



Sujet :

Codage parcimonieux de descripteurs locaux pour la reconnaissance d'objets dans des images

Niveau : Bac + 5 - 3^{ème} année d'école d'ingénieur – Master

Durée du stage : 4 à 6 mois

Stage pouvant se poursuivre en thèse : Oui

Contexte

Le laboratoire d'ingénierie de la connaissance multimédia et multilingue développe des outils et méthodes pour le traitement automatisé et robuste de flux numériques multimédia. Il est composé de chercheurs et d'ingénieurs de recherche dont les spécialités vont du traitement d'image au traitement du langage naturel, tous réunis autour des thématiques de l'analyse sémantique des contenus multimédia. Les applications de ces recherches concernent la recherche d'information (moteurs de recherche, sites d'entreprises), la veille (technologique, média, sources ouvertes), la gestion du patrimoine, le commerce électronique (tourisme, VoD, biens culturels) ou encore le développement de ressources linguistiques et sémantiques.

Le thème de ce stage concerne la détection puis la reconnaissance d'objets dans des images. Il s'agit d'une tâche fondamentale qui est l'objet de recherches très actives au niveau mondial [1,2]. Les techniques les plus récentes associent des descripteurs locaux agrégés selon des dictionnaires de "mots visuels" à différentes échelles associés à des algorithmes d'apprentissage (SVM, boosting...). Certains travaux proposent d'utiliser des dictionnaires de grande taille [4] impliquant ainsi un possible codage parcimonieux des images, c'est-à-dire tel que les unités actives soient une suite discrète de valeurs dont seules un petit nombre ont de grandes amplitudes, la plupart ayant une amplitude faible voire nulle.

Objectifs du stage

Les objectifs de ce stage seront les suivants :

- Effectuer un état de l'art des techniques de reconnaissance d'objets et de scènes
- Appropriation des outils existants au laboratoire
- Recherche et développement d'un nouvel algorithme de reconnaissance d'objet exploitant le codage parcimonieux de descripteurs locaux agrégés.
- Eventuelle validation dans le cadre d'une campagne d'évaluation internationale [1,5]

La participation à une campagne d'évaluation internationale [1,5] permet de comparer le système développé à l'état de l'art, en s'attaquant à des problèmes réels et en utilisant des corpus réalistes. Les algorithmes utilisés seront codés en C++ sous Linux.

Bibliographie :

[1] <http://pascallin.ecs.soton.ac.uk/challenges/VOC/>

[2] Sivic, Russell, Zisserman, Freeman, and Efros, "Unsupervised Discovery of Visual Object Class Hierarchies", CVPR 2008.

[3] Lazebnik, Schmid, Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, CVPR 2006.

[4] Campagne ImageCLEF : <http://ir.shef.ac.uk/imageclef/>

Domaines de spécialité requis :

Traitements de l'information, traitement des images, architecture logicielle, C++

Moyens informatiques mis en œuvre :

C++, et éventuellement des langages de script (perl...). Travail sous linux



COMMISSARIAT A L'ENERGIE ATOMIQUE
DRT/DTSI/SRCI
Relation avec les écoles : MARC MERGY
Tél : 01 46 54 81 58
E-Mail : marc.mergy@cea.fr

Autres moyens mis en œuvre (expériences, méthodes d'analyses, autres...)

Centre : Fontenay-aux-Roses

Pôle ou Direction : DRT - LIST

Dépt/Service/Labo : DTSI/SRCI/LIC2M :

Laboratoire d'Ingénierie de la Connaissance Multimédia Multilingue

Web : <http://www-list.cea.fr/>

Encadrement

Nom de l'ingénieur responsable : Hervé LE BORGNE Tél. : +33 1 46 54 85 31

E-mail : herve.le-borgne@cea.fr Fax : +33 1 46 54 75 80

Secrétariat : DTSI/SRCI Tél. : +33 1 46 54 91 17

Nom du chef de laboratoire : Olivier Mesnard Tél. : +33 1 46 54 73 38