



Université Lille Nord de France
Pôle de Recherche
et d'Enseignement Supérieur

Ecole doctorale régionale Sciences Pour l'Ingénieur Lille Nord-de-France - 072



Titre : Méthodes informées de co-factorisation non négatives pour la séparation de sources. Application à la recherche de sources de pollution issues de l'activité humaine terrestre et maritime

Financement prévu : 50% Région

Cofinancement éventuel : 50 % PMCO

(Co)-Directeur de thèse : Gilles ROUSSEL

E-mail : gilles.rousseau@lisic.univ-littoral.fr

Encadrants : Gilles DELMAIRE

E-mail : gilles.delmaire@lisic.univ-littoral.fr

Laboratoire : LISIC (Laboratoire d'Informatique Signal et Image de la Côte d'Opale, EA 4491)

Equipe : SPECIFI

Descriptif :

Les méthodes de type Factorisation Matricielle Non négative (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 et d'interprétation des résultats concernant les différents facteurs. Dans certaines situations, en effet, les lignes de la matrice des sources peuvent être géométriquement proches, ce qui rend les méthodes classiques peu efficaces. Par ailleurs, les données réelles sont entachées de mesures aberrantes. A cette fin, nous avons proposé, dans des travaux récents, des approches dites informées dans lesquelles les informations des experts du domaine sont considérées comme des contraintes dures du problème de NMF [1-4]. Ces informations peuvent provenir de la connaissance experte de certaines sources contenues dans la matrice des sources (contraintes égalité [1,2] ou de borne [3]) ainsi que de la connaissance partielle d'un modèle contenu dans certaines composantes de la matrice de contribution [4].

Extrêmement flexibles, ces approches de NMF informées ont été appliquées à la séparation des polluants atmosphériques par l'équipe du Professeur Courcot (UCEIV-ULCO). En particulier, dans le cadre sa thèse de doctorat, C. Roche a utilisé nos outils pour mesurer l'impact du trafic maritime sur la pollution globale affectant la population vivant sur la Côte d'Opale (projet ECUME financé par la DREAL, ayant notamment fait l'objet de communications au Campus de la Mer).

De nouvelles problématiques émergent lorsque non-pas un mais plusieurs capteurs sont déployés. La co-factorisation se présente comme une alternative avantageuse pour tenir compte de l'hypothèse d'une matrice de sources communes à plusieurs capteurs répartis géographiquement. Certains auteurs se proposent de relaxer cette hypothèse afin de tolérer de faibles fluctuations entre les deux facteurs au sens de la norme L_1 ou L_2 .

Nous proposons dans le cadre de cette thèse d'étendre ces notions de factorisations collectives à des situations où la matrice des sources contient uniquement une partie commune, conduisant à considérer



des couplages plus flexibles entre les différentes matrices à co-factoriser tout en permettant l'ajout de connaissances expertes telles que proposées dans nos précédentes approches de NMF informée [1-4]. D'un point de vue applicatif, ces méthodes sont appliquées à la pollution de l'air dans l'environnement marin et terrestre. L'enjeu principal consiste en un diagnostic fin de la matrice des sources de pollution. Cette matrice rassemble les signatures chimiques des différentes sources en présence. L'apport de la nouvelle méthodologie permettra de prendre en compte plusieurs sites urbains ou ruraux de manière simultanée en considérant que certaines sources sont communes à l'ensemble des sites.

Ce projet s'inscrit dans la continuité du projet ECUME incluant de nombreux partenaires dont l'INERIS (O. Favez), le laboratoire UCEIV (D. Courcot), l'École des Mines de Douai (L. Alleman) et l'association ATMO Nord Pas de Calais (T. Delaunay). Ce projet visait à étudier la contribution du trafic maritime (détroit de Calais) sur la pollution atmosphérique globale de la Côte d'Opale à partir de mesures situées au Cap Gris-Nez et Calais port, mais en ne considérant qu'un seul site. L'exploitation simultanée des deux sites de mesure (ou plus) nécessitera donc la mise en œuvre de ces nouvelles méthodes de co-factorisation.

Mots clés :

Factorisation Matricielle non-négative

Publications récentes de l'équipe en lien avec la thématique.

- [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] M. Plouvin, A. Limem, M. Puigt, G. Delmaire, G. Roussel, D. Courcot. Enhanced NMF initialization using a physical model for pollution source apportionment. *Proc. of ESANN*, pp. 261-266, Bruges, Belgium, April 23-25, 2014.
- [5] C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, Blind Calibration of Mobile Sensors Using Informed Nonnegative Matrix Factorization, in *Proc. of LVA/ICA*, vol. LNCS 9237, pp. 497-505, 2015.
- [6] C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, Blind Mobile Sensor Calibration Using an Informed Nonnegative Matrix Factorization With a Relaxed Rendezvous Model, in *Proc. of ICASSP 2016*, Shanghai, China, March 20-25, 2016.