

修士論文概要書

2010年 2月提出

専攻名 (専門分野)	情報理工学	氏名	佐藤 元昭	指導 教員	甲藤 二郎 印
研究指導名	画像情報研究	学籍番号	5108B059-4 ^{CD}	教員	
研究 題目	一般物体認識における画像の学習効率および認識率向上に関する検討				

1 はじめに

近年、デジタルカメラの普及などによる画像データの爆発的な増加に伴い、人手による画像のカテゴリ分類が困難になっており、画像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[1]。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、認識に不要であると考えられる情報を事前に取り除くことで、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

図 1 に、提案手法の概要図を示す。以下、関連技術の説明と、本研究への応用の詳細を述べる。

2.1 関連技術

• Bok (Bag of keypoints) 法

Bag-of-keypoints 法 [1] は、統計的言語処理における Bag-of-words のアナロジーで、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。

一般的な処理の流れとしては、まず SIFT による局所特徴量の抽出を行い、そこで学習画像より抽出された全ての SIFT 特徴量を k-means によってベクトル量子化を行う。続いて、k-means の各クラスタ重心を一つの visual word として割り当て、codebook を構築する。この codebook に基づいて、各画像に一つの特徴ベクトルを生成し、それらを入力として SVM による学習と識別を行う。

• SIFT (Scale Invariant Feature Transformation)

SIFT [2] 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スペース極値検出、キーポイントのローカライズ、記述段階はオリエンテーション割り当てと、SIFT Descriptor による特徴量の記述からなる。

• Seam Carving

Seam Carving [3] 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 seam は、サイズが $n \times m$ の画像 I の垂直 seam を s^x とすると、(1) 式の様に定義される。

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

• Saliency Map

Saliency Map とは、画像や映像を提示した際に、人間の視覚的注意が画像や映像のどの領域に最初に向けられるのかを表現した視覚モデルである。

人間の視覚情報処理の初期段階では、いくつかの単純な視覚的特徴が処理され、複数の特徴マップ(feature maps)として表現される。その後、それらの特徴マップが統合され、顕著性マップ(Saliency Map)として表現されるという特徴統合理論に基づき、Saliency Map はモデル化されている。本研究では、モデルの実装に Itti-Koch らによって提案されたモデル[4]を用いる。

2.2 本研究への応用

• SIFT

Bok 法における SIFT 特徴点のサンプリング方法には、オリジナルの SIFT で使われている Difference of Gaussian を用いて特徴点を検出する sparse sampling や、ランダムもしくは固定間隔の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。本研究では、画像の格子点上を特徴点とする 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。

• Seam Carving

Seam Carving は画像中の全 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を設けて、その閾値を越えるまで 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。ここで、 k は閾値、 $e(I(s_i))$ は画像 I 中の seam s_i が持つエネルギー値を表す。

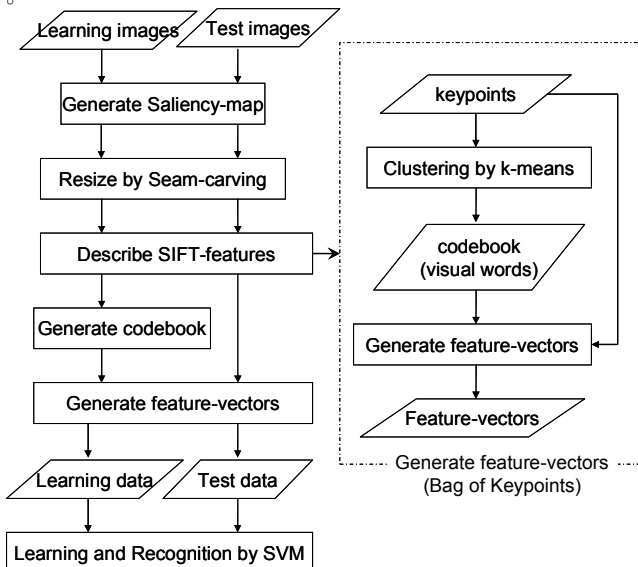


図 1: 提案手法の概要図
Figure 1: Flowchart of proposal

$$s^* = \min_s E(s) = \min_s \sum_{i=1}^n e(I(s_i)) < k \quad (2)$$

• Saliency Map

Seam Carving は、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。ここで、物体領域は背景領域と比較して顕著度が高いと考えられるため、各ピクセルのエネルギー値に Saliency Map の輝度値を割り当てることで、顕著度の高い物体領域以外の、認識に不要であると思われる領域を除去することを目的とする。

同時に、SIFT 特徴量の記述領域を、顕著度の高い領域のみに限定することで、画像の学習効率および認識率を目指す。

3 評価実験

• 3.1 実験概要

実験用画像データベースには、caltech-101 を使い、ここから 101 種類、1 種類につき、学習画像として 20 枚ずつ、テスト画像として 20 枚ずつ利用した。学習、識別方法は Bok 法[1]を用い、k-means によるクラス数 500 に設定した。

また、評価実験は以下の 4 種類行った。

- ① Bok 法単独によるもの。
- ② 画像に事前に Seam Carving 処理を施したもの[5]。
- ③ Seam Carving のエネルギー関数として Saliency Map を使い、また、SIFT 特徴量の記述領域を顕著度の高い領域のみに限定したもの。
- ④ Saliency Map を生成し、SIFT 特徴量の記述領域を顕著度の高い領域のみに限定したもの。

①、②の手法で、画像に対し SIFT 特徴量の記述を行った例を図 2、図 3 に示す。



図 2: ①の手法で SIFT 特徴量を記述した画像例
Figure 2: A result of method ① applied to an image.



図 3: ①、②の手法で SIFT 特徴量を記述した画像例
Figure 3: A result of method ①, ② applied to an image.

• 3.2 実験結果と考察

実験結果を図 3 に示す。従来手法と比較して、どの提案手法も良い識別結果を示した。また、Saliency Map と Seam Carving の両方を用いた手法が、全体的に最も高い識別率を示した。Saliency Map と Seam Carving を単体で用いた場合と比較すると、分類対象カテゴリ数が少ない内は Seam Carving の方が良い識別率を示し、101 種類のカテゴリ分類となると、ほぼ同様の識別率を示した。Saliency Map と Seam Carving の両方を用いた手法と従来手法と比較すると、特にカテゴリ数が増えた場合についての違いがより顕著で、10 種類のカテゴリ分類の場合には、どの手法も大差無いが、101 種類のカテゴリ分類となると、提案手法の方が従来手法に比べ、4%程識別率が向上する結果となり、本手法の有効性を確認できた。

識別率が向上した理由としては、画像の格子点上に SIFT 特徴量を記述する方式の場合、背景領域などの認識対象である物体領域以外の特徴量の記述も多く行っているのに対し、物体領域により重点的に特徴量を記述することが出来たためであると考えられる。認識対象カテゴリ数が少ない場合は、背景領域も物体認識のための重要な情報となるが、カテゴリ数が増加してくると背景情報の重要度が低下してくるため、そこで顕著度の高い領域に限定して SIFT 特徴量を記述したことで、物体領域の情報をより多く学習することが出来、画像の学習効率、および識別率が向上したと考えられる。

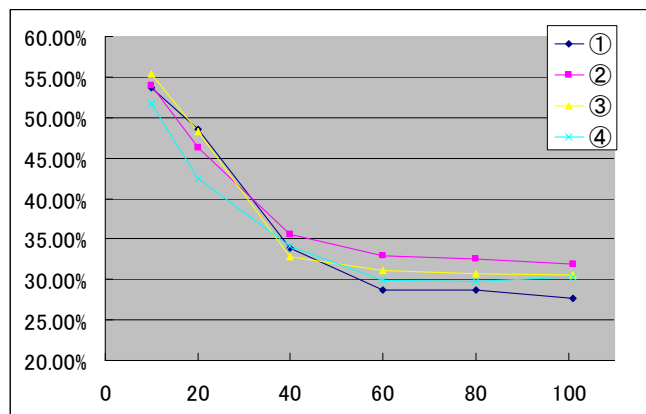


図 3: SVM による識別結果
Figure 3: Recognition result by SVM.

4 むすび

本稿では、Seam Carving と Saliency Map を用いて物体認識に不要であると思われる領域を除去することで Bok 法の特性改善を試み、実験結果としてその有効性を示した。今後の課題として、現在は Saliency Map、Seam Carving、Bok 法とそれぞれ独立して機能しているが、Saliency Map と SIFT のどちらも多重解像度表現を用いていることに注目した多重解像度解析方法の統合についても検討していきたい。

文献

- [1] G. Csurka, C. Dance, L. Fan, and C. Bray, "Visual categorization with bags of keypoints," ECCV, pp. 1-22, 2004.
- [2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol.60, No.2, pp.91-110, Jan.2004.
- [3] S. AVIDAN, A. SHAMIR, "Seam carving for content-aware image resizing," SIGGRAPH, 2007.
- [4] Laurent Itti, Christof Koch, and Ernst Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, No. 11, November 1998.
- [5] 佐藤元昭, 甲藤二郎, "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," 情報処理学会 AVM 研究会, AVM-64-7, Mar.2009.