し，アルゴリズム/特徴選択の自動化が必要となる．

最後に，対象認識技術の新領域への適用に際しては，要求精度は事前に判明しないため，市場に技術を出して実際のユーザに評価を受けることが重要である．本技術は現在，市場にて評価を受けている段階である [日本 b, 日本 a, NTT ア]．

謝辞

本研究を博士論文としてまとめるにあたり，懇切なご指導，ご鞭撻を賜りました奈良先端科学技術大学院大学横矢直和教授，千原國宏教授，山澤一誠准教授に謹んで深謝いたします．

本論文は，筆者が日本電信電話株式会社NTTサイバースペース研究所にて行った研究をまとめたものです．日本電信電話株式会社に在職しながら，奈良先端科学技術大学院大学の博士後期課程への入学という機会をいただき，研究をまとめることができました．現所属長である南部明所長，現上司である中尾昌善プロジェクトマネージャ，石橋聡プロジェクトマネージャ，小池秀樹ディレクタに感謝いたします．

本研究は着手から5年間という長い年月を経てまとめられました．様々な組織変更を経ており様々な方から助言をいただきました．研究内容について有益な御討論をいただいた有川知彦氏，荒川賢一氏，安野貴之博士，鈴木章博士，伊藤直己氏，安藤慎吾博士に感謝いたします．また，製品化に関して多大な貢献をいただいた宮地寿人氏，曾根岡昭直博士，上村弥生氏，林經正氏，市河研一氏，塩昭夫氏，戸井田徹氏，宮武隆之氏，田中久子氏に感謝いたします．論文の執筆過程で多くの助言をいただいた当時および現在の画像メディア通信プロジェクトの皆様感謝いたします．また，実験等に協力していただいた株式会社レッド・エンタテインメントの皆様，株式会社セガの皆様感謝いたします．また，快く素材の掲載許可をいただいた株式会社創通，株式会社サンライズに感謝いたします．

また，これまでの研究人生を支援，応援していただいた両親および兄に感謝いたします．

最後に多大なる苦勞をかけつつも温かい目で筆者を見守ってくれた妻 有理果，娘 葉音に感謝いたします．

参考文献

- [Aka73] Akaike, H.: Information theory and an extension of the maximum likelihood principle, *in Proc. of the Second Int'l Symposium on Information theory*, pp. 267–281, 1973.
- [And06] Ando, S., Y. Kusachi, A. Suzuki, K. Arakawa, and T. Yasuno: Kanji Recognition in Scene Images Using Distortion Parameter Estimation Based on Support Vector Regression, *in Proc. of the 8th IASTED Int'l Conf. on Signal and Image Processing*, No. 534-055, Aug. 2006.
- [Cha99] Chang, P. and J. Krumm: Object Recognition with color cooccurrence histograms, *in Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition*, Vol. 2, No. p.2498, 1999.
- [Che04] Chen, X. and A. L. Yuille: Detecting and reading text in natural scenes, *in Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 366–373, 2004.
- [Clo71] Clowes, M.: On seeing things, *Artificial Intelligence*, Vol. 2, No. 1, pp. 79–116, 1971.
- [Dra89] Draper, B., R. Collins, J. Brolio, A. Hanson, and E. Riseman: The Schema System, *Int'l Journal of Computer Vision*, Vol. 3, No. 2, pp. 209–250, 1989.
- [Har01] Haritaoglu, I.: InfoScope: Link from Real World to Digital Information Space, *in Proc. of the Int'l Conf. on Ubiquitous Computing*, pp. 247–255, 2001.
- [Kim97] Kimura, F., T. Wakabayashi, S. Tsuruoka, and Y. Miyake: Improvement of Handwritten Japanese Character Recognition Using Weighted Direction Code Histogram, *Pattern Recognition*, Vol. 30(8), pp. 1329–1337, 1997.

- [Kus04] Kusachi, Y., N. Ito, A. Suzuki, and K. Arakawa: Kanji Recognition in Scene Images without Detection of Text Fields, *in Proc. of the Int'l Conf. on Pattern Recognition*, Vol. 2, pp. 457–460, 2004.
- [Lie96] Lienhart, R.: Automatic Text Recognition for Video Indexing, *in Proc. ACM Multimedia*, pp. 11–20, 1996.
- [Lop00] Lopresti, D. and J. Zhou: Locating and Recognizing Text in WWW Images, *Information Retrieval*, Vol. 2, No. 2-3, pp. 177–206, 2000.
- [Lu06] Lu, Z., S. Ando, K. Kataoka, Y. Kusachi, A. Suzuki, Y. Takahashi, and T. Yasuno: Text Locating Algorithm for Scene Images Based on Gray-Scale Connected Component Analysis, *画像の認識・理解シンポジウム (MIRU2006)*, pp. 565–570, 2006.
- [Lu07] Lu, Z., S. Ando, K. Kataoka, Y. Kusachi, A. Suzuki, Y. Takahashi, and T. Yasuno: Text Locating and Verification Algorithm for Scene Images Based on Gray-Scale CC Analysis, *画像電子学会*, Vol. 36, No. 4, pp. 509–519, 2007.
- [Luc03] Lucas, S. M., A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young: ICDAR 2003 Robust Reading Competition, *in Proc. of the Int'l Conf. on Document Analysis and Recognition*, pp. 682–687, 2003.
- [Nag00] Nagy, G.: Twenty year of Document Image Analysis in PAMI, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 38–61, 2000.
- [NTT ア] NTT アイティ株式会社：物体認識技術を利用した携帯応用システム『MagicFinder』の販売を開始, <http://www.ntt-it.co.jp/press/2006/061004/061004bcj.html>.
- [Oht85] Ohta, Y.: *Knowledge-Based Interpretation of Outdoor Natural Color Scenes*, Pitman Advanced Publishing Program, Boston, 1985., ISBN:

0-273-08673-1.

- [Ohy94] Ohya, J., A. Shio, and A. Akamatsu: Recognition of characters in scene images, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 2, pp. 214–220, 1994.
- [Oja83] Oja, E.: *Subspace Methods of Pattern Recognitions*, Research Studies Press, 1983, ISBN: 0-863-80010-6.
- [Pop94] Pope, A.: Model-Based Object Recognition - A Survey of Recent Research, Technical Report TR-94-04, 1994.
- [Sat99] Sato, T., T. Kanade, E. Hughes, M. Smith, and S. Satoh: Video OCR: Indexing Digital News Libraries by Recognition of Superimposed Caption, in *Proc. of the ACM Multimedia Systems Special Issue on Video Libraries*, Vol. 7(5), pp. 385–395, 1999.
- [Sch98] Schneiderman, H. and T. Kanade: Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition, in *Proc. of Computer Vision and Pattern Recognition*, pp. 45–51, 1998.
- [Ten77] Tenenbaum, J. and H. Barrow: Experiments in Interpretation Guided Segmentation, *Artificial Intelligence*, Vol. 8, pp. 241–274, 1977.
- [Tur91] Turk, M. and A. Pentland: Face recognition using eigenfaces, in *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.
- [Vap95] Vapnik, V. N.: *The Nature of Statistical Learning Theory*, Springer-Verlag, 1995, ISBN: 0-387-94559-8.
- [Vio01] Viola, P. and M. Jones: Rapid object detection using a boosted cascade of simple features, in *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition*, pp. 511–518, 2001.

- [Wak01] Wakahara, T., Y. Kimura, and A. Tomono: Affine-Invariant Recognition of Gray-Scale Characters Using Global Affine Transformation Correlation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 4, pp. 384–395, 2001.
- [Yan01] Yang, J., J. Gao, Y. Zahng, and A. Waibel: Towards Automatic Sign Translation, in *Proc. of the Human Language Technology*, pp. 1–6, 2001.
- [Yos00] Yoshimura, H., M. Etoh, K. Kondo, and N. Yokoya: Grey-scale Character Recognition by Gabor Jets Projection, in *Proc. of the Int'l Conf. on Pattern Recognition*, Vol. 2, pp. 335–338, 2000.
- [Zah02] Zahng, J., X. Chen, A. Hanneman, J. Yang, and A. Waibel: A Robust Approach for Recognition of Text Embedded in Natural Scenes, in *Proc. of the Int'l Conf. on Pattern Recognition*, Vol. 3, pp. 204–207, 2002.
- [Zhu05] Zhu, K., F. Qi, R. Jiang, L. Xu, M. Kimachi, Y. Wu, and T. Aizawa: Using Adaboost to Detect and Segment Characters from Natural Scenes, in *Proc. of the Int'l Workshop on Camera-based Document Analysis and Recognition*, pp. 52–59, 2005.
- [デン] デンソーウェーブ: QR Code.com, <http://www.denso-wave.com/qrcode/>.
- [メデ 06] メディアシーク: 携帯電話のカメラをかざすだけで英単語を翻訳する携帯アプリ『カメラ辞典』の提供を開始!, <http://www.enfour.co.jp/news/2006083101.pdf>, 2006.
- [安藤 06] 安藤, 草地, 鈴木, 荒川: サポートベクトル回帰を用いた3次元物体の姿勢推定法, *電子情報通信学会論文誌*, Vol. J89-D, No. 8, pp. 1840–1847, Aug. 2006.

- [黄瀬 04] 黄瀬, 大町, 内田, 岩村: カメラを用いた文字認識・文書画像解析の現状と課題, 電子情報通信学会 技術研究報告, パターン認識・メディア理解, Vol. PRMU2004-246, No. 742, pp. 85–90, 2004.
- [岡 82] 岡: 図形からのセル特徴の抽出, 電子情報通信学会, Vol. J65-D, No. 10, pp. 1219–1226, 1982.
- [岡 83] 岡: セル特徴を用いた手書き漢字の認識, 電子情報通信学会, Vol. J66-D, No. 1, pp. 17–24, 1983.
- [高橋 84] 高橋, 内藤, 増田: 外郭構造情報を利用したストローク抽出法による手書き漢字認識, 電子情報通信学会, Vol. J67-D, No. 9, pp. 1052–1059, 1984.
- [高橋 04] 高橋, 田中, 鈴木, 塩, 大塚: 多値テンプレートマッチングを用いたナンバープレート認識法 雑音抑制フィルタと文字配置規則の効果的利用, 電子情報通信学会論文誌, Vol. J87-D-II, No. 7, pp. 1451–1461, 2004.
- [坂上 03a] 坂上, 尺長: 正規化固有空間への部分射影の幾つかの応用について, 情報処理学会研究報告, CVIM, Vol. 2003, No. 2, pp. 155–162, 2003.
- [坂上 03b] 坂上, 尺長: 正規化固有空間への部分射影問題の解法, 情報処理学会論文誌 (コンピュータビジョンとイメージメディア), Vol. 44, No. SIG_17, pp. 100–108, 2003.
- [山口 02] 山口, 福井: パターンハッシング: 部分画像と不変量索引を用いた分散アピアランスモデル, 情報処理学会論文誌 (コンピュータビジョンとイメージメディア), Vol. 44, No. SIG_5, pp. 64–73, 2002.
- [山田 84] 山田: 輪郭 DP 整合法と手書き漢字認識への適用, 電子情報通信学会, Vol. J67-D, No. 3, pp. 351–358, 1984.

- [山本 73] 山本, 中島, 中田: 階層的パターンマッチングによる漢字認識の基礎 - 印刷漢字認識の研究 -, 電子情報通信学会, Vol. 56-D, No. 12, pp. 714-721, 1973.
- [山本 82] 山本: 弛緩整合法による手書教育漢字認識, 電子情報通信学会, Vol. J65-D, No. 9, pp. 1167-1174, 1982.
- [新井 00] 新井, 桑野, 倉掛, 杉村: 映像中のテロップ表示フレーム検出方法, 電子情報通信学会論文誌, Vol. J83-D-II, No. 6, pp. 1477-1486, 2000.
- [森 01] 森, 澤木: 低品質文字の認識手法とその応用に関するサーベイ, 電子情報通信学会技術研究報告, パターン認識・メディア理解, Vol. PRMU2001-275, No. 713, pp. 25-40, 2001.
- [草地 98] 草地, 山澤, 竹村, 横矢: 動画像中の歩行者追跡による映像入れ換え合成, 映像情報メディア学会, Vol. 52, No. 9, pp. 1357-1365, 1998.
- [草地 00] 草地, 加藤, 奥平: 自動倣い作業ロボットシステムのための断面形状認識, 日本ロボット学会誌, Vol. 18, No. 6, pp. 882-889, 2000.
- [村瀬 94] 村瀬, シュリー: 2次元照合による3次元物体認識 - パラメトリック固有空間法 -, 電子情報通信学会論文誌, Vol. J77-D-II, No. 11, pp. 2179-2187, 1994.
- [村瀬 95] 村瀬, シュリー: 多重解像度と固有空間表現による3次元物体のイメージスポッティング, 情報処理学会論文誌(コンピュータビジョンとイメージメディア), Vol. 36, No. 10, pp. 2234-2243, 1995.
- [村瀬 05] 村瀬: 画像認識のための生成型学習, 情報処理学会論文誌(コンピュータビジョンとイメージメディア), Vol. 46, No. SIG_15(CVIM12), pp. 35-42, 2005.
- [大場 97] 大場, 池内: 局所固有空間手法による金属物体の安定認識, 電子情報通信学会論文誌, Vol. J80-D-II, No. 12, pp. 3147-3154, 1997.

- [中村 04] 中村, 片山, 山室, 曾根原: カメラ付携帯電話機を用いたアナログ画像からの高速電子透かし検出方式, 電子情報通信学会論文誌, Vol. J87-D-II, No. 12, pp. 2145–2155, 2004.
- [津雲 88] 津雲, 浅井: 文字認識技術の最近の動向, 電子情報通信学会技術研究報告, 画像工学研究会, Vol. IE88-5, No. 14, pp. 31–38, 1988.
- [鶴岡 87] 鶴岡, 粟田, 原田, 木村, 三宅: 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, Vol. J70-D, No. 7, pp. 1390–1397, 1987.
- [田中 86] 田中, 青田, 塩野, 真田, 手塚: 手書き漢字の部分パターン抽出の一手法, 電子情報通信学会, Vol. J69-D, No. 6, pp. 911–920, 1986.
- [土橋 02] 土橋, 岡崎, 高木: 顔照合セキュリティシステム *FacePass_{TM}*, 東芝レビュー, Vol. 57, No. 8, pp. 48–51, 2002.
- [日本 a] 日本電信電話株式会社, MTV ジャパン株式会社: ユビキタス・アクセス新技術を活用したプロモーションを開始ケータイで撮影した映像がそのままリンクになる!, <http://www.ntt.co.jp/news/news07/0703/070319a.html>.
- [日本 b] 日本電信電話株式会社, 株式会社セガ: 実世界の2次元 / 3次元対象物とネット上のサービスを連携できるオブジェクト認識技術の共同実験開始について, <http://www.ntt.co.jp/news/news05/0512/051213.html>.
- [梅田 94a] 梅田: 手書き漢字認識手法の個性的なサーベイ, 手書き文字認識技術の過去・現在・未来シンポジウム公演論文集, pp. 54–63, 1994.
- [梅田 94b] 梅田, 佐瀬: 高速日本語文字読み取り装置 N6370R モデル70, NEC技報, Vol. 45, No. 3, pp. 51–56, 1994.
- [目黒 84] 目黒, 梅田: マルチフォント印刷漢字認識装置, 電子情報通信学会, Vol. J67-D, No. 8, pp. 908–915, 1984.

研究業績

学術論文誌

1. Z. Lu, S. Ando, K. Kataoka, Y. Kusachi, A. Suzuki, Y. Takahashi, T. Yasuno, "Text Locating and Verification Algorithm for Scene Images Based on Gray-Scale CC Analysis," 画像電子学会誌, vol.36, No.4, pp.509-519, 2007
2. 草地良規, 鈴木章, 伊藤直己, 荒川賢一, 安野貴之, "景観画像中の文字候補群による画像インデクシング/検索法," 電子情報通信学会論文誌, Vol.J90-D, No.9, pp.2562-2572, Sept. 2007. (本論文の2, 3章に関連)
3. 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, "サポートベクトル回帰を用いた3次元物体の姿勢推定法," 電子情報通信学会論文誌, Vol.J89-D, No.8, pp.1840-1847, Aug. 2006.
4. 草地良規, 加藤晃市, 奥平雅士, "自動倣い作業ロボットシステムのための断面形状認識," 日本ロボット学会誌, vol.18, No.6, pp.882-889, Sept. 2000.
5. 草地良規, 山澤一誠, 竹村治雄, 横矢直和, "動画像中の歩行者追跡による映像入れ換え合成," 映像情報メディア学会, Vol.52, No.9, pp.1357-1365, Sept. 1998.

国際会議

1. Y. Kusachi, A. Suzuki, K. Arakawa, T. Yasuno, "Object Recognition Using Don't Care Dimensions and Its Application to Information Retrieval System on Cellular Phone," in *Proc. of the 8th IASTED Int'l Conf. on Signal and Image Processing (SIP2006)*, 534-077, Aug. 2006. (本論文の2, 4章に関連)
2. S. Ando, Y. Kusachi, A. Suzuki, K. Arakawa, T. Yasuno, "Kanji Recognition in Scene Images Using Distortion Parameter Estimation Based on

Support Vector Regression,” in *Proc. of the 8th IASTED Int’l Conf. on Signal and Image Processing (SIP2006)*, 534-055, Aug. 2006.

3. S. Ando , Y. Kusachi , A. Suzuki , K. Arakawa, “Appearance-based Pose Estimation of 3D Object Using Support Vector Regression,” in *Proc. of the IEEE Int’l Conf. on Image processing (ICIP 2005)*, I-341-4, Sept. 2005.
4. Y. Kusachi, N. Ito, A. Suzuki, K. Arakawa, “Kanji Recognition in Scene Images without Detection of Text Fields,” in *Proc. of the 17th Int’l Conf. on Pattern Recognition (ICPR2004)*, Vol. 2, pp. 457-460, 2004. (本論文の 2 , 3 章に関連)
5. Y. Kusachi , K. Kato , M. Okudaira, “Tracking a Feature Point in Profile Data Using an Articulated Object Model and Tuning Parameters Using the Genetic Algorithm,” in *Proc. of the IEEE Int’l Conf. on Emerging Technologies and Factory Automation (ETF A1999)*, pp.145-150, Oct, 1999.

研究会・シンポジウム

1. 安藤慎吾 , 片岡香織 , 草地良規 , 鈴木章 , 安野貴之 , “物体検出・識別のための複数基準による絞込み方式,” 画像電子学会モバイルイメージ研究会 2 , pp.5-8, Oct. 2006.
2. 安藤慎吾 , 草地良規 , 鈴木章 , 荒川賢一 , 安野貴之 , “サポートベクトル回帰による変形パラメータ推定を利用した景観中文字認識の高精度化,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU106-291) , pp47-52, May, 2006.
3. Z. Lu , S. Ando , K. Kataoka , Y. Kusachi , A. Suzuki , Y. Takahashi , T. Yasuno, “Text Locating Algorithm for Scene Images Based on Gray-Scale Connected Component Analysis,” 画像の認識・理解シンポジウム (MIRU2006), pp.565-570, 2006.

4. 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “サポートベクトル回帰による3次元物体の姿勢推定,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU104-291), pp.9-14, Oct. 2004.
5. 草地良規, 伊藤直己, 鈴木章, 荒川賢一, “画像インデクシングを目的としたテキスト領域抽出不要の景観中文字認識 ~ 階層的認識辞書の適用による処理量の削減 ~,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU104-291), pp.37-42, Oct. 2004. (本論文の2, 3章に関連)
6. 草地良規, 鈴木章, 荒川賢一, “局所的な隠蔽, 照明反射, 複雑背景に対処するアピアランスベースの3次元物体認識,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU104-290), pp.137-144, Sept. 2004. (本論文の2, 4章に関連)
7. 草地良規, 加藤晃市, 奥平雅士, “関節物体モデルを用いた断面形状認識,” 第5回画像センシングシンポジウム (SSI) 講演論文集, pp.33-38, Aug. 1999.
8. 草地良規, 加藤晃市, 奥平雅士, “関節物体モデルを用いた特徴点認識とGAによるパラメータチューニングの検討: 溶接ロボットへの適用,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU98-609), pp.59-66, Feb. 1998.
9. 草地良規, 山澤一誠, 竹村治雄, 横矢直和, “段階的なモデルマッチングによる動画像における歩行者の追跡と映像合成への応用,” 電子情報通信学会 技術研究報告, パターン認識・メディア理解研究会 (PRMU96-435), pp.41-48, Aug. 1996.

全国大会

1. 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “局所パターンの統計的性質に着目した複雑背景下の非剛体物体認識,” 画像電子学会年次大会, 2005年6月.

2. 草地良規，加藤晃市，奥平雅士，“断面形状変化に適応する関節物体モデルを用いた特徴点認識,” 電子情報通信学会，情報・システムソサエティ大会，pp.276, 1998.
3. 草地良規，山澤一誠，竹村治雄，横矢直和，“段階的なモデルマッチングを用いた人間歩行の解析と画像合成への応用,” 電子情報通信学会，情報・システムソサエティ大会，pp.427, 1996.

解説記事

1. 上村弥生，宮地寿人，草地良規，“オブジェクト認識技術のプロデュース,” NTT 技術ジャーナル， pp.18-21, vol.19, No.3, Mar. 2007.

報道発表

1. 日本電信電話株式会社，株式会社セガ，“実世界の2次元 / 3次元対象物とネット上のサービスを連携できるオブジェクト認識技術の共同実験開始について,” <http://www.ntt.co.jp/news/news05/0512/051213.html>， 2005.
2. 日本電信電話株式会社，MTV ジャパン株式会社，“ユビキタス・アクセス新技術を活用したプロモーションを開始．ケータイで撮影した映像がそのままリンクになる!,” <http://www.ntt.co.jp/news/news07/0703/070319a.html>， 2006.

メディア

1. “TVからサイトへ誘導,” 日経産業新聞，2007年5月21日
2. “カメラ携帯使い動画認識の実験,” 日経産業新聞，2007年3月20日
3. “広告サービス新手法で実験,” 日刊工業新聞，2007年3月20日
4. “携帯で動画認識の実験,” 日本経済新聞，2007年3月19日

5. “パチリで飛びます！” WBSトレンドたまご，2005年12月19日
6. “立体物画像から接続,” 日経産業新聞，2005年12月14日

展示会（公開分のみ）

1. “オブジェクト認識PF：OBJe,” NTT R&D フォーラム 2007，2007年2月24日～2月25日
2. “キャラクタースタンド,” NTT グループコミュニケーション EXPO，2005年12月20日～12月22日
3. “携帯カメラによる実世界言語インタフェース,” NTT R&D フォーラム 2005，2005年2月24日～2月25日

特許

1. 目黒眞一，表勇三，大野一英，梅野光，高杉欣光，神好人，柴田央人，遠藤公誉，中山丈二，草地良規，太田崇博，“物品情報管理方法及び物品情報管理装置並びに物品情報管理プログラムを記録した記録媒体,” 特許 3872439
2. 草地良規，加藤晃市，手塚博久，“画像認識システム性向上方法、装置、および画像認識システム性能向上プログラムを記録した記録媒体,” 特許 3545624
3. 草地良規，加藤晃市，“遠隔監視方法、装置、および遠隔監視プログラムを記録した記録媒体,” 特許 3538112
4. 加藤晃市，草地良規，奥平雅士，“2次元走査型レンジセンサ投光器走査方法及びシステム装置並びに2次元走査型レンジセンサ投光器走査プログラムを記録したコンピュータで読取り可能な記録媒体,” 特許 3511474
5. 草地良規，加藤晃市，“形状認識装置及び方法並びに形状認識プログラムを記録した記録媒体,” 特許 3501647

6. 目黒眞一，表勇三，大野一英，梅野光，高杉欣光，神好人，柴田央人，遠藤公誉，中山丈二，草地良規，太田崇博，“物品情報管理装置及び物品情報管理プログラムを記録した記録媒体,” 特許 3451245

公開特許

1. 安藤慎吾，草地良規，鈴木章，荒川賢一，“パターン認識装置、パターン認識方法，パターン認識プログラム，及びパターン認識プログラムの記録媒体,” 特許公開 2006 - 252504
2. 草地良規，鈴木章，荒川賢一，安藤慎吾，“文字認識装置，文字認識方法，及び文字認識プログラムの記録媒体,” 特許公開 2006 - 251920（本論文の 2, 3 章に関連）
3. 草地良規，鈴木章，荒川賢一，安藤慎吾，“文字列撮影補助装置，文字列撮影補助方法，及び文字列撮影補助プログラム,” 特許公開 2006 - 235818
4. 安藤慎吾，草地良規，鈴木章，荒川賢一，“文字認識装置，文字認識方法，及び文字認識プログラムの記録媒体,” 特許公開 2006 - 235817
5. 鈴木章，草地良規，安藤慎吾，荒川賢一，“文字列探索装置，探索方法およびこの方法のプログラム,” 特許公開 2006 - 106931
6. 安藤慎吾，草地良規，鈴木章，荒川賢一，“パラメータ推定装置，パラメータ推定方法，パラメータ推定プログラム，および，パラメータ推定プログラムの記録媒体,” 特許公開 2006 - 079321
7. 草地良規，鈴木章，荒川賢一，“物体識別装置，物体識別方法，およびこの方法のプログラム,” 特許公開 2006 - 059284（本論文の 2,4 章に関連）
8. 草地良規，鈴木章，荒川賢一，“辞書構成装置，辞書構成方法，および，辞書構成プログラム,” 特許公開 2006 - 059018（本論文の 2 章に関連）

9. 草地良規, 鈴木章, 伊藤直己, 安藤慎吾, 荒川賢一, “特徴抽出装置, 特徴抽出方法, および, プログラム,” 特許公開 2005 - 339398 (本論文の2,3章に関連)
10. 草地良規, 鈴木章, 荒川賢一, “物体識別のための辞書生成方法とその装置, プログラム及び記録媒体,” 特許公開 2005 - 339397 (本論文の2,4章に関連)
11. 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “パターン認識学習装置と識別装置, パターン認識学習処理方法と識別処理方法及びパターン認識プログラム並びにそのプログラムを記録した記録媒体,” 特許公開 2005 - 339186
12. 伊藤直己, 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “オブジェクト領域抽出装置, 抽出方法, この方法のプログラム及びこのプログラムを記録した記録媒体,” 特許公開 2005 - 339076
13. 伊藤直己, 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “画像探索装置, 画像探索方法, この方法のプログラム及びこのプログラムを記録した記録媒体,” 特許公開 2005 - 339075
14. 安藤慎吾, 草地良規, 鈴木章, 荒川賢一, “ナビゲーション方法とその装置, プログラム及び記録媒体,” 特許公開 2005 - 331737
15. 片山淳, 中村高雄, 草地良規, 鈴木章, 荒川賢一, “カメラ付き携帯端末レンズ歪曲歪補正方法及び装置及びカメラ付き携帯端末レンズ歪曲歪補正システム及びカメラ付き携帯端末レンズ歪曲歪補正プログラム,” 特許公開 2005 - 182515
16. 草地良規, 鈴木章, 荒川賢一, “物体パラメータ推定装置, 物体パラメータ推定方法, プログラムおよび記録媒体,” 特許公開 2005 - 149167 (本論文の2,4章に関連)
17. 安藤慎吾, 鈴木章, 荒川賢一, 草地良規, “パラメータ推定方法とその装置, プログラム及び記録媒体,” 特許公開 2005 - 148988

18. 草地良規，鈴木章，伊藤直己，荒川賢一，安藤慎吾，“オブジェクト識別方法とその装置，プログラム及び記録媒体，” 特許公開 2005 - 148987（本論文の 2,3 章に関連）
19. 伊藤直己，草地良規，鈴木章，荒川賢一，“画像処理方法及びその装置，プログラム並びにこの記録媒体，” 特許公開 2005 - 141458
20. 伊藤直己，安藤慎吾，草地良規，鈴木章，荒川賢一，“画像処理装置，画像処理方法，プログラム，および，記録媒体，” 特許公開 2005 - 134949
21. 草地良規，鈴木章，荒川賢一，“パターン認識装置，パターン認識方法，パターン認識プログラムおよびパターン認識プログラムを記録した記録媒体，” 特許公開 2005 - 115432（本論文の 2,4 章に関連）
22. 草地良規，鈴木章，伊藤直己，荒川賢一，“オブジェクト識別装置，識別方法，この方法のプログラム，およびこのプログラムを記録した記録媒体，” 特許公開 2005 - 078555（本論文の 2,3 章に関連）
23. 伊藤直己，鈴木章，杵淵哲也，草地良規，有川知彦，荒川賢一，“データ保存装置とその方法，及びそのプログラムとそのプログラムを記録した記録媒体，” 特許公開 2004 - 363871
24. 杵淵哲也，伊藤直己，草地良規，鈴木章，荒川賢一，有川知彦，“文字・図形領域抽出装置，抽出方法，この方法のプログラムおよびこのプログラムを記録した記録媒体，” 特許公開 2004 - 355184
25. 草地良規，鈴木章，杵淵哲也，荒川賢一，伊藤直己，有川知彦，“オブジェクト学習装置と方法，オブジェクト識別装置と方法，およびこれらのプログラムとこれらのプログラムを記録した記録媒体，” 特許公開 2004 - 355183（本論文の 2 章に関連）
26. 草地良規，杵淵哲也，鈴木章，荒川賢一，有川知彦，伊藤直己，“物体識別装置，方法，プログラムおよびプログラムを記録した媒体，” 特許公開 2004 - 318627（本論文の 2 章に関連）

27. 草地良規，鈴木章，杵渕哲也，荒川賢一，伊藤直己，有川知彦，“オブジェクト学習装置，オブジェクト識別装置，これらの方法，これらのプログラムおよびこれらのプログラムを記録した媒体，” 特許公開 2004 - 192310（本論文の 2 章に関連）
28. 草地良規，鈴木章，荒川賢一，杵渕哲也，伊藤直己，有川知彦，“広告情報提供方法，装置，広告情報提供プログラムおよびこのプログラムを記録した記録媒体，” 特許公開 2004 - 185591（本論文の 4 章に関連）
29. 伊藤直己，鈴木章，草地良規，杵渕哲也，“情報表示方法，装置，プログラム，および記録媒体，” 特許公開 2004 - 178297
30. 草地良規，鈴木章，杵渕哲也，荒川賢一，伊藤直己，有川知彦，“物体識別装置，物体識別方法，この方法のプログラム，およびこのプログラムを記録した記録媒体，” 特許公開 2004 - 145818（本論文の 2,3 章に関連）
31. 杵渕哲也，伊藤直己，草地良規，鈴木章，“画像認識用サーバ，画像認識用携帯端末装置，画像認識方法，画像認識用プログラム，画像認識用プログラムを記録した記録媒体，” 特許公開 2004 - 145416
32. 草地良規，鈴木章，“物体識別機能付き画像撮影・蓄積・検索方法，装置，プログラム，および該プログラムを記録した記録媒体，” 特許公開 2004 - 070717
33. 草地良規，数藤恭子，“プログラムモジュール流通方法，プログラムモジュール流通装置，プログラムモジュール流通用プログラムおよびプログラムモジュール流通用プログラムを記録した記録媒体，” 特許公開 2004 - 038745
34. 草地良規，数藤恭子，高橋裕子，奥平雅士，“画像認識アルゴリズム生成方法，画像認識アルゴリズム生成装置，画像認識アルゴリズム生成プログラムおよび画像認識アルゴリズム生成プログラムを記録した記録媒体，” 特許公開 2004 - 038744

35. 鈴木章，伊藤直己，杵淵哲也，草地良規，“画像処理装置及びその方法と，画像処理プログラム及びそのプログラムを記録した記録媒体，” 特許公開 2003 - 323616
36. 目黒眞一，表勇三，大野一英，梅野光，高杉欣光，神好人，柴田央人，遠藤公誉，中山丈二，草地良規，太田崇博，“物品情報管理方法及び物品情報管理装置並びに物品情報管理プログラムを記録した記録媒体，” 特許公開 2003 - 323472
37. 太田崇博，草地良規，遠藤公誉，“負荷分散システム，” 特許公開 2002 - 245018
38. 目黒眞一，表勇三，大野一英，梅野光，高杉欣光，神好人，柴田央人，遠藤公誉，中山丈二，草地良規，太田崇博，“物品情報管理方法及び物品情報管理装置並びに物品情報管理プログラムを記録した記録媒体，” 特許公開 2002 - 149804
39. 草地良規，加藤晃市，“遠隔監視方法，装置，および遠隔監視プログラムを記録した記録媒体，” 特許公開 2001 - 268657
40. 草地良規，加藤晃市，“形状モデル及び形状認識方法及び形状認識装置及び形状認識プログラムを格納した記憶媒体，” 特許公開 2000 - 353245
41. 草地良規，加藤晃市，手塚博久，“システム性能向上方法，装置，およびシステム性能向上プログラムを記録した記録媒体，” 特許公開 2000 - 207376
42. 加藤晃市，草地良規，奥平雅士，“2次元走査型レンジセンサ投光器走査方法及びシステム装置並びに2次元走査型レンジセンサ投光器走査プログラムを記録した記録媒体，” 特許公開 2000 - 180132
43. 草地良規，加藤晃市，“形状認識装置及び方法並びに形状認識プログラムを記録した記録媒体，” 特許公開平 11 - 250257

付録

部分空間法は、識別対象の学習パターンが複数ある場合に、学習パターンの分布をモデル化し、入力ベクトルとモデルの距離を計算して、入力ベクトルのカテゴリを識別する方法である [Oja83]。部分空間法には、CLAFIC 法と投影距離法があり、その違いは、後者が主成分分析を行う前に学習パターンの平均を減算することである。ここでは、投影距離法についてのみ説明する。投影距離法では、カテゴリ c の学習パターンの共分散行列 Σ_c を求め、主成分分析する。共分散行列は、以下の式で定義される。

$$\Sigma_c = \frac{1}{N-1} \sum_{\vec{v} \in V} (\vec{v}_n - \tilde{v})(\vec{v}_n - \tilde{v})^t \quad (16)$$

ただし、カテゴリ c の学習パターンをベクトル \vec{v}_n 、 V は \vec{v}_n の集合、 n は学習パターン番号であり、学習パターン数を N として、 $1 \leq n \leq N$ である。また、 \vec{v}_n の平均をベクトル \tilde{v} で表す。次に、求められた共分散行列 Σ_c を主成分分析し、固有値の大きさ順に主成分ベクトル $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k, \dots, \vec{u}_K$ を得る。ただし、 K は学習パターンの次元数であり、 $1 \leq k \leq K$ である。上記主成分を用いて入力ベクトルを識別する。カテゴリ c のモデルと入力ベクトル \vec{x} の距離を以下の式で定義する。

$$l_c = |\vec{x} - \tilde{v}|^2 - \sum_{d=1}^{K'} \{\vec{u}_d \cdot (\vec{x} - \tilde{v})\}^2 \quad (17)$$

ただし、 K' は固有値の寄与率により算出するか、システム設計者により定められる正整数であり、 $K' < K$ である。距離 l_c が一番小さいカテゴリ c を求め、入力ベクトル \vec{x} をそのカテゴリと判定する。画像認識においては、画像を 1 ベクトルとみなすことで部分空間法を適用できる。部分空間法は、特徴次元を圧縮し、計算コストを削減できるという利点と、パターン分布を超平面でモデル化するため学習パターン同士の間を補間できるという利点がある。