

# パーティクルフィルタによる適応的重み推定と Adaboost による判定基準マイニングを用いたシーン識別

堀田一弘†

† 名城大学理工学部電気電子工学科

〒 468-8502 名古屋市天白区塩釜口 1-501

E-mail: †kazuhotta@meijo-u.ac.jp

あらまし 本論文ではパーティクルフィルタによる適応的重み推定と Adaboost による判定基準マイニングを用いたシーン識別法を提案する．シーン識別には構図のような位置に依存した情報が有効である．しかしその一方で、画像内の対象の位置は一定でないので位置に依存しない類似度も必要である．そこで、本論文ではこれら 2 つを局所カーネル主成分分析 (KPCA) と大局 KPCA により表現し、それらをテストサンプルに応じて適応的に重みづけ統合する．パーティクルフィルタを用いて各クラス毎に重みを選択した場合、正解のクラスが 1 位になる重みを持つ割合がかなり高い．しかし、複数のクラスが 1 位になる重みを持つ場合の明確な判定基準がないため、候補の絞りこみが高い精度でできるにも関わらず、最終的に正しく識別できないことがある．そこで、適応的な重み推定により得られる情報を基に Adaboost を用いて最終的な判定基準のマイニングを行った．適応的な重み推定と判定基準のマイニングを行うことにより、局所または大局 KPCA だけを用いた場合と比べて識別率が 14% 以上向上した．13 シーンのデータベースを用いた実験では、従来法と比較して最も高い識別精度である 91.06% が得られた．

キーワード 適応的重み推定, 判定基準のマイニング, シーン識別, 局所 KPCA, 大局 KPCA

## 1. はじめに

もしシステムが入力画像のシーンを認識できれば、その画像中に含まれる可能性が高い対象を推測できるかもしれない．画像全体の情報だけから特定対象の位置や大きさを推定する研究もある [1], [2]．しかし、それらの方法は画像全体の情報を用いているだけであり、シーンカテゴリの情報を用いていない．シーンカテゴリとそこに含まれる対象には強い相関関係があるので、シーンカテゴリの情報は単なる画像全体の情報よりも良い文脈情報になり得る．従って、シーンカテゴリの情報は対象検出の高速化や対象認識の高精度化に役立つだけでなく、画像理解の第一段階としても重要である．

そうした背景からシーンカテゴリの識別法は盛んに研究されている [3] ~ [11]．最も有名な方法は局所領域と bag-of-visual words を用いたものである [3], [4]．これらの論文ではグリッドサンプリングを用いて局所特徴を抽出することにより精度が向上することを報告している．また、Lazebnik らも積極的に位置情報を用いることにより精度が向上することを示している [5]．これはシーン識別には構図のような位置に依存した情報が有効であることを示している．しかし、画像内での対象の位置は変化するので、シーン内に含まれる対象の位置に依存しない類似度も重要である．つまり、シーン識別では位置に依存する類似性と位置に依存しない類似性という相反する 2 つの類似性が重要となる．

位置に依存した大まかな構図情報を表現するために再

構成が可能である大局 KPCA を用い、位置に依存しない局所的な情報の表現として局所 KPCA を用いた．そして、相反する 2 つの要求に答えるために各々の再構成誤差を統合することを考えた [11]．しかし、各テストサンプル毎に統合の際の最適な重みは異なるので、状況に応じて適応的に重みを決める必要がある．そこで、本論文ではパーティクルフィルタを用いて統合の際の重みを適応的に決める [12]．各パーティクルを統合の際の重みベクトルに割り当て、再構成誤差を事後確率に関連づければ、そのサンプルを識別するのに有効な重みベクトルを得ることができる．シーン識別は多クラスの識別問題であるので、パーティクルフィルタによる重み推定は各クラス毎に独立に行い、そのクラスが 1 位になるような重みを選択する (もちろん、そのような重みが存在しない場合も多い)．13 シーンのデータベース [13] を用いた実験により、適応的な重み推定により正しいクラスが 1 位になる重みを持つ割合が 97% と高いことが分かった．

しかし、クラス毎に重み推定を行うので、サンプルによっては複数のクラスが 1 位になる重みを持つことがある．その場合には再構成誤差を用いて判定することが 1 つの手であるが、重みがクラス毎に異なるので、この方法により正しいクラスに識別される保証はない．そこで、適応的重み推定を候補をしばり込む方法と位置づけ、適応的な重み推定の際に得られる重みベクトルの和や再構成誤差等の情報を基に、複数クラスが 1 位になる重みを持った場合に有効な判定基準をマイニングした．これには特徴選択 [14] に有効である Adaboost [15] を用いた．

実験では従来法 [3] ~ [5], [7] ~ [11] で使われている 13 シーンのデータベース [13] を用いて評価した。まず、識別基準のマイニング結果であるが、複数のクラスが 1 位になる重みを持つような場合には、構図のような大局情報はあまり役に立たず、局所 KPCA の再構成誤差と識別に関連する情報が重要であることが分かった。次に、識別精度であるが、提案手法は局所もしくは大局 KPCA を単独で用いた場合よりも精度が 14% 以上も向上した。また、同じデータベースを用いた従来法と比べて最も高い精度が得られた。これにより、テストサンプルに応じた適応的重み推定法及び判定が難しいサンプルに対する識別基準のマイニングが有効であることが実証できた。

本論文の構成は以下の通りである。まず、2. 節では提案手法の概略を説明する。識別の基になる特徴量に関しては 3. 節、局所 KPCA と大局 KPCA については 4. 節で説明する。5. 節ではパーティクルフィルタを用いた適応的重み推定について述べる。判定が難しいサンプルに対する判定基準のマイニング法を 6. 節で説明する。実験結果を 7. 節に示し、最後に 8. 節で結論と今後の課題を述べる。

## 2. 提案手法の概略

提案手法は以下の 4 つのステップから構成される。第一ステップはガボール特徴に基づく特徴量抽出 [8] である。本論文では従来法 [3], [8] により有効性が報告されているグリッドサンプリングと方向ヒストグラムを用いる。画像から 3 段階の大きさで 8 方向のガボール特徴を抽出した後、特徴量を  $16 \times 16$  のグリッドに分割し、各グリッドで方向ヒストグラムを作成する [9]。

第二ステップは局所 KPCA と大局 KPCA である [11]。シーンの大まかな構図情報をモデル化するために再構成可能である大局 KPCA を利用し、シーン中の細かい対象を表現するために局所 KPCA を用いた。各 KPCA により得られた再構成誤差を重みづけ統合することにより識別を行うが、テストサンプル毎に有効な重みは異なる。そこで、本論文ではパーティクルフィルタを用いて適応的に重みを選択する。この適応的重み推定が第三のステップである。

適応的な重み推定により正しいクラスが 1 位になる重みを持つ割合はかなり高い。1 つのクラスだけが 1 位になる重みを持つ場合にはそのクラスに識別すれば良いが、複数のクラスが 1 位になる重みを持つ場合には最終的な判定基準がない。この際に再構成誤差を用いるのは 1 つの手であるが、クラス毎に異なる重みベクトルを用いるのであまりうまく働かないことが分かった。我々人間にはどの基準が有効であるかは分からないので、適応的な重み推定の際に得られる 10 個の値の全クラス内での順位を基に Adaboost により最終的な判定基準のマイニングを行う。これが第四のステップである。これらの 4 つのステップを順を追って説明していく。

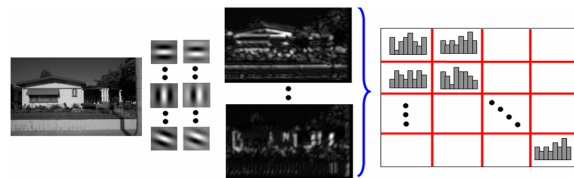


図 1 ガボール特徴から方向ヒストグラムを作成する方法

## 3. ガボール特徴の方向ヒストグラム

近年、方向ヒストグラムの有効性が報告されている [5], [16]。従来手法では単純な方向微分のヒストグラムが用いられることが多かったが、ここでは局所的な見えの情報を抽出できるガボール特徴を基に方向ヒストグラムを作成する [8]。

本論文では 3 つの大きさパラメタと 8 つの方向パラメタのガボールフィルタを用いる。ガボール特徴を抽出した後、 $M \times M$  のグリッドに分割し、各グリッドで方向ヒストグラムを作成する。様々なパラメタのガボールフィルタの出力値から方向ヒストグラムを作成する方法について図 1 を用いて説明する。図の例では分割数がいまいであるが、実験では  $16 \times 16$  の領域に分割する。また、図中では 1 つの大きさパラメタのガボール特徴から方向ヒストグラムを作成しているが、実際には 3 つの大きさパラメタのガボールフィルタを用いる。まず、8 方向の実数部と虚数部のガボールフィルタを画像に適用し、各点毎に実数部と虚数部のノルムを計算する。次に、グリッド内の各画素で抽出した 8 方向のガボール特徴の中で最大出力方向に投票していくことにより、8 次元の方向ヒストグラムを作成する。この処理を大きさパラメタ毎に行う。つまり、 $16 \times 16$  に分割した各領域が 24 次元 (= 3 (大きさパラメタ)  $\times$  8 (方向パラメタ)) の方向ヒストグラムを持つ。

しかし、各グリッドで得られた 24 次元の方向ヒストグラムよりも  $2 \times 2$  のような複数グリッド上の方向ヒストグラムを用いた方が高い精度が得られることが報告されている [10]。そこで、本論文では  $2 \times 2$  のグリッド上の 4 つの方向ヒストグラムを 1 つの局所特徴として用いる。つまり、各局所特徴の次元数は  $96 (= 24 \times 2 \times 2)$  となる。重なりがないように分割すれば、1 枚の画像から  $64 (= 8 \times 8)$  の局所特徴が得られる。

局所 KPCA では位置情報を捨てた 96 次元の局所特徴が使われ、大局 KPCA では位置情報を持った全ての局所情報が利用される。

## 4. 局所 KPCA と大局 KPCA

本節では局所 KPCA と大局 KPCA を説明する<sup>(注1)</sup>。局所 KPCA は画像の細部を表現し、大局 KPCA は画

(注1): 誌面の都合により KPCA の説明を省略するので、文献 [17] を参照して下さい。

像の大まかな情報を表現することになる [11] . また , 各 KPCA の再構成誤差の計算法についても本節で説明する . 局所 KPCA の再構成誤差は位置不変な類似度となり , 大局 KPCA の再構成誤差はそのシーンの構図情報に関する類似度となる .

#### 4.1 局所 KPCA

位置情報を捨てた局所特徴の集合は様々なパターンを含むので , その分布は非線形となる . 従って , それを表現するためには KPCA の非線形性が有効である [18] . 本論文では , 位置情報を捨てた  $2 \times 2$  のグリッド上にある 4 つの方向ヒストグラムの集合に KPCA を適用したものを「局所 KPCA」と呼び , クラス毎に 1 つの空間を作成する . 特徴のノルムの正規化は精度を向上させることが報告されている [19] ので , ここでは各グリッド上の 24 次元の各方向ヒストグラムのノルムが 1 になるように正規化した .

局所 KPCA の概略を図 2 に示す . まず , 局所特徴  $x$  を非線形写像を用いて  $\phi(x)$  に射影する . この空間で線形主成分分析を行うことにより , 固有ベクトル行列  $V$  を得ることができる . これを用いて KPCA 空間内での特徴量  $y_i = V^T \phi(x)$  を得ることができ , 再構成した特徴量は  $VV^T \phi(x)$  となる . つまり , 局所 KPCA における  $\phi(x)$  の再構成誤差は

$$\begin{aligned} \|\phi(x) - VV^T \phi(x)\|^2 &= \phi(x)^T \phi(x) - \phi(x)VV^T \phi(x) \\ &= K(x, x) - \|V^T \phi(x)\|^2, \quad (1) \end{aligned}$$

のように計算できる . ここで ,  $K(x, x)$  はカーネル関数であり ,  $v^T \phi(x) = \sum_i \alpha_i K(x_i, x)$  ( $\alpha$  は KPCA により得られた固有ベクトル) である . 本論文では正規化多項式カーネル [20] を用いた . このカーネル関数の定義は

$$K(x, y) = \frac{(1 + x^T y)^d}{\sqrt{(1 + x^T x)^d (1 + y^T y)^d}}, \quad (2)$$

である . このカーネルは RBF カーネルと比べてパラメタ依存性が低い上 , 精度は RBF カーネルと同等であることが知られている [20] . 実験では経験的に  $d = 5$  を用いた . また , 上の式から分かるように  $K(x, x) = 1$  である . 従って ,  $i$  番目の局所特徴  $x_i$  の局所 KPCA による再構成誤差  $\epsilon_{li}$  は  $K(x_i, x_i) - \|V^T \phi(x_i)\|^2 = 1 - \|V^T \phi(x_i)\|^2$  となり , 誤差の値は 0 ~ 1 の値を取る .

局所 KPCA だけを用いた識別では , 各クラスの KPCA を用いて画像内の全ての局所特徴の再構成誤差の和  $\sum_i^N \epsilon_{li}$  を計算し , 誤差の和が最小となるクラスに識別すれば良い .

#### 4.2 局所カーネルの和統合を用いた大局 KPCA

1 節で述べたように , シーン識別には構図のような大局的な類似度と位置に依存しない局所的な類似度の相反する 2 つの類似度が重要である . 画像によっては大局的

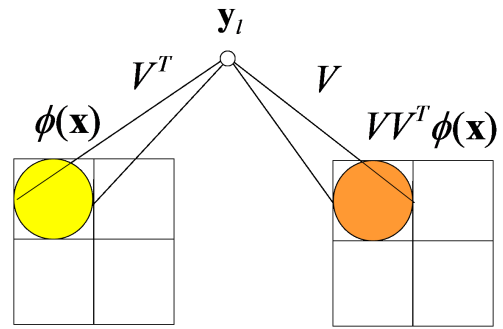


図 2 局所 KPCA

な観点から類似度を測った方が識別しやすいものもあれば , 構図情報があまり役に立たない場合もあるだろう . つまり , どちらを利用すべきかは事前には分からず , 大局的な構図情報のような観点および局所的な観点からの同じ特徴量の誤差を計算し , 適応的にどちらの類似度を優先すべきかを定めるべきである . そのためには同じ特徴量の誤差を 2 つの異なる観点から測りたいが , もし特別な工夫がない通常の KPCA を大局特徴の集合に適用した場合 , 画像全体の再構成誤差  $\epsilon_g$  しか得ることができない . つまり , 各局所特徴 ( $2 \times 2$  グリッド上の 4 つの方向ヒストグラム) に対する大局的な観点からの再構成誤差  $\epsilon_{gi}$  を得ることはできない . ここでは , 局所特徴量の再構成誤差を大局的な観点から得るために , 局所カーネルの和統合を利用する [11], [18] .

局所カーネルの和統合は以下のように定義される .

$$\begin{aligned} K_{sum}(x, y) &= \sum_i^N \phi(x_i)^T \phi(y_i) = \sum_i^N K(x_i, y_i) \\ &= \phi_g(x)^T \phi_g(y). \quad (3) \end{aligned}$$

ここで ,  $\phi_g(x) = (\phi(x_1)^T, \dots, \phi(x_N)^T)^T$  ,  $x = (x_1^T, \dots, x_N^T)^T$  である . つまり , 各局所特徴  $x_i$  を非線形写像により  $\phi(x_i)$  に射影し , それらを結合させることにより大局特徴  $\phi_g(x)$  を作る . そして , 大局特徴  $\phi_g(x)$  の集合に線形主成分分析を適用すれば良い . この様子を図 3 に示す . 固有ベクトル  $W$  を用いて KPCA 空間での新しい特徴量  $y_g$  を得ることができ , 再構成した特徴量を  $WW^T \phi_g(x)$  のように計算できる .

もし局所特徴に適用するカーネル関数を 2 次の正規化多項式カーネルとすれば , 非線形写像により射影した高次元空間での特徴量  $\phi(x)$  を明示的に記述することができる . 特に , ここでは局所カーネルの和統合を用いているので , KPCA により得られる固有ベクトルの次元もそれほど高くならずに明示的に書くことができる [18] . 従って , 大局 KPCA による再構成誤差は ,  $\phi_g(x)$  と  $\widehat{\phi_g(x)} = WW^T \phi_g(x)$  の差を計算すれば良いので簡単である . 再構成誤差  $\|\phi_g(x) - \widehat{\phi_g(x)}\|^2$  は全ての局所特徴における再構成誤差の和  $\sum_i^N \|\phi(x_i) - \widehat{\phi(x_i)}\|^2$  となる .  $i$  番目の局所特徴に対する再構成誤差を  $\epsilon_{gi}$  と記述する .



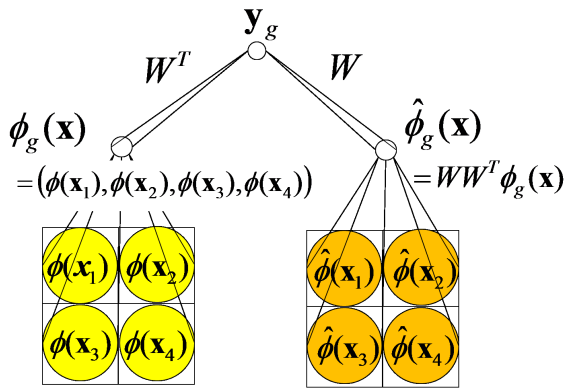


図3 大局 KPCA

局所 KPCA では位置情報を持たない局所特徴だけから固有ベクトルを決定するのに対し，大局 KPCA では位置情報を持った画像全体から固有ベクトルが決まる．上述のように大局 KPCA は  $\phi_g(x)$  の線形主成分分析により得られるので，周辺領域との相対的な情報も入っている．これらが 2 つの KPCA の違いであり，同じ局所特徴の再構成誤差を局所的な観点と大局的な観点から測り，それらを適応的に統合する．

大局 KPCA だけを用いて識別するには，クラス毎に各局所特徴の再構成誤差を計算し，それらを  $\sum_i^N \epsilon_{gi}$  のように和統合する．そして，最も小さい再構成誤差を持つクラスに識別すれば良い．

## 5. パーティクルフィルタを用いた適応的重み推定

上述のように，同じ局所特徴の再構成誤差を局所的な観点および大局的な観点から計算し，それらを適応的に統合したい．本論文では，パーティクルフィルタを用いて適応的に重みを推定する方法を提案する．まず，パーティクルフィルタについて 5.1 節で簡単に説明し，5.2 節で適応的な重み推定について説明する．

### 5.1 パーティクルフィルタ

時刻  $t$  での状態と観測をそれぞれ  $X_t, Z_t$  とする．この時，事後確率  $p(X_t|Z_t)$  はベイズの定理を用いて

$$p(X_t|Z_t) \propto p(Z_t|X_t) \int p(X_t|X_{t-1}) p(X_{t-1}|Z_{t-1}) dX_{t-1},$$

のように計算できる．

パーティクルフィルタでは， $p(X_t|Z_t)$  を  $N$  個のパーティクル  $\{(s_t^k, \pi_t^k) | k = 1, \dots, N\}$  を用いて推定する．ここで， $s_t^k, \pi_t^k$  は  $k$  番目のパーティクルの状態とその重要度である．確率分布は離散パーティクルにより表現されるので，任意の非線形分布を表現することができる．

事後確率  $p(X_t|Z_t)$  は以下のように再帰的に求めることができる．

- (1) パーティクルの状態  $s_{t-1}^k$  を  $P(X_{t-1}|Z_{t-1})$  に応じ

て選択する．

- (2) パーティクル  $s_t$  の状態を遷移確率  $p(X_t|X_{t-1} = s_{t-1}^k)$  を用いて変化させる．

- (3) 各パーティクルにおける尤度  $p(Z_t|X_t = s_t^k)$  を計算し，それを基に事後確率  $p(X_t|Z_t)$  を推定する．

### 5.2 適応的な重み推定法

本論文では，局所および大局的な観点から得られた局所特徴の再構成誤差を以下のように統合する．

$$E = \sum_{j=1}^N (\gamma_j \epsilon_{lj} + (1 - \gamma_j) \epsilon_{gj}). \quad (4)$$

ここで， $\gamma_j$  は  $j$  番目の局所特徴に関する重みであり， $0 \sim 1$  の値をとる．また， $\epsilon_{lj}$  と  $\epsilon_{gj}$  は局所および大局 KPCA による再構成誤差を表している．

当然，構図情報が有効なサンプルもあれば，そうでないサンプルもある．従って，識別に有効な重みベクトル  $\gamma$  はテストサンプル毎に異なり，事前には分からない．そこで，本論文ではパーティクルフィルタを用いてそのテストサンプルにとって有効な重みを決定する．識別に有効な重みを得るために，パーティクルの状態を 2 つの KPCA により得られた再構成誤差を統合する際の重み  $s_t = \{\gamma_{t1}, \dots, \gamma_{tN}\}$  に割り当てる．重み空間での事後確率分布をパーティクルフィルタを用いて推定し，その分布から重みベクトルを選択すれば，そのサンプルを識別するために有効な重みベクトルを得ることができる．

各パーティクルの  $p(Z_t|X_t)$  の下での識別に関連したスコアを以下のように定義する．

$$Score_c(s_t^k) = \frac{\min_{i \neq c} E_i^k - E_c^k}{\sum_{i=1}^N E_i^k}. \quad (5)$$

ここで， $E_c^k$  は  $k$  番目のパーティクル (重みベクトル) により計算されたクラス  $c$  の再構成誤差であり， $\min_{i \neq c} E_i^k$  はクラス  $c$  を除く全クラス中での最小の再構成誤差である．他のクラスの最小距離と比較することにより，もしクラス  $c$  が最小の再構成誤差を持てばスコアは正の値となり，クラス  $c$  以外が最小の誤差を持てばスコアは負の値となる．また，パーティクルフィルタを各クラスに独立に適用するので，重みベクトルの大きさがスコアに影響を与えるかもしれない．そこで，全クラスの再構成誤差の和で正規化したスコアを用いた．

次に，パーティクルの尤度と重要度について説明する．クラス  $c$  にとって有効な重みを得るためにはクラス  $c$  と他のクラスとの違いを強調すれば良い．従って， $k$  番目のパーティクルの尤度を先程のスコアを用いて以下のように定義する．

$$\tilde{\pi}_t^k = p(Z_t|X_t = s_t^k) \propto \exp(a \times Score_c(s_t^k)). \quad (6)$$

指数関数を用いることにより，大きなスコアを持つパーティクルが選択されやすくなる．また， $a$  は，式 (5) を

$\sum_{i=1}^N E_i^k$  で正規化することにより、スコアが小さくなり過ぎるのを補正するための定数である。実験では  $a$  の値を  $\sum_{i=1}^N E_i^k$  の大まかな値である 300 とした。この尤度  $\hat{\pi}$  は正規化されていないが、実際には全パーティクルの重要度の和が 1 になるように正規化する  $\pi_t^k = \hat{\pi}_t^k / \sum_{k=1}^M \hat{\pi}_t^k$ 。このパーティクルの重要度を用いることにより、各クラスにとってスコアが大きくなるようなパーティクル、即ち重みベクトルを選択することができる。

適応的な重み推定アルゴリズムは以下のようになる。

- (1) パーティクルの状態  $s'_{t-1}$  (重みベクトル) を前の事後確率  $P(X_{t-1}|Z_{t-1})$  に応じて選択する。  $t = 0$  では  $P(X_0|Z_0)$  を一様分布と考え、 $s'_0$  をランダムに選択した。
- (2) パーティクルの状態  $s_t$  を遷移確率  $p(X_t|X_{t-1} = s'_{t-1})$  を用いて移動させる。本論文ではランダムウォークモデル  $X_t = X_{t-1} + n_t$  を用いる。ここで、 $n_t$  はガウシアンノイズである。実験ではガウシアンノイズの分散として 3 つの分散値を用意し、そこから毎回ランダムに選択した。
- (3) 式 (5) を用いて各パーティクルのスコアを計算し、パーティクルの尤度  $p(Z_t|X_t = s_t^k)$  を式 (6) により計算する。パーティクルの重要度  $\pi_t$  を和が 1 になるように正規化する。

(4) (1) ~ (3) のステップを  $t$  が既定回数に達するまで繰り返す。実験では  $t$  の上限を 20 回とした。

このアルゴリズムにより各クラスのスコアが大きくなるような重みベクトルを得ることができる。テストサンプルがどのクラスに属するかは事前には分からないので、このアルゴリズムを全クラスに適用し、クラス毎にスコアが大きくなる重みベクトルを推定する。この方法により、かなり高い割合で正解のクラスが 1 位になる重みを持つことが分かった。しかし、当然、複数のクラスが 1 位になる重みを持つ場合も出てくる。この場合には最終的に 1 つのクラスを選択しなければならないが、クラス毎に異なる重みベクトルが推定されるので、式 (5) のスコアはあまり有効に働かないことが実験により分かった。複数クラスが 1 位になる重みを持つ場合に有効な判定基準は人間には分からないので、適応的な重み推定により得られる 10 個の情報から判定基準を自動的にマイニングすることにした。

## 6. Adaboost を用いた判定基準のマイニング

本節では適応的な重み推定により複数クラスが 1 位になる重みを持った場合に対する最終的な判定基準のマイニングについて説明する。ここでは、特徴選択 [14] で有効性が知られている Adaboost [15] を利用する。

マイニングに利用する情報および弱識別器について 6.1 節で説明する。Adaboost に関する説明は紙面の都合により省略するので、適宜文献を参照されたい [15]。

### 6.1 判定基準のマイニングに利用する情報

適応的な重み推定により再構成誤差や重みベクトル等

のいくつかの情報を得ることができる。しかし、我々人間にはどの情報が複数クラスが 1 位になる重みを持つ場合の最終判定に有効であるのか分からない。そこで、重み推定の際に得られる 10 個の情報から Adaboost を用いて最終判定に有効な基準を自動的に得る。ただし、重み等の値そのものはサンプルに依存して変わってしまうため、全クラス内での順位を用いて弱識別器を構成する。弱識別器の例を挙げれば、再構成誤差が小さい方から 4 位以内とか重みベクトルの和が小さい方から 3 位以内のようになる。

以下に判定基準マイニングに利用する 10 個の情報を示す。ただし、以下は Adaboost をクラス  $c$  に適用することを想定している。また、10 個の情報をループ中のいつ収集するかであるが、パーティクルフィルタの最終ループの中で式 (5) に示すスコアが最大となるパーティクル (重みベクトル) の情報を用いた。

- (1) 式 (4) に示すクラス  $c$  の再構成誤差。
- (2) 上の値を全クラスの再構成誤差の総和  $\sum_{i=1}^N E_i^k$  で正規化した値。これは重みベクトルの大きさに依存しない。
- (3) クラス  $c$  の再構成誤差とクラス  $c$  以外の再構成誤差の最小値との差 ( $\min_{i \neq c} E_i^k - E_c^k$ )。
- (4) 式 (5) に示すクラス  $c$  のスコア。これは重みベクトルの大きさに依存しない。
- (5) 局所 KPCA のみを用いた場合のクラス  $c$  の再構成誤差  $\sum_{i=1}^N \epsilon_{li}$ 。
- (6) 大局 KPCA のみを用いた場合のクラス  $c$  の再構成誤差  $\sum_{i=1}^N \epsilon_{gi}$ 。
- (7) 全クラスの再構成誤差の和  $\sum_{i=1}^N E_i$ 。
- (8) 重みベクトルの和  $\sum_{i=1}^N \gamma_i$ 。
- (9) 重みベクトルの各要素の 2 乗和  $\sum_{i=1}^N \gamma_i^2$ 。
- (10) パーティクルフィルタを用いた適応的な重み推定の最後ループで 1 位になる重みが得られた回数。

Adaboost への入力はこの 10 個の値の全クラス内での順位なので、上述のように弱識別器は情報毎に順位を閾値としたものになる。

適応的な重み推定により 1 つのクラスだけが 1 位になる重みを持つ場合にはそのクラスに識別すれば良い。問題は複数クラスが 1 位になるような場合である。全クラスの学習サンプルの中でそのような場合が起こった場合のみサンプルを収集し、それを基に Adaboost を用いて判定基準のマイニングを行う。複数クラスが 1 位になる重みを持つ場合の最終的な判定基準はクラス間でそれほど違いがないと考えたことと、このような場合が少ないクラスと多いクラスがあることから、クラス毎に基準を変えずに全クラスで 1 つの判定基準を作成した。この判定基準は複数クラスが 1 位になる重みを持つ場合のサンプルから作成したので、そのような場合にのみ適用する。

## 7. 評価実験

本節では 13 シーンデータベース [13] を用いた評価実験の結果を示す。まず初めに実験で用いるデータベースと評価方法を 7.1 節で説明する。次に、マイニングの結果について 7.2 節で述べる。7.3 節では識別結果を示す。従来法との比較結果も併せて示す。

### 7.1 画像データベースと評価方法

本論文では、最近提案されたシーン識別法 [3], [4], [7], [8], [21] と比較をするために、それらの論文で用いられている 13 種類のシーンカテゴリデータベース [13] を利用する。13 種類のシーンカテゴリの例を図 4 に示す。データベースに含まれる画像は濃淡画像であり、画像サイズは一定でない。各シーンカテゴリに含まれる画像枚数も異なっている。室内から屋外シーンまで様々なカテゴリが含まれている。カメラアングル、画像内に含まれる対象やその位置が変化するので、クラス内分散は大きい。

ここでは、各シーンカテゴリの画像を学習用とテスト用の 2 つの画像セットをランダムに選択する。具体的には、各カテゴリでランダムに選択した 100 枚の画像を学習用、残りをテスト用とした。これは従来法 [3], [4], [7], [8], [21] と同じやり方である。

ただし、提案手法では判定基準マイニングのためのサンプルも必要となる。そこで、各カテゴリ 100 枚の学習用画像をさらに 2 つのセットに分割した。最初のセットは各カテゴリ 80 枚の画像を含み、局所および大局 KPCA に利用される。各カテゴリ 20 枚を含む画像セットは判定基準のマイニング用として利用する。つまり、KPCA と Adaboost 用に各クラス計 100 枚の学習画像を用いる。これは従来法と同じである。パーティクルフィルタを用いた適応的加重み推定法を各カテゴリ 20 枚のセットに適用し、不正解のクラスが 1 位になる重みを持つ場合に正と負のサンプルを収集する。この処理により、全クラスで 150 枚以上の正のサンプルと 300 枚以上の負のサンプルが集まった。これらのサンプルから前節で説明した 10 個の情報の順位を抽出し、Adaboost により判定基準を作成する。

各カテゴリの学習以外の画像、即ち 100 枚以外をテストに利用する。しかし、カテゴリ毎にテスト画像の枚数が異なる (最小は 110 枚、最大は 310 枚のテスト画像)。その影響を低減するために各カテゴリ毎の識別率の平均値を評価値とした。この評価方法も従来法と同じである。ここでは、乱数の初期値を変えながらこの評価を 3 回行い、その平均値を最終的な評価値とした。

### 7.2 判定基準のマイニング結果

本節ではマイニングの結果について述べる。前節で述べたように、乱数の初期値を変えながら 3 回の評価を行うので、3 つの判定基準が得られることになる。しかし、

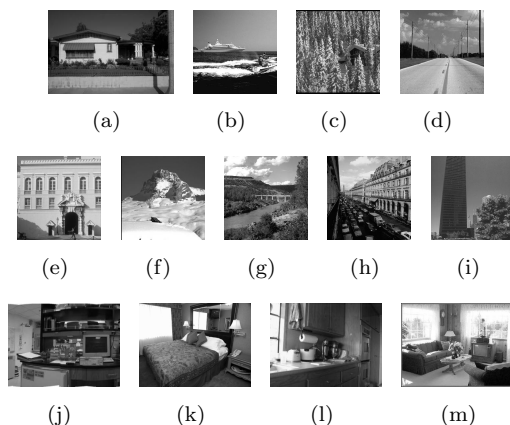


図 4 13 シーンの例 (a) suburb (b) coast (c) forest (d) highway (e) inside-city (f) mountain (g) open-country (h) street (i) tall-building (j) office (k) bedroom (l) kitchen (m) living-room

その中身を解析してみると Adaboost により得られた最初の弱識別器は 3 回とも同じであった。最初の弱識別器に利用された情報は 6.1 節の (7) に示した「全クラスの再構成誤差の和」であり、弱識別器は全クラスの再構成誤差の和が小さい方から 4 位以内であった。クラス毎にパーティクルフィルタを用いて重み推定を行っているので、クラス  $c$  に適用した時に得られる最大スコアに対応する重みベクトルを用いた全クラスの再構成誤差の和と他のクラスにパーティクルフィルタを適用して得られた全クラスの誤差の和は異なる。これが全クラス中で小さい方から 4 位以内という意味である。

2 番目や 3 番目の弱識別器で利用された情報は 3 回の評価で一定ではなかったが、6.1 節の (1) に示した「全クラスの再構成誤差の総和  $\sum_{i=1}^N E_i^k$  で正規化したクラス  $c$  の再構成誤差の値」が 3 回中 2 回の評価で 3 番目までに選択された。弱識別器はともに正規化した再構成誤差が 1 番小さいというものであった。その他に上位で選択されているのは、「最終ループの全パーティクルの中でそのクラスが 1 位になる重みの数が多い」であった。

ここで、得られた判定基準の意味を考える。最初の弱識別器では「全クラスの再構成誤差の和」が使われたが、これは大局 KPCA よりも局所 KPCA を優先することを意味している。なぜなら局所 KPCA による再構成誤差の方が大局 KPCA による再構成誤差よりも小さいからである。つまり、複数クラスが 1 位になる重みを持つような判定が難しいサンプルの場合には大局的な構図のような情報があまり有効ではないことを意味している。

それとは逆に「正規化した再構成誤差の値が小さい」は直接的に識別に関連したものである。つまり、まずは局所 KPCA の誤差を重要視するが、それだけではうまく識別ができないので、識別に関連した指標も利用すると考えることができる。同様に、「全パーティクルの中で 1 位になる重みの個数が多い」というのも識別に深く関

連するものである．その他の下位の弱識別器は誤識別率が大きく，あまり最終判定には影響を与えないことと，3回のマイニングで統一の見解ができないことから説明を省く．

この判定基準は複数クラスが1位になる重みを持つような判定が難しいサンプルに対してだけ有効であり，全てのサンプルに有効なわけではない．適応的重み推定の時点で1意にクラスが決まる場合にはこれを使う必要はない．

### 7.3 評価結果

本節では評価結果及び従来法との比較結果を示す．まず，提案手法を局所 KPCA のみもしくは大局 KPCA のみを用いた場合と比較する．提案手法では全ての  $\gamma_i$  を 0 とすれば大局 KPCA のみを用いることに等しく，全ての  $\gamma_i$  を 1 とすれば局所 KPCA だけを用いた結果が得られる．各手法の評価結果を表 1 に示す．適応的重み推定ではパーティクルの数を 5000 とし，Adaboost によるマイニング結果を利用する提案手法では最大繰り返し回数を 20 とした．また，表中の識別率は乱数の初期値を変えて 3 回評価を行った際に得られる平均値である．

局所 KPCA もしくは大局 KPCA だけを用いた場合には識別精度が 77% に達しないことが分かる．しかし，パーティクルフィルタを用いて適応的に重みを推定することにより識別精度が 81% 以上となった．ただし，複数クラスが 1 位となった場合には式 (5) に示したスコアが最大となるクラスに識別した．この結果から適応的な重み推定の有効性が確認できるのだが，適応的重み推定により正解のクラスが 1 位になる重みを持つ割合は 97.01% (3 回の評価の平均値) である．これは適応的重み推定により正解クラスが最終候補まで残るが最終的に正しく識別される回数が少ないことを表している．つまり，適応的な重み推定は識別を行う方法ではなく，候補を絞り込む方法と位置づけると良い．

適応的な重み推定により候補を絞りこみ，Adaboost により得られた最終判定基準を用いる方法の識別率は 91.06% となり，Adaboost によるマイニングの効果が確認できる．各々の KPCA を単独で用いた場合よりも実に 14% 以上も識別精度が向上した．また，判定基準をマイニングしない場合と比べても 9% 以上の精度が向上している．これによりパーティクルフィルタを用いて適応的に重みを推定し，判定が難しいサンプルに対してはマイニングにより得られた判定基準を利用することの有効性が実証できた．

提案手法の有効性を示すために，シーン識別でよく利用されている bag-of-visual words [22] と線形サポートベクターマシンを組み合わせた場合の評価を行う．基準となる局所特徴量は提案手法と同じ  $2 \times 2$  グリッド上の 4 つの方向ヒストグラムである．この評価結果も表 1 に示す．bag-of-visual words 法は 73% 以下の識別率であり，シー

表 1 提案手法の評価結果

Method	Classification rate
提案手法	91.06%
適応的重み推定のみ	81.33%
局所 KPCA のみ	76.62%
大局 KPCA のみ	74.71%
linear SVM of bag-of-words	72.66%

表 2 従来法との比較結果

Method	Classification rate
提案手法	91.06%
[7] (PAMI2008)	85.9%
[10] (ICPR2010)	84.33%
[9] (Patt.Rec.2011)	82.86%
[4] (ECCV2006)	73.4%
[3] (CVPR2005)	65.2%

ン識別でよく利用される方法よりも提案手法の方が圧倒的に高い精度であることが分かる．

最後に，提案法と従来法 [3], [4], [7], [9], [10] の比較を行う．比較結果を表 2 に示す．ただし，各手法の識別精度は各論文から抜粋した数値である．2 つの方法 [3], [4] は bag-of-visual words を用いているので，ここで作成した bag-of-visual words 法の結果とあまり変わらない．このことは特徴量にそれほどの差がないことを表している．

提案手法は visual words の KPCA 空間内での自己相関特徴を用いた方法 [9] よりも圧倒的に勝っている．また，局所共起特徴を用いた方法 [10] よりも精度が高い．この方法は位置不変性と周辺との関係性を併せ持つ方法であるが，提案手法も性質の異なる 2 つの KPCA を組み合わせることにより 2 つの要求を満たしている．また，従来法の中で最も識別精度が高かった spatial pyramid probabilistic Latent Semantic Analysis を用いた Boschらの方法 [7] よりも高い精度が得られている．

識別精度は使用する特徴量と識別器に依存しているため，両方共に異なる方法とは直接的な比較が難しい．しかし，標準的なシーン識別法である bag-of-visual words 法を提案手法と同じ特徴量を用いて識別したものに比べ，提案手法は圧倒的に高い識別率であり，また同じデータベースを用いて評価を行ったどの方法よりも高い精度が得られていることから，適応的な重み推定により候補を絞りこみ，判定が難しいサンプルには Adaboost により得られた判定基準を用いることが有効であったと言える．

## 8. おわりに

シーン識別では位置に依存した大まかな構図情報と位置不変な細かい類似性が重要である．そこで，提案手法では位置不変な細かい対象を表現する方法として局所 KPCA を用い，大まかな構図情報を表現するために大局

KPCA を用いた . そして , 2 つの KPCA により得られた再構成誤差をテストサンプルに応じて適応的に統合する方法を提案した . しかし , パーティクルフィルタを用いて各クラスが 1 位になるような重み選択を行うと , 複数クラスが 1 位になる重みを持つ場合がある . その際にはクラス間で重みベクトルが異なるため , 単純に再構成誤差を用いてもあまり精度の良い識別はできない . そこで , このような場合の判定基準を Adaboost を用いてマイニングした .

13 シーンのデータベースを用いた実験により , 適応的な重み推定を用いると局所もしくは大局 KPCA を単独で用いるよりも 5% 程度精度が向上することが分かった . さらに , 判定が難しいサンプルに対してマイニングの結果を用いると 14% 以上も識別精度が向上した . 同じデータベースを用いて評価を行った最も精度の高い従来法 [7] と比較しても提案手法は 5% 以上高い識別精度が得られ , 適応的な重み推定と判定基準のマイニングを用いた提案手法の有効性が実証できた .

本論文では提案手法をシーン識別に利用したが , 提案手法自体は汎用的な枠組であり , 識別対象に依存しない . 例えば , 対象カテゴリ識別では大局 KPCA が文脈情報を表し , 局所 KPCA が位置不変な対象の類似度を表すことになり , 有効に働くと考えられる . また , 同じような枠組が行動認識等でも利用できると期待される . これらの問題への応用は今後の課題である .

謝 辞

本研究の一部は人工知能研究振興財団研究助成および科学研究費 (21700189) の補助によるものである .

## 文 献

- [1] A. Torralba, "Contextual priming for object detection," *International Journal of Computer Vision* **53**(2), pp. 169–191, 2003.
- [2] T. Ishihara, K. Hotta, and H. Takahashi, "Estimation of object position based on color and shape contextual information," in *Proc. International Conference on Image Analysis and Processing, LNCS Vol.5716*, pp. 57–62, 2009.
- [3] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in *Proc. IEEE CS Conference on Computer Vision and Pattern Recognition*, pp. 524–531, 2005.
- [4] A. Bosch, A. Zisserman, and X. Munoz, "Scene classification via plsa," in *Proc. European Conference on Computer Vision*, pp. 517–530, 2006.
- [5] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE CS Conference on Computer Vision and Pattern Recognition*, pp. 2169–2178, 2006.
- [6] J. C. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. Smeulders, "Kernel codebooks for scene classification," in *Proc. European Conference on Computer Vision*, pp. 696–709, 2008.
- [7] A. Bosch, A. Zisserman, and X. Munoz, "Scene classification using a hybrid generative/discriminative approach," *IEEE Trans. Pattern Analysis and Machine*

*Intelligence* **30**(4), pp. 712–727, 2008.

- [8] K. Hotta, "Scene classification based on multi-resolution orientation histogram of gabor features," in *Proc. International Conference on Computer Vision Systems, LNCS Vol.5008*, pp. 291–301, 2008.
- [9] K. Hotta, "Local autocorrelation of similarities with subspaces for shift invariant scene classification," *Pattern Recognition* **44**(4), pp. 794–799, 2011.
- [10] K. Hotta, "Scene classification using local co-occurrence feature in subspace obtained by kpca of local blob visual words," in *Proc. International Conference on Pattern Recognition*, pp. 4230–4233, 2010.
- [11] K. Hotta, "Integration of reconstruction error obtained by local and global kernel pca with different role," in *Proc. International Conference on Computer Graphics, Visualization and Computer Vision*, pp. 91–97, 2011.
- [12] K. Hotta, "Adaptive weighting of local classifiers by particle filter for robust tracking," *Pattern Recognition* **42**(5), pp. 619–628, 2009.
- [13] *13 Scene categories database*. <http://vision.cs.princeton.edu/Datsets/SceneClass13.rar>.
- [14] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision* **57**(2), pp. 137–154, 2004.
- [15] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences* **55**(1), pp. 119–139, 1997.
- [16] K. Grauman and T. Darrell, "Discriminative classification with sets of image features," in *Proc. International Conference on Computer Vision*, pp. 1458–1465, 2005.
- [17] B. Schölkopf, C. Burges, and A. Smola, *Advances in kernel methods: support vector learning*, MIT Press, 1998.
- [18] K. Hotta, "Non-linear feature extraction by linear principal component analysis using local kernel," *Pattern Recognition Recent Advances*, pp. 99–109, 2010.
- [19] K. Hotta, "Local normalized linear summation kernel for fast and robust recognition," *Pattern Recognition* **43**(3), pp. 906–913, 2010.
- [20] R. Debnath, N. Takahide, and H. Takahashi, "A decision based one-against-one method for multi-class support vector machine," *Pattern Analysis and Applications* **7**(2), pp. 164–175, 2004.
- [21] K. Hotta, "Scene classification based on local autocorrelation of similarities with subspaces," in *Proc. IEEE International Conference on Image Processing*, pp. 2053–2056, 2009.
- [22] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proc. ECCV Workshop on Statistical Learning in Computer Vision*, pp. 1–16, 2004.