

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 5 日現在

機関番号：62615

研究種目：挑戦的萌芽研究

研究期間：2013～2014

課題番号：25540076

研究課題名(和文)画像検索の良さの推定

研究課題名(英文) Estimation of Goodness of Image Retrieval

研究代表者

佐藤 真一 (SATO, Shin'ichi)

国立情報学研究所・コンテンツ科学研究系・教授

研究者番号：90249938

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)：本研究では、画像検索において、ある問い合わせが「良い」画像検索結果を出力するかどうかを、実際に検索結果を算出することなく、高速に推定する手法、さらには良い画像検索結果を出力する問い合わせを推定する手法について検討した。実際に画像検索の「良さ」を定義し、実際の画像データセットに基づき高速・スケーラブルな画像レコメンデーションを実装した。100万画像を有するデータベースに対してもスケーラブルに動作し、有効なレコメンデーションが実現できたことを確認している。

研究成果の概要(英文)：We studied the estimation of goodness of image retrieval, namely, a method to judge if a given query returns good image retrieval results, without performing actual retrieval. On top of this, we implemented image recommendation system which can recommend good query with given image database. We tested the system with the database contains 1 million images and confirmed that our method can effectively recommend good query region with scalable processing.

研究分野：マルチメディア検索

キーワード：画像データベース 画像検索 画像レコメンデーション 転置索引 分枝限定法

1. 研究開始当初の背景

画像を問い合わせとした画像検索は、query by example と言われ、盛んに研究されている。1990年代に始まった画像の内容検索の研究では、主として色ヒストグラムに基づく類似度が用いられており、色合いの似た画像の検索くらいしかできなかった(夕陽の画像の検索程度)。しかし、最近では、局所特徴量を用いた Bag of Features 等の表現に基づく画像検索手法が盛んに検討されており、特定物体や同一種別・シーンの検索等が精度よく実現可能となってきた。これは Google Image の類似画像検索の完成度の高さ等にも表れている。最近の研究の主流は、こうした手法に基づいた、大規模画像データに対するスケーラビリティや高速性の研究、データベース中の画像中の部分領域の効率よい探索等となっている。

本提案では、一つ、従来の研究に決定的に欠けている視点を指摘する。いずれの研究も、問い合わせは与えられることが前提であり、与えられた問い合わせ画像に対してより良い検索結果を生成すること、あるいはより高速に検索を実行すること等が主眼であった。しかし、ある問い合わせ画像をシステムに与えた際、検索結果が芳しくないケースが少なからずある。検索したいような画像がそもそもデータベースに存在しない場合、あるいは確かに類似した画像は存在するが、大量にありすぎて意味をなさない場合(森林のテクスチャ等)が相当する。適切な問い合わせ画像を「見つける」問題は、それほど単純な問題ではないと考えられる。Google 検索エンジンでは、問い合わせ文を途中まで入力すると、適切に補完してくれるレコメンドの機能が存在するが、画像を問い合わせとした画像検索では同等の機能は存在しない。本研究では、画像検索において、あるクエリが「良い」画像検索結果を出力するかどうかを、実際に検索結果を算出することなく、高速に推定する手法について検討する。これにより、画像検索におけるレコメンド等も可能となる。

2. 研究の目的

本研究では、画像検索の「良さ」の定義、その推定手法の実現、並びにそれが実際に期待するような応用につながる可能性があることについて明らかにする。しかる後に、実用規模の高速アルゴリズムの実現、並びに実際のリアルタイム画像検索システム等の実用の実現について、検討を進める。画像検索の「良さ」の推定という切り口は、国内外に例がなく、きわめて独創的である。これが実現できれば、画像検索におけるレコメンドーションをはじめとする新たな応用が期待できる。また、従来の画像検索研究やその高速化研究とは検討の方向が異なり、全く新たな学問分野の創出にもつながる可能性がある。

3. 研究の方法

本研究では、適当な画像データベースを対象として、

- 1) 画像検索の「良さ」の定義とその振る舞いの検証、
- 2) 画像検索の良さの高速な計算手法の検討と実装、
- 3) 良い問い合わせ画像の高速な探索手法の検討と実装、
- 4) 応用技術の実現(問い合わせ画像のレコメンドーション)を行う。

なお、画像から得る特徴量としては、基本的には局所特徴量に基づく Bag of Visual Word (BoVW) 表現を使うことを想定している。

4. 研究成果

研究成果の概要

定義した画像検索の「良さ」に基づき、実際の画像データセットに基づき高速・スケーラブルな画像レコメンドーションを実装し、研究発表を行った。データセットとしては、既存のデータセットである Oxford Building 105k(10万画像)と MQA-1M(100万画像)を用い、複数回出現する物体を良い問い合わせとすることで本研究に用いた。分枝限定法と転置索引を適切に組み合わせた高速でスケーラブルな技術を実現し、上記のデータセットを用いて100万画像を有するデータベースに対してもスケーラブルに動作することを確認した。

研究成果の詳細

図1に構築した画像レコメンドーションシステムの概要を示している。本システムでは、相応規模の画像データベースが与えられることを想定している。その上で、入力画像(図中左端)に対し、その画像中の任意の矩形領域のうち、最も「良い」データベース検索結果を返す領域を推薦する。最も「良い」検索結果の定義だが、種々検討した結果、入力画像中の任意の矩形領域と、データベース中のすべての画像中の任意の矩形領域の間で、領域内に対応する BoVW ヒストグラムの間で Normalized Histogram Intersection(NHI) を最大化するような場合、最も良い検索結果を返すとした。また、上記で、任意の矩形領域を用いると述べているが、実際の実装では、探索する矩形の数をある程度制限するため、各画像から Object Proposal(Vandelande, ICCV 2011)を算出し、その外接矩形を用いた。それでも、各画像から数千程度の候補領域を検討することになる。

入力画像とデータベース中の画像について、すべての候補矩形領域間で NHI を求めるのは、組み合わせ数が多すぎて実用的な時間では計算ができない。そのため、すべての矩形領域については包含関係に基づく階層化を行い、その階層に基づく分枝限定法により効率の良い検索を実現している。また、NHI

の算出のため、階層化した各ノードに対応する BoVW ヒストグラムを転置索引により構造化しておき、高速計算を可能としている。より具体的に述べると、問い合わせ画像については、その都度 Object Proposal を算出し、領域間の包含関係に基づき階層構造を算出する。これは、入力画像一枚だけの処理なので、高速に処理可能である。

データベース中の画像については、まずは画像レベルの階層化を行い、ついて画像内の領域の階層化を行っている。画像レベルについては、画像から得た BoVW ヒストグラムに基づくクラスタリングにより実現している。次いで、各画像については、Object Proposal を算出し、包含関係に基づいて階層化を実現している。最終的に、画像データベース(内包する画像すべてを含む)全体の階層を構築し、各ノードに対応する BoVW を算出しておいて転置索引に登録しておく。この処理には時間がかかるが、事前に処理しておくことができる。

これらに基づき、入力画像が与えられたときに、効率よく領域を提案することができる。実際、100万画像を有するデータベースによる実験でも、実用的な処理速度を実現しており、単純に転置索引のみを使った方法に比べて10-30倍の高速化を実現している。図2に本手法が生成した提案画像領域の例を示す。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計2件)

1. Nozha Boujemaa, Alexander G. Hauptmann, and Shin'ichi Satoh, "The Future of Multimedia Analysis and Mining: Visions from the Shonan Meeting," Media Impact Section, IEEE Multimedia, Vol.20, No.2, pp.4-12, doi: 10.1109/MMUL.2013.30, April-June, 2013. (査読なし)
2. 佐藤真一, "大規模データに基づく画像・映像意味解析," 画像ラボ, Vol.24, No.10, pp.12-17, 2013. (査読なし)

[学会発表](計3件)

1. Thanh Duc Ngo, Sang Phan, Duy-Dinh Le, and Shin'ichi Satoh, "Recommend-Me: Recommending Query Regions for Image Search," ACM Symposium on Applied Computing (SAC), Information Access and Retrieval (IAR) Track, Full Paper (Oral), pp. 913-918, Gyeongju (Korea), 2014.3.24-28. (査読あり)
2. [特別講演]佐藤真一, "社会分析ツールとしての大規模放送映像解析,"

"Large-Scale Broadcast Video Analysis as Social Analysis Tools," 電子情報通信学会技術研究報告, PRMU, 幕張メッセ(千葉県・千葉市), 10月3-4日, 2013. (査読なし)

3. [招待講演]佐藤真一, 画像認識技術に基づく画像・映像意味解析, 日本音響学会2014年秋季研究発表会, 北海学園大学(北海道・札幌市), 9月3-5日, 2014. (査読なし)

[図書](計0件)

[産業財産権]
なし

[その他]
なし

6. 研究組織

(1)研究代表者

佐藤 真一 (SATOH Shin'ichi)

国立情報学研究所・コンテンツ科学研究系・教授

研究者番号: 90249938

(2)研究分担者
なし

(3)連携研究者
なし

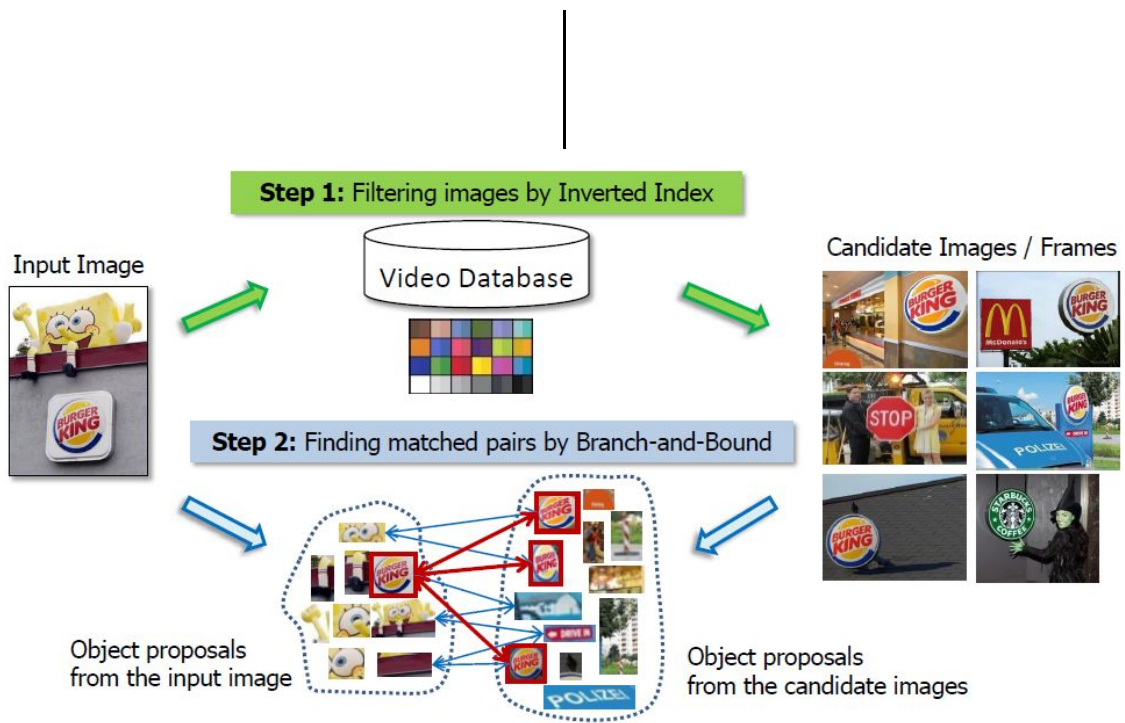


図1 画像レコメンデーションシステムの概要

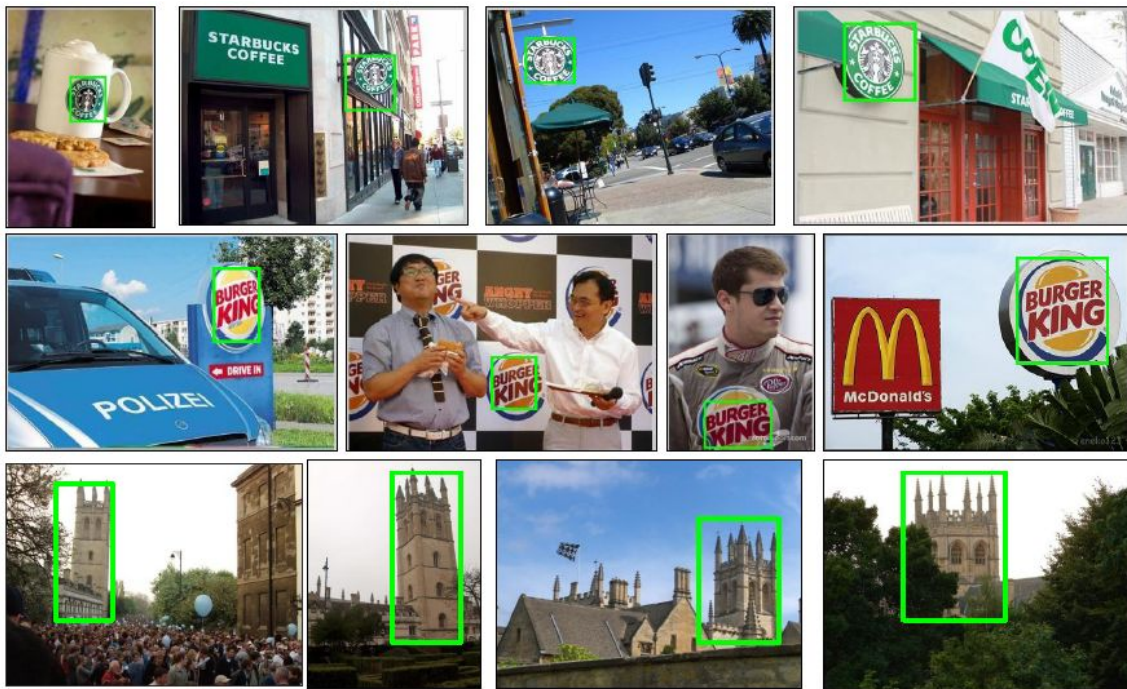


図2 画像レコメンデーション結果の例