

距離・画像情報を統合した ロボットのための屋内環境のカテゴリ識別

マルティネス モゾス オスカル^{†2} 水谷 仁^{†1}
倉爪 亮^{†2} 岩下 友美^{†2} 長谷川 勉^{†2}

人と共生するサービスロボットにとって、ロボット周囲の空間の認識は高度で重要な機能の一つである。本論文では、Kinect など RGB-D カメラから得られる RGB 画像と距離画像を用いて、オフィス、台所など空間の認識を行う手法を提案する。提案手法では、まず RGB 画像を濃淡画像に変換し、その濃淡画像および同時刻に得られた距離画像を、それぞれ局所二値パターンと短縮局所二値パターン、および局所ヒストグラムを用いて、近傍画素間の関係を表現した特徴ベクトルに変換する。その後、得られた濃淡および距離情報の特徴ベクトルを結合し、サポートベクトルマシンなどの識別器により、空間の種類を認識する。5つの異なるカテゴリに属する部屋に対して認識実験を行った結果、提案手法により高い認識性能が得られることを確認した。

Indoor place categorization for service robots using camera and depth images

OSCAR MARTINEZ MOZOS^{†2} HITOSHI MIZUTANI^{†1}
RYO KURAZUME^{†2} YUMI IWASHITA^{†2}
and TSUTOMU HASEGAWA^{†2}

An important capability for service robots working with humans in indoor environments is their ability to categorize the different places where they are located. In this paper we present an approach to categorize different areas in indoor environments using an RGB-D sensor like the Kinect camera. First, RGB images are transformed into grey scale images. Then, grey scale and depth images are transformed into histograms of local features that incorporate neighboring relations by applying local binary patterns and also a short version of this pattern. The feature vectors corresponding to grey scale and depth images are combined and categorized into different places using supervised classifiers like for example support vector machines. We apply this method to distinguish five different place categories and obtain high recognition rates.

1. はじめに

人と共生するサービスロボットにとって、ロボット周囲の空間の認識 (*place recognition*) は高度かつ重要な機能の一つである。空間の認識では、ロボットには正確な位置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。ロボットがこの能力を獲得できれば、その情報は場面に応じた適切な行動選択^{1),2)} や、人との自然なインタラクションにおける事前情報またはコンテキストとして用いることができる^{3),4)}。例えば、ロボット自身が現在台所にいると認識できれば、調理や食事に関する指示が多く出されると予測でき、また物体はその多くが空間種別と強く結びついていることから、物体認識においても空間種別から物体種別や用途、用法を限定できる。また、識別された空間の情報は、セマンティックマップ^{5),6)} の構築や空間の概念的表現^{7),8)} などにも利用できる。さらに空間の認識において、ロボットは自身が置かれた、または見ている空間に対し、それぞれラベルを割り当てることができる。そのとき、“office”や“laboratory”のように、人が一般的に用いる名前をラベルとして用いれば、人とロボットの直感的なコミュニケーションが実現できる^{8),9)}。

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 (*instance classification*) と空間の識別 (*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 965 号室など) を、センサ情報から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 956 号室は、他の類似の部屋とともに、より一般的な概念である “オフィス” として識別される。本論文では、後者の問題を取り扱う。

空間の識別では、これまでに様々なセンサが用いられている。例えば、レーザレンジファインダ^{5),10),11)} や画像センサ^{3),12),13)}、あるいはこれらの2つのセンサを統合した手法も提案されている^{14),15)}。一方、ごく最近では様々なロボットアプリケーションにおいて、Kinect センサ¹⁶⁾ が盛んに用いられつつある。このセンサは低価格で、かつ高速に面状の距離情報を得ることができ、我々もこのセンサを用いて空間の識別を行うことを試みる。

†1 九州大学大学院システム情報科学府
Kyushu University, Graduate School of Information Science and Electrical Engineering
†2 九州大学大学院システム情報科学研究院
Kyushu University, Faculty of Information Science and Electrical Engineering

本論文では, Kinect センサにより集められた RGB および距離画像を用いて, 空間の識別を行う手法を提案する. 提案手法では, まず RGB 画像を濃淡画像に変換し, その濃淡画像および同時刻に得られた距離画像を, それぞれ局所二値パターン¹⁷⁾, 短縮局所二値パターン, および局所ヒストグラムを用いて, 近傍画素間の関係を表現した特徴ベクトルに変換する. その後, 得られた濃淡および距離情報の特徴ベクトルを結合し, 両者を統合した特徴ベクトルとする. この特徴ベクトルを用いて, サポートベクトルマシンやランダムフォレストなどの識別器により, 空間の種別を認識する.

第 2 章から第 4 章では, それぞれ関連手法, 濃淡および距離画像を用いた空間の識別, 識別実験について述べる.

2. 関連手法

近年, 移動ロボットによる空間の識別技術が向上し, ロボットは様々なセンサにより典型的な屋内空間が識別できるようになった. 例えば, レーザレンジファインダにより得られる 2D 画像は, 屋内環境で異なる空間を識別するのに用いることができる. Mozos らは 2D 画像をいくつかの幾何学的特徴を含むベクトルに変換し, それらを AdaBoost を用いて分類した^{5), 18)}. さらに, Friedman らは, ボロノイ図の各点に条件付き確率場を適用した. 各点はそれぞれの地点でレーザ計測により得られる空間の幾何学的特徴も表している. さらに, Brunskill らは Spectral Clustering を用いて, 屋内環境の 2D 下位マップを作成した¹¹⁾. その下位マップは局所的な 2D スキャンから得られる特徴ベクトルを用いて分類される. また, Shi らは全体をスキャンせず, それぞれの独立したスキャンを意味的ラベルへ分類する手法を提案している¹⁹⁾.

一方, 視覚センサも空間の認識に多く用いられてきた. 例えば, Wu と Rehg による手法では, 家の中の複数の空間の識別に CENTRIST 記述子を用いた¹²⁾. この記述子は, Ojala らが提案している局所二値パターンと同様である¹⁷⁾. さらに, Ranganathan は空間の識別のための PLIS システムを提案している¹³⁾. この手法では変化点を見つけることにより, 一続きの画像列を分割する. そのあと, 分割された画像はそれぞれ視覚カテゴリに分類される. Torralba らはウェアラブルカメラから得られる画像の局所特徴量と大域特徴量を用いて, 隠れマルコフモデルにより屋内・屋外環境の特定の空間の認識および空間の識別を行った³⁾.

空間を識別するために, 異なる種類のセンサを組み合わせる試みも行われてきた. Rottmann らは 2D レーザスキャンと視覚による物体検出を組み合わせ, ブースティングを用いて屋内

b_6	b_7	b_8
b_5		b_1
b_4	b_3	b_2

図 1 $b_j(i)$ ($j = 1, \dots, 8$) の配置

環境の空間を識別した¹⁴⁾. Pronobis らは複数の視覚とレーザによる手がかりを組み合わせ, サポートベクトルマシンを用いて, 屋内環境の空間を識別した¹⁵⁾.

3. 濃淡および距離画像を用いた空間の識別

3.1 局所二値パターンを用いた濃淡画像からの特徴量画像生成

局所二値パターン (Local Binary Patterns, LBP) 演算子¹⁷⁾ 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. 本稿ではそれを濃淡画像と距離画像に適用し, 典型的な屋内空間の識別を行う.

濃淡画像を用いた空間の識別では, まず Kinect センサから得られる RGB 画像を濃淡画像に変換し, 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 $p_1(i), \dots, p_8(i) \in \mathcal{N}(p(i))$ 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.

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

ただし, 周囲 8 近傍の画素値 $b_j(i)$ ($j = 1, \dots, 8$) の配置は図 1 の通りである. 得られた周囲画素の二値 $b_j(i)$ は時計回りに $\{b_0(i), \dots, b_8(i)\}$ の順番で並べられ, 図 2 に示すように 0 から 255 までのラベル $v^{\text{LBP}}(i)$ に変換される. 注目画素の値は, このラベル $v^{\text{LBP}}(i)$ により置き換えられ, LBP に変換された特徴量画像 T_{gray} が生成される.

3.2 局所二値パターンを用いた距離画像からの特徴量画像生成

同様に Kinect センサから得られる距離画像に, LBP を適用して特徴量画像を得ることができる. しかし, Kinect センサでは, 計測不能な個所および計測範囲外の距離値に nan

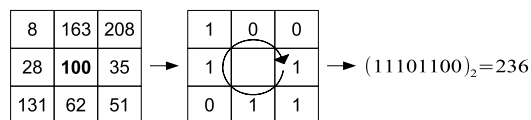


図 2 中心画素に対する LBP の計算例. この例では 236 がセットされる

がセットされるため, 距離画像に対して単純に LBP の計算ができない. そこで, 距離値が nan である画素, あるいは前章の LBP の計算で周囲画素に nan が含まれる場合には, LBP の出力を 256 とした. 従って, LBP を用いて得られた特徴量画像 T_{depth} の画素値 $v^{\text{LBP}+}(i)$ は以下のように計算される.

$$v^{\text{LBP}+}(i) = \begin{cases} 256, & \delta_{\text{nan}}(p(i)) \vee \exists \delta_{\text{nan}}(\mathcal{N}(p(i))) \\ v^{\text{LBP}}(i), & \text{otherwise} \end{cases} \quad (2)$$

ここで $\delta_{\text{nan}}(\cdot)$ は, もし距離値に nan が設定されていた場合には 1 を, それ以外は 0 を返す関数であり, $v^{\text{LBP}}(i)$ は LBP により得られた特徴量である (図 2).

濃淡画像および同時刻に得られた距離画像と, それらの局所二値パターンの例を図 3 に示す.

3.3 局所ヒストグラムおよび特徴量ベクトルの生成

3.1, 3.2 項の方法で得られた特徴量画像 T_{gray} , T_{depth} を以下の方法でそれぞれ特徴ベクトル \mathbf{x}_{gray} , $\mathbf{x}_{\text{depth}}$ に変換する.

まず, 特徴量画像 T (T_{gray} および T_{depth}) に対して空間ピラミッド²⁰⁾ を生成する. ただし, 解像度のレベル l を図 4 に示すように定義する. すなわち, 初期画像を $p = 2^l \times 2^l$ 個の領域に分割したものをレベル l の画像とする. 次に分割したそれぞれの領域で局所ヒストグラム H_p^l を計算する. ここで局所ヒストグラム H_p^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, 特徴ベクトル \mathbf{x}_T は, 各レベルの局所ヒストグラムを結合したものと表される.

$$\mathbf{x}_T = \{\{H_1^0\}, \{H_1^1, \dots, H_{2^1 \times 2^1}^1\}, \dots, \{H_1^L, \dots, H_{2^L \times 2^L}^L\}\} \quad (3)$$

3.4 濃淡情報と距離情報の統合

濃淡情報と距離情報を統合するため, 前項の方法で得られた特徴ベクトル \mathbf{x}_{gray} と $\mathbf{x}_{\text{depth}}$ 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

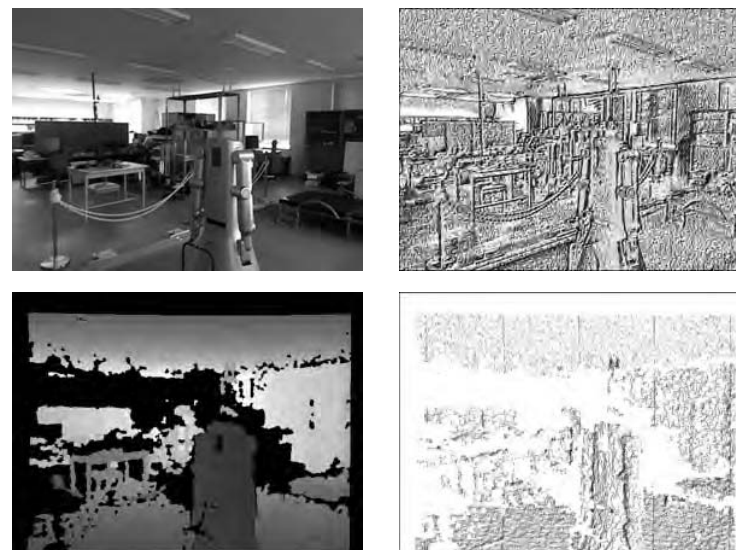


図 3 左列: "laboratory" の濃淡画像と距離画像, 右列: それぞれの局所二値パターン画像

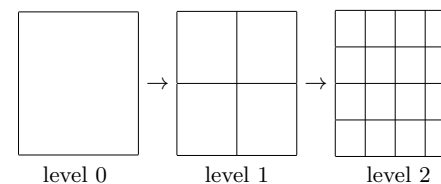


図 4 3 層空間ピラミッド

いられる.

$$\mathbf{x} = \{\mathbf{x}_{\text{gray}}, \mathbf{x}_{\text{depth}}\} \quad (4)$$

3.5 短縮局所二値パターンを用いた特徴ベクトルの次元削減

LBP の局所ヒストグラムを用いた手法では, 特徴ベクトルの次元が, 空間ピラミッドの

解像度のレベルが高くなるにつれて指数関数的に増加する問題がある。

そこで本論文では特徴ベクトルの次元削減を目的に、Ojala ら²¹⁾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) を新たに用いる。SLBP では、特定の代表的な LBP のみを用いることで、次元削減をはかる。

特定の代表的な LBP を見つけるため、式 (1) を用いて得られる二値 $b_j(i)$ の 0, 1 間の遷移の回数 $U(i)$ に注目する。

$$U(i) = |b_1(i) - b_8(i)| + \sum_{j=1}^7 |b_j(i) - b_{j+1}(i)| \quad (5)$$

例えば 00000000_2 , 11111111_2 , 01010101_2 では、0, 1 間の遷移の回数 U はそれぞれ 0, 0, 8 となる。これより、濃淡画像に対し、SLBP の値 $v^{\text{SLBP}}(i)$ は以下のように表される。

$$v^{\text{SLBP}}(i) = \begin{cases} v^{\text{LBP}}(i), & U(i) \leq U_{\text{Th}} \\ \text{nan}, & \text{otherwise} \end{cases} \quad (6)$$

同様に、距離画像に対しては、SLBP の値 $v^{\text{SLBP}^+}(i)$ は以下のように表される。

$$v^{\text{SLBP}^+}(i) = \begin{cases} v^{\text{LBP}}(i), & U(i) \leq U_{\text{Th}} \\ 256, & \delta_{\text{nan}}(p(i)) \vee \exists \delta_{\text{nan}}(\mathcal{N}(p(i))) \\ \text{nan}, & \text{otherwise} \end{cases} \quad (7)$$

ここで U_{Th} は $U(i)$ に対する閾値である。ただし、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 3.3, 3.4 項で示した通常の LBP と同様の方法で行う。第 4 章では、 $U_{\text{Th}} = 2$ および 4 の SLBP の実験結果を示す。この識別実験で用いた $U_{\text{Th}} = 2$ および 4 の SLBP では、256 種類の LBP のうち、それぞれ 58 種類、198 種類のみが用いられ、その結果特徴ベクトルの次元も 22.8%、77.4% に大幅に短縮できた。

3.6 識別方法

各画像を特徴ベクトルに変換し、Support Vector Machine (SVM)^{22),23)} を用いて識別を行う。SVM には M 対の濃淡および距離画像の組から得られた M 個の m 次元特徴ベクトル

ル $\mathbf{x}_i \in \mathcal{R}^m (i = 1, \dots, M)$ とともに、ラベル $y_i \in \{1, -1\}$ (すなわちそのカテゴリに属するか、属さないか) が入力される。SVM の目的は 2 クラス $\{1, -1\}$ の学習データとの距離が最大となる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。

$$\min_{\mathbf{w}, b, \xi} C \sum_{i=1}^M \xi_i + \frac{1}{2} \|\mathbf{w}\|^2 \quad (8)$$

$$\text{subject to } y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \quad (9)$$

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。

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \quad \gamma > 0 \quad (10)$$

ここで、 γ は分布の半径を制御するパラメータである。

多クラス問題の識別は一方式²⁴⁾で行い、実験では LIBSVM ライブラリ²⁵⁾を用いた。パラメータ C および γ は、 $C \in [2^{-5}, 2^{-2}, \dots, 2^{16}]$ および $\gamma \in [2^{-12}, 2^{-9}, \dots, 2^3]$ の範囲でクロスバリデーションを用いた総当たり探索により決定した²⁶⁾。

4. 識別実験

4.1 データセット

提案手法の性能評価のために、屋内環境で実験を行った。実験では、まず移動台車の上部、高さ 125cm の位置に Kinect センサを固定し、様々な空間で連続的に計測を行い、RGB および距離画像のデータセットを構築した。対象とした空間のカテゴリは “corridor”, “kitchen”, “laboratory”, “office” および “study room” の 5 つである。まず、それぞれのカテゴリ毎に、いくつかの異なる空間 (部屋) で RGB および距離画像の取得を同時に行った。例えば、表 1 ではカテゴリ “laboratory” は 4 つの異なる laboratory の情報を含んでおり、同一のカテゴリには属するが実際には異なる空間 (部屋) の情報が含まれている。また、それぞれの空間で Kinect センサを異なる軌道で移動させ、RGB および距離画像列の対を複数得た。このあと、各 RGB 画像列とそれに対応する距離画像列の同期をとり、同時刻の RGB および距離画像を得た。表 1 は実験で用いた空間毎の画像の組 (RGB 画像および距離画像) の数を示す。また、データセットに含まれる各カテゴリの RGB および距離画像の一例を図 5 に示す。

4.2 識別性能

構築した RGB および距離画像のデータセットを用い、空間のカテゴリ識別実験を行った。

表 1 各カテゴリの RGB および距離画像の枚数

category	place	number of image sets
corridor	corridor 1	68
	corridor 2	42
	corridor 3	70
	corridor 4	99
kitchen	kitchen 1	73
	kitchen 2	65
	kitchen 3	53
laboratory	laboratory 1	99
	laboratory 2	99
	laboratory 3	81
	laboratory 4	78
study room	study room 1	71
	study room 2	70
	study room 3	49
	study room 4	62
office	office 1	57
	office 2	45
	office 3	47

表 2 実験 1 の識別結果

Confusion matrix for $L = 0$, average correct **89.37%**

%	corridor	kitchen	laboratory	study room	office
corridor	95.66	2.02	2.32	0.00	0.00
kitchen	2.45	76.85	3.21	12.40	5.10
laboratory	0.00	0.12	99.02	0.23	0.63
study room	0.00	8.86	5.57	84.64	0.93
office	0.00	3.51	5.79	0.00	90.70

表 3 実験 2 の識別結果

Confusion matrix for $L = 0$, $U_{Th} = 2$, average correct **87.27%**

%	corridor	kitchen	laboratory	study room	office
corridor	95.05	0.20	3.84	0.91	0.00
kitchen	2.64	69.43	4.15	22.64	1.13
laboratory	0.25	1.24	95.51	2.26	0.75
study room	0.00	3.29	10.57	85.82	0.32
office	0.00	4.39	5.09	0.00	90.53

Confusion matrix for $L = 0$, $U_{Th} = 4$, average correct **92.61%**

%	corridor	kitchen	laboratory	study room	office
corridor	96.47	0.91	2.02	0.61	0.00
kitchen	2.64	85.88	1.51	7.89	2.08
laboratory	0.00	0.20	97.91	0.77	1.12
study room	0.00	5.14	2.29	91.93	0.65
office	0.00	3.51	5.61	0.00	90.88

実験では、leave-one out cross-validation 法を N 回適用し、提案手法の識別性能を評価した。ただし、それぞれのカテゴリに含まれる空間数が異なることから、学習データと評価データの N 個の組を以下のように決定した。まず各カテゴリに含まれるある一つの空間を評価データとする。この空間はそれぞれのカテゴリ内でランダムに決定される。また同じカテゴリの残りの空間のデータは、学習データとして使用される。もし選択された評価データの組が既に評価に使用されていた場合には、それを破棄し、それ以前とは異なる学習、評価データの組が得られるまで選択処理を続ける。これにより、評価データは常に学習データには含まれない空間となる。なお、本実験では $N = 10$ とした。また、それぞれの評価時には、学習データを式 (10) で示したカーネルにより SVM に学習させ、評価データをそれぞれ学習後の SVM で評価して識別率を求めた。

実験 1 では LBP を用いた場合の識別性能を評価した。ただし、空間ピラミッドのレベルは 3 種類 ($L = 0; 1; 2$) とした。平均で最も高い識別性能が得られた $L = 0$ に対する結果を表 2 に示す。これより、LBP を用いた手法により “kitchen” 以外は 85~99% の高い識別性能が得られたが、“kitchen” と “study room” は混同しやすいことがわかる。

実験 2 では SLBP を用いた場合の識別性能を評価した。ただし、0, 1 間の遷移の回数 U

の閾値 U_{Th} は 2 および 4 を用いた。平均で最も高い識別性能が得られた $L = 0$ に対する結果を表 3 に示す。これより、特に $U_{Th} = 4$ の場合で、LBP に比べて SLBP の識別性能が向上していることがわかる。

実験 3 では SVM との比較のため、識別器として Random Forest (RF) を用いて SLBP に対する識別性能を評価した。ただし、 U_{Th} は RF で最も高い識別性能が得られた 4 とした。最も識別精度が高かった $L = 2$ のときの混同行列を表 4 に示す。これより、RF では “kitchen”, “study room” の識別精度の低下が大きく影響したことがわかる。一方、RF では SVM に比べ “office” の識別率は向上した。

表 5 に LBP と SLBP, SVM と RF の比較をまとめて示す。表 5 より、最も識別率が高

表4 実験3の識別結果
Confusion matrix for $L = 2$, average correct **88.43%**

%	corridor	kitchen	laboratory	study room	office
corridor	99.90	0.00	0.10	0.00	0.00
kitchen	0.75	70.57	2.64	23.58	2.45
laboratory	0.00	0.00	96.41	2.60	0.99
study room	0.00	3.71	15.00	79.29	2.00
office	0.00	4.04	0.00	0.00	95.97

表5 識別性能の比較

L	SVM			RF
	LBP(%)	SLBP($U_{Th} = 2$)(%)	SLBP($U_{Th} = 4$)(%)	SLBP($U_{Th} = 4$)(%)
0	89.37	87.27	92.61	85.74
1	85.68	85.53	88.10	87.59
2	84.59	82.46	88.95	88.43

かった手法は $U_{Th} = 4$ の SLBP を用いて SVM で識別した場合であり、SLBP により特徴ベクトルの次元を削減することで、LBP よりも高い精度で空間の識別が可能であることがわかる。さらに、 $U_{Th} = 2$ の SLBP では、LBP の特徴ベクトルに対し 22.8% のサブセットしか用いられないにも関わらず、LBP と比べて識別精度に大きな差はなかった。なお、SVM および Random Forest を用いた場合、最も識別精度が高かったレベル L はそれぞれ 0 および 2 であった。局所ヒストグラムを用いることは、各カテゴリの局所的な特徴の違いを表現するために有効であると考えられ、今後は画像の水平方向のみの分割（例えば天井や床など）などにより、局所ヒストグラムの有効性をより詳細に検討する予定である。

5. 結論と今後の課題

本稿では、Kinect センサから得られた RGB および距離画像を用いて、空間を識別する手法を提案した。提案手法では、濃淡および距離画像から LBP および SLBP を用いてそれぞれ特徴ベクトルを抽出し、それらを組み合わせた特徴ベクトルを SVM および RF により識別した。識別実験の結果、特に新たに提案した SLBP を用いることにより、5つの空間のカテゴリにおいて高い識別精度を示すことが確認された。

今後は、より適当な記述子について検討するとともに、学習、評価データのカテゴリ数や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) Stachniss, C., Mozos, O.M. and Burgard, W.: Speeding-up multi-robot exploration by considering semantic place information, *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp.1692–1697 (2006). Orlando, FL, USA.
- 2) Zender, H., Jensfelt, P. and Kruijff, G.-J.: Human- and situation-aware people following, *Proc. of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp.1131–1136 (2007).
- 3) Torralba, A., Murphy, K.P., Freeman, W.T. and Rubin, M.A.: Context-based vision system for place and object recognition, *International Conference on Computer Vision (ICCV)* (2003). Nice, France.
- 4) Kollar, T. and Roy, N.: Utilizing object-object and object-scene context when planning to find things, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (2009). Kobe, Japan.
- 5) Mozos, O.M., Stachniss, C. and Burgard, W.: Supervised learning of places from range data using AdaBoost, *IEEE International Conference on Robotics and Automation (ICRA)* (2005). Barcelona, Spain.
- 6) Pronobis, A., Jensfelt, P., Sjö, K., Zender, H., Kruijff, G.-J.M., Mozos, O.M. and Burgard, W.: Semantic modelling of space, *Cognitive Systems, ser. Cognitive Systems Monographs, H. I. Christensen, A. Sloman, G.-J. M. Kruijff, and J. Wyatt, Eds. Germany: Springer*, pp.165–221 (2010).
- 7) Zender, H., Mozos, O.M., Jensfelt, P., Kruijff, G.-J.M. and Burgard, W.: Conceptual spatial representations for indoor mobile robots, *Robotics and Autonomous Systems*, Vol.56, No.6, pp.493–502 (June 2008).
- 8) Christensen, H.I., Kruijff, G. and J.Wyatt, E.: *Cognitive Systems, ser. COSMOS. Springer Verlag* (May 2010).
- 9) Zender, H., Jensfelt, P., Mozos, O.M., Kruijff, G.-J.M. and Burgard, W.: An integrated robotic system for spatial understanding and situated interaction in indoor environments, *Proceedings of the Conference on Artificial Intelligence (AAAI)* (2007). Vancouver, British Columbia, Canada.
- 10) Friedman, S., Pasula, H. and Fox, D.: Voronoi random fields: Extracting the topological structure of indoor environments via place labeling, *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)* (2007). Hyderabad, India.

- 11) Brunskill, E., Kollar, T. and Roy, N.: Topological mapping using spectral clustering and classification, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (October 2007). San Diego.
- 12) Wu, J., Christensen, H.I. and Rehg, J.M.: 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.
- 13) Ranganathan, A.: PLISS: Detecting and Labeling Places Using Online Change-Point Detection, *Robotics: Science and Systems VI* (2010).
- 14) Rottmann, A., Mozos, O.M., Stachniss, C. and Burgard, W.: Semantic place classification of indoor environments with mobile robots using boosting, *National Conference on Artificial Intelligence (AAAI)* (2005). Pittsburgh, PA, USA.
- 15) Pronobis, A., Mozos, O.M., Caputo, B. and Jensfelt, P.: Multi-modal semantic place classification, *International Journal of Robotics Research*, Vol.29, No.2-3, pp. 298-320 (2010). February-March.
- 16) : *Microsoft Kinect*. <http://www.xbox.com/en-us/kinect>.
- 17) Ojala, T., Pietikainen, M. and Harwood, D.: A comparative study of texture measures with classification based on featured distributions, *Pattern Recognition*, Vol.29, No.1, pp.51-59 (1996).
- 18) Mozos, O.M. and Burgard, W.: 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.
- 19) Shi, L., Kodagoda, S. and Dissanayake, G.: 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.
- 20) Lazebnik, S., Schmid, C. and Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2006).
- 21) Ojala, T., Pietikäinen, M. and Mäenpää, T.: Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, Vol.24, No.7, pp.971-987 (July 2002).
- 22) Cortes, C. and Vapnik, V.: Support-vector network, *Machine Learning*, Vol.20, pp. 273-297 (1995).
- 23) Bishop, C.M.: *Pattern Recognition and Machine Learning*, Springer (2006).
- 24) Knerr, S., Personnaz, L. and Dreyfus, G.: 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.
- 25) Chang, C.-C. and Lin, C.-J.: 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>.
- 26) Hsu, C.-W., Chang, C.-C. and Lin, C.-J.: *A practical guide to support vector classification* (2010). <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.



図 5 各カテゴリ (corridor, kitchen, laboratory, office, study room) での RGB および距離画像の例