



ARCUS E2D2

Sujet de Thèse ARCUS E2D2 2017

sélection de la thématique : (cochez une ou plusieurs cases)

- Sp1 « Ville, Aménagement et Développement Durable »:
- Sp2 « Modélisation et Infrastructure pour l'Environnement »:
- Sp3 « Expertise et Traitement en Environnement »:
- Sp4 « Calcul Scientifique »:

Partenaire proposant le Sujet :

Laboratoire d'accueil :

Laboratoire d'Informatique Signal Image de la Côte d'Opale

Responsable(s) : Gilles ROUSSEL, Gilles DELMAIRE

Axe SPECIFI (Système de Perception et Fusion d'Information) , Groupe SENIH (Surveillance en Environnement Naturel et Industriel et Humain)

Université d'accueil : Université du Littoral Côte d'Opale

Partenaire potentiel pour la collaboration et la co-tutelle :

- Si le partenaire n'est pas défini, veuillez sélectionner les partenaires potentiels :

FRANCE LIBAN MAROC PALESTINE

- Si un partenaire est déjà identifié, veuillez compléter les informations suivantes (si disponible) :

Laboratoire d'accueil : Department of Computer Engineering, Faculty of Engineering

Responsable(s) : Antoine ABCHE

Université d'accueil : Université de Balamand, Tripoli, Liban

Mots clés : Factorisation Matricielle Non-négative, Pollution de l'air, optimisation non linéaire.

Points particuliers : (précisez les points particuliers que le candidat devra considérer, langue, compétences)

- Connaissance en analyse matricielle recommandée.
- Connaissance des langages de programmation appréciée

TITRE DE LA THÈSE

Méthodes de NMF Bayésiennes informées. Application à la recherche de sources de pollution

SUJET DE LA THÈSE

(Une page maximum)

Les méthodes de type Factorisation Matricielles Non négatives (NMF) sont actuellement très en vogue dans la communauté traitement du signal de par la diversité des applications qu'elles peuvent concerner. Les applications potentielles vont du traitement audio-phonique, l'imagerie multi- et hyper-spectrale, l'analyse de spectres biomédicaux, ou de manière plus générale la séparation de sources. Ces méthodes consistent à approcher la matrice de données en un produit de deux facteurs de dimensions plus faibles, vérifiant la non-négativité de leurs éléments, appelés respectivement la matrice de sources et la matrice de contribution. Cependant, la plupart de ces méthodes souffrent de problèmes de stabilité, d'initialisation et d'interprétation des résultats concernant les différents facteurs. De plus, ces méthodes suscitent toujours dans la communauté de forts intérêts pour comprendre l'impact d'une forte corrélation dans la matrice de sources et la présence de données aberrantes dans la matrice de données. Récemment, des travaux ont montré l'intérêt pour des méthodes de NMF informées afin de résoudre les problèmes de stabilité. Ces méthodes informées [1-4] prennent en compte les connaissances expertes de certaines entrées de la matrice des sources ainsi que des contraintes de somme à 1.

Une alternative consiste dans cette thèse à développer des méthodes Bayésiennes de NMF informées qui, par conséquence, ne souffriraient pas de problèmes d'initialisation. Cependant, l'enjeu de ce travail consiste à formaliser les contraintes expertes sous forme de densité de probabilité, à savoir les valeurs connues des éléments de la matrice de sources ainsi que la somme à 1 de ses lignes. Des travaux récents ont permis de tenir compte uniquement de la non-négativité des variables, mais actuellement, aucun travail à notre connaissance [5, 6] ne tient compte à la fois des forçages de la matrice de sources et de la somme à 1 de ses lignes. Notre objectif est donc d'essayer de lever ce verrou scientifique lors de cette thèse.

D'un point de vue applicatif, ces méthodes sont appliquées à la séparation des sources de la pollution de l'air ou plus récemment à la séparation spectrale des images terrestres (type d'agriculture, forêt, plan d'eau, ...) par l'imagerie multispectrale dans l'environnement marin et terrestre. En ce qui concerne sa séparation de source de pollution, le laboratoire a participé à différents projets privés ou publics comme par exemple le projet ECUME incluant de nombreux partenaires dont l'INERIS, le laboratoire UCEIV, l'École des Mines de Douai et l'association ATMO Nord Pas de Calais). Ce dernier projet visait à étudier la contribution du trafic maritime (déroit de Calais) sur la pollution atmosphérique globale de la Côte d'Opale à partir de mesures situées au Cap Gris-Nez et Calais port.

L'enjeu principal de cette thèse consiste ici en un diagnostic fin de la matrice des sources de pollution (ou spectre des pixels en imagerie hyperspectrale). Cette matrice rassemble les signatures chimiques des différentes sources en présence. L'apport de la nouvelle méthodologie permettra donc de caractériser la confiance et l'incertitude dans la matrice des sources. Les calculs seront à effectuer en utilisant soit les moyens de calcul massif du laboratoire LISIC ou de la future plateforme.

Références :

- [1] A. Limem, G. Delmaire, M. Puigt, G. Roussel, and D. Courcot, *Non-negative matrix factorization under equality constraints—a study of industrial source identification*, Applied Numerical Mathematics (APNUM), Volume 85, pp. 1-15, November 2014.
- [2] R. Chreiky, G. Delmaire, M. Puigt, G. Roussel, D. Courcot, A. Abche. Split Gradient Method for Informed Non-negative Matrix Factorization, Proc. of LVA/ICA, pp. 376-383, Liberec, Czech Republic, August 25-28, 2015.
- [3] A. Limem, M. Puigt, G. Delmaire, G. Roussel, D. Courcot. Bound constrained weighted NMF for industrial source apportionment. Proc. of MLSP, Reims, France, September 21-24, 2014.
- [4] R. Chreiky, G. Delmaire, C. Dorffler, M. Puigt, G. Roussel, A. Abche. Robust Informed Split Gradient NMF Using Alpha Beta-Divergence For Source Apportionment. *26th IEEE International Workshop on Machine Learning for Signal Processing, (MLSP 2016), Vietri Sul Mare, Salerno, Italy, September 13-16, 2016.*
- [5] Dobigeon N., S. Moussaoui, J.Y.Tourneret, C.Carteret, Bayesian separation of spectral sources under non-negativity and full additivity constraints, Signal Processing, 2009, vol 89, pp 2667-2669.
- [6] Altmann Y., S. McLaughlin, A. Hero, Robust Linear Spectral Unmixing using Anomaly Detection, IEEE Trans. Computational Imaging, vol. 1, no. 3, pp 174-185, Sept. 2015.