ISPRS Journal of Photogrammetry and Remote Sensing 114 (2016) 92–101

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory

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

ARTICLE INFO

Article history:
Received 8 September 2015
Received in revised form 26 January 2016
Accepted 28 January 2016
Available online 18 February 2016

Keywords:
Mobile mapping
Laser scanning
Traffic sign inventory
Traffic sign recognition
Point cloud segmentation

ABSTRACT

Nowadays, mobile laser scanning has become a valid technology for infrastructure inspection. This technology permits collecting accurate 3D point clouds of urban and road environments and the geometric and semantic analysis of data became an active research topic in the last years. This paper focuses on the detection of vertical traffic signs in 3D point clouds acquired by a LYNX Mobile Mapper system, comprised of laser scanning and RGB cameras. Each traffic sign is automatically detected in the LiDAR point cloud, and its main geometric parameters can be automatically extracted, therefore aiding the inventory process. Furthermore, the 3D position of traffic signs are reprojected on the 2D images, which are spatially and temporally synced with the point cloud. Image analysis allows for recognizing the traffic sign semantics using machine learning approaches. The presented method was tested in road and urban scenarios in Galicia (Spain). The recall results for traffic sign detection are close to 98%, and existing false positives can be easily filtered after point cloud projection. Finally, the lack of a large, publicly available Spanish traffic sign database is pointed out.

1. Introduction

The visibility and quality of traffic signs is one of the most important factors for road safety. In Spain, driver distractions are the leading cause of fatal car accident (RACE et al., 2013). The developments of intelligent vehicles that can warn and inform the driver are crucial to increasing driver safety. Furthermore, the European Directive on Road Infrastructure Safety Management (European Commission, 2013) remarks that Member States should ensure periodic inspections of roads in operation. Currently, road safety inspection is typically done by a qualified inspector who evaluates a number of safety parameters and writes up a report where the elements that need maintenance are remarked. Therefore, decisions are based upon the knowledge of the inspector and a subjective diagnosis. A semi-automatic or fully automatic inspection will reduce its subjectivity and will save public resources while improving the road safety.

The most common source of information for this inventory task is RGB images taken by one or several cameras installed in a vehicle. The literature on traffic sign detection and recognition on images is numerous, and a broad variety of computer vision techniques have been applied to this problem. Illumination changes, occlusions, cluttered scenes or vandalized traffic signs are some of the challenging problems to deal with. Different color spaces have been used, for instance HSI-HSV (Fleyeh, 2006; Gomez-Moreno et al., 2010), YUV (Shadeed et al., 2003) or Gaussian color model (Li et al., 2015) as a visual feature to define a traffic sign region. Shape features have also been studied, such as Hough Transform (Barrile et al., 2012), Local Contour Pattern (Landesa-Vázquez et al., 2010), or Local Binary Patterns (Liu et al., 2014). Usually, both color and shape features are combined in order to obtain a better description. Contextually, there are two possible objectives in these works: traffic sign detection, this is, computing the regions on an image that contain a traffic sign (Salti et al., 2015; Wu et al., 2013); and traffic sign recognition, where the specific semantics of the detected traffic signs are obtained using complex machine learning techniques, such as Convolutional Networks (Sermanet and Lecun, 2011) or Support Vector Machines (Wang et al., 2013a). Some of the cited articles are based on the German Traffic Sign Detection Benchmark (GTSDB) and the German Traffic Sign Recognition Benchmark (Houben et al., 2013; Stallkamp et al., 2012) which are respectively a single-image detection and a multi-class classification challenge. Wang et al. (2013b) obtained the best

* Corresponding author.
E-mail address: msoilan@uvigo.es (M. Soilán).

http://dx.doi.org/10.1016/j.isprsjprs.2016.01.019
0924-2716/

results for the GTSDB with detection rates of almost 100%. For GTSRB, Cireşan et al. (2012) got a recognition accuracy of 99.5% with a committee of Convolutional Neural Networks, improving the human performance. However, geometric inventory parameters and contextual properties of traffic signs (inclination of the panel, angle and distance with respect to the trajectory of a vehicle, etc.), which are relevant features to take into account during inventories, are not considered.

Nowadays, Mobile Mapping Systems (MMS) equipped with laser scanners are able to collect accurate and reliable 3D point clouds. A survey of an urban or a road environment provides geometric and radiometric information of infrastructure facilities. Research is focused on the automation of detection and classification processes for these elements. For example, Zhou and Vosselman (2012) detect curbstones and therefore road sides in urban areas, Yu et al. (2015) detect street lights using a pairwise 3D shape context that classifies clustered 3D points, and Serna and Marcotegui (2014) detect and classify 20 urban objects (cars, bollards, traffic lights, etc.) using mathematical morphology and Support Vector Machines (SVM), and they present an extensive review of methodologies employed for urban object classification in point clouds. Recently, Yang et al. (2015) classified several objects (including traffic signs) segmenting multi-scale supervoxels and applying an hierarchical and heuristic classification process.

This paper intends to combine two sources of information, namely 3D point clouds and RGB imagery, both installed in a Mobile Mapping System. In our previous work (Riveiro et al., 2015), only 3D data was processed and the imagery collected by the MMS was omitted. Consequently, in this work the previous methodology is improved not only by adding RGB imagery but by slightly modifying the point cloud processing workflow. Therefore, the proposed method has proven to be efficient for the task of detection, classification and extraction of geometric properties of vertical traffic signs. In Section 2, the proposed methodology is described. In Section 3, the MMS features and the study case are presented. Section 4 will summarize the results and their discussion by comparing them with the results in (Riveiro et al., 2015), and Section 5 will outline the conclusions.

2. Methodology

The proposed method aims to identify both geometric and semantic properties of traffic signs in urban and highway environments. For that purpose, two main sources of data are used: point clouds collected with a mobile laser scanner and imagery collected with RGB cameras, both of them integrated in a MMS.

The general workflow is shown in Fig. 1. First, the 3D point cloud is segmented using ground removal and intensity filter techniques. The segmented cloud is organized using a DBSCAN-based clustering, and further filtered in order to isolate each traffic sign in the road. Second, the geometric parameters of each detected sign are extracted. Subsequently, the 3D points of each traffic sign are reprojected on 2D images, using the trajectory of the vehicle, the time stamp of both 3D points and images, and the relative orientation between the vehicle and the cameras. Finally, a hierarchical classification method is carried out to obtain the meaning of the traffic sign.

2.1. Point cloud segmentation

González-Jorge et al. (2013) state that the intensity attribute of a point cloud can be used for segmentation processes, and therefore for the evaluation of road signs. Furthermore, Pu et al. (2011) classify basic traffic sign shapes using the intensity attribute to support the process. Similarly to these methods, our point cloud segmentation method focuses on using the intensity data recorded by the scanner.

First, the original point cloud is preprocessed. Using the trajectory data, the distance from the 3D points to the sensor is computed, and points further than 20 meters are filtered out. This way, only the information of the road or street where the MMS is traveling will be kept for further processing.

Let $P = (\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{I}, \mathbf{ts})$ be the point cloud after preprocessing, where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are the 3D coordinates, $\mathbf{I}$ is the intensity, and $\mathbf{ts}$ is the time stamp of each point. A ground removal process is used to separate off-ground points from ground points in P. This step is motivated by two reasons: ground removal significantly reduces the number of points in the cloud, consequently reducing the computational load and saving memory resources; the second advantage is that road markings, which present high intensities in the cloud, are removed too so that they do not distort the filtering based on intensity. Note that the present methodology focuses only on vertical traffic sign detection, and this way the problem of dealing with road marking detections is avoided.

The ground removal process consists of a rasterization of P onto the horizontal plane. This space reduction is motivated by two requirements: improving the computational performance of the algorithm; and, given that mobile laser scanning point clouds are leveled, the projection of 3D points onto the XY plane produce an appropriate plan view of the environment. During rasterization, it is necessary to define a grid size, g_s, that defines the resolution of the raster structure. Grid sizes from 20 cm to 1 m were tested and trade-off between efficiency and quality of the segmentation was obtained for $g_s = 0.5$ m (for smaller sizes, the quality is similar while the execution time grows exponentially, and bigger sizes lose local information). For each point p_i in the cloud, a cell coordinate (x_i^R, y_i^R) and an index id_i are assigned following Eq. (1):

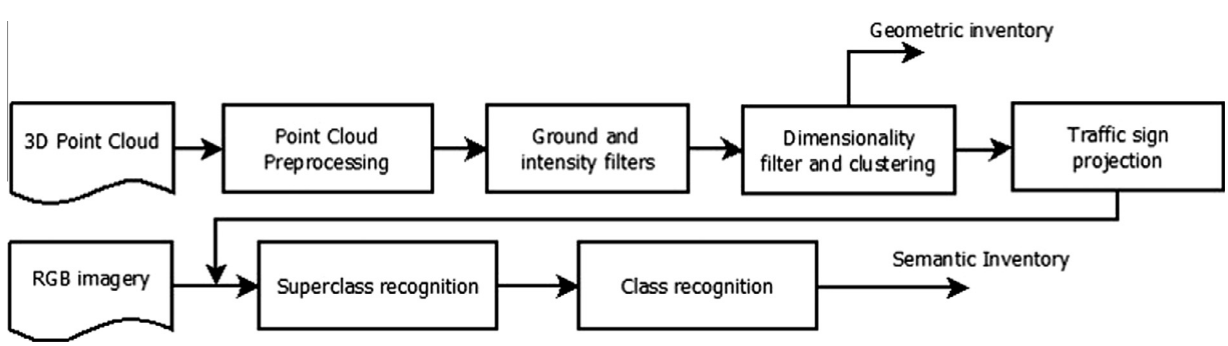

Fig. 1. General workflow.