

## Reconnaitances des objets dans la vidéo égocentrique: aide au contrôle de neuro-prothèses

Encadrants : Pr. Jenny Benois-Pineau, Pr. Pascal Desbarats, équipe Image et Son  
En collaboration avec l'INCIA, Dr. Aymar de Rugy.

Due au développement rapide des dispositifs vidéo de faibles taille et poids, l'analyse de la vidéo égocentrique acquise par les caméras portées par des personnes trouve les applications dans plusieurs secteurs et notamment la santé[1]. Les problèmes classiques d'analyse/reconnaissance/vision se posent pour ces contenus complexes[2]. Dans le cadre du stage master recherche nous nous intéressons à la reconnaissance des objets d'intérêt à saisir par des sujets amputés portants des neuro-prothèses des membres supérieurs.

Le dispositif du contrôle visuel consiste en outils de suivi du regard sur des lunettes (Tobii). Il enregistre la vue égocentrique du sujet. Dans la période d'enregistrement qui précède l'action de saisie, ces données traduisent l'attention visuelle et l'intention du sujet. Les cartes d'attention visuelle sont calculées à partir de ces données et indiquent les zones d'intérêt comportant les objets d'intérêt. L'outil de reconnaissance des objets d'intérêt est la composante centrale du système de contrôle hybride Vision – EMG qui pilotera la prothèse. Un premier système de reconnaissance des objets dans la vue égocentrique a été élaborée au sein du LABRI dans le cadre du projet PEPS CNRS-Idex Suvipp avec le laboratoire de bio-physique l'INCIA[3]. Une architecture des réseaux profonds comme dans [4] est à la base de cet outil.

Dans le cadre du stage master il s'agit d'étendre la reconnaissance des objets aux corpus naturels écologiques, enregistrés avec des lunettes TOBII et de l'inscrire dans le cadre de l'apprentissage incrémental : en effet les catégories des objets sont répétables dans des scénarios égocentriques écologiques, mais leur forme et apparence changent. Un outil pre-entraîné doit s'adapter à l'évolution de l'environnement.

[1]S.KARAMAN, J.BENOIS-PINEAU, V. DOVGALECS, R.MÉGRET, J.PINQUIER, R.ANDRÉ-OBRECHT, Y.GAËSTEL AND J.-F. DARTIGUES, "Hierarchical Hidden Markov Model in Detecting Activities of Daily Living in Wearable Videos for Studies of Dementia", Multimedia Tools and Applications, 69(3): 743-771 (2014)

[2] V. BUSO. I. GONZALEZ-DIAZ. J. BENOIS-PINEAU, « Goal-oriented top-down probabilistic visual attention model for recognition of manipulated objects in egocentric videos », Signal Processing : Image Communication, 2015, doi:10.1016/j.image.2015.05.006

[3] [Philippe Pérez de San Roman](#), [Jenny Benois-Pineau](#), [Jean-Philippe Domenger](#), [Florent Paquet](#), [Daniel Cataert](#), [Aymar de Rugy](#), « Saliency Driven Object recognition in egocentric videos with deep CNN », [arXiv:1606.07256](#), submitted to CVIU

[4] R. B. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2016) 142-158. doi:10.1109/TPAMI.2015.2437384. URhttp://dx.doi.org/10.1109/TPAMI.2015.2437384

[5]. <https://www.labri.fr/projet/AIV/dossierSiteSuVIPPP/presentationFr.php>

Environnement de travail : Tobii, C++, OpenCV Linux, Caffe

## Object recognition in egocentric video : assistance to neuro-prostheses wearers

Tuteurs : Pr. Jenny Benois-Pineau, Pr. Pascal Desbarats, Image and Son research department

In collaboration with INCIA, Dr. Aymar de Rugy.

Due to the rapid development of wearable video acquisition devices which are light and low-weight, egocentric video analysis becomes popular in various application domains including Health[1]. The classical problem of analysis pattern recognition, computer vision arise on these complex content. [2]. In the master internship project we are interested in recognition of objects to grasp with the goal of assistance of amputees wearing neuro-prostheses of upper limbs.

The visual control set-up consists of an eye-tracker and a scene camera mounted on the glasses ( Tobii). It records the egocentric view of the subject. During the time frame preceding the grasping action, these data express visual attention and intention of the subject. Visual attention maps are computed with these data and delimit the regions of interest (ROI) comprising the objects-of-interest (OI). OI recognition toll is the central component of a hybrid Vision-EMG control system which is being designed for the prostheses control.

The first object recognition framework has been developed in LABRI during the project PEPS CNRS-Idex Suvipp time-frame [5] together with bio-physical Lab INCIA.

A Deep CNN architecture as in [4] is the basis of this framework.

During the master internship project we expect to further develop the object recognition approach focusing on the real-life objects in the wild, that is in an ecological situation for subjects. But also the new methodology of incremental learning has to be implemented, indeed the categories of objects in the everyday life scenarios are repeatable, but their shapes and appearances change. A pre-trained recognition tool needs to be adaptable to evolving environment.

Working environment: Tobii, C++, OpenCV Linux, Caffe

[1]S.KARAMAN, J.BENOIS-PINEAU, V. DOVGALECS, R.MÉGRET, J.PINQUIER, R.ANDRÉ-OBRECHT, Y.GAËSTEL AND J.-F. DARTIGUES,"Hierarchical Hidden Markov Model in Detecting Activities of Daily Living in Wearable Videos for Studies of Dementia", Multimedia Tools and Applications, 69(3): 743-771 (2014)

[2] V. BUSO. I. GONZALEZ-DIAZ. J. BENOIS-PINEAU, « Goal-oriented top-down probabilistic visual attention model for recognition of manipulated objects in egocentric videos », Signal Processing : Image Communication, 2015, doi:10.1016/j.image.2015.05.006

[3] [Philippe Pérez de San Roman](#), [Jenny Benois-Pineau](#), [Jean-Philippe Domenger](#), [Florent Paclet](#), [Daniel Cataert](#), [Aymar de Rugy](#), « Saliency Driven Object recognition in egocentric videos with deep CNN », [arXiv:1606.07256](#), submitted to CVIU

[4] R. B. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2016) 142{158. doi:10.1109/TPAMI.2015.2437384. UR<http://dx.doi.org/10.1109/TPAMI.2015.2437384>

[5]. <https://www.labri.fr/projet/AIV/dossierSiteSuVIPP/presentationFr.php>