前景と背景情報を利用した物体認識の高精度化に関する研究

伊原 有仁　　指導教授：藤吉 弘亘

1. はじめに

近年，一般物体認識の研究として画像分類問題が盛んに取り組まれており，Bag-of-features[1] が有効とされている．この手法は，カテゴリ (前景) を認識する画像分類に対して，画像全体から抽出した局所特徴量を用いているため，背景情報がノイズとされている．しかし，背景情報のみでも高精度なシーン認識ができることも報告されており，画像分類に有効な情報が何であるか明確ではない．

そこで，本研究では，背景情報の画像分類への有効性を調査し，局所特徴量の識別を用いた前景情報と背景情報の共起表現による画像分類法を提案する．

2. Bag-of-features による画像分類

2.1 Bag-of-features

画像分類とは，画像中の物体やシーンのカテゴリを求める問題である．この問題に対して，画像を局所特徴量の集合として表現し画像分類を行う手法として，Bag-of-features が提案されている．Csurka らによって提案された Bag-of-features は，画像をキーポイントが持つ局所特徴量の集合と考え，画像 1 枚をベクトル量子化ヒストグラムによって表現し，画像分類を行う．局所特徴量には，画像の拡大・縮小，回転，照明変化に頑健な SIFT 特徴量 [2] が多く用いられている．ベクトル量子化には，visual word dictionary と呼ばれる辞書を参照する．visual word dictionary は，画像から抽出した局所特徴量をクラスタリングして得られた代表ベクトル群で表現される．作成したベクトル量子化ヒストグラムを学習，識別することで画像分類を行う．

2.2 Bag-of-features における背景情報の有効性

画像分類には前景情報だけでなく，背景情報も必要であることが考えられる．そこで，画像全体，前景のみ，前景と背景領域の 3 つの特徴量抽出法を使用して調査を行う．局所特徴量として，10pixel 毎のグリッドでキーポイントを与え，10 から 50pixel まで 10pixel 間隔のスケールで SIFT 特徴量を抽出する．識別器には傾向調査が容易である AdaBoost を用いて画像分類を行う．このときの弱識別器の数を 100 個とする．使用するデータは画像分類のためのデータベースである Caltech256 から，butterfly, elephant, hawksbill, helicopter, motorbikes, airplanes, car-side, faces-easy, toad の 9 カテゴリを用いる．前景と背景を分けるために，マスク画像を手動で用意した．各カテゴリから 30 枚を学習用画像とし，残りの 2,411 枚を評価用画像とする．visual word の数は 100 とし，評価には f-measure を用いる．

認識率から得られる有効な特徴の傾向

各特徴抽出法の認識率を比較し，背景情報の有効性を確認する．図 1(a) に認識結果を示す．画像全体を用いた場合の精度が最も悪く，前景と背景を分けた場合に最も精度が良い結果となった．従って，前景と背景に分けて特徴を抽出することが有効であると言える．

背景情報の有効性

前景と背景の両方が混在，投票が無い，前景のみ，背景のみ，以上の各パターンで構築される visual word の比率を求め，背景情報が利用されているか確認する．図 1(b) に，各特徴の比率を示す．図 1(b) から，前景と背景の両方が混在した visual word が選択されていることがわかる．従って，画像全体から特徴記述をしても，前景情報だけでなく背景情報も用いて画像分類をしていることがわかる．

以上のことから，高精度な画像分類には背景情報も必要であることが判明した．従って，次に前景と背景情報を用いた画像分類手法を提案する．

3. 提案手法：前景と背景の共起表現による画像分類

提案手法では，キーポイントを自動で前景と背景に識別する．前景と背景の共起表現として，2 次元ベクトル量子化ヒストグラムを作成し，画像分類における精度向上を図る．提案手法の流れを図 2 に示す．学習では，キーポイント識別器とカテゴリ識別器の 2 つの識別器を構築する．キーポイント識別器は，マスク画像により分けたキーポイントを学習用データとして使用する．カテゴリ識別器は，マスク画像により分けたキーポイントから，2 次元ベクトル量子化ヒストグラムを作成し，学習用データとして使用する．未知入力画像に対しては，予め構築したキーポイント識別器を用いて，キーポイントを前景と背景に識別する．そして，識別された前景と背景のキーポイントから，2 次元ベクトル量子化ヒストグラムを作成し，予め構築したカテゴリ識別器により，画像のカテゴリを識別する．

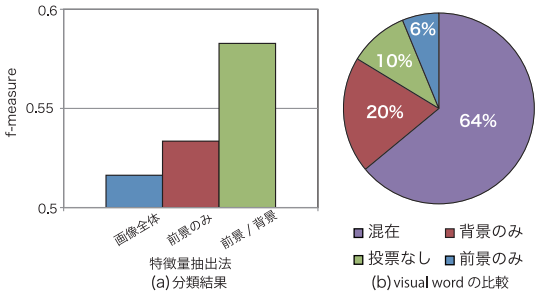

図 1：分類結果と visual word の比率

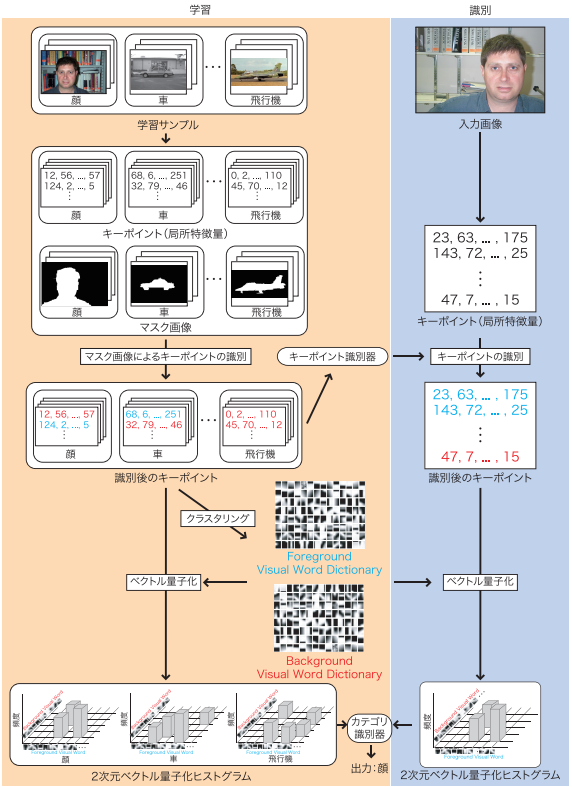

図 2：提案手法の流れ

3.1 キーポイント識別器の構築

キーポイントを自動で前景と背景に分けるために，キーポイント識別器の構築を行う．識別器の構築には，2.2 で用いたデータセットと同様のものを使用する．各カテゴリからランダムに 30 枚を選び，合計 270 枚から検出したキーポイントを学習用データとして使用する．識別器には SVM を使用し，キーポイント識別器を構築する．図 3 に，butterfly と airplanes のキーポイント識別例を示す．青色が前景，赤色が背景と識別されたキーポイントである．誤識別があるが，大まかに識別できていることがわかる．評価用データ 2,411 枚を用いた実験より，約 77.98%の識別率を得た．

入力画像　　キーポイント識別画像

図 3：キーポイント識別例 (青：前景，赤：背景)

3.2 前景と背景の visual word dictionary の作成

前景と背景の visual word dictionary の作成には，マスク画像により識別した前景と背景のキーポイントを使用する．作成には k-means クラスタリングを使用する．

3.3 2 次元ベクトル量子化ヒストグラムの作成

作成した前景と背景の visual word dictionary を参照して 2 次元ベクトル量子化ヒストグラムを作成する．前景のキーポイントは前景の visual word dictionary を参照し，背景のキーポイントは背景の visual word dictionary を参照することで，2 次元空間上で交差したビンに投票を行う．このとき，前景のキーポイントが持つスケールの 6 倍の領域に存在する背景のキーポイントと対応付けて投票を行う (図 4)．学習では，マスク画像で識別したキーポイントを用いるのに対し，未知入力画像では，キーポイント識別器を用いて 2 次元ベクトル量子化ヒストグラムを作成する．

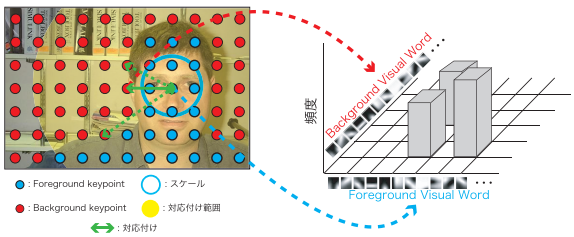

図 4：キーポイントの対応付け

さらに，前景内部や前景から離れた背景のキーポイントは対応付けが困難である．そのため，対応付けの範囲内にある前景と背景のキーポイントの数を調べ，その比率によって 1 次元ベクトル量子化ヒストグラムを作成し，2 次元ベクトル量子化ヒストグラムと組み合わせる．

3.4 カテゴリ識別器の構築

2 次元ベクトル量子化ヒストグラムを用いてカテゴリ識別器を構築する．学習用データは，キーポイント識別器の構築で用いたものを使用する．識別器には，画像分類問題で多く使用されている SVM を使用する．

3.5 カテゴリ識別

未知入力画像に対して，カテゴリの識別を行う．まず，入力画像にグリッドでキーポイントを与え，予め構築したキーポイント識別器を用いて，前景と背景のキーポイントに識別する．次に，識別した前景と背景のキーポイントを対応付けて，2 次元ベクトル量子化ヒストグラムを作成し，カテゴリ識別器によって，画像のカテゴリを識別する．

4. 評価実験

提案手法の有効性を示すために画像分類実験を行う．

4.1 画像分類実験

画像分類実験では，2,411 枚の評価用データを用いて画像分類を行う．提案手法では，対応付け範囲内のキーポイントのみを用いる方法を提案手法 1，提案手法 1 と比率による 1 次元ベクトル量子化ヒストグラムを組み合わせた方法を提案手法 2 とする．比較手法として，従来法である 1 次元ベクトル量子化ヒストグラムを用いる．従来法の特徴量抽出法として，前景と背景に分けて 1 次元ベクトル量子化ヒストグラムを作成する方法を用いる．これにより，共起表現を用いる提案手法の有効性を見ることができると考えられる．評価方法として，f-measure により評価する．

4.2 画像分類結果

図 5 に画像分類結果を示す．従来法に比べ，提案手法 1 では約 5.1%精度が向上した．さらに，提案手法 2 では，前景内部や前景と離れた背景の情報を捉えることができたため，提案手法 1 に比べ約 0.2%精度が向上している．

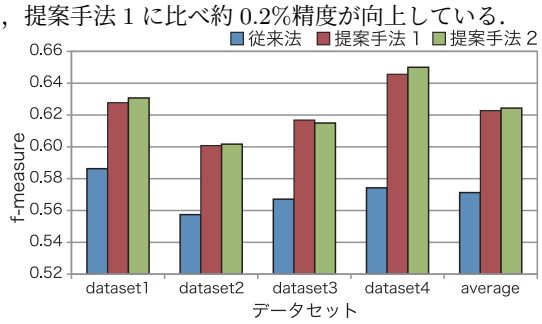

図 5：画像分類結果

図 6 に，学習に用いた 2 次元ベクトル量子化ヒストグラムの可視化例を示す．黒いほど確率が高くなっている．butterfly と toad は背景が似ているため，1 次元では前景の visual word の差が識別に有効だが，2 次元では共起表現により，前景の差が背景の visual word の頻度にも反映される．従って，カテゴリ間の差をより強く表現した特徴量を抽出することができるため精度が向上すると考えられる．

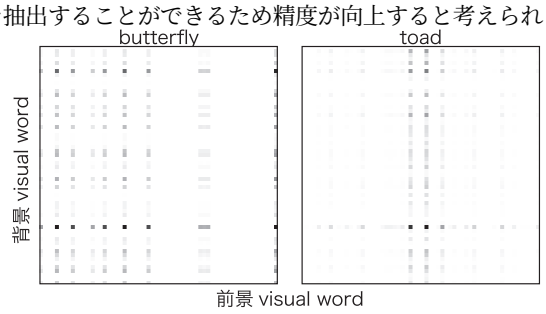

図 6：2 次元ベクトル量子化ヒストグラムの可視化例

4.3 対応付ける領域の変化による精度の違い

提案手法では，キーポイントを対応付ける範囲を前景キーポイントが持つスケールの 6 倍とした．ここでは，対応付ける範囲を 2 倍，4 倍，6 倍，8 倍，10 倍と変化させ，さらに全ての点と対応させた場合の画像分類精度を調べる．図 7 に結果を示す．結果から，局所的な制限が有効であることがわかる．しかし，2 倍，4 倍においても精度が低下しているため，制限を強くすると背景のコンテキストが有効に使用できないと言える．

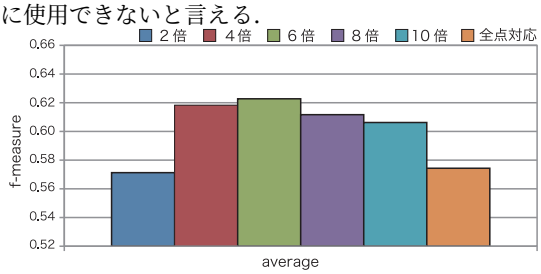

図 7：対応付けの領域の変化による精度の違い

5. おわりに

本稿では，背景情報も有効であることを示し，前景と背景情報の共起表現による画像分類手法を提案した．提案手法を用いた画像分類実験により，共起表現による提案手法の有効性を確認した．今後の課題として，キーポイント識別器の精度向上が挙げられる．

参考文献

[1] G. Csurka, C. Bray, C. Dance, and L. Fan, “Visual Categorization with Bags of Keypoints”, Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pp. 1-22, 2004.

[2] D. Lowe, “Distinctive image features from scale-invariant keypoints”, International Journal of Computer Vision(IJCV), 60(2), pp. 91-110, 2004.

研究業績

[1] 伊原有仁, 藤吉弘亘, 高木雅成, 公文宏明, 玉津幸政, “ 交通道路標識認識における異なる部分空間を用いた対応点マッチングの高精度化 ”, 電学論C, Vol. 129, No. 5, pp.893-900, 2009 .

（他 国際会議論文 1 件，研究会技術報告 2 件）