

画像認識用データ圧縮空間の自動構成方式

安藤 祐樹 渡辺 俊典 菅原 研
電気通信大学 大学院 情報システム学研究科

概要

本研究室では昨年、新たな画像認識手法として、画像内の認識対象の特徴を多次元直方体で表現し、複数の直方体同士間の相関量によって類似物を認識する手法 FRCM を提案した。今回、この認識手法において基本的役割を果たす認識用特徴量空間の自動生成方式を開発した。これによりシステム構築コストの削減及び認識率の向上が確認された。また、輝度値の正規化を行うことで照明条件の変化に対して認識率の向上が見られた。

Automatic Construction of

Data Compression Feature Space for Image Recognition

Masaki Ando Toshinori Watanabe Ken Sugawara
The Graduate School of Information Systems,
The University of Electro-Communications

Abstract

As a new image recognition method, we proposed a new method FRCM that expresses the feature of a to-be-recognized object by a multi-dimensional rectangle and recognizes it by using its similarities to pre-instructed objects' feature rectangles. In this recognition technique, we newly developed an automatic feature space construction method. Reduction of systems configuration cost and improvement in the accuracy of recognition were ascertained by this method. Moreover, improvement in the accuracy of recognition under changes of lighting conditions through the normalization of a luminosity value is illustrated also.

1 はじめに

2次元の画像からもとの3次元情報を導出することは一般には不可能である。しかし、われわれ人間は、目で捉えた画像から外界の情報を把握することができる。これは、左右2つの眼による立体視、自分と外界の相対運動による像の動き、射影による像の歪みや明暗変化など様々な手掛かりを用いていると考えられる。このような機能の原理的な解明と効率よいアルゴリズムの開発の為に研究が活発に行われている[1][2]。この分野は一般に、単眼視とステレオ視に基づくものとに大別される[3][4][5]。特に単眼画像からの認識では、対象画像から得られる情報に制限があるため、多くの解決しなければならない問題があるが、その反面、汎用性や画像を得るためのコストなどを考えると非常に有用なテーマであるといえる。

2 関連研究

表1に単眼画像を用いた認識手法を示す[1]～[7]。画像認識は原理的には事前登録したモデルと未知画像とのマッチングを基礎としているが、柔軟性の高い認識手法を実現するには次の諸課題の解決が重要であると考えられる。

- (1) 認識に先立つオブジェクトの切り出しを行わない
- (2) 画像を直接教示に用いることができる
- (3) 少量の教示で多様な画像に対応できる

表1 単眼画像の認識手法

方式	内容	システム構築コスト	計算コスト	多様なシーン
グラフ ツリー 知識ベース	モデル用意 切り出し マッチング	高い	高い	低い
Shape-from- Shading	陰影情報から曲率抽出	低い	低～中	低い
事例ベース (FRCM)	事例教示 類似検索	低～中	低～中	中～高

これらの課題を解決する為に我々は、オブジェクトを多次元特徴量直方体で表現し、これの相関関係を利用したオブジェクト認識手法 FRCM(Feature Rectangle Correlation Method)[6]を提案した。

2.1 関連研究の概要

我々は、単眼モノクロ画像による事例ベースに基づくオブジェクト認識システムの作成を行ってきた。オブジェクトの切り出しを行わないようにする為に、まず、画像を局所領域に分割する。ここで得られた局所領域を最小全域木を利用してテキスト化を行い[8]、そのテキストを処理して特徴量を求めていく[6][9][10]。このときの処理には階層文法データ圧縮法を用い、局所画像から選ばれた基底画像によって得られる基底生成文法[6][11]を用いて処理を行う事により、圧縮率、基底生成文法の利用頻度、基底生成文法の標準偏差を特徴量として算出する。この特徴量からオブジェクトの多次元特徴量直方体を作成する。また、画像を直接教示として用いる為に、既知オブジェクト画像からあらかじめ多次元特徴量直方体を作成しておき、これを教示として登録しておく。これと未知のオブジェクトから作成した多次元特徴量直方体の相関関係を階層的に比較することで認識を行う。これにより、少量の教示で認識を行うことが出来るようになった。この認識手法について詳しく説明を加える。

2.2 多次元特徴量直方体による画像表現

2.2.1 画像の局所領域化

対象画像を認識するための前処理として画像全体を一定サイズの局所画像に分割し(図1)、各局所領域ごとに処理を行っていく。複雑なシーンでも局所領域には繰り返し性が高いことから、少数の教示情報で多くの未知情報を処理することができると期待される。

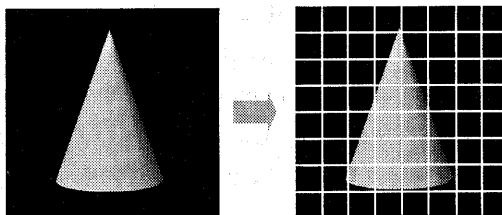


図1 画像の局所領域化

2.2.2 MSTによる局所領域の1次元符号化

画像は、その各画素を接点、近隣の画素間を枝と見なことでグラフ表現が可能となる。ここでは、各枝に隣接接点間の輝度の差(>0)を重みとして与える。局所領域に対して最小全域木(MST=Minimum Spanning Tree)化を行い、MST をトラバースすることで1次元テキストを作成する[6][8][9][10]。なお、本研究では陰影画像内の面の形状、陰影の変化などの特徴情報をより的確に表現するために8方向MSTを用いて1次元テキスト化した(図2)。

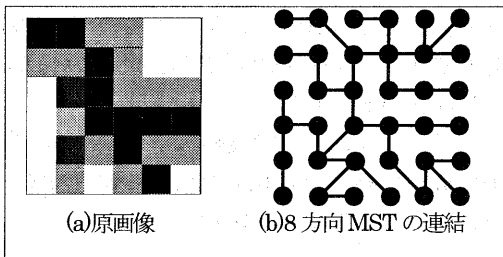


図2 8方向MSTの生成

2.2.3 階層文法抽出データ圧縮法

特徴量を求めるため、得られた1次元テキストを圧縮する。ここでは階層文法抽出データ圧縮法を用いた[11]。この圧縮法はテキスト内に繰り返し現れるパターンを文法形式で抽出し、得られた文法を用いてテキストを圧縮表現するものである。与えたテキストが圧縮されると同時に、そのテキストの内部構造を強く反映した生成文法の集合(階層構造をもっている)が得られる。

2.2.2 特徴ベクトルの生成

図3に特徴量ベクトル生成の流れを示す。まず初めに、規範として利用する幾つかの局所領域を与え、それを基底画像と呼ぶ。基底画像から抽出したテキストを基底データと呼ぶ(図3の $T_1, T_2, \dots, T_n$)。基底データをこの圧縮法で処理すると基底データ毎に生成文法集合(図3の $R_1, R_2, \dots, R_n$)が得られる。これを基底生成文法ルール集合と呼ぶ。

次に、未知の局所領域から作成した入力テキストXをこの基底生成文法ルール集合内の文法群を用いて圧縮し、以下の3つの特徴量を算出する。

- 文法を用いた圧縮率 (C)
- 文法の重み付き使用頻度 (S)
- 使用頻度の標準偏差 ($S.D$)

このうち S はつぎの方法で求める。各ルール集合についてその要素文法ルールが使用された回数の多い順に重み付けを行い、それを基底生成要素文法ルールの順に並び替える。生成された階層文法数 n 、 i 番目のルールの重みを W_i 、 i 番目のルールの使用回数 C_i とすると、基底生成文法の利用頻度 S は、

$$S = \sum_{i=1}^n (W_i \times C_i) \quad (1)$$

と求め、これを用いて S の標準偏差 $S.D$ を求める。また圧縮率 C は、圧縮後のテキスト長を圧縮前のテキスト長で割った値とする。

2.2.3 多次元特徴量直方体による表現

(画像内矩形領域の表現)

2.2.2 で得られた特徴量ベクトルと画像平面を用い、多次元のオブジェクト検索空間を生成する(図4)。この空間内であらかじめ登録しておいたオブジェクトの多

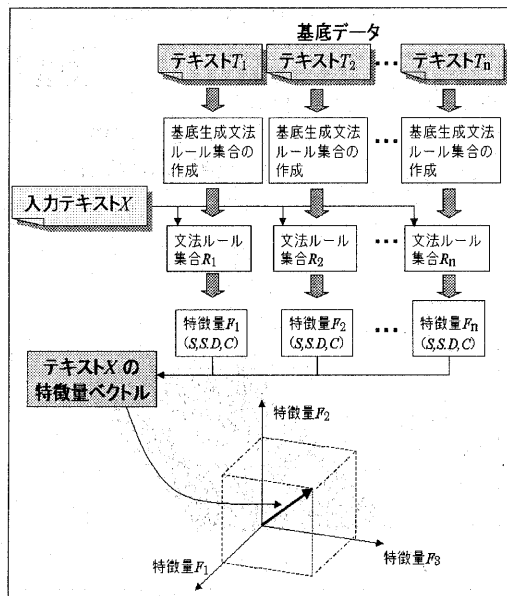


図3 特徴量ベクトルの生成

次元直方体と未知の入力オブジェクトの多次元直方体とを相互比較することにより入力オブジェクトの認識を行う。まず画像内の局所領域の集合(矩形領域)において、これに含まれる局所領域の特徴量ベクトルの最大値と最小値を求める。この得られた多次元領域の積を求めることで矩形領域の多次元特徴量直方体表現が可能である。

2.2.4 検索空間内での

多次元直方体の類似度

登録オブジェクトの多次元特徴量直方体を V_{case} 、入力オブジェクトの多次元特徴量直方体を V_{in} とするとこれらの相関関係を表す値 FRC (Feature Rectangle Correlation) を式(2)のように定義する。

$$FRC_{case,in} = -\log \left(\frac{V_{case} \cap V_{in}}{V_{case} \cup V_{in}} \right) \quad (2)$$

FRC 値が小さいほど特徴量空間内における登録オブジェクトの多次元直方体内に、入力オブジェクトの特徴量直方体が含まれている重複度が高いことになる。このことを利用して、未知入力オブジェクトと複数の登録オブジェクトとの間の FRC 値を求めていき、FRC 値が

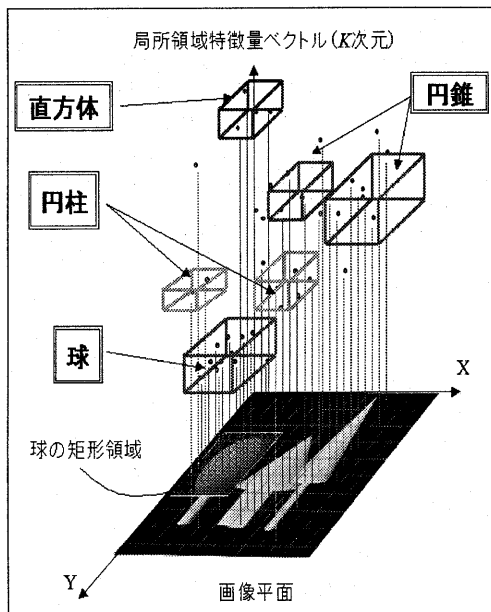


図4 オブジェクト検索空間

最小となった時の登録オブジェクトが未知入力オブジェクトと最も似ていると考えることが出来る。

3.1 関連研究の課題と解決法

3.1.1 輝度値

従来法では、局所領域を1次元符号化する際に、現画像の輝度値をそのまま等間隔に分類した量子化値を用いていた。しかし、これでは光源の強さや方向の変化による影響を強く受けってしまうと思われる。

そこで本研究では、輝度値を正規化することによって光源による影響を回避できるようにする。1次元符号化のとき輝度値を8段階の符号語に分類するが、その分類を等間隔に行うのではなく、正規化を行ってから広範囲に符号語を使用する。

まず、オブジェクトの輝度値の最低値($wmin$)と最高値($wmax$)を求めておく。画素の輝度値を w_{xy} とすると、正規化された輝度値 w'_{xy} は次のようになる。

$$w'_{xy} = \begin{cases} \frac{(w_{xy} - wmin) \times (255 - 32)}{wmax - wmin} + 32 & 0 < w_{xy} \leq 255 \\ 0 & (w_{xy} = 0) \end{cases} \quad (3)$$

この正規化により、オブジェクト部分の符号を広範囲に使うことができるようになる。よって、従来法よりもオブジェクトの陰影情報をより明確に表現することが可能になる。

3.1.2 特徴量の見極め

認識に用いることの出来る特徴量は圧縮率、利用頻度、利用頻度の標準偏差である。これは、基底から得られる特徴量をすべて用いる事によって認識能力が向上すると思えた為である。しかし、実際にすべての特徴量を使い、FRC 値を求めることが認識率を高めることになるのかわかっていない。

そこで、本研究では、有効な特徴量がどれなのかを認識実験を行い比較することで決定していく。事例を用いて多様な未知オブジェクトを認識するという本システムの特性から、出来るだけ少数の事例で多様な未知入力を正しく認識できることが望まれる。なお、比較する際には基底数と事例数とを一定にすることが必要となる。

これらが異なると認識能力は当然異なってくる。よって、基底画像及び事例画像を固定して、認識能力を比較することで、どの特徴量が認識に重要なのかを決定する。

3.1.3 特徴量空間の自動生成

従来法では、特徴量空間を生成する基底画像を、認識実験の結果を考慮し人手で選択していた。そのため、別のオブジェクトを認識するために新たな特徴量空間を作成する必要があるとき、多くの実験を行い、適切な基底画像を選択する必要がある。これでは常に適切な基底画像を選択できる可能性は低く、また多くの手間を必要としてしまう。実画像などに本手法を適用する場合を考えると、自動で基底画像を選択し、適切な特徴量空間を作成するシステムが必要となってくる。

同じような陰影情報を持つ局所領域は、同じようなテキスト列を持ち、同じような特徴量を示す。つまり、特徴量空間内では、同じような陰影情報を持つ局所領域どうしは、近い位置にあり、違う陰影情報を持つ局所領域間の距離は、遠くなると考えられる。よって、特徴量空間のクラスタ解析を行い作成されたクラスタ内から代表画像を選択し、新たな基底画像とすることでより良い特徴量空間が作成されると考えられる。

最初に、基底画像の候補となる局所領域を画像から作成するが、局所領域内に多くの背景部分が含まれていれば、その局所領域は、オブジェクトの特徴を示すというよりも背景の特徴を示すことになってしまう。よって、背景部を多く含んだ局所領域は、オブジェクトの特徴をあまり持たないので基底画像には適さないと考えられる。そこで、 50×50 ピクセルの局所領域のうち500ピクセル以上が背景部ならば、基底画像の候補からはずすことにした。

次に、得られた局所領域から仮の基底画像を用いて特徴量ベクトルを求める。この特徴量ベクトルを仮の特徴量空間内に配置し、ベクトル間のユークリッド平方距離を非類似度とし、群平均法を用いてデンドログラムを作成する。次に、このデンドログラムを切断し、クラスタを作成する。次に、このクラスタから代表ベクトルを求めるのであるが、各クラスタの重心に最も近い特徴量ベクトルをそのクラスタの代表ベクトルとした。このベクトルを生成した局所領域を新たな基底画像とする。この

基底画像を用いて新たな特徴量空間を作成する。

ここで、使用する仮の基底画像は、偏った特徴を示す可能性がある。対応策として次の方式を用いる。

まず、仮の基底画像を用いてクラスタ解析を行い、新たな基底画像の選択を行う。ここで、すべての局所領域の特徴量ベクトルが生成され、空間外データが存在しない、つまり、仮の基底画像が偏っていない場合、この新たな基底画像を用いて新たな特徴量空間とする。しかし、空間外データが存在したならば、新たな基底画像にこの局所領域を加え、新たな特徴量空間を作成し、この空間内でクラスタ解析を行いもう1度基底画像を選択する。このシステムのフローチャートを図5に、概念図を図6に示す。

4 実験結果及び考察

4.1 特徴量の見極め

最初に、特徴量の見極めの為の実験を行った。基底画像、登録画像を図7,8に示す。登録画像はオブジェクトごとに1件のみとする。また、表2に認識に用いた特徴量の組み合わせを示す。これらを用いて形状を変化させた未知画像を22個(図9)認識し、その認識率を表3に示す。

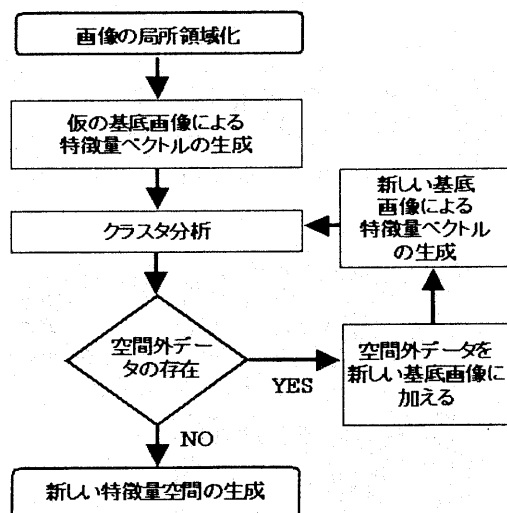


図5 特徴量空間の自動生成システム

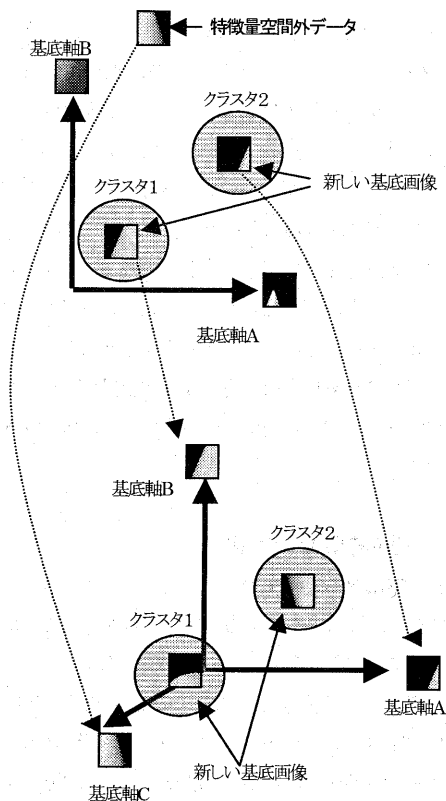


図6 特微量空間の自動生成システムの概念図

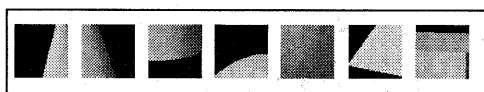


図7 基底画像

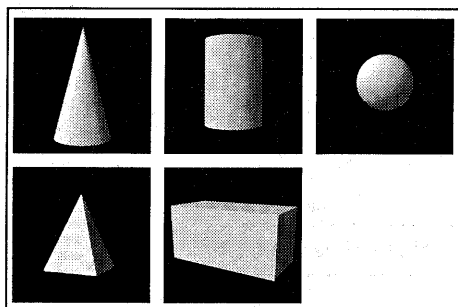


図8 登録画像(各1件のみ使用)

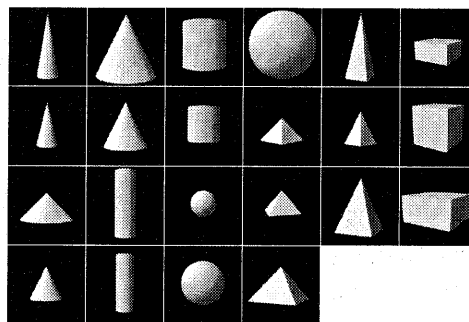


図9 未知入力画像

表2 実験に用いる特微量

特微量 実験	圧縮率	利用頻度	利用頻度の 標準偏差
実験1	○	×	×
実験2	×	○	×
実験3	○	○	×
実験4	○	○	○

表3 認識率(登録画像1件)

実験	認識率
実験1	9/22 = 40.9%
実験2	14/22 = 63.6%
実験3	11/22 = 50.0%
実験4	11/22 = 50.0%

この結果を比較すると基底生成文法の利用頻度(2.2.2 参照)が最も高くなった。文法利用頻度というのは、少数事例で多用な形状変化に対応するのに好都合と考えられる。圧縮率は、形状の変化による1次元符号の変化の影響を受け易く、認識率が低くなったと思われる。

また、文法利用頻度の標準偏差については、この値を比較してみるとどのようなオブジェクトでもほぼ一定になってしまっている。よって、FRC 値への影響が少ない。これらのことが、認識率向上に寄与しない原因となっていると考えられる。

4.2 特徴量空間の自動生成

次に、特徴量空間の自動生成の実験を行った。4.1の実験の結果から基底生成文法の利用頻度だけを特徴量として使用していく。まず、12種類の画像から140の局所領域を作成した。

次に図7で示した基底画像を仮の基底画像として用いて、140個の局所領域の特徴量ベクトルを生成し、得られた140個の特徴量ベクトルに対しクラスタ解析を行い、7個のクラスタを作成した。

この結果、図10のような7個の基底画像が選択されたが、この仮の特徴量空間に入りきらないデータが6個あった(図11)。よって、新たな基底画像にこの空間外データを加え、基底数13の特徴量空間内で140個の局所領域を再びクラスタ解析した。その結果、図12の局所領域が基底画像として選択された。この時、空間外データは存在しなかったため、これを新たな基底画像として使用することにした。

この基底画像を用いて新しい基底空間を作成し、認識実験を行った。登録画像及び未知画像は図7,8と同様なものを用いた。認識結果は表4である。やはり、利用頻度を用いた実験での認識率が高く、円錐、四角錐、立方体では認識率は100%であった。

4.3 異なる照明条件下での認識実験

ここでは、5つのオブジェクトに対して照明の向きを変化させて認識実験を行う。実験で用いた照明は、平行光線を用いる。この認識実験で用いた登録オブジェクトを図13に示す。未知入力オブジェクトとして、登録オブジェクトと同じものを使い、照明条件の影響だけを見る。登録オブジェクトの照明の向きは、ヘディング45度、ピッチを35度の向きにした。それに対して未知入力オブジェクトの照明の向きは、ヘディングを60度、30度、0度、-30度、-60度、ピッチを20度、35度、50度、65度にした。つまり、1つのオブジェクトに対して20通りの照明条件で認識を行う。円錐についての例を図14に示す。その認識結果が表5である。

その結果を見ると改善は見られたが、認識率はまだ低く、照明条件の多様性には対応できていない。そこで、事例を増やして認識実験を行ってみた。その結果が表6である。少量の事例追加によって多様な照明条件のもと

での認識率が確実に増加していくことが確認できた。

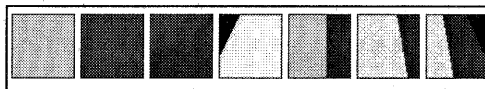


図10 択された局所領域

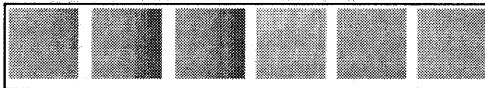


図11 空間外データ

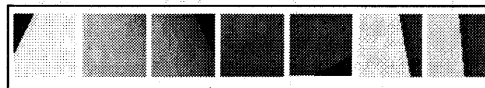


図12 新しい基底画像

表4 新しい基底画像による認識結果

実験	認識率
実験1	13/22 = 59.1%
実験2	17/22 = 77.2%
実験3	13/22 = 59.1%
実験4	14/22 = 63.6%

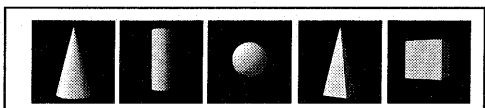


図13 登録画像

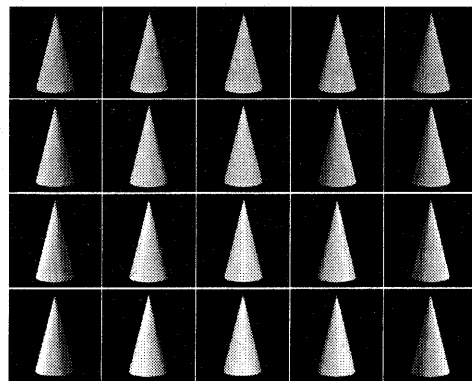


図14 照明を変化させた未知入力画像(円錐)

表5 照明条件の変化に対する認識結果
(登録事例数1件)

オブジェクト	認識率	認識率
円錐	9/20=45.0%	11/20=55.0%
円柱	1/20=5.0%	11/20=55.0%
球	17/20=85.0%	9/20=45.0%
四角錐	6/20=30.0%	9/20=45.0%
直方体	9/20=45.0%	13/20=65.0%

表6 事例数の増加による認識率の変化

教事例数	1個	3個	5個
円錐	55.0%	75.0%	90.0%
円柱	55.0%	65.0%	80.0%
球	45.0%	65.0%	80.0%
四角錐	45.0%	70.0%	75.0%
直方体	65.0%	70.0%	75.0%

5 まとめ

本研究では、オブジェクトの特徴表現手法として既存の多次元特徴量直方体の相関関係を利用したオブジェクト認識手法に以下の改良を導入し、その有効性を示した。

- ・ 認識に有効な特徴量は、基底生成文法の利用頻度である事を明らかにした
- ・ 自動的に基底画像を選択し特徴量空間を生成するシステムを完成し、システム構築コストの削減を実現した
- ・ 輝度を正規化することで照明条件の変化への対応能力を向上させた

参考文献

- [1] L.G.Brown, "A survey of Image Registration Techniques", ACM Computing Surveys, vol.24, no.4, pp.325-376, 1992.
- [2] 大田友一, 池内克史, "3次元情報の再構成とモデ

ル化", O plus E, pp.86-105, 1985.

- [3] K.Ikeuchi, B.K.P.Horn, "Numerical Shape from Shading and Occluding Boundaries", Artificial Intelligence, vol.17, pp.141-181, 1981.
- [4] B.K.P.Horn, "Hill-Shading and the Reflectance Map", in Proc. of DARPA Image Understanding Workshop, pp.79-120, 1979.
- [5] J.Weng, N.Ahuja, T.S.Huang, "Matching Two Perspective Views", IEEE TRANS.on PAMI, vol.14, no.8, pp.806-825, 1992.
- [6] 品田岳史, 安藤祐樹, 渡辺俊典, 菅原研, "多次元特徴量直方体の相関関係を用いたオブジェクト認識手法 FRCM", 情報処理学会研究報告, 2001-CVIM-127, 25-32, 2001.
- [7] Philip L.Worthington, Edwin R.Hancock, "Object Recognition Using Shape-from-Shading", IEEE TRANS.on PAMI, vol.23, no.5, pp.535-542, 2001.
- [8] J.R.Leasch, A.E.Iverson, B.N.Webb, K.F. West, "Segmentation of Multiband, Imagery using Minimum Spanning Trees", Igorithm for Multi-spectral and Hyperspectral Imagery II, Proc. SPIE 2758, pp.10-18, 1996
- [9] 渡辺俊典, "データ圧縮を用いたパターンの万能特徴表現方式 UPRDC", 日本写真測量学会 平成12年度年次学術講演会発表論文集, pp.253-258, 2000.
- [10] T.Watanabe, K.Sugawara, and H.Sugihara, "A new pattern representation scheme using data compression", IEEE Trans. PAMI, Vol. 24, No. 5, 2002 (to appear).
- [11] C.G.Nevill-Manning, I.H.Witten, "Compression and Explanation using Hierarchical Grammars",