BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Volumul 67 (71), Numărul 1, 2021
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ
DOI:10.2478/bipie-2021-0003

PREPROCESSING METHODS FOR GC-IMS IMAGE ANALYSIS

BY

SIMONA CONDARAGIU[1,*], NICOLAE LUCANU[1], PHILIPP LEBHARDT[2], JENS LANGEJUERGEN[2] and IULIAN B. CIOCOIU[1]

[1]“Gheorghe Asachi” Technical University of Iaşi, Romania
[2]Depart. for Automation in Medicine and Biotechnology, Fraunhofer IPA, Mannheim, Germany

Received: May15, 2021
Accepted for publication: July 26, 2021

Abstract. The paper presents a cascade of preprocessing procedures aiming at enhancing the quality of images generated by gas-chromatography ion mobility spectrometers (GC-IMS). The methods successfully eliminate the Reactant Ion Peak (RIP) lines, while additionally filtering out noisy components. The resulting images enable better localization and matching of the peaks associated with the presence of distinct analytes in the samples under study.

Keywords: ion mobility spectrometry; gas chromatography; filtering.

1. Introduction

Ion mobility spectrometry (IMS) has been developed in the 60’s as a fast trace gas detector solution, and since then it found numerous applications in a broad range of scientific and industrial tasks. More specifically, IMS related technologies are currently used in medicine for identifying pulmonary or gastric diseases, in food quality control for discriminating between original and counterfeited products, or in security/military environments for drugs or explosives detection, based on the analysis of specific combinations of volatile

*Corresponding author; *e-mail*: scondaragiu@etti.tuiasi.ro

organic components (VOC) present in the testing samples. More recently, the technology has also been used for non-invasive COVID-19 detection.

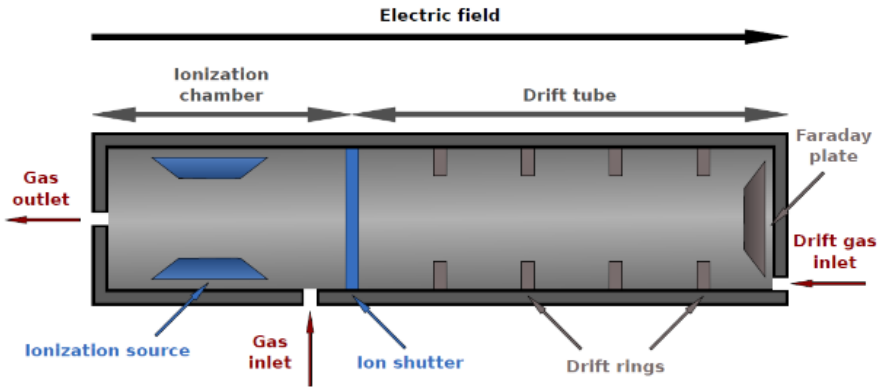

Fig. 1 – Constructive principle of IMS devices (Hauschild *et al.*, 2003).

The basic principle behind IMS is described in Fig. 1. An IMS consists of an ionization chamber separated by an ion shutter from a drift tube, and a detector. In the ionization chamber the samples molecules under study are collected and are ionized (typically using radioactive β- sources, or, more recently, electron emitters or X-rays). Depending on the applied voltage, the generated (positive and/or negative) ions are moved through the drift chamber under atmospheric pressure with a homogenized electrical field in the direction of the detector. Two effects determine the averaged drift time t_D of the ions in the drift chamber: induced by the external electric field the ions are accelerated depending of charge and mass of the ion, due to collisions of the ions with the molecules of a neutral drift gas (*e.g.*, nitrogen, purified air) the ions are decelerated depending on the cross-section of the molecules in their specific orientation. Finally, the ions reaching the end of the drift tube are sensed on a Faraday plate detector. In order to identify the analytes under study, a set of normalization steps are applied in order to compute the so- called reduced ion mobility K_0 given by (Vautz *et al.*, 2018):

$$K_0 = \frac{L}{E\,t_D}\frac{p}{p_0}\frac{T_0}{T}\ [\mathrm{cm^2\,V^{-1}\,s^{-1}}] \tag{1}$$

Simple inspection of Eq. (1) reveals that the K_0 parameter is dependent of the drift tube length L, the intensity of the accelerating electric field E, and the specific experimental temperature T and pressure p of the drift gas (p_0 and T_0 are the normal pressure and temperature, respectively).

While the low cost, fast response and the possibility of constructing portable, hand-held devices are significant advantages of the approach, the

limited resolution in identifying distinct chemical components has led to a significant number of instrumental or operational developments. The present IMS related landscape includes the Field Asymmetric Ion Mobility Spectrometer (FAIMS), time-of-flight secondary ion mass spectrometers (TOF-SIMS) and traveling wave IMS (TW-IMS) among the most used solutions.

In order to characterize the motion of the ions within the drift tube and the associated resolution limitations of the measurement process, analytical models have been proposed. The input parameters are the temporal width w_{Inj} of a Gaussian-shaped ion packet injected into the drift tube, the drift length L (the distance to the detector plate), and the voltage U_{D} generating the electric field. The ions travel along the drift tube with a velocity that is proportional to their specific mobility K and the packet widens due to diffusion, depending on the temperature T, and the Boltzmann's constant k_{B}. The separation between various components is measured by the resolution R, defined as the ration between the drift time t_{D} and the full width at half maximum ($w_{0.5}$) upon arrival at the detector. Since both t_{D} and $w_{0.5}$ can be estimated we may compute R as in Eq. (2) below (e is the elementary electric charge) (Kirk *et al.*, 2013):

$$R = \frac{t_D}{w_{0.5}} = \frac{\dfrac{L^2}{KU_D}}{\sqrt{w^2_{Inj} + \dfrac{16 k_B T \ln 2}{eU_D}\left(\dfrac{L^2}{KU_D}\right)^2}} \tag{2}$$

The resolution R reaches a maximum value R_{opt} at an optimum drift voltage U_{opt} given by:

$$U_{opt} = \sqrt[3]{\frac{8 k_B T \ln 2}{eK^2}\frac{L^4}{w^2_{Inj}}} \tag{3}$$

for which we finally get:

$$R_{opt} = \sqrt[3]{\frac{e}{24\sqrt{3}\ln 2 k_B T K}\frac{L^2}{w_{Inj}}} \tag{4}$$

For the optimal drift voltage U_{opt} we additionally get the width of the ion packet when reaching the detector as $w_{0.5|U_{opt}} = \sqrt{3}\, w_{Inj}$.

Practical implementations are subject to various non-idealities, including the noise of the amplifiers, inhomogeneities of the drift electric field, or Coulomb repulsion, to name a few. The sensitivity of the measurement (reliable detection of low concentration analytes) and the resolution impose conflicting

requirements in terms of the sample volume, hence small samples of about 1 mL should be used for analyzing highly complex compounds, while larger volumes of about 10 mL should be necessary in very low concentrations scenarios (Vautz *et al.*, 2018).

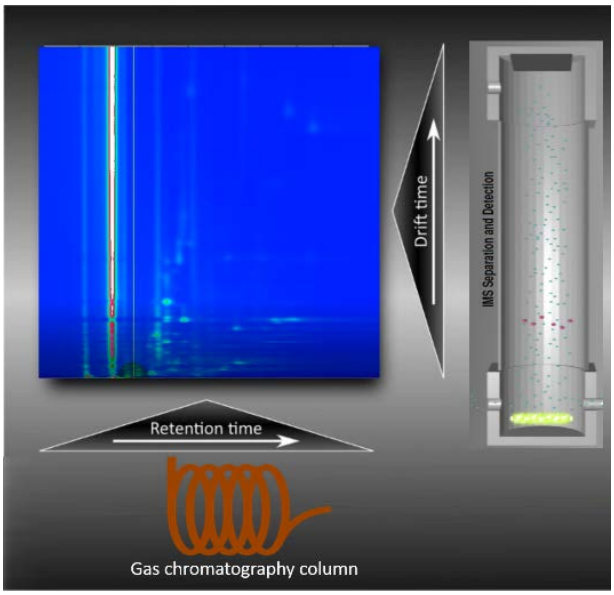

Fig. 2 – Constructive principle of GC-IMS devices.

The typical output of an IMS measurement session reveals the presence of a set of peaks that correspond to distinct VOC components present in the samples under test. In complex, realistic measurement scenarios, we may face the problem of peak superposition due to a clustering phenomenon that takes place in the ionization and drift regions. This would make extremely difficult a clear separation between components, hence modern approaches use a pre-separation step, usually based on gas-chromatography (GC), as in Fig. 2. The GC consists of a long capillary with two phases. A mobile phase with the carrier gas leading the sample under study to the detector and the stationary phase, typically a polymeric layer. Depending on the specific adhesion and dissolution of each analyte, the sample is split up and the analyst exit the GC after a distinct retention time. Both common and multi-capillary columns (MCC) are considered in order to efficiently split components with almost identical ion mobility characteristics that would nevertheless exhibit distinct retention times.

The typical output of a GC-IMS equipment may be represented in 2D or 3D dimensions called chromatograms and reveals a combination of peaks defined on a pair of axes corresponding to the drift time and retention time, respectively, as in Fig. 3.

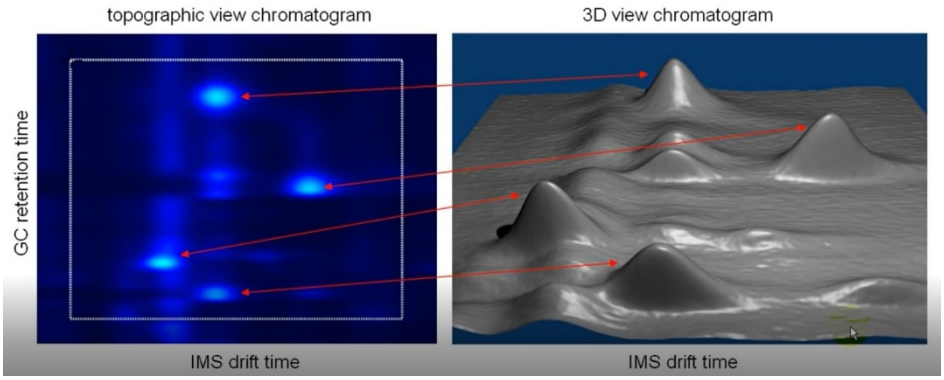

Fig. 3 – Topographic 2D maps and 3D representations of GC-IMS data (GAS, 2011).

2. Preprocessing Methods

Images obtained from practical use of GC-IMS equipments are affected by a series of nonidealities, superimposed noise and nonlinear effects that require the design of a proper processing flow in order to enable accurate detection of the various components present in the test sample, especially for low concentration scenarios.

To start with, it is worth mentioning that even in the absence of actual samples to be measured, ionization may still produce ions originating from the supporting atmosphere (typically, H3O+ or NH4+). As a consequence, those will travel along the drift tube and finally generate specific lines in the image that are collectively known as the Reactant Ion Peak (RIP). Those lines, along with other sources of noise, should be eliminated through proper filtering. While simple spectral averaging may prove useful in simple cases, more performant solutions include Savitzky-Golay and wavelet-based filtering (Szymanska, 2018).

On the other hand, the measurement instruments may degrade over time and call for proper baseline estimation and variation compensation, that use various (polynomial or wavelet) fitting and optimization strategies. Further on, peaks superposition due to clustering phenomena should also be corrected, and existing solutions rely on blind source separation techniques, Independent Components Analysis (ICA), or redundant representations (Zhang *et al.*, 2020).

When performing repeated experiments using the same samples multiple images are obtained. The mobility of the ions suffers mild deviations due to modifications of the pressure and temperature, gas flow variation or the presence of impurities. After noise removal and baseline compensation, matching peaks associated to the same analytes in various images is critical in order to obtain reliable results.

3. Proposed Preprocessing Pipeline

In order to eliminate the presence of the RIP lines and the various noise components, we propose a pre-processing flow that includes the following steps:

a) Savitzky-Golay filtering: a third order bi-dimensional filter that basically implements a smoothing operation aiming at eliminating the low-level background noise from the original images;

b) Removing RIP lines and contaminations in the vertical direction as well as linear artifacts in the horizontal direction by parallel combination of the morphological top-hat operator acting on both dimensions (Meyer, 1977);

c) Calculate a threshold based on the maximal intensity value in a signal-free RIP region and set all intensity values below this threshold to zero. As a consequence, only peaks originating from the true analytes will have an intensity greater zero in the RIP region. As an additional benefit, this step also corrects the baseline drift;

d) Whittaker smoothing to further filter out remaining noise traces from previous operations;

e) apply the watershed transform: this represents a segmentation procedure that enables clear separation of the regions corresponding to individual peaks from the baseline level (Beucher and Lantuejoul, 1979).

The results of applying the various filtering steps above are indicated in Figs. 4-8. The images clearly show effective RIP lines removal, while preserving and separating the patches that include components present in the testing sample. The Savitzky-Golay filter merely removes the background low-level noise from the original image. The combined effect of top-hat filtering, threshold-based cleaning and Whitakker smoothing eliminates most of the unwanted artefacts, while the watershed transform clearly separates the regions corresponding to the various components of the analytes that will further enable matching the corresponding peaks visible in different measurement sessions.

4. Conclusions

The proposed pre-processing flow clearly yields the elimination of the various sources of contaminants of the images provided by the GC-IMS device, while additionally enabling proper separation of the regions that correspond to useful components of the analytes. The position, shape, and area of those regions may be further used in order to implement a peak-matching procedure of corresponding points from distinct images successively acquired during an experimental session. Further study will be devoted to an innovative solution for this task that will use not only the geometric position of the peaks, but also feature-based descriptors showing robustness against mild variations of their location and shape.

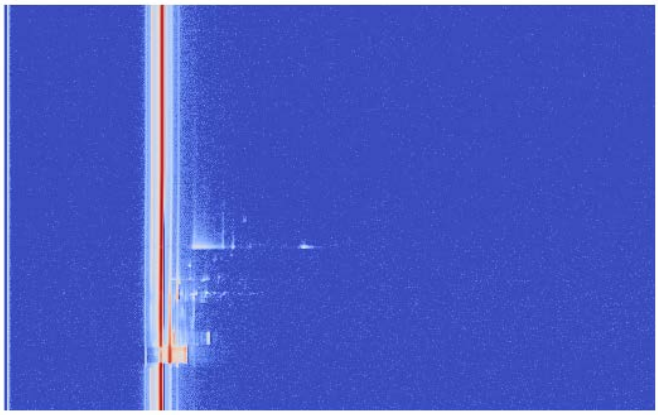

Fig. 4 – Original GC-IMS image to be processed.

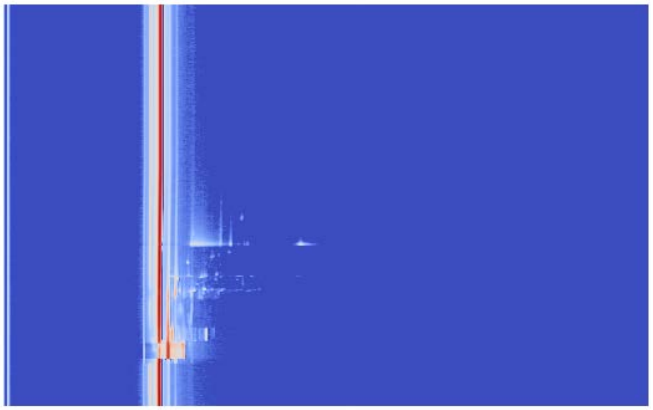

Fig. 5 – Result of Savitzky-Golay filtering.

Fig. 6 – Result of applying the top-hat operator.

Fig. 7 – Combined effect of thresholding and Whitakker smoothing.

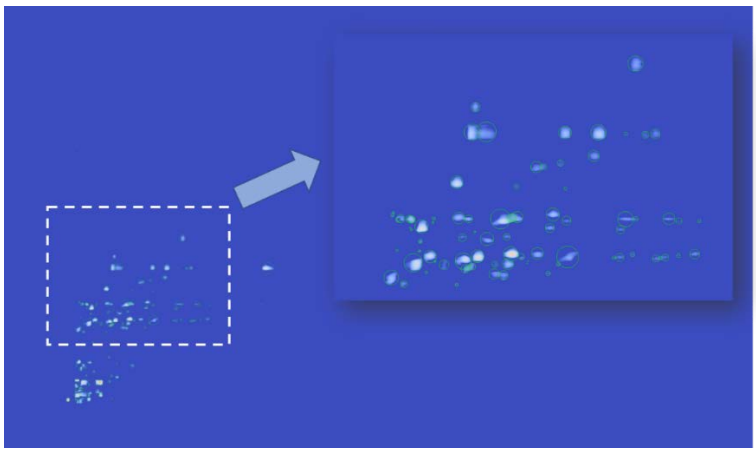

Fig. 8 – Final result after applying the complete preprocessing pipeline.

Acknowledgements. This paper is supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 952378, project BrainTwin (Development of a World-Class Neuroengineering Research Centre by European Twinning) and by a research grant of the TUIASI, project number PN-III-CEI-SUPORT-PO-2020.

REFERENCES

Beucher S., Lantuejoul C., *Use of Watersheds in Contour Detection*, Proc. Internat. Workshop on Image Proc., CCETT/IRISA 1979, Rennes, France, 432-480.

Hauschild A-C., Kopczynski D., D'Addario *et al.*, *Peak Detection Method Evaluation for Ion Mobility Spectrometry by Using Machine Learning Approaches*, Metabolites, **3**, 277-293 (2003).

Kirk AT., Allers M., Cochems P., Langejuergen J., Zimmermann S., *A Compact High Resolution Ion Mobility Spectrometer for Fast Trace Gas Analysis*, Analyst, **138**, 5200-5207 (2013).

Meyer F., *Contrast Feature Extraction*, Quantitative Analysis of Microstructures in Material Sciences, Biology and Medicine, Riederer-Verlag, 1977.

Szymanska E., *Modern Data Science for Analytical Chemical Data - A Comprehensive Review*, Anal. Chim. Acta, **1028**, 1-10 (2018).

Vautz W., Franzke J., Zampolli S., Elmi I., Liedtke S., *On the Potential of Ion Mobility Spectrometry Coupled to GC Pre-Separation e - A Tutorial*, Anal. Chim. Acta, **1024**, 52-64 (2018).

Zhang G., Peng S., Xie Q., Yang L., Cao S., Huang Q., *Multiscale Orthogonal Matching Pursuit Algorithm Combined with Peak Model for Interpreting Ion Mobility Spectra and Achieving Quantitative Analysis*, Anal. Chim. Acta, **1110**, 181-189 (2020).

*** GAS, Tutorial LAV Introduction to GC IMS Chromatograms, https://www.youtube.com/watch?v=v_HfoL6aaFg.

METODE DE PREPROCESARE PENTRU ANALIZA IMAGINILOR GENERATE DE SPECTROMETRELE DE TIP GC-IMS

(Rezumat)

Lucrarea prezintă un set de proceduri de preprocesare menite să îmbunătățească calitatea imaginilor generate de spectrometrele cu mobilitate ionică cu cromatografie în fază gazoasă (GC-IMS). Metodele reușesc să elimine cu succes liniile de vârf ale ionilor reactivi (RIP), filtrând în plus componentele zgomotoase. Imaginile rezultate permit o mai bună localizare și potrivire a vârfurilor asociate cu prezența unor analiți distincți în probele studiate.