

建物画像検索のための特徴点選択手法の一検討 A method for corresponding point selection for similar building image retrieval

上野 智史[†] 橋本 真幸[†] 米山 暁夫[†]
Satoshi UENO Masayuki HASHIMOTO Akio YONEYAMA

1. はじめに

GPS データを取得できる携帯電話を用いて、ユーザをリアルタイムで道案内をするナビゲーションシステムが一般的に普及している。しかしながら高層ビル街など GPS の位置取得精度が不安定で誤差が生じやすいエリアがあり、このようなエリアでも正確な位置情報の取得が求められている。このような環境で精度の高い位置情報を取得する手法として、位置情報が付加されている過去に撮影された街並画像と、現在のユーザ付近の撮影画像とを画像マッチングをすることで位置情報を推定する手法が考えられる。

筆者らはこれまでに、画像内の局所特徴量に基づき画像マッチングを用いて上記方式が実現できる可能性を示した [1]。しかしながら複数のデータベース画像からクエリ画像と同一の正解領域を検索する上記手法において、SIFT など局所特徴量を用いる場合、画像内の特徴点が多いことから対応点探索速度、データベースの特徴量保存容量に関して課題があった。

そこで本論文は、撮影角度の変化に耐性の高い特徴点のみを残しつつ特徴点の数を削減する手法を提案する。本手法により、撮影画像と同エリアの画像を検索する精度を保ったまま特徴点を削除することが可能となり、検索速度の向上、データベース容量の削減の効果がある。実験を通して提案手法の有効性を検証する。

2. 画像マッチングに基づく位置・方向推定

過去に撮影された位置情報付き街並画像データベースとの画像マッチングに基づく位置情報・方向情報推定システムの概要を図 1 に示す。ユーザが撮影した画像に付随する GPS データに基づき対象エリアを限定して、画像特徴量を用いて対象エリア内からユーザ撮影画像と類似する領域を探索する。画像局所特徴量として SURF(Speeded Up Robust Features)[2]を用いる。SURF は SIFT[3]と同様に画像スケール、回転、輝度変化、オクルージョンに耐性のある特徴量で特徴量生成手法と特徴次元数の観点で SIFT より高速な手法である。また画像データベースは Google マップストリートビュー[4]などの位置情報付き街並画像データベースを用いる。

SURF は上記特性により同一オブジェクトを精度よく検出する手法であるが、検索対象画像データが多い場合は画像特徴点が多くなるために次のような問題が発生していた。
(i) 全ての特徴点に対して類似度を算出する時間がかかる。
(ii) 特徴量の蓄積容量が膨大になる。

そこで本論文では画像データベース内の撮影画像を基にして局所特徴点の数を削減する手法を提案する。位置情報付き街並画像データベースは全方位画像で構成されており、任意の視点方向画像を生成できる。そのため視点方向をずらした画像を合成することによりある建物に対して若干撮影角度が異なる画像を作成できる。これらの画像を用いて撮影角度が異なっても対応関係がとれる撮影角度に耐性の高い特徴点を選択し、それ以外の特徴点を削除する。

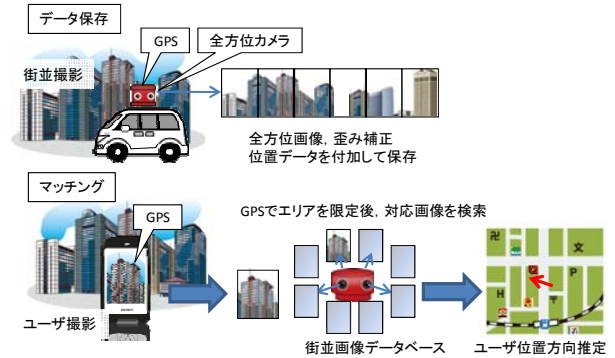


図 1 システムの概要

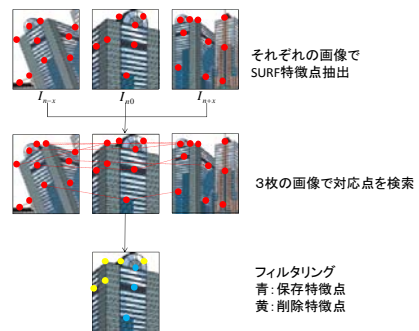


図 2 提案手法の流れ

3. 提案手法

提案手法は、街並画像データベース内で撮影角度差が元で対応点関係が抽出できない特徴点を削除することを基本原理とする。同手法により対応領域検索精度を落とさずに特徴点数を低減でき検索速度の向上が期待できる。

街並画像データベースは全方位画像で構成され、まず撮影画像データと同様の環境を作成するために全方位画像の領域を分割し歪みを補正した画像を生成する。さらに、ある建物画像 I_{n0} に注目した際にその建物画像の中心から x 度移動した画像 $I_{n\pm x}$ を作成する。 I_{n0} と両隣の画像 $I_{n\pm x}$ の局所特徴量間の類似度を算出し、対応点関係を生成する。この対応点探索の結果、対応関係を結べなかった I_{n0} の特徴点をフィルタリングする(図 2)。これら対応関係を結べなかった特徴点は SURF 特徴点の中でも幾何変換に耐性の少ない特徴点であると考えられる。実際の撮影画像はデータベース画像と同じ向きで撮影されていることは少なく、幾何変換に耐性の少ない特徴点は検索時に有効ではないと判断できるため削除する。特徴点をフィルタリングする手法としては以下の 4 手法を利用した。

- 1) 対応関係がない特徴点を削除
両隣の画像 $I_{n\pm x}$ から I_{n0} の各特徴点の対応点を探求して、対応関係がない I_{n0} 内の特徴点を削除する
- 2) 誤対応点削除

両隣の画像 $I_{n\pm x}$ から I_{n0} の各特徴点の対応点を探索して、誤対応であった対応点である I_{n0} 内の特徴点を削除する

- 3) 1)と 2)を併用した手法
1)と 2)の削除点を両方削除する
- 4) 閾値判定により対応が解消された点を削除
3)に加えて、対応点の類似度閾値 T_h に基づき対応点組と判断されなかった特徴点をさらに削除する

上記の誤対応の判定は、各画像間の対応点の相対位置をプロットして、平均から大きく外れる対応点を誤対応とする。類似度閾値 T_h は [3] で利用されているものと同様で、ある特徴点に対して最近傍の対応点と対象画像内の 2 番目に類似度の高い対応点との類似度の比に対する閾値である。画像内の類似する特徴点と比較することにより、窓や壁など似た形状のために局所特徴量が類似する特徴点の誤対応を減少させる。[3]では $T_h = 0.80$ を用いている。 T_h は小さい値ほど画像内に類似の特徴点がある場合に対応点をフィルタする程度が高くなる。

また画像 I_{n0} に対する画像 $I_{n\pm x}$ の設定方法として、 $x = 1$ 度から 15 度の間の値で複数の角度差の画像を利用した。 I_{n0} 内の各特徴点に対して、複数の画像 $I_{n\pm x}$ の特徴点からいずれの条件も満たさなかった特徴点を削除することにした。

4. 実験

4.1 実験環境

ユーザ撮影画像として携帯電話カメラで撮影した都内のある駅周辺の建物画像 137 枚を利用した。またユーザ撮影画像サイズは 640×480 ピクセルである。画像データベースは全方位画像で構成されるが、画像 I_{n0} は撮影位置から仰角上向きに 40 度、下向きに 20 度程度の領域を方位角 30 度ずつ、15 度刻みで分割していき、ひずみを補正した画像を利用した。画像 I_{n0} の両隣の画像 $I_{n\pm x}$ は $x = 1, 3, 5, 10, 15$ 度の値の画像を利用した。

評価基準としてデータベース特徴点削除率と両隣の画像 $I_{n\pm x}$ との対応点を元に RANSAC [5] を用いて射影変換行列が生成できた枚数を用いた。また最終的な評価は、ユーザ撮影画像を用いて特徴点を削減した画像データベース中の画像検索で第一位に表示された正解領域画像の割合を用いた。

4.2 実験結果

実験結果を以下に示す。表 1 に各フィルタリング手法と特徴点類似度閾値 T_h を変化させた場合の特徴点の削減率の関係を、表 2 に両隣の画像の対応点を元に射影変換行列が生成できた枚数を示す。表 1 の特徴点の削減率は SURF をそのまま用いた場合を基準にした削減率を表している。また表 2 の従来手法は SURF を利用した場合を示しており、ランダム手法は特徴点を 10%ランダムに削減したものを繰り返した平均の結果を意味する。類似度閾値 T_h は 0.6-1.0 に変化させて実験を行った。対応ペアは両隣の画像であるため、実験画像の 2 倍の画像が対象である(最大で $137 \times 2 = 274$ 枚)。

表 2 より全ての手法において従来手法より対応領域の射影変換が成立する枚数が同等以上となっていることが分かる。また同様にランダムに特徴点を削減した場合より精度が向上しており本提案手法の有効性が確認できる。特に 4) の手法で $T_h = 0.80$ を用いた時に対応領域を検出できる割合

表 1 特徴点削減率

	類似度閾値 T_h				
	0.6	0.7	0.8	0.9	1.0
手法(1)	2.0	2.0	2.0	2.0	2.0
手法(2)	2.6	3.9	6.3	9.2	11.7
手法(3)	4.7	6.0	8.3	11.2	13.8
手法(4)	6.7	8.7	12.2	16.7	20.8

表 2 提案手法と対応精度の関係

	類似度閾値 T_h				
	0.6	0.7	0.8	0.9	1.0
手法(1)	203	203	203	203	203
手法(2)	210	219	202	217	220
手法(3)	221	213	203	211	220
手法(4)	207	209	224	219	207
従来手法	198	198	198	198	198
ランダム	197	197	197	197	197

が高く、この時の特徴点の削減率は 12.2% となり、データベース画像から特徴点を効果的に削減できることを示した。

さらに上記手法 4) と閾値 $T_h = 0.80$ を用いて、実際のユーザ撮影画像を検索した結果を示す。ユーザ撮影画像に付属する GPS データ半径 50m 以内の画像に対し、提案手法、従来手法(SURF)、ランダムに 10%特徴点を削除した手法と比較する。類似度が一番高い画像が正解画像であった検索精度は、それぞれ 0.50, 0.48, 0.43 であった。提案手法は、ランダムに特徴点を削除することより効果的に特徴点を削除することができており、また特徴点を削除しても従来手法とほぼ同等の検索結果であることを確認した。計算速度は特徴点の削減率と同様に 10%程度改善されていることを確認した。

5. まとめと今後の課題

本論文では、街並画像データベース画像に対して、撮影角度の変化に耐性の高い特徴量のみを残しつつ特徴点の数を削減する手法を提案した。本手法により、データベース画像に対して、撮影角度差のある撮影画像と同一の領域をデータベース画像から検索する精度を保ったまま画像局所特徴点を 10%削減することが可能となり、検索速度の向上、データベース容量の削減の効果があることを確認した。

街並画像データベース画像は解像度が低く特徴点抽出しにくい領域が存在する。今後はこのようなデータベース画像から比較的特徴点を抽出しやすい看板領域を効果的に抽出すること、特徴点抽出しにくい壁領域などで効果的な特徴点を抽出することを検討する。

参考文献

- [1] 上野, 橋本, 米山, “建物画像検索のための対応点検出精度の改善手法に関する一検討,” ITE2009 年次大会予稿集, 2009.
- [2] Herbert Bay, Tinne Tuytelaars and Luc Van Gool "SURF: Speeded Up Robust Features", Proceedings of the 9th European Conference on Computer Vision, Springer LNCS volume 3951, part 1, pp 404--417, 2006.
- [3] D.G.Lowe, “ Distinctive Image Features from Scale Invariant Keypoints, ” International Journal of Computer Vision, Vol.60, no. 2, pp. 91-110, 2004.
- [4] <http://www.google.co.jp/help/maps/streetview/>
- [5] M. A. Fischler, R. C. Bolles. Random Sample Consensus: Paradigm for Model Fitting with Applications to Image analysis and Automated Cartography. Comm. of the ACM, Vol.24, p 381-395, 1981.