Segmentation and classification of road markings using MLS data.

Mario Soilán[a,*], Belén Riveiro[b], Joaquín Martínez-Sánchez[a], Pedro Arias[a]

[a]*Department of Natural Resources and Environmental Engineering, School of Mining Engineering, University of Vigo, 36310, Vigo, Spain*

[a]*Department of Materials Engineering, Applied Mechanics and Construction, School of Industrial Engineering, University of Vigo, 36310, Spain*

Abstract

Traffic signs are one of the most important safety elements in a road network. Particularly, road markings provide information about the limits and direction of each road lane, or warn the drivers about potential danger. The optimal condition of road markings contributes to a better road safety. Mobile Laser Scanning technology can be used for infrastructure inspection and specifically for traffic sign detection and inventory. This paper presents a methodology for the detection and semantic characterization of the most common road markings, namely pedestrian crossings and arrows. The 3D point cloud data acquired by a LYNX Mobile Mapper system is filtered in order to isolate reflective points in the road, and each single element is hierarchically classified using Neural Networks. State of the art results are obtained for the extraction and classification of the markings, with F-scores of 94% and 96% respectively. Finally, data from classified markings are exported to a GIS layer and maintenance criteria based on the aforementioned data are proposed.

Keywords: Mobile laser scanning, point cloud segmentation, road marking, classification, smart cities.

1. Introduction

While driving a vehicle, a human driver perceives numerous visual information that indicates potential dangers, speed limits, or road layout among others. Even though road markings may not be the most important type of traffic sign, they are crucial for pedestrian safety in populated environments. An optimal positioning and condition of pedestrian crossings will definitely help to reduce and prevent run-over accidents.

The automatic detection of road markings has been used for autonomous vehicle guidance (Vacek et al., 2007; Hoffmann et al., 2007) or for Lane Departure Warning (LDW) systems, which are a common active safety feature in commercial cars (Kibbel et al., 2005) that prevent involuntary lane departure (Rimini-Doering et al., 2005). Furthermore, this automation may assist the maintenance and inventory tasks reducing both the cost of the process and the subjectivity of the results, as these activities are normally carried out by inspectors. The inspectors are who subjectively report the maintenance needs of a surveyed road based on a technical specification. For Spanish roads, road marking specifications are defined by Ministerio de Fomento (1987).

There are several studies in the literature that deal with road marking detection. Typically, computer vision techniques have been applied to line tracking (McCall and Trivedi, 2006) and road object classification (Danescu and Nedevschi, 2010) using RGB images and video. Road marking detection in images is a challenging task, as there are a number of factors that determine the robustness of an algorithm, namely weather conditions, pavement material and color, shadowing and lighting changes, etc. Therefore, there is a need to study the robustness of different data sources for this task.

* Corresponding author.

E-mail address: msoilan@uvigo.es

Nowadays, one of the most popular sources of data for infrastructure surveying is laser scanning. These instruments are able to collect geometric and radiometric properties of their surroundings in a cost-efficient, reliable manner. When mounted on a Mobile Mapping System (MMS), laser scanners collect dense, accurate 3D point clouds driving at standard speeds (Puente et al., 2013; Ussyshkin, 2009).

Although it is still an active research topic, promising results have been achieved using Mobile Laser Scanning (MLS) technology, transforming unorganized 3D data into a meaningful set of segments. Yang and Dong (2013) segment point cloud scenes using Support Vector Machines (SVM) for classifying each point as linear, planar or spherical, based on features extracted from a Principal Component Analysis (PCA) of the optimal neighborhood of each point. Points within the same class are subsequently merged and refined into significant segments. Vo et al. (2015) developed an octree-based region growing which segments a point cloud in a coarse-to-fine process. First, it groups together points on a voxelized space that have similar normal vectors and residual values, and then performs a refinement that studies the unallocated points and the boundaries of the previously computed segments. Different classification techniques are used for giving a semantic meaning to each point cloud partition. Serna and Marcotegui (2014) computed geometric, contextual and color features of segmented objects and classified them in 20 classes (cars, bollards, fences, etc.) using hierarchically trained SVM. Although SVM are extensively used for remote sensing applications (Mountrakis et al., 2011), other authors use heuristics for classifying point cloud segments. In (Yang et al., 2015) seven urban objects are classified from the saliency and the geometry of each segment, and in (Pu et al., 2011) pole objects and five classes of shapes are distinguished by computing their Minimum Bounding Rectangle (MBR), Minimum Bounding Circle (MBC) and convex hull, and studying their geometrical relationships. The segmentation-classification approach is employed for road marking extraction as well. The only relevant segment in this case is the road, therefore ground segmentation is the usual first step. Douillard et al. (2011) computed height and vertical variance of each voxel in a 3D grid, and segmented the ground based on local differences between adjacent voxels. Serna and Marcotegui (2013) use the λ-flat zones labeling algorithm to segment the ground and defines it as the largest flat zone in a range image that contains height information of the point cloud. Guan et al. (2014) detect curbs in consecutive profiles of a point cloud and delineate the limits of the road using the curbs as road limits. Although it performs robustly in roads with curbs, the performance of this segmentation algorithm in roads without curbs remains unclear. In Guan et al. (2014) road markings are extracted from geo-referenced intensity images using Inverse Distance Weighted (IDW) interpolation and a point density dependent thresholding. As road markings are reflective surfaces, the intensity feature is commonly used for creating raster images where road markings can be extracted. For instance, Kumar et al. (2014) applied a range dependent thresholding process and morphological operations over an intensity image, and Chen et al. (2009) used Hough Transform Clustering for detecting lane markings. Recently, Yu et al. (2015) extracted road markings directly from the 3D point cloud via multisegment thresholding and Otsu binarization, classified them in four classes using Deep Boltzmann Machines (DBM) and, in a second level of hierarchy, in seven classes of road markings comprising arrows, pedestrian crossings or stop lines.

In this paper, a method for road marking extraction and classification from MLS data is proposed. Furthermore, with the aim at assessing and updating the state of the road markings in the context of a Smart City, the outputs of the proposed method are exported to a Geographic Information System (GIS) layer, and criteria for updating the road marking inventory is proposed, moving a step forward with respect to related work. The methodology is described in detail in Section 2. Section 3 presents the study case and the Mobile Mapping System characteristics. The performance of the method and its comparison with related work is summarized in Section 4, and finally the conclusions and future work are outlined in Section 5.

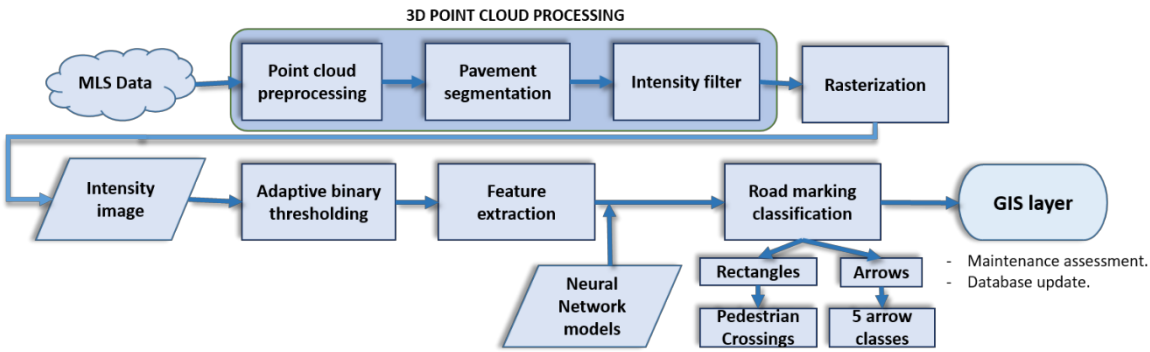

Figure 1. Overview of the method.

2. Methodology

An overview of the method is shown in Figure 1. First, the pavement is segmented and an intensity filter is applied in order to extract reflective pavement points from the point cloud. Subsequently, a raster image is created and road markings are detected after several image processing steps. Two different sets of features are computed for each marking, which is classified in two levels of hierarchy, distinguishing pedestrian crossings and five classes of arrows. Finally, the data from each classified road marking is exported to a GIS layer, where maintenance assessment can be conducted.

2.1. Pavement segmentation

In Section 1 different road marking extraction approaches were remarked, all of them having in common a ground segmentation step. The road pavement is the only segment that contains relevant information. Furthermore, filtering out off-ground information reduces computational load and saves memory resources.

First, the raw 3D point cloud is preprocessed, filtering out points further than 10 meters from the trajectory. That is, distant and noisy points are removed from the point cloud and only the data that are close to the trajectory are kept. This step reduces the size of the point cloud and it will likely remove data for adjacent streets, therefore the analysis will cover only the streets on which the MMS is moving along.

Let $\boldsymbol{P} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{I}, \boldsymbol{\alpha}, \boldsymbol{t_s})$ be the preprocessed point cloud, where $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ are the 3D coordinates, $\boldsymbol{I}$ is the reflected laser pulse intensity, $\boldsymbol{\alpha}$ is the angle of the laser beam and $\boldsymbol{t_s}$ is the time stamp; and let $\boldsymbol{T} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t_s})$ be the trajectory recorded by the navigation system of the vehicle.

First, the point cloud is partitioned in several sections $\boldsymbol{S}_i (i = 1 \dots n)$. Defining $\boldsymbol{L_s}$ as the section length, a subset $\boldsymbol{N}_s = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ with n coordinates of $\boldsymbol{T}$ is computed so that the distance between each $\boldsymbol{N}_s(i)$ and $\boldsymbol{N}_s(i + 1)$ in the direction of the trajectory is approximately $\boldsymbol{L_s}$ (the exact distance will be determined by the spatial resolution of the trajectory). Then, the direction of the section is defined as $\boldsymbol{v} = \boldsymbol{N}_s(i + 1)$ - $\boldsymbol{N}_s(i)$, and $\boldsymbol{S}_i$ is extracted as a subset of points in $\boldsymbol{P}$ within the transversal sections defined by $\boldsymbol{N}_s(i)$ and $\boldsymbol{N}_s(i + 1)$ along the vector $\boldsymbol{v}$ (Figure 2a).

Secondly, a curb-based segmentation approach based on Wang et al. (2015) is proposed. This method constructs a saliency map on 3D point clouds by computing the dominant normal vector of the cloud via K-means clustering, and then projecting the distance between each normal vector and the dominant normal vector

into a hyperbolic tangent function space, so the difference between a salient and a non-salient point is large enough to be easily classified. In order to efficiently apply the method, each section $\boldsymbol{S}_i \subset \boldsymbol{P}$ is voxelized, that is, a cubic grid with size g_s is defined and a voxel coordinate $(x^V, y^V, z^V)_i$ is assigned to each point $(x_i, y_i, z_i) \in \boldsymbol{S}_i$ following Equations (1)-(4).

$$x_i^V = \frac{x_i - \min(\boldsymbol{x})}{g_s} \tag{1}$$

$$y_i^V = \frac{y_i - \min(\boldsymbol{y})}{g_s} \tag{2}$$

$$z_i^V = \frac{z_i - \min(\boldsymbol{z})}{g_s} \tag{3}$$

$$id_i^V = x_i^V + N_x y_i^V + N_x N_y z_i^V \tag{4}$$

where indices id_i^V link each point in the cloud with its coordinate in the voxel grid space. Therefore, several features based on the points within each populated voxel can be computed. Specifically, the centroids of the points within each voxel define a point cloud where a 3D point represents a single voxel. Salient points (i.e. curbs, walls, poles, etc.) are found in this cloud following the method in Wang et al. (2015) (Figure 2b).

Finally, the pavement is defined as a group of non-salient points delimited by salient points at both sides of the trajectory. Each section $\boldsymbol{S}_i$ is partitioned in a number of transversal profiles of length L_p, and the closest salient points with respect to the trajectory on its right and its left are selected. Outliers from each group of points are removed using KSE method (Jirachan and Piromsopa, 2015), and two polynomial curves are fitted, defining the pavement limits on each side of the road (Figure 2c).

There are two parameters which influence this process. The section length L_s has an impact on the amount of memory resources needed during the implementation of the algorithm, and L_p influences on the adjustment of the pavement limits. The parameters for the case study presented in Section 3, which have been experimentally chosen, are $L_s = 10m, L_p = 50cm$. Regarding the cubic grid size g_s, note that the proposed

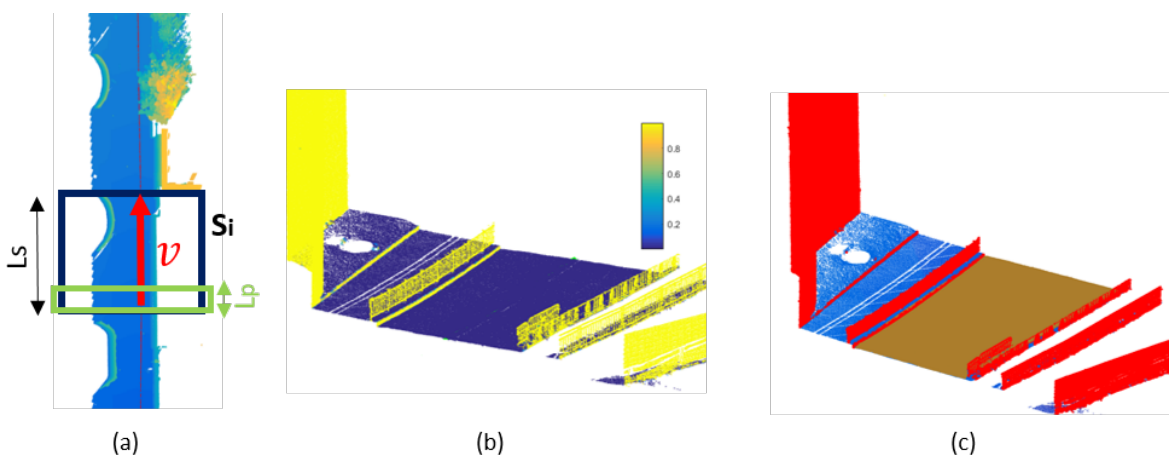

Figure 2. Pavement segmentation. (a) The point cloud is partitioned in several sections S_i of length L_s along the trajectory. (b) Saliency map of the voxelized point cloud of a section S_i. The algorithm enlarges the difference between salient and non-salient points. (c) The pavement (brown points) is segmented as a group of non-salient points between salient points (coloured in red) at both sides of the trajectory.

method computes the normal vectors for each point in the cloud. The voxelization aims for the downsampling of the input point cloud. Therefore, in order to speed up the calculation without losing local information (e.g. normal vectors of points in curbs) a grid size of $5cm$ is chosen.

2.2. Intensity filter

Once the pavement segment is extracted, the next step aims to isolate the road markings, which are reflective surfaces due to the properties of the materials used for painting them. Therefore, the reflectivity difference between asphalt and road markings is considered for this filtering step. The intensity distribution of a road that contains road markings can be separated in two main classes that approximate to Gaussian distributions. A class with the majority of the points and low intensity values, and a second class with a lower number of points and higher intensities. A Gaussian Mixture Model (GMM) (McLachlan and Peel, 2000) with two components is estimated for the intensity values. Each point in the cloud is assigned to the component with the largest posterior probability, and those points assigned to the component with lower intensity are filtered out from the cloud.

The purpose of this step is to obtain a point cloud with the minimum amount of data for creating intensity images. Otherwise, the pavement points that are removed in this step would introduce noise in the image. This statement is proven in Figure 3. Figures 3a-b show the raw pavement segment and the resulting intensity image (which is created as explained in Section 2.3). Figures 3c-d show the filtered point cloud and an intensity image where road markings are distinguishable in a clearer way.

2.3. Intensity image processing

In order to create an intensity based image, the filtered point cloud is projected onto the horizontal plane where a square grid is defined. Similarly to the voxelization process, each point is assigned to a cell in the grid, and the intensity values are used to compute a feature that can be visualized in an image. In this case, the selected feature is the normalized sum of the intensities for the set of points contained in each cell. The resulting grayscale image, as shown in Figure 3d, clearly differentiates road markings.

Before extracting features from markings, the grayscale image is binarized. Although the recorded intensity for two objects of the same material depends on the distance between the laser and the object, the angle of incidence and the atmospheric conditions (Höfle and Pfeifer, 2007) this work uses the intensity value directly provided by the measurement system, as the need for normalized intensity values (defining a reflectance model) is not as obvious as for Airborne Laser Scanner (ALS) data with strong elevation differences. However, some markings may have slightly smaller intensity values due to the aforementioned dependences of the intensity value, so a simple process is performed in order to correct this effect. Instead of computing a single threshold for the entire image, an adaptive thresholding based on the angle of the laser beam is used. First, the angle mean is computed for the set of points in each cell. Subsequently, several masks are computed selecting cells with similar angles, and the grayscale image is binarized independently for each mask using Otsu method (Otsu, 1979). Figure 4 shows this process and the improvement of an adaptive thresholding with respect to a single thresholding process.

The binary image is further processed in order to improve the quality of the markings. First, a connected components labelling algorithm is used. The number of '1' pixels in the binary component i, n_{true}^{i} is directly related to the area of the component given the 2D grid size of the raster, g_{s2D}. That is, the area (in units of g_{s2D}) of the component i will be approximately $area^{i} = n_{true}^{i} \cdot g_{s2D}^{2}$. All the road marking classes considered for classification have areas between 1.2 m^2 and 4 m^2. Thus, an area filter is set to remove elements in the binary

image such that $1\ m^2 < \ area^i < 5\ m^2$. Furthermore, the ratio between area and perimeter (measured in pixels) is used for filtering lane markings.

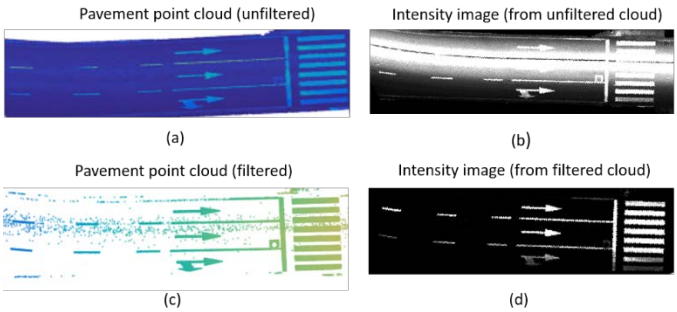

Figure 3. Intensity filter. The pavement point cloud (a) is projected into a 2D plane and an intensity based image is obtained (b). Road markings are not clearly distinguishable. By contrast, if an intensity filter is applied in the point cloud (c) the quality of the resulting image (d) increases significantly.

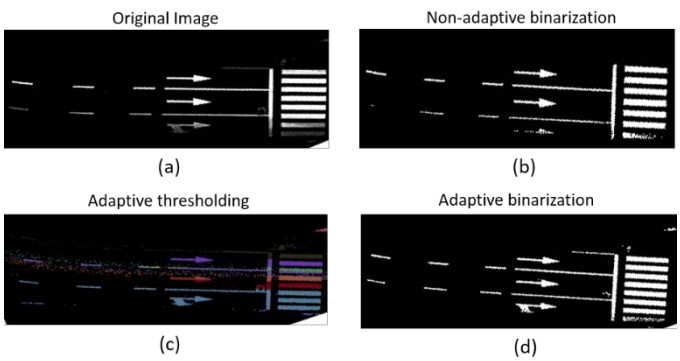

Figure 4. Intensity image processing. (a) Intensity based image. (b) Binary image using a single threshold. (c) Adaptive thresholding. Each group of pixels that contains points with similar scan angles will be analysed separately, with different thresholds. (d) Binary image using an adaptive threshold.

2.4. Feature extraction and road marking classification

The connected components in the resulting binary image (Figure 4d) can be analyzed and classified individually. For each marking, several discriminative features are computed based on the pixel distribution and on geometric parameters. Therefore, every road marking is defined with a feature vector that can be the input of a supervised machine learning algorithm which classifies the marking. The classification result has a semantic meaning that complements the road marking spatial information.

Based on the number of markings of the same class in the available dataset, and on their frequency of occurrence on the road, the following classes were defined:

- *Rectangular marking*. A group of parallel rectangular markings forms a pedestrian crossing. A single rectangular marking which is transversal to the trajectory forms a stop line.
- *Arrow*. They indicate the flow direction and the allowed turns in intersections. There exist several arrow classes.

- *Others / negatives*: Any other road marking or spurious element that does not belong to the two main classes. This class contains chevron markings, written information, give ways, and elements that are not road markings, e.g. manhole covers or other reflective surfaces that may lay on the pavement surface.

In order to distinguish between these three classes, a set of features is computed for each element in the binary image. Although limited, the information available in a binary set of pixels can be used to extract shape and geometric information. A Geometry Based Feature (GBF) is proposed, so that for each image J of a single road marking, its feature vector contains:

- *Area (A).* The number of '1' pixels in J can be converted to the approximated area of the road marking in square meters as stated in Section 2.3.
- *Bounding ellipse measurements* (l_1, l_2, r). The ellipse with the same second central moments than J is used for computing three features, namely the length of major and minor axes measured in meters (l_1, l_2), and the ratio between major axis and minor axis r.
- *Shape measurements (E, T, R).* The first affine moment invariant I_1 (Flusser and Suk, 1993) is used to compute the ellipticity (Equation 5) and the triangularity (Equation 6) of the binary element in J. Furthermore, the rectangularity is measured as the ratio between the area of the element and the area of the minimum bounding rectangle (Rosin, 2003).
- *Pixel distribution* ($\boldsymbol{pix}_{feat}$). The horizontal and vertical distribution of the pixels in the image is collected in two normalized histograms of 20 bins. Two semantically equivalent road marking images may have different size and orientation, thus some processing steps are needed to standardize this feature. This calculation is summarized on Algorithm I.

$$E = \begin{cases} 16\pi^2 I_1 & if\ I_1 \leq 1/(16\pi^2) \\ 1/(16\pi^2 I_1) & otherwise \end{cases} \quad (5)$$

$$T = \begin{cases} 108 I_1 & if\ I_1 \leq 1/108 \\ 1/(108 I_1) & otherwise \end{cases} \quad (6)$$

ALGORITHM I.
PIXEL DISTRIBUTION CALCULATION.

Data: Input image (*J*), major axis of the bounding ellipse angle with respect to x axis (*orientation*).
Output: Pixel distribution feature (pix_{feat}).
Functions: rotate(*I*, angle) rotates *angle* degrees the image I.
size(*I*) returns the number of rows and columns of the image I.

```
oriented_J = rotate(J, orientation);
[numberRows, numberCols] = size(J);
for each row i of oriented_J
        rowPix_i ← (number of white pixels in row i / numberRows);
end
for each column j of oriented_J
        colPix_j ← (number of white pixels in column j / numberCols);
end
rowHist ← normalized 20 bin histogram of rowPix;
colHist ← normalized 20 bin histogram of colPix;
rowFeature ← rowHist bin values;
colFeature ← colHist bin values;

return pix_feat = append(rowFeature, colFeature);
```

A 47-dimensional feature vector $\boldsymbol{GBF}_i$ is obtained for each marking (Equation 7). Six first features are extracted from shape measurement and forty features are extracted from pixel distribution.

$$\boldsymbol{GBF}_i = \left(A, l_1, l_2, r, E, T, R, \boldsymbol{pix}_{feat}\right)_i \tag{7}$$

For this work, Neural Networks (Bishop, 1995) are chosen as classification method. Specifically, a two-layer feedforward network with a hidden layer and an output layer is applied as it is one of the most common designs for pattern recognition. It is a supervised machine learning algorithm, thus a training dataset is required. The features extracted from two of the study case areas, which are defined in Section 3, are manually labeled and used for testing and cross-validating the Neural Network models. The models are built following a one-vs.-all multiclass perceptron strategy. Furthermore, the models are cross-validated in order to reduce the risk of over-fitting. Two different models were trained, to compare the performance of two different features: The aforementioned GBF, and Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2004).

Finally, every new road marking is automatically labelled by the Neural Network. This step develops a first level of hierarchy within the classification process, where the classified elements represent three generic classes. In a second hierarchy level, a proper semantic meaning is given to the road markings classified as rectangular markings or as arrows. First, rectangular markings are contextually analyzed. Let $\boldsymbol{P}_r = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\alpha}, \boldsymbol{t_s})$ be the 3D point cloud of a rectangular marking. The director vector of the marking, $\boldsymbol{v}_r$, is computed as the eigenvector that corresponds to the largest eigenvalue from the Principal Component Analysis (PCA) of $\boldsymbol{P}_r$. The direction of the vehicle trajectory $\boldsymbol{v}_t$ is analogously defined, selecting only those points within the time stamp range of $\boldsymbol{P}_r$ for PCA (Figure 5a). With these vectors, the angle between the road marking and the trajectory (α_{rt}) and the azimuth (angle with respect to the north) of the marking (α_{rn}) can be computed. Rectangular markings are merged into pedestrian crossings if the angle between two markings is less than a threshold (set to $th_{rn} = 10º$) and the distance between their centroids (d_r) is closer than $th_r = 2.5\ m$. For any marking that remains isolated, the angle α_{rt} is taken into account. The marking is considered a stop line if it is perpendicular with respect to the trajectory ($90º \pm \alpha_t$), and a false positive otherwise.

Regarding the arrows, the second level of hierarchy classifies them in five classes: Straight, Straight-left, Straight-right, Left and Right. First, the direction of the trajectory $\boldsymbol{v}_t$ is obtained with the same procedure than for rectangular markings. The angle of the trajectory with respect to the north (α_{an}) is used as rotation angle for the intensity image of each single arrow. Then, the resulting image is compared against a set of arrow models. Two parameters are used for this comparison: Correlation coefficient (Equation 8) and Structural Similarity Index (SSIM) (Wang et al., 2004).

$$corr = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \tag{8}$$

where $\bar{A}$ and $\bar{B}$ are the means of images A and B, and m, n are the number of rows and columns of A and B.

Both correlation and SSIM return a single measure for each comparison between the marking and the models. Two thresholds were empirically defined for each parameter using a small subset of arrows. The largest measure among all the comparisons that is above the threshold values, defines the class of each arrow marking (Figure 5b). It also defines the orientation of the arrow (that is, if the arrow points in the same direction than the

trajectory, or in the opposite direction). The performance of correlation and SSIM is discussed in Section 4.2. The selected thresholds were 0.4 for correlation and 0.987 for SSIM.

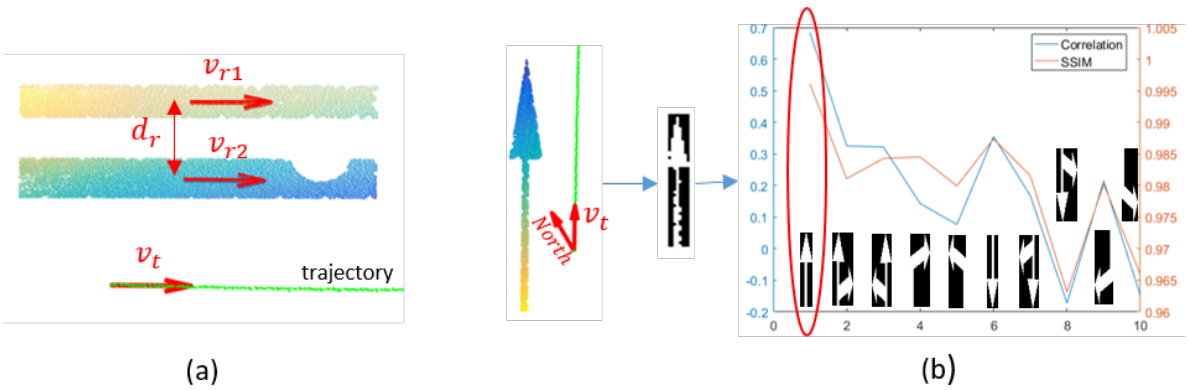

Figure 5. Road marking classification. (a) Rectangular markings are considered as elements of the same pedestrian crossing if they have similar orientations (v_{r1}, v_{r2}) and the distance between their centroids (d_r) is less than a threshold. (b) For each arrow, an oriented intensity image is compared with ten arrow models of five classes, using correlation and SSIM as similarity measures.

2.5. GIS layer creation and maintenance criteria

Once the classification concludes, the data of each marking can be gathered within an object. A class was defined for storing rectangular markings and arrows (Figure 6), with the following properties:

- Class: String that defines the class of the marking (from the first classification level).
- Sublcass: String that defines the specific meaning of the marking (from the second classification level).
- Cloud: 3D points of the marking in the original point cloud.
- bBox: 3D bounding box of the marking.
- Image: Binary image of the marking.
- Azimuth: Angle from North.
- nElements: Number of single elements of the marking (number of rectangles for pedestrian crossings).

These data can be easily exported to a GIS vectorial layer and compared with previous inventories. An automatic process between the acquired GIS layer L_i and the previous layer L_{i-1} (e.g. overlapping of bounding boxes) will create three possible scenarios, which correspond to true positives, false positives, and false negatives. For each scenario, a maintenance criteria is proposed:

1. True positives: The same marking is detected in L_i and L_{i-1}. No maintenance activity would be required, and the marking is updated automatically.
2. False positives: A new marking appears in L_i which has no overlap with any marking in L_{i-1}. This will occur if the marking has been recently painted, or if it is an actual false positive. If it is a recently painted marking, the database should be updated. If it is an actual false positive, the marking would be deleted from the layer.
3. False negatives: There is no overlap L_i for a road marking in L_{i-1} . This may occur if the marking has lost its retroreflective properties, or if it was occluded during the survey. A warning is created for the marking, as it may have lost its retroreflective properties, or it could have been occluded.

In this way, the decision-making in the inventory process is simplified, and the costs minimized. The vast majority of the markings will be automatically updated, and a small percentage would require manual decisions in order to cope with processing errors (which may result in both false positives and false negatives), recently painted markings (false positives) and occlusions during the survey (false negatives).

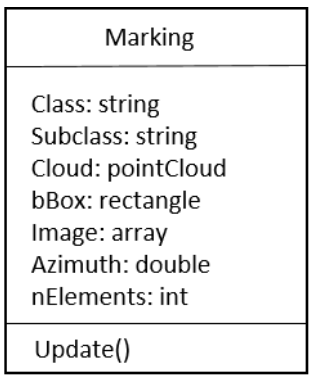

Figure 6. Class that defines a road marking with its properties.

3. Case study

The data in this work were collected using the LYNX Mobile Mapper by Optech (Optech Inc., 2012), Figure 7a. The system is composed of two LiDAR sensor heads with 500,000 measurements per second and a field of view (FOV) of 360º. The angle between their rotational axes is 90º and they have an angle of 45º with respect to the trajectory of the vehicle. Note that the methodology described in Section 2 process the data of each sensor independently. The navigation system integrates an Inertial Measurement Unit (IMU) with a two-antenna heading measurement system (GAMS). Furthermore, imagery data is registered by four 5-MPix JAI cameras that are synchronized by the Lynx Survey software. A complete accuracy study and system review can be found in Puente et al. (2013).

Three different scenarios were chosen for the validation of the method. The first one, located in the city of Lugo, in northwest Spain, comprises a 2.5km three-lane avenue that encircles a roman wall in the city center. The second and third scenarios are located in two different areas in the city of Vigo, close to the port and to the coastline (Figures 7b-d).

The survey was conducted using consecutive strips of data in order to avoid excessively large point clouds. The road marking detection methodology was applied to each strip independently. Data from Vigo-Port and Vigo-Coast areas were collected for training the classification models as described in Section 2.4. The information of the study areas is summarized in Table 1.

Table 1. Case study data.

Area	Strips	Points
Lugo (City center)	5	129,553,905
Vigo-Port	6	82,419,049

Vigo-Coast	3	50,986,844

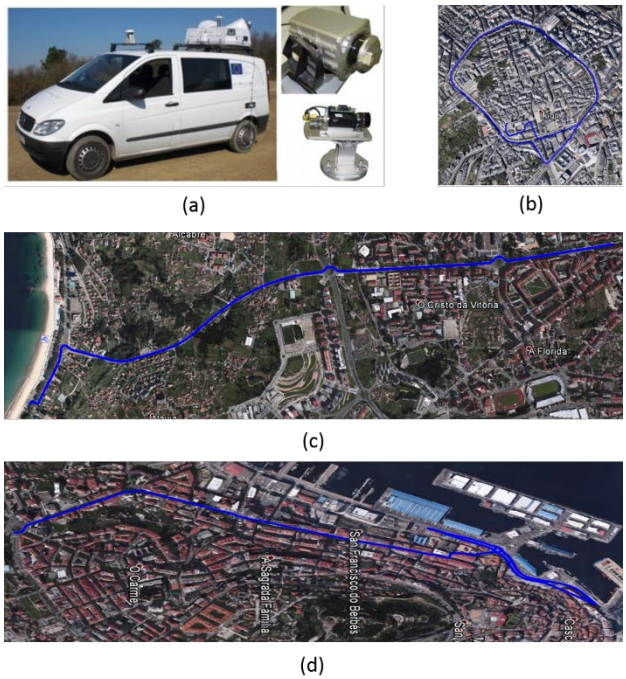

Figure 7. Case study. (a) Mobile Mapping System (MMS) used for the survey. (b) - (d): Trajectories of the study case, in the cities of Lugo (b) and Vigo (c), (d).

4. Results and discussion

The proposed methodology was tested in the roads described in Section 3. In this section, the efficiency of the road marking detection algorithm is shown and compared with other state of the art results. Subsequently, the performance of the road marking classification stage is shown. Finally, causes of error are analyzed and the drawbacks of this methodology are pointed out.

4.1. Road marking detection

In order to evaluate the road marking detection method, two representative road sections were selected (Figure 8). As depicted in Section 2.3, an intensity filter is applied, followed by a rasterization and an intensity based binarization of the raster image. The resulting image is compared with a manually annotated ground truth that contains the actual road markings to be extracted. The following performance metrics are defined:

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F_{score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{11}$$

where TP, FP and FN are the number of true positives, false positives and false negatives.

The pixel by pixel comparison between the manually annotated image and the resulting images from the road marking detection method leads to results over 90% as shown in Table 2.

Table 2. Road marking extraction results.

	Precision	Recall	Fscore
Section 1	0.936	0.913	0.925
Section 2	0.986	0.922	0.953
Average	0.961	0.917	0.939

The analyzed sections intend to be similar to the data used in Yu et al. (2015), so a comparison with state of the art results can be made. Specifically, Table 3 compares the results of the presented method with results from Guan et al. (2014) and Yu et al. (2015), extracted from the latest reference.

Table 3. Comparison with state of the art methods.

Method	Precision	Recall	Fscore
Guan et al. [18]	0.905	0.875	0.89
Yu et al. [21]	0.91	0.93	0.925
Proposed	0.961	0.917	0.939

The proposed method shows a similar performance to other state of the art methods. Although it does not detect road markings as completely as the Yu et al. (2015) method, the overall performance of the method proposed in this paper is slightly better according to this comparison. However, the data used by each author for the application of their methods are different. This fact has to be pointed out as the quantitative comparison

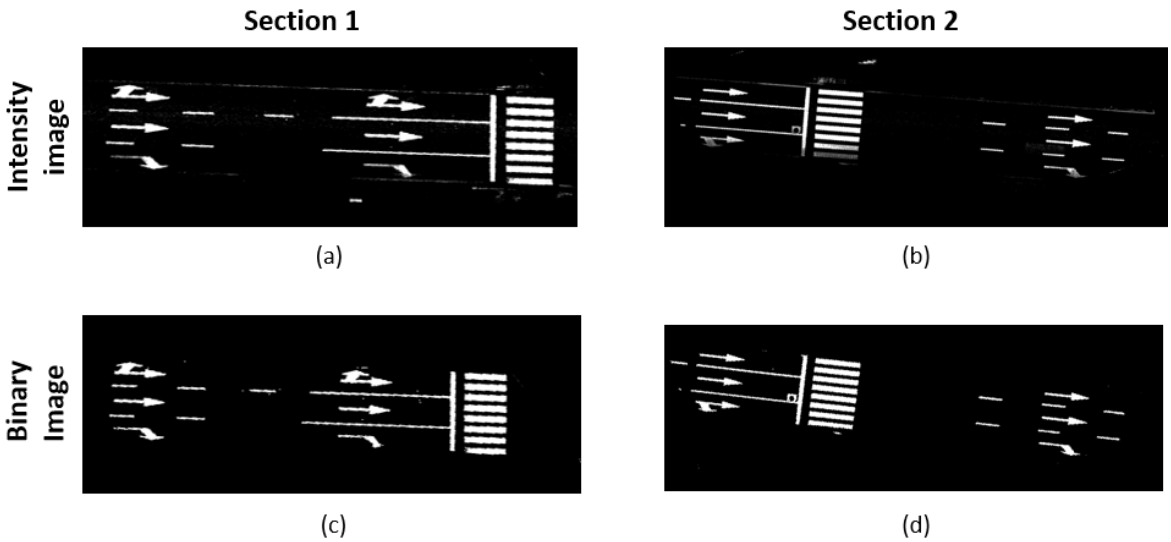

Figure 8. Representative road sections used for the evaluation of the method. (a)-(b): Intensity images. (c)-(d): Binary images.

may not be totally reliable. Anyhow, from a qualitative point of view, the performance of the presented method is similar to other state of the art methods regarding road marking detection.

4.2. Road marking classification

Once the road markings are detected, each individual marking is classified, in a first level of hierarchy, in two main classes, namely arrows and rectangular markings. In Section 2.4 the creation of classification models using Neural Networks is explained. Models using two different features (GBF and HOG) were considered and are compared on Table 4. Tables 5-6 show the confusion matrices for GBF and HOG features respectively. It can be seen that F score is almost the same for both features, although the proposed GBF classifies correctly more markings than HOG at the expense of increasing the number of false positives. These results correspond to Lugo (City Centre) MLS dataset, as data of Vigo-Coast and Vigo-Port MLS sets were used for training and cross-validation.

Table 4. Road marking classification results, first hierarchy level. Comparison of the proposed GBF feature with HOG feature.

Feature	Precision	Recall	Fscore
GBF	0.938	0.982	0.960
HOG	0.959	0.965	0.962

Table 5. Confusion matrix using GBF feature for classification.

GT/predict	Rectangle	Arrow	Negative
Rectangle	**104**	2	0
Arrow	3	**60**	1
Negative	4	3	**14**

Table 6. Confusion matrix using HOG feature for classification.

GT/predict	Rectangle	Arrow	Negative
Rectangle	**104**	1	1
Arrow	1	**63**	0
Negative	8	3	**10**

Finally, in a second level of hierarchy, a proper semantic meaning is given to the road markings. The analysis of rectangular markings as described in Section 2.4 merges them into pedestrian crossings. For those markings predicted as arrows, they are classified in five classes: Straight, Straight-left, Straight-right, Left and Right. Two different metrics were proposed to classify arrow markings, image correlation and SSIM. In Table 7, a comparison between both measures is shown, using GBF as feature for the first level of hierarchy. SSIM provides the best overall score, although the results are quite similar. Finally, Table 8 shows the confusion matrix for the final classification, including pedestrian crossings.

Table 7. Road marking classification results, second hierarchy level. Comparison of correlation and SSIM metrics.

Measure	Precision	Recall	Fscore
Correlation	0.953	0.953	0.953
SSIM	0.969	0.954	0.961

Table 8. Confusion matrix using SSIM as similarity measure in the second level of hierarchy.

GT/ predict	Ped. Crossing	Straight	Straight-right	Straight-left	Right	Left	Negative
Ped. Crossing	**16**	0	0	0	0	0	0
Straight	0	**46**	0	0	0	0	2
Straight-right	0	0	**5**	0	0	0	0
Straight-left	0	0	0	**3**	0	0	0
Right	0	0	0	0	**2**	0	0
Left	0	2	0	0	0	**1**	0
Negative	0	0	0	0	0	0	**6**

4.3. Inventory and maintenance

The road marking classification step outputs a set of objects that contain spatial, contextual and semantic information for each classified marking. As stated in Section 2.5, a GIS layer can be defined with these information.

In Figure 9, several road marking detections are drawn over a georreferenced orthophoto of the studied area, together with a manually annotated ground truth layer which pretends to simulate a previous inventory condition.

Excluding false positives and false negatives, the average distance between the centroids of the road marking bounding boxes retrieved by the proposed method and the ground truth is 0.51m. Note that the orthophoto in Figure 9 has a pixel resolution of 0.3m.

In Figure 9a there is a misdetection of one of the markings. Following the criterion defined in Section 2.5, if there exist previous knowledge on the position of the road markings, a warning appears indicating the misdetection, so it can be manually assessed. Misdetections may appear due to paint fading, algorithm failures or, as it is the case, an occlusion produced by a double parked car.

4.4. Error analysis

There are different sources of error that influence the robustness of the method. Road markings whose paint is in bad conditions due to an extended exposure to adverse weather and vehicle transit, eventually lose their reflective properties. Therefore, the intensity based filtering will likely fail. However, in the context of a road inspection application, the absence of road markings (false negatives) can be meaningful, since it would indicate the need of maintenance activities. Furthermore, the presence of parked vehicles may cause occlusions (as shown in Section 4.3), and manhole covers, waste or other reflective objects are sources of false positives or noise in the intensity image that influence the results in Table 2.

The detection and classification results show a state of the art performance. However, small occlusions or areas of a single marking where the reflectance is lower may cause the detection method to output two markings instead of one, or the classification method to misclassify the marking.

As a final remark, note that the data of each MLS LiDAR head is processed independently. If the data of both sensors were combined, the intensity values would be more consistent and the overall result could be slightly improved.

Figure 9. The GIS layer obtained with the proposed algorithm (yellow bounding boxes) can be visualized over an ortophoto, and compared with respect to a ground truth or data from a previous survey (red bounding boxes). (a) Misdetections due to occlusions or algorithm failures are automatically found in order to assist the maintenance process.

5. Conclusions and future work

In this paper a method for the automatic detection and classification of road markings from 3D point clouds collected by a Mobile Mapping system, which aims to assess the state of road signs is presented. The method relies on the reflective properties of the road markings in order to properly filter the point cloud and to create binary raster images. The performance of two different sets of features was compared for the classification of markings in a first level of hierarchy (rectangular markings and arrows) using a Neural Network. In a second level of hierarchy, rectangular markings are merged into pedestrian crossings using contextual data, and a semantic meaning is given to each arrow. Finally, the classified markings data are exported to a GIS layer, where it can be automatically and semi-automatically interpreted. This work relies on existing point cloud segmentation and classification techniques and explores new features for classification of road markings. A comparison with state of the art results provides a similar performance regarding road marking detection and classification, with overall results of 94% and 96% respectively. Furthermore, this work moves a step forward with respect to previous works, with the application of the classification methodology into the context of a Smart City, creating a GIS layer from the LiDAR data extracted by the method, which can be merged with a road network database, and simplifying the decision-making within inventory and maintenance activities.

The novelty of this work is twofold. First, the processing chain, although a combination of existing algorithms, includes new classification features which are experimentally compared with state of the art features, obtaining slightly better global results. Second, it offers a tool for establishing a road marking inventory. Road markings are an important road network asset, therefore simplifying the decision making process will save resources while providing a clear picture of the condition of the road markings in the network more efficiently.

Future work should provide more insight on the applications of the presented methodology. For instance, the surroundings of a pedestrian crossing can be analyzed in order to qualitatively assess its quality in terms of security for a pedestrian, using 3D and imagery data. Another application could be the usage of the contextual information of the arrows to define automatically the direction of traffic in each road lane or in the intersections. Furthermore, the road marking database should be enriched with more classes (give way, stop or written symbols) in order to offer a more complete definition of the road network, and a robust implementation of the method should include mechanisms to adapt its parameters whenever the percentage of erroneous detections

when comparing to the previous inventory exceeds a certain threshold given by the experimental results obtained in this work. Finally, the usage of the proposed method for Intelligent Transportation Systems (ITS) which collect 2D perfilometer data could be explored as future work.

Acknowledgements

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness through the project HERMES:S3D – Healthy and Efficient Routes in Massive Open-data based Smart Cities (Ref.: TIN201346801-C4-4-R) and Human Resources program FPI (Grant BES-2014-067736).

References

Bishop, C.M., 1995. Neural networks for pattern recognition, Journal of the American Statistical Association. Oxford Univ. Press. doi:10.2307/2965437

Chen, X., Kohlmeyer, B., Stroila, M., Alwar, N., Wang, R., Bach, J., 2009. Next generation map making: Geo-referenced ground-level LIDAR point clouds for automatic retro-reflective road feature extraction. ACM Adv. Geogr. Inf. Syst. 488. doi:10.1145/1653771.1653851

Dalal, N., Triggs, W., 2004. Histograms of Oriented Gradients for Human Detection. doi:10.1109/CVPR.2005.177

Danescu, R., Nedevschi, S., 2010. Detection and classification of painted road objects for intersection assistance applications, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 433–438. doi:10.1109/ITSC.2010.5625261

Douillard, B., Underwood, J., Kuntz, N., Vlaskine, V., Quadros, a., Morton, P., Frenkel, a., 2011. On the segmentation of 3D lidar point clouds. Proc. - IEEE Int. Conf. Robot. Autom. 2798–2805. doi:10.1109/ICRA.2011.5979818

Flusser, J., Suk, T., 1993. Pattern recognition by affine moment invariants. Pattern Recognit. 26, 167–174. doi:10.1016/0031-3203(93)90098-H

Guan, H., Li, J., Yu, Y., Wang, C., Chapman, M., Yang, B., 2014. Using mobile laser scanning data for automated extraction of road markings. ISPRS J. Photogramm. Remote Sens. 87, 93–107. doi:10.1016/j.isprsjprs.2013.11.005

Hoffmann, G.M., Tomlin, C.J., Montemerlo, M., Thrun, S., 2007. Autonomous automobile trajectory tracking for off-road driving: Controller design, experimental validation and racing. Proc. Am. Control Conf. 2296–2301. doi:10.1109/ACC.2007.4282788

Höfle, B., Pfeifer, N., 2007. Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS J. Photogramm. Remote Sens. 62, 415–433. doi:10.1016/j.isprsjprs.2007.05.008

Jirachan, T., Piromsopa, K., 2015. Applying KSE-test and K-means clustering towards Scalable Unsupervised Intrusion Detection, in: 12th International Joint Conference on Computer Science and Software Engineering (JCSSE). pp. 82–87. doi:10.1109/JCSSE.2015.7219775

Kibbel, J., Justus, W., Fürstenberg, K., 2005. Lane estimation and departure warning using multilayer laserscanner, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 777–781. doi:10.1109/ITSC.2005.1520147

Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T., 2014. Automated road markings extraction from mobile laser scanning data. Int. J. Appl. Earth Obs. Geoinf. 32, 125–137. doi:10.1016/j.jag.2014.03.023

McCall, J.C., Trivedi, M.M., 2006. Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation. IEEE Trans. Intell. Transp. Syst. doi:10.1109/TITS.2006.869595

McLachlan, G., Peel, D., 2000. Finite Mixture Models, Technometrics. doi:10.1198/tech.2002.s651

Ministerio de Fomento, 1987. Norma de carreteras 8.2-IC. Marcas Viales.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 66, 247–259. doi:10.1016/j.isprsjprs.2010.11.001

Otsu, N., 1979. Threshold selection method from grey-level histograms. IEEE Trans Syst Man Cybern.

Pu, S., Rutzinger, M., Vosselman, G., Oude Elberink, S., 2011. Recognizing basic structures from mobile laser scanning data for road

inventory studies. ISPRS J. Photogramm. Remote Sens. 66, S28–S39. doi:10.1016/j.isprsjprs.2011.08.006

Puente, I., González-Jorge, H., Riveiro, B., Arias, P., 2013. Accuracy verification of the Lynx Mobile Mapper system. Opt. Laser Technol. 45, 578–586. doi:10.1016/j.optlastec.2012.05.029

Rimini-Doering, M., Altmueller, T., Ladstaetter, U., Rossmeier, M., 2005. Effects of lane departure warning on drowsy drivers' performance and state in a simulator, in: PROCEEDINGS of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. pp. 88–95.

Rosin, P.L., 2003. Measuring shape: Ellipticity, rectangularity, and triangularity. Mach. Vis. Appl. 14, 172–184. doi:10.1007/s00138-002-0118-6

Serna, A., Marcotegui, B., 2014. Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning. ISPRS J. Photogramm. Remote Sens. 93, 243–255. doi:10.1016/j.isprsjprs.2014.03.015

Serna, A., Marcotegui, B., 2013. Urban accessibility diagnosis from mobile laser scanning data. ISPRS J. Photogramm. Remote Sens. 84, 23–32. doi:10.1016/j.isprsjprs.2013.07.001

Ussyshkin, V., 2009. Mobile Laser Scanning Technology for Surveying Application: From Data Collection to End-Products. FIG Work. Week, Surv. Key Role Accel. … 1–13.

Vacek, S., Schimmel, C., Dillmann, R., 2007. Road-marking analysis for autonomous vehicle guidance. Proc. Eur. Conf. Mob. Robot. 1–6.

Vo, A.-V., Truong-Hong, L., Laefer, D.F., Bertolotto, M., 2015. Octree-based region growing for point cloud segmentation. ISPRS J. Photogramm. Remote Sens. 104, 88–100. doi:10.1016/j.isprsjprs.2015.01.011

Wang, H., Luo, H., Wen, C., Cheng, J., Li, P., Chen, Y., Wang, C., Li, J., Member, S., 2015. Road Boundaries Detection Based on Local Normal Saliency From Mobile Laser Scanning Data 12, 2085–2089.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612. doi:10.1109/TIP.2003.819861

Yang, B., Dong, Z., 2013. A shape-based segmentation method for mobile laser scanning point clouds. ISPRS J. Photogramm. Remote Sens. 81, 19–30. doi:10.1016/j.isprsjprs.2013.04.002

Yang, B., Dong, Z., Zhao, G., Dai, W., 2015. Hierarchical extraction of urban objects from mobile laser scanning data. ISPRS J. Photogramm. Remote Sens. 99, 45–57. doi:10.1016/j.isprsjprs.2014.10.005

Yu, Y., Li, J., Guan, H., Jia, F., Wang, C., 2015. Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 709–726.