周辺固有ベクトルを利用した画像の直交展開と物体認識への応用

北海道大学院工学研究科　○下野浩，近藤司，金子俊一，五十嵐悟

要旨

2次元画像で3次元物体を認識するには物体の種類、姿勢、照明条件などによって大量の画像を学習しなければならず効率よく情報を圧縮して認識を行う必要がある．画像集合の情報を圧縮する手法の一つに，周辺固有ベクトル[1]を用いて直交展開するものがある．本研究では2次元画像を用いた3次元物体の認識にこれを利用し，学習・認識時間について検討する．

1. 始めに

画像の周辺固有ベクトルを利用した画像の圧縮には，列—周辺固有ベクトルを用いて展開，行—周辺固有ベクトルを用いて展開，行，列両方の周辺固有ベクトルを同時に用いて展開（2次元直交展開[2]）の3種類がある．2次元直交展開を用いる手法は，行あるいは列方向のみについて展開するものよりも圧縮率は一般に高い．しかし，行あるいは列方向のみについて展開するときの固有値計算は1回でよいが，2次元直交点展開するときの固有値計算は2回しなければならず，展開するのに時間がかかる．そこで本研究では列方向のみについて展開することによる学習，認識について検討を試みる．以下では，行—周辺固有ベクトルを用いた展開について述べる．

2. 周辺固有ベクトル

画像集合 P を s 個の正規直交ベクトル $\{\tilde{\boldsymbol{w}}^T{}_i\}_{i=1}^s$ $\left(\text{ただし}\tilde{\boldsymbol{w}}_i^T\tilde{\boldsymbol{w}}_j=\delta_{ij}\right)$ で

$$\boldsymbol{P}\cong\hat{\boldsymbol{P}}_s=\sum_{i=1}^{s}\hat{\boldsymbol{v}}_i\tilde{\boldsymbol{w}}_i' \tag{1}$$

のように近似する．このとき $\hat{\boldsymbol{v}}_i$ は展開係数ベクトルであり，これを行—周辺固有ベクトルと呼ぶ．

3. 本手法の概要

物体認識のアルゴリズムは大きく分けて，学習段階，認識段階の2つのプロセスに分けることが出来る．

3. 1 学習処理

複数の既知対象物の複数の見え方に対応する学習画像群を扱う．それぞれの学習画像は物体の種類，姿勢，照明条件を表す2つのパラメーターb，θ，ϕを持つものとする．

3.1.1 学習画像の前処理

学習画像を正規化し，平均画像を差し引く．まず入力された学習画像 P_Oから物体部分を切り出し，に大きさの正規化を行う．次にカメラによる輝度のばらつきをなくすために，輝度の正規化を行う．

$$\boldsymbol{P}_N^{(b)}(\theta,\phi)=\frac{\boldsymbol{P}_O^{(b)}(\theta,\phi)}{\left\|\boldsymbol{P}_O^{(b)}(\theta,\phi)\right\|} \tag{2}$$

ここで，パラメータはそれぞれ物体 b，姿勢 θ，照明条件 ϕ を表すものとする．次に，各学習画像から次式で計算される全学習画像の平均画像を差し引く．

$$\begin{aligned}\boldsymbol{C}&=\frac{1}{B\Theta\Phi}\sum_{b}^{B}\sum_{\theta}^{\Theta}\sum_{\phi}^{\Phi}\{\boldsymbol{P}_N^{(b)}(\theta,\phi)\}\\&=\underset{P}{E}\{\boldsymbol{P}_N^{(b)}(\theta,\phi)\}\end{aligned} \tag{3}$$

3.1.2 平均列—相関行列

各学習画像の列方向の相関を表す列—相関行列の標本平均をとる．

平均列—相関行列　$$\boldsymbol{R}_c=\underset{P}{E}(\boldsymbol{P}^T\boldsymbol{P}) \tag{4}$$

3.1.3 固有ベクトルの計算

次の固有値問題を解く．

$$\boldsymbol{R}_c\tilde{\boldsymbol{w}}_i=\lambda_i\tilde{\boldsymbol{w}}_i \tag{5}$$

固有値 $\lambda_1,\lambda_2,\cdots,\lambda_s$ $(\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_s,1\leq s\leq m)$ に対応する固有ベクトル $\tilde{\boldsymbol{w}}_1,\tilde{\boldsymbol{w}}_2,\cdots,\tilde{\boldsymbol{w}}_s$ を算出する．

3.1.4 展開係数ベクトル算出

それぞれの学習画像に対して，平均 2 乗誤差

$$\varepsilon_s^2=\underset{p}{E}\left\|\hat{\boldsymbol{P}}_s-\boldsymbol{P}\right\|^2 \tag{6}$$

$$=\sum_{i=1}^{s}\underset{p}{E}\hat{\boldsymbol{v}}_i^T\hat{\boldsymbol{v}}_i+tr\boldsymbol{R}_c-2\sum_{i=1}^{s}\underset{p}{E}\hat{\boldsymbol{v}}_i^T\boldsymbol{P}\tilde{\boldsymbol{v}}_i \tag{7}$$

を最小にするという意味で最適な展開係数ベクトルは

$\frac{\partial\varepsilon_s^2}{\partial\hat{\boldsymbol{v}}_i}=0$から

$$\hat{\boldsymbol{v}}_j=\boldsymbol{P}\tilde{\boldsymbol{w}}_j \tag{8}$$

により求まる．

3.1.5 展開係数ベクトル間の補間

同じ物体どうしで，姿勢あるいは照明条件の変化が少ない画像間では，画像どうしの相関が高いため，展開係数同

士の相関も高いものになる．そこで姿勢に関して離散的な学習画像間を何らかの手段で補間する事で，学習画像に含まれない姿勢の認識が可能となる．[3]そこで $\hat{\nu}_i$ を，それぞれ3次スプライン曲線により補間し，それぞれの学習画像に対応する点の間を1/10刻みの点で代表させることで，離散的ではあるがより細かく姿勢を表現した．

3．2 認識処理

3.2.1 入力画像の前処理・展開係数ベクトル算出

入力画像に関しても，3.1.1 と同様の前処理を施したうえで，3.1.4 の式(8)と同様の方法で展開係数ベクトルを求める．

3.2.2 最小距離識別による認識

入力画像と全ての学習画像との距離を次式により測る

$$D^{(b)}(\theta,\phi)=\left\|\hat{\boldsymbol{P}}_s-\hat{\boldsymbol{P}}_s^{(b)}(\theta,\phi)\right\| =\sqrt{\sum_i^s\{\nu_i-\nu_i^{(b)}(\theta,\phi)\}^2} \qquad (9)$$

ここで $\hat{P}_s,\hat{P}_s^{(b)}(\theta,\phi)$ は入力画像および学習画像の近似表現である．次に，$\min\{D^{(b)}(\theta,\phi)\}$ を満たすパラメーター b^*，θ^*，ϕ^*，を求めることによって入力画像に対応する物体の識別し，その姿勢や照明方向を推定することが出来る．

4．実験

2次元直交展開を用いる手法と比較するために，数種類の物体の姿勢を変化させた画像を用いて，物体の種類を識別し，その姿勢を認識する実験を行った．なお，光源は1方向からの照明に固定した．実験に用いた画像は物体部分を 100×100 に切り出した 256 階調の濃淡画像で，学習用として 3 物体×36 姿勢(10(deg)おきにターンテーブルを回転させて得られたもの)を 108 枚，認識性能評価用としては学習したものと 5(deg)ずつ姿勢の異なった同数の画像を用意した．

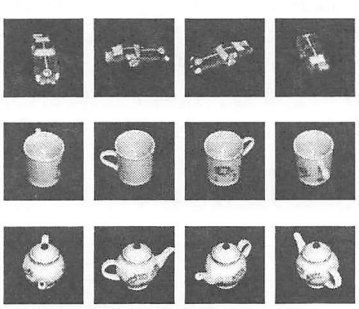

図1 実験に使用した物体 (それぞれ4姿勢分)

学習，認識にかかった時間を表1に示す．

表1 車の場合の処理時間(単位:秒 Athlon700MHz)

展開ベクトル係数の次数		s=1	s=3	s=5	s=t=5
学習処理	前処理	30.204	30.233	30.384	46.537
	固有ベクトル算出	2.563	2.593	2.574	4.266
	展開係数ベクトル算出・補完	0.181	0.281	0.370	0.380
	計	32.948	33.108	33.328	51.183
認識処理		0.041	0.124	0.258	0.019

物体の種類の認識においては，誤認識を起こした例はなかった．

姿勢認識誤差の平均と展開係数ベクトル次数 s との関係は図2のようになった．次数 s が 2 以上ならば誤差平均は1度以内に収まっている．

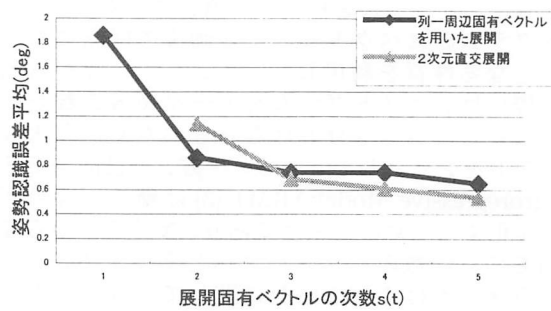

図2 s と姿勢認識誤差の関係

5．まとめ

今回の実験で行一周辺固有ベクトルのみを用いた画像の展開しても物体の認識に使えることが確認できた．2次元直交展開に比べ学習時間を短縮することが出来た．しかし，認識時間は展開係数ベクトルが大きいため逆に時間がかかった．また，2次元直交展開と比べ姿勢認識誤差が少し大きくなっている．これは展開係数ベクトルを補間するときに誤差が生じるものと思われる．

参考文献

[1] 大津展之，栗田多喜夫，関田巌:"パターン認識"，朝倉書店，1996.

[2] 小林武史，金子俊一，五十嵐悟:"画像の 2 次元直交展開に基づく物体認識"，電子情報通信学会技術報告，PRMU99-215，2000

[3] 村瀬洋，シュリー ナイヤー:"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，D-Ⅱ，Vol.J77-D-Ⅱ，No.11，pp.2179-2187，1994.