

Proposition de stage M2

Sujet :

Apprentissage profond pour l'asservissement visuel

Laboratoire : IRISA, INRIA Rennes-Bretagne Atlantique

Département : DS5, Signaux et images numériques, robotique

Équipe de recherche : Lagadic (<http://www.irisa.fr/lagadic>)

Responsable : Eric Marchand (marchand@irisa.fr)

Mots-clés : vision robotique, asservissement visuel, deep learning

Les techniques d'asservissement visuel [1] consistent à contrôler les mouvements d'un système robotique à l'aide d'informations visuelles acquises par un système de vision et intégrées au sein de lois de commande en boucle fermée. Classiquement ces informations visuelles sont de type géométrique (point, droite, cercle, etc) et des algorithmes d'extraction et de suivi (résultant d'une phase de traitement d'image) doivent être mis en œuvre. Nous souhaiterions pouvoir supprimer cette étape à l'occasion de cette thèse.

En effet, une nouvelle approche est en train de voir le jour. Elle consiste à travailler sur l'ensemble de l'image en utilisant l'information photométrique de l'ensemble des pixels [2]. Cette approche est séduisante puisqu'elle permet d'obtenir une excellente précision de positionnement et permet de se passer de l'étape de suivi et de mise en correspondance (en ce sens on parlera **d'asservissement visuel direct**). Elle souffre par contre d'un domaine de convergence plus local et d'une faible robustesse aux modifications de l'environnement.

Récemment, il a été démontré qu'il était possible d'utiliser des réseaux de neurone convolutionnel (CNN – Convolutional Neural Network) pour calculer la position d'une caméra par rapport à une scène [3]. Cette position peut ensuite être utilisée pour développer une loi de commande permettant de positionner une caméra montée sur l'effecteur d'un robot [4]. Cette méthode fonctionne correctement si le réseau est entraîné avec des images (simulées) de la scène considérée. Pour aller plus loin, plusieurs questions restent à étudier :

- Peut-on entraîner le réseau avec des images de scène quelconques ? (capacité de généralisation des CNN)
- Peut-on contraindre la trajectoire du robot ?
- Peut-on définir une commande qui évite l'étape de localisation ?
- Quelle information doit-on donner en entrée du réseau (images, différence entre deux images, image de gradients, etc. ?)
- Comment améliorer la robustesse aux modifications de l'environnement ?

Finalement, des études précédentes ont montré la dualité entre l'asservissement visuel et la localisation d'une caméra (pouvant le plus souvent se restreindre à un problème de suivi d'objet dans des séquences d'image). Se basant sur les mêmes bases méthodologiques, les méthodes développées dans le contexte de l'asservissement visuel pourront ainsi être appliquées au problème de la localisation et du suivi 3D. Une approche unifiée entre la détection/reconnaissance d'objet et le suivi/localisation pourra donc être étudiée.

Les développements méthodologiques réalisés seront intégrés dans la bibliothèque logicielle ViSP (C++) et seront validés en simulation puis via des expérimentations réelles sur un système de vision robotique (en particulier sur un robot mobile et sur un robot à 6 degrés de liberté mais aussi éventuellement sur des drones de type quadrotor).

Compétences souhaitées :

- connaissances en vision par ordinateur et traitement d'images
- goût pour les mathématiques, optimisation
- Machine learning
- programmation sous C++
- goût pour l'expérimentation

Bibliographie :

- [1] F. Chaumette, S. Hutchinson. **Visual servo control, Part I: Basic approaches**. IEEE Robotics and Automation Magazine, 13(4):82-90, Décembre 2006.
- [2] C. Collewet, E. Marchand. **Photometric visual servoing**. IEEE Trans. on Robotics, 27(4):828-834, Août 2011.
- [3] Kendall, A., Grimes, M., & Cipolla, R. (2015). **Posenet: A convolutional network for real-time 6-dof camera relocalization**. In *Proceedings of the IEEE international conference on computer vision* (pp. 2938-2946)
- [4] Q. Bateux, E. Marchand, J. Leitner, F. Chaumette, P. Corke. **Visual Servoing from Deep Neural Networks**. In *RSS workshop on New Frontiers for Deep Learning in Robotics*, Boston, Ma, Juillet 2017.