

Proposition de stage de Master

Tracking et intégration temporelle de primitives de haut niveau issues de réseaux convolutionnels pour la localisation par l'image

Equipe: TANGRAM(ex MAGRIT): <http://magrit.loria.fr>, <http://team.inria.fr/tangram>

Encadrants : Marie-Odile BERGER, Gilles SIMON\\

Lieu : INRIA/LORIA, 615 rue du jardin botanique, 54600 Villers-les-Nancy

Contact : marie-odile.berger@loria.fr gilles.simon@loria.fr

Homepages : <https://members.loria.fr/gsimon> <https://members.loria.fr/moberger/>

Motivations :

Les méthodes basées sur les réseaux convolutionnels (CNN) permettent désormais la détection de primitives de haut niveau de manière largement indépendante des conditions d'illumination ou du point de vue adopté pour la prise de vue. La détection d'objets [yolo] a largement bénéficié de ces méthodes tandis que des méthodes prometteuses émergent pour la détection de lignes d'horizon ou de points de fuite [workman] et commencent à concurrencer les méthodes traditionnelles [simon]. Alors que le calcul de la pose de la caméra reposait il y a encore peu de temps uniquement sur les points d'intérêt détectés dans l'image, il est maintenant possible d'utiliser les objets détectés dans l'image pour la localisation [gaudilliere] ce qui est particulièrement intéressant dans les environnements peu texturés. Par ailleurs, la détection de la ligne d'horizon et des points de fuite permet d'apporter des informations sur la composante rotationnelle de la pose.

Dans beaucoup d'applications, notamment de réalité augmentée, l'utilisateur est libre de se déplacer dans la scène et on peut ainsi avoir à traiter des vues globales de la scène ou au contraire des vues rapprochées de certains objets. Le nombre d'indices disponibles peut alors être petit et il importe que la localisation puisse tirer partie au mieux de ces indices, même s'ils sont relativement imprécis. Par ailleurs la localisation doit être faite de manière continue temporellement, ce qui implique de pouvoir mettre en correspondance ces indices au cours du temps.

Sujet

Le sujet consistera à poser les bases d'un système permettant l'intégration temporelle de ces indices de haut niveau avec les classiques points d'intérêt au sein d'un système de localisation par l'image. Les indices considérés dans ce travail seront les objets détectés, les lignes d'horizon et des points de fuite (PF). Les points suivants seront abordés:

- mise en correspondance temporelle des objets: Le problème est plus complexe que dans le cas de points car il faut comparer des structures dont la forme n'est pas précisément définie puisque la détection fournit seulement une boîte englobante. La similarité des descripteurs CNN [sunderhauf] pourra être utilisée entre deux images ainsi que la stabilité temporelle de la trajectoire des objets.
- mise en correspondance des points de fuite: les détecteurs de PF génèrent souvent de nombreux points qu'il est difficile d'apparier sur la seule base de leur coordonnées. Des stratégies visant à rectifier les images à partir des PF afin de valider leur ressemblance, et donc la pertinence de la mise en correspondance des PF, seront étudiées
- Des stratégies d'intégration de ces indices dans la procédure de localisation, avec les traditionnels points d'intérêt, devront ensuite être proposées, l'objectif étant d'intégrer de manière préférentielle les indices de haut niveau apportant une information complémentaire de celle des points présents. Des techniques d'*explainability* [gradcam] pourront être utilisées dans ce but.

Cadre du travail

Le ou la stagiaire sera intégré dans l'équipe TANGRAM (anciennement MAGRIT) et bénéficiera des compétences de l'équipe sur la localisation en général, et sur la localisation à partir d'objets. Il sera aussi intégré au projet MOVEON, en collaboration avec le DFKI (Kaiserslautern), qui vise à développer de nouvelles méthodes de localisation visuelle.

Le sujet proposé est à la confluence de l'apprentissage et de la géométrie. Nous recherchons un ou une candidate avec des connaissances (ou un fort intérêt) dans les domaines suivants: traitement d'image, vision par ordinateur et apprentissage.

Les candidatures doivent être envoyées à et marie-odile.berger@loria.fr et gilles.simon@loria.fr

Bibliographie

[yolo] J. Redmon, S. Divvala, R. Girshick, A. Farhad. *You Look Only Once: Unified, real-time object detection*. CVPR, 2016.

[workman] S. Workman, M. Zhai, N. Jacobs. *Horizon Lines in the Wild*. BMVC, 2016.

[simon] G. Simon, A. Fond, M.-O. Berger. *A Contrario Horizon-first Vanishing Point Detection Using Second Order Grouping Laws*. ECCV, 2018.

[gaudilliere] V. Gaudillière, G. Simon, M.-O. Berger. *Perspective-2-Ellipsoid: Bridging the GAP Between Object Detection and 6-DoF Camera Pose*. IEEE Robotics and Automation Letter, 2020.

[sunderhauf] N. Sunderhauf et al. *Place Recognition with ConvNet Landmarks: Viewpoint-Robust, Condition-Robust, Training Free*. Robotics Sciences and Systems Conference, 2015.

[gradcam] Ramprasaath et al. Grad-CAM: CVPR, 2017.