Advanced Engineering Informatics 32 (2017) 263–274

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Multi-class US traffic signs 3D recognition and localization via image-based point cloud model using color candidate extraction and texture-based recognition

Vahid Balali [a,*], Arash Jahangiri [b], Sahar Ghanipoor Machiani [b]

[a] Department of Civil Engineering and Construction Engineering Management, California State University, Long Beach, CA 90840, United States
[b] Department of Civil, Construction, and Environmental Engineering, San Diego State University, San Diego, CA 92182, United States

ARTICLE INFO

Article history:
Received 21 September 2016
Received in revised form 16 March 2017
Accepted 24 March 2017

Keywords:
Traffic signs
Condition assessment
3D recognition
Localization
Image-based
3D point cloud

ABSTRACT

Continuous condition monitoring and inspection of traffic signs are essential to ensure that safety and performance criteria are met. The use of 3D point cloud modeling by the construction industry has been significantly increased in recent years especially for recording the as-is conditions of facilities. The high-precision and dense 3D point clouds generated by photogrammetry can facilitate the process of asset condition assessment. This paper presents an automated computer-vision based method that detects, classifies, and localizes traffic signs via street-level image-based 3D point cloud models. The proposed pipeline integrates 3D object detection algorithm. An improved Structure-from-Motion (SfM) procedure is developed to create a 3D point cloud of roadway assets from the street level imagery. In order to assist with accurate 3D recognition and localization by color and texture features extraction, an automated process of point cloud cleaning and noise removal is proposed. Using camera pose information from SfM, the points within the bounding box of detected traffic signs are then projected into the cleaned point cloud by using the triangulation method (linear and non-linear) and the 3D points corresponding to the traffic sign in question are labeled and visualized in 3D. The proposed framework is validated using real-life data, which represent the most common types of traffic signs. The robustness of the proposed pipeline is evaluated by analyzing the accuracy in detection of traffic signs as well as the accuracy in localization in 3D point cloud model. The results promise to better and more accurate visualize the location of the traffic signs with respect to other roadway assets in 3D environment.

Published by Elsevier Ltd.

1. Introduction

Most State Departments of Transportations (DOTs) have no automated system for inventorying and maintaining records of roadway assets. Over the past few years, the DOTs have proactively looked into videotaping roadways –using inspection vehicles equipped with high resolution cameras– to provide accurate information on location and condition of high quantity and low cost roadway assets. The inventory data collection involves a survey cycle of one year for critical highways and completely neglects all other local and regional roadways due to their high costs and manual analysis. Such data collection and analysis have to be done for millions of miles of roads and the practice needs to be repeated every so often. The potential of applying object detection algorithms has enabled development of unique computer vision methods for detection and classification of traffic signs to minimize the need for manual inventorying and maintaining records of the transportation assets for enhanced condition assessment [1–3]. To make these methods useful, both False Positive (FP) and False Negative (FN) rates should be very low. That is a major reason why current traffic sign inventory and condition assessment practices are still carried out manually.

Traffic sign recognition is among the important components of roadway asset management systems. On-demand access to as-is condition of traffic signs has a significant potential for improving decision-making on transportation asset management. Visualizing 3D models with photographs provides an unprecedented opportunity for inspection crew to visually interact with type of traffic signs, geo-localize potential errors or issues, and quickly disseminate this information to other users across the project. It can also facilitate field reporting and quality inspections as it allows iterations of work-in-progress and inspections to be properly logged.

* Corresponding author.
E-mail addresses: vahid.balali@csulb.edu (V. Balali), ajahangiri@mail.sdsu.edu (A. Jahangiri), sghanipoor@mail.sdsu.edu (S.G. Machiani).

http://dx.doi.org/10.1016/j.aei.2017.03.006
1474-0346/Published by Elsevier Ltd.

Some of the research projects, particularly for sign detection have been motivated by DARPA's (Defense Advanced Research Project Agency). Google has also been focusing on application of laser scanners and cameras together for purpose of autonomous navigation of google car [4,5]. In addition, recent advances in computing have created a great opportunity to run 3D image-based reconstruction at large scales. A few research groups [6,7] have already demonstrated high density and accurate image-based reconstruction results. The potential for creating and visualizing a comprehensive inventory of traffic signs using online databases such as Google Street View images has been explored by [8]. However, the geo-information of traffic signs detected using these images is not very accurate for inventory and 3D localization purposes. The location information of traffic signs is extracted from the GPS (Global Positioning System) tag of the images which may not reflect the exact location of signs.

Traffic sign 3D recognition and localization is a longstanding and challenging problem in computer vision. The main challenge is that the appearance of traffic signs in images is affected by a number of factors, such as illumination, scale, camera viewpoint, intra-class variability, and the changing geometry/appearance of roadway assets. These images are traditionally unordered and un-calibrated, and usually include significant amount of occlusion (both dynamic and static), which makes the application of 3D reconstruction algorithms difficult. Furthermore, current approaches are designed to work with hundreds of photographs of a static scene in which there is very high spatial density among the camera positions and view directions. Photogrammetry data usually contains noise that appears as fuzziness in the point cloud. Asset condition assessment involves preprocessing the point cloud data, for which point clouds have to be cleaned of any unwanted or occluding objects, and noises. This process is time consuming and labor intensive and often subject to human errors. Errors in trimming out portions of point clouds may result in loosing important data and consequently effecting the data computation during the post-processing stage. In order to have an accurate 3D point cloud and to minimize the errors for different purposes, such as 3D object detection, there is a need to automate the process of point cloud cleaning and noise removal to reduce human errors, and to decrease the time and cost associated with the manual process. How to take all parameters that could affect accuracy in 3D recognition process is still an open problem.

To address these challenges, the efforts in building 3D representations for traffic sign recognition and localization is presented in this paper. Compared to 2D appearance-based traffic sign representations, 3D representations can capture the 3D nature of traffic signs and better handle viewpoint variation, occlusion and truncation in recognition and localization process. This paper proposes a pipeline for image-based detection and classification of traffic signs criteria. The 3D image-based reconstruction module builds upon the newly proposed SfM and is tested in the context of sequentially captured images for traffic signs. A non-deterministic algorithm is used to remove the noises from the point clouds automatically. And finally the traffic signs are recognized and localized via image-based 3D point cloud model. The following sections describe the review of the state-of-the-art recognition and reconstruction methods, discuss the proposed framework for 3D traffic sign recognition and localization via point cloud in detail; and evaluate the framework's performance using real-life data.

2. Related work

The state DOTs use a variety of methods for inventorying roadside assets. Based on the underlying technologies and the sensing platform, the methods used by many state and local agencies to collect roadway inventory data can be listed as field inventory, photo/video log, integrated GPS/GIS mapping systems, aerial/satellite photography, virtual photo tourism, terrestrial laser scanners, and mobile mapping systems [9]. Vehicle-mounted LiDAR, a relatively new type of mobile mapping system, is capable of collecting large amounts of detailed 3D roadway inventory data, but it requires expensive equipment and significant data reduction to extract the desired roadway inventory data.

2.1. 2D detection and classification

There are more than 670 types of traffic signs which come in hundreds of variations, such as in dimension, color, text, and font. Ai and Tsai [10] presented a sign recognition algorithm that primarily focuses on speed limit signs. Their algorithm benefits from a probabilistic color model, the likelihood of sign locations in the images along with the traditional sign features (shape, color, and content features). Good sign features are especially crucial to develop successful sign recognition methods. Sign color, shape, texture, etc., are the most commonly used features. Brkic [11] reviewed the popular traffic sign detection methods in three categories: color-based [12], shape-based [13,14], and learning-based [1]. Many researchers developed different sign color models for sign color segmentation and validation, such as deterministic color models in [15], statistic color models in [16], and HIS color model [17]. Sign shape is another important feature for sign recognition. Many shape detectors are reported in the literature, such as polygon detectors [18], active contour detector [19], line detectors [20], and angle detectors [21,22]. Sign textures or texts are also investigated. Integration of color, shape, and texture for sign recognition is reported in [1,23]. So far the most popular features are edges and gradients, but other options such as HOG (Histogram of Oriented Gradient) and Haar wavelets, which have been investigated [1,24,25]. There is also a wealth of options and the choice is made in conjunction with the choice of a detection method. Learning-based methods are based on machine learning such as the Viola-Jones detector, in which the method is trained to classify different objects. These approaches are considered suitable solutions, especially if many observations can be labeled (i.e. supervised learning approach).

2.2. 3D reconstruction

With the advent of visual remote sensing technologies, 3D scene reconstruction has been employed by the construction industry for several different uses such as 3D as-built model reconstruction [26], project control and progress monitoring [27], MEP clash detection and as-built reconstruction [28,29], 3D thermal modeling [30], and infrastructure surveying and inspection [31]. The precise and non-contact measurement ability of 3D point cloud models has facilitated monitoring of infrastructure systems. Recent research studies have also focused on applying image processing techniques and laser scanner technology for highway asset data collection and road construction [32,33]. Performances of these technologies in creating realistic 3D models have been also compared in many studies [34].

The problem of registering large numbers of unordered ground photos, time-lapse videos, and aerial imagery has received tremendous interest in the civil engineering and computer vision communities. Significant success has been reported with semi-automated systems for registering 3D point cloud models with time-lapsed videos [35,36] and more recently commodity smartphones [37,38]. While photogrammetry can provide precise geometry data, the utilization of point clouds without adequate preprocessing may result in inaccurate assessments.