UNIVERSITAT DE BARCELONA

Multivariate Signal Processing for Quantitative and Qualitative Analysis of Ion Mobility Spectrometry data, applied to Biomedical Applications and Food Related Applications

Ana Verónica Guamán Novillo

FACULTAT DE FÍSICA

Departament d'Electrònica

MEMÒRIA PER OPTAR AL TÍTOL DE DOCTOR PER LA UNIVERSITAT DE BARCELONA

Doctorat en Enginyeria i Tecnologies Avançades (RD 99/2011)

Multivariate Signal Processing for Quantitative and Qualitative Analysis of Ion Mobility Spectrometry data, applied to Biomedical Applications and Food Related Applications

by

Ana Verónica Guamán Novillo

Director:

Dr. Antonio Pardo

Codirector:

Dr. Josep Samitier

Tutor:

Dr. Antonio Pardo

OBJECTIVES

Objectives

This work deals with signal processing methodologies applied to biomedical and food applications (bio-related applications) analyzed by Ion Mobility Spectrometry (IMS).

During the development of this work, different solutions have been established for providing answers to open issues in the field of IMS signal processing -on its applicability to bio-related applications.

More precisely, the goals of this thesis are presented below:

To show the necessity of performing multivariate signal processing on bio-related datasets measured with IMS instruments. In this thesis, univariate and multivariate algorithms have been applied to different datasets measured with IMS with the main goal of enriching the discussion about the need of using multivariate techniques, instead of univariate techniques by default. The variety of the samples analyzed by IMS has increased in the last years, thus different behaviors such as non linarites and mixture effect in the spectra needs to be tackled. These behaviors are not usually discussed in terms of univariate analysis, but in this thesis a deep understanding about the proper use of univariate and multivariate is debated. In order to understand this problematic issues synthetic and real samples were analyzed with different commercial Ion Mobility Spectrometers. Among of the multivariate algorithms variety, algorithms for resolving the whole spectra and bilinear decomposition techniques were tested. Moreover, different strategies were proposed and explained in the different clinical and food applications.

To develop strategies for IMS calibration. In relation to the previous point, in this thesis different calibration algorithms are probed. The useless of univariate strategies for IMS calibration will be presented and the necessity of multivariate strategies for the quantification of samples will be demonstrated.

To establish validation strategies for qualitative and quantitative analysis of Ion Mobility Spectrometry measurements. Not long ago, the typical IMS application was the detection of a compound over a threshold. New IMS applications implies to have a considerable amount of compounds- some of them unknown, together with many samples for either select a set of compounds or for determining the quantitative values of them. In these new scenarios, to get overfitting on the results and to provide overoptimistic results must be avoided. Thus, it is necessary to establish a proper validation methodology for obtaining reliable results.

To study the adaptation of specific figures of merit for instrumentation in the IMS field. The figures of merit (as sensibility or limit of detection) has to be adapted to specific characteristics of the IMS allowing the comparison between different IMS commercial devices. In this thesis, three IMS devices, with different operating principles have been used. As far as I know, there are a few published works about the comparison of performances between different IMS. This study have a big impact in concepts as

standardization, recalibration, development of libraries of spectra, transferability of results and calibrations. In this thesis, the key characteristics of each spectrometer are discussed and limit of detection of a set of compounds are calculated for establishing main differences.

To develop strategies for compound selection, classification and pattern recognition problems using IMS. The lack of standard libraries of compounds in the IMS field is a trouble for the use of this kind of instruments. Therefore, it is important to develop strategies for guarantee accurately results based on the selection of relevant compounds- even if they are unknown.

The present doctoral dissertation has the following structure:

- **Introduction:** provide a general introduction about the content of this thesis. It is explained the interest of measuring volatile organic compounds in different fields, among of them clinical and food applications. An overview of the typical chemical sense technologies are explained together with the data analysis of them.
- **Objectives:** summarize the objectives of the thesis
- **Chapter One:** explains the nonlinearities and mixture effects in IMS, especially in applications that provides a high dimensional data.
- **Chapter Two:** tackles the signal processing algorithms focusing in Ion Mobility Spectrometry, going through univariate techniques to multivariate techniques.
- **Chapter Three:** explains the different dataset that were studied in this thesis for quantitative and qualitative analysis. It is also explained the performance of the commercial IMS used during this thesis and signal processing methodologies for obtaining reliable results.
- **Chapter Four:** details the qualitative results of the different applications.
- **Chapter Five:** shows the quantitative results of the different datasets and the main implications of using univariate and multivariate techniques.
- **Conclusions:** summarizes relevant conclusions of this doctoral dissertation.
- **Summary in Spanish:** Review the thesis in Spanish