High Definition Collaborative Mapping

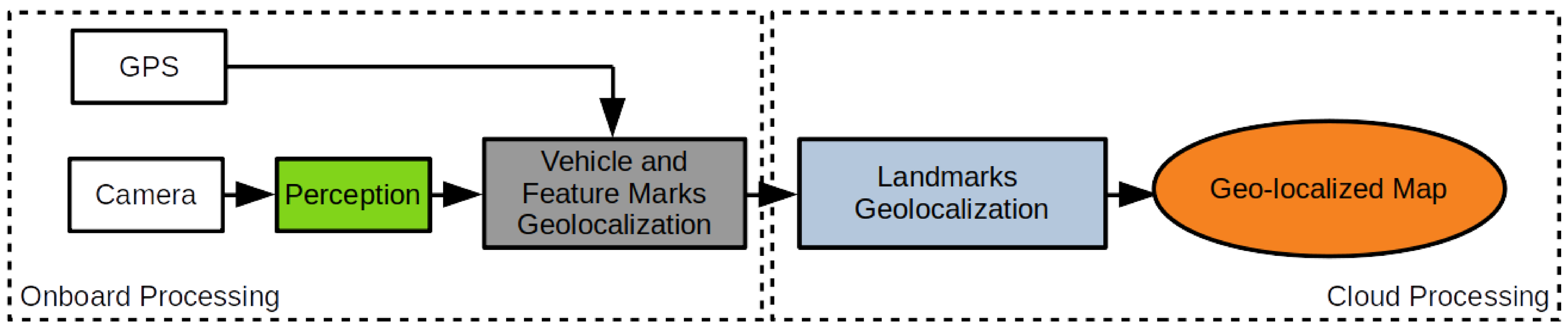
Alexis Stoven-Dubois[1], Kuntima Kiala Miguel[1], Aziz Dziri[1], Bertrand Leroy[1], Roland Chapuis[2]

[1]Vedecom - Mobility Department, [2]Université Clermont Auvergne - Institut Pascal

Introduction

- For connected vehicles to have a significant effect on road safety, it is required that they can be accurately geo-positioned within a common frame.
- While GNSS receivers lack of precision, another strategy consists in using visual sensors, and matching images over a map of accurately positioned landmarks.
- Major actors in the field have tried building maps by using fleets of vehicles equipped with high-quality sensors, but are now facing [1]:
  - Strong logistical costs for maintaining the fleets.
  - Slow rates for updating the maps.
- Instead, we intend to use production vehicles equipped with standard sensors, and crowdsource their individual observations.

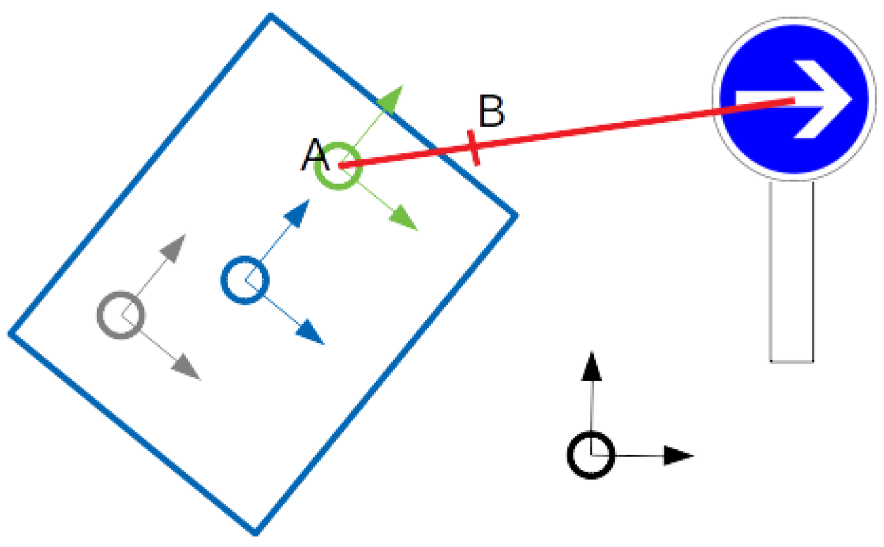
Methods



- First, the *Perception* block receives images from the camera, uses a CNN-derived architecture to detect traffic signs and establish bounding boxes [2], and outputs their descriptions.
- Next, the *Vehicle and Feature Marks Geolocalization* block receives positions from the GPS receiver, and traffic signs descriptions as inputs.

  The position and orientation of the camera is estimated directly from GPS readings. As a traffic sign is detected, a projection line is established linking its bounding box center to the camera center.

  Traffic signs observations, consisting each of a description and a geo-positioned projection line, are outputted.



- Cloud servers receive traffic signs observations from potentially several vehicles as inputs, and match them with their corresponding traffic sign in the map.

  For each traffic sign, a new estimation $\hat{X}$ of its geo-position is computed, using all of its associated projection lines Z, and applying a least-squares optimization:

$$\hat{X} = \min_{X} \sum dist(X, Z) \tag{1}$$

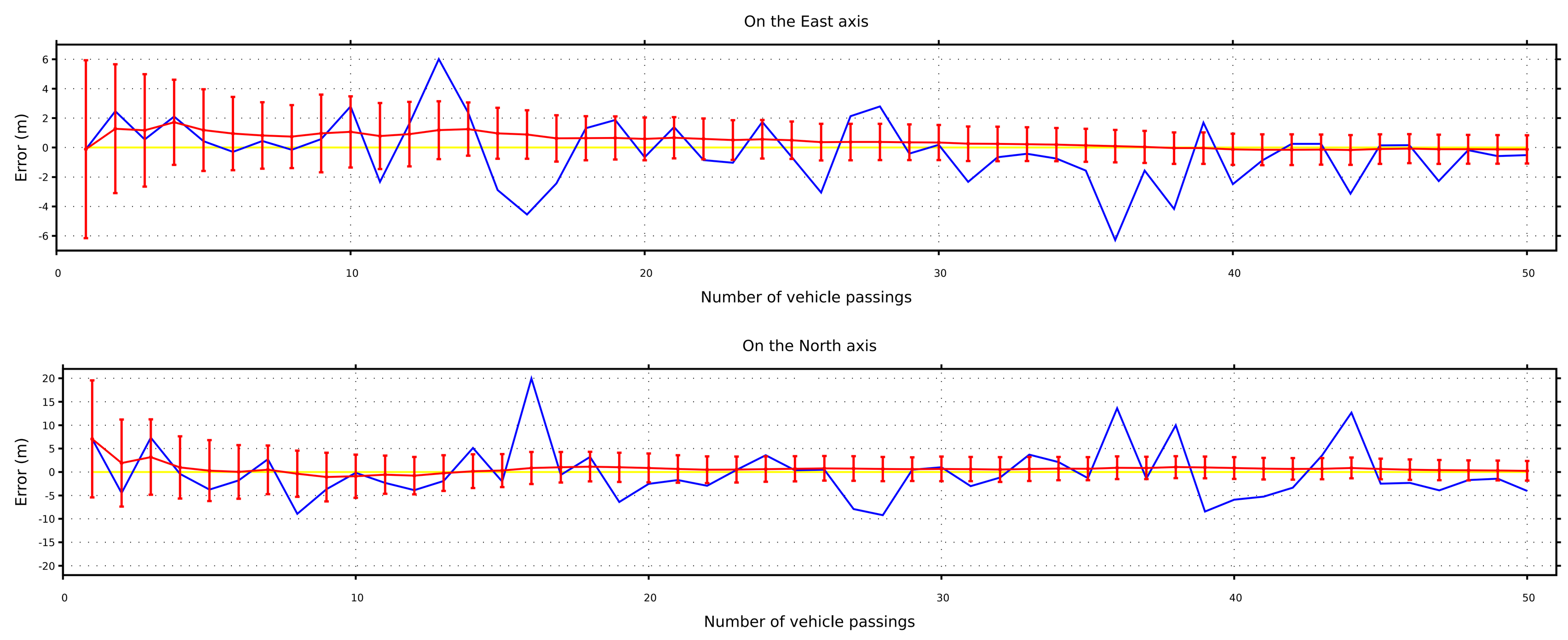
  with $dist(X, Z)$ being the orthogonal distance between the geo-location X and the projection line Z.

Conclusion

- Our simulation confirmed the hypothesis holding that the map accuracy converges towards a null error, as more vehicles detect the traffic signs.
- Our real experiments, despite a limited number of passings, could show a better performance in average than single-passing measurements.
- Future works include:
  - The implementation of deviations calculations for the regular optimization applied by the *Landmarks Geolocalization* block.
  - The extension of our solution to other types of landmarks, such as road markings or buildings.
  - The dynamic management of the map's landmarks.

Simulation Results

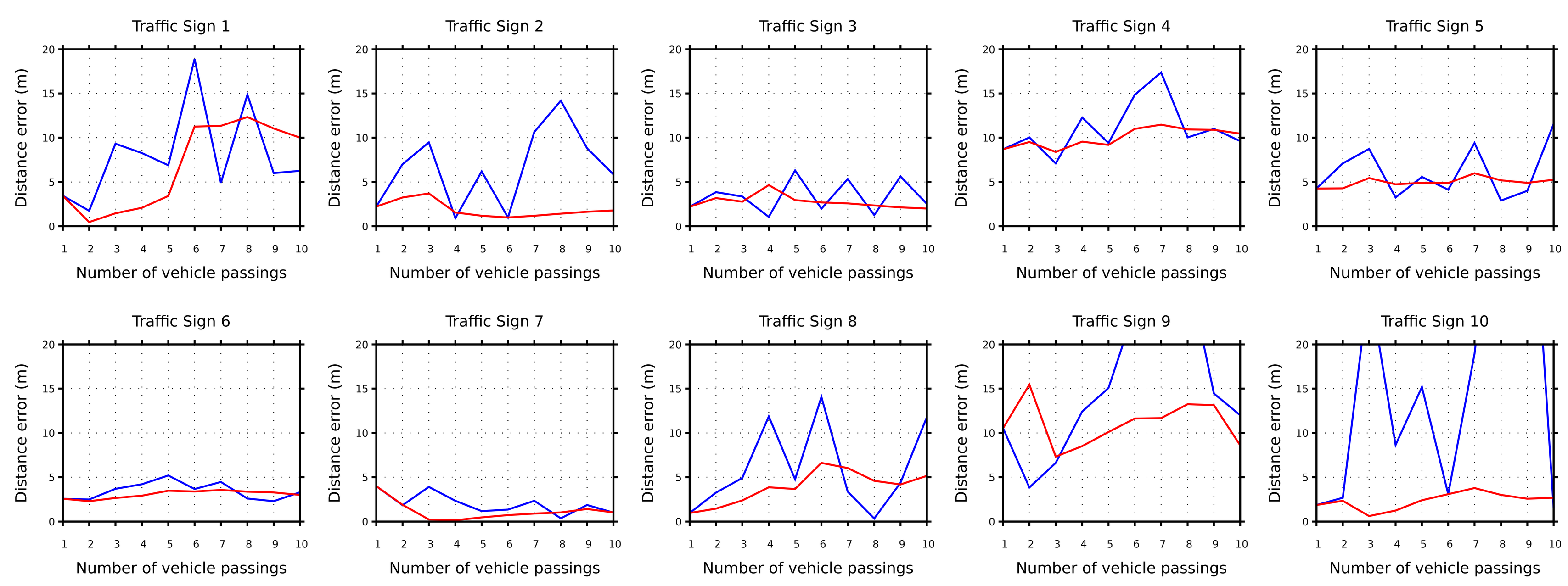
- A 2D simulation of our solution was implemented, in which a traffic sign was defined along a straight road, and vehicles trajectories were computed for several passings.
- Random, white noises of amplitude **5** m in position and **0.35** rd in orientation were applied around true vehicles trajectories to generate outputs from the GPS receiver.
- Random, white noises of amplitude **5** $pixels$ were applied around true bounding box centers to generate outputs from the *Perception* block.
- At each passing, a simplified optimization based on the yaw angles of projection lines was applied by the *Landmarks Geolocalization* block, enabling to compute associated deviations [3].



Simulation Results - Errors for single-passing measurements (blue) and for estimations of our approach (red) are shown, as well as the groundtruth (yellow). Deviations related to estimations of our approach are depicted as $[-2\delta; +2\delta]$ ranges.

Early Results

- A field-experiment was performed, in which a vehicle equipped with a standard GPS receiver and a mono-visual camera was driven for **4** hours on a **7** km loop, enabling to collect data for **10** passings along the loop.
- The geo-positions of **10** traffic signs were measured with an RTK-GPS receiver, constituting a groundtruth to compare our results with.
- At the end of each passing, the regular optimization was applied by the *Landmarks Geolocalization* block to estimate the geo-positions of all traffic signs:



Real Results - Distance errors for single-passing measurements (blue) and for estimations of our approach (red) are shown.

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, pp. 159–162, jun 2016.

[2] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-Sign Detection and Classification in the Wild," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2110–2118, IEEE, jun 2016.

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

alexis.stoven-dubois@vedecom.fr