Open internship positions in the IMAGINE Group (spring-summer 2018)

The IMAGINE Group

IMAGINE is a research group at École des Ponts ParisTech (ENPC). It located at Champs-sur-Marne, 25 min from Paris center. It is part of the Gaspard-Monge Computer Science lab (LIGM). IMAGINE is also supported by the French Scientific and Technical Centre for Building (CSTB).

The research domains of IMAGINE are **computer vision, machine learning,** and **optimization.** In particular, the IMAGINE group is well known internationally for its work on calibration and multi-view stereovision for **3D reconstruction**. This knowledge has been transferred in 2011 to the successful startup company Acute3D, created by former members of the group. IMAGINE also won in 2011 the PRoVisG Mars 3D Challenge, aiming at testing and improving the state of the art in visual odometry and 3D terrain reconstruction in planetary exploration. However, the domains of interest of IMAGINE are wider than just 3D geometry. Another common interest in the group include semantic aspects, in particular related to **scene understanding** tasks, with a strong expertise in **deep learning**. Members of the group have been holding in 2015 the first position in the PASCAL VOC2012 comp4 object detection leaderboard, and second position in the MS COCO object detection leaderboard in a Facebook team. Since then, the group continues to obtain best results using deep learning techniques on a variety of benchmarks.

Open internship positions (spring-summer 2018)

Below is a list of (4-6 month) internships open in the IMAGINE Group in spring-summer 2018. Note that only a subset of these positions will be filled, depending on the profile of applicants.

N° INT18-01: [MSc 2] Large scale Art analysis

N° INT18-02: [MSc 2] 3D reconstruction from historical data

N° INT18-03: [MSc 2] Semantic & volumic 3D reconstruction of indoor scenes from range data

N° INT18-04: [MSc 2] Neural networks for semantic segmentation of temporal 3D data

N° INT18-05: [MSc 2] Plane detection with complete parameter estimation using Q-intersection

To apply to one of these internships, please send an email to the contact people mentioned in the internship description, with:

- the number(s) and title(s) of the internship(s) you are interested in,
- your résumé/CV,
- a transcript of your grades/marks for last year and this year (even if incomplete),
- a short statement of your interest and appropriate competence with respect to the proposed subject(s).

Incomplete applications will not be considered.

We also have a number of **PhD positions** to be filled in 2018, some of which are related to these internships. Making a good internship in the group offers very good chances to obtain a PhD fellowship afterwards.

Internship in the IMAGINE Group — spring-summer 2018 – http://imagine.enpc.fr/

N° INT18-01

Type of Internship

Research (Master Sc, last year)

Title

Large scale Art analysis

Context

Computer vision recently made huge progress by leveraging large-scale data for image analysis and applications of these idea to new area are still open problems. In particular, there are relatively few works applying these idea to artwork analysis.

Objectives

This internship could be oriented in several direction depending on the interests of the student:

- relationship of the novelty and appearance of paintings to market prices. This project could be in collaboration with Christophe Spaenjers at HEC (spaenjers@hec.fr) and would be based on a very large scale dataset of transactions history in art market. The two main challenges will be to define mathematically a good notion of novelty and to learn to establish relationships between appearance and price.

- inter-painting relationships discovery for Art history. Imagine being able to select a painting and automatically see in front of your eyes the painting it has been influenced by, in term of content, painting style or composition, and potentially all the paintings it has influenced.

- style representations. Several important challenges are related to style: learning style invariant representations, learning to evaluate style independently of the content, and transferring style between representations.

Profile

Programming experience in C++, Matlab or Python, basic knowledge of computer vision and machine learning. Interest in the subject. Some knowledge or experience with Deep Learning would be a plus.

References

[1] Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A Neural Algorithm of Artistic Style. *arXiv preprint arXiv:1508.06576*.

[2] Mensink, T., & van Gemert, J. (2014, April). The Rijksmuseum Challenge: Museum-Centered Visual Recognition. In *Proceedings of International Conference on Multimedia Retrieval* , 2014

[3] Saleh, B., Abe, K., Arora, R. S., & Elgammal, A. (2014). Toward automated discovery of artistic influence. *Multimedia Tools and Applications*, 2014

Contact

Mathieu Aubry (mathieu.aubry@imagine.enpc.fr)

N° INT18-02

Type of Internship

Research (Master Sc, last year)

Title

3D reconstruction from historical data

Context

We are now able to perform large scale and high accuracy 3D reconstruction [4]. However, this is only true if a very large quantity of high quality images is available or can be acquired. This is not the case for historical photographs or drawings which depict places that do not exist anymore, or that may have changed over time.

Objectives

This internship could be oriented in several directions depending on the interests of the student:

- multi-view 3D reconstruction from historical photographs. The first challenge will be to align depictions with very different styles and then to perform reconstruction using only a few available low quality photographs.

- single-view 3D reconstruction from historical or non realistic depiction. Many cue in an image allow to estimate depth from a single image. In particular one could try to extend the recent learning approaches to depth estimation [3] to historical photographs.

- geo-localization/grouping of historical depictions. The main challenge would be to increase the robustness of existing algorithm for geo-localization to be able to handle historical photograph and possibly non-realistic depictions.

Profile

Programming experience in C++, Matlab or Python, basic knowledge of computer vision and machine learning. Interest in the subject.

References

[1] Aubry, M., Russell, B. C., & Sivic, J. Visual geo-localization of non-photographic depictions via 2D-3D alignment. *Visual Analysis and Geolocalization of Large-Scale Imagery*, 2015

[2] Cléry, I. Valorisation géométrique et radiométrique d'un patrimoine de photographies anciennes scannées. IGN.

[3] Wang, X., Fouhey, D. F., & Gupta, A. (2014). Designing deep networks for surface normal estimation. *arXiv preprint arXiv:1411.4958*.

[4] Snavely, N., Seitz, S. M., & Szeliski, R. . Photo tourism: exploring photo collections in 3D. *ACM transactions on graphics (TOG),* 2006

Contact

Mathieu Aubry (mathieu.aubry@imagine.enpc.fr)

Pascal Monasse (monasse@imagine.enpc.fr)

N° INT18-03

Type of Internship

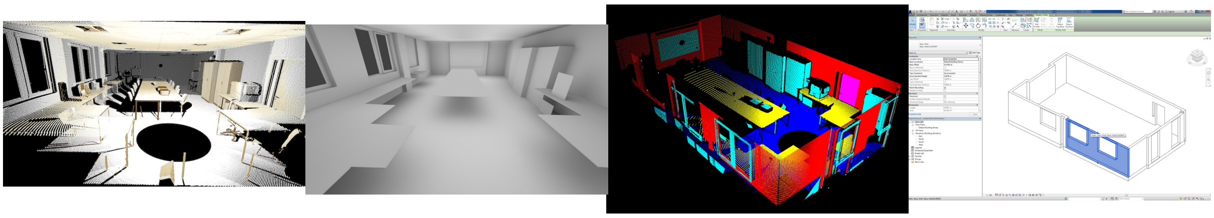
Research (Master Sc, last year)

Title

Semantic & volumic 3D reconstruction of indoor scenes from range data

Context

Due to technology leaps and cost reduction, capturing 3D range data (e.g., using laser scans, photogrammetry or Kinect-like sensors) has become available to a wide class of users, with numerous applications. One of our goals is to develop a method to reconstruct a geometric and semantic representation of scenes acquired with such range sensors, in particular indoor scenes.



Objectives

This ambitious goal cannot be reached with a single internship or even a PhD. We however propose a number of subtasks to work on, leading to different internships:

- adapting [1] from laser scans to low-cost range sensors (e.g., Kinect), which are more noisy, have a different geometry and possibly require the registration/fusion of several acquisitions,
- extending [1] to work on large scenes, using strategies to limit surface extension hypotheses so that reconstruction can be performed on smaller fractions of the data and later merged,
- decomposing the reconstructed volume into a small number of non-intersecting cuboids corresponding for instance to wall portions (extending 2D results such as [5] to 3D), possibly with holes (for door and window openings) and using gravity constraints as well as visual cues,
- inferring semantic label for these cuboids using domain-specific knowledge, possibly learning priors such as adjacency and size constraints from datasets of semantized 3D building models.

There are strong chances that these internship topics will be continued with a PhD thesis.

Profile

C++ proficiency, basic computational geometry skills, computer vision basics. Python is a plus.

References

[1] Alexandre Boulch, Martin de La Gorce, Renaud Marlet. Piecewise-planar 3D reconstruction with edge and corner regularization. Symposium on Geometry Processing/Comp. Graph. Forum, 2014.

[2] J. Guerry, A. Boulch, B. Le Saux, J. Moras, A. Plyer and D. Filliat. SnapNet-R: Consistent 3D Multi-View Semantic Labeling for Robotics. Workshop 3D Reconstruction Meets Semantics at ICCV 2017.

[3] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. NIPS 2017.

[4] T. van Lankveld, M. J. van Kreveld, R. C. Veltkamp. Identifying rectangles in laser range data forurban scene reconstruction. Computers & Graphics, vol 35, 2011.

[5] C. Nogueira de Meneses, C. Carvalho de Souza. Exact solutions of rectangular partitions via integer programming. Int. J. Comput. Geom. Appl., 2000.

[6] T. Shao, A. Monszpart, Y. Zheng, B. Koo, W. Xu, K. Zhou, N. J. Mitra. Imagining the unseen: stability-based cuboid arrangements for scene understanding. SIGRAPH Asia, 2014.

Contact

Renaud Marlet (renaud.marlet@enpc.fr), Alexandre Boulch (alexandre.boulch@onera.fr)

Internship in the IMAGINE Group — spring-summer 2018 – http://imagine.enpc.fr/

N° INT18-04

Type of Internship

Research (Master Sc, last year)

Title

Neural networks for semantic segmentation of temporal 3D data

Context

With the emergence of affordable sensors for the general public (e.g., Kinect, iPhone X) or for professionals (laser scanners) and with the development of robust photogrammetry techniques, point clouds have become the input data of many algorithms for surface reconstruction, semantic understanding, or 3D animation, replacing images for the perception of the environment in many situations: robotics, autonomous cars, urban mapping, biometrics... Besides, these sensors not only allow accurate spatial acquisition but also repeated captures of the same scene at different times.

Thanks to the development of machine learning techniques, in particular based on convolutive neural networks, many advances have been made in vision-based robotics, both in navigation and in the analysis of the 3D environment, online or offline [Boulch et al. 2017, Guerry et al. 2017]. New applications are also emerging, such as maintenance aids in difficult environments with augmented reality where 2D or 3D imagery, allowing the recognition of objects of interest, must be robust to the sensor movement (temporal robustness).

Objectives

We want to improve point cloud understanding for robotic uses. Several issues are raised:

1. How to make semantic segmentation of 3D data more efficient?

Understanding 3D point clouds, including object detection and semantic segmentation, is an essential step in navigation and decision-making for an autonomous agent. Current techniques are inherited from image segmentation [Su et al. 2015], or use voxelic approaches [Maturana & Scherer 2015], or work on unstructured data [Qi et al. 2017]. We want to improve this kind of techniques by leveraging on additional information available for data acquired by odometry or SLAM (e.g., in robotics).

2. How to use low-level geometric information to improve object segmentation?

Semantics is generally correlated with the scale of the object (wall, door, furniture, road, trees...). The search for geometric primitives (plans, cylinders...) can be a means of regularizing semantic segmentation and providing information on the consistency of a set of points as an object. For example, in indoor scenes, furniture and rooms consist of planes (e.g., tables, walls, etc.) and/or cylinders (pillars) and/or cuboids (closets). In addition, this geometric information can also be used to complete a point cloud when it contains noise or holes (missing data).

3. How to identify and characterize changes for data captures at different dates?

Point clouds are generally treated statically. However, with the emergence of low-cost sensors, it is now possible to have *temporal* point clouds, e.g., captured at different times on a construction site or by a robot operating in a dynamic environment. Variability in the data, either in the scene or due to the acquisition mode, raise new issues to identify and characterize the changes. Different strategies can be studied: directly on the data, or after a first step of semantization.

Profile

Proficiency in machine learning, computer vision, computer graphics. Programming skills.

Contact

Renaud Marlet (renaud.marlet@enpc.fr),

Alexandre Boulch (alexandre.boulch@onera.fr), Bertrand Le Saux (bertrand.le_saux@onera.fr)

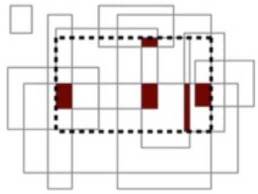
N° INT18-05

Type of Internship

Research (Master SC, last year)

Title

Plane detection with complete parameter estimation using Q-intersection



Context

Shape detection is a basic task in image processing and there exist randomized algorithms like RANSAC that are able to find, in a 3D point cloud, a plane (or another shape as a sphere or a cylinder) maximizing the number of points (inliers) being at a given distance of it. Some extensions of RANSAC have been built to find several shapes in a greedy manner. We propose to study an alternative innovative technique, based on a complete tree search in the parameter space, to find all the planes having as support a given number of points. Promising preliminary results have been obtained on artificial point clouds, but a validation and generalization of the method on real scenes is still to be done.

Objectives

A plane is described by a set of parameters. The complete search builds a search tree in the parameter space, and uses the Q-intersection contractor that contracts a current box in the search tree into a box compatible with at least Q points of the cloud, i.e., these Q points fit the equation of a plane within a given tolerance interval for each parameter.

The first part of the internship will consist of the study of the existing Q-intersection algorithms and the existing parameterizations. The main part of the internship will consist in adapting the method to real scenes, determining the limits of the approach (in percentage of inliers), studying the sensitivity to noise in the image, and precision issues: how to stop the splitting in order not to obtain too many small boxes representing near solutions, and how to regroup solutions containing the same inliers. One will also study how to tune the different parameters of the algorithm (the threshold Q, the tolerance in the equations, the stopping criterion). The implementation will be done in C++, using the Ibex interval solver library.

Profile

Knowledge on algorithms, C++ programming, and combinatorial optimization is required. Some knowledge on image and 3D processing is desirable but not mandatory.

References

[1] B. Neveu, M. de La Gorce, G. Trombettoni. Improving a Constraint Programming Approach for Parameter Estimation. ICTAI 2015, Nov 9-11, Vietri sul mare, Italy.

[2] G. Trombettoni, I. Araya, B. Neveu, G. Chabert. Inner Regions and Interval Linearizations for Global Optimization. Proc. of AAAI 2011, pages 99-104, San Francisco, CA, USA.

[3] C. Carbonnel, G. Trombettoni, P. Vismara, and G. Chabert. Q-intersection algorithms for constrained-based robust parameter estimation. Proc. of AAAI 2014.

[4] A. Bethencourt and L. Jaulin. 3D Reconstruction Using Interval Methods on the Kinect Device Coupled with an IMU. Int. Journal of Advanced Robotic Systems, 10,2013.

[5] L. Jaulin and S. Bazeille. Image Shape Extraction using Interval Methods. Proc.of the 16th IFAC Symposium on System, pages 378–383, 2009.

[6] R. Schnabel, R. Wahl, and R. Klein. Efficient RANSAC for point-cloud shape detection. Computer Graphics Forum, 26(2) :214–226, 2007.

[7] G. Chabert, L. Jaulin. Contractor Programming. Artificial Intelligence,173:1079–1100, 2009.

Contact

Bertrand Neveu (bertrand.neveu@enpc.fr)