

# Categorization of Indoor Places by Combining Local Binary Pattern Histograms of Range and Reflectance Data from Laser Range Finders

Oscar Martinez Mozos<sup>a</sup>      Hitoshi Mizutani<sup>b</sup>,      Ryo Kurazume<sup>b</sup>  
and Tsutomu Hasegawa<sup>b</sup>

<sup>a</sup>*School of Computer Science, University of Lincoln, Lincoln LN6 7TS, UK,*

*E-mail: omozos@lincoln.ac.uk*

<sup>b</sup>*Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan,*

*E-Mails: mizutani@irvs.ait.kyushu-u.ac.jp, kurazume@ait.kyushu-u.ac.jp, hasegawa@ait.kyushu-u.ac.jp*

## Abstract

This paper presents an approach to categorize typical places in indoor environments using 3D scans provided by a laser range finder. Examples of such places are offices, laboratories, or kitchens. In our method, we combine the range and reflectance data from the laser scan for the final categorization of places. Range and reflectance images are transformed into histograms of local binary patterns and combined into a single feature vector. This vector is later classified using support vector machines. The results of the presented experiments demonstrate the capability of our technique to categorize indoor places with high accuracy. We also show that the combination of range and reflectance information improves the final categorization results in comparison with a single modality.

*keywords:* place categorization, laser scanner, range image, reflectance image

## 1 INTRODUCTION

An important capability for service robots acting and working indoors is their ability to categorize the different places in which indoor environments are typically divided like for example corridors, laboratories or kitchens. Service robots able to distinguish the previous indoor places, and able to assign them the corresponding label, i.e. *office* or *laboratory*, will greatly improve their communication capabilities with humans [1].

In the place categorization problem the robot takes one observation at some location and it assigns this observation the label of the corresponding place. The corresponding labels represent general categories of places and not particular instances. That means that the robot assigns the same label to

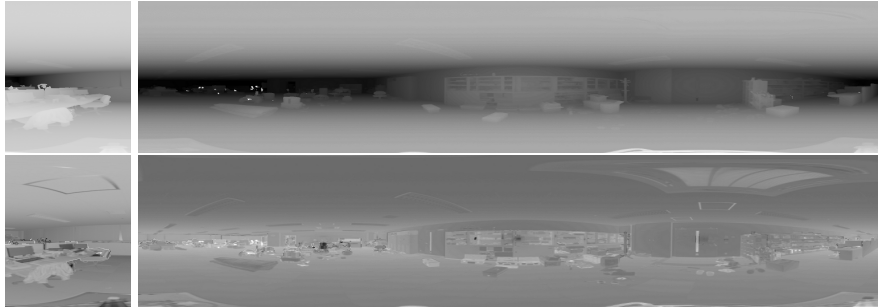


Figure 1: The top images depict a partial (left), and panoramic (right) range images in which darker values indicate farther distances. The corresponding reflectance images are shown in the bottom with darker values indicating lower reflectances. Images have been transformed into gray scale, and their color curves have been manipulated for a better visualization.

different places that pertain to the same category. For example the service robot can assign the label *office* to any office located in an building independently of its location, since all offices belong to the same category. An advantage of place categorization is that a service robot can recognize new instances of a known category even if this new instance was not seen before. That means that the robot can assign the term *office* to new previously unseen offices.

The categorization of places constitutes an additional piece of information that allows a service robot to improve its behavior in other high-level tasks. The categorical label assigned to a place can be used as entry point to a knowledge database in which each place label has assigned some predetermined behavior. For example, in the work presented in [2], the label *corridor* implies that this place leads to other offices in the environment and thus it is a good candidate to start exploring. Moreover, in [3] a place recognized as a *corridor* implies that the robot can move faster. In the previous works, the implicit information about the different behaviors related to the label *corridor* are specified in the system beforehand. In addition to behaviors, a system can also link to the labels information about typical objects found in each place. This information can be used to improve the object searching task [4, 5]. A final application of place categorization constitutes the high level representation of indoor environments by autonomous robots [6, 7].

In this paper, we present an approach to categorize places in indoor environments using 3D scans from a laser range finder installed on a mobile robot. Laser sensors provide highly accurate distance measurements, and they also provide information about reflectance values at each measured point. An important advantage of laser sensors over other sensor modalities is that the provided range and reflectance data are robust under illumination changes. To the best of our knowledge this is the first work which combines range and reflectance data from a laser range finder to categorize laser scans into typical indoor places.

In our method we first divide the panoramic scans into a set of overlapping partial scans which are categorized independently. Partial scans cover only a restricted field of view of approximately 10 degrees

around the robot, whilst complete panoramic scans cover 360 degrees around the robot. Examples of partial and panoramic scans are shown in Fig. 1. The final categorization for a panoramic scan can be obtained by selecting the place with highest probability according to the categorization of the partial scans.

Our categorization approach combines the range and reflectance information provided by a laser scanner to categorize indoor places. Range and reflectance images are transformed into histograms of local binary patterns (LBP) whose dimensionality is further reduced by applying a uniformity threshold. The histograms resulting from each modality are then combined into a single feature vector, which is categorized using support vector machines. In addition, we apply spatial pyramids to improve the accuracy of the final classification. Since our method is supervised, the set of categories is predefined in advance.

In this paper we present categorization results using both partial and panoramic scans for five categories: corridor, kitchen, laboratory, study rooms, and offices. Using only partial scans we obtain an average categorization accuracy of over 93%. This result demonstrates that we can categorize place with high accuracy by taking a partial scan covering only 10 degrees around the robot. This capability can be very advantageous in situations where the robot does not have enough time to register and analyze a complete panoramic scan. In addition, we present categorization results of complete panoramic scans in which we obtain an average accuracy of 96%. This accuracy increases to 100% in corridors, laboratories and study rooms from our specific environment.

The rest of this paper is organized as follows. After presenting related work in Sect. 2 we describe our laser range finder sensor in Sect. 3. Our main method for place categorization is introduced in Sect. 4. In Sect. 5 we describe the application of our multi-modal method to classify partial laser scans. In Sect. 6 we present our approach to classify complete panoramic scans. Finally, in Sect. 7 we show our experimental results.

## 2 Related Work

The problem of place categorization has been addressed by researches using different types of sensors. For example, 2D laser scans are categorized into different indoor places in [8] by transforming each scan into a vector of geometrical features that are fed into a supervised classifier. In [9], similar features from 2D laser scans are used to classify locations in *Voronoi Random Fields*. Moreover, the work in [10] is able to determine the place category for each beam in the scan using logistic regression techniques. In comparison to these previous works we used 3D scans as observations which provides a richer representation of the places and allows us to increase the number of categories to be recognized.

Other techniques categorize places using camera images only. For example, in [11] local and global features from images taken by a wearable camera are classified using a hidden Markov model. Moreover, the PLISS system introduced in [12] represents the images using a bag of words approach and it further detects change-points in sequences of images. In [13] images representing different indoor places in office

environments are categorized by combining multiple visual cues. Further works have applied local binary patterns (LBPs) [14] to camera images for scene categorization. In [15], the CENTRIST descriptor is applied to images to classify different scenes and objects. Color based CENTRIST is applied in [16] by combining the values in HSV color space. Finally, in [17], histograms of oriented uniform LBPs are extracted from images to categorize places indoors and outdoors. Finally, visual cues have been used in combination with 2D laser to improve the final classification in [18]. However, visual images strongly depend on illumination conditions while scans obtained with a laser range finder are more robust under illumination changes. For this reason we use laser scans as main observations in our system.

Depth data have also been used to categorized places indoors. In [19], a set of global geometrical features are extracted from 3D point clouds obtained using several Kinect cameras covering 360 degrees around the robot. In comparison to this work we use a laser range finder which has a higher accuracy and working range. Moreover, we use more categories and we are also able to classify partial scans covering just 10 degrees around the robot.

Our work is close to [20], where a Kinect camera is used to categorize pairs of depth and gray scale images into indoor places using histograms of local binary patterns. However, the working range of the Kinect camera is very short. Moreover, the RGB images provided by the Kinect camera are less robust to changes in illumination making this cue less reliable. In contrast, in this paper we use laser range finders as main sensor since they are more stable under illumination changes. Moreover, we combine range and reflectance data for the final categorization of places, and present a method to classify complete panoramic scans by using the categorization of partial ones. We believe the results presented in this paper are of great interest in the robotics community since a large number of robotics applications use laser range finders as main sensor for indoor mapping and localization.

Finally, reflectance data from laser range finders has been applied to solve different tasks. For instance, in [21] reflectance data is used as part of a trilateral filter for denoising range data. In [22] laser reflectance data is used to localize a robot outdoors. Moreover, in [23] several techniques are presented for 3D city modeling using range and reflectance data from laser range finders. To the best of our knowledge our work is the first including reflectance data from a laser range finder for indoor place categorization.

### 3 Range and Reflectance Sensor Data

In our approach we use a laser range finder (LRF) situated on a mobile robot as shown in Fig. 2. This kind of sensor provides range data by measuring the round-trip time of a laser pulse reflected by an object. Our specific sensor is a SICK LMS151 Laser with maximum range of 50 meters and angular resolution of 0.25 degrees. The measurements have a systematic error of  $\pm 30\text{mm}$  with 1 sigma equal to  $\pm 12\text{mm}$ .

In our configuration, we rotate the laser around the vertical axis to obtain a complete panoramic range image around the robot. In addition to range data, laser range finders can measure the strength of



Figure 2: The left picture shows the robot used in our experiments. A white arrow indicates the laser range finder used to take scans. The right picture presents a closer look at the laser range finder.

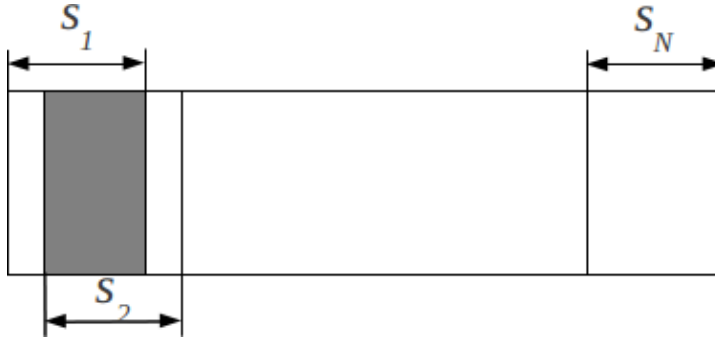


Figure 3: Partial scans in a panoramic image with gray areas indicating the overlapping area. Each partial scan  $s_i$  has a resolution of  $624 \times 760$  points. A complete panoramic scan has a resolution of  $3753 \times 760$  points.

the reflected laser pulse. This reflectance value indicates the intensity of the point on the target surface under a single-frequency light source. We obtain a unique reflectance value for each laser beam which means that the range image and the reflectance image obtained by a LRF are precisely and fundamentally aligned. The bottom image of Fig. 1 shows the panoramic reflectance image corresponding to the top panoramic range image in the same figure. In this work we do not apply any correction to the reflectance values provided by the laser range finder.

In a further step, each panoramic scan is divided into overlapping sub-scans as shown in Fig 3. The reason to divide each panoramic scan is twofold. First, we have a bigger number of training and test scans for our experiments. Second, we want to test whether we can obtain good classification results with partial scans only, since in some real situations the robot cannot spend much time waiting for a complete panoramic scan.

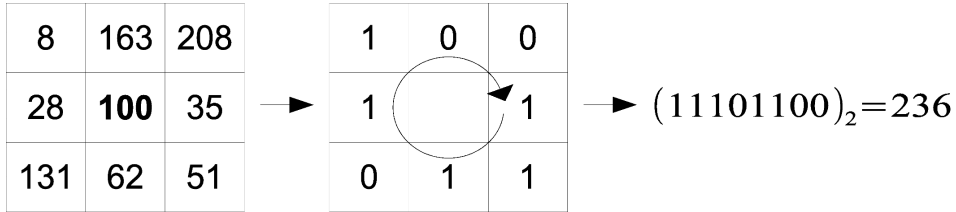


Figure 4: The left image presents a toy example with a center pixel (marked in bold) and the corresponding 8-neighbors. The middle image presents the binary results after the comparisons. The right image shows the resulting decimal value.

## 4 Histograms of Local Binary Patterns in Depth and Reflectance Images for Partial Scans

As explained above, we take complete panoramic scans that are divided into partial ones. In this section, we explain how to obtain histograms of local binary patterns (LBP) for depth and reflectance images corresponding to each partial scan.

We store the range and reflectance measurements of each partial scan in the matrices  $D$  and  $R$  respectively. The measurements in each matrix are ordered so that neighboring values in the matrices correspond to real neighboring measurements in the environment. In this way the matrices  $D$  and  $R$  can be treated as images in which each pixel contains the corresponding range and reflectance value.

In a second step we apply a local binary transformation to each matrix  $D$  and  $R$  as follows [14]. We compare each value  $v(i)$  at position  $(x_i, y_i)$  in the matrix with each of the values  $v(j)$  corresponding to the 8-neighboring pixels  $j \in N_8(i)$  as shown in Fig. 4. Then for each comparison we obtain a binary value  $b_j \in \{0, 1\}$  indicating whether the value  $v(i)$  was bigger or smaller than the neighboring value  $v(j)$  as

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

The resulting binary values are ordered clockwise starting on the value which is to the right to the reference pixel  $v(i)$ . The corresponding binary string is then transformed into a decimal value  $d$  in the range  $[0, \dots, 255]$ . This decimal value  $d$  will be the new value for pixel  $i$ . An example of this process is shown in Fig. 4. The local binary operator that we apply is equivalent to the  $LBP_{8,1}$  operator in [14] with the sole difference that we do not interpolate values at the diagonals.

Finally, this transformation is applied to every pixel in the range and reflectance matrices obtaining the transformed matrices  $D^T$  and  $R^T$  respectively. In a further step we represent each transformed matrix into a histogram in which each bin indicates the frequency of appearance of each decimal value  $d \in [0, \dots, 255]$ . Thus each histogram contains 256 bins. Following the ideas in [20], we further reduce the dimensionality of each histogram by selecting bins according to a uniformity measurement  $U$  [14]. Each bin in a histogram represents a decimal value  $d$  which, in turn, corresponds to a binary string (see Fig. 4). The uniformity measurement  $U(d)$  of a decimal value  $d$  indicates the number of transitions

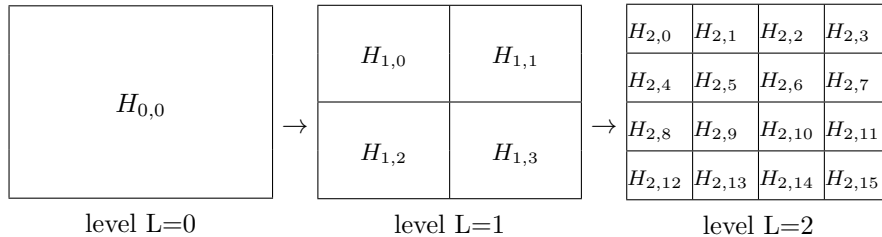


Figure 5: Spatial pyramid with 3 levels. Each partial image contains one histogram  $H_{L,i}$  with level  $L \in \{0, 1, 2\}$  and  $i \in \{0, \dots, 2^L \times 2^L - 1\}$ .

between 0/1 values in its corresponding binary string and can be obtained as

$$U(i) = |b_1(i) - b_8(i)| + \sum_{j=1}^7 |b_j(i) - b_{j+1}(i)|. \quad (2)$$

As an example,  $U(236) = 4$  as can be corroborated in Fig. 4.

After obtaining the uniformity measurement of the decimal value represented by each bin in the histogram, we keep only the bins that represent decimal values whose uniformity measurement is equal or below a certain threshold  $\theta$ . If  $H$  is the histogram representation of a transformed image, then the new reduced histogram is obtained as

$$H(\theta) = \{h_d \in H \mid U(d) \leq \theta\}, \quad (3)$$

where  $h_d$  indicates the bin in the histogram  $H$  for the LBP decimal value  $d$ . Lower values of  $\theta$  result in histograms with lower dimensionality. By changing the threshold  $\theta$  we are selecting local binary patterns with different uniformity, i.e. representing different local structures. According to previous works on LBP [14], [20] features with low uniformity threshold usually correspond to more discriminative local binary patterns. Another advantage of this selection is the reduction in the dimension of the final feature vectors representing the scans. This reduction significantly decreases the time for training and classifying. In the experimental section we will show results that indicate that lower values of  $\theta$  provide slightly better classification results while significantly reducing the size of the final feature vectors.

Finally, we apply this reduction to the histograms representing the range and reflectance data  $H^D$  and  $H^R$ , and we obtain the corresponding reduced histograms  $H^D(\theta)$  and  $H^R(\theta)$  respectively.

In order to represent each range and reflectance matrix by a final feature vector we use an extra step which consists of the application of spatial pyramids [24]. In this step we recursively divide each matrix into smaller parts. We treat each matrix division as an independent matrix and calculate its corresponding LBP histogram  $H_{l,i}$ , where  $l$  indicates the level and  $i$  identifies a partial image. An example of this process is shown in Fig. 5. The final feature vector representing the complete original matrix is obtained by concatenating the histograms of each local part in each level

$$\vec{x} = \{\{H_{0,0}\}, \{H_{1,0}, \dots, H_{1,3}\}, \dots, \{H_{L,0}, \dots, H_{L,2^L \times 2^L - 1}\}\}. \quad (4)$$

Finally, by applying a particular uniformity threshold  $\theta$  to each local histogram following (3), and by concatenating them according to (4), we obtain the feature vectors  $\vec{x}^D(\theta)$  and  $\vec{x}^R(\theta)$  for the range

and reflectance images respectively as

$$\begin{aligned}\vec{x}^D(\theta) &= \{\{H_{0,0}^D(\theta)\}, \{H_{1,0}^D(\theta), \dots, H_{1,3}^D(\theta)\}, \dots, \{H_{L,0}^D(\theta), \dots, H_{L,2^L \times 2^L - 1}^D(\theta)\}\} \\ \vec{x}^R(\theta) &= \{\{H_{0,0}^R(\theta)\}, \{H_{1,0}^R(\theta), \dots, H_{1,3}^R(\theta)\}, \dots, \{H_{L,0}^R(\theta), \dots, H_{L,2^L \times 2^L - 1}^R(\theta)\}\}.\end{aligned}\quad (5)$$

## 5 Multi-Modal Categorization of Partial Scans

In this paper, we combine two modalities, range and reflectance data, at the feature level. Thus, for each pair of feature vectors  $\vec{x}^D(\theta)$  and  $\vec{x}^R(\theta)$  representing a partial scan in one particular place, a final vector  $\vec{x}^P(\theta)$  is obtained by their concatenation as

$$\vec{x}^P(\theta) = \{\vec{x}^D(\theta), \vec{x}^R(\theta)\}.\quad (6)$$

Once we have a final multi-modal feature vector  $\vec{x}^P(\theta)$  for representing partial scans of indoor places we use a support vector machine (SVM) [25, 26] for the final categorization. Support vector machines take as input a set of  $n$  feature vectors  $\vec{x}_i$  together with their labels  $y_i \in \{1, -1\}$ . The idea behind SVMs is to find the hyperplane that maximizes the distance between the examples of the two classes  $\{1, -1\}$ . This is done by finding a solution to the optimization problem

$$\min_{\vec{w}, b, \xi} C \sum_{i=1}^n \xi_i + \frac{1}{2} \|\vec{w}\|^2,\quad (7)$$

subject to the condition

$$y_i (\vec{w}^T \phi(\vec{x}_i) + b) \geq 1 - \xi_i,\quad (8)$$

where  $\vec{w}$  is the normal to the hyperplane, and  $\xi_i \geq 0$  are slack variables that measure the error in the misclassification of  $\vec{x}_i$ . In addition, we use a radial basis function (RBF) kernel

$$K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma \|\vec{x}_i - \vec{x}_j\|^2), \gamma > 0\quad (9)$$

Multi-class classification is performed using the ‘‘one-against-one’’ approach [27]. In our experiments we use the LIBSVM library [28]. Following the method in [29], the parameters  $C$  and  $\gamma$  are selected by grid-search using cross-validation in the ranges  $C \in [2^{-5}, \dots, 2^{15}]$  and  $\gamma \in [2^{-12}, \dots, 2^3]$  with grid step 3 in both cases.

## 6 Categorization of Complete Panoramic Scans

In this section, we present a method to categorize complete panoramic scans using their corresponding partial scans. In addition, we introduce a probabilistic representation over categories for each panoramic scan.

We apply our previous categorization method to each partial scan, and we construct a categorization vector  $\vec{z}$  representing the whole panoramic image as

$$\vec{z} = \{C(s_1), \dots, C(s_N)\},\quad (10)$$



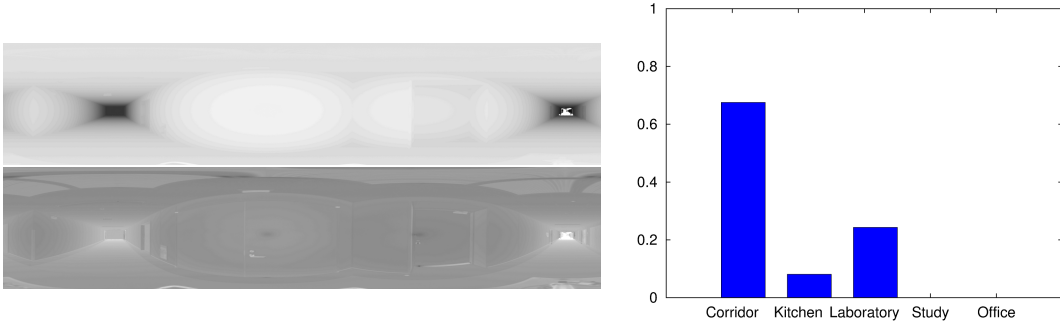


Figure 6: The left images correspond to panoramic range and reflectance data in a corridor environment. Darker values indicate farther distances and lower reflectances respectively. Both images have been transformed into gray scale for a better visualization. The right image depicts the corresponding probability distribution over places.

where  $C(s_i)$  is the categorization result of partial scan  $s_i$ , and  $N$  is the total number of partial scans. Thus, the vector  $\vec{z}$  is transformed into a probabilistic distribution over categories

$$\vec{P}(\vec{z}) = \{P(\text{label}_j | \vec{z}), \dots, P(\text{label}_M | \vec{z})\}, \quad (11)$$

where each  $\text{label}_j$  corresponds to one of our categories, i.e.  $\text{label}_1 = \text{Corridor}$ ,  $\text{label}_2 = \text{Kitchen}$ ,  $\text{label}_3 = \text{Laboratory}$ ,  $\text{label}_4 = \text{Study Room}$ , and  $\text{label}_5 = \text{Office}$ . Finally,  $M$  is the total number of categories, which is 5 in our case. The probability for a specific place  $P(\text{label}_j | \vec{z})$  is obtained as

$$P(\text{label}_j | \vec{z}) = \frac{\sum_i^N I(C(s_i) = \text{label}_j)}{N}, \quad (12)$$

where  $I(\cdot)$  is the indicator function which returns 1 if the argument is true, and 0 otherwise. Finally, the categorization of a panoramic scan  $S$  is given by the label with maximum probability

$$\text{label}_S = \arg \max_{\text{label}_j} P(\text{label}_j | \vec{z}), \quad \forall \text{label}_j. \quad (13)$$

An example panoramic scan taken in a corridor together with its corresponding probability distribution is shown in Fig. 6. As we can appreciate the histogram indicates a higher probability for the label *corridor*.

## 7 Experiments

In this section, we present results demonstrating that our system based on a laser range finder is able to categorize places with high precision. We first introduce the dataset we use in our experiments. Then we present results on the categorization of partial scans by comparing single and combined modalities. Finally, we show results on complete panoramic scans.

## 7.1 Dataset of Indoor Places

For our experiments, we have created a dataset of different indoor places. This dataset is composed of five indoor places: corridors, kitchens, laboratories, study rooms, and offices. At each place we took 15 panoramic 3D laser scans covering 360 degrees around the robot with laser positioned 95cm above the floor. To take the panoramic scans we situated the robot in 15 different locations inside each place. The locations were spatially distributed inside each place covering most of the possible situations. For example we put the robot in the centre of each place, close to the corners, close to the doorways and windows, etc. Some example of panoramic scans are shown in Fig. 1. Each panoramic scan has a resolution of  $3753 \times 760$  points.

We further divide a panoramic scan into overlapping sub-scans of  $624 \times 760$  points starting every 104 pixels (which corresponds to approx. 10 degrees) by following the steps presented in Sect. 6. Finally, our dataset is composed of a total of 285 panoramic scans, and a total of 10545 partial scans. A summary of our dataset together with some examples are shown in Fig. 7.

## 7.2 Categorization of Partial Scans

In the next experiments we apply our categorization approach as follows. We randomly select one place from each category in our dataset, i.e. one corridor, one kitchen, one study room, one office, and one laboratory. Then we use the partial scans of these places for testing purposes and use the rest of data to train the classifier. In this way, the training data does not contain any scan of a place from the test data. This approach allows us to test the generalization capabilities of our method on previously unseen places. We repeat this process 10 times and average over final the categorization results.

In a first experiment we search for the level of spatial pyramids  $L$  and uniformity threshold  $\theta$  which provide the best classification results using combined range and reflectance feature vectors (Sect. 5). For this experiment we have used a subset of the dataset presented in the previous section with the aim of reduce the searching time. The results of this search are shown in Table 1. The values in this table represent average correct categorization results over all place categories. We can see that we obtain best overall results using the parameters  $L = 2$  and  $\theta = 4$ . Moreover, using a threshold of  $\theta = 4$  we reduced in 22.6% the dimension of the final feature vector.

In the next experiment we have used the previous best parameters to classify the whole dataset. The confusion matrix for this experiment is presented in Table 2. The worst result is obtained for the category *Kitchen*. This is probably due to our dataset, which contains two different types of kitchens with very different furniture distributions. Examples of such kitchens are shown in Fig. 8. One possible solution can be two create two different kitchen sub-categories, however, in this paper we want to focus on general classes.

Finally, in Table 3 we present a comparison of categorization results when using single and combined modalities. In this table, we show the average correct classification over all place categories. We can see that the average categorization rate improves when combining modalities.

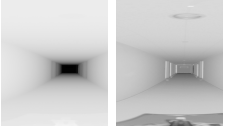


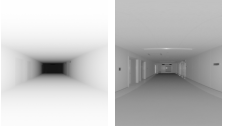
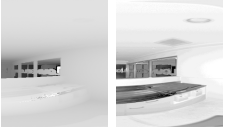
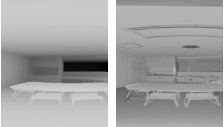
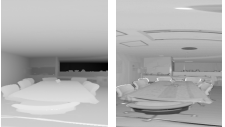
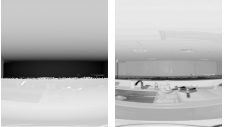
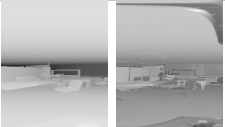
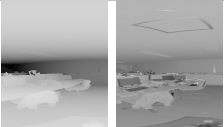
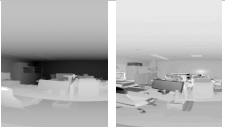
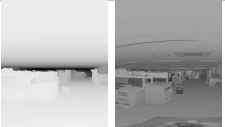
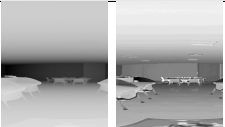
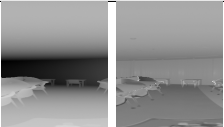

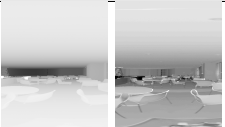
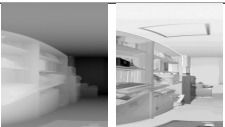

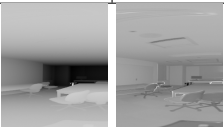
			
corridor 1 (15,555)	corridor 2 (15,555)	corridor 3 (15,555)	corridor 4 (15,555)
			
kitchen 1 (15,555)	kitchen 2 (15,555)	kitchen 3 (15,555)	kitchen 4 (15,555)
			
Laboratory 1 (15,555)	Laboratory 2 (15,555)	Laboratory 3 (15,555)	Laboratory 4 (15,555)
			
Study Room 1 (15,555)	Study Room 2 (15,555)	Study Room 3 (15,555)	Study Room 4 (15,555)
			
Office 1 (15,555)	Office 2 (15,555)	Office 3 (15,555)	

Figure 7: Summary of the data contained in our dataset. Each cell contains one example of range and corresponding reflectance image obtained in a particular place. The images have been converted into gray scale for better visualization purposes. Darker colors indicate farther places and lower reflectances respectively. Under the images we find the name of the place and, in brackets, the number of panoramic scans taken in that place and the total number of partial scans after dividing all corresponding panoramic scans.

### 7.3 Categorization of Complete Panoramic Scans

As explained in Sect. 6, we divide the panoramic scans into partial ones. Then we categorize each partial scan  $s_i$  and obtain a categorization vector  $\vec{z}$  representing the panoramic image (see Sect. 6). To test the categorization of complete scans we have applied the previous method to the panoramic scans corresponding to the 10 test sets of Sect. 7.2. For each test set, we have categorized the complete scans corresponding to each place and we have averaged over the 10 sets. The resulting confusion matrix is presented in Table 4. Using complete panoramic scans improve the classification in all places as we can see by comparing Table 4 and Table 2. In our specific environment, we reach perfect recognitions in corridors, laboratories, and study rooms by using complete panoramic scans.

Table 1: Categorization for different  $L$  and  $\theta$  parameters using a combination of range and reflectance data.

%	L=0	L=1	L=2
$\theta = 2$	$83.70 \pm 14.43$	$86.32 \pm 14.16$	$89.06 \pm 12.88$
$\theta = 4$	$88.19 \pm 14.31$	$91.78 \pm 13.09$	<b><math>93.46 \pm 10.59</math></b>
$\theta = 6$	$88.27 \pm 14.47$	$91.78 \pm 12.04$	$93.05 \pm 10.68$
$\theta = 8$	$88.11 \pm 14.75$	$91.76 \pm 12.14$	$92.95 \pm 10.69$

Table 2: Confusion matrix for the categorization of places using  $L = 2$  and  $\theta = 4$  and a combination of range and reflectance data.

%	Corridor	Kitchen	Laboratory	Study Room	Office
Corridor	<b><math>99.37 \pm 0.85</math></b>	$0.31 \pm 0.79$	$0.07 \pm 0.17$	$0.00 \pm 0.00$	$0.25 \pm 0.52$
Kitchen	$4.11 \pm 9.27$	<b><math>77.20 \pm 40.54</math></b>	$0.13 \pm 0.27$	$1.23 \pm 2.62$	$17.33 \pm 32.23$
Laboratory	$0.31 \pm 0.53$	$0.00 \pm 0.00$	<b><math>99.33 \pm 1.27</math></b>	$0.00 \pm 0.00$	$0.36 \pm 0.76$
Study Room	$0.00 \pm 0.00$	$0.09 \pm 0.28$	$0.00 \pm 0.00$	<b><math>99.91 \pm 0.28</math></b>	$0.00 \pm 0.00$
Office	$0.56 \pm 0.80$	$6.22 \pm 9.21$	$0.05 \pm 0.17$	$0.00 \pm 0.00$	<b><math>93.17 \pm 9.89</math></b>

## 8 Conclusion

In this paper, we have presented a novel approach for place categorization in indoor environments by combining distance and reflectance measurements from a laser range finder. For each modality, we calculate the corresponding histogram of local binary patterns and we additionally reduce the dimensionality of each histogram using a uniformity criterion. The results of our experiments show that we can classify complete panoramic scans with high reliability. We also have shown that we can obtain high classification rates with individual partial scans. Finally, our experiments indicate that the combination of both modalities, distance and reflectance information, improve the final categorization of laser scans.

As future work, we want to study the application of new integration methods and modalities for place categorization. Moreover, we want to increase the size of our data set by including more categories. Finally, we plan to study the effect that specific locations where the observation can have in the final classification results. In this way we can analyze which positions inside the rooms are more advantageous to obtain a better confidence in the classification results.

## 9 Acknowledgments

This work has been partially supported by Grants-in-Aid for Foreign Fellows from the Japan Society for the Promotion of Science (22-00362) and by Grant-in-Aid for Scientific Research (B) (22300069).



Figure 8: The two different types of kitchens in our dataset.

Table 3: Comparison of results between single and combined modalities with  $L = 2$  and  $\theta = 4$ .

range	reflectance	range + reflectance
$91.88 \pm 9.19$	$91.36 \pm 8.47$	<b><math>93.79 \pm 10.56</math></b>

## REFERENCES

- [1] K. Sjöo, H. Zender, P. Jensfelt, G.-J. M. Kruijff, A. Pronobis, N. Hawes, and M. Brenner, “The Explorer system,” in *Cognitive Systems* (H. I. Christensen, G.-J. M. Kruijff, and J. L. Wyatt, eds.), vol. 8 of *Cognitive Systems Monographs*, ch. 10, pp. 395–421, Berlin/Heidelberg, Germany: Springer Verlag, 2010.
- [2] C. Stachniss, O. M. Mozos, and W. Burgard, “Efficient exploration of unknown indoor environments using a team of mobile robots,” *Annals of Mathematics and Artificial Intelligence*, vol. 52, pp. 205–227, April 2008.
- [3] H. Zender, P. Jensfelt, and G.-J. Kruijff, “Human- and situation-aware people following,” in *Proc. of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 1131–1136, August 2007.
- [4] A. Aydemir, M. Göbelbecker, A. Pronobis, K. Sjöo, and P. Jensfelt, “Plan-based object search and exploration using semantic spatial knowledge in the real world,” in *Proc. of the 5th European Conference on Mobile Robots (ECMR’11)*, (Örebro, Sweden), Sept. 2011.
- [5] T. Kollar and N. Roy, “Utilizing object-object and object-scene context when planning to find things,” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, (Kobe, Japan), 2009.
- [6] H. Zender, P. Jensfelt, O. M. Mozos, G.-J. M. Kruijff, and W. Burgard, “An integrated robotic system for spatial understanding and situated interaction in indoor environments,” in *Proceedings of the Conference on Artificial Intelligence (AAAI)*, (Vancouver, British Columbia, Canada), 2007.

Table 4: Confusion matrix for the categorization of panoramic images

%	Corridor	Kitchen	Laboratory	Study Room	Office
Corridor	<b>100.00</b>	0.00	0.00	0.00	0.00
Kitchen	4.00 ± 10.52	<b>80.00 ± 42.16</b>	0.00	0.00	16.00 ± 34.31
Laboratory	0.00	0.00	<b>100.00</b>	0.00	0.00
Study Room	0.00	0.00	0.00	<b>100.00</b>	0.00
Office	0.00	1.33 ± 2.81	0.00	0.00	<b>98.67 ± 2.81</b>

- [7] A. Pronobis and P. Jensfelt, “Large-scale semantic mapping and reasoning with heterogeneous modalities,” in *Proc. of the 2012 IEEE International Conference on Robotics and Automation (ICRA)*, (Saint Paul, MN, USA), May 2012.
- [8] 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.
- [9] S. Friedman, H. Pasula, and D. Fox., “Voronoi random fields: Extracting the topological structure of indoor environments via place labeling,” in *In Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, (Hyderabad, India), 2007.
- [10] L. Shi, S. Kodagoda, and G. Dissanayake, “Laser range data based semantic labeling of places,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Taipei, Taiwan), October 2010.
- [11] 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.
- [12] A. Ranganathan, “PLISS: Detecting and Labeling Places Using Online Change-Point Detection,” in *Proceedings of Robotics: Science and Systems*, (Zaragoza, Spain), June 2010.
- [13] A. Pronobis and B. Caputo, “Confidence-based cue integration for visual place recognition,” in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (San Diego, CA, USA), pp. 2394–2401, Oct. 2007.
- [14] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, jul 2002.
- [15] J. Wu and J. 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.

- [16] W.-T. Chu and C.-H. Chen, “Color centrist: a color descriptor for scene categorization,” in *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval (ICMR)*, (Hong Kong), 2012.
- [17] E. Fazl-Ersi and J. K. Tsotsos, “Histogram of oriented uniform patterns for robust place recognition and categorization,” *The International Journal of Robotics Research*, vol. 31, no. 4, pp. 468–483, 2012.
- [18] A. Rottmann, O. M. Mozos, C. Stachniss, and W. Burgard, “Semantic place classification of indoor environments with mobile robots using boosting.,” in *National Conference on Artificial Intelligence (AAAI)*, (Pittsburgh, PA, USA), 2005.
- [19] L. Shi, S. Kodagoda, and R. Ranasinghe, “Fast indoor scene classification using 3d point clouds,” in *Australasian Conference on Robotics and Automation (ACRA)*, (Melbourne, Australia), December 2011.
- [20] O. M. Mozos, H. Mizutani, R. Kurazume, and T. Hasegawa, “Categorization of indoor places using the kinect sensor,” *Sensors*, vol. 12, pp. 6695–6711, May 2012.
- [21] S. Oishi, R. Kurazume, Y. Iwashita, and T. Hasegawa, “Denoising of range images using a trilateral filter and belief propagation,” in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (San Francisco, CA, USA), pp. 2020–2027, Sept. 2011.
- [22] D. Zhang, R. Kurazume, Y. Iwashita, and T. Hasegawa, “Appearance and map-based global localization using laser reflectivity,” in *Proc. of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, (Phuket, Thailand), pp. 1010–1016, Dec. 2011.
- [23] P. Axelsson, “Processing of laser scanner data algorithms and applications,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2-3, pp. 138–147, 1999.
- [24] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006.
- [25] C. Cortes and V. Vapnik, “Support-vector network,” *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [26] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [27] 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.
- [28] 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>.

- [29] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, “A practical guide to support vector classification.” <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>, 2010.