

Détection d'alignements dans les cartes de caractéristiques d'un CNN : application au calcul des points de fuite d'une image

1 Informations générales

Encadrants Marie-Odile BERGER, Gilles SIMON
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone 03 83 59 20 67
Email marie-odile.berger@inria.fr gilles.simon@loria.fr
Bureau C 034

2 Motivations

La détection automatique des points de fuite (PF) d'une image est un pré-requis à la résolution de nombreux problèmes de vision par ordinateur tels que le calcul de l'orientation d'une caméra, l'orthorectification d'une image, la reconstruction 3D d'une scène à partir d'une photographie, etc. Pour cette raison, ce problème a fait l'objet de nombreux travaux au cours des vingt dernières années. Les premières méthodes [1], essentiellement basées sur l'algorithme d'espérance-maximisation, ont été confrontées à la difficulté de fixer des valeurs de seuil généralisables, produisant fréquemment des PF dupliqués, manqués ou erronés. Un saut qualitatif a été réalisé récemment grâce à des travaux de l'équipe Magrit [2] et de l'Université du Kentucky [3]. Ces travaux ont montré qu'il est possible de détecter la ligne d'horizon indépendamment des PF, et ainsi de limiter la zone de détection des PF horizontaux à cette ligne, évitant de nombreuses fausses détections. Dans [2], nous avons montré que la ligne d'horizon peut être détectée en tant qu'alignement significatif des contours de l'image orthogonaux à la ligne zénithale. Dans [3], un réseau de neurones convolutif (*Convolutional Neural Network* – CNN) dit “end-to-end” permet de régresser directement les paramètres de la ligne d'horizon, après avoir été entraîné sur un grand nombre d'images urbaines.

Malgré ces avancées, la détection automatique de PF peut générer des résultats aberrants dans certaines situations : absence ou rencontres fortuites de segments avec notre méthode, scènes ou points de vue d'observation mal représentés dans les données d'apprentissage avec la méthode de Zhai et al. Ces problèmes devraient pouvoir être surmontés en restant dans un cadre géométrique et stochastique, mais en considérant des alignements de primitives de niveau sémantique supérieur à celui des contours d'une image. Une solution basique consisterait à utiliser un CNN pour détecter des objets communément alignés dans une image (par exemple, pour des scènes urbaines, les fenêtres d'un même étage ou des voitures garées en file indienne), puis de détecter les PF à partir des tels alignements. Cette méthode ne permettrait cependant pas de tirer profits de structures non reconnues par un CNN standard, telles que des motifs décoratifs propres à un édifice particulier.

3 Sujet

Plusieurs auteurs ont remarqué que les points de fuite jouent un rôle clé dans la résolution de certains problèmes de vision par ordinateur à l'aide d'un CNN end-to-end, tels que le calcul de pose ou l'estimation d'une carte de profondeur. Une étude préalable (menée par nos soins) a par ailleurs montré que des alignements de primitives de niveau sémantique élevé émergent naturellement dans certaines couches d'un CNN entraîné à régresser des points de fuite. Le sujet visera dans un premier temps à conforter cette observation, en explorant les couches profondes de différents CNN afin de déterminer quelles couches et quels types de CNN sont les plus à même de générer ce type de primitives. Dans un second temps, nous tenterons de détecter les alignements proprement dits dans les cartes de caractéristiques des couches identifiées, à l'aide de techniques d'analyse d'image classiques telles que la transformée de Hough ou la méthode stochastique utilisée dans [2].

4 Cadre du travail

Des réseaux de neurones convolutifs pré-entraînés seront analysés. L'environnement Matlab a notre préférence mais n'est pas obligatoire si le stagiaire se sent plus à l'aise avec d'autres langages. Une réunion hebdomadaire est prévue avec les encadrants et des discussions plus informelles seront bien-sûr bienvenues tant avec les permanents qu'avec les doctorants de l'équipe Magrit.

5 Compétences requises

Connaissances élémentaires en géométrie, vision et apprentissage.

Références

- [1] Kosecká J., Zhang W. Video Compass. *European Conference on Computer Vision*, 2002.
- [2] G. Simon, A. Fond, M.-O. Berger. A-Contrario Horizon-First Vanishing Point Detection Using Second-Order Grouping Laws. *European Conference on Computer Vision*, 2018.
- [3] M. Zhai, S. Workman, N. Jacobs. Detecting vanishing points using global image context in a non-manhattan world. *Computer Vision and Pattern Recognition*, 2016.