リアルタイム全方向レーザスキャナを用いた屋外環境の種別推定

Place Categorization for Outdoor Environment using Real-time Omni-directional Laser Scanner

大音 雄輝(九大) 鄭 好政(九大) 〇 河村 晃宏(九大) 岩下 友美(九大) 正 倉爪 亮(九大)

Yuki OTO, Kyushu University, y_oto@irvs.ait.kyushu-u.ac.jp Hojung JUNG, Kyushu University, Akihiro KAWAMURA, Kyushu University Yumi IWASHITA, Kyushu University, Ryo KURAZUME, Kyushu University

This paper presents a place categorization technique using a real-time omni-directional laser scanner for outdoor environment. Firstly, we created a large database consisting of range data of six categories, that is, coast, forest, indoor parking, outdoor parking, residential area, and urban area. Two kinds of features, spin-image and local binary pattern, are extracted from range data and used for the estimation of the category which the range data belongs to by SVM. The performance of the proposed technique is evaluated through the experiments using the developed database.

Key Words: Outdoor Place Categorization, Laser Scanner, Range Image, Spin-Image, Local Binary Patterns

1 緒言

移動ロボットや自動走行車に求められる重要な機能として、ロボットが置かれた環境の空間の認識が挙げられる。空間の認識は高度な位置推定問題であり、ロボットに対して、正確な座標の推定ではなく、より抽象度の高い空間の種類を認識する能力が求められる。ロボットは空間の種類に関する情報を得ると、次に取るべき行動を決定できる [1]. さらに、空間の種類に関する情報は、典型的な物体の認識において事前情報またはコンテキストとして用いることができ [2]、またセマンティックマップ [3] や空間の概念的表現 [4] などにも利用できる.

空間の認識は、2つの問題に分けて考えることができる。すなわち、特定の空間の認識 (classification) と識別 (categorization)である。特定の空間の認識は、ロボットが以前に訪れたある特定の空間 (例えば 928 室など)を、センサ等から得られる情報を用いて認識するものである。一方、空間の識別は、ロボットがそれぞれの空間のカテゴリ (クラス)を識別する問題であり、例えば先ほどの 928 室は、他の部屋と同様の一般的な概念である"オフィス"として識別される。本研究では、後者の問題を取り扱う。

空間の認識に対して、これまでに様々なセンサを用いた手法が 提案されている。例えば、文献 [5] では 2D レーザスキャナが入力として用いられ、文献 [2][6] では画像センサが用いられた。また、これらの 2 つのセンサを統合した手法も提案されている [7]. 一方、近年 Focus 3D (FARO) や LMS 151 (SICK) のような、3D レーザスキャナが普及してきたことで、周囲環境の高精度且つ高密度なデータを容易に入手できるようになり、移動ロボットに関する研究が盛んに行われている。

本稿では、特に屋外環境の空間の認識に注目する.屋外環境は屋内環境と比べて通行人や駐輪、駐車車両などの動的に変化する障害物が数多く存在し、時刻や気象条件の変動に伴って照明条件が大きく変化するため、一般に、識別はより困難になると考えられる。加えて、動的な空間内での安全な動作を実現するためには、処理のリアルタイム性も求められる。コンピュータによる識別では、識別対象から、その対象の特徴を表現する識別特徴量を抽出する必要がある.この際、周囲の環境そのものを一つの物体として取り扱うことで、通行人の移動等の小さな変化に対する頑健性の向上が期待できる.計算コストの少ない特徴量を用いることで、処理のリアルタイム性も併せて実現できる.また、レーザスキャナはレーザ光の走査により周囲の計測を行うため、照明変化に頑健である.そこで本研究では、レーザスキャナから得られる全周距離画像を用いて、屋外環境における全周距離画像データ

ベースの構築及びそれを用いた屋外環境の種別推定実験を行う.

2 屋外環境の全周距離画像データベース

本研究では、屋外環境の識別を行うために、まずは屋外環境における全周距離画像データベースを構築する。本章では、構築された全周距離画像データベースの詳細を述べる。全方向レーザスキャナは、リアルタイムな計測が可能な Velodyne 社の LiDAR 全方位レーザーライダー・イメージングユニットである HDL-32e を用いる。HDL-32e の最大計測距離は 70m, 測定角度は水平 360° , 垂直 41.3° である。

レーザスキャナを車載し (図 1),様々な環境を走行しながら全周距離画像を取得した.この時の車両の移動速度は $30\sim50 \mathrm{km/h}$ であり,距離画像は 0.5 秒間隔で出力される.データは,Coast(沿岸),Forest(森),ParkingIn(屋内駐車場),ParkingOut(屋外駐車場),Residential(住宅街),Urban(市街地) の 6 種類のカテゴリについて取得し,カテゴリ毎に 10 箇所でデータを取得した.得られる全周距離画像と実際に走行した場所の一例を図 2 に示す.また,データ総数は 34200 枚,データ容量は $59.23 \mathrm{GB}$ である.その詳細を表 1 に示す.

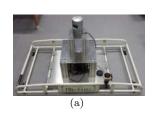




Fig.1 Experimental setup

3 屋外環境の識別手法

本章では、全周距離画像データベースから特徴を抽出するため の特徴記述子について説明する.特徴記述子には、3次元形状記

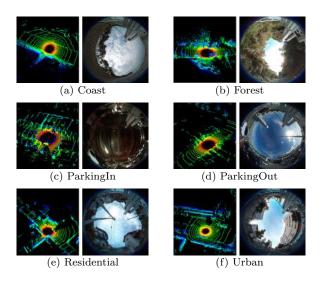


Fig.2 Example of range data

Table 1 The number of range data

		Set									
	1	2	3	4	5	6	7	8	9	10	TOTAL
Coast	511	254	571	221	314	376	872	506	386	287	4298
Forest	440	824	980	707	730	720	439	311	797	531	6479
ParkingIn	520	357	274	873	583	343	466	592	344	428	4780
ParkingOut	874	579	388	370	477	536	581	563	460	617	5445
Residential	674	787	667	724	563	973	717	720	977	662	7464
Urban	490	572	587	487	410	566	712	565	606	739	5734
TOTAL	3509	3373	3467	3382	3077	3514	3787	3257	3570	3264	34200

述子としてよく用いられている Spin-Image と,屋内の環境認識で高い識別性能を示す LBP[10] を用いる.

3.1 Spin-Image

Spin-Image は、形状モデルの頂点の法線ベクトルを軸とした円柱面に、周囲の頂点を投票することで生成される 2 次元画像を用いた、類似形状物体の識別手法である [8]. Spin-Image は、注目頂点の法線方向を基準として周囲の頂点を 2 次元パラメータ (α,β) で表し、 (α,β) を添字とする 2 次元配列上に射影される点の数を投票することで得られる (図 3). 3 次元から 2 次元への射影には式 (1) を用いる.

$$S_O: R^3 \to R^2$$

 $S_O(x) \mapsto (\alpha, \beta) = (\sqrt{\|x - p\|^2 - (n \cdot (x - p))^2}, n \cdot (x - p))$ (1)

但し、p は基準となる頂点 O の位置、n は頂点 O における法線 ベクトルである.

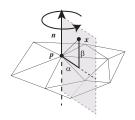


Fig.3 Spin-Image

Spin-Image は本来,複数の頂点に対して計算されるが,本研究では,レーザスキャナの設置位置を注目頂点とし,地面に鉛直な法線ベクトルを軸とする円柱面に,1 データ内の点群全てを投票することで環境全体の特徴を1 枚の Spin-Image で記述する.生成される Spin-Image の一例を図4 に示す.

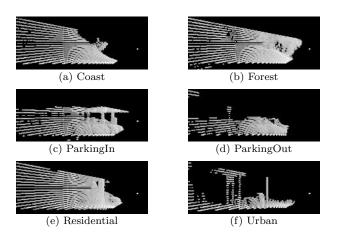


Fig.4 Spin-Image generated from range data

3.2 LBP

局所二値パターン(Local Binary Patterns, LBP)演算子 [9] は、濃淡画像のテクスチャ分析と分類を目的に Ojala により提案された.この演算子は、参照ピクセルとその近傍ピクセルの濃淡値の大小関係をベクトル形式で表すものである.

距離画像を用いた空間の識別では、まずレーザスキャナから得られる距離値を基に濃淡画像を生成し、その濃淡画像を LBP を用いて特徴量画像に変換する。 LBP は近傍画素の情報を用いて注目画素に新たなラベルを付加するものであり、今回は濃淡画像に対して以下のように適用した。まず、濃淡画像のそれぞれの画素 i に対して、その画素値 p(i) を周囲 8 近傍の画素値 $p_1(i)$ 、…、 $p_8(i) \in N(p(i))$ と比較する。次に、それぞれの近傍画素ごとに二値 0.1 を以下の式で決定した。

$$b_j(i) = \begin{cases} 1, & p(i) > p_j(i) \\ 0, & otherwise \end{cases} (j = 1, ..., 8)$$
 (2)

ただし、周囲 8 近傍の画素値 $b_j(i)(j=1,\cdots 8)$ の配置は図 5 の通りである。得られた周囲画素の二値 $b_j(i)$ は反時計回りに $b_1(i),\cdots,b_8(i)$ の順番で並べられ、図 6 に示すように 0 から 255 までのラベルに変換される。注目画素の値は、このラベルにより置き換えられ、LBP に変換された特徴量画像が生成される。HDL-32e では、計測不能あるいは計測範囲外の地点の距離値に 0 がセットされる。そこで、距離値が 0 であるか、近傍画素に 0 が含まれる場合には LBP の出力を nan とした。



Fig.5 Placement of $b_j(i)(j=1,\cdots 8)$

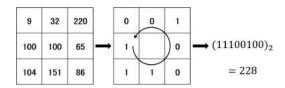


Fig.6 Example of LBP

4 全周距離画像を用いた屋外環境の識別実験

本章では、全周型レーザスキャナから得られる全周距離画像を 用いて屋外環境の識別実験を行った結果を示す.

4.1 屋外環境識別実験

本節ではまず、環境識別の際に用いる識別器について説明する. その後、第2章で構築した全周距離画像データベースを用いて行った実験の結果を示す.

4.1.1 SVM

本研究では、特徴量に対する識別器に SVM を用いた。SVM には、M 個の m 次元教師データ $\mathbf{x}_i \in \Re^m (i=1,\ldots,M)$ に対し、それぞれクラスラベル $y_i \in \{1,-1\}$ (あるクラスに属するか、属さないか)が付加され入力される。SVM の目的は、分離超平面とそれに最も近い教師データとの距離(マージン)が最大となる最適超平面を見つけることであり、次の最適化問題で表される。

$$\min_{\mathbf{w},b,\xi} \quad C \sum_{i=1}^{M} \xi_i + \frac{1}{2} ||\mathbf{w}||^2$$
 (3)

subject to
$$y_i \left(\mathbf{w}^T \phi(\mathbf{x}_i) + b \right) \ge 1 - \xi_i$$
 (4)

ここで,C はマージンパラメータ, \mathbf{w} は超平面に対する垂線, $\xi_i \geq 0$ は \mathbf{x}_i の誤分類に伴う超平面からの逸脱量, $\phi(\cdot)$ は非線形写像,b はバイアス項を表している.実験では radial basis function (RBF) カーネル(式 (5))を用いた.カーネルの式を以下に示す.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma ||\mathbf{x}_i - \mathbf{x}_j||^2\right) \tag{5}$$

多クラス問題への拡張は一対一方式 [11] を用い、識別実験では LIBSVM ライブラリ [12] を用いた。 カーネルのパラメータ C, γ の最適値は、総当たり法 [13] により $C \in [2^{-5},\dots,2^{12}]$, および $\gamma \in [2^{-16},\dots,2^3]$ の範囲で求めた。

4.1.2 実験方法

本実験では、構築したデータセットを用いて各特徴量の屋外環 境識別性能を評価する.実験の手順を以下に示す.

- 1. 各カテゴリについて 10 セットの内から 1 つずつランダム に選択し、選択したセットのデータ全てをテストデータと オス
- 2. 各カテゴリの残り 9 セットのデータ全てを学習データとする.
- 3. 各学習データの距離画像から特徴量を抽出し、SVM を用いて学習を行う.
- 4. 各テストデータからも同様に特徴量を抽出し, SVM により 識別を行う.
- 5. 1.~4. を 10 回繰り返す.

また、SVM へのデータの入力は以下のように行う.

Spin-Image は、全周距離画像から生成される画像 (図 4) をベクターに変換して SVM に入力する. LBP は、まず、全周距離画像を基に生成されるパノラマ画像から LBP 画像を生成する. その後、LBP 画像の各輝度値からヒストグラムを作成し、SVM への入力データとする.

4.1.3 実験結果

まず Spin-Image 特徴量で、4.1.2 節の通りに実験を行った後、LBP 特徴量で、全く同じデータセットの組み合わせで実験を行った。 SVM のパラメータの最適値及びその時の正答率を表 2 に示す。また、各カテゴリ毎の識別率を表 3、表 4 に示す。

Table 2 Classification results

Table 2 Classification results									
特徴量	C	γ	正答率 [%]	標準偏差 [%]					
Spin-Image	7	-14	79.23	4.51					
パノラマ画像 + LBP	1	1	83.98	4.59					

各カテゴリの識別率に注目すると、Spin-Image を用いた識別では、Coast と Urban の識別率が他 4 つのカテゴリと比べて低くなっている。Coast の識別率の低下は、レーザスキャナの性質上、海の点群が取得できないことに原因があると考えられる。一般に、海岸沿いには木や建物が存在することが多く、今回走行した場所においても家屋等が立ち並ぶ場所が幾つか見受けられた。このよ

Table 3 Confusion matrix(Spin-Image using voting)

	Coast	Forest	ParkingIN	ParkingOut	Residential	Urban
Coast	65.56%	8.48%	0.96%	4.34%	8.43%	12.23%
Forest	10.16%	86.30%	0.18%	0.06%	2.94%	0.36%
ParkingIN	1.67%	1.15%	81.84%	2.06%	7.97%	5.31%
ParkingOut	0.81%	1.36%	0.67%	86.26%	6.53%	4.37%
Residential	1.56%	3.10%	1.52%	9.37%	82.95%	1.50%
Urban	6.29%	0.98%	1.23%	24.05%	3.13%	64.31%

Table 4 Confusion matrix(Panorama image + LBP using voting)

	Coast	Forest	ParkingIN	ParkingOut	Residential	Urban
Coast	67.92%	12.89%	0.12%	8.95%	5.77%	4.34%
Forest	11.00%	88.31%	0%	0.12%	0.57%	0%
ParkingIN	0.28%	0.79%	91.04%	4.50%	3.17%	0.22%
ParkingOut	2.83%	0.28%	3.98%	82.67%	4.99%	5.25%
Residentia1	1.27%	0.61%	0.45%	1.90%	86.95%	8.82%
Urban	0.54%	0%	0.05%	6.96%	10.03%	82.42%

うな、道路を挟んで片方が Coast 以外の特徴を持っているときに 反対側(Coast 側)の特徴が抽出できない場合には、認識が困難 になる. Urban は、識別率が低いだけでなく、ParkingOut への 誤識別率が非常に高くなっている. このことから、単なる交通量 の多さに加え、多発する信号待ちの状態が、ParkingOut に近い 状況を作り出していることが分かる. また、Urban に限らず、例 えば計測中に後続車がいる場合等、長時間車両の点群を計測して しまうと ParkingOut に誤識別してしまうと考えられ、実際に、 車通りのほとんど無い Forest のみ ParkingOut への誤識別が低 くなっている.

パノラマ画像と LBP を用いた識別では Spin-Image と比べる と特に Urban の識別性能が優れていることに加え、全体的に誤識別率が低くなっている. しかし、Coast に関しては Spin-Image の場合と同様の理由で識別率が低下していると考えられる. また、ParkingIn の識別率が高いことや、Urban を Residential と誤識別する確率の高さから、レーザスキャナのスキャンラインに基づいたパノラマ画像を生成したことで屋内駐車場の天井や、市街地のアパートやビル等の高層建築物の特徴を抽出できていることが分かる.

今回構築した全周距離画像データベースに対して、Spin-Image が 79.23%、LBP が 83.98%の識別率を示すことが確認できた. 特に、LBP は屋内環境に限らず屋外環境の識別においても高い識別性能を示すことが分かった. 一方、どちらの特徴量も見えの情報を用いていないため、点群分布が似ている環境の識別は困難であった.

4.2 投票による識別手法

次に、投票を用いた識別手法により識別性能の向上を図った. 具体的には、過去数フレーム分の連続画像に対する多数決で識別 結果を選択する.

投票による実験は、4.1.3節で得られた結果を用いて行う、4.1.3節の実験では、正答率とは別に、各入力データに対する SVM の出力結果を羅列したファイルを出力している。このファイルのデータは、出力の正誤と関係無くカテゴリ順に並んでいるため、先頭からデータを順番に読み出していくことは、6種類の環境を順番に走行する場合と同等と考えられる。

投票数 n を $\{10,20,\ldots,90,100\}$ と変化させていった時の正答率を図 7,表 5 に示す.尚,n=1 は,4.1.3 節と同様の結果である

図7より、投票数が増加するにつれて初めは正答率も高くなっていくが、投票数が50を過ぎた辺りから徐々に低下していく傾向が見られる。全体の正答率が最も高かったのはどちらも投票数が40の時である。それぞれの特徴量について、投票数が40の時のカテゴリ毎の識別率を表6、表7に示す。Spin-ImageがUrbanをParkingOutと識別してしまう確率はあまり下がらなかったが、どちらの特徴量も識別率の上昇が確認できた。このことから、誤識別の多くは不連続なノイズのように発生していることが分かり、投票による識別を行うことでこれらの影響による識

別率の低下を抑えることができた.

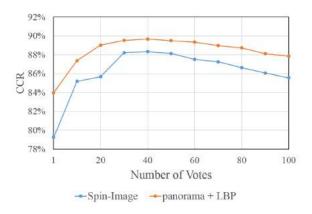


Fig.7 Correct Classification Rate for various numbers of votes

Table 5 Correct Classification Rate

Table 9 Correct Classification Rate									
特徴量	投票数	正答率 [%]	投票数	正答率 [%]					
	10	85.18	60	87.50					
	20	85.68	70	87.25					
Spin-Image	30	88.23	80	86.62					
	40	88.34	90	86.05					
	50	88.14	100	85.55					
	10	87.37	60	89.34					
	20	89.03	70	88.97					
パノラマ画像 + LBP	30	89.52	80	88.73					
	40	89.67	90	88.12					
	50	89.49	100	87.84					

Table 6 Confusion matrix(Spin-Image)

	(1 0)							
/	Coast	Forest	ParkingIN	ParkingOut	Residential	Urban		
Coast	90.16%	2.27%	0%	1.38%	2.19%	3.99%		
Forest	10.04%	89.95%	0%	0%	0.01%	0%		
ParkingIN	0%	3.63%	84.95%	0.69%	5.51%	5.23%		
ParkingOut	0%	0%	3.29%	93.17%	3.17%	0.37%		
Residential	0%	0.19%	0%	4.54%	95.27%	0%		
Urban	1.86%	0%	0.64%	19.40%	3.97%	74.13%		

5 結言

本論文では、全周型レーザスキャナから得られる全周距離画像を用いて屋外環境のデータベースを構築し、その全周距離画像データベースを使用して屋外環境の種別推定実験を行った。まず、レーザスキャナを車載し、Coast、Forest、ParkingIn、ParkingOut、Residential、Urban の 6 種類の屋外環境を計測して全周距離画像データベースを構築した。

次に、全周距離画像データベースから特徴を抽出するための特徴記述子について説明した、特徴記述子には、3次元形状記述子としてよく用いられている Spin-Image と、屋内の環境認識で高い識別性能を示す LBP を用いた.

最後に、これら二つの特徴量を用いて、屋外環境の識別実験を行った。Spin-Image は一枚の画像内に環境の点群分布を全て記述する手法、LBP は全周距離画像から生成されたパノラマ画像の特徴を記述する手法であり、点群分布の似た環境の識別は困難であったが、全体では Spin-Image が 79.23%、LBP が 83.98%の識別性能を示した。さらに、得られた結果を基にして、投票による多数決で識別結果を選択する手法を提案した。投票数を増やす

Table 7 Confusion matrix(Panorama image + LBP)

/	Coast	Forest	ParkingIN	ParkingOut	Residential	Urban
Coast	82.99%	5.20%	0%	7.15%	1.50%	3.16%
Forest	8.64%	90.87%	0%	0.48%	0%	0%
ParkingIN	0%	3.31%	93.95%	2.16%	0.59%	0%
ParkingOut	0.96%	0%	5.25%	89.16%	2.44%	2.19%
Residentia1	0.45%	0%	0%	2.70%	93.23%	3.63%
Urban	0%	0%	0%	3.84%	8.76%	87.39%

ことで誤識別率が低下し、過去 40 フレームの画像を投票に用いた場合には、Spin-Image が 88.34%, LBP が 89.67%まで識別率が向上した.

今後は、環境に応じて特徴量を選択、あるいはそれぞれの特徴量を組み合わせる等して識別性能のさらなる向上を図る.また、さらに多くのカテゴリを含む屋外環境識別も行えるようにしていく必要がある.

謝 辞本研究は文部科学省科学研究費補助金基盤研究 (A) (課題番号 26249029) の支援を受けた.

参考文献

- C. Stachniss, O. M. Mozos, and W. Burgard, "Speedingup multi-robot exploration by considering semantic place information", in Proc. of the IEEE International Conference on Robotics and Automation (ICRA), Orlando, FL, USA, 2006, pp.1692-1697.
- [2] A. Torralba, K. P. Murphy, W. T. Freeman, and M. A. Rubin, "Context-based vision system for place and object recognition", in International Conference on Computer Vision (ICCV), Nice, France, 2003.
- [3] O. M. Mozos, C. Stachniss, and W. Burgard, "Supervised learning of places from range data using AdaBoost", in IEEE International Conference on Robotics and Automation (ICRA), Barcelona, Spain, 2005.
- [4] H. Zender, O. M. Mozos, P. Jensfelt, G.-J. M. Kruijff, and W. Burgard, "Conceptual spatial representations for indoor mobile robots", Robotics and Autonomous Systems, Vol. 56, No. 6, pp.493-502, June 2008.
- [5] Oscar Martinez Mozos and Wolfram Burgard: "Supervised Learning of Topological Maps using Semantic Information Extracted from Range Data", IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006.
- [6] Jianxin Wu, James M. Rehg, "CENTRIST: A Visual Descriptor for Scene Categorization", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.33, No.8, pp.1489-1501, 2011.
- [7] Andrzej Pronobis, Oscar Martinez Mozos, Barbara Caputo, and Patric Jensfelt: "Multi-modal Semantic Place Classification", International Journal of Robotics Research, 29, 2-3, pp.298-320, 2010.
- [8] A. Johnson, "Spin-Images: A Representation for 3-D Surface Matching", Ph.D. Dissertation, The Robotics Institute, Carnegie Mellon University, 1997.
- [9] T. Ojala, M. Pietikäinen and D. Harwood, "A Comparative Study of Texture Measures with Classification based on Feature Distributions", Pattern Recognition, 29, 1, pp.51-59, 1996.
- [10] Oscar Martinez Mozos, Hitoshi Mizutani, Hojung Jung, Ryo Kurazume, Tsutomu Hasegawa, "Categorization of Indoor Places by Combining Local Binary Pattern Histograms of Range and Reflectance Data from Laser Range Finders", Advanced Robotics, Vol.27, No.18, pp.1455-1464, 2013.
- [11] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: a stepwise procedure for building and training a neural network", in Neurocomputing: Algorithms, Architectures and Applications, J. Fogelman, Ed. Springer-Verlag, 1990.
- [12] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines", ACM Transactions on Intelligent Systems and Technology, vol. 2, pp. 27:1-27:27, 2011, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/.
- [13] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "Apractical guide to support vector classification", http://www.csie.ntu.edu.tw/~ cjlin/papers/guide/guide.pdf, 2010.