

Etude de la réduction de la dimension et extraction des tendances évolutives pour les données multi-temporelles dans un contexte de High Performance Computing

Nom, prénom du proposant CNES : Julien Michel, Manuel Grizonnet, Jordi Inglada

Sigle du proposant : DCT/SI/AP, CB Email du proposant : julien.michel@cnes.fr

Laboratoire d'accueil envisagé : CESBIO

Directeur de Recherche envisagé : Jordi Inglada

Profil du candidat : Master en traitement du signal et des images

Descriptif du sujet :

Contexte :

Avec l'arrivée des missions à forte revisite temporelle, un nouveau cap est franchi en terme de volume de données à exploiter. De nombreuses initiatives de centres de traitements (Kalideos, Pôle Surfaces Continentales, segment sol S2 ...) existent pour pré-traiter ces données. Les laboratoires de traitement d'image et de télédétection recevront de longues séries d'images parfaitement superposées et calibrées, dont l'exploitation va demander une évolution des algorithmes d'analyse d'image. Du point de vue des algorithmes et méthodes, l'analyse des séries multi-temporelles d'images superspectrales connaît de nombreuses avancées ces dernières années. Parmi les techniques d'intérêt, on peut citer l'interpolation temporelle, la réduction de dimension dans la direction temporelle (parfois conjointement aux dimensions spatiale et spectrale), l'extraction de phénomènes évolutifs temporels (évolution par opposition à détection de changement), et la synthèse de produits exempts de certaines perturbations (nuages, ombres, traînées issues de l'aviation ...).

Concernant la mise en oeuvre efficace des méthodes, de nombreux travaux sont menés actuellement sur le thème du High Performance Computing, et notamment concernant la mise en oeuvre d'algorithmes sur les GPUs (Graphics Processing Unit). Les composants pour les segments sols du CNES (outils communs) ont par exemple été portés sur GPU et les gains de temps sont substantiels pour certaines méthodes. Par ailleurs une bibliographie très fournie existe, ainsi que de nombreuses librairies de traitement d'image dédiées aux GPU. On peut notamment citer les travaux de l'université d'Extremadure (Espagne) concernant le traitement d'images hyperspectrales.

Objectifs généraux :

L'objectif de cette thèse est d'adapter les méthodes d'analyse de données multi-variées à l'analyse de séries multi-temporelles longues dans un contexte de High Performance Computing. Il s'agit de lever les verrous scientifiques liés au portage des méthodes classiques sur une architecture massivement parallèle, afin de répondre à la problématique du traitement efficace du volume de données, mais également de concevoir de nouvelles méthodes et algorithmes rendus possibles par la puissance de calcul disponible. De plus, ces méthodes devront permettre d'exploiter la complémentarité avec des données de type THR, hyperspectrales ou radar ponctuelles. Le cadre applicatif sera fourni par le CESBIO (à préciser). Un co-encadrement avec le laboratoire HPC de l'université d'Extremadure permettra de bénéficier de leur expérience dans le domaine.

La première partie de la thèse sera consacrée à l'étude de la réduction de la dimension et à l'extraction des tendances évolutives pour les données multi-temporelles. Toutes ces techniques sont d'un grand intérêt pour l'analyse des séries multi-temporelles dans la mesure où elles réduisent le nombre de variables à observer tout en concentrant le contenu informationnel. Dans le cadre de cette thèse, il faudra ajouter la dimension temporelle : il faut formaliser et généraliser ces méthodes pour prendre en compte de cette dimension et proposer de nouveaux algorithmes permettant l'analyse jointe de cet hyper-cube spatio-spectro-temporel. La mise au point de ces nouvelles méthodes devrait favoriser l'émergence d'algorithmes d'extraction de tendances évolutives à partir des séries multi-temporelles.

L'apport du High Performance Computing et notamment de la technologie GPU sera évalué pour les nouvelles méthodes mises au point. Le gain de temps est ici un élément important mais non-central. En

effet il s'agit en premier lieu d'étudier l'impact de l'utilisation du HPC selon deux aspects : l'adaptation des algorithmes aux architectures massivement parallèles et l'impact de la puissance algorithmique pour l'émergence ou l'amélioration d'algorithmes (exploration exhaustive d'espaces de recherche de paramètres ou d'optimisation par exemple).

Les recherches menées au CESBIO et en collaboration avec des laboratoires partenaires donneront un cadre applicatif et de validation très riche. Le CESBIO dispose de sites instrumentés et d'une grande variété de données terrain qui permettront la validation quantitative des méthodes mises en oeuvre dans la thèse. Le CESBIO dispose aussi de longues séries d'images multi-temporelles (Formosat-2 et Landsat) dont les caractéristiques sont proches de celles de la mission Sentinel-2, ce qui permettra, dès le début de la thèse de réaliser des expériences réalistes.