



## PhD thesis: Vision-based localization with discriminative features from heterogeneous visual data

### Project's description

This PhD thesis will be part of a national project funded for next four years by the French National research Agency (ANR). The project, pLaTINUM (Long Term MappINg for Urban Mobility), involves INRIA Sophia Antipolis (Lagadic Team), LITIS Rouen, IGN Paris (MATIS lab.) and Le2i UMR CNRS 6306.

The main topic of these researches is related to autonomous navigation in complex environments, a topic that has attracted a lot of attention in the past decades, with now many proposals such as autonomous vehicles. A safe navigation is based on an accurate map of the environment, which is usually obtained offline and used during navigation. In a large and dynamic environment, such as a city, the map keeps changing, making essential the update and enhancement of the 3D map. In addition, to navigate and to be accurately localized, the moving agent/user, a vehicle or a person, should be able to request the map with precision and in near real-time, thus processing time must be reduced.

### Thesis work and objectives

In this thesis, we focus on the localization of the agent/user, and propose to study vision-based methods from a previously obtained 3D map of the scene. This 3D map contains a 3D reconstruction of the scene as well as RGB and semantic information. We assume that the 3D map is coded as a set of Spherical Images, a compact representation with RGB, depth and semantic information, well positioned in the environment.

Thus, the objective is to accurately localize a moving user equipped with a vision sensor (RGB or RGB-D camera) in this environment. This is a difficult task even if an approximate location of the user is known, due to large changes in viewpoint and appearance between the 3D model and the user's view (RGB or RGB-D image). Local visual features (keypoints) based methods are most likely to fail. Moreover, as the spherical images contain complementary information such as RGB, depth and other semantic information, it is relevant - but challenging - to exploit these heterogeneous data to extract meaningful descriptors.

Therefore, this thesis is focused on the **automatic extraction of discriminative features from heterogeneous visual data (color images, depth images, etc.)**. Literature on localization based on the extraction of visual features is rich, see for example [1-8]. In the thesis, we will follow recent developments in machine learning research and investigate deep learning strategies for this task, such as in [1,2,4], which proved their relevance for the extraction of discriminative features. In

particular, we will be interested in: i) the best way to represent the data according to the heterogeneous data available; ii) the automatic extraction of discriminative features from heterogeneous data; iii) a matching method taking into account the diversity of the data; and iv) a reduced processing time.

## **Application**

The thesis will take place between laboratories Le2i-UMR CNRS 6306 of Burgundy University (Le Creusot) and Matis of IGN (Paris area).

Interested candidates must send to the contacts, **in a single PDF before 30<sup>th</sup> April 2016**:

- a detailed résumé,
- a motivation letter explaining the interest in the topic and suggesting ideas for solutions,
- the names and contacts of at least two references.

## **Requirements for application**

- Candidates must hold, or be about to obtain before September 2016, a master degree in computer science, mathematics, computer vision or related fields.
- Candidates must show excellent programming skills in Matlab, C++, and be familiar with OpenCV/PCL libraries.
- **Knowledge of French is not a prerequisite.**

## **Contacts**

- dro-desire.sidibe[at]u-bourgogne.fr (Le2i)
- cedric.demonceaux[at]u-bourgogne.fr (Le2i)
- bahman.soheilian[at]ign.fr (Matis)
- valerie.gouet[at]ign.fr (Matis)

## **References**

- [1] '*Painting-to-3D model alignment via discriminative visual elements*', M. Aubry, B. Russel, J. Sivic. ACM Trans. on Graphics 2014, Vol. 33 Issue 2, March 2014.
- [2] '*Data-driven visual similarity for cross-domain image matching*', A. Shrivastava, T. Malisiewicz, A. Gupta, A.E. Efros. ACM Trans. on Graphics 2011, Vol. 30 Issue 6, December 2011.
- [3] '*Image Based Geo-Localization in the Alps*', O. Saurer, G. Baatz, K. Köser, L. Ladicky, M. Pollefeys. Int. Journal of Computer Vision 2015, Vol. 116, Issue 3, February 2015, pp 213-225.
- [4] '*Discriminative Unsupervised Feature Learning with Convolutional Neural Networks*', Alexey Dosovitskiy, Jost T. Springenberg, Martin Riedmiller and Thomas Brox, Advances in Neural Information Processing Systems, pages 766-774, 2014
- [5] '*Augmenting vehicle localization accuracy with cameras and 3D road infrastructure database*', L. Wei, B. Soheilian, V. Gouet-Brunet. ECCV workshop on Computer Vision in Vehicle Technology 2014, Zurich, Switzerland, September 6-12, 2014
- [6] '*Vehicle localization using mono-camera and geo-referenced traffic signs*', X. Qu, B. Soheilian, N. Paparoditis. IEEE Intelligent Vehicles Symposium (IV2015), Seoul, South Korea, 28 June - 1st July 2015
- [7] '*Robust and optimal SoS-based point-to-plane registration of image sets and structured scenes*' D.P. Paudel, A. Habed, C. Demonceaux, P. Vasseur. IEEE/CVF International Conference on Computer Vision (ICCV 2015), Santiago, Chile, December 2015.
- [8] '*LMI-based 2D-3D registration: from uncalibrated images to euclidean scene*' D.P. Paudel, A. Habed, C. Demonceaux, P. Vasseur. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), Boston, MA, USA, June 2015.