

Haar-Like 特徴を用いた屋外自律移動ロボットのナビゲーションのためのシーンマッチング

塩谷 敏昭^{†‡} 太田 直哉[†]

[†] 群馬大学大学院工学研究科 〒376-8515 群馬県桐生市天神町 1-5-1

[‡] ミツバ 〒376-8555 群馬県桐生市広沢町 1-2681

E-mail: shioya@ail.cs.gunma-u.ac.jp

あらまし ビューシーケンスに基づく自律移動ロボットのナビゲーションにおいて、シーン探索が重要な課題である。特に屋外移動ロボットでは、環境光の変化や、見る位置の変化による、テクスチャの変化が問題となる。そこで、細かいテクスチャ変化にロバストなおおまかな特徴を記述でき、かつ細かいテクスチャも表現できる Haar-Like 特徴を用いた特徴記述方法を提案する。また、多くのミスマッチを含む対応点の組から、最適な画像方向を推定する方法として、2 次元確率分布を計算する方法を示し、有用性を確認する。さらに、遊歩道や、ペDESTリアンデッキなどの人が日常通行する環境において、晴れの日と曇りの日でシーン探索の評価をし、SURF 特徴との比較において、提案手法の有用性を示す。

キーワード Haar-Like 特徴, ビューシーケンス

1. はじめに

自律移動ロボットのナビゲーションにおいて、松本らは、教示走行により、ある間隔で連続的に撮影した画像の列を記憶（ビューシーケンス）しておき、自律走行時に得られる画像とビューシーケンス中の画像とのマッチングを行い現在位置を認識する方法を提案している^[1]。この方法においては、シーン全体の見え方を用いるために、人工的にランドマークを敷設したり、画像中からランドマークを抽出するという処理が必要なくなるという利点があるが、画像のマッチングを行うことが重要な課題である。金井らは、画像マッチングの方法において、SURF 特徴を用いている^[2]。

画像のマッチング方法としては、パターンマッチングのように、画像全体の見え方を比較する方法と、SIFT^[3], SURF^[4]といった局所の特徴量を用いた方法がある。局所の特徴量を用いた方法では、画像全体の見え方を比較する方法に比べ、隠れや、全体的な見え方の変化に対して影響を受けにくいことが期待される。

移動ロボットのナビゲーションにおいては、歩行者など、他の移動物体による隠れや、ロボットの移動による視点位置の変化による全体的な見え方の変化が起こりうるので、局所特徴を用いたシーンマッチングは有効であると考えられる。

局所特徴を用いた画像のマッチングでは、特徴点どうしの対応関係を得る段階と、誤りを含む対応関係の情報からマッチする画像が含まれるかどうか、どこに含まれるかの判定を行う段階とがある。

特徴点どうしの対応関係を得る段階において、SIFT と SURF は、特徴点の周辺に 4×4 のタイル状に配した小領域において、 x 方向・ y 方向のグラディエントを特徴ベクトルの要素とし、それぞれの領域内で、要素ベクト

ルの大きさを 1 に正規化している。これは小領域内に収まる比較的高い周波数の空間的情報を保持しているが、領域間にまたがる低い周波数の情報は失われることになる。この結果、特徴点の識別能力は低下する。

特徴点のマッチング判定は、特徴ベクトルのユークリッド距離を比較し、さらに、2 番目に近い特徴点との距離の比を用いるのがよいとされている。これは、ミスマッチを避けるために有効である。しかし、正しく対応する特徴点間の特徴ベクトル距離が近くても、特徴ベクトル間距離が近い、対応しない特徴点が存在すると、マッチング判定されなくなるという問題がある。これは、特徴記述子の識別能力の問題である。

SIFT・SURF ではロバストな物体検出のために、スケールインバリエント、回転インバリエントな機能を備えている。これは、スケールピラミッドから 3 次元的にピークをとることによって、特徴点のサイズを決定し、方向ヒストグラムから特徴点の方向を決定し、これらを元に特徴記述のための小領域のサイズ・位置を決定することにより実現している。しかし、地上を移動するロボットでは、カメラが進行方向を軸に回転することはほとんどなく、カメラの進行方向への移動によって、記憶されたシーンと近いスケールでのマッチングが可能であることから、これらのロバスト性向上のための機能は不要である。

一方、物体認識では、Viola・Jones らが、Haar-Like 特徴を、シンプルな基本的特徴として用いている^[5]。これは、Haar wavelet 変換から、物体を効率的に識別できる要素を多くのサンプルからの学習によって抽出していると考えられる。この方法では、高速に、学習サンプルに含まれる変化に対応するロバストな検出をすることができ、多くの学習サンプルを必要とするため、自律移動

ロボットのナビゲーションにおいて、何度も教示走行をする必要があり、ユーザーにとっての負担が大きいという問題がある。

誤りを含む対応関係の情報からマッチする画像が含まれるかどうか、どこに含まれるかの判定を行う段階においては、金井らは、RANSAC により画像間の透視変換を得て、記憶したシーンとのマッチ判定、位置認識をおこなっていた。しかし、この方法ではミスマッチが多いと誤った透視変換が出力されるという問題があった。

我々は、これらの問題に対して、3つの方法を提案する。

ひとつは、低い周波数の情報を保持する、サイズの異なる小領域を用いること。ふたつ目は、あらかじめ選択された、いくつかの Haar-Like 特徴を、特徴ベクトルの要素として用いること。最後に、特徴ベクトルの比較から得られた、対応関係から、2次元確率分布を計算し、そのピークをとることで判定する方法を提案する。

特徴点の抽出について、SURF では、Fast-Hessian 検出器を用いている。我々の実験では、OpenCV に含まれる cvSurf.cpp に実装された Fast-Hessian 検出器を使用して、特徴点検出を行う。ここで、抽出された点に対して、SURF 特徴による特徴点記述により、対応点検出をした場合と、Haar-Like 特徴による特徴点記述により対応点検出をした場合とで比較し、提案手法の有用性について述べる。さらに、2次元確率分布のピークから視点の移動量を検出を試み、より多くのシーンにおいて、正しく検出できることを示す。

2. Haar-Like 特徴

SURF では、fig.1 に示すように、特徴点の周囲の領域を 4×4 の小領域に分割し、それぞれの領域において、 x 方向・ y 方向のグラディエントおよびその絶対値、 $v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$ の4要素をもつ特徴ベクトルを用いている。したがって、ひとつの特徴点に対し、64要素の特徴ベクトルとなる。また、これを拡張し、128要素のベクトルを用いる方法も提案されている。

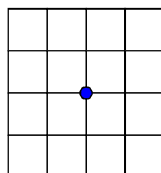


fig. 1 sub-regions for the extraction of the descriptor.

我々は、fig.2 に示す 9つの Haar-Like 特徴を使用した。

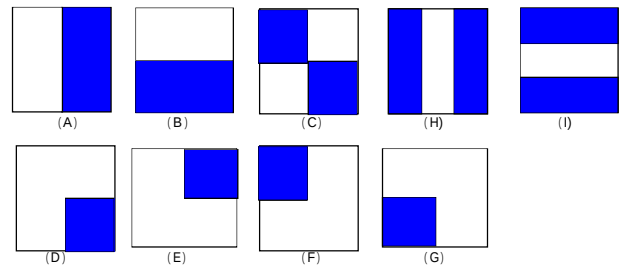


fig. 2 Haar wavelet types used for our process.

Haar-Like 特徴のサイズは、 24×24 ピクセル、 12×12 ピクセル、 6×6 ピクセルとし、特徴点の周囲に 12×12 は4つ、 6×6 は16の領域をタイル状に並べる。したがって、21のサイズ・位置に9の Haar-Like 特徴を適用することで、189要素の特徴ベクトルとなる。

fig.2の(D)~(G)および、(H),(I)は、-1, 1の面積が違うため、本来応答しない、一様な画像に対してオフセットを生じる。そこで、オフセットを打ち消すようにそれぞれの領域に重みを付ける。

fig.2(D)では、fig.3に示すように $S_{11} \sim S_{22}$ の4領域に分け、(1)式のように計算する。

s_{11}	s_{12}
s_{21}	s_{22}

fig. 3 sub-regions for calculate(D) ~ (G).

$$\{(S_{11}+S_{12}+S_{21})+3S_{22}\}/54 \quad (1)$$

(E)~(G)は、(D)を回転させて、同様に計算する。

(H)では、fig.4に示すように $S_1 \sim S_3$ の3領域に分け、(2)式のように計算する。

s_1	s_2	s_3
-------	-------	-------

fig. 4 sub-regions for calculate (H),(I).

$$\{(S_1+S_3)+2S_2\}/48 \quad (2)$$

(I)では、(H)を回転させて、同様に計算する。

同じカメラ位置で、時間をおいて撮影した2枚の画像 image1, image2 について、recall vs. 1-precision^[6] グラフを作成し、SURF 特徴と、提案する Haar-Like 特徴の比較をすることで、提案手法が正しくマッチングしているこ

とを確認する．fig.5 は，画像上の距離 2 ピクセル以下のものを正しい対応点として，image1 と image2 のすべての特徴点の組み合わせについて，特徴ベクトル間のユークリッド距離によりマッチ判定した場合の recall vs. 1-precision グラフにおいて，Haar-Like 特徴と 4 つの SURF 特徴と比較した結果を示す．surf64 は，標準の SURF 特徴，u-surf64 は，スケールインバリエント・回転インバリエント機能をスキップした SURF 特徴，surf128 は，拡張 SURF 特徴，u-surf128 はスケールインバリエント・回転インバリエント機能をスキップした拡張 SURF 特徴をそれぞれ表している．ここでは，特徴ベクトルのユークリッド距離でマッチング判定している．

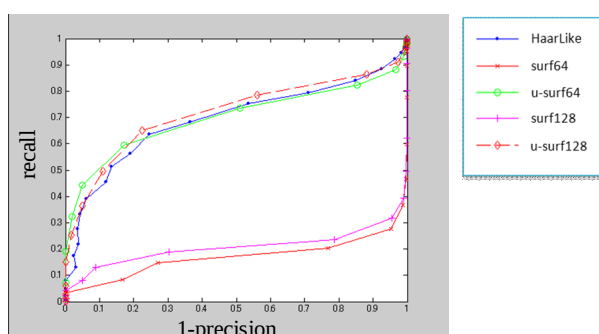


fig. 5 The recall vs. (1-precision) graph for match by the Euclidean distance between the feature vectors.

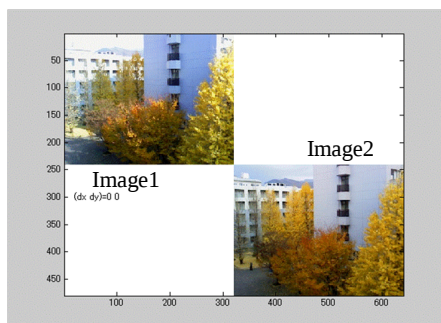


fig. 6 Images for evaluation

fig.5 から，マッチング判定を特徴ベクトルのユークリッド距離でした場合，提案する Haar-Like 特徴は，u-surf，u-surf128 とほぼ同程度の性能であることが確認された．fig.6 に，評価に使用した画像を示す．

fig.7 に，2 番目に近い特徴点との比により，マッチ判定をした場合の recall vs. 1-precision グラフを示す．この結果から，2 番目に近い特徴点との比でマッチング判定した場合は，すべての手法で性能向上が見られるが，提案する Haar-Like 特徴が，最も良いことが確認された．

3. 2次元確率分布

自律移動ロボットのナビゲーションにおいて，ビュー

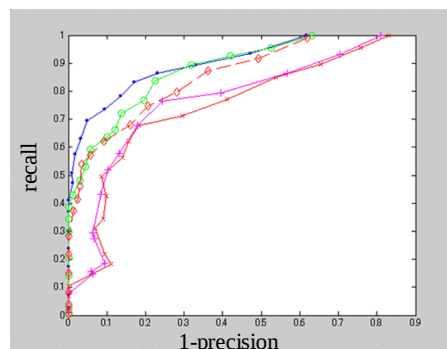


fig. 7 The recall vs. (1-precision) graph for match by the ratio with the second nearest neighbor.

シーケンスに記憶された画像と入力画像とのマッチングを行い，記憶されている画像が見えているかどうか，また，入力画像のどの部分に見えているかを知ることが課題となる．特に，どの部分に見えているかを知ることが，カメラが，記憶された画像を撮影した位置・方向に対して，現在どの方向を向いているかを知るために，非常に重要である．これは，特徴点どうしの対応関係を知ることができれば得られる．

対応関係の評価は，特徴ベクトル空間でのユークリッド距離を比較し，最も小さい特徴点の組みと 2 番目に小さい特徴点の組の比で示す．評価値が小さいほど，強い対応関係を示している．

対応関係にはエラーが多く含まれていることが問題である．対応しない特徴点どうしの画像上のズレ位置が偶然一致する確率はきわめて低いと言える．したがって，画像上のズレ位置が集中する点は正しい対応点を示している確率が高い．fig.8 に，ほぼ同じ位置から撮影した 2 枚の画像について調べた特徴点と対応関係を示す．

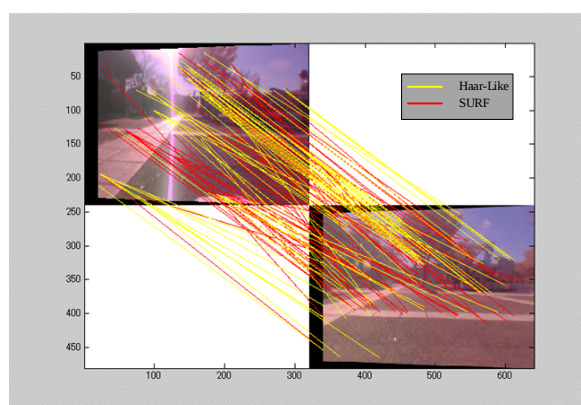


fig. 8 Correspondency between the feature points

シーケンスに記憶された画像と入力画像とのマッチングそれぞれの特徴ベクトルから得られた対応関係を線で結んで示している．正しい対応関係はそれぞれ平行で同

じ長さの線になっている．ほぼ同じ位置から撮影した画像なので，正しい対応関係の特徴点どうしは，画像上で近くにあるはずである．fig.9 に対応関係を示された特徴点の画像上の距離と，対応関係の評価値を示す．

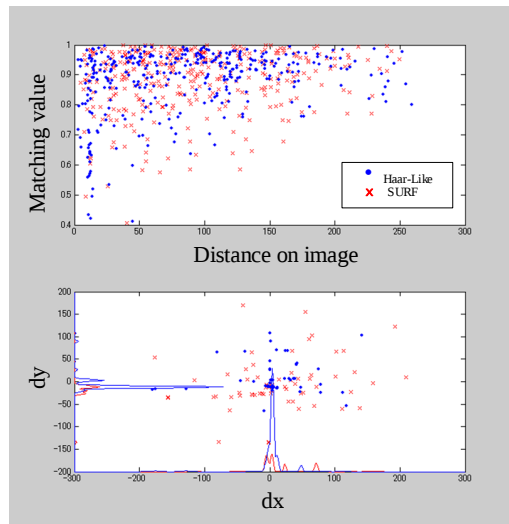


fig.9 The upper section : Distance of matched feature on image. The lower section : Difference vector on image.

fig.9 上段は対応関係を示された特徴点間の画像上の距離を横軸に，対応関係評価値を縦軸に示したものである．この図で，左下に近い点は，強い対応関係を示された特徴点の組が，それぞれの画像上の近い位置から検出された点であり，正しくマッチングされていることを示している．右下に近い点は，画像上で遠い点どうしが強い対応関係を示していることになり，ミスマッチであると判断できる．下段は対応関係評価値 0.8 以下の組の画像上の位置の差を点で示している．この図で，点が集中しているところが，2つの画像間のずれを示している．

これらの点から，対応関係評価値で重みづけた2次元確率分布を計算する．点(x,y)の確率は式(3)で計算される．

$$P(X,Y) = \sum \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(X-x_i)^2 + (Y-y_i)^2}{2\sigma^2} w_i \right\} \quad (3)$$

ここで， x_i ， y_i は，特徴点位置の差， w_i は対応関係評価値の逆数， $\sigma=3$ とした．

fig.10 に確率分布を画像化したものを示す．上段は Haar-Like 特徴の結果，下段は SURF の結果である．この結果から最も高いピークを検出し，その位置を画像のずれ量と判定する．fig.9 下段の下端，左端に，fig.10 の確率分布のピークを通る水平・垂直断面を示す．

このピークの高さで，2枚の画像間のマッチ度を評価することができる．ここで得られた画像間のズレ位置を元に，2枚の画像を重ねて示し，目で見て正しいずれ位

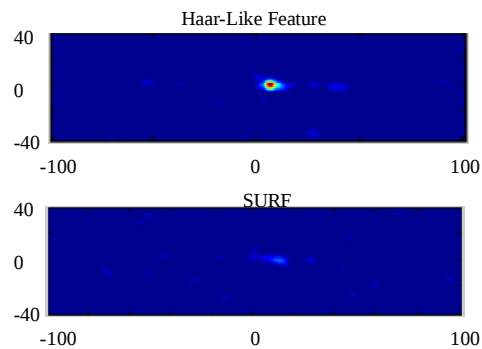


fig. 10 Two-dimensional probability distribution with the weight.

置が検出されていることを確認した．

確率分布のピークの変化から，ビューシーケンスに記憶された画像を撮影した地点に対して最も近づいた地点を知ることができる．

ビューシーケンスに記憶された画像を撮影した地点に対して近づいていくときは，正しくマッチングできる確率が上がるため，確率分布のピークは次第に高くなっていき，遠ざかっていくときは，確率分布のピークは次第に低くなっていく．fig.11 に，fig.8 に示したシーンの前後でカメラが移動したときの確率分布のピーク高さの変化を示す．

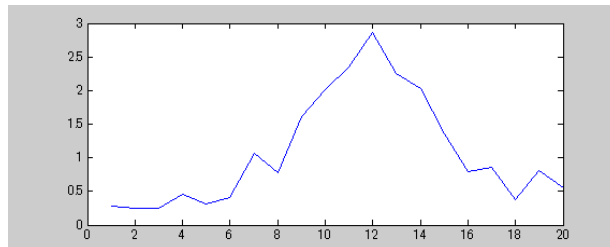


fig.11 A change of the peak height

4. 実験方法と結果

人が歩く一般的な環境での画像による実験を行う．つくば市中央公園遊歩道を中心につくばエクスプレス駅周辺ペディストリアンデッキを含めて，全長 1.2km のコースが，つくばチャレンジ 2010^[7]の規定コースとして設定されている．このコースを fig.12 に示す．このコースから多様なシーンを含むように 20 のシーンを選択しマッチングの実験をした．選択したシーンの画像を fig.13 に示す．このコースを晴れの日に 2 回，曇りの日に 1 回教示走行したときの画像を用い，この中から手作業で，ほぼ同じ位置から撮影された組を抽出した

これらの画像の組に対して，マッチングを行い，2次元確率分布のピークが得られることを確認する．また，得られたピークの位置から，画像をずらして照合し，正し



fig. 12 Course of Tsukuba Challenge2010



fig. 13 The image which was took in each scene

い位置に認識されたか人の目で確認する。
 晴れの日と曇りの日でマッチングした結果の例を fig.14 に示す。
 シーン 9 は、良くマッチングできた例である。提案する Haar-Like 特徴でも、SURF 特徴でも、良く確率分布にピークが現れている。シーン 13 は全体的にミスマッチが多く、正しい対応点がほとんどない結果である。この場合でもピークは見られるが、正しい位置ではない。シーン 18 は、提案する Haar-Like 特徴は良くマッチングできているが、SURF ではうまくいかない例であ

る。SURF の場合でも、ピークは見られるが、正しい位置ではない。fig.15 に、20 シーンそれぞれで得られた 2 次元確率分布のピーク値を示す。
 fig.15 は、横軸にシーン番号、縦軸に 2 次元確率分布のピーク値をプロットした。×印は、確認の結果、ただしずれ位置を検出できていないシーンである。ほとんどのシーンで、提案する Haar-Like 特徴の方が、大きなピークを得ることができた。シーン 13、シーン 18 では、SURF 特徴によるマッチングに失敗している。同様に、晴れの日と曇りの日に撮影した 2 通りの画像からマッ

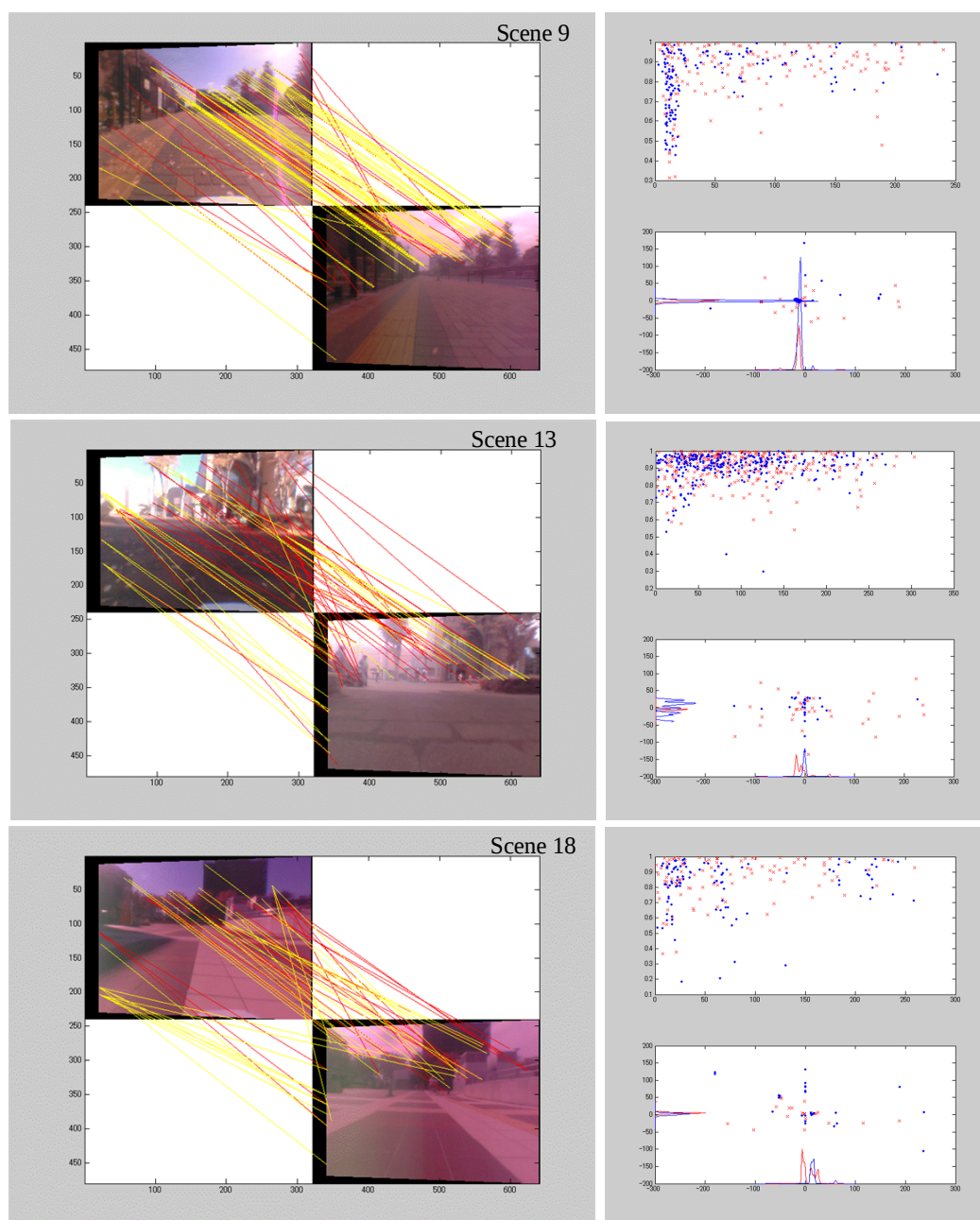


fig. 14 The matching result of a fine day and the cloudy day.

ングを行った結果について、20 シーンそれぞれで得られた2次元確率分布のピーク値を fig.16 に示す。晴れの日同士の日データでは、すべてのシーンで Haar-Like 特徴が良い結果となっている。ただし、Haar-Like 特徴ではシーン 13 で失敗、SURF ではシーン 5, 13, 17 で失敗している。

これらの結果から、晴れの日どうしの方が、よりマッチングが難しいことが分かる。

fig.17, fig.18 に、それぞれ、晴れの日と曇りの日、晴れの日どうし、曇りの日と雨の日の 20 シーンの画像に

おいて、recall vs. 1-precision グラフを示す。

5. おわりに

屋外型自律移動ロボットナビゲーションのためのビュースケッチ画像のマッチングについて、特徴記述方法として、いくつかの異なるサイズで 9 パターンの Haar-Like 特徴を用いる方法を提案し、SURF 特徴マッチングの方法と比較して、有用性を確認した。特徴ベクトルのユークリッド距離でのマッチ判定では、スケール・回転インバリエント機能をスキップした SURF 特徴と同程度の性能であったが、2 番目に近い点との

比でマッチ判定する方法では、最も良い結果が得られた。

また、多くのミスマッチを含む特徴点对応関係から、もっともらしい画像のずれ量を求める方法として、2次元確率分布を計算する方法を提案した。得られた画像間のズレ位置から手作業で確認した結果、ほとんどのシーンで正しいマッチング結果が得られることが確認された。

一般的に人が歩くようなシチュエーションで撮影された画像を用いてマッチングの実験を行い、ほとんどのシーンにおいて SURF よりよいマッチングができることを確認した。

今後の課題として、特徴ベクトルの計算に用いる Haar-Like 特徴のパターンについて、最適なパターンを採用できているか、最適な数であるかを確認できていない。さらによりマッチングができるパターンがあることも考えられるので、この部分について、さらに検討する必要がある。

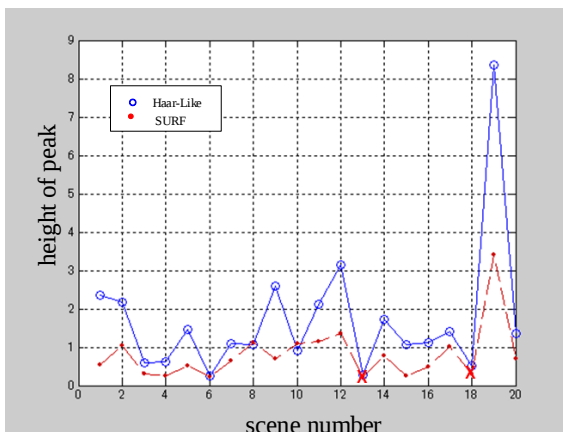


fig. 15 The height of the peak in a fine day and the cloudy day.

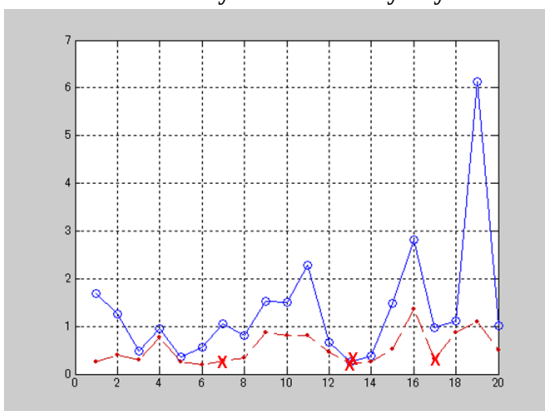


fig. 16 The height of the peak in a fine day and a fine day.

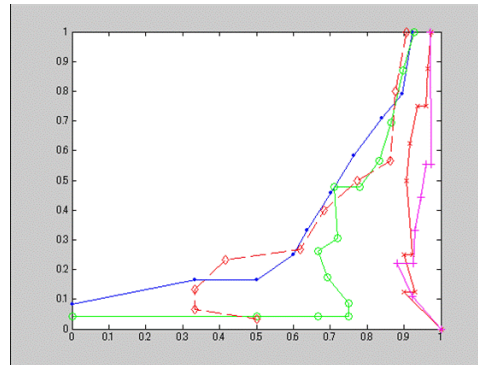


fig. 17 The recall vs. (1-precision) graph for a fine day and a cloudy day.

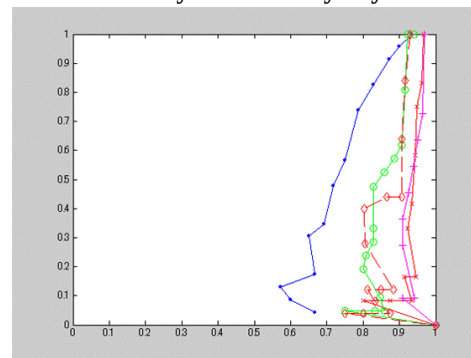


fig. 18 The recall vs. (1-precision) graph for a fine day and a fine day.

文 献

- [1] 松本吉央, 稲葉雅幸, 井上博允, “ビューベーストアプローチに基づく移動ロボットナビゲーション”, 日本ロボット学会誌 Vol. 20 No. 5, pp.506 ~ 514, 2002
- [2] 金井克友, 浦野和典ほか, “画像認識による自己位置認識と障害物検出を搭載した自律移動ロボットの開発”, 第 10 回計測自動制御学会システムインテグレーション部門講演会 (SI2009) 予稿集, 東京, 2009.12, pp.1009-1012
- [3] D. G. Lowe, “Distinctive image features from scale-invariant key points”, Internal Journal of Computer Vision, 60, 2, pp.91-110, January 2004.
- [4] H. Bay, T. Tuytelaars, and L. V. Gool, “SURF: Speeded Up Robust Features”, Proc. of the 9th European Conference on Computer Vision, May 2006.
- [5] P. Viola, M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, IEEE CVPR, 2001, pp.511-518
- [6] Ke, Y., Sukthankar, R.: PCA-SIFT: A more distinctive representation for local image descriptors. In: CVPR (2). (2004) 506 – 513