

モバイル端末を用いた画像検索のための特徴点選択手法に関する一検討 A Method for Feature Selection for Image Retrieval on Mobile Devices

松崎 康平†
Kohei Matsuzaki

酒澤 茂之†
Shigeyuki Sakazawa

1. はじめに

SIFT(Scale Invariant Feature Transform)[1]等の局所特徴を用いた画像検索手法が高精度な特定物体認識を実現するとして注目されている[2]. また、カメラを内蔵したモバイル端末の普及と、処理能力の向上に伴い、モバイル端末を用いた画像認識サービスも可能となっている[3]. SIFTのような局所特徴抽出処理はモバイル端末においては計算コストが高いため、モバイル端末を用いた画像認識は外部サーバと連携する方式が一般的であった. しかし、近年では、BRIEF(Binary Robust Independent Elementary Features)[4]やORB(Oriented FAST and Rotated Brief)[5]といった高速かつ省メモリな局所特徴抽出処理を実現するバイナリ特徴量が提案されており、モバイル端末上でのローカルな画像検索が可能になっている. このようなモバイル端末上での画像検索は通信環境に依存しない高速な検索を可能にする一方で、検索データベースを端末内に保持する必要がある.

モバイル端末のような計算資源の限られた状況では検索時の処理時間とメモリ消費量を抑えることが重要となる. 検索対象となるデータベースへ登録するオブジェクト数が増加するにつれ、検索に要する局所特徴量の数が膨大となるため、大量のメモリが必要となり、また、処理時間も増加する.

そこで本研究では、モバイル端末を用いた画像検索の観点で、寄与度の大きい局所特徴点を選択することにより、データベースに登録する局所特徴点の数を削減する手法を提案する. 本手法により、検索精度を保ったままデータベースサイズを削減することが可能になり、メモリ消費量の削減、検索速度向上を実現する.

2. 関連技術

局所特徴を用いた画像検索では、一般的に Bag-of-Visual Words(BoVW)と呼ばれる手法によりデータベースに登録されたリファレンス画像とクエリ画像の類似度を計算する. BoVWでは、事前に Visual Words(VWs)と呼ばれるベクトル集合を用意する. そして、画像から局所特徴ベクトルを抽出し、最も類似した Visual Word(VW)に量子化する. この量子化によって画像は VWs のヒストグラムで表現され、画像間の類似度はデータベース内の画像とクエリ画像を表現するヒストグラムの内積計算によって算出される.

上記のように、ヒストグラムの内積表現によって類似度が定義されるため、クエリ画像に存在しない VW は画像間の類似度に影響を及ぼさない.

また、局所特徴は照明変化、スケール変化、回転に対して頑健な特徴量とされているが、射影変化に対応できないとされており[6]、またバイナリ特徴は SIFT と比べた場合

にスケール不変性に乏しいと言われている[7].

モバイル端末を用いた画像検索においては、一般的にクエリ画像はユーザがオブジェクトをカメラで撮影したものとなるため、リファレンス画像に対するスケール変化、射影変化のような見えの変化が生じる. 前述の通り BoVWによる画像検索ではクエリ画像に存在する VW しか画像間の類似度に影響を及ぼさないため、目的となるオブジェクトから検出された特徴点を量子化した VW をデータベースに登録しても、クエリ画像に存在する VW と一致しなければ検索に寄与しないという課題がある.

3. 提案手法

前述の課題に対し、本稿では、リファレンス画像から抽出される局所特徴の中から、見えの変化に対して頑健と推定されるものを選別し、それらのみをデータベースに登録する手法を提案する. BoVW においては、クエリ画像に写るオブジェクトの領域から抽出された局所特徴の内、正解画像と同一座標系における同一の座標から検出された点であり、尚且つ同一の VW に量子化される局所特徴しか画像間の類似度算出に寄与しないといえる. 従って初めにリファレンス画像から局所特徴を抽出し、量子化を行った後、見えの変化を模擬した人工画像においても同様の処理を施すことによって、上記の条件を満たす確率の高い特徴点を識別する. 具体的には、以下の処理によって局所特徴毎のオブジェクトの見えの変化に対する頑健性を推定する.

初めに、登録したい画像から局所特徴を抽出し、特徴点毎に ID を付与する. 以後、これを初期局所特徴集合と呼ぶものとする. 予め用意された VWs を用いて初期局所特徴集合を量子化し、それぞれの ID と紐づける. 次に、登録したい画像を様々な仮想視点から撮影した様に変換した人工画像群を作成する. ここで、画像の変換には 3 次元座標系に仮想視点を設置し、ピンホールカメラモデルを用いて算出した射影変換行列集合 H_1, \dots, H_n を用いる. 仮想視点は、モバイル端末を持ったユーザがオブジェクトを撮影する際のカメラの視点を模擬することを目的とするため、例えば登録したい画像を原点に置き、空間的に均等に配置されるとされる Geodesic Dome の各頂点座標や 3 次元格子状構造の各頂点座標を用いるものとする. その後、作成された人工画像群のそれぞれから局所特徴を抽出し、量子化を行う. それぞれの人工画像から抽出された局所特徴集合毎に、その画像を作成する際に用いた射影変換行列の逆行列を用いて、特徴点の座標を射影する. これにより、それぞれの人工画像から抽出された局所特徴集合の座標系と、初期局所特徴集合の座標系を一致させることが出来る. 座標系を一致させた後、全ての人工画像から抽出された局所特徴について、次の 2 つの条件の両方を満たすか否かの判定を行う.

- (1)初期局所特徴と同一座標から検出される
- (2).初期局所特徴と同一の VW に量子化される

† 株式会社 KDDI 研究所 メディア・HTML5 応用グループ

上記の条件を満たす局所特徴が存在した場合、初期局所特徴の ID にスコアを投票する。また、スケール変化に対する頑健性を推定するため、登録した画像を $\sqrt{2}$ 倍ずつ縮小した画像を4層作成し、上述の処理を画像のスケール毎に繰り返して、投票スコアを累計する。

最終的に、初期特徴点の ID に投票されたスコア上位 N 件を選択することにより、オブジェクトの射影変化およびスケール変化に対して頑健な特徴点を選択出来る。

4. 評価実験

本節では、提案手法の有用性を評価するため、データベースに登録する特徴点数を変更した際の Accuracy 及び、Matching Score を調査する。Matching Score は、Geometric Verification を行った際の inlier の数と登録された特徴点数の割合とする。ここでは inlier の条件として、同一の VW にマッチングされ、尚且つ幾何的な拘束条件を満たすものとする。また、ベースライン手法として、特徴点検出時に検出する点数を指定する方式で作成したデータベースを用意し、比較する。

評価用データセットとして、モバイル端末での画像検索を想定した Stanford Mobile Visual Search データセットの CD カテゴリを用いる。上記データセットは、CD ジャケット等複数のカテゴリについてリファレンス画像及びモバイル端末で撮影されたクエリ画像が提供されている。

提案手法における初期局所特徴点数、及びクエリ特徴点数は常に 900 とし、局所特徴抽出には ORB アルゴリズムを用いた。また、提案手法における仮想視点の配置には Geodesic Dome モデル $z>0$ の頂点座標を用いた。

提案手法及びベースライン手法を用いてデータベースを作成し、画像検索を行った際の認識率を図 1 に、認識した画像における Matching Score の平均値を図 2 に示す。図 1 より、データベースに登録する特徴点数を減少させた場合においても、提案手法では Accuracy の低下を抑え、精度を維持できることが分かる。また、図 2 はデータベースに登録された特徴点の内、画像検索に寄与した特徴点の割合を示すものであり、ベースライン手法と比べて提案手法の方がより多くの inlier となる特徴点を残し、良い結果を得られていることが分かる。なお、特徴点数が増加するほど Matching Score が低下する傾向が見られるが、これは登録した特徴点数に対する inlier の割合の低下を意味し、検索精度の低下を示すものではない。

5. まとめ

本稿では、モバイル端末を用いた画像検索において、検索に有用な特徴点を選択することによってデータベース容量を削減する手法を提案した。提案手法は、登録画像から複数の人工画像を作成して特徴抽出を行い、それらの特徴と登録画像の特徴を比較することにより、見えの変化に対して頑健な特徴点を選択する仕組みであり、その有用性を実験により評価した。その結果、提案手法を用いることによって精度を維持したままデータベースに登録する特徴点数を削減できることが確認できた。

今後は、データベースの大規模化への対応と、特徴点数を更に削減した場合の高精度化について検討する予定である。

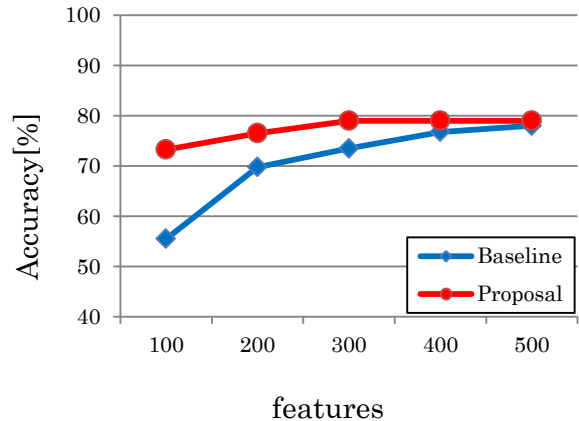


図 1 特徴点数毎の認識率評価

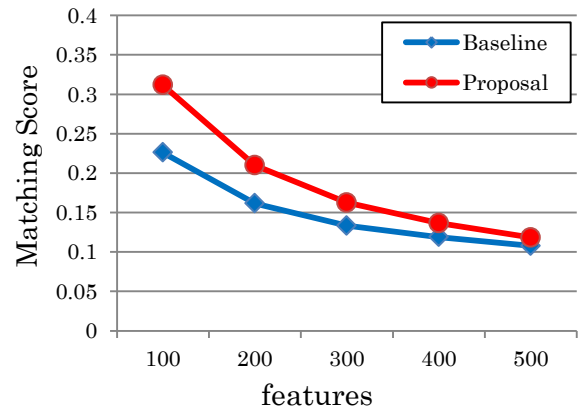


図 2 特徴点数毎のマッチング評価

参考文献

- [1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision Volume 60, Issue 2, pp.91-110, 2004.
- [2] H. Jégou, M. Douze, and C. Schmid, "Improving bag-of-features for large scale image search", In Proc. International Journal of Computer Vision, Volume 87, Issue 3, pp.316-336, 2010.
- [3] <http://www.a9.com/whatwedo/mobile-technology/>
- [4] M. Calonder, V. Lepetit and C. Strecha and P. Fua "BRIEF: Binary Robust Independent Elementary Features", In Proc. European Conference on Computer Vision, 2010
- [5] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: an efficient alternative to SIFT or SURF", In Proc. International Conference on Computer Vision, 2011.
- [6] 藤吉弘亘, 安倍満, "局所勾配特徴抽出技術 sift 以降のアプローチ", 精密工学会誌, vol.77, no.12, pp.1109-1116, 2011
- [7] J. Heinly, E. Dunn, and J.-M. Frahm, "Comparative evaluation of binary features", In Proc. European Conference on Computer Vision, 2012.