Transportation Research Part C 63 (2016) 1–22

Contents lists available at ScienceDirect

Transportation Research Part C

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

Virtual 3D city model as *a priori* information source for vehicle localization system

Maya Dawood [a,b,*], Cindy Cappelle [c], Maan E. El Najjar [a], Mohamad Khalil [b], Bachar El Hassan [b], Denis Pomorski [a], Jing Peng [a]

[a] Lille 1 University, LAGIS, UMR CNRS 8164, Lille, France
[b] Lebanese University, Doctoral School for Sciences and Technology, AZM Center for Research in Biotechnology, Tripoli, Lebanon
[c] Université de Technologie de Belfort-Montbéliard, Laboratoire Systèmes et Transports, Belfort, France

A R T I C L E I N F O

Article history:
Received 13 August 2011
Received in revised form 4 December 2015
Accepted 4 December 2015
Available online 29 December 2015

Keywords:
Localization
Data fusion
Image processing
Intelligent vehicle
GPS
3D-GIS

A B S T R A C T

This paper aims at demonstrating the usefulness of integrating virtual 3D models in vehicle localization systems. Usually, vehicle localization algorithms are based on multi-sensor data fusion. Global Navigation Satellite Systems GNSS, as Global Positioning System GPS, are used to provide measurements of the geographic location. Nevertheless, GNSS solutions suffer from signal attenuation and masking, multipath phenomena and lack of visibility, especially in urban areas. That leads to degradation or even a total loss of the positioning information and then unsatisfactory performances. Dead-reckoning and inertial sensors are then often added to back up GPS in case of inaccurate or unavailable measurements or if high frequency location estimation is required. However, the dead-reckoning localization may drift in the long term due to error accumulation. To back up GPS and compensate the drift of the dead reckoning sensors based localization, two approaches integrating a virtual 3D model are proposed in registered with respect to the scene perceived by an on-board sensor. From the real/virtual scenes matching, the transformation (rotation and translation) between the real sensor and the virtual sensor (whose position and orientation are known) can be computed. These two approaches lead to determine the pose of the real sensor embedded on the vehicle. In the first approach, the considered perception sensor is a camera and in the second approach, it is a laser scanner. The first approach is based on image matching between the virtual image extracted from the 3D city model and the real image acquired by the camera. The two major parts are: 1. Detection and matching of feature points in real and virtual images (three features points are compared: Harris corner detector, SIFT and SURF). 2. Pose computation using POSIT algorithm. The second approach is based on the on–board horizontal laser scanner that provides a set of distances between it and the environment. This set of distances is matched with depth information (virtual laser scan data), provided by the virtual 3D city model. The pose estimation provided by these two approaches can be integrated in data fusion formalism. In this paper the result of the first approach is integrated in IMM UKF data fusion formalism. Experimental results obtained using real data illustrate the feasibility and the performances of the proposed approaches.

* Corresponding author at: Beirut, Lebanon. Tel.: +961 3593709.
E-mail address: dawood-maya@hotmail.com (M. Dawood).

http://dx.doi.org/10.1016/j.trc.2015.12.003
0968-090X/

1. Introduction

The geo-localization system is an essential component of intelligent transportation systems (ITS). Nowadays, outdoor positioning systems often rely on GPS because of its affordability and convenience (Abbott and Powell, 1999). However, GPS suffers from satellite masks occurring especially in urban environments, under bridges, tunnels or in forests. To provide continuous, accurate, and high integrity position data, GPS-based localization systems should then incorporate additional sensors (as proprioceptive sensors or environment perception sensors) and/or database (for example 2D digital map) (El Badaoui El Najjar and Bonnifait, 2005). Nevertheless, using only incremental encoders placed on the rear wheels and gyroscopes is not sufficient in case of long GPS outages, because the dead-reckoning (DR) localization is subject to drift due to error accumulation (Se and Lowe, 2002; Dawood and Cappelle, 2011). Variety of approaches studied to compensate this problem. Map-matching algorithms is one example. It produces the vehicle's position using road map and an input generated from GPS or GPS integrated with DR sensors. Map matching process based on finding the road where the vehicle is traveling then project current positioning point to the road of vehicle traveling (Bernstein and Kornhauser, 1988; Renault and LeMeur, 2005). provide good survey of existing map matching. Another paper highlights the potential impacts of the forthcoming European Geostationary Overlay Service (EGNOS) on the performance of map matching algorithm (Quddus and Ochieng, 2007). In addition, two major approaches of vision based localization are SLAM (Simultaneous Localization And Mapping) and visual odometry. The goal of SLAM is to localize a robot in the environment while mapping it at the same time (Davison, 2003). Real time hierarchical outdoor SLAM based on stereovision and GPS fusion is proposed in Schleicher and Bergasa (2009). Marais and Meyrie (2014) deals with combination of video and GNSS for localization. In this paper, the proposed approach consists in determining where satellites are located in the fisheye image, and then excluding those located in non-sky regions when calculating the position. Visual odometry for vehicle geo-localization has been an object of great interest during last decades. An approach for SLAM is proposed in Dissanayake and Newman (2001), which is based on the extended Kalman filter. In Scaramuzza and Seigwart (2008), an appearance-guided monocular omni-directional visual odometry for outdoor ground vehicles is proposed.

In this paper, new approaches to back up GPS limitations and drift of dead-reckoning sensors based localization are proposed. The originality of these proposed approaches is the integration of a virtual 3D model of the environment in the absolute localization process. The principle is to compare the scene perceived by an on-board sensor and the scene generated thanks to the virtual 3D model. In the first case, the considered exteroceptive sensor is a video camera. In the second case, it is a laser scanner (Peng and Najjar, 2009).

The first approach is based on image matching (Se and Lowe, 2001, 2002; Dawood and Cappelle, 2011). Features points are detected in the real image acquired by the an-board video camera and in the virtual image generated by the virtual 3D model. Then the two sets of detected points are matched. Three methods are often used for feature detection and matching: Harris corner detector (Peng and Chen, 2009; Harris and Stephens, 1988), SIFT (Scale Invariant Feature Transform) (Se and Lowe, 2001; Peng and Chen, 2009; Lowe, 2004) and SURF (Speed Up Robust Features) (Bay and Tuytelaars, 2006; Bay and Ess, 2008). Harris corner extraction algorithm is one of the effective and widely used algorithms for feature detection. However, Harris corners are not scale invariant. Due to this, we turn to SIFT algorithm whose idea is to represent image parts by histograms of gradient location and orientation. Although, SIFT is currently the most used method for scale/rotation invariant feature detection, it is too computationally expansive to be used in real time or large scale evolution environments. The results of SURF is comparable to the SIFT one. But the additional advantage is that SURF is less cost computational. Finally, to estimate the vehicle position and orientation, POSIT algorithm is used in David and Dementhon (2004) and DeMenthon and Davis (1995).

It permits to avoid the intrinsic non linearity of the geometrical constraints that arises when the imaging process is modeled as a perspective transformation.

Considering the laser scanner based second approach (Peng and El Badaoui El Najjar, 2009), the geo-pose is computed by scan matching between real scan data provided by the real laser scanner and virtual scan data obtained using the virtual 3D model. ICP (Iterative Closest Point) algorithm (Zhang, 1994) is used to determine the transformation R and T between the two scans data. The pose can then be computed with the kinematic model of unicycle (Gutmann and Schlegel, 1996).

Each of the pose approach estimation has advantages and drawbacks. The laser scanner based approach is more efficient in term of computation time than the vision based approach. Moreover, high resolution textured images are not needed contrary to the vision based method (Lothe and Bourgeois, 2009). In addition, in the laser scanner based approach, it is not affected by the time of day (day or night) as the vision approach. Nevertheless, information provided by laser scanner can be insufficient in certain configurations of the environment, vision based approach is then useful as images provide more information.

The pose estimated using virtual 3D model can be integrated in a multi-sensor data fusion framework (with GPS, dead-reckoned sensors ...). Many researches use multi-sensors fusion for better vehicle tracking. Salameh and Challita (2013) proposes two different applications in the area of V2V communications. First application presents a method for better car tracking using GPS information shared through V2V communication and vision system. Second application presents a new simulated framework for prototyping the whole process by combining embedded data, vision, and V2V simulation. Extended Kalman Filter (Grewal and Weill, 2011; Welch and Bishop, 2002), Unscented Kalman Filter (Julier and Uhlmann, 1997; Wan and Van der Merwe, 2001; Julier and Uhlmann, 2004), and Particular Filtering (Gustafsson and Bergman,