# Object Classification with Range and Reflectance Data from a Single Laser Scanner

Shuji Oishi[a], Naoaki Kondo[b] and Ryo Kurazume[b]

[a]Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi, Japan
[b]Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, Japan

## ABSTRACT

This paper presents a new object classification technique for 3D point cloud data acquired with a laser scanner. In general, it is not straightforward to distinguish objects that have similar 3D structures but belong to different categories based only on the range data. To tackle this issue, we focus on laser reflectance obtained as a side product of range measurement by a laser scanner. Since laser reflectance contains appearance information, the proposed method classifies objects based on not only geometrical features in range data but also appearance features in reflectance data, both of which are acquired by a single laser scanner. Furthermore, we extend the conventional Histogram of Oriented Gradients (HOG) so that it couples geometrical and appearance information more tightly. Experiments show the proposed technique combining geometrical and appearance information outperforms conventional techniques.

**Keywords:** 3D object classification, Laser scanner, Laser reflectance, HOG

## 1. INTRODUCTION

The recent development of range measurement technology has provided accurate 3D modeling of real objects, and a variety of range sensors have been released including RIEGL VZ-400 (RIEGL GmbH), SwissRanger SR4100 (MESA Imaging AG), HOKUYO TOP-URG (HOKUYO), and Kinect (Microsoft). At the same time, much attention has been attracted to 3D object classification[1].[2] 3D object classification is the task of annotating objects in 3D point cloud data captured by range sensors, and it is one of the important issues in the field of robotics and computer vision. For example, semantic labeling of 3D point cloud data, which classifies input 3D point clouds into several categories, enables us to generate an environmental map with a label for each point cloud in it, and robots can perform service tasks safely and efficiently referring to this labeled environmental map.

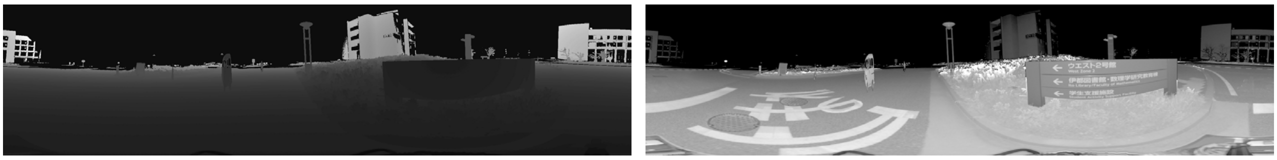
For landscape surveying or digital 3D modeling in outdoor environments, laser scanners such as TOPCON GLS-1500 (TOPCON) and Leica Scan Station 2 (Leica Geosystems AG) have been widely used. Though these laser scanners (LIDAR) provide high-resolution and accurate range data, it is not straightforward to distinguish objects that have similar 3D structures but belong to different categories based only on the range data. A simple solution to this problem is to utilize appearance information captured with a digital camera. However, since range and camera images are generally captured with individual sensors, a calibration process between these sensors must be performed. In addition, camera images are susceptible to variation of illumination, and the image quality is sometimes low due to bad light conditions.

In this paper, as an alternative to a texture image taken with a digital camera, we exploit **"laser reflectance"** from laser scanners for appearance information. Laser scanners generally obtain the range data of a target object according to the round-trip time of emitted laser pulses. As a side product of the range measurement, the laser reflectance that indicates the strength of the laser pulse reflected on a surface of a target object is available. The laser reflectance contains appearance of a target object under a single-frequency light source. Moreover, since range and reflectance images that depict range and reflectance values as gray scale images are captured in the same optical system, both images are fundamentally aligned.

Focusing on the property of the reflectivity, we introduce a novel object classification method with a single laser scanner. The proposed method describes objects appearances with the Histogram of Oriented Gradients (HOG) feature[3]

---

Further author information: (Send correspondence to S.Oishi)
S.Oishi: oishi@cs.tut.ac.jp
R.Kurazume: kurazume@ait.kyushu-u.ac.jp



(a) Range image (b) Reflectance image

Figure 1. Panoramic range and reflectance images taken with a laser scanner

from reflectance images. In addition, we develop **"3DHOG"** extending the conventional HOG feature according to the corresponding range data so that the appearance of a target object can be represented in 3D space.

The paper is structured as follows: A brief survey of previous approaches will be presented in Section 2. In Section 3, we will describe the proposed approach with the laser reflectivity in detail. In Section 4, the validity of the proposed method is demonstrated through some experiments using actual range and reflectance images obtained with a laser scanner.

## 2. RELATED WORK

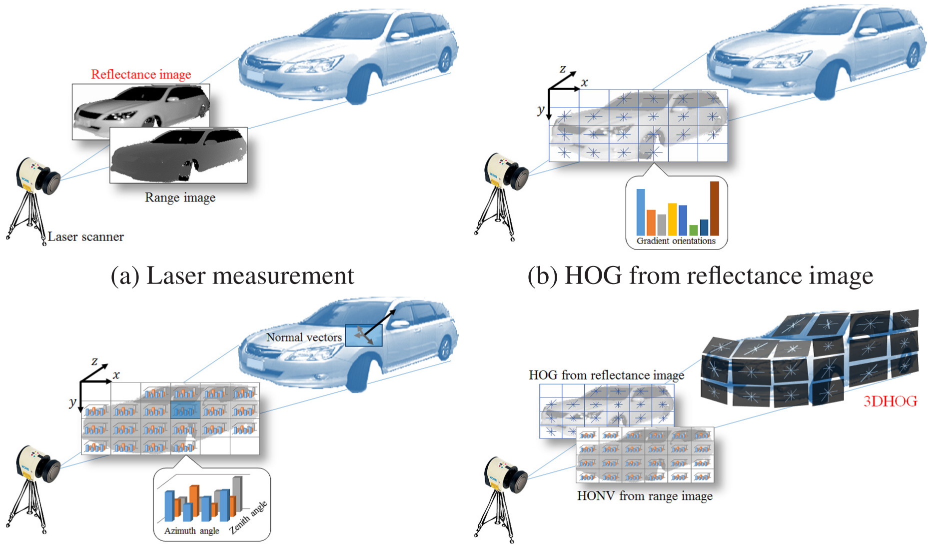
Scene analysis via object classification is an important task and has been widely studied for navigation robots or autonomous vehicles. Zhu et al.[4] proposed a segmentation and classification method for range data obtained with a laser scanner to annotate class labels to the point clouds. They segment each object in an input range image applying the Felzenszwalb and Huttenlocher (FH) algorithm,[5] and then extract statistical features such as variance of normal vectors and size of the segment, and classify the objects with SVM and CART. Xiong et al.[6] also introduced a semantic labeling method to 3D laser scans using a sequential parsing procedure. The proposed technique uses k-means++[7] for point cloud clustering, and classifies objects in point cloud data based on their structures and contextual information from neighbor clusters. They also developed a multi-stage inference procedure to capture these relationships instead of a graphical model to suppress computational time and provided good results in various environments even with different scanning systems.

On the other hand, it is also effective to combine the point cloud data and texture images taken with color sensors. Aijazi et al.[8] developed a simple segmentation and classification technique using range data and color values. In this technique, the input point clouds are first clustered as small voxels, and then each voxel is transformed into a super-voxel by integrating these voxels based on their properties such as the centroids, variances and color values. Each super-voxel is classified based on the properties with a set of standard, predefined thresholds. Posner et al.[9] also combined geometric features from a 3D laser scanner and appearance features from a digital camera for semantic annotation of urban maps. They first divide the 3D point cloud data into local plane patches applying RANSAC, and then segment each patch into several areas again based on color similarities. This process provides visually similar image patches with geometric structures. Finally, they define an efficient contextual inference framework with a generative bag-of-words classifier, and recognize objects based on a graph over the patches. These techniques achieved high classification accuracies by utilizing range and texture data simultaneously. However, range and color images are basically measured with independent sensors from different points of view, and the registration error between these images could have negative effects on object classification. Additionally, it is difficult to capture similar images under different lighting conditions even in the same scene.

In this paper, we utilize reflectance images as alternatives to texture images for appearance information, and further define the "3DHOG" feature to extract edges on a target object in 3D space. Since reflectance images are obtained as side products of range measurement, they don't need any registration processes and aren't affected by lighting conditions.

Note that several researchers have already proposed methods called "3DHOG" that extend the HOG feature into 3D domain[10].[11] A method in[11] was proposed for traffic surveillance in video data and extracts 3D appearance of a target object, but it needs to prepare full 3D models for all categories in advance. A 3DHOG extraction method for 3D mesh models was also proposed in.[11] However, this method was designed to describe histograms of "geometric" gradients of the target object in 3D domain, and it doesn't provide the appearance information.

In addition, the reflectance information has been used for other applications such as vegetation detection[12] or localization[13] so far. On the other hand, this paper addresses multi-class object classification focusing on the reflectance image. The major contribution of this paper is to represent target objects' appearances in a 3D domain combining the range data and the laser reflectivity, and achieve higher recognition accuracy exploiting it.




(a) Laser measurement (b) HOG from reflectance image

(c) HONV from range image (d) 3DHOG

Figure 2. Feature extraction

## 3. OBJECT CLASSIFICATION USING RANGE AND REFLECTANCE DATA

The proposed method combines the range and reflectance images from a laser scanner for object classification. In our method, range and reflectance images are transformed into histograms of oriented normal vector (HONV)[14] and oriented gradient (HOG),[3] and combined into a single feature vector.

In addition, we extend the conventional HOG so that the appearance of the target object can be represented in 3D domain. The conventional HOG feature is able to represent 2D object appearances robustly and provides good recognition performance. However, we should consider a geometric structure of the target object in order to describe the appearance more accurately because objects with similar appearances do not always have similar structures. In this paper, we develop the **"3DHOG"** feature by transforming the HOG feature into a 3D domain referring to HONV.

In this section, we first go over the property of reflectance image, the overview of the proposed method, and the algorithms of the conventional HOG and HONV features. Then the concept and the algorithm of the proposed "3DHOG" are described.

### 3.1 REFLECTANCE IMAGE

A time-of-flight range sensor, such as a laser scanner, acquires range data by measuring the round-trip time of a laser pulse reflected on an object surface. Figure 1(a) shows an example of a range image acquired by a laser scanner. On the other hand, most optical range sensors record the strength of the reflected laser pulse (reflectivity).

Reflectance values indicate an intensity on the surface points of the targets under a single-frequency light source. Figure 1(b) shows an example of a reflectance image that depicts reflectance values as a gray-scale image. As mentioned above, the reflectance value is determined uniquely for each pixel in the range image. In other word, the range image and the reflectance image are fundamentally aligned.

### 3.2 PROPOSED METHOD

We propose two kinds of classification methods for 3D objects utilizing range and reflectance images (Fig.2(a)). The overview of the proposed methods is summarized as follows.

Method 1 : Appearance description with HOG and reflectivity

1. Measure range and reflectance data with a laser scanner.

2. Create gray-scale range and reflectance images in which gray values of each pixel are proportional to the measured range and reflectance values.

3. Calculate HONV (Section 3.4) from the range image to extract the geometric feature.

4. Calculate HOG (Section 3.3) from the reflectance image to describe the appearance feature.

5. Classify the target object using both features with a Support Vector Machine (SVM).

Method 2 : 3D appearance description extending HOG

To represent the appearance of the target object in 3D domain, an additional procedure is exploited before the classification (5th step in the Method 1).

4. Calculate 3DHOG (Section 3.5) which transforms the HOG feature into 3D domain according to Histogram of Oriented Normal Vectors (HONV) derived from the corresponding range image.

### 3.3 APPEARANCE DESCRIPTION WITH HOG

Histogram of Oriented Gradients (HOG) is a descriptor introduced by Dalal et al.[3] for pedestrian detection, and it has been widely used as one of the most popular features for object recognition. It is able to describe object appearances robustly as local distributions of gradient orientation of the intensities (Fig.2(b)).

As the first step to extract HOG feature, the gradient and its strength at each pixel in an input image is calculated. The second step of the calculation includes decomposing an input image into $k \times k$ small rectangular cells and creating a histogram of oriented gradients at each cell. The gradient orientation is discretized into $N_\alpha$ possible angles (bins) evenly over 0 to 180 degrees, and we compute the histogram at each cell by casting a weighted vote for the corresponding bin according to the gradient and strength of each pixel within the cell. Finally, HOG features are obtained through block-normalization where $q \times q$ cells are groped into a block and all histograms in it are normalized. By repeating the block normalization for overlapping blocks, HOG can extract a feature invariant to lighting condition and small deformation.

### 3.4 NORMAL VECTOR EXTRACTION WITH HONV

Histogram of Oriented Normal Vector (HONV) proposed by Tang et al.[14] was designed as a feature which described geometric characteristics of a 3D object captured with a depth sensor (Fig.2(c)). Since HONV represents the local 3D geometry of the object as a local distribution of the normal vector orientations, we first have to calculate a normal vector at each pixel $p = (x, y)$ in a range image. The normal vector $N$ can be obtained as follows:

$$N = \begin{pmatrix} -\frac{\partial d(x,y)}{\partial x} \\ -\frac{\partial d(x,y)}{\partial y} \\ 1 \end{pmatrix} \tag{1}$$

where $d(x, y)$ denotes a depth value at pixel $p$. Then, the direction of the extracted normal vector is represented in spherical coordinates with the zenith angle $\theta$ and the azimuth angle $\varphi$:

$$\theta = tan^{-1}\sqrt{\left(\frac{\partial d(x,y)}{\partial x}\right)^2 + \left(\frac{\partial d(x,y)}{\partial y}\right)^2} \tag{2}$$

$$\varphi = tan^{-1}\left(\frac{\partial d(x,y)}{\partial y} / \frac{\partial d(x,y)}{\partial x}\right) \tag{3}$$

Finally, HONV can be calculated by the same procedure as HOG extraction; Decomposing the range image into $k \times k$ cells, and creating a histogram by voting for the corresponding bin in $N_\theta \times N_\varphi$ discretized angles according to the zenith angle $\theta$ and the azimuth angle $\varphi$ of each pixel within the cell.

## 3.5 THREE-DIMENSIONAL HISTOGRAM OF ORIENTED GRADIENTS : 3DHOG

This section introduces the new descriptor for 3D object recognition in point cloud data obtained with a laser scanner. In recent years, a variety of descriptors has been developed by many researchers for 2D object recognition in an image, such as SIFT,[15] SURF,[16] HOG,[3] and so on. These features give excellent results in object detection/recognition in the 2D image domain. However, every object has not only its appearance but also 3D shape, and it is not straightforward to describe the appearance with 2D features.

Therefore, in this paper, we extend the HOG feature so that the appearance of a target object can be represented as a distribution of the oriented gradient in 3D space combining range and reflectance images obtained with a laser scanner (Fig.2(d)). The key idea behind this approach is as follows: Since range and reflectance images are fundamentally aligned, we first extract the conventional HOG feature from a reflectance image to describe the appearance of a target object, and then transform the feature into 3D domain according to Histogram of Oriented Normal Vectors (HONV) derived from the corresponding range image.

Algorithm 1 shows the pseudo code of the proposed 3DHOG. HONV holds a normal vector distribution at each small cell, therefore it is able to project HOG at the corresponding cell in the reflectance image onto the surface of the target object by rotating the normal vector of the reflectance image $N_i$ to the normal vectors in HONV $N_{\theta,\varphi}$. Note that the normal vector $N_i$ is set as (0, 0, -1) (Fig.2(b)).

## 3.6 CLASSIFICATION

In this work, we adopt a supervised learning approach and utilize a linear support vector machine as a classifier. Based on the feature vector extracted from range and reflectance data, it determines the category of the target object and gives the corresponding label.

# 4. SEMANTIC LABELING TO ENVIRONMENTAL MAP

This section describes experimental results of 3D object classification for semantic labeling. Objects in this experiments consist of 5 classes, that is, Building, Car, Human, Tree, and Pole (Fig.3). We tested three kinds of feature combinations: {HONV}, {HONV, HOG}, and {3DHOG}. Note that we repeated learning and recognition processes using the leave-one-out cross-validation method to evaluate the performances.

Range and reflectance images were captured with a laser scanner (SICK, LMS151) and a rotating table, and each image had a resolution of $760 \times 1135$ points. The apramelers are set as $N_c = 8 \times 8$, $S_b = 2 \times 2$, $N_\alpha = 18$, $N_\theta = 9$, $N_\varphi = 9$. Consequently, the dimensions of the feature combinations are {{HONV}, {HONV, HOG}, {3DHOG}} = {15876, 19404, 15876}.

## 4.1 GROUND REMOVAL AND SEGMENTATION

Before extracting features of target objects in the scan data, segmentation of point clouds must be performed to determine a point cloud cluster of each object. As the first step of the segmentation, we remove the ground part in the input point clouds. By assuming the ground is flat, we can extract it by fitting a plane to the points lower than the position of the laser scanner. This process enables us to isolate objects in the point cloud data and make it easy to distinguish them.

After the ground removal, we then detect each object's cluster utilizing the enhanced radially bounded nearest neighbor.[17] In this method, point clouds are assigned to a cluster if they are located in a sphere of a certain radius. The radius is proportional to the distance between the scanned objects and a laser scanner because the density of point cloud data reduces with the distance. The distance-varying radius enables us to suppress under-segmentation or over-segmentation compared with the case with a fixed radius. Example range and reflectance images segmented with this method are shown in Fig.3.

## 4.2 CLASSIFICATION

Table 1 shows the classification results for semantic labeling. As shown in this table, HONV feature provides good recognition performance, However, it failed to recognize objects which have similar shapes but belong to different categories. On the other hand, the combination of HONV and HOG (proposed method 1) gives higher performance than only HONV. Moreover, the proposed 3DHOG (proposed method 2) successfully classified objects in the input data with the highest recognition accuracy. The environmental map labeled by 3DHOG is shown in Fig.4. As shown in Table 1 and Fig.4, we can see that the 3DHOG has quite high classification performance by combining range and reflectance information.

**Algorithm 1:** Three-dimensional HOG

**Input**:

– Range and reflectance images
– The number of cells $N_c (= k \times k)$
– Block size $S_b (= q \times q)$
– The number of bins for HOG $N_\alpha$
– The number of bins for HONV $N_\theta, N_\varphi$

**Output**:

– Three-dimensional HOG which contains:
1) 3dgradArr - array of discretized 3D gradient vectors
2) 3dgradStr - array of the strength of each 3D gradient

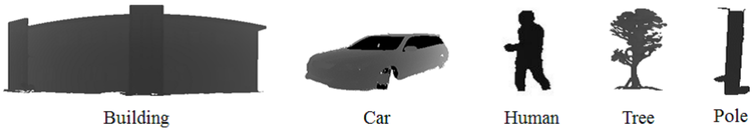
**Definision:**

– A HOG structure contains:
1) gradArr - array of discretized gradient vectors
2) gradStr - array of the strength of each gradient

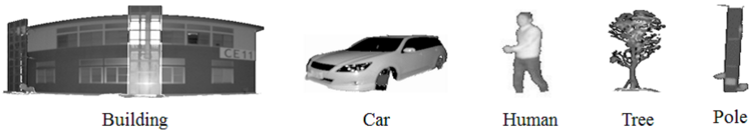
– A HONV structure contains:
1) normArr - array of discretized normal vectors
2) normStr - array of the strength of each normal vector

**The algorithm:**

1) divide range and reflectance images into small cells
2) create HOG[$N_c$] and HONV[$N_c$] structures for each cell
3) **for** *each cell $c$* **do**
    A) calculate HOG[$c$] from the reflectance image
    B) calculate HONV[$c$] from the range image
    C) **for** *each bin $u_{\theta,\varphi}$ in HONV[c]* **do**
        a) calculate a projection matrix A from $N_i$ to HONV[$c$].normArr[$u_{\theta,\varphi}$]
        b) **for** *each bin $u_\alpha$ in HOG[c]* **do**
            i) calculate 3D edge by multipling the matrix A and HOG[$c$].gradArr[$u_\alpha$]
            ii) create 3DHOG histogram by casting a weighted vote according to
            HOG[$c$].gradStr[$u_\alpha$]$\times$HONV[$c$].normStr[$u_{\theta,\varphi}$]
            for the corresponding bin
            3DHOG[$c$].3dgradStr[$u'_{\theta,\varphi}$]
4) normalize 3DHOG at each block




(a) Range images




(b) Reflectance images

Figure 3. Categories of objects in environmental maps

Table 1. Classification results of semantic labeling

| | HONV | | | | | |
|---|---|---|---|---|---|---|
| | Building | Car | Human | Tree | Pole | Total |
| Images | 103 | 104 | 124 | 131 | 28 | 490 |
| Detected | 105 | 101 | 124 | 131 | 29 | |
| Correct | 101 | 98 | 121 | 127 | 27 | 474 |
| Precision | 96.2% | 97.0% | 97.6% | 96.9% | 93.1% | |
| Recall | 98.1% | 94.2% | 97.6% | 96.9% | 96.4% | 96.7% |

| | HONV and HOG | | | | | |
|---|---|---|---|---|---|---|
| | Building | Car | Human | Tree | Pole | Total |
| Images | 103 | 104 | 124 | 131 | 28 | 490 |
| Detected | 102 | 103 | 124 | 132 | 29 | |
| Correct | 99 | 100 | 121 | 129 | 27 | 476 |
| Precision | 97.1% | 97.1% | 97.6% | 97.7% | 93.1% | |
| Recall | 96.1% | 96.2% | 97.6% | 98.5% | 96.4% | 97.1% |

| | 3DHOG | | | | | |
|---|---|---|---|---|---|---|
| | Building | Car | Human | Tree | Pole | Total |
| Images | 103 | 104 | 124 | 131 | 28 | 490 |
| Detected | 105 | 99 | 127 | 131 | 28 | |
| Correct | 102 | 99 | 124 | 129 | 27 | 481 |
| Precision | 97.1% | 100% | 97.6% | 98.5% | 96.4% | |
| Recall | 99.0% | 95.2% | 100% | 98.5% | 96.4% | 98.2% |

## 5. CONCLUSION

This paper introduced novel object classification techniques combining range and reflectance information obtained with a single laser scanner. Since the laser reflectivity, which is obtained as a side-product of the range measurement, shows an intensity on the surfaces points of a target under a single-frequency light source, we can extract the appearance feature by taking advantage of it. Moreover, since range and reflectance images are considered to be fundamentally aligned, we developed the 3D HOG that can describe appearance of a target object in 3D space by transforming HOG from a reflectance image into the 3D domain based on the corresponding normal vector derived from the range image. We carried out experiments using actual range and reflectance dataset, and the results showed that the laser reflectivity enabled us to achieve higher recognition accuracy.

Though the proposed method utilized reflectance images to describe 3D texture edges on objects, it can be extended to texture images taken with digital cameras if they are calibrated to the range images. Therefore, as future work, we will apply the proposed methods to RGB-D datasets[18] to evaluate the performances.

## ACKNOWLEDGMENTS

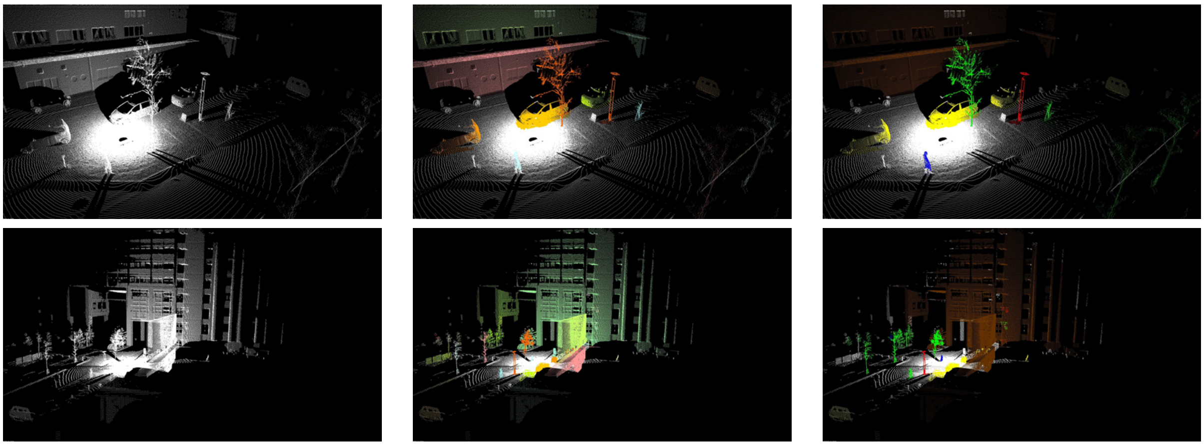
The present study was supported in part by a Grant-in-Aid for Scientific Research (A) (26249029).

## REFERENCES

[1] Johnson, A. and Hebert, M., "Using spin images for efficient object recognition in cluttered 3d scenes," in [*IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*], **21**(5) (1999).

[2] Triebel, R., Kersting, K., and Burgard, W., "Robust 3d scan point classification using associative markov networks," in [*IEEE International Conference on Robotics and Automation*], (2006).

[3] Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection," in [*Computer Vision and Pattern Recognition*], **Vol.1** (2005).



(a) Point cloud data (b) Segmentation (c) Semantic labeling

Figure 4. Semantic labeling using 3DHOG (proposed method 2) : (a) Point cloud data obtained with a laser scanner and a rotation table. (b) Segmentation result applying the enhanced radially bounded nearest neighbor[17] (c) Semantic labeling result using the proposed method 2. Note that each object in the scene is colored according to the classification results; Building:Brown, Car:Yellow, Human:Blue, Tree:Green, Pole:Red, White:The ground and under-segmented clusters.


[4] Zhu, X., Zhao, H., Liu, Y., Zhao, Y., and Zha, H., "Segmentation and classification of range image from an intelligent vehicle in urban environment," in [*IEEE/RSJ International Conference on Intelligent Robots and Systems*], (2010).

[5] Felzenszwalb, P. F. and Huttenlocher, D. P., "Efficient graph-based image segmentation," *Int. J. Comput. Vision* **59**, 167–181 (sep 2004).

[6] Xiong, X., Munoz, D., Bagnell, J. A. D., and Hebert, M., "3-d scene analysis via sequenced predictions over points and regions," in [*IEEE International Conference on Robotics and Automation (ICRA)*], (May 2011).

[7] Arthur, D. and Vassilvitskii, S., "K-means++: The advantages of careful seeding," in [*Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*], 1027–1035 (2007).

[8] Aijazi, A. K., Checchin, P., and Trassoudaine, L., "Segmentation based classification of 3d urban point clouds: A super-voxel based approach with evaluation," *Remote Sensing* **5(4)**, pp.1624–1650 (2013).

[9] Posner, I., Cummins, M., and Newman, P., "Fast probabilistic labeling of city maps," *Robotics: Science and Systems* (June 2008).

[10] Scherer, M., Walter, M., and Schreck, T., "Histograms of oriented gradients for 3d object retrieval," in [*18th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*], pp.41–48 (2010).

[11] Buch, N., Orwell, J., and Velastin, S. A., "3d extended histogram of oriented gradients (3dhog) for classification of road users in urban scenes.," in [*British Machine Vision Association*], (2009).

[12] Wurm, K. M., Kretzschmar, H., K • merle, R., Stachniss, C., and Burgard, W., "Identifying vegetation from laser data in structured outdoor environments," *Robotics and Autonomous Systems* **5(4)**, pp.1624–1650 (2012).

[13] Levinson, J. and Thrun, S., "Robust vehicle localization in urban environments using probabilistic maps," in [*IEEE International Conference on Robotics and Automation*], 4372–4378 (2010).

[14] Tang, S., Wang, X., Lv, X., Han, T. X., Keller, J., He, Z., Skubic, M., and Lao, S., "Histogram of oriented normal vectors for object recognition with a depth sensor," in [*The 11th Asian Conference on Computer Vision*], (2012).

[15] Lowe, D. G., "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision* **60**, 91–110 (nov 2004).

[16] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L., "Speeded-up robust features (surf)," *Comput. Vis. Image Underst.* **110**, 346–359 (jun 2008).

[17] Choe, Y., Ahn, S., and Chung, M. J., "Fast point cloud segmentation for an intelligent vehicle using sweeping 2d laser scanners," in [*9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*], pp.38–43 (2012).

[18] Lai, K., Bo, L., Ren, X., and Fox, D., "A large-scale hierarchical multi-view rgb-d object dataset.," in [*IEEE International Conference on Robotics and Automation*], 1817–1824, IEEE (2011).