

レーザ距離画像と反射率画像を用いた屋内環境のカテゴリ識別

水谷 仁* マルティネス モゾス オスカル** 倉爪 亮 岩下 友美** 長谷川 勉**

* 九州大学大学院システム情報科学府 ** 九州大学大学院システム情報科学府

E-mail: {mizutani,omozos}@irvs.ait.kyushu-u.ac.jp, {kurazume,yumi,hasegawa}@ait.kyushu-u.ac.jp

1. はじめに

日常生活環境で人間に対して様々なサービスを行うロボットにとって、ロボット周囲の空間の識別は重要な機能である。空間の識別は、例えば高次の概念的表現も含め空間情報を獲得するセマンティックマッピング [1-6] や、物体認識のための事前知識としての利用 [7, 8], あるいは空間毎に実行可能な行動計画の立案 [9-11] など、様々な応用が考えられる。

空間の識別では、ロボットはセンサ情報を基に、それぞれの場所に対応したラベルを割り付ける。ここで、オフィスや台所、研究室など、人間がそれぞれの場所に対して与えるものと同様なラベルを使用することで、人とロボットの円滑なコミュニケーションが実現できる [12, 13]。

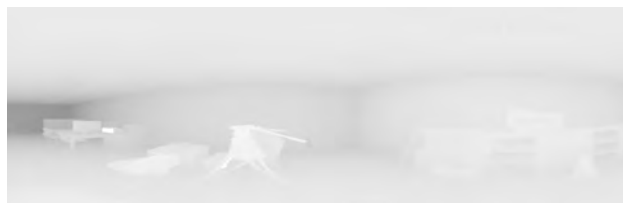
本稿では、レーザレンジファインダを用いた新たな屋内環境の識別手法を提案する。レーザレンジファインダは、センサから能動的に発射されたレーザ光が、対象物表面に反射して再びセンサに戻るまでの時間を計測することで、対象物までの距離を得るものである。レーザレンジファインダでは、レーザ光の反射時間とともにレーザ光の強度（反射率）を計測することができる。そのため、距離データおよび反射率データが原理的に厳密に位置合わせされている。一例として、距離データおよび反射率データを濃淡値として 2 次元表示した距離画像および反射率画像を図 1 に示す。

提案手法では、まず距離データおよび反射率データを、それぞれ局所二値パターン (Local Binary Pattern, LBP) [6] により近傍計測点間の関係を表現した特徴量画像に変換する。そして、それら特徴量画像からそれぞれ格子状に異なる領域サイズで分割された空間ピラミッド [12] と呼ばれる画像列を生成し、各分割領域から局所ヒストグラムを抽出・結合する。その後、得られた距離データおよび反射率データのヒストグラムを統合し、特徴ベクトルとする。最後に、サポートベクトルマシンにより、撮影された空間の種類を認識する。ただし空間ピラミッドでは、分割領域が小さくなるにつれて特徴次元が指数関数的に増加する問題がある。そこで Ojala ら [11] によって提案された評価値 U を参考に、局所二値パターンのサブセットである短縮局所二値パターン (Short Local Binary Pattern, SLBP) を導入し、特徴次元の削減を行う。

2. 提案手法

2.1 局所二値パターンを用いた特徴量画像生成

局所二値パターン (Local Binary Patterns, LBP) 演算子 [6, 11] は、濃淡画像のテクスチャ分析と分類を目的に Ojala らにより提案された。この演算子は、参照



(a) 距離画像



(b) 反射率画像

図 1 レーザレンジファインダによる計測画像

画素とその近傍画素の濃淡値の大小関係をベクトル形式で表す。

本稿では、LBP を距離データおよび反射率データの各計測点 i に対して以下のように適用する。まず、各計測点 i に対して、その計測値 (距離値および反射率) $p(i)$ を周囲 8 近傍の計測値 $p_1(i), \dots, p_8(i) \in \mathcal{N}(p(i))$ と比較する。その際、それぞれの近傍点ごとに二値 $\{0, 1\}$ を以下の式で決定する。

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

ここで、周囲 8 近傍の計測値 $p_j(i)$ ($j = 1, \dots, 8$) の配置は図 2 に示す通りである。その後、近傍点に対して得られた二値 $b_j(i)$ を時計回りに $\{b_0(i), \dots, b_8(i)\}$ の順番で並べ、2 進数を生成する。そして、得られた 2 進数を図 3 に示すように $[0, 1, \dots, 255]$ の 10 進数のラベル $v^{\text{LBP}}(i)$ に変換し、LBP 値を得る。これより、距離データおよび反射率データからそれぞれ LBP に変換された特徴量画像 T_{depth} および $T_{\text{reflectance}}$ が生成される。

p_6	p_7	p_8
p_5		p_1
p_4	p_3	p_2

図 2 $p_j(i)$ ($j = 1, \dots, 8$) の配置

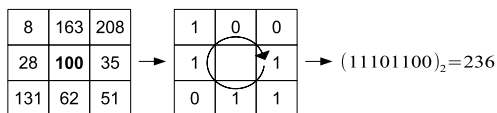


図3 参照点に対するLBPの計算例．この例では236がセットされる．

2.2 空間ピラミッドを用いた特徴ベクトルの生成

2.1 項の方法により得られた特徴量画像 T_{depth} , $T_{\text{reflectance}}$ を以下の方法を用いてそれぞれ特徴ベクトル $\mathbf{x}_{\text{depth}}$, $\mathbf{x}_{\text{reflectance}}$ に変換する．

まず特徴量画像 T (T_{depth} および $T_{\text{reflectance}}$) に対して、図4に示すような空間ピラミッド [12] と呼ばれる階層的な画像列を生成する．空間ピラミッドでは、各レベル l において特徴量画像 T を格子状に均一に $d = 2^l \times 2^l$ 個の領域に分割する．次に各分割領域において局所ヒストグラム H_d^l を計算する．ここで局所ヒストグラム H_d^l のそれぞれのピンは、レベル l の p 番目の画像において対応する画素値の画素数である．最後に、各レベルの局所ヒストグラムを結合し、特徴量画像 T に対する特徴ベクトル \mathbf{x}_T を得る．

$$\mathbf{x}_T = \{ \{H_1^0\}, \{H_1^1, \dots, H_{2^1 \times 2^1}^1\}, \dots, \{H_1^L, \dots, H_{2^L \times 2^L}^L\} \} \quad (2)$$

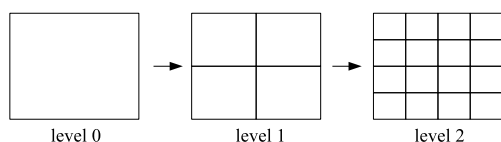


図4 3層空間ピラミッド

2.3 距離情報と反射率情報のモダリティ統合

第2.2項の方法により、距離データおよび反射率データに対する特徴ベクトル $\mathbf{x}_{\text{depth}}$ および $\mathbf{x}_{\text{reflectance}}$ がそれぞれ得られる．距離情報と反射率情報を統合するため、得られた特徴ベクトル $\mathbf{x}_{\text{depth}}$ と $\mathbf{x}_{\text{reflectance}}$ を結合し、最終的な特徴ベクトル \mathbf{x} とする．この特徴ベクトル \mathbf{x} が識別器の入力として用いられる．

$$\mathbf{x} = \{ \mathbf{x}_{\text{depth}}, \mathbf{x}_{\text{reflectance}} \} \quad (3)$$

2.4 短縮局所二値パターンを用いた特徴次元の削減

空間ピラミッドでは、レベル L が高くなるにつれて特徴次元が指数関数的に増加する問題がある．そこでOjalaら [11] によって提案された評価値 U を参考に、局所二値パターンのサブセットである短縮局所二値パターン (Short Local Binary Pattern, SLBP) を導入し、特徴次元の削減を行う．

SLBP では、式 (1) を用いて得られる二値 $b_j(i)$ の0, 1間の遷移回数 $U(i)$ に着目し、空間を構成する単調なLBPを選択する．

$$U(i) = |b_1(i) - b_8(i)| + \sum_{j=1}^7 |b_j(i) - b_{j+1}(i)| \quad (4)$$

例えば 00000000_2 , 11111111_2 , 01010101_2 では、0, 1間の遷移回数 U はそれぞれ0, 0, 8となる．この0, 1間の遷移回数 U に閾値 θ を設けることで、SLBPの値 $v^{\text{SLBP}}(i)$ を以下のように決定する．

$$v^{\text{SLBP}}(i) = \begin{cases} v^{\text{LBP}}(i), & U(i) \leq \theta \\ \text{nan}, & \text{otherwise} \end{cases} \quad (5)$$

ただし、 $\theta = 8$ のときは通常のLBPと等しく、最終的な記述子はCENTRIST [5] と等価である．ここで、短縮局所二値パターンを用いて得られた特徴量画像を局所ヒストグラムに変換する際、nanは考慮に入れない．それ以外の特徴ベクトルの計算方法は第2.2, 2.3項で示した通常のLBPと同様の方法で行う． $\theta = 2, 4$ および6のSLBPでは、256種類のLBPのうちそれぞれ58種類, 198種類, 254種類が選択される．その結果、LBPに比べて特徴ベクトルの次元は22.7%, 77.3%, 99.2%に削減される．

2.5 識別方法

距離データおよび反射率データの対から得られた特徴ベクトルは、サポートベクトルマシン (Support Vector Machines, SVMs) [13, 14] により識別される．

M 個の m 次元教師データ $\mathbf{x}_i \in \mathbb{R}^m (i = 1, \dots, M)$ に対し、それぞれクラスラベル $y_i \in \{1, -1\}$ (あるカテゴリに属するか、属さないか) が付加されSVMに入力される．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．

$$\min_{\mathbf{w}, b, \xi} C \sum_{i=1}^M \xi_i + \frac{1}{2} \|\mathbf{w}\|^2 \quad (6)$$

$$\text{subject to } y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \quad (7)$$

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．実験では次のラジアル基底関数 (Radial Basis Function) カーネルを用いた．

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \quad \gamma > 0 \quad (8)$$

ここで、 γ は分布の半径を制御するパラメータである．

多クラス問題への拡張は一対一方式 [15] を使い、実験ではLIBSVMライブラリ [16] を用いた．パラメータ C および γ は、それぞれ $C \in [2^{-5}, 2^{-2}, \dots, 2^{16}]$ および $\gamma \in [2^{-12}, 2^{-9}, \dots, 2^3]$ の範囲で交差検定を用いた総当たり探索により決定した [17] ．

3. 識別実験

3.1 データセット

全周囲の距離データおよび反射率データを取得するため、2次元レーザレンジファインダを高さ131cmの回転テーブル上に固定し、テーブルを回転させながらスリット状レーザにより計測を行った．その後、全周データを 10° ずつ回転させながら、 625×760 のサイズで切り出し、データセットを構築した．

対象とした空間のカテゴリは屋内環境である“corridor”, “kitchen”, “laboratory (lab.)”, “office”および“study room (study)”の5つである。これらカテゴリにおいて、それぞれいくつかの異なる空間(部屋)で距離データおよび反射率データを取得した。例えば、表1のカテゴリである“laboratory”では4つの異なる研究室を含んでいる。また、同一の空間において、いくつかの異なる地点で計測を行った。表1に識別実験で用いた空間毎のデータ対(距離データおよび反射率データ)の数を示す。一例として、各空間における距離画像および反射率画像を図5に示す。

表1 空間毎の計測データ対の数(合計 1110)

category	place	number of image sets
corridor	corridor 1	74
	corridor 2	74
	corridor 3	74
	total	222
kitchen	kitchen 1	37
	kitchen 2	74
	kitchen 3	74
	kitchen 4	37
	total	222
laboratory	laboratory 1	74
	laboratory 2	74
	laboratory 3	37
	laboratory 4	37
	total	222
study room	study room 1	74
	study room 2	74
	study room 3	37
	study room 4	37
	total	222
office	office 1	74
	office 2	74
	office 3	74
	total	222

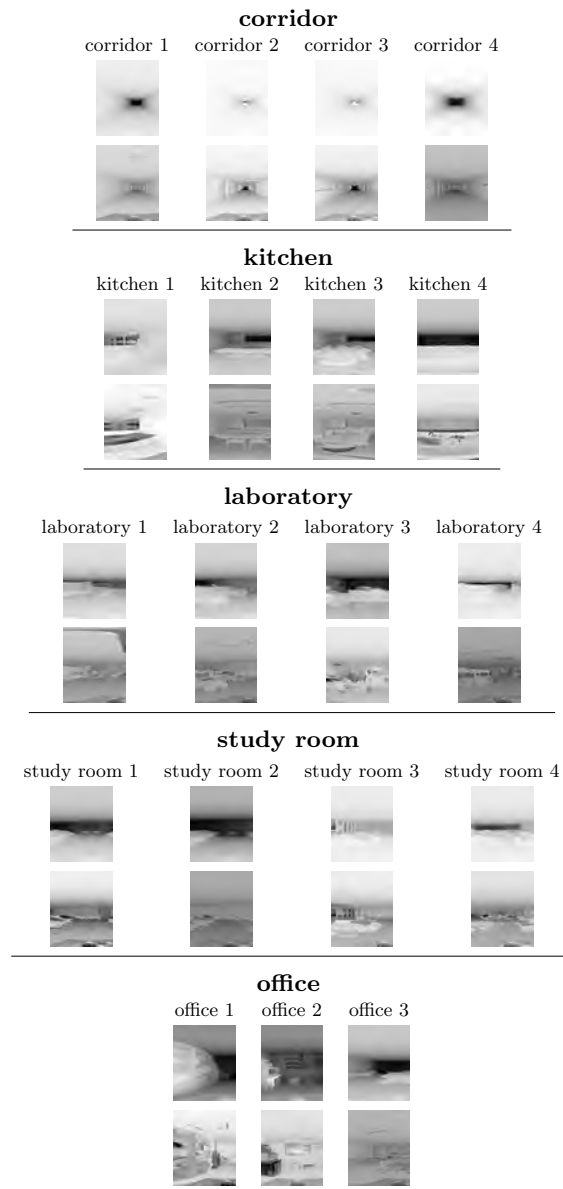


図5 距離画像(上)および反射率画像(下)の例

3.2 実験結果

実験では、教師データとテストデータの10個の組を用いて、提案手法を評価した。これら教師データとテストデータの生成方法は以下の通りである。まず各カテゴリから一つの空間をランダムに選択し、テストデータとする。そして、残りの空間は教師データとして用いる。これにより、テストデータは常に教師データには含まれない空間が選択される。また、選択された空間の組が既にテストデータとして存在する場合には、それを破棄し、異なるテストデータが得られるまで選択処理を続ける。これを10回繰り返し、教師データとテストデータの10個の組を生成した。

提案手法の評価方法は以下の通りである。教師データを用いて識別器の学習を行い、学習後テストデータの認識率を求める。これを教師データとテストデータの10個の組に対し行い、その平均を取り最終的な認識率とした。空間ピラミッドのレベル L およびSLBPの閾値 θ は、それぞれ[12],[18]で高い認識率が報告さ

れている $L = 2$ および $\theta = 4$ を用いた。

実験ではまずマルチモダリティ(距離情報および反射率情報の対)の性能評価のため、モダリティをそれぞれ単体で用いた場合との認識率の比較を行った。表2はその結果であり、最も高い認識率96.00%を示したマルチモダリティの有効性を確認した。また反射率情報単体のときの標準偏差が14.01%であるのに対し、マルチモダリティの標準偏差は6.19%であり、モダリティの統合により安定した認識率が得られた。

次に、最も高い認識率を示したマルチモダリティにおいて各カテゴリの認識率を検証した。表3はマルチモダリティにおける混同行列であり、行方向、列方向に並んだラベルはそれぞれ予測カテゴリ、正答カテゴリを表す。表3より、“kitchen”は他のカテゴリに比べて認識率が低く、識別が難しいカテゴリであることがわかる。特に、“kitchen”と“office”を混同するケースが最も多いことを確認した。

表 2 $L = 2, \theta = 4$ のときのモダリティ間の性能比較 (%)

depth	reflectance	depth+reflectance
90.76±5.82	88.70±14.01	96.00±6.19

表 3 $L = 2, \theta = 4$ のときのマルチモダリティにおける混同行列 (%) (認識率 96.00%)

%	corridor	kitchen	lab	study	office
corridor	98.11	0.54	1.35	0.00	0.00
kitchen	2.43	85.14	0.00	0.00	12.43
lab	2.97	0.00	97.03	0.00	0.00
study	0.00	0.27	0.00	99.73	0.00
office	0.00	0.00	0.00	0.00	100.00

4. 結論と今後の課題

本稿では、レーザレンジファインダから得られた距離データおよび反射率データを用いて、空間を識別する手法を提案した。提案手法では、まず距離データおよび反射率データから SLBP を用いてそれぞれ特徴量画像を生成する。そして特徴量画像から、格子状に異なる領域サイズで分割された空間ピラミッドと呼ばれる画像列を生成し、各分割領域から局所ヒストグラムを抽出・結合する。その後、距離データおよび反射率データに対するヒストグラムを統合し、特徴ベクトルを生成する。最後に、特徴ベクトルを SVMs により識別した。識別実験の結果、提案手法の高い認識率とマルチモダリティの有効性が確認された。

今後は、より適当な記述子について検討するとともに、データセットのカテゴリ数とデータ数を増やすことを目標とする。

謝辞

本研究の一部は、日本学術振興会外国人特別研究員事業、文部科学省科学研究費補助金基盤研究(B)(課題番号 23360115)の支援を受けた。また実験にご協力いただいた九州大学内田誠一教授、谷口倫一郎教授、諸岡健一准教授に感謝いたします。

参考文献

- [1] O. M. Mozos, C. Stachniss and W. Burgard: "Supervised learning of places from range data using AdaBoost", IEEE International Conference on Robotics and Automation(ICRA) (2005). Barcelona, Spain.
- [2] O. M. Mozos and W. Burgard: "Supervised learning of topological maps using semantic information extracted from range data", IEEE/RSJ International Conference on Intelligent Robots and Systems (ICRA) (2006). Beijing, China.
- [3] E. Brunskill, T. Kollar and N. Roy: "Topological mapping using spectral clustering and classification", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (October 2007). San Diego.
- [4] L. Shi, S. Kodagoda and G. Dissanayake: "Laser range data based semantic labeling of places", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (October 2010). Taipei, Taiwan.
- [5] J. Wu, H. I. Christensen and J. M. Rehg: "Visual place categorization: Problem, dataset, and algorithm", Proc of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2009). St. Louis, MO, USA.
- [6] T. Ojala, M. Pietikäinen and D. Harwood: "A comparative study of texture measures with classification based on featured distributions", Pattern Recognition, **29**, 1, pp. 51–59 (1996).
- [7] A. Ranganathan: "Pliss: Detecting and labeling places using online change-point detection", Robotics: Science and Systems VI (2010).
- [8] A. Torralba, K. P. Murphy, W. T. Freeman and M. A. Rubin: "Context-based vision system for place and object recognition", International Conference on Computer Vision (ICCV) (2003). Nice, France.
- [9] A. Rottmann, O. M. Mozos, C. Stachniss and W. Burgard: "Semantic place classification of indoor environments with mobile robots using boosting", National Conference on Artificial Intelligence (AAAI) (2005). Pittsburgh, PA, USA.
- [10] A. Pronobis, O. M. Mozos, B. Caputo and P. Jensfelt: "Multi-modal semantic place classification", International Journal of Robotics Research, **29**, 2–3, pp. 298–320 (2010). February-March.
- [11] T. Ojala, M. Pietikäinen and T. Mäenpää: "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Trans. on Pattern Analysis and Machine Intelligence, **24**, 7, pp. 971–987 (2002).
- [12] S. Lazebnik, C. Schmid and J. Ponce: "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2006).
- [13] C. Cortes and V. Vapnik: "Support-vector network", Machine Learning, **20**, pp. 273–297 (1995).
- [14] C. M. Bishop: "Pattern Recognition and Machine Learning", Springer (2006).
- [15] S. Knerr, L. Personnaz and G. Dreyfus: "Single-layer learning revisited: a stepwise procedure for building and training a neural network", Neurocomputing: Algorithms, Architectures and Applications (1990). J. Fogelman, Ed., Springer-Verlag.
- [16] C.-C. Chang and C.-J. Lin: "Libsvm: A library for support vector machines", ACM Transactions on Intelligent Systems and Technology, **2**, pp. 27:1–27:27 (2011).
- [17] C.-W. Hsu, C.-C. Chang and C.-J. Lin: "A practical guide to support vector classification" (2010).
- [18] O. M. Mozos, H. Mizutani, R. Kurazume and T. Hasegawa: "Categorization of indoor places using the kinect sensor", Sensors, **12**, pp. 6695–6711 (May 2012).