

弁別性と持続性の評価に基づく 移動ロボットのためのビジュアルランドマーク選択 Quality Visual Landmark Selection Based on Distinctiveness and Repeatability

下田 真宙 正山崎 公俊 (信州大)

Masamichi SHIMODA and *Kimitoshi YAMAZAKI (Shinshu Univ.), kyamazaki@shinshu-u.ac.jp

This paper describes a method of landmark selection from image streams captured by a camera mounted on a mobile robot. To select stable visual landmarks for mobile robots, two measures to evaluate the landmark “visibility” are considered. After that, under the assumption that robots are able to seek their feasible landmark actively, repeatability of landmarks is evaluated. Weighting techniques using feature position relationships is proposed, and landmark selection criteria using variation coefficient is defined. Proof experiments using a real mobile robot shows the effectiveness of the proposed method.

Key words: visual landmark, mobile robot

1 はじめに

移動ロボットは屋内外で広い行動範囲を持ちうる存在である。例えば目的地への長距離移動をおこなわせる場合、ロボットには走行中の自身の存在位置を適切に知るための機能を搭載する必要がある。従来、カメラやレーザー距離計などによる外界センシングの結果を利用して、自己位置推定を行う手法が提案されている [1, 2]

本研究の目的は、自己位置推定の手掛かりとすべき視覚特徴を移動ロボットに自動で収集させる仕組みの構築である。本稿では、カメラを搭載した移動ロボットが取得した画像群から視認性の高い画像領域を切り出し、ランドマークとして記録していく手法を提案する。ここでの視認性とは、他の景色との区別が容易であるか否か（弁別性）と、オクルージョン等が起きにくい場所にあり頻繁に観測できるか否か（持続性）の両方に対する性質を意味する。

視認性の評価は、所定のコースを複数回走行させた結果を用いておこなう。例えば、ある走行時に見つけたランドマークが弁別性の高いものであったとしても、のちの走行において走行経路がほとんど変わらないのにもかかわらず同じランドマークを捉えられないのであれば、それは視認性が低いとみなす。ランドマークをそのような判断のもとに自動で収集することができれば、少量の環境知識によって適切に自己位置推定を行う移動ロボットの実現が期待できる。

2 ランドマークの表現

2.1 ランドマークの視認性

ランドマークとするべき画像領域を選ぶとき、ランドマークとしての適切さを次のように考える。

- (1) 周囲のものと区別がしやすい
- (2) 動かない

- (3) 周囲のものに隠れない

- (4) 視点により見え方が大きく変化しない

(1) は周囲に類似したランドマークが存在した場合、異なるランドマークを認識する可能性やどちらのランドマークも認識できない可能性があるためである。(2) は、ロボットの走行する経路は毎回同じとは限らないので、ランドマークが動いていると認識できない可能性が高いためである。(3) は周囲のものにランドマークの一部あるいは全体が隠れると認識できなくなるためである。(4) は視点によって見え方が変わり、認識できなくなる可能性があるためである。本研究では(1)を弁別性、(2)、(3)、(4)を持続性と定義し、この2つをあわせて視認性と定義する。

2.2 特徴領域の抽出

ランドマークの一つの単位を、複数の画像特徴点を密に含んだ矩形領域とする。この生成方法を次に説明する。まず、入力画像から特徴点を検出する。そして、ある特徴点の位置から別の異なる特徴点の位置までの距離が一定以下であった場合、2つの特徴点を1つのグループとする。同様にこのグループ内のいずれかの特徴点からの距離が一定以下の特徴点があった場合、3つを1つのグループとする。この処理を続けることで特徴点が密集している領域をひとつのグループとえる。最後に、このグループ内の特徴点に外接する四角形を生成し、特徴領域とする。

特徴領域は一枚の画像から多数抽出される。それらのうち2.4節の処理によりランドマーク候補が選出され、さらにそこから、2.5節の処理によりランドマークが選出される。全体の処理の構成を図1に示す。

2.3 SIFT 特徴量のマッチング

周囲に類似した特徴領域が存在するかを判定する場合や、ランドマークが入力画像内に存在するかを判定する

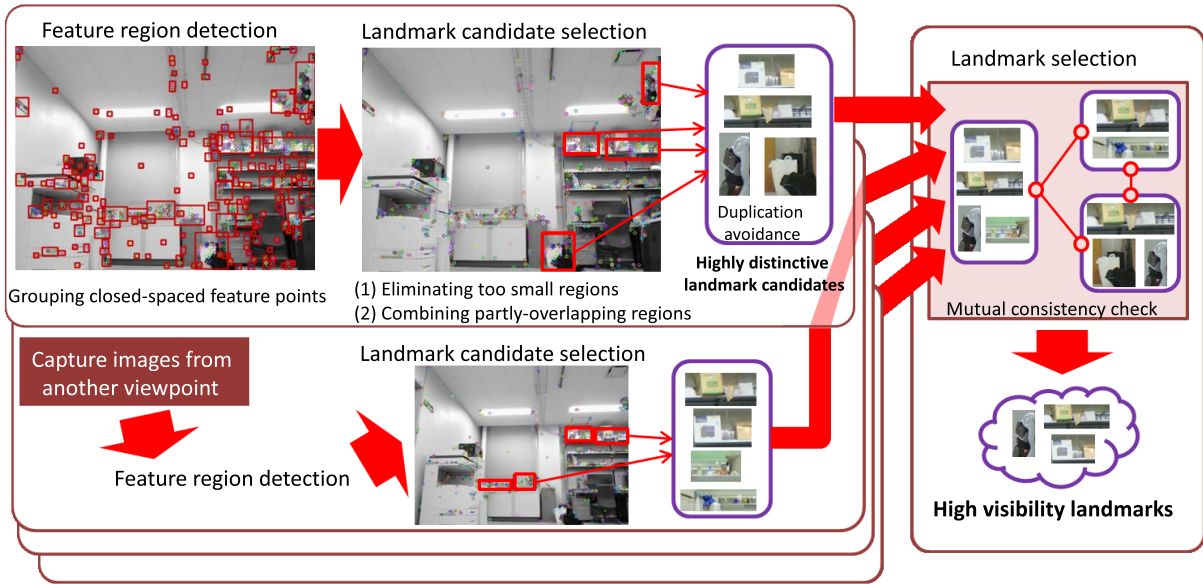


Fig. 1: Landmark selection pipeline

場合に SIFT 特徴量 [3] のマッチングを用いる。SIFT は各特徴点に対し整数型の 128 次元の特徴量ベクトルを記述でき、その特徴量は拡大・縮小、回転、照明変化に頑強である。このことから、ロボットの走行位置や時間に影響されにくい特徴量といえる。この SIFT 特徴量を 2 枚の画像間で比較することにより、画像の類似およびランドマーク等の認識の判定をおこなう。

本研究では、SIFT 特徴量同士の比較によりマッチングをおこなうが、条件によりそれを 2 つに大別する。1 つ目はランドマーク等の 1 枚の入力画像内から切り出した画像領域同士のマッチングである。以降これをマッチング I とし、周囲に類似した領域が存在するかを判定する場合などに用いる。2 つ目は画像領域と入力画像のマッチングである。以降これをマッチング II とし、入力画像内にランドマークが存在するかを判定する場合などに用いる。

マッチング I ではまず、2 枚の画像から SIFT 特徴量を求める。次に、2 枚の画像 (A : 特徴点数 k_1 , B : 特徴点数 k_2 。ただし $k_1 \leq k_2$) の間で、 A の各特徴量ベクトルと B のすべての特徴量ベクトルとのユークリッド距離 b_1 を計算する。 b_1 が最も小さい組み合わせで、かつ b_1 の値が閾値以下であればその 2 つの特徴量ベクトルをマッチしたとする。これを画像 A のすべての特徴量ベクトルに対して行い、マッチした数 m_1 が一定以上だった場合この画像 A, B は類似しているとする。

マッチング II ではマッチング I と同様の処理をおこなうが、特徴量ベクトル間のユークリッド距離の閾値を別の値 b_2 に設定し、マッチの誤対応の削除をおこなう。まず、ランドマークの中心点を基準点と考える。次に、ランドマークの各特徴点と基準点の位置ベクトルから、その特徴点とマッチした入力画像中の特徴点における基準点を求める。このため、入力画像中の基準点 (X, Y) を以

下のように計算する [4]。

$$\begin{aligned} X &= x_i - \frac{\sigma_i}{\sigma_l} \times \sqrt{\Delta x^2 + \Delta^2} \times \cos(\theta + \theta_l - \theta_i), \\ Y &= y_i - \frac{\sigma_i}{\sigma_l} \times \sqrt{\Delta x^2 + \Delta^2} \times \sin(\theta + \theta_l - \theta_i), \\ \theta &= \tan^{-1} \frac{\Delta y}{\Delta x}, \end{aligned} \quad (1)$$

ここで、 x_l, y_l はランドマークに含まれる特徴点の座標、 σ_l はスケール、 θ_l は輝度勾配の方向である。一方で、 x_i, y_i は入力画像に含まれる特徴点の座標、 σ_i はスケール、 θ_i は輝度勾配の方向である。 $\Delta X, \Delta Y$ は位置ベクトルである。

これにより、入力画像中の基準点の座標を求めることができる。入力画像中にランドマークとなるものが存在するならば、求められる基準点は同じ位置に存在する。入力画像中の各基準点の座標を中心とした半径 d の円内に存在する基準点の数を求め、それが閾値 m_2 を超えていれば、入力画像中のランドマークを認識したとする。

2.4 ランドマーク候補の抽出

入力画像から 2.2 節に示す方法で複数の特徴領域を得ることができたとする。この中で特徴領域内の特徴点数 k が一定以下のものを削除する。その後、面積の一定以上が重なっている特徴領域同士を 1 つに統合する。

ここで、周囲に類似した特徴領域が存在するかを確認するため、特徴領域同士のマッチング I を総当りでおこない、類似している特徴領域が存在するかを判定する。類似した特徴領域が入力画像中に存在した場合、それらの特徴領域を削除し、残った特徴領域をランドマーク候補とする。これにより、周囲に類似したものがない弁別性のよいランドマーク候補が得られる。

2.5 ランドマークの選定

上述の方法で得られたランドマーク候補から 2.1 節の条件 (2), (3), (4) に従いランドマークの選定をおこなう。ある地点で位置を少しづつ変えながら撮影した画像を n

枚用意する．この n 枚の画像群からそれぞれランドマーク候補を抽出し，以下の 4 つの方法により抽出されたランドマークの選定をおこなう．

- A ランドマーク候補同士の単純なマッチング
- B 入力画像へのランドマークのマッチング
- C 重み特徴点法
- D 重み係数法

A では自身を含むすべてのランドマーク候補画像に対しマッチング I を行い，それぞれのランドマーク候補の類似しているランドマーク候補数 l_c を求める．その後，類似しているランドマーク候補数 l_c が一定数以下の画像を削除する．これにより，類似している画像の組み合わせが残り，この中から最も l_c が多いランドマーク候補をランドマークと決定する． l_c が同数の場合，特徴点が最も多いランドマーク候補をランドマークと決定する．

B はランドマーク候補と全ての入力画像とのマッチング II をおこない，ランドマーク候補の認識の可否を判定し，認識した回数 l_i を求める．この認識回数 l_i を全てのランドマーク候補について求める． l_i が閾値以上のランドマーク候補の中でマッチング II を総当りでおこない，類似しているランドマーク候補が存在しない場合，そのランドマーク候補をランドマークとする．類似したランドマーク候補が存在した場合， l_i が大きいランドマーク候補をランドマークとする．

C は B と同様，ランドマーク候補と入力画像のマッチングをおこなうが，ここでランドマーク候補の各特徴点のマッチ数を記録する．これは，視点を変えた複数の入力画像中でその特徴点を見つけることができた回数であり，これをその特徴点の重み g_i とする． g_i が閾値以上の特徴点数を g とし，重みのある特徴点とする．この g をそのランドマーク候補の特徴点数 k で割った値を良特徴点密度 h としこの値が大きいランドマーク候補がランドマークとなる．

$$h = \frac{g}{k} \quad (2)$$

D は C と同様に特徴点の重み g_i を求め， g_i の総和を特徴点数で割った値を重み係数 G とし，この値が大きいランドマーク候補がランドマークとなる．

$$G = \frac{\sum g_i}{k} \quad (3)$$

他のものに隠れていたり，場所により見え方が変わるランドマーク候補は l_c ， l_i ， g ， G の値が小さくなる．このことから，これらの値が大きいランドマーク候補を選定することにより，2 章に示した条件 (2)，(3)，(4) を満たす持続性のよいランドマークを選定することができる．

3 実験

3.1 実験条件

車輪型移動ロボットに 1 台のカメラを搭載し，実験をおこなった．移動ロボットは T-frog project [5] の “i-Cart mini” であり，カメラは BUFFALO Inc. の “BSW32KM” を用いた．一台のノート型計算機をロボットに搭載し，

VGA サイズの画像の取得と，ロボットへの走行命令送信に利用した．

3 章で述べた手法により，実際にランドマークの記録をおこなう．図 7 に示すように，ある基準地点 (i) を設け，その地点からロボットの進行方向およびそれに直行する方向に $\pm 0.2[m]$ ごとに場所を移動した合計 9 箇所画像を撮影した．これを信州大学工学部キャンパスの 10 の地点でおこない，2.5 節に示す 4 つの方法によるランドマーク選択処理を適用した．

3.2 ランドマークの記録と評価方法

各種パラメータの閾値は次のようである．特徴点間の距離 $a \leq \sqrt{10}$ ，特徴領域を削除する際の特徴点数 $k \leq 20$ ，マッチング I のユークリッド距離 $b_1 \geq 150$ ，類似判定のマッチ数 $m_1 \geq (k_1 + k_2)/10$ ，マッチング II のユークリッド距離 $b_2 \geq 250$ ，認識判定のマッチ数 $m_2 \geq k/5$ ，基準点を数える際の半径 $d \leq \sqrt{10}$ ，視点を変えた入力画像の枚数 $n = 9$ ．手法 A のパラメータは，ランドマークの類似数 $l_c \geq 3$ ，ランドマーク候補の類似数 $l_c \leq 2$ とした．手法 B のパラメータは，ランドマークの認識数 $l_i \geq 9$ ランドマーク候補の認識数 $l_i \leq 8$ とした．手法 C のパラメータは，ランドマークの特徴点の重み $g \geq 9$ とした．A，B のランドマーク候補は，条件を満たすものの中からランダムに 8 つを選択した．C，D では，それぞれ h ， G の値が各地点ごとに大きい順に 8 つを選択した．

全ての走行経路が異なった入力映像でランドマークの選択をおこない，認識回数 n の平均 μ と偏差 σ を求める．また，認識回数の変動係数 σ/μ を求め，これを視認性の評価値として用いる．変動係数の値が小さいと，走行経路が異なった入力映像での認識回数のばらつきが小さいことを意味しており，視認性の高いランドマークであるといえる．

3.3 実験結果

A について，類似数が 3 以上のランドマークは 2 以下のランドマーク候補と比べて変動係数の値が大きいとは限らなかった．変動係数が大きいランドマークが複数ある地点を見ると，走行経路が進行方向に対して垂直に変化したときの映像での認識回数が少ない傾向がある．これはロボットの走行方向に対して，垂直方向への視点の見え方の変化が激しいためといえる．

B については，こちらも A と同様，認識数が 9 のランドマークは認識数が 8 以下のランドマークと比べ，変動係数の値が大きいとは限らない．変動係数が大きい原因も A と同様のことが言える．

C については，重みのある特徴点のみを使用した場合は全ての特徴点を使用した場合に比べ，平均認識回数が大きく，変動係数が小さい．この結果より，見つけやすい特徴点を選択できたといえる．変動係数の関係から走行経路が異なる映像においても，毎回安定して認識できる視認性の高い特徴点を選択できていた．また，重みのある特徴点のみを使用した場合は処理時間が 100 フレームあたり平均 23.89[s] であり，全ての特徴点を使用した場合の処理時間 26.06[s] に比べ 8.33 % 軽減できた．これらのことから，処理時間を削減し，マッチングの精度も向上できているといえる．

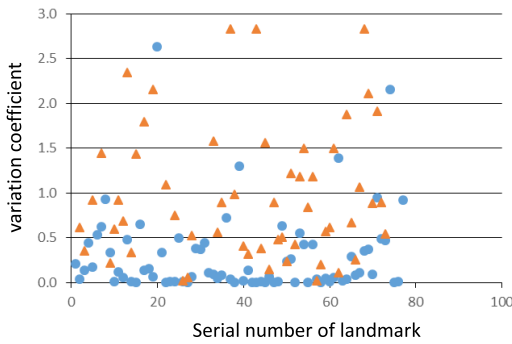


Fig. 2: Serial number of landmark vs. variation coefficient

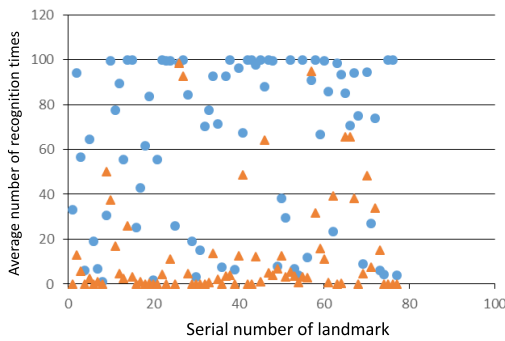


Fig. 3: Serial number of landmark vs. average number of correspondence

D については、重み係数が大きくなるにつれて平均認識回数が大きくなる傾向が見られた。このことから、見つけやすい特徴点に対して重みをおくことができたといえる。また、重み係数が大きくなるにつれ、変動係数が小さくなる傾向が見られた。このことから、重み係数が高いランドマークは、視認性の高いランドマークといえる。特に、重み係数が5.0以上のランドマークは変動係数の値が小さく、特に視認性に優れているといえる。

C, D についてさらに考察を行う。図2は、選出されたランドマークの通し番号と変動係数の関係を表したものである。“all”はランドマークに含まれるすべての特徴点を用いた場合，“only”は重み係数 g を持つ特徴点のみを利用した場合である。明らかに後者のほうが変動係数が小さいので、視認性の高いランドマークの選出に有用と言える。図3は、ランドマークの通し番号と正しい対応が取れた回数との関係を表したものである。重み係数を定義できた画像特徴は、多数回の検出が可能であった。

図4は、重み係数と正しい対応の回数を表したものである。重み係数が大きな特徴ほど、正しい対応の回数が多いことがわかる。本実験においては、例えば重み係数が5.0以上のものだけを選べば、それが視認性の高いランドマークの選出と等しいと言える。同様の傾向は図5からも見て取れる。

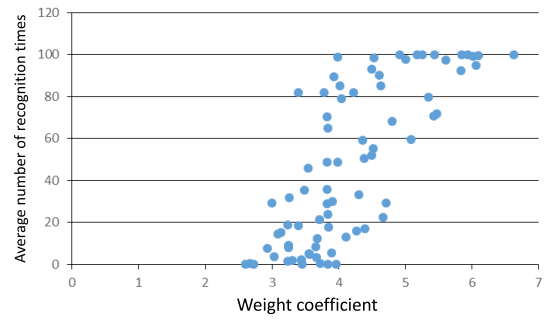


Fig. 4: Weight coefficient vs. average number of correspondence

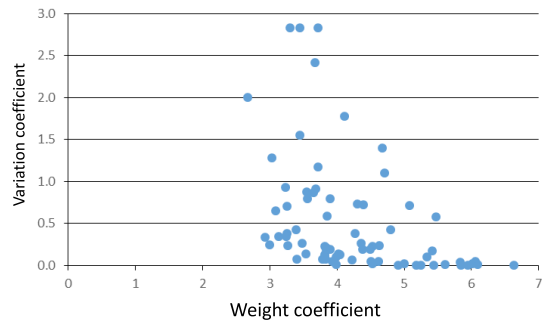


Fig. 5: Weight coefficient vs. variation coefficient

4 まとめ

本稿では、視認性に基づくランドマークを記録する手法を提案した。また、実験により提案手法の課題と有効性を確認した。4種類のランドマーク選択の方法を示し、所定の基準を満たす特徴量のみを使用すれば視認性の高いランドマークを選出できることを示した。

今後の課題として、輝度変化に対する検証や四季変化のように広い時間軸での視認性の評価が挙げられる。

参考文献

- [1] Y. Matsumoto, M. Inaba, and H. Inoue: “Visual navigation using view-sequenced route representation,” in Proc. of Int’l Conf. on Robotics and Automation, pp. 83 – 88, 1996.
- [2] Y. Kaneko and J. Miura: “View Sequence Generation for View-Based Outdoor Navigation,” in Proc. of 1st Asian Conf. on Pattern Recognition, pp. 139 – 144, 2011.
- [3] D. Lowe: “Distinctive Image Features from Scale-Invariant Keypoints,” Int’l Journal of Computer Vision, Vol. 60, No. 2, pp. 91 – 110, 2004.
- [4] A. Ihara, H. Fujiyoshi, M. Takagi, H. Kumon, Y. Tamatsu: “Improved Matching Accuracy in Traffic Sign Recognition by Using Different Feature Subspaces,” in Proc of Int’l Conf. on Machine Vision Applications, pp. 130 -133, 2009.
- [5] <http://t-frog.com/en/> (19/12/2014)