



M2NUM is co-financed by the European Union with the European regional development fund (ERDF, HN0002137) and by the Normandie Regional Council

Stage M2 ou PFE : Analyse 3D de scènes routière par réseaux de neurones profonds

Durée du stage : 6 mois à partir du mois de février ou mars

Contact : Envoyer un CV et une lettre de motivation à Samia Ainouz, LITIS (samia.ainouz@insa-rouen.fr) et à Anastasia Zakharova, LMI (anastasia.zakharova@insa-rouen.fr).

Compétences requises : des compétences en traitement d'images et en apprentissage statistiques seront nécessaires. Des compétences en mathématiques et techniques d'optimisation seront appréciées.

Langages maîtrisés : Il serait souhaitable de maîtriser un des langages suivant : Python, C/C++, Open CV.

Ce stage est dans le cadre du projet M2NUM*. Il sera effectué à l'INSA de Rouen au sein du LITIS (Laboratoire Informatique, Traitement de l'Information et des Systèmes) et du LMI (Laboratoire des Mathématiques de l'INSA de Rouen Normandie)

Dans le cadre du projet M2NUM (<http://lmi2.insa-rouen.fr/~m2num/>) et dans le contexte des systèmes d'aide à la conduite, il est important pour un conducteur d'appréhender l'environnement qui l'entoure. Cela l'aidera sans doute à éviter les obstacles qu'il peut rencontrer et d'avancer en toute sécurité. La reconstruction 3D de l'environnement est un moyen incontournable pour connaître une scène routière et estimer la distance des dangers de la route par rapport au conducteur [5]. La nature des scènes routières fait que les images stéréo acquises, bien qu'elles correspondent à la même scène, comportent des différences de luminance qui peuvent être parfois importantes. Ces différences proviennent principalement des points de vues des caméras utilisées pour avoir les scènes stéréos à mettre en correspondance, et peuvent nuire à toute interprétation et analyse efficace de la scène [4]. De même, les bruits d'acquisition et les réflexions parasites rendent les algorithmes de détection ou de stéréovision obsolètes.



UNION EUROPEENNE

D'un autre coté, la recherche actuelle a connu un engouement sans précédent pour les méthodes basées sur les perceptrons, à savoir les réseaux de neurones profonds. Ces méthodes ont rendu à l'heure actuelle les meilleures performances dans toutes les applications de vision par ordinateur où elles ont été appliquées (analyse de scènes [1,2], détection d'obstacles [1,3], stéréovision [6]). Dans le cadre de ce stage, nous voulons étudier le meilleur module de détection 3D qui combinera en même temps la détection d'obstacles et la reconstruction 3D de l'environnement routier. Nous souhaitons pour ce faire, utiliser des architectures tirées des apprentissages multi-tâche (« multi-task learning »). Le travail demandé lors de ce stage est d'appliquer des architectures existantes dans la littérature, en adaptant certains de leurs paramètres pour l'appliquer à l'analyse 3D des scènes routières. Les images à l'entrée du réseau peuvent être de modalités différentes (visible, infrarouge, polarimétriques) afin de maximiser le nombre d'informations utiles caractérisant les scènes.

[1] R. Benenson, M. Omran, J. Hosang, and B. Schiele. “Ten years of pedestrian detection, what have we learned?” In Proceedings of ECCV Workshop, 2014.

[2] Juan C. Caicedo et Svetlana Lazebnik “Active object localization with deep reinforcement learning“. In : *Proceedings of the IEEE International Conference on Computer Vision*. 2015. p. 2488-2496.

[3] F. Chabot, M. Chaouch, J. Rabarisoa, C. Teulière, T. Chateau. *Deep MANTA : A Coarse-to-fine Many-Task Network for joint 2D and 3D vehicle analysis from monocular image*. CVPR, 2017.

[4] Alina Miron, S. Ainouz, A. Rogozan, A. Benschraï. Towards a robust and fast color stereo matching for intelligent vehicle application, IEEE International Conference on Image Processing 2012

[5] Sylvie Chambon. Mise en correspondance stéréoscopique d'images couleur en présence d'occultations. Thèse de doctorat, Université Paul Sabatier, Toulouse, décembre 2005.

[6] J. Zbontar and Y. LeCun. “Computing the stereo matching cost with a convolutional neural network”. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1592-1599, 2015

