

Algorithme de classification hiérarchique multi-critère

Stage de formation de niveau Master 2

Développement, implémentation et expérimentation d'un algorithme de classification hiérarchique multi-critère pour le traitement d'images spectrales combinant similarités spectrales et proximité spatiale.

I Encadrement

Responsables du stage

- Agnès Grimaud, LMV UMR8100 (CNRS, UVSQ) agnes.grimaud@uvsq.fr
- Serge Cohen, IPANEMA USR3461 (CNRS, MIC, UVSQ), serge.cohen@synchrotron-soleil.fr
- Gilles Celeux, équipe SELECT de l'INRIA (INRIA, UPSud), gilles.celeux@inria.fr

II Descriptif du stage

Intitulé du stage

Classification (*Clustering*) hiérarchique hybride pour grand jeux de données issus d'images spectrales

Description du stage

La caractérisation des matériaux anciens appelle souvent à l'exploitation de données d'imagerie spectrale, c'est-à-dire d'images pour lesquelles chaque pixel est caractérisé par un spectre complet. En effet la richesse d'information fournie par l'acquisition et l'exploitation d'un spectre complet est souvent déterminante pour caractériser les détails des matériaux hétérogènes et d'une grande complexité que sont les matériaux rencontrés en archéologie, paléontologie ou conservation du patrimoine culturel. De plus les jeux de données obtenus par imagerie font souvent plusieurs GO, voire dizaine de GO et il importe d'avoir des algorithmes d'analyse performant en termes de coût de calcul.

L'objectif du stage de master est de mettre au point et d'implémenter un algorithme de classification hiérarchique multi-critère se basant à la fois sur les similarité spectrales et la proximité spatiale des pixels et des régions candidates à l'agrégation. Pour pouvoir être efficace l'algorithme ainsi développé devra utiliser une approche permettant de ne calculer les similarités spectrales que des pixels satisfaisant les contraintes spatiales induites par les critères de classification et mettre en place des règles de mise à jour bénéficiant des calculs déjà effectués avant chaque étape d'agrégation de classe. Il s'agit donc d'un algorithme de classification hiérarchique pour lequel la matrice de similarités s'avère creuse.

Le stagiaire devra également effectuer l'implémentation de l'algorithme développé et l'expérimenter sur des données réelles. Le codage sera effectué en utilisant l'environnement R (<http://www.r-project.org/>), et intégrera éventuellement des parties en langage de plus bas niveau si ceux-ci permettent d'accélérer significativement les calculs : C ou OpenCL (<https://www.khronos.org/opencv/>, à travers le paquet éponyme de R).

Le cas échéant, la recherche menée durant ce stage sera à la base d'un sujet de thèse.

Dates, localisation, rémunération du stage et contact

Le stage débutera début avril et il est souhaitable qu'il dure 6 mois. Sa rémunération est de 550 euros par mois.

Les personnes à contacter pour candidater à ce stage sont Agnès Grimaud (agnes.grimaud@uvsq.fr) et Serge Cohen (serge.cohen@synchrotron-soleil.fr).

Le stagiaire sera localisé au laboratoire IPANEMA et au LMO par alternance.