I-030

画像特徴に基づく複数基底セットを用いた画像符号化方式
A Study of Image Coding using Multiple Bases Sets based on Image Features

王 冀† 八島 由幸† 高木 基宏‡ 藤井 寛‡ 清水 淳‡
Ji Wang Yoshiyuki Yashima Motohiro Takaki Hiroshi Fujii Atsushi Shimizu

1． はじめに

従来の画像符号化技術では，入力画像ごとに個別に冗長性を除去するのが一般的である．しかし，世の中にはスポーツやニュースのように類似性のある画像が多く存在する．たとえば同一人物が違う場所で撮影した画像のように異なる画像間にも冗長性が存在する．このような画像間の類似性を除去するためには，多くの類似画像をあらかじめ学習し共通基底で表現できるようにしておくことで，符号量の削減を図ることができると考えられる[1]．本研究では DSIFT 特徴によって分類された画像セットに対してそれぞれのカテゴリの性質を反映した複数の基底セットを設計し，これを符号化の際に適応的に利用する画像符号化方式を提案する．今回，基底セットを主成分基底（以降 PCA 基底）としてシミュレーション実験を行い，基底作成のためのブロックサイズおよび基底セット数と圧縮効率の関連についての検討結果を報告する．

2． 処理手順

2.1 処理の流れ

図 1 に符号化処理の流れを示す．まず，種々の画像に対して基底セットを設計する．基底学習においては，Dense SIFT(DSIFT) [2]により小ブロックごとに局所特徴量を算出してこれらを N 個のクラスに分類した後，クラスごとにその性質を反映した基底を設計する．設計された基底セットおよび各クラスの特徴ベクトル代表値は，エンコード側およびデコード側で共有しておく．符号化時には，入力画像をブロックに分割し，符号化対象ブロックの DSIFT 特徴量を各クラスの特徴ベクトル代表値と比較してクラスを決定した後，符号化対象ブロックを該クラスで設計された基底の線形和として表現しその重み係数を符号化する．また，選択されたクラス番号も同時に符号化する．復号側では重み係数を復号した後，該当するクラスの基底を用いて画素値を復号する．以下それぞれの処理の内容を説明する．

2.2 局所特徴適応基底セットの作成

画像分類でよく用いられている特徴算出手法とする SIFT[3]では，画像中の特徴点（キーポイント）のみの画素勾配情報を利用する．しかし画像符号化等への応用においては，特徴点のみではなくパターンの少ない平坦部やエッジ部の情報も必要であるので，本実験では一定の間隔および範囲で特徴計算をする DSIFT を利用する．特徴計算を行う点(u,v)に関して勾配強度m(u, v)と勾配方向θ(u, v)を以下の式により求める．ここで，$f_u(u,v)$および$f_v(u,v)$はそれぞれ水平・垂直方向の画素勾配を示す．

$$m(u,v) = \sqrt{f_u(u,v)^2 + f_v(u,v)^2} \quad (1)$$

$$\theta(u,v) = tan^{-1}(f_v(u,v)/f_u(u,v)) \quad (2)$$

†（学）千葉工業大学大学院 情報科学研究科
‡（社）NTT メディアインテリジェンス研究所

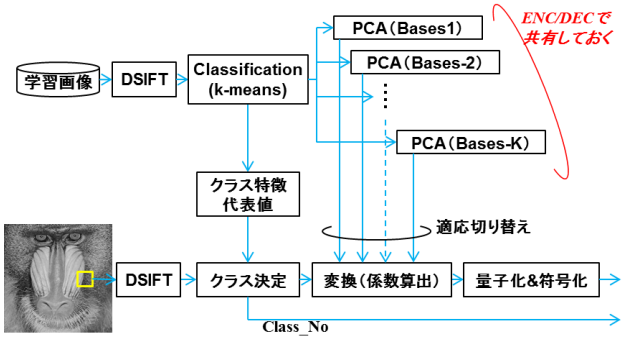

図 1 符号化処理の流れ

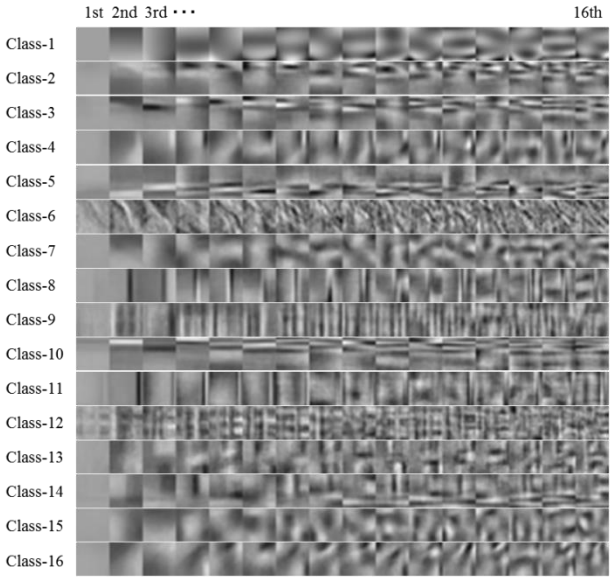

図 1 基底画像セットの例（上位 16 個）

符号化ブロック毎に縦横それぞれ 4 分割した 16 の領域それぞれに対して勾配方向θ(u, v)を 8 方向に量子化し，それぞれの方向に勾配強度m(u, v)を投票したヒストグラム（4 × 4 × 8 =128 次元のベクトル）で特徴量を記述する．

次に，多くの学習用画像を準備し，これらから得られた特徴量ベクトルを基に基底セットを作成する．図 1 に示すように，学習画像から得られた全 N 個のブロックに対して，前述の DSIFT を用いて，特徴ベクトル$\mathbf{V}(128 \times N)$を抽出する．抽出された$\mathbf{V}$に対して k-means クラスタリングを実行し，各クラスに分類された特徴ベクトル$\mathbf{V_k}$の平均ベクトルをクラス代表特徴（Visual words）とする．

最後に，各クラスに分類された画像ブロックデータを主成分分析する．あるクラスに属するデータを$\mathbf{x}_i$，データの個数をN_c，$\mathbf{x}_i$の平均ベクトルを$\mathbf{m}$とし，$\mathbf{x}_i$の分散共分散行列$\mathbf{X}$を作成し，$\mathbf{X}$の固有値と固有ベクトルを算出する．

$$\boldsymbol{X} = {}^{1}/_{N_c} \sum_{i=1}^{N_c} (\boldsymbol{x}_i - \boldsymbol{m})(\boldsymbol{x}_i - \boldsymbol{m})^T \quad (3)$$

$$\boldsymbol{X}\boldsymbol{d}_k = \lambda_k \boldsymbol{d}_k \quad (4)$$

表 1 実験条件

学習用画像	2688 枚（256×256）
ブロックサイズ	8×8，16×16，32×32
クラス数	1，64，128，256
量子化ステップ	0.15，0.1，0.05，0.03，0.02

ここで，λ_kは行列$\mathbf{X}$の k 番目に大きい固有値，$\boldsymbol{d}_k$はλ_kに対する固有ベクトルであり，固有ベクトル$\boldsymbol{d}_k$がそのクラスの基底画像になる．図 2 は，ブロックサイズ16×16，クラス数 256 として設計した基底画像セットの例である（256 クラスのうち 16 クラスのみを掲載）．

2.3 符号化

符号化対象画像をブロック分割し，分割されたブロック画像ごとに PCA 基底を選択して変換する．具体的には，ブロック$\boldsymbol{x}$毎に DSIFT 特徴量を算出し，得られた特徴ベクトルとすべてのクラスの Visual words との二乗誤差を計算する．誤差の一番小さい Visual words に対応する PCA 基底を選択し，式(5)により重み係数$\mathbf{z}_k$を算出する．

$$\boldsymbol{z}_k = \boldsymbol{d}_k^T \boldsymbol{x} \qquad (5)$$

$\mathbf{z}_k$を量子化して選択された基底番号とともに符号化し復号側に送る．復号側では，重み係数と基底番号がわかれば，基底セットと平均ベクトルはあらかじめ共有しているため，画像の再構成が可能となる．

3. 実験と考察

提案方式の性能を把握するため，符号化ブロックサイズおよび基底クラス数をパラメータとしてシミュレーション実験を行った．実験条件を表 1 に示す．いずれの場合も，クラス数が 1 の場合には DCT とほぼ同様の基底画像が作られており，クラス数を増やすごとに，ブロック特徴に応じた基底が作成されることを確認した．

図 3 に，入力画像（学習データ外，画像サイズ 2144×1424）に対して，ブロックサイズを 16×16 とし，基底セット数（クラス数）を 1，64，128，256 と変化させた場合の符号化特性を示す．符号量は基底に対する重み係数のエントロピーで計算し，クラス切り替え情報も含まれている．加えて，図 4 には量子化ステップ 0.05 の場合の復号結果例を示した．これらの実験結果から，基底切り替えの有効性とともに，使用されたクラス数が 128 程度でその効率は収束していることがわかる．

一方，図 5 は，クラス数を 256 に固定した場合の，ブロックサイズ（8×8，16×16，32×32）と圧縮効率の関係の測定結果を示している．広い範囲にわたってブロックサイズ 16×16 が良い結果を示しているが，レートが高くなるほど小さなブロックサイズが有効になり，レートが低くなると大きなブロックサイズが有効である傾向が見られる．

4. まとめ

DSIFT 特徴によって分類された画像セットに対してそれぞれのカテゴリの性質を反映した複数の基底セットを設計し，これを符号化の際に適応的に利用する手法を提案し性能を評価した．今回は PCA 基底を用いたが，より効率的な画像表現を狙いスパースコーディングを適用し，複数のスパース基底セットを用いる手法も検討する予定である．

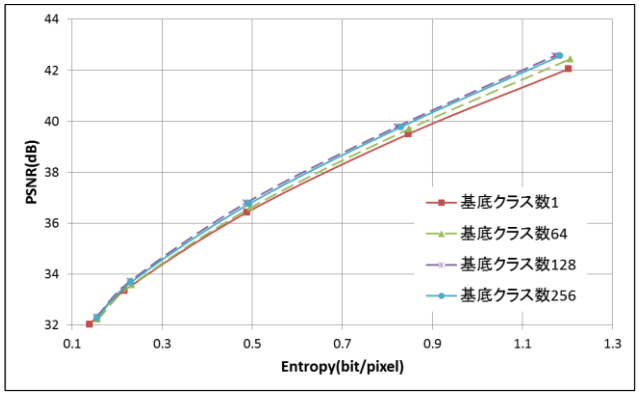

図 2 クラス数の変化による符号化特性

図 3 復号結果の比較（左：原画像，中：クラス数 1 の復号画像，右：クラス数 128 の復号画像）

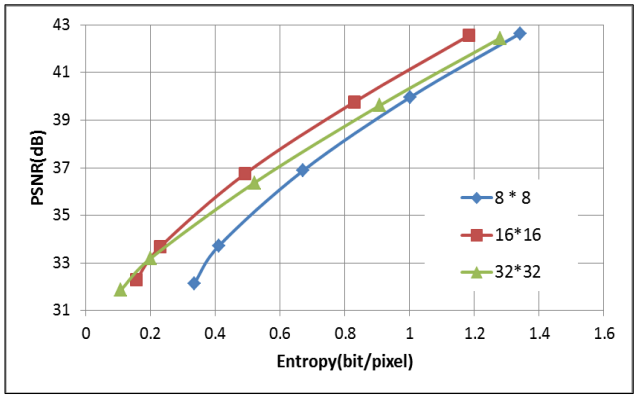

図 4 ブロックサイズの変化による符号化特性

参考文献

[1] 塚原 正人，長谷山 美紀，北島 秀夫．画像符号化におけるエッジ保存を考慮した適応 KL 基底．電子情報通信学会技術研究報告．IE，画像工学 97(429)，45-49，1997-12-11

[2] A. Vedaldi and B. Fulkerson．Vlfeat：An open and portable library of computer vision algorithms．In Proc. int. conf. on Multimedia，pp. 1469-1472，2010．Available：www.vlfeat.org/.

[3] D. Lowe：Distinctive image features from scale-invariant keypoints．Int. J. Computer. Vision，60(2)：91–110，2004.