ウェーブレット変換と固有空間法を用いた物体認識に関する研究

北海道大学　○小林武史, 金子俊一, 五十嵐悟

要旨

2 次元画像を用いて 3 次元物体認識及び物体のポーズ認識を行うパラメトリック固有空間法[1]は, 画像の部分的な特徴を抽出する処理が不要であるという特徴があるが, 前処理として多方向の視点から見た画像を学習し, 画像集合から生成した多次元の共分散行列の固有値・固有ベクトルを求める事が必要で, この処理には計算時間及び使用メモリの両面から見て高い計算コストを要する. そこで本研究では, ウェーブレット変換を用いて画像データ中の認識に大きな影響を及ぼさない周波数成分を取り除き, 行列の次元を下げ, 学習にかかる処理時間を軽減することを目的としている.

1. パラメトリック固有空間法

この方法は学習段階と認識段階の 2 段階に分けられる.

学習段階では学習サンプル画像の集合から固有値・固有ベクトルを計算し, それぞれの学習サンプル画像を固有空間上の点に投影する.

認識段階では入力画像をこの固有空間に投影し, その点に最も近い学習サンプル画像の投影された点を探すことにより物体の認識を行う.

正方形に切り出された画像をラスタースキャンしベクトル

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_N]^T \quad (N\text{は画素数})$$

として表現し, これを正規化したベクトルを $\mathbf{x}$ と表す.
ここでポーズを変化させた 1 種類の物体を学習させる場合を考える. すべての画像集合の平均

$$\mathbf{c} = \frac{1}{R}\sum_{r=1}^{R}\mathbf{x}_r \quad (R\text{はポーズ数})$$

を各学習サンプルから差し引いた行列

$$\mathbf{X} \equiv [\mathbf{x}_1 - \mathbf{c}, \mathbf{x}_2 - \mathbf{c}, \cdots, \mathbf{x}_R - \mathbf{c}]$$

によりその物体を表現する.

次に K-L 展開により画像集合を圧縮する. これは主要な固有値に対応する固有ベクトル(主成分)が構成する固有空間により元の画像を表現しようとする手法である. 共分散行列は $\mathbf{Q}=\mathbf{X}\mathbf{X}^T$ により計算される. $\mathbf{Q}$ の固有方程式 $\lambda_i\mathbf{e}_i=\mathbf{Q}\mathbf{e}_i$ を解き, 最大の固有値から k 個の固有値 $\lambda_1 \geq \cdots \geq \lambda_k (\geq \cdots \geq \lambda_N)$ に対応する固有ベクトル $\mathbf{e}_1 \cdots \mathbf{e}_k$ を基底ベクトルとすることにより k 次元の固有空間が得られる.
次に学習サンプルをこの固有空間上の点 $\mathbf{g}$ に次式により投影する.

$$\mathbf{g}_r = [\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_k]^T(\mathbf{x}_r - \mathbf{c}) \qquad (r=1,2,\cdots,R)$$

これにより学習サンプルは k 次元空間上の点に投影される.
近い姿勢の画像同士は固有空間上の近い点に投影されるため, 1 つの物体を系列として扱う事ができ, 投影点の集合を曲線で補間する事により学習集合にない姿勢の推定が可能である. 本研究では, 3 次スプライン曲線を用いている. (図 1)

同様に入力画像 $\mathbf{y}$ を

$$\mathbf{z} = [\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_k]^T(\mathbf{y} - \mathbf{c})$$

により同じ固有空間上の点 $\mathbf{z}$ に投影する. 同じ物体で姿勢の近い画像同士は相関が強いため, 固有空間上の近い点に投影される. したがって学習集合から生成した曲線と $\mathbf{z}$ の距離が最小となる場所が認識結果である.

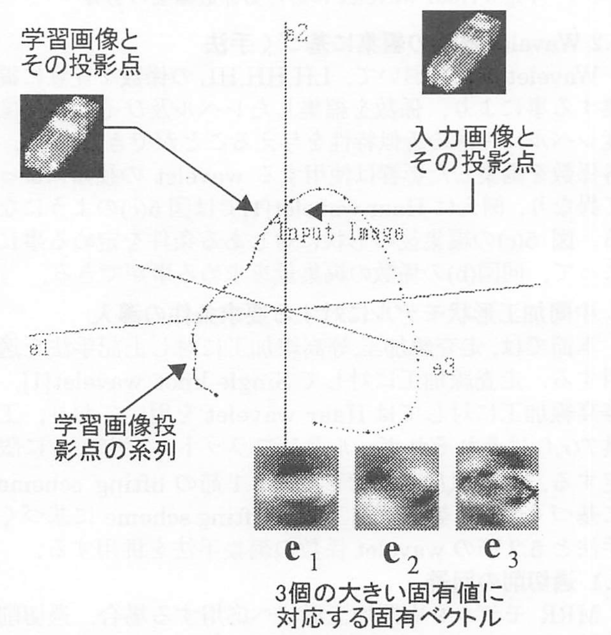

図 1. 主成分となる固有ベクトルとその固有空間

2. ウェーブレット変換による画像圧縮

固有空間法における画像の共分散行列 $\mathbf{Q}$ は, 画像の画素数を N とした場合 $N \times N$ の行列となり, 固有値を算出する際に大きなコストを要する.

また, 大きな固有値 $\lambda_1 \cdots \lambda_k$ に対応する固有ベクトル $\mathbf{e}_1$

…e_k は, 物体の大まかな形状を表すであろうと考えられ, 固有値・固有ベクトルを計算する前に高周波成分を除いておくことは認識にさほど大きな影響を与えないであろうと思われる. このような理由から, 本研究では学習の前に学習サンプルに対してウェーブレット変換[2](以下 WT)により周波数分解を行い, 低周波成分のみを取り出す事でデータ圧縮を行う.

2 次元での分解アルゴリズムにおける係数列{ $c_{LL}^{(n)}$ }の低周波成分を{$c_{LL}^{(n-1)}$}とすれば

$$c_{LL,k_1,k_2}^{(n-1)} = \sum_{l_1,l_2} a_{l_1-2k_1} a_{l_2-2k_2} c_{LL,l_1,l_2}^{(n)}$$

により{$c_{LL}^{(n)}$}の低周波成分のみを分離することができる. ここで{a_k}はウェーブレット分解列であり, 今回は 2 階の B スプラインウェーブレット分解列を用いた. 上式を画像に適用すれば画像の低周波成分, すなわち大まかな変化を表す係数列が得られる. {$c_{LL}^{(n-1)}$}は{$c_{LL}^{(n)}$}の 1/4 のデータ量となる.

これをそれぞれの学習サンプルに対して行うことにより学習段階での処理の軽減を図る.

3. 実験

図 2 のような 3 つの物体を 10 度おきに 360 度回転させて撮影した画像を学習し, 物体認識実験を行った. WT 圧縮を用いた場合と用いない場合を比較した. なお, 評価用の画像は学習画像と 5 度ずつ姿勢を変化させた同枚数の画像を用いた.

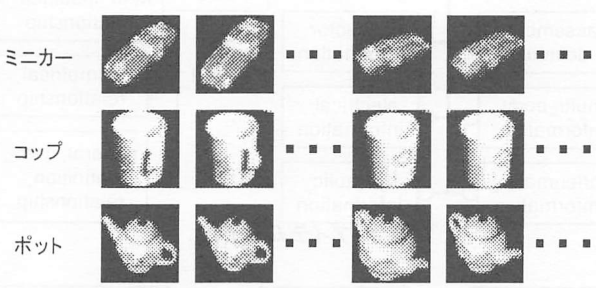

図 2.実験に使用した物体

今回の実験では光源は 1 点に固定した. 画像のサイズは 32×32 である. 固有空間の次元数は 8 次元とした. この際の結果は図 4, 図 5 の通りである.

WT による圧縮を取り入れると, 学習にかかる処理コストはメモリの使用量において約 1/4 程度, 計算時間(そのほとんどは固有値計算に費やされる)にして 1/16 程度(図 4)まで節減された. また, この例においては, WT 圧縮を使用すると, 物体の種類を認識する際には差がないが(どちらも全て正認識), 約 1 割程度姿勢の認識精度が低下する(図 5). この原因としては, 主成分となる固有ベクトル(図 3)に本来含まれるべき成分も学習段階において切り捨ててしまっていることが考えられる. 今後はこれを解決することが課題である.

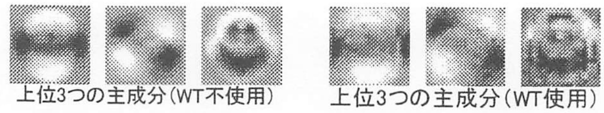

図 3. 主成分となる固有ベクトル

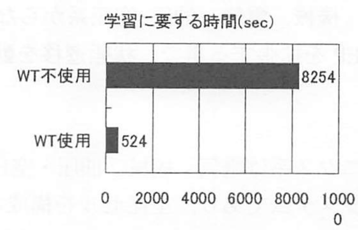

図 4. 学習処理時間

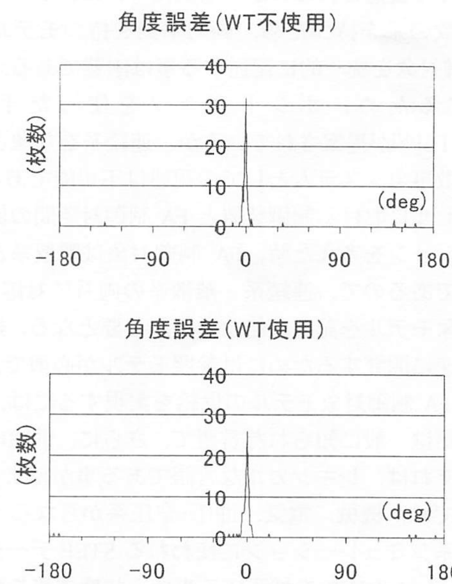

図 5. 角度誤差

4. まとめ

今回の実験で学習における処理コストを大幅に軽減することができた. 今後は認識性能を上げることが課題である.

参考文献

[1] 村瀬洋, パラメトリック固有空間法による 3 次元物体の認識とスポッティング, MIRU"94, pp. II-49-56, 1994.

[2] チャールズ K. チュウイ, ウェーブレット応用, 東京電気大出版局, 1997.

[3] 小林武史, 金子俊一, 五十嵐悟, ウェーブレット変換による多重解像度解析と固有空間法を用いた物体認識に関する研究, 精密工学会春季大会論文集, pp. , 1999.