

1. はじめに

近年、画像分類問題において Bag-of-features モデルが提案され、多くの研究が取り組まれている。Bag-of-features は、画像を局所特徴量の集合とみなし、局所特徴量のヒストグラムをその画像の特徴量としてカテゴリ分類する手法である。そのため、分類性能は局所特徴の抽出法や記述法の影響を受けやすい。そこで、本研究では、局所特徴量のサンプリング手法の違いによる分類性能への影響について明らかにすることを目的とする。

2. Bag-of-features による画像分類

Bag-of-features[1] による画像のカテゴリ分類では、はじめに、局所特徴量を抽出し、それをベクトル量子化する。ベクトル量子化に用いるコードブックは、visual word dictionary と呼ばれ、予め学習画像から作成しておく。ベクトル量子化された特徴量を用いて、ベクトル量子化ヒストグラムを作成する。得られたベクトル量子化ヒストグラムをその画像の特徴量として、識別器に入力することでカテゴリに分類する。図 1 に、Bag-of-features の流れを示す。

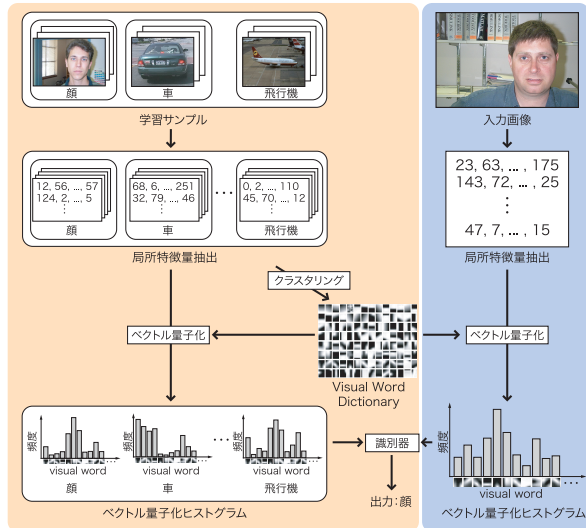


図 1: Bag-of-features による画像分類の流れ

3. 局所特徴量とサンプリング手法

本手法は、局所特徴量として SIFT 特徴量 [2] を用いる。以下に、SIFT 特徴量とそのサンプリング手法について述べる。

**SIFT 特徴量** SIFT 特徴量 [2] は、ある位置とスケールからなる特徴点（キーポイント）の代表輝度勾配方向を決定し、その方向を基準とした輝度勾配ヒストグラムを用いて多次元ベクトルで特徴を記述する（図 2）。SIFT 特徴量

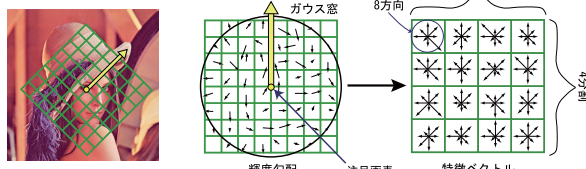


図 2: SIFT 特徴量の記述

は、画像の回転や輝度変化に対して頑健な局所領域の記述が可能のため、Bag-of-features の局所特徴に有効とされている。

**キーポイントのサンプリング手法** Bag-of-features に用いられるキーポイントのサンプリング法として以下の手法がある。

**DoG** Difference of Gaussian の極値をキーポイントとし、その領域から SIFT 特徴量により記述する。

**グリッドサンプリング** 10pixel ごとのグリッド点をキーポイントとして 10 から 30 の間でランダムでスケールを選択し、その点に対し SIFT 特徴量により記述する。

**マルチスケールグリッドサンプリング** 10pixel ごとのグリッド点をキーポイントとして 10, 20, 30, 40, 50 のスケールから SIFT 特徴量により記述する。

本研究では、上記のキーポイントのサンプリング法を比較する。

4. 評価実験

**実験概要** 画像データベース Caltech256 を用いてキーポイントのサンプリング法を評価する。対象カテゴリは、butterfly, elephant, hawkbill, helicopter, motorbikes, air-planes, car-side, faces-easy, toad の 9 カテゴリとする。各カテゴリの 30 枚を学習用画像として選択し、残りを評価用画像とする。また、ベクトル量子化に用いる visual word dictionary のサイズは 100, 500, 1000 とする。識別器には、AdaBoost を使用し、弱識別器の数を 100 とする。評価方法として、f-measure を用いる。

**実験結果** 各サンプリング手法による分類率を図 3 に示す。図 3 より、マルチスケールのグリッドサンプリングが

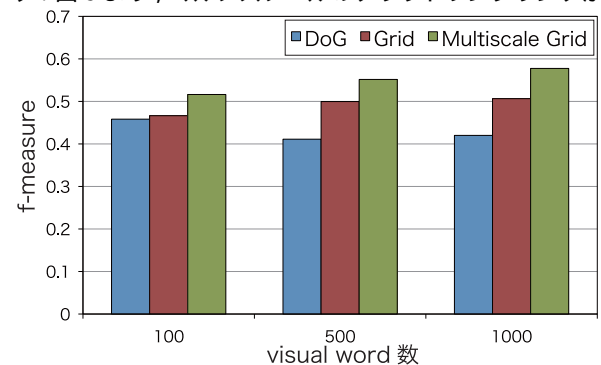


図 3: 各サンプリング手法の分類率

最も高い分類率が得られ、visual word 数が 1000 のとき約 58%であった。次に、AdaBoost の弱識別器からカテゴリ分類に有効な特徴量を解析し、可視化することで局所特徴の位置に対する傾向を調査する。図 4 に、DoG とマルチスケールグリッドサンプリングの可視化結果を示す。可視

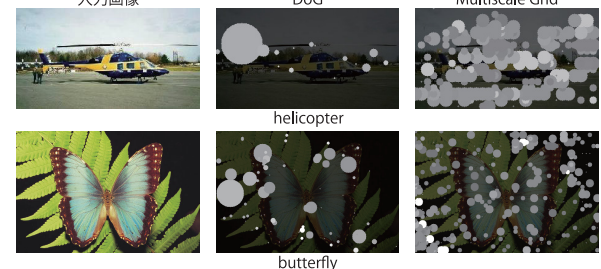


図 4: カテゴリ分類に有効な特徴量の可視化結果

化結果から、グリッドサンプリングは DoG より多くの局所特徴量が選択されていることがわかる。グリッドサンプリングは、DoG では検出できないテクスチャがない位置からも特徴抽出するため、分類結果が向上したといえる。

5. おわりに

Bag-of-features における局所特徴のサンプリング手法を評価した。評価の結果、マルチスケールグリッドサンプリングが最も良い結果が得られた。これは、DoG では検出されないテクスチャがない位置からも特徴抽出できるため、分類結果の向上に貢献することが判明した。

参考文献

[1] G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray: “Visual categorization with bags of keypoints”, ECCV In Workshop on Statistical Learning in Computer Vision, pp. 1–22 (2004).  
 [2] D. G. Lowe: “Distinctive image features from scale-invariant keypoints”, IJCV, 60, 2, pp. 91–110 (2004).