Models and Methods for Material and Environmental Sciences

XXXVII Cycle

PhD Student: Alessandra Olarini

New chemometric approaches to investigate the spatial – spectral interplay in Hyperspectral imaging

Hyperspectral imaging and multimodal spectroscopic are mature non-destructive tools for the analyst, as well as suitable to monitor chemical systems in time, enabling characterization of composition/structure and their spatial and temporal evolution. From the data analysis standpoint, they match the paradigm of Big Data and pose new challenges to state of art chemometrics methodology.

The research project is aimed at exploiting the integration of machine learning, image processing and chemometrics to develop new approaches suitable to be tuned on the basis of sought information while balancing modeling capability and chemical interpretation.

Main focus will be on hyperspectral imaging data, where two different approaches may be currently envisioned.

On the one hand, multivariate classification methods aiming at assigning a class label to every pixel. On the other hand, spectral unmixing approaches aiming at identifying individual sources of spectral variations, with each measured pixel now described as a linear mixture of the pure spectra characteristic of those unknown individual sources. For classification, the underlying assumption is that the spectral signature measured at one pixel is characteristic of one type of object (class) only. Thus, mixed pixels represent the major problem for classification approaches. By contrast, spectral unmixing is more difficult due to the inverse nature of the problem and can benefit from the knowledge of the presence of unmixed pixels. However, classification requires (some) pixels to be labelled with class membership for model training. On the other hand, there is as well methodology, based on tensor analysis where spatial and spectral dimension are dealt simultaneously, in this case a bilinear model at pixel level is not sought but the hyperspectral data cube is decomposed as the sum of outer product of a component matrix and a vector.

Thus, the thesis project, developed jointly with the University of Lille, will evaluate pros and cons of these approaches, in connection with the structure of the data hand, investigated in both spectral and spatial modes.