Applied Soft Computing 37 (2015) 799–808

Contents lists available at ScienceDirect

Applied Soft Computing

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

Generation of road maps from trajectories collected with smartphone – A method based on Genetic Algorithm

George H.R. Costa, Fabiano Baldo*

Santa Catarina State University, Joinville, Brazil

ARTICLE INFO

Article history:
Received 29 April 2015
Received in revised form 18 June 2015
Accepted 21 July 2015
Available online 25 September 2015

Keywords:
Digital road maps
Map generation
Smartphones
Genetic Algorithm

ABSTRACT

Smartphones and automotive GPS have considerably boosted the use of digital road maps. For this reason, they must be updated regularly with accurate new data. The methods currently used to generate maps — photogrammetry and collaborative editing — have low frequency of update because they depend on manual intervention. By using an automated method it should be possible to improve map update speeds while maintaining similar level of accuracy. The literature presents some approaches for automatic road map creation using moving objects, but none of them is prepared for continuous update. Therefore, this work aims to propose a new automated method that uses trajectories provided by GPS receivers integrated in smartphones. It is assumed that the points that represent the center of the roads can be found through approximations provided by Genetic Algorithm. After that, these points are combined to generate the road map. However, the use of trajectories collected with smartphones provides some challenges, such as: elimination of data with bad accuracy, identification of the means of transport used and reduction of the volume of data processed. Thus, the objective of this work is to propose a method that cleans, analyzes and enriches data from smartphones to generate accurate road maps that can be continuously updated, using Genetic Algorithm. Tests indicate that the proposed method can generate maps with quality similar to the reference maps with less than 2 m of difference in average. Additionally, a comparison between the Fuzzy C-Means algorithm and the Genetic Algorithm shows that the later is a little slower but generates more accurate results.

1. Introduction

Due to the popularization of devices with integrated GPS (Global Positioning System) receiver, digital road maps are becoming important tools used everyday by a considerable number of people. The more they are used, the better they should reflect the existing road structure in order to avoid path mismatches, traffic accidents and so forth. Therefore, this means that digital road maps must be updated regularly using high accuracy data.

Usually maps are generated and updated by photogrammetric methods applied on pictures taken by airplanes and, more recently, on satellite images [11]. The results are adequate, but need manual adjustments in areas blocked by treetops or buildings. Another disadvantage is that these maps are not regularly updated. After some time they might contain inconsistencies compared with the actual roads. Due to these shortcomings different methods are being proposed.

An interesting approach is to use GPS as source of data. When GPS data is collected while an object is moving it represents the object's localization history, also called the trajectory of the moving object [3]. This trajectory is composed of points that indicate where the moving object was in each time instant, and can serve as source of traveled paths to generate maps. Assuming that the moving objects are motor vehicles, the map generated from GPS data will therefore be a road map.

One initiative to build maps from GPS data is to use collaborative editing tools such as OpenStreetMap [21]. This collaborative project allows users to upload GPS trajectories and use them to create or update maps. However, since the map editing is done manually, maps created with this method can also become inconsistent soon. There is also the possibility of errors caused by people editing the map of an area they do not know [9]. Therefore, an automatic solution would be more effective, since it could allow maps to be updated more quickly. Studies like [4,5] support this approach.

To collect the needed GPS data any device with an integrated GPS receiver could be used. However, smartphones also contain

* Corresponding author. Tel.: +55 4740097849.
E-mail addresses: geohrc@gmail.com (G.H.R. Costa), fabiano.baldo@udesc.br (F. Baldo).

http://dx.doi.org/10.1016/j.asoc.2015.07.017
1568-4946/

other sensors, which combined with its GPS receiver can provide more detailed information about how, when and where the moving objects (smartphone users) are moving. Since smartphones are generally carried close to their users during the whole day, the volume of data collected tends to be high. In combination with the fact that the number of people with these devices has been increasing fast over the years, smartphones seems to be a suitable choice for gathering georeferenced data to build road maps. Despite that, it is necessary to observe the quality of the collected trajectories since they might contain a lot of noise, like points with low accuracy and/or with wrong coordinates. Therefore, it is necessary to apply techniques to reduce these noises. Identifying the mean of transport used along each trajectory is also necessary, since data of different means could be mixed together and just those from vehicles are relevant for road maps. Additionally, compression of trajectories can be used to reduce the quantity of less meaningful points and thus reduce the time for processing the whole set of collected GPS data.

The main challenge in automatically generating road maps using GPS data is related to the accuracy of collected points, since it interferes directly with the exact identification of the road center. Empirical tests have found points with an accuracy radius ranging from 2 m up to 50 m, which is too much for generating maps. For this reason, this work assumes that none of the collected points are correctly positioned at the center of the road and, thus, new ones should be created combining the data of nearby points. Due to the high volume of collected points it is not possible to deterministically create these center points. Therefore, it is necessary to use an approximation method, and this work assumes that a Genetic Algorithm can be a suitable choice.

Each point created by the Genetic Algorithm represents the center of a small portion of the road. So, after processing the whole set of collected points to find all center points, they are connected sequentially to finally generate the road map.

As new trajectories are collected, the map needs to be updated periodically. To do this, the approximation method should recreate the center points using both old and new collected points, but prioritizing more recent data. Therefore, the method must take into consideration not only the accuracy of each point, but also the date when each trajectory has been recorded.

Based on the challenges and assumptions presented, the objective of this work is to propose a structured method capable to cleaning, analysing and enriching moving objects' trajectories in order to generate accurate digital road maps using Genetic Algorithm.

The methodological procedure utilized in this work begins with a literature review of related work. After that, the following subjects are reviewed: techniques for identifying transportation means, reducing noise, compressing trajectories and Genetic Algorithm. In parallel, a system to collect and store trajectories for testing purposes is developed. Based on the knowledge acquired from related works an initial solution is defined. Then, this solution evolves cyclically with the gradual implementation of reviewed techniques and analysis of test results. The evaluation of the maps generated by the proposed method is performed qualitatively by comparing them with Google Maps [8], and quantitatively by calculating metrics between them and map data from OpenStreetMap [21]. In all cases the test scenarios are places from the city of Joinville-Brazil. Qualitative test results indicate that the generated maps are visually similar to the reference ones. Likewise, quantitative test results reinforce this conclusion, since the average difference compared to OpenStreetMap is only 1.91 m, with a standard deviation of 1.52 m.

Additionally, a modified version of the method using Fuzzy C-Means instead of Genetic Algorithm is used to compare both algorithms. Results show that the choice of Genetic Algorithm was correct, since although it is a little slower, it is more accurate.

The paper is structured as follows. Section 2 presents related work and compares them with the solution proposed here. Section 3 details the proposed method. Section 4 presents the evaluation of results. Finally, conclusion and possible future work are discussed.

2. Other proposals to generate road maps

The literature related to generation of maps contains many different proposals on how to create center points. Brüntrup et al. [4] do it with trajectories from any GPS receiver, but assume that only motor vehicles were used. After a client device collects the trajectories, they are sent to a server to remove noise based on limits of speed, acceleration, and relation between distance and time for two sequential points. After that, it divides each trajectory into segments and, for each of them, uses a clustering technique based on Artificial Intelligence to identify which points should be added to the map. This method does not depend on an initial map, so it can create maps from scratch. Also, the authors mention that the map can be updated using the same method.

Cao and Krumm's [5] work uses GPS data recorded with in-vehicle GPS loggers. The trajectories are filtered by the application server based on limits of distance, time and angle of movement between two sequential points. After that the center points are created using a technique developed by them which is based on principles of physical attraction and repulsion. That is, nearby trajectories likely to represent the same road are grouped together, while points with opposed angle of movement (e.g. two-way roads) are repelled from each other. As in the previous work, this method does not depend on an initial map. The authors do not discuss about updating the map.

Jang et al. [11] collect trajectories using mobile devices equipped with GPS receivers. Like Brüntrup et al. [4], they also assume that only motor vehicles were used. No steps for filtering the trajectories have been described. To generate the map, they first divide the area where trajectories have been recorded in squares of 1 m. After that, the points on each square are clustered based on the distance to nearby squares. The clusters are then connected to make the road network, and the analysis of the shape and angle of the streets is used to correct any problems. This step is also not detailed by the authors. As in the previous works, they do not use an initial map. Finally, the authors comment that the method to update the map has not been automatized yet.

Zhang et al. [27] use trajectories publicly available at OpenStreetMap as source of data for their method. In contrast to the previous works, this one aims to update existing maps instead of generating one from scratch. Initially, their solution divides trajectories in segments according to the velocity of the points and the distance between them. After that, perpendicular lines are drawn every few meters for each road of an initial map. The set of points where each trajectory crosses a given line is filtered by the distance of each point to the road and their angle of movement. Finally, the resulting set serves as input to a robust estimation method. The points with at least 95% of confidence are considered the new center points.

Niu et al. [20] use smartphones (Blackberry and iPhone) as client devices to collect trajectories, but assume that these trajectories are only from motor vehicles. Their method to create center points begins by filtering trajectories based on limits of accuracy and angle of movement. Points with speed zero or repeated points in the same coordinates are discarded. Finally, they use a combination of Robust Loess and subtractive clustering. As in the previous work, this one depends on an initial map as reference.

Although each work proposes a different method to create center points, most of them follow common ideas. For instance, the independence of initial maps. Only Niu et al. [20] and Zhang et al.