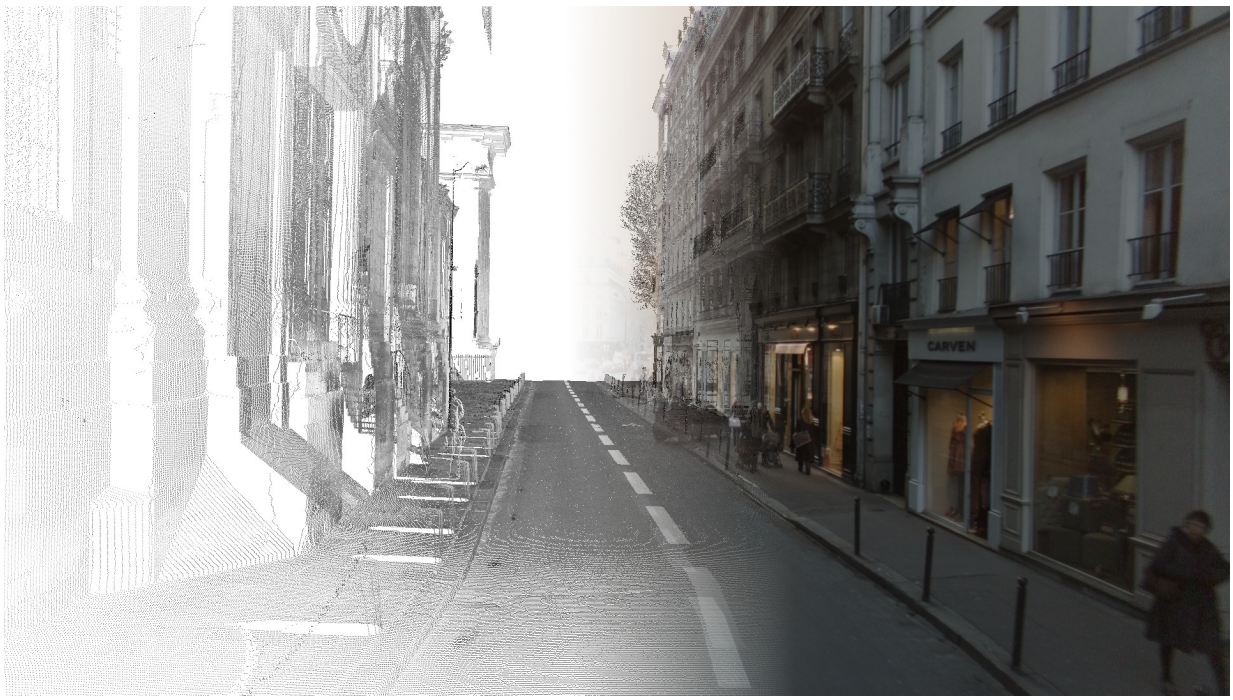
PhD proposal : Hybrid urban scene analysis from mobile mapping images and laser scan



Superimposed image and laser scan of the same street

1 Context

This PhD proposal, in the context of the National Project ANR PlaTINUM, aims at developing a geolocalized global frame, stored on a cloud and made of a set of 3D representations based on geometric, photometric and semantic information. This geographical information system (SIG) must be able both to update and to enrich automatically its content thanks to data transmitted by remote agents and also to provide a navigation service. This double objective in this collaborative application consists in using information stored in the cloud in order to help the agents during their navigation tasks, which in return will inform the cloud about the local detected differences which will be treated afterwards in order to update the geolocalized central reference.

Mobile mapping is a data acquisition technique that relies on sensors mounted on a terrestrial vehicule equipped with a georeferencing system allowing a very precise geolocalization of the acquired data. The interest in this technique has been increasing over the past decade as it allows for the collection of very accurate data at the scale of an entire city. However, processing such data is limited by various difficulties specific to this type of acquisition :

- Very high data volumetry (1 To by day of acquisition) which calls for very efficient processing tools in terms of number of operations and memory footprint.
- Scene complexity : whereas many mapping techniques rely on a 2D or 2.5D scene analysis, this becomes impossible when the sensor is within the scene. Mobile mapping data needs to be analysed in full 3D, and by handling the specificities of the acquisition geometry (occlusions, high level of detail, ...)
- Acquisition geometry : this type of acquisition induces a variable and anisotropic sampling.

Because mobile mapping collects data that is both detailed and precisely georeferenced, it can be used to precisely localise other sources of data with much less accurate (or no) localisation, which is of utmost importance for navigation of people and autonomous vehicles and augmented reality. In particular, the robotics community has shown the importance of constructing a visual map of the environment in which a robot is supposed to operate [Meilland11]. Recent works also show that semantising these maps enhances their usability for localisation [Drouilly15]. Moreover, such a semantisation of the environment helps a robot perform its task. However, building a structured semantic representation an urban scene is a challenging task. It has been extensively studied from images by the computer vision community, but in this case the analysis is 2D and not 3D, and not necessarily consistent between various images of the same scene. Conversely, the problem of semantisation of urban laser scans is an active field of research in the laser scanning community [Weinmann15]. The main goal of this PhD is to combine information from both images and a laser point cloud of an urban environment to semantize it in 3D and in a consistent manner.

2 PhD goals

To work jointly with oriented images and laser scans acquired by a mobile mapping system, the first requirement is that both data are perfectly aligned [Miled16]. A first task of the PhD will be to build a 3D mesh representation from the mobile laser scans by adapting a reconstruction method from the state of the art [Berger14]. The RGB images will then be merged into a mesh texture by choosing the best image to project on each triangle based on visibility, resolution and regularity criteria [Waechter14]. Once this is done, a semantic classification will have to be performed on the textured mesh representation taking advantage of the photometry and the geometry of the scene. Relying on a deeper analysis of the state of the art in machine learning and computer vision, different approaches will be selected, implemented and compared on available representative databases. A particular focus will be done on the recent and promising deep learning methods based on DCNN (Deep Convolutional Neural Networks) [Eigen14] and on the use of pre-trained CNN built from generic databases [Krizhevsky12]. Based on this comparison study, the most appropriate method will be selected and validated on the IGN database. Finally, the PhD will investigate higher level representations of the scene such as :

- Semantic segmentation : the scene is segmented into individual objects with semantic labels (trees, cars, poles).
- Relationship graph : in addition, a graph can be constructed to encode the geometric relationships between these objects (above, below, in front, between, ...)

3 Bibliography

[Berger14] Matthew Berger, Andrea Tagliasacchi, Lee Seversky, Pierre Alliez, Joshua Levine, Andrei Sharf, Claudio Silva. State of the Art in Surface Reconstruction from Point Clouds. Eurographics 2014 - State of the Art Reports, Apr 2014, Strasbourg, France. 1 (1), pp.161-185, 2014, EUROGRAPHICS star report.

[Drouilly15] R. Drouilly, P. Rives, B. Morisset. Hybrid Metric-Topological-Semantic Mapping in Dynamic Environments. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'15, Hamburg, Germany, Septembre 2015.

[Eigen14] D. Eigen et R. Fergus. "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture". In : CoRR abs/1411.4734 (2014).

[Krizhevsky12] A. Krizhevsky, I. Sutskever et G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." In : NIPS. 2012, p. 1106–1114.

[Meilland11] M. Meilland, A.I. Comport, P. Rives. Dense visual mapping of large scale environments for real-time localisation. In IEEE Conference on Intelligent Robots and Systems, IROS'11, San Fransisco, USA, Septembre 2011

[Miled16] Mourad Miled, Bahman Soheilian, Emmanuel Habets, Bruno Vallet. Hybrid online mobile laser scanner calibration through image alignment by mutual information. XXIII Congress of the International Society of Photogrammetry and Remote Sensing, Prague, July 2016 (to appear)

[Waechter14] Michael Waechter, Nils Moehrle, and Michael Goesele. Let There Be Color ! Large-Scale Texturing of 3D Reconstructions. European Conference on Computer Vision (ECCV 2014), Zürich, Switzerland, 6-12 Sept. 2014.

[Weinmann15] Martin Weinmann, Boris Jutzi, Stefan Hinz, Clément Mallet. Semantic point cloud interpretation based on optimal neighbourhoods, relevant features and efficient classifiers. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 105, July 2015, Pages 286 ?304

4 Keywords

Classification, Machine learning, computer vision, laser scanning.

5 Technical Environnement

For practical reasons, the PhD work will be conducted in C++ under Ubuntu.

6 Requirements

- Masters degree in computer vision, machine learning or image (or signal) processing.
- Strong experience in C++ development
- Autonomy, pragmatism
- Strong interest in scientific research

7 Conditions

- The position will be for 3 years, and will start between september and december 2016.

— Main localisation : Laboratoire MATIS - Institut Géographique National, 73 avenue de Paris 94165 Saint Mandé, France
— As the thesis will be co-directed between IGN-Matis and INRIA-Lagadic group, the student will spend a significant part of his time at INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 Sophia Antipolis Cedex, France

Contacts

Applications consist of a cover letter describing how your research experience is relevant to the position and a resume sent to both advisors. You can also contact both advisors for more information about any aspect of the position and research topic :

— Bruno VALLET (IGN-Matis) , Phone : 01 43 98 80 81, Mail : bruno.vallet@ign.fr, Website : recherche.ign.fr

— Patrick RIVES (INRIA-Lagadic), Mail : Patrick.Rives@inria.fr, Website : Patrick.Rives-eng.html