HAL
open science

A Collaborative Framework for High-Definition Mapping

Alexis Stoven-Dubois, Kuntima Kiala Miguel, Aziz Dziri, Bertrand Leroy, Roland Chapuis

To cite this version:

Alexis Stoven-Dubois, Kuntima Kiala Miguel, Aziz Dziri, Bertrand Leroy, Roland Chapuis. A Collaborative Framework for High-Definition Mapping. IEEE Intelligent Transportation Systems Conference (ITSC 2019), Oct 2019, Auckland, New Zealand. pp.1845-1850, 10.1109/ITSC.2019.8917292 . hal-02908399v2

HAL Id: hal-02908399

https://hal.science/hal-02908399v2

Submitted on 9 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Collaborative Framework for High-Definition Mapping

Alexis Stoven-Dubois[1], Kuntima Kiala Miguel[1], Aziz Dziri[1], Bertrand Leroy[1] and Roland Chapuis[2]

***Abstract*— For connected vehicles to have a substantial effect on road safety, it is required that accurate positions and trajectories can be shared. To this end, all vehicles must be accurately geo-localized in a common frame. This can be achieved by merging GNSS (Global Navigation Satellite System) information and visual observations matched with a map of geo-positioned landmarks. Building such a map remains a challenge, and current solutions are facing strong cost-related limitations.**

We present a collaborative framework for high-definition mapping, in which vehicles equipped with standard sensors, such as a GNSS receiver and a mono-visual camera, update a map of geo-localized landmarks. Our system is composed of two processing blocks: the first one is embedded in each vehicle, and aims at geo-localizing the vehicle and the detected feature marks. The second is operated on cloud servers, and uses observations from all the vehicles to compute updates for the map of geo-positioned landmarks. As the map's landmarks are detected and positioned by more and more vehicles, the accuracy of the map increases, eventually converging in probability towards a null error. The landmarks' geo-positions are estimated in a stable and scalable way, enabling to provide dynamic map updates in an automatic manner.

***Index Terms*— Collaborative Techniques & Systems, Accurate Global Positioning, Sensing, Vision and Perception**

I. Introduction

Intelligent communications will soon be enhancing all standard and autonomous vehicles on the roads. This new development is expected to have a significant impact on road safety, by allowing collective and real-time exchange of positions and observations, including infrastructure status (e.g. construction works) and locations of unplanned events (e.g. road accidents), between vehicles. Nevertheless, achieving such a safety improvement requires that vehicles and other road data are accurately geo-localized.

Geolocalization performed using GNSS (Global Navigation Satellite System) information cannot guarantee accurate positioning, especially in "urban canyons". Another localization strategy consists in using embedded visual sensors, and matching images over accurately positioned landmarks, hence requiring the preliminary development of a high-definition map containing geo-localized landmarks. Recently, major actors in the field have tried building high-definition maps by deploying fleets of vehicles equipped with high-end sensors. Having elaborated high-definition maps for the major highways, they are now facing strong logistical and economical limitations, and are not considering to register entire road networks in the near future [1]. Furthermore, due to the usage of dedicated vehicles fleets equipped with high-end sensors, live update of the maps can not be considered.

Instead of relying on fleets of dedicated vehicles, we intend to efficiently use production vehicles equipped with GNSS receivers and front mono-visual cameras, which will soon be part of the standard equipment. These vehicles, whether they are man- or self-driven, will collaborate in the build-up of a geo-localized map by visually identifying landmarks and measuring their geo-positions. The accuracy of such a map will be the outcome of the crowdsourcing of a vast amount of geo-position measurements. In turn, these vehicles could make use of this map to enhance their safety by accurately geolocalizing themselves within the infrastructure, and sharing accurate positions and trajectories.

The main contribution of this paper is the proposition of a scalable collaborative framework for high-definition mapping which overcomes the economic lock attached to the making of high-definition maps on large territories, and allows for constant updates of the map. We show that the incremental process of refining landmarks' geo-positions through measurement crowdsourcing allows for a refinement of the map accuracy.

This paper is organized as follows: in the next section, we discuss previous methods for mapping areas and geolocalizing the maps. Next, we present the framework for our crowdsourced mapping solution. Then, we show results of the first experiments of our geo-localized map-building application. Finally, we discuss the results and detail next directions for the improvement of our system.

II. State-of-the-Art

Connected vehicles and autonomous ones require accurate geolocalization, eventually up to the decimeter-level, for efficient data-sharing [2]. The use of GNSS receivers is not sufficient to achieve such an accuracy, as even more expensive RTK-GNSS technology suffer from multi-path and unavailability issues in urban areas [3]. On the other hand, the use of highly accurate landmarks allows the vehicle to geo-localize itself using its embedded sensors, even in urban environments, as shown in [4] and [5], where a map of pole-shaped landmarks and a map of traffic signs are respectively used.

Those two maps are only built in an unscalable way and designed to be applied within restricted experiments zones. The building of a large map made of accurate geo-localized landmarks remains a real challenge, as the map must be stable to new updates by the different vehicles, and scalable, i.e. have low computations and storage requirements.

[1]Mobility Department, VEDECOM, F-78000 Versailles, France alexis.stoven-dubois@vedecom.fr

[2]Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, F-63000 Clermont-Ferrand, France roland.chapuis@uca.fr

Photogrammetry applications seek to build geo-localized maps of various objects or landscapes with the highest possible accuracy [6]. To achieve this, feature points are extracted and matched within massive amounts of images, additional sensing outputs are added if available, and a global optimization (bundle adjustment) is operated [7]. geo-localized data can be included through the use of GNSS sensors [8], making the accuracy sensitive to GNSS flaws, and by manually installing geo-localized anchor points in the scene [9], which is a tedious and expensive operation. This process must be performed offline, and is unscalable for producing large maps of accurate landmarks.

SLAM (Simultaneous Localization and Mapping) techniques have aimed to build maps as a support for various robots' localization and navigation algorithms [10]. Although the process is similar to the photogrammetry one, filtering [11] or local optimization [12], which provide less accurate maps than with a global optimization, are generally preferred due to real-time requirements. Recently, C-SLAM methods have enabled different vehicles to map and position themselves within a common reference frame. However, such techniques induce intensive communication and computation requirements, and are not suited for building large geo-localized maps. Moreover, SLAM-based maps are often built with regard to the vehicle's reference, and can only be geo-referenced through the use of GNSS sensors, again making the map's accuracy sensitive to GNSS weaknesses [3].

Our proposition is an incremental process which crowdsources measurements from standard vehicles to update a map of geo-localized landmarks through a centralized optimization process. Our framework starts with an empty map, and does not require any prior knowledge regarding the landmarks' configuration. Collecting observations from a vast amount of vehicles provides dynamic updates of the map, decoupling the localization and mapping warrants stability to communications outages, and strongly selecting the landmarks to register ensures long-term scalability.

III. System Overview

For our framework to be fully cooperative, we consider all connected vehicles equipped with the following set of standard sensors: a GNSS receiver and a mono-visual camera. Fig. 1 illustrates this framework, which can be split into onboard perception and localization tasks, and a cloud processing.

A. Onboard Processing

The onboard processing operations described below are implemented on each vehicle to process raw data delivered by the sensors. They aim to localize both the vehicle and the feature marks, and to communicate such information to cloud servers.

1) Perception: The *Perception* block receives images from the camera as inputs, detects and describes feature marks within them, and provides their descriptions as outputs. As these descriptions will be matched on cloud servers with the descriptions of the map's landmarks, they must be robust to variable imaging conditions, including scale and environmental and illumination changes. This ensures that vehicles driving through the same areas detect similar feature marks.

Urban areas include many roadside elements that can be used as feature marks. For this purpose, we have chosen to use traffic signs, as they are semantic objects which can be robustly matched [13], and frequently observed especially in urban environments. For detecting traffic signs, a CNN-derived architecture is used [14], which provides corresponding bounding boxes. Each traffic sign observed in an image is described using both its bounding box pixel-position, and its semantic information. It can be noted that other types of feature marks can be considered simultaneously and added to the *Perception* block in the future.

2) Vehicle and Feature Marks Geolocalization: This block's inputs are the geo-positions from the GNSS receiver and the descriptions of feature marks from the *Perception* block. Based on this information, it computes the geolocalization of both the vehicle and the traffic signs, and provides geo-localized feature marks observations as outputs. Those latters are registered on the vehicle and uploaded to cloud servers whenever possible.

All vehicles and feature marks are to be geo-localized within a world frame. The vehicle and its sensors relate to three different frames, as depicted in Fig 2: its own frame situated at its center, the GNSS receiver frame, and the camera frame. Relations between these frames are obtained from an extrinsic calibration procedure, which provides:

- The transform matrix T_V^G linking the vehicle frame to the GNSS receiver frame.
- The transform matrix T_V^C linking the vehicle frame to the camera frame.

The true states X^V, X^G and X^C of the vehicle, GNSS receiver and camera are defined by their respective geo-position and heading. The GNSS receiver provides an observation Z^G of its state X^G, including both its geo-position and heading (as the vehicle is moving):

$$Z^G = g(X^G, w) \tag{1}$$

with g being the GNSS receiver's observation model, and w being its noise. Making use of T_V^G linking the GNSS frame to the vehicle frame, we compute the estimation $\hat{X}^V$ of the vehicle state directly from the GNSS observation Z^G. As for the camera, an estimation $\hat{X}^C$ of its state is computed from $\hat{X}^V$, making use of T_V^C linking the camera frame to the vehicle frame.

Each feature mark description is associated with an observation of its bounding box pixel-position $Z^D = \begin{pmatrix} u & v \end{pmatrix}^T$.

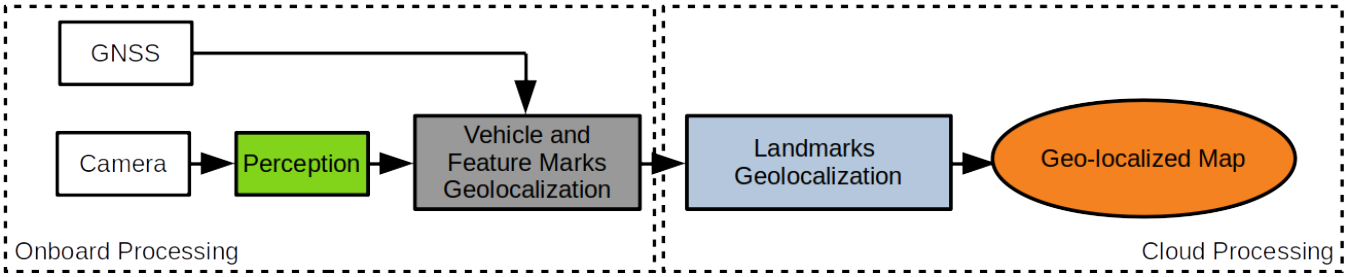

Fig. 1. Collaborative High-Definition Mapping Pipeline - Each vehicle visually detects feature marks. Using its embedded sensors, it geo-localizes both itself and the feature marks. Cloud servers, receiving geo-positioned feature marks from several vehicles, can update the map of geo-localized landmarks.

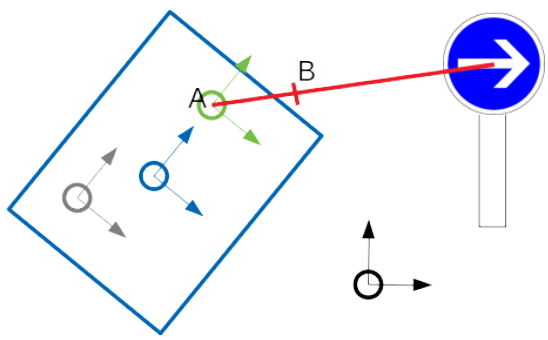

Fig. 2. 2D Representation of Frames and Projection Lines (View from the top) - The world frame is depicted (black), along with the vehicle frame (blue), the camera frame (green), and the GNSS receiver frame (grey). The projection line and its two points A and B originating from the detection of the traffic sign are also shown (red).

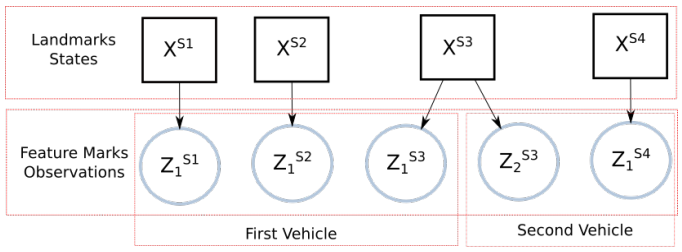

Fig. 3. Landmarks Geolocalization Example - The cloud server has received feature marks observations from two different vehicles, and has matched them with various traffic signs. The known observations are represented as round factors, the states to estimate as square factors, and the joint constraints as directed arrows.

Knowing the camera intrinsic calibration matrix K, a projection line linking the camera center to the traffic sign, as depicted in Fig. 2, can be established and modeled as passing through two points A and B [7]:

$$A = \hat{X}^C \tag{2}$$

$$B = \hat{X}^C + (T_W^C)^{-1} K^{-1} \begin{pmatrix} u & v & 1 \end{pmatrix}^T \tag{3}$$

with T_W^C the transform matrix linking the camera frame to the world frame. The state X^S of a traffic sign is defined by its geo-position, thus each projection line (A_i, B_i) originating from a detection of this sign constitutes an observation Z^S of its geo-position.

Finally, the traffic sign observation Z^S, along with the traffic sign description, must be uploaded to cloud servers.

B. Cloud Processing

The cloud servers receive feature marks observations Z^S as inputs, use them to update and improve the map of geo-localized traffic signs, and provide map updates as outputs. Such map-building operation does not have to be processed in real-time, but has to be stable to new updates and scalable. No prior knowledge on the map is required, as map updates are able to initialize the map on their own.

First, feature marks observations are matched with the map's traffic signs, by comparing their respective descriptions. Our decision to focus on the detection of traffic signs greatly facilitates this task, as the descriptions contain strongly distinguishable semantic information [13]. Furthermore, the limited number of traffic signs warrants that this operation remains scalable.

Then, the map is updated, using the feature marks observations to either confirm or infirm the landmarks' geo-positions. This operation can be modeled as an optimization problem, as shown in Fig. 3, with:

- The traffic signs' geo-positions X^{S_i} to estimate for each traffic sign i.
- The observations $Z_j^{S_i}$ of the traffic signs' geo-positions, with $Z_j^{S_i}$ being the jth projection line linking a camera state estimation $\hat{X}_C$ to the traffic sign i.

Considering a traffic sign i, all of its previous and new observations $Z_j^{S_i}$ are registered and used within the optimization. The new best estimation $\hat{X}^{S_i}$ for the traffic sign geo-position can be obtained through a least-squares triangulation-based optimization, minimizing:

$$\hat{X}^{S_i} = \underset{X^{S_i}}{\operatorname{argmin}} \sum_j dist(X^{S_i}, Z_j^{S_i}) \tag{4}$$

with $dist(X^S, Z^S)$ being the orthogonal distance between the geo-position X^S and a projection line Z^S.

IV. Experiments

To assess the effectiveness of the proposed solution, a two-step experimentation was performed. First, we validate our solution through a simulation. Then, we confront it to real data.

A. Simulation

Each vehicle detecting a traffic sign i provides observations $Z_j^{S_i}$ of its geo-position. Both the detection of the traffic sign and the vehicle's geo-localization suffer from some noise, leading to noisy projection lines $Z_j^{S_i}$. However, as more and more measurements are received, cloud servers are able to estimate the geo-position $\hat{X}^{S_i}$ of the traffic sign with an increasing accuracy, eventually converging towards a null error. To validate this, a simulation of our solution was implemented, without any loss of generality, as a 2D simulation on the North-East plane.

- First, a traffic sign $i = 1$ and its true position X^{S_1} are defined along a straight road.
- For a chosen number n of different vehicle passings, vehicles true states X^V are generated.
- For each vehicle state X^V, an image is associated, and the corresponding true pixel-position X^D for the bounding box of the traffic sign is computed.
- Next, GNSS observations Z^V are generated by applying a random, white noise of $5.0\,m$ for the positions and $0.35\,rd$ for the orientations around the vehicles states X^V.
- Similarly, feature marks descriptions Z^D outputted from the *Perception* block are generated applying a random, white noise of $5\,pixels$ around the pixel-positions X^D.
- GNSS observations Z^V and feature marks descriptions Z^D are fed as inputs for the *Vehicle and Feature Marks Geolocalization* block, which establishes feature marks observations, i.e. projection lines $Z_j^{S_1}$.

The error associated to GNSS measurements is not white and can be affected by strong biases, due to atmospheric conditions and multi-path issues [15]. Nevertheless, our collaborative approach estimates the landmarks geo-positions using many different measurements acquired during a large time span (with different atmospheric conditions), and by different vehicles with slightly different positions on the road (leading to different multi-path effects). Therefore, we can consider that our assumption for a white distribution error still stands.

At each vehicle passing, the *Landmarks Geolocalization* block is activated, and an estimation $\hat{X}^{S_1}$ of the traffic sign's geo-position is computed. During the optimization, all feature marks observations $Z_j^{S_1}$ received at the current vehicle passing and at the previous ones are used. Within our simulation, we simplified the function to optimize:

$$\hat{X}^{S_1} = \underset{X^{S_1}}{\operatorname{argmin}} \sum_j dist(X^{S_1}, Z_j^{S_1}) \tag{5}$$

with:

$$dist(X^S, Z^S) = ||head(proj(X^S, Z^S)) - head(Z^S)|| \tag{6}$$

where $head(Z^S)$ is the heading (angle to the North axis) of Z^S, and $proj(X^S, Z^S)$ is the projection line passing through X^S and the camera position $\hat{X}^C$ associated with Z^S. Such a simplification enabled us to estimate at each vehicle passing, not only the traffic sign's geo-position $\hat{X}^{S_1}$, but also its covariance Σ^{S_1} and deviations σ^{S_1}, by applying directly the method from [16].

Results of our simulation are shown in Fig. 4. For each vehicle passing, a single-passing measurement is also computed using only projection lines $Z_j^{S_1}$ from that passing in the optimization described upper. Errors on the East and North axes for single-passing measurements and for estimations of our collaborative approach are computed as simple differences from the groundtruth X^{S_1} , and shown. Also, deviations related to the estimations of our approach are depicted as $[-2\sigma^{S_1}; 2\sigma^{S_1}]$ ranges centered around our solution's estimations.

The results of this simulation confirmed the theoretical assumption of convergence as we see that the deviations for the $East$ and $North$ axes effectively decrease as the number of vehicle passings increases, and that they even decrease with the square root of the number of observations. This indicates that, as more observations are received, the estimation of the traffic sign's geo-position converges towards a specific location. Further, we observe that the groundtruth (i.e. $error = 0$) is always comprised within the deviations ranges, indicating that the estimation of the traffic sign's geo-position effectively converges towards its true location, hence towards a null error.

B. Real Experiments

Having verified the convergence hypothesis of our crowd-sourced approach through a simulation, we have also performed a field-experiment to confirm the effectiveness of the solution in real conditions. The experiment consisted in driving a vehicle, equipped with a standard GNSS sensor and a front-looking mono-visual camera, on a 4 km loop within 4 hours, allowing to acquire data for 10 vehicle passings on the loop. Data from the 10 vehicle passings was shuffled randomly, so as to avoid any bias due to the drawing order. Furthermore, the positions of 10 traffic signs along the loop were acquired manually using an RTK-GPS receiver, providing a groundtruth with which the results of our approach have been compared.

For each considered traffic sign $i = 1, ..., 10$, a number of detections occured at each vehicle passing, leading to feature marks descriptions outputted from the *Perception* block. Feeding both those and GNSS observations Z^V as inputs for the *Vehicle and Feature Marks Geolocalization* block, feature marks observations (projection lines) $Z_j^{S_i}$ are outputted. Finally, at the end of each vehicle passing, the regular optimization of the *Landmarks Geolocalization* block is processed, giving an estimation $\hat{X}^{S_i}$ of the geo-position of each traffic sign i.

The results obtained for the different traffic signs are depicted in Fig. 5. As previously for the simulation, single-passing measurements are also computed at each vehicle passing using only the feature marks observations $Z_j^{S_i}$ from that passing in the regular optimization. Errors for estimations of our collaborative approach and for single-passing measurements are computed as distances from the groundtruth X^{S_i}, and shown.

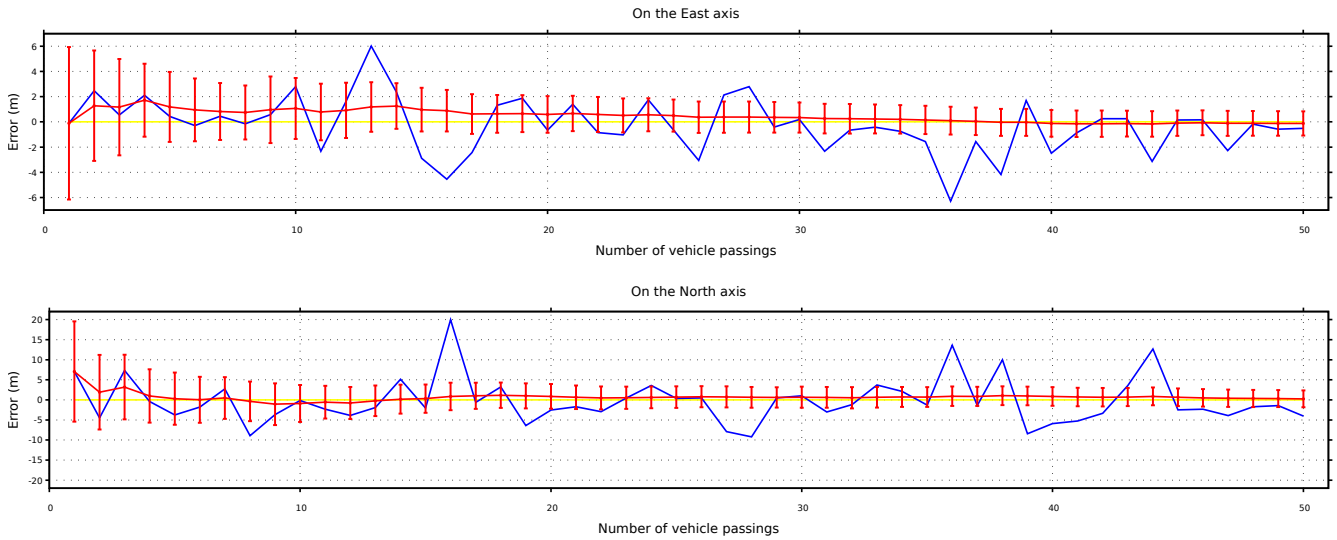

Fig. 4. Simulation Results - Errors on the $East$ and $North$ axes for single-passing measurements (blue) and for estimations of our collaborative approach (red) are shown, as well as the groundtruth (i.e. $error = 0$, yellow). Deviations related to the estimations of our solution (red) are also depicted as $[-2\sigma^{S_1}; 2\sigma^{S_1}]$ ranges.

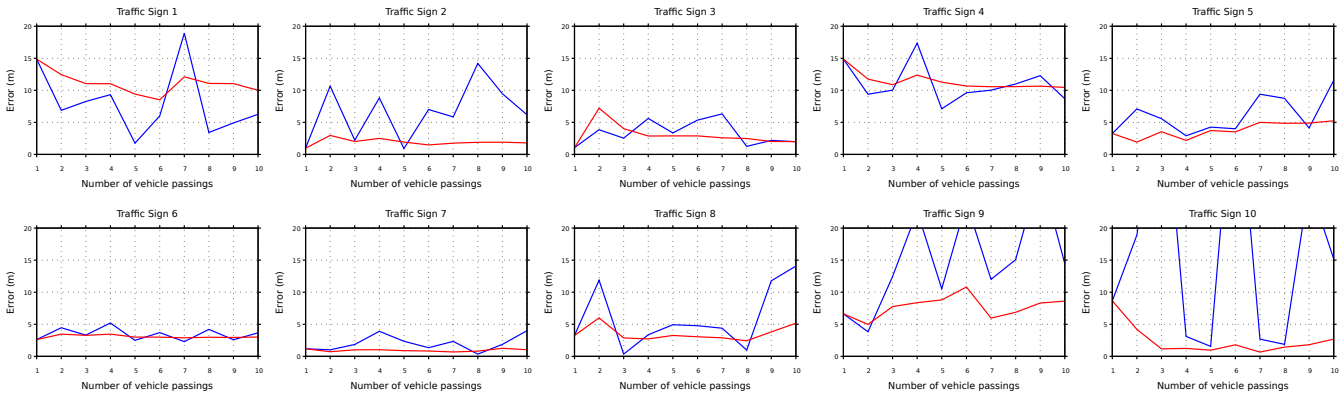

Fig. 5. Field-tests Results - Distance errors for single-passing measurements (blue) and for estimations of our collaborative approach (red). Distance errors for single-passing measurements of the traffic signs $i = 9$ and $i = 10$ are high in some cases. For visibility purposes, they are not depicted.

The results show that single-passing measurements for the traffic signs' geo-positions may be extremely inaccurate in some cases. This corresponds to a low quality of vehicle geolocalization provided by the GNSS receiver, as our vehicle was driven in a condensed urban zone. In the meanwhile, estimations of our collaborative approach quickly surpass the average accuracy of single-passing measurements, indicating that our crowdsourced approach gives better performances. However, the convergence in probability of the error confirmed by simulation is not observable here, due to the small amount of vehicle passings acquired during the tests.

In order to observe the convergence using the data acquired during the field-test, we propose to build a dataset according to the two following principles:

- We take the first traffic sign as a reference, and translate all the geo-positions observations $Z_j^{S_i}$ of the other traffic signs $i = 2, ..., 10$ to the reference geo-position. This allows to transform the 10 traffic signs with 10 passings into 1 traffic sign observed during 100 passings.

$$t_l^1 = t(X^{S_l} - X^{S_1}) \ ; \ l = 2, ..., 10 \tag{7}$$

  where t_l^1 is the translation allowing to superpose observations of the traffic sign $i = l$ to the reference traffic sign $i = 1$. X^{S_l} is the groundtruth position of the traffic sign $i = l$.
- Furthermore, to avoid any bias due to drawing order, we compute 1000 permutations of these 100 vehicle passings, and average at each passing the error resulting of the geo-position estimation of our collaborative approach.

The results obtained using this new dataset are depicted in Fig. 6, and show that our collaborative approach outperforms single-passing measurements. As the number of vehicle passings increases, our method converges to an error of 3 meters

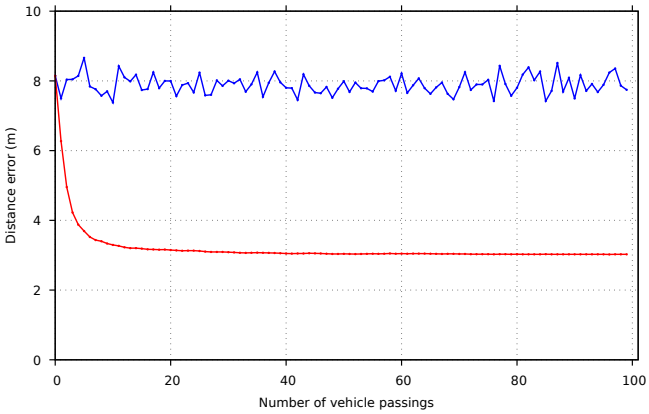

Fig. 6. Distance errors for single-passing measurements (blue) and for estimations of our collaborative approach (red), when the observations of all the traffic signs are superposed on the first traffic sign, generating 100 passings.

while single-passing measurements provide an average error of 8 meters. The magnitude of this error can be explained by an imprecision of the setup of the camera and GNSS receiver in the test-vehicle, due mostly to an inaccuracy of synchronization and extrinsic calibration of the camera with regard to the GNSS receiver. Nevertheless, this experiment confirms the hypothesis of convergence in probability and shows the effectiveness of the proposed approach.

V. Conclusion

In this paper, we have presented a collaborative mapping framework which will allow all connected vehicles equipped with a GNSS receiver and a mono-visual camera to position landmarks with accuracy. The map is collaboratively built out of landmark geo-position measurements performed by standard vehicles. This crowdsourcing approach, while being scalable and cost-effective, allows for continuous updates of the map. Finally, a field-test has been presented, showing that our collaborative approach has a better positioning accuracy in average than independent geo-position measurements. While field experiments are still going on in order to gather more data, several enhancements can be envisioned:

- Our proposition only considers traffic signs for landmark detection. This could be generalized to other types of features marks and objects, which will be especially useful in areas where traffic signs are scarce.
- The present implementation assumes that the landmarks in the map are manually selected during initialization. An enhancement could consist in implementing a dynamic management of the feature marks, where new landmarks could be proposed by vehicles and obsolete ones could be revoked by the cloud servers. Doing so will allow to automatically obtain a higher density of landmarks, and result in a more efficient updating of the map.

References

[1] H. G. Seif and X. Hu, "Autonomous Driving in the iCity—HD Maps as a Key Challenge of the Automotive Industry," *Engineering*, vol. 2, no. 2, pp. 159 – 162, 2016.

[2] J. Ziegler et al., "Making Bertha Drive—An Autonomous Journey on a Historic Route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, pp. 8–20, Summer 2014.

[3] G. Bresson et al., "Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving," *IEEE Transactions on Intelligent Vehicles*, vol. 2, pp. 194–220, Sep. 2017.

[4] R. Spangenberg, D. Goehring, and R. Rojas, "Pole-based localization for autonomous vehicles in urban scenarios," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2161–2166, Oct 2016.

[5] M. Buczko and V. Willert, "Efficient global localization using vision and digital offline map," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, Oct 2017.

[6] M. J. Westoby et al., "'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications," *Geomorphology*, vol. 179, pp. 300 – 314, 2012.

[7] O. Özyeşil et al., "A survey of structure from motion," *Acta Numerica*, vol. 26, p. 305–364, 2017.

[8] M. Rehak and J. Skaloud, "Applicability of New Approaches of Sensor Orientation to Micro Aerial Vehicles," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, pp. 441–447, June 2016.

[9] S. Harwin and A. Lucieer, "Assessing the Accuracy of Georeferenced Point Clouds Produced via Multi-View Stereopsis from Unmanned Aerial Vehicle (UAV) Imagery," *Remote Sensing*, vol. 4, no. 6, pp. 1573–1599, 2012.

[10] B. Suger and W. Burgard, "Global outer-urban navigation with OpenStreetMap," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1417–1422, May 2017.

[11] P. Tiefenbacher, T. Schulze, and G. Rigoll, "Off-the-shelf sensor integration for mono-SLAM on smart devices," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 15–20, June 2015.

[12] K. Pirker, M. Rüther, and H. Bischof, "CD SLAM - continuous localization and mapping in a dynamic world," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3990–3997, Sep. 2011.

[13] T. Botterill, S. Mills, and R. Green, "Correcting Scale Drift by Object Recognition in Single-Camera SLAM," *IEEE Transactions on Cybernetics*, vol. 43, pp. 1767–1780, Dec 2013.

[14] Z. Zhu et al., "Traffic-Sign Detection and Classification in the Wild," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2110–2118, June 2016.

[15] Z. Tao and P. Bonnifait, "Modeling L1-GPS Errors for an Enhanced Data Fusion with Lane Marking Maps for Road Automated Vehicles," in *2015 European Navigation Conference (ENC)*, Apr. 2015.

[16] A. Eudes and M. Lhuillier, "Error propagations for local bundle adjustment," in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2411–2418, June 2009.