

Seam-Carving を用いた画像の学習効率および認識率向上に関する検討

佐藤 元昭[†] 甲藤 二郎[‡]

[†]早稲田大学 理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: † { satoh , katto } @katto.comm.waseda.ac.jp

あらまし 近年、デジタル画像の爆発的な増大によって、一般画像認識技術への要求が高まっている。そこで、認識の手法として G.Csurka らによって提案された bag-of-keypoints 手法 (Bok 法) を用いて、画像分類実験を行った。この時、学習画像、テスト画像の両方に対して、事前に seam-carving を用いて認識に不要であると思われる領域の除去を行っておくことで、画像の学習効率および認識率の向上が図れるのではないかと考え、実験を行った。その結果、事前に何も処理をしなかった場合に比べ、4.5%程識別率を向上させることが出来た。

キーワード 物体認識、画像分類、シームカービング

A Study on Improvement of Learning Efficiency and Recognition Rate of Images by using Seam-Carving

Motoaki SATO[†] and Jiro KATTO[‡]

[†] Dept. of Computer Science, Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555

E-mail: † { satoh , katto } @katto.comm.waseda.ac.jp

Abstract Recently, a demand for generic image recognition is increasing due to explosive increase of digital images. In this paper, we carried out an experiment by using bag-of-keypoints method (Bok method) proposed by G.Csurka as a technique of image recognition. In this experiment, we tried with an assumption that, by removing an area that seemed to be unnecessary for recognition by applying seam-carving to both the learning images and the test images, learning efficiency and recognition rate could be improved. As a result, the recognition rate can be improved by about 4.5% compared with the case where nothing is processed beforehand.

Keyword object recognition, image classification, seam-carving

1. はじめに

近年、デジタルカメラの普及などにより、画像中に撮影されている物体の内容を計算機に自動で認識させる「一般物体認識」に対する要求が高まっている。この一般物体認識においては、事前に内容が既知である大量の学習画像が必要となるが、このような学習データセットを事前に人手で構築することは困難であるため、認識対象となる物体の種類が制限された学習データセットを用いて研究を行うことが一般的となっている。しかし、このようなデータセットに含まれている画像の中には、認識対象以外の背景などを多く含んだ画像も多く存在している。本研究では、このことが物体の認識を困難にしている一つの要因であると考

え、このような画像から seam-carving[1]を用いて無駄な背景を取り除くことで、画像の学習効率および認識率の向上させることを目的とする。

2. 提案手法

2.1. 提案手法の処理の流れ

以下に提案手法の処理の流れを示す。また、処理の概要図を図1に示す。

- ① 学習画像群とテスト画像群を用意する。
- ② ①で用意した画像群に、seam-carvingによるリサイズを行い、seam-carvingを行った画像と行っていない画像を用意する。

- ③ ②で用意した画像全てに対して、Bok法[2]による特徴ベクトルの生成を行う。
- ④ 学習画像から生成した特徴ベクトルを学習データ、テスト画像から生成した特徴ベクトルをテストデータとして SVM (サポートベクターマシン) によって学習と識別を行う。

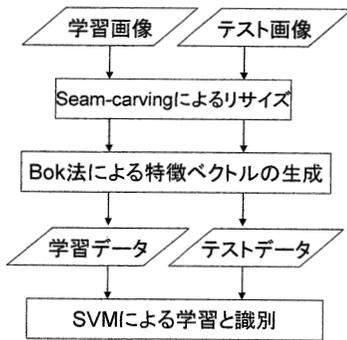


図 1. 提案手法の流れ

2.2. Bag-of-keypoints 法 (Bok 法) [2]

Bag-of-keypoints [2] 法は、統計的言語処理における bag-of-words のアナロジーで、テキスト表現手法の bag-of-words の考え方を画像に適用させた手法となる。bag-of-words が、語順を無視して文章を単語の集合と考えるのと同様に、bag-of-keypoints では位置を無視して、画像を局所特徴 (keypoints) の集合として考える。Bag-of-keypoints は Bag-of-visual-words とも呼ばれる。

全体の流れは、

- ① 画像の局所特徴を表現する局所特徴ベクトルである keypoints を画像から抽出する。この局所特徴ベクトルの表現には SIFT[3]法が用いられることが一般的となっている。
- ② 得られた keypoints に対して、ベクトル量子化を行う。本研究では、ベクトル量子化に単純で最もよく使われて

いる k-means 法を用い、クラスタリングされた各クラスターの重心を、代表ベクトルとして生成する。この各クラスターの重心である代表ベクトルを visual word と呼ぶ。

- ③ visual word をまとめたものを code book と呼び、codebook を用いて画像の特徴ベクトルを生成する。
- ④ 学習画像から生成された画像の特徴ベクトルを学習データ、テスト画像から生成された画像の特徴ベクトルをテストデータとして、分類器による学習と識別を行う。本研究では、学習、識別に SVM(Support Vector Machine) を用いる。

という流れになる。

つまり、bag-of-keypoints では、画像の特徴量は、画像の局所特徴 (keypoints) から生成した代表ベクトル (visual words) の出現頻度のヒストグラムによって表現される。Bag-of-keypoints の概要図を図 2 に示す。

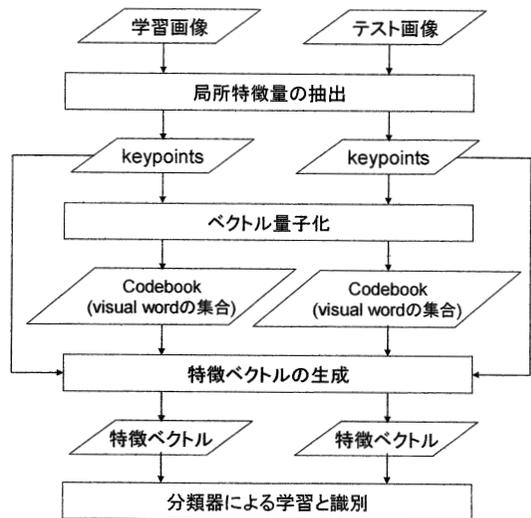


図 2. Bag-of-keypoints 法の概要図

2.2.1 SIFT [3]

SIFT (Scale Invariant Feature Transformation) [3] は、ある特徴点の代表輝度勾配方向を決定し、その方向を基準として、他方向の輝度勾配ヒストグラムを作成し、多次元ベクトルで特徴量を記述する。これにより、回転・スケール変化・照明変化に頑強であるという特徴を持つ。

SIFT 特徴点抽出の流れは、特徴の抽出に適した点の検出と、特徴量の記述の 2 段階からなり、検出段階はスケール空間極値検出、キーポイントのローカライズからなり、記述段階はオリエンテーション割り当てと、SIFT Descriptor により特徴量の抽出からなる。スケール空間極値検出によりスケール変化に、オリエンテーションの割り当てにより回転に、SIFT Descriptor による特徴量の記述により照明変化に、それぞれ不変であるという特徴を得る。

2.2.1.1 Bok 法における SIFT

Bok 法における SIFT 特徴点のサンプリング方法には、SIFT の手法である Difference of Gaussian を用いて特徴点を検出する sparse sampling や、ランダムや画像の格子点上を特徴点とする dense sampling などが主に挙げられる。今回は、画像の縦横 10pixel 毎の格子点を特徴点とする dense sampling 法を用いる。このため、本研究では、基本的に SIFT の特徴量記述段階である SIFT Descriptor のみを利用することになる。

2.2.2 SVM(Support Vector Machine)

SVM はニューロンのモデルとして最も単純な線形しきい素子を用いて、2 クラスの識別器を構成する手法である。1960 年代に Vapnik らが考案した Optimal Separating Hyperplane を起源とし、カーネル学習法と組み合わせると非線形の識別器になる。この拡張はカーネルトリックと呼ばれる手法で、このカーネルトリックの功績で、現在知られている多くの手法の中でも認識性能の優れた学習モデルの一つであると考えられている。

SVM は基本的には 2 クラス識別にしか

対応していなく、複数クラスで識別する場合には、複数の SVM を組み合わせるなどの工夫が必要となる。SVM を内部的に組み合わせ、複数クラス分類を行ってくれるプログラムとして、LIBSVM や、SVMlight の複数クラス分類版である SVMmulticlass などがあり、本研究では、LIBSVM を用いた。

2.3. seam-carving [1]

Seam-carving[1]とは、周囲と混ざって目立たないピクセルの集合となる seam の除去、またはその seam の隣への新たな seam の追加を、目標となる解像度になるまで繰り返すことで、画像内容を考慮したリサイズを行う技術である。このため、全てのピクセルについて各々が持つエネルギー値を定義し、それらのエネルギーの総和が最小になるような seam を決定する問題となる。画像 I の各ピクセルが持つエネルギー値を定義する簡単なエネルギー関数が (1)式となる。

$$e_1(I) = \left| \frac{\partial}{\partial x} I \right| + \left| \frac{\partial}{\partial y} I \right| \quad (1)$$

seam とは、垂直 seam の場合は座標 $(x(i), i)$ に、水平 seam の場合は座標 $(j, y(j))$ に位置するピクセルを端から端まで繋いだもので、この seam は、サイズが $n \times m$ の画像 I の垂直 seam を s^x とすると、(2)式の様に定義される。

$$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, s, t, \forall i, |x(i) - x(i-1)| \leq 1 \quad (2)$$

また、水平 seam を s^y とすると、(3)式の様に定義される。

$$s^y = \{s_j^y\}_{j=1}^m = \{(j, y(j))\}_{j=1}^m, s, t, \forall j, |y(j) - y(j-1)| \leq 1 \quad (3)$$

そして、(2),(3)式で定義された seam の中からエネルギーが最小となるものを選ぶ。これを式で表すと(4)式のようになる。

$$s^* = \min_s E(s) = \min_s \sum_{i=1}^n e(I(s_i)) \quad (4)$$

ここで $e(I(s_i))$ とは画像 I 中の seam s_i が持つエネルギー値となる。

図 3 に、画像中の全 seam の中で、エネルギー総和が最小であると決定された seam の例を示す。画像中の赤線が seam になり、縦方向の赤線が垂直 seam、横方向の赤線が水平 seam に相当する。



図 3. エネルギーが最小となる seam の例 ([1]より)

2.3.1 seam-carving の本研究への応用

Seam-carving は(2),(3)式で定義された画像中の全 seam 中から、エネルギーが最小となる seam を除去していくことで縮小を行う技術となる。

この時、エネルギー値に閾値 k を設けて、その閾値 k を越えるまで seam の除去を繰り返すことで、画像内容に関係ないと思われる領域の除去を行う。これを式で表すと以下の(5)式ようになる。

$$s^* = \min_s E(s) = \min_s \sum_{i=1}^n e(I(s_i)) < k \quad (5)$$

今回の実験では $k=1000$ で固定とした。

3. 評価実験

3.1. 実験環境

学習、テスト画像には、この分野における標準的なデータセットである caltech-256 から 20 種類(1 番～20 番)、1 種類につき、学習画像として 50 枚ずつ、テスト画像として 40 枚ずつ利用した。学習画像は合計で 1000 枚、テスト画像は合計で 800 枚となっている。学習、識別方法は Bok 法[2]を用い、実験の概要は以下のようになっている。また、SIFT 特徴量は 10×10 pixel の格子点毎に記述し、k-means によるクラスタ数(visual word の数)は 500 に設定した。実験に用いた画像の一部を次ページの図 6 に示す。

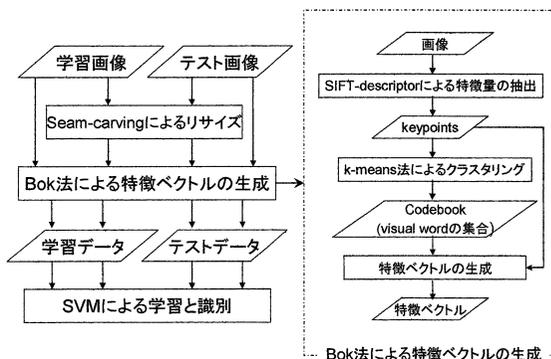


図 4. 実験の概要図

3.2. 実験結果

実験結果を図 5 と表 1 に示す。

図 5 は横軸が caltech-256 におけるカテゴリー番号を表していて、縦軸がそれぞれのカテゴリーに属するテスト画像の識別結果を表している。例えば、横軸が 6 で、縦軸が 05 であれば、カテゴリー番号 6(バスケットゴール)に属するテスト画像の内、50%は正しく識別されたということになる。

表 1 は、全てのテスト画像の識別率の平均を表した表となっている。

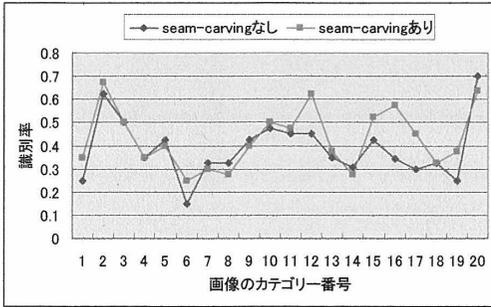


図 5.SVM による識別結果

表 1.SVM による識別結果(認識率の平均)

seam-carving の有無	なし	あり
識別率 (%)	38.75%	43.20%

3.3. 考察

全体としては識別率を 4.5%程向上させることが出来た。また、特定の 카테고리番号 1,6,12,15,16,17,19 などについて、特に顕著な効果が見られた。

カテゴリ番号 1,6,12,15,16,17,19 は、それぞれ AK(銃),バスケットのゴール、双眼鏡、盆栽、ラジカセ、ボーリングの玉、ボクシンググローブの画像群となっていて、比較的、空白や単純な背景の画像が多いカテゴリとなっている。これらのカテゴリについては 10~20%ほど識別率を向上させることが出来、本手法の有効性を確認することが出来た。しかし、6 番の浴槽などのカテゴリに属する画像からは、顕著なエッジがあまり見られなかったことなどから、seam-carving の閾値が上手く機能せず、オブジェクトを大きく歪めてしまっているケースが見られた。また、細かいエッジが画像全体に見られるような画像に対しては、seam-carving によるリサイズはほとんど効果が見られなかった。

4. おわりに

本研究では、一般画像分類の前処理として、seam-carving を利用した学習、識別に不要であると思われる領域の除去を行うことで、画像の学習効率および識別率の向上が出来るのではないかと考え、実験を行った。画像の学習識別は Bok 法を用い、特徴点抽出には SIFT、特徴ベクトル量子化には k-means、識別器には SVM を用いて実験を行った。実験では、この分野における標準的なデータセットである caltech-256 から 20 種類のカテゴリを選択してマルチクラス分類を行った。結果としては、事前に seam-carving による学習、識別に不要であると思われる領域の除去を行っていない場合に比べて、全体として 4.5%程、一部のカテゴリについては 10~20%程識別率を向上させることが出来た。

今後の課題として、今回の実験では、空白や背景などが多いカテゴリに対してしか、効果を上げることが出来なかったということが挙げられる。今後は、画像全体の特徴を考慮した可変閾値などを設けて seam-carving を適用させることで、より多様な画像に対して有効な seam-carving の適用手法を考え、更なる学習、識別効率の向上を目指す。

参 考 文 献

- [1] S. AVIDAN, A. SHAMIR, "Seam carving for content-aware image resizing," SIGGRAPH, 2007.
- [2] G. Csurka, C. Dance, L. Fan, and C. Bray, "Visual categorization with bags of keypoints," ECCV, pp. 1-22, 2004.
- [3] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol.60, No.2, pp.91-110, Jan.2004.

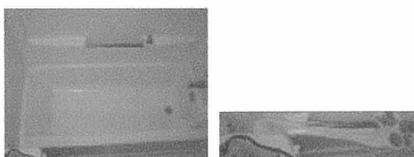
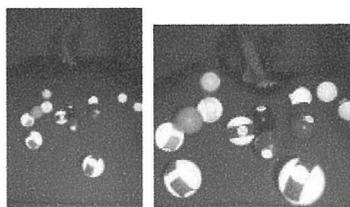
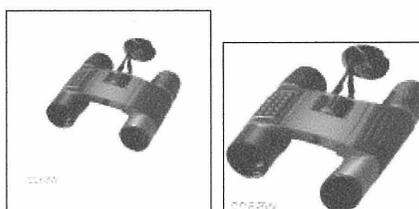
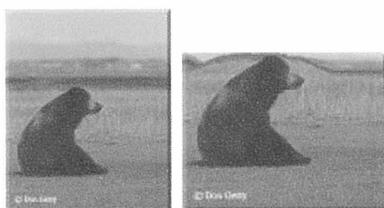
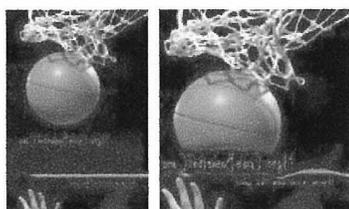
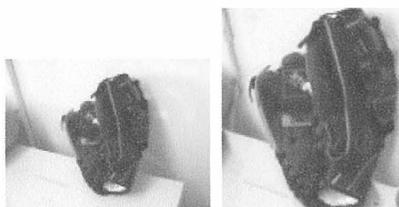
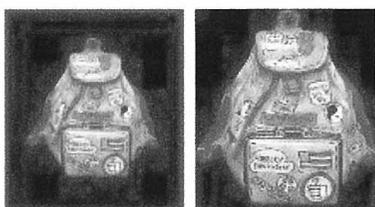
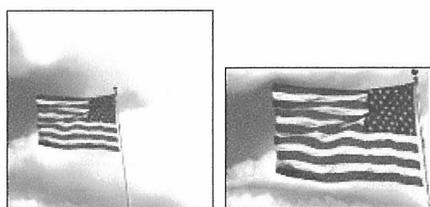


図 6.実験に用いた画像の一部
 (左. 元画像)
 (右. Seam-carving 適用後の画像)