

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 1 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26540079

研究課題名(和文)大規模データを用いたBag of Visual Wordsからの画像再構成

研究課題名(英文)Image Reconstruction from Bag of Visual Words using Large-scale Image Data

研究代表者

原田 達也 (Harada, Tatsuya)

東京大学・情報理工学(系)研究科・教授

研究者番号：60345113

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究の目的は、Bag-of-Visual-Words (BoVW)画像特徴から、元の画像を再構成することである。BoVWでは、局所記述子が量子化されており、かつ空間情報が失われているため元画像の再構成は困難である。本研究では、局所記述子の最適な配置を推定するため、この隣接の自然さと、大域的な位置の選好性を取り入れた評価関数を提案し、大規模画像データを用いてこれらを獲得する手法を提案した。また、効果的な最適化手法を提案した。提案手法の有用性を示すため、約100カテゴリから収集した画像に対してBoVWからの元画像の再構成を行った。

研究成果の概要(英文)：The objective of this study is to reconstruct images from Bag-of-Visual-Words (BoVW). BoVW is defined here as a histogram of quantized local descriptors extracted densely on a regular grid at a single scale. This task is challenging for two reasons: 1) BoVW includes quantization errors. 2) BoVW lacks spatial information of local descriptors. To tackle this difficult task, we use a large-scale image database to estimate the spatial arrangement of local descriptors. Then this task creates a jigsaw puzzle problem with adjacency and global location costs of visual words. Solving this optimization problem is also challenging because it is known as an NP-Hard problem. We propose a heuristic but efficient method to optimize it. To underscore the effectiveness of our method, we apply it to BoVWs extracted from about 100 different categories and demonstrate that it can reconstruct the original images.

研究分野：知能機械情報学

キーワード：コンピュータビジョン 機械学習

1. 研究開始当初の背景

画像特徴量からの画像の再構成については、いくつかの先行研究が存在する。例えば、Weinzaepfel らは SIFT 記述子とキーポイントの座標、向き、大きさといった幾何情報から画像の再構成を行った。D'Angelo らは BREF や FREAK とその座標情報から画像再構成を実現している。Vondrick らは HOG 特徴からの画像の再構成を行った。Oliva らは GIST 特徴の可視化を試みている。HOG、や GIST は局所記述子として用いられることもあるため、上記の研究は全て局所記述子から画像を再構成する研究であると言える。

しかし、一般に画像検索や画像認識などのタスクでは局所記述子をそのまま扱うことは少なく、コーディングとプーリングと呼ばれる処理過程を経て最終的な画像特徴量とすることが多い。Bag-of-Visual-Words (BoVW) は、最も基本的なコーディングとプーリングの手法であり、画像検索や画像認識への有用性が示され、画像特徴量のデファクトスタンダードの一つとなっている。BoVW は局所記述子をベクトル量子化した後に画像内に存在する量子化後の局所特徴の平均を計算することで得られる。この前半の処理をコーディングといい、後半の処理を(平均値)プーリングと呼ぶ。しかしながら、コーディングとプーリング後の画像特徴量、つまり空間情報が欠落した画像特徴量から画像を再構成する手法は未だ提案されていない。

2. 研究の目的

そこで、本研究の目的は、コーディングとプーリング後の画像特徴から元画像が復元可能なアルゴリズムを開発し、その復元精度を適切に評価する手法を提案することである。本研究の特色として、空間情報の欠落した画像特徴から元画像を再構成する手法を世界に先駆けて開発する点であり、これにより、例えば画像特徴同士の重みづけ和による新たな画像生成といった自動コンテンツ生成への重要な第一歩となる。

具体的には、画像特徴量として Bag of Visual Words (BoVW) から画像を再構成することを目標とする。BoVW から画像を再構成する際に大きな困難となるのは次の2点である。

- (1) BoVW を構成する局所記述子の画像中での位置情報が失われている点。
- (2) 局所記述子がコードワードに割り当てられるときに量子化誤差が生じる点。

本研究では、これら2つの問題を解決する新規なアルゴリズムを開発し、BoVW からの画像の再構成といった困難な問題に挑戦する。BoVW から画像を再構成するアルゴリズムの概要を図1に示す。

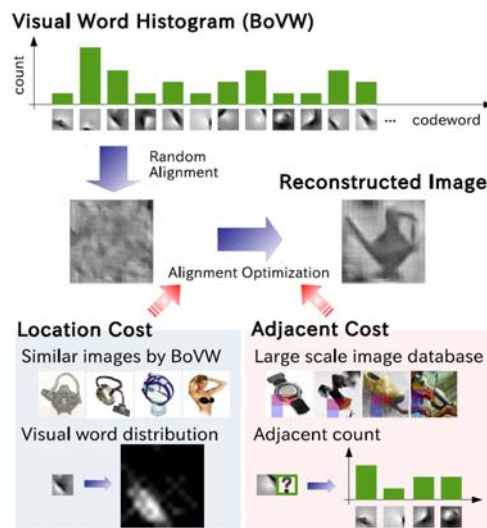


図1 : BoVW からの画像再構成の概要

3. 研究の方法

ここでは BoVW から画像を再構成する具体的な手法を説明する。これは以下の4ステップから構成される。

- (1) BoVW から画像を再構成する問題を局所記述子の最適空間配置問題として定式化
- (2) 大規模画像データベースを援用したパラメータ推定
- (3) 上記問題の評価関数の二次配置問題と最適化
- (4) 大規模な画像の BoVW 特徴からの画像再構成評価実験

それぞれのステップを以下に説明する。

- (1) BoVW から画像を再構成する問題を局所記述子の最適空間配置問題として定式化

局所記述子の幾何情報の復元を行うことができれば、BoVW からの画像の再構成は局所記述子とその幾何情報から画像を再構成する問題に帰着され、様々な既存手法を活用することができる。BoVW における局所記述子の抽出には特徴的な点を検出する方法と等間隔に密にサンプリングする方法があるが、物体認識や画像検索において密なサンプリングの優位性が報告されているため、本研究では密なサンプリングを用いる。

密なサンプリングでは、局所記述子はグリッド状に抽出され、それらはすべて同じ向き、大きさを持つ。その場合、復元すべき局所記述子の幾何情報はグリッド上の位置のみである。ここで、グリッドの間隔、画像サイズ、局所記述子の抽出サイズは既知であるとする。

局所記述子をグリッド上に配置する問題について考える。局所記述子を配置するにあたって、考えられる方法が2つある。ひとつ

は、局所記述子の隣り合いの適合度を考えて配置する方法である。これは、ジグソーパズルに例えるならば、隣り合うピースのエッジや色が一致するようにピースを並べることに相当する。もうひとつは、局所記述子が画像中のどのあたりに配置されやすいかを大まかに考え、概形を作ることである。これは、ジグソーパズルに例えるならば、空のピースは上方に、地面のピースは下方に配置することに相当する。そこで局所記述子の隣接パターンコストと、局所記述子の大域的な位置コストの双方を考慮した評価関数を考える。

この評価関数を最小化するグリッド上のコードワード配置を求めることにより、局所記述子の幾何情報が復元される。局所記述子の空間的な配置が決定されれば、任意の局所記述子画像化手法を用いることにより、画像全体を復元することができる。

(2) 大規模画像データベースを援用したパラメータ推定

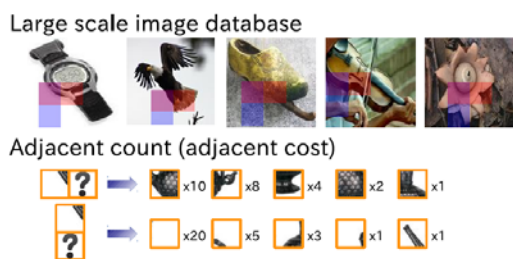


図2：隣接パターンコスト

隣接コストは、コードワード同士が隣り合う場合のコストを意味する。これは、再構成される画像のエッジ等の一貫性を規定する。コードワード間の隣接コストは、コードワード同士の隣り合いやすさより定義する。ここでコードワード同士の隣り合いやすさの定義が問題となるが、本提案では、ImageNet等の大規模画像データベースを活用し、隣り合いやすさを統計的にモデリングする(図2参照)。

Similar images retrieved by BoVW



Visual word distribution (location cost)



図3：大域位置コスト

大域的な位置コストは、あるコードワードが画像中の特定の位置に置かれるコストを意味する。これは、再構成される画像の大域的な画像らしさを規定する。具体的にはコードワードが画像中のある位置に置かれるコストは、その位置への置かれやすさの負の対数を取ることで定義する。ここで、置かれやすさの定義が問題となるが、本提案では、

BoVW が類似画像検索の性能に優れることを活用し、入力された BoVW を用いて類似する画像を大規模画像データベースから近傍探索し、その類似画像中でコードワードがそれぞれの位置に置かれる回数を調べることで置かれやすさと考える(図3参照)。

(3) 評価関数の二次配置問題と最適化

評価関数は、一連のジグソーパズルを解く研究に用いられる評価関数の一般化にあたる。ジグソーパズルにおいては各ピースの上下左右の隣接のみが考慮されるが、本研究では上下左右だけでなく斜めや2つ以上離れている場合についてもコストを考慮する点異なる。

評価関数は適切な変形により二次配置問題を一般化した問題に帰着される。二次配置問題はメタヒューリスティクスを組み合わせた解法が主流であり大規模な実験調査により遺伝的アルゴリズムとタブー探索を組み合わせたアプローチの有効性が示されている。遺伝的アルゴリズムとタブー探索を組み合わせると計算量が膨大なものとなるため、本提案では遺伝的アルゴリズムと山登り法を組み合わせた手法により最適化問題を解くことにする。

(4) 大規模な画像の BoVW 特徴からの画像再構成実験

上記(1)~(3)の手順によって、BoVW からの画像再構成のアルゴリズムのプロトタイプが完成する。この有効性を検証するために大規模データの BoVW から画像再構成実験を試みる。生成された画像の質の評価は大変困難な課題であり、クラウドソーシングを利用した不特定多数の人の定性的な評価を活用した画像評価を行う。

4. 研究成果

提案手法の定性的、定量的な評価のため、以下の3手法を用いた結果を示す。

- (1) 提案手法による BoVW から画像再構成
- (2) Invert HOG を BoVW に適用
- (3) BoVW による大規模画像データベースからの類似画像検索

実験に利用した画像を図4に、それぞれの手法を用いて再構成された画像の例を図5に示す。提案手法により再構成された画像は、前項で得られた画像とよく類似しており、局所記述子の幾何情報が精度よく推定されていることがわかる。Invert HOG を用いたものは極度に不鮮明であり、もはや何の画像であるか読み取ることはできない。大規模データベースを用いた最類似画像は、クエリ画像と内容が大きく異なるものが多く、元画像の

把握は困難である。提案手法を用いて得られた画像も元画像に比べると細かいテクスチャが失われている。しかしながら、原因は局所記述子の幾何情報の推定ではなく、局所記述子群から画像を構成する部分にある。鮮明さの向上は、コーディングされた特徴量からの画像の再構成ではなく、局所記述子から画像を再構成する研究の発展に依存すると考えられる。



図4：再構成に用いた画像の例

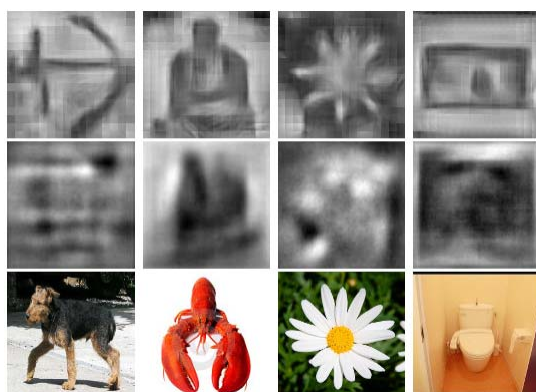


図5：BoVW から再構成された画像の例。上段：提案手法による結果。中段：Invert HOG の BoVW への適用。下段：大規模画像データベースからの最近傍検索。

提案手法では、画像から得られた特徴量から画像を再構成するだけでなく、人工的に得られた特徴量から画像を構成することも可能である。このことは、将来的には特徴量空間での操作を利用した画像生成のアプリケーションへと発展し得る。ここでは、2つの画像特徴を特徴空間内で補間することにより、画像のモーフィングを試みる実験を行った。図6に結果を示す。これは、特徴空間上での特徴量の操作による画像生成の可能性を示すものである。

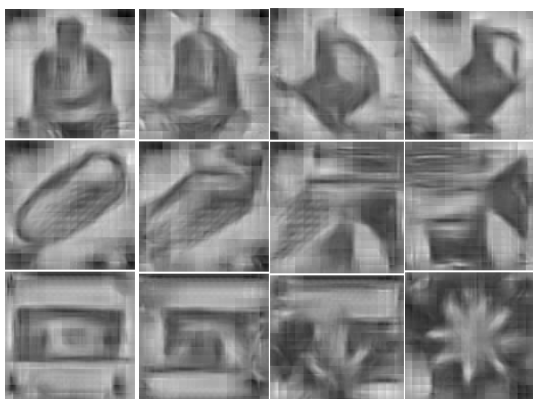


図6：BoVW からの画像生成によるモーフィング。

5. 主な発表論文等

〔雑誌論文〕(計1件)

- ① Hiroharu Kato and Tatsuya Harada. Image Reconstruction from Bag-of-Visual-Words. The Twenty-Seventh IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2014), pp.955-962, 2014.

〔学会発表〕(計1件)

- ① Hiroharu Kato, Tatsuya Harada. Image Reconstruction from Bag-of-Visual-Words. 画像の認識・理解シンポジウム (MIRU2014), 7/30/2014.

〔図書〕(計0件)

〔産業財産権〕

○出願状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

<http://www.mi.t.u-tokyo.ac.jp/#researchdetail-generation>

6. 研究組織

(1) 研究代表者

原田 達也 (HARADA, Tatsuya)
東京大学・大学院情報理工学系研究科・教授
研究者番号：60345113

()

研究者番号：

(3) 連携研究者

()

研究者番号：