PRELIMINARY RESULTS ON QUANTITATIVE GC-IMS ANALYSIS OF ARABICA AND ROBUSTA COFFEES IN MIXES

Kateryna Trach[1], Oleksandr Prystopiuk[1], Martin Sabo[2], Štefan Matejčík [1, 2]

[1]*Department of Experimental Physics, Comenius University, Bratislava, Slovak Republic*
[2]*MaSaTech s.r.o., Sadová 3018/10 91501 Stará Turá, Slovak Republic*
E-mail: trach1@uniba.sk

We have applied MCC-GC-IMS (Multicapillary - Column-Gas-Chromatography - Ion Mobility Spectrometry) technique to qualify the composition of the coffee samples. In total 41 coffee samples were processed and analyzed, (22 samples Arabica, 8 Robusta, and 11 mixtures). Machine learning (ML) and chemometrics methods were applied to the 2D MCC-GC-IMS spectra of the samples. We have achieved 96% accuracy of the composition prediction, which is sufficient value for practical application.

1. Introduction

Coffee is among the most popular beverages worldwide and simultaneously coffee market belongs to the biggest agricultural markets (82,441.77 million U.S. dollars of worldwide revenue in the year 2019) [1]. According to International Coffee Organization [2], the EU consumption of coffee in the year 2019 was about 45 million bags (one bag is equivalent to 60kg).

Over the last decade, consumers' habits towards coffee have changed significantly. Nowadays so-called "third wave" of coffee culture is observed [3]. It can be characterized by the process of coffee transformation from a regular commodity to a valuable handicraft product. This is manifested in form of emergence of the small roasteries and coffee shops, where unique mixes with emphasis on the recognized region of the beans origin and the author's recipe of mixing/brewing are created [4].

Despite habit changes, general consumer behavior remains the same: the most important demand is to get a high-quality product. Studies have investigated factors that can be considered as subjective indicators of quality thus can stimulate or limit coffee consumption and purchase. There is evidence that two main groups of such factors are "sensory preferences" and "functional motives" [5]. The former includes sensory qualities of coffee: taste and smell, while the latter group comprises positive emotions, feeling of being aroused, focused mental state, etc. In terms of objective qualities, both groups of factors are related to the chemical composition of ground coffee that is used to brew the beverage. Sensory qualities are provided by the specific aroma of volatile organic compounds (VOCs), while flavor depends on substances that are extracted from coffee powder with boiling water. Finally, functional motives are fully related to the quantity and bioavailability of caffeine contained in the beverage. Keeping this in mind, coffee quality can be assessed with modern analytical techniques. For instance, laser-induced-breakdown spectroscopy [6], high-performance liquid chromatography [7], gas chromatography [8] were reported to be successful in the analysis of coffee in different forms.

Considering the aforementioned change, the quality assessment remains an issue, especially at coffee shops and roasteries that do not belong to retail networks thus have no access to quality control laboratories. Obviously, such small enterprises can not utilize above stated techniques due to the high cost of equipment and the need for highly-qualified personnel to operate it and to interpret the results. Very few options remain to provide high-quality coffee: either to purchase directly from trusted farmers or believe to the label information provided by gross retailers. Unfortunately, both have been found unreliable. Fraud in coffee mixes is a widespread problem [8-10]. In the most often case Arabica is replaced with more cheap Robusta species. Due to the huge difference in compounds in these species, a poor mix affects beverage quality, alters its taste and aroma.

The ion mobility spectrometry (IMS) method is successfully used for VOCs identification. IMS has numerous advantages due to which it shows a significant growth in use over the last decade. In particular, it is used for the characterization of VOCs in coffee [7, 11]. The method needs very simple sample preparation and no consumables are needed. Also modern IMS devices are table-top or portable and have a reasonable price. However, IMS spectra interpretation demands special knowledge and identification of individual peaks with reference compounds. As was mentioned before, this drawback limits the use of the method in small enterprises.

The presented investigation aimed to find out whether the software based on the machine learning approach can determine the quantitative composition of coffee powder mix containing various proportions of Arabica and Robusta species. In the case of satisfactory performance, such software may eliminate the need for IMS spectra interpretation. With this IMS analysis of coffee mixes may become attractive for small coffee-focused enterprises.

2. Materials and Methods

Total 41 samples were processed and analyzed. Among them 22 samples were pure Arabica species, 8 were pure Robusta and 11 were mixes. Most of the samples were originally in form of roasted beans in various packages of 100-250g weight. Beans were ground by an electric grinder. Mixes that were used in the experiment were created by adding 10-90% (by mass) of Robusta to 90-10% of Arabica. Each mix contained a single kind of each species.

To perform headspace sampling 1.0g of freshly ground coffee in form of powder was put into a 5ml glass vial. Afterward, the vial was heated in the oven at 90 °C for 20 minutes. 2.25ml of headspace vapor was automatically sampled and introduced into the MCC-GC-IMS injector with a maintained temperature of 132 °C. All samples were analyzed by MCC-GC-IMS Peakmachine (MaSaTECH, SK) [12], which consists of the automatic sampler, MCC-GC column, and IMS device. The device was pre-set as follows: drift tube temperature 100°C; drift gas flow was 700 ml/min, sample speed was 30 ml/min and injection speed was 100 ml/min.

All spectra were measured in positive polarity. In total, 41 ion mobility spectra and 41 2D maps were obtained. For spectra visualization and processing MaSaTECH data post-processing software was used [13]. Analysis of obtained spectra was performed with Chemometrics software [13]. It has several machine learning functions that can be used for unknown IMS spectra classification. Random forest architecture (a kind of neural network that is often used in solving classification problems) was used in the present investigation.

3. Results

Original 2D spectra of Arabica and Robusta (Fig.1, b) have differences in peak positions, its intensity, and some peaks are absent. For the relevant example, one can see the 2D IMS spectrum of pure Arabica (Fig.1, a) and pure Robusta (Fig.1, b).

However, it is hard to find out the mix composition from the 2D spectrum (Fig.2, c). Because of this obtained spectra were classified by random forest method provided in the program Chemometrics. As can be seen from Tab. 1, where the part of analyzed mixes is specified, the classification of samples is accurate enough. The accuracy average was 0.96, which is sufficient for practical application. For instance, the mixture containing 90% of Arabica and 10% of Robusta was predicted as 89.9% of Arabica and 10.1% of Robusta. For other entries in the table i.e. for other coffee mixes relevant results also were obtained. The lowest matching value among the presented results is 70.4% against the real value of 90%. This may be due to the geographical origin of the coffee, as the other two coffee samples of this brand, showed higher Arabica content than was actually blended: the revealed value of 83.3% against real content of 80% and revealed value 85% against real content of 80%, respectively.

Tab. 1. Results of the coffee samples classification.

Given composition		Composition revealed with ML	
90% Arabica Bozin Brazil	10% Robusta Caffe Gourmet India	89,9% Arabica Bozin Brazil	10,1% Robusta Caffe Gourmet India
80% Arabica Bozin India	20% Robusta Hardy 3 countries	83,3% Arabica Bozin India	16,7% Robusta Hardy 3 countries
80% Arabica Bozin Guatemala	20% Robusta Trieste	85% Arabica Bozin Guatemala	15% Robusta Trieste
20% Arabica Trieste	80% Robusta CP Guatemala	14,4% Arabica Trieste	85,6% Robusta CP Guatemala
20% Arabica Trieste	80% Robusta CP Guatemala	14,4% Arabica Trieste	85,6% Robusta CP Guatemala

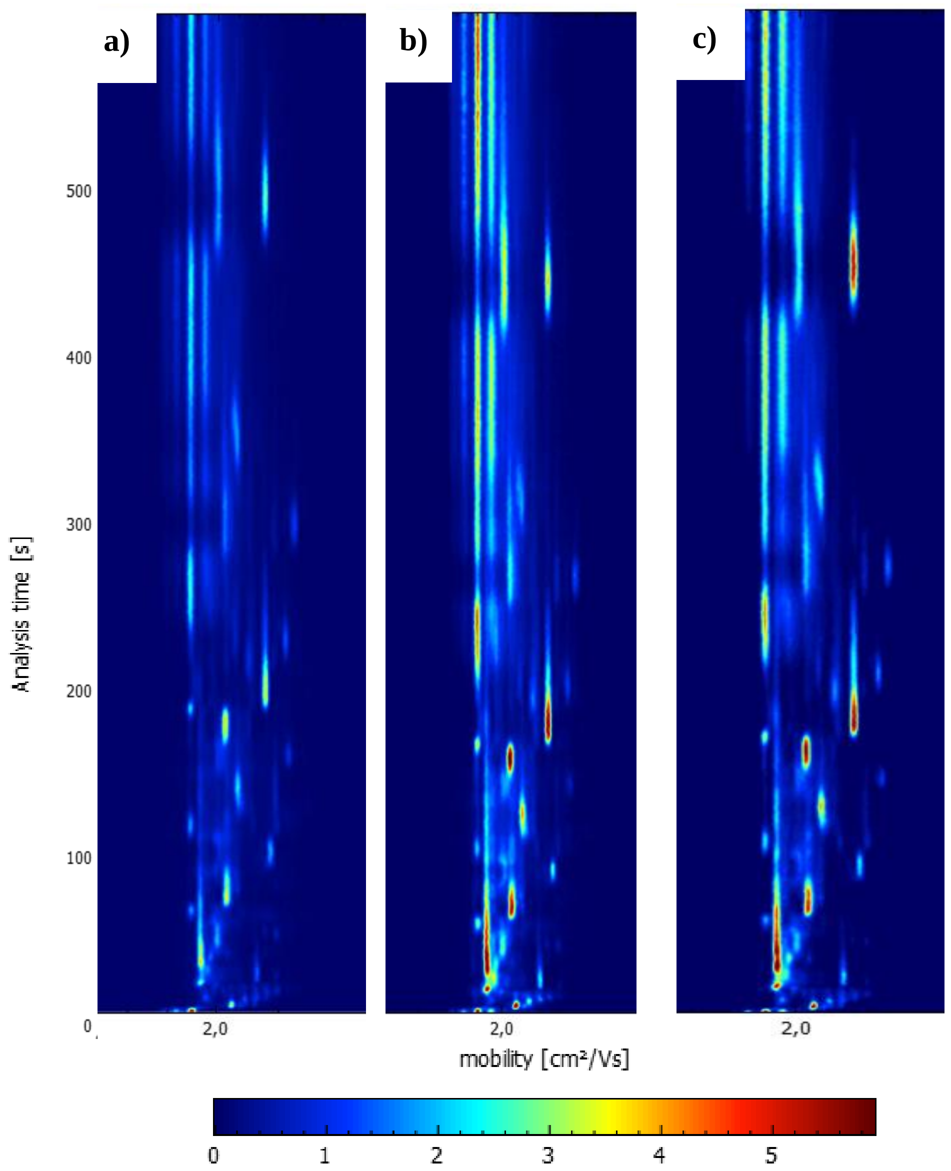

Fig.1. 2D MCC-GC-IMS spectra of VOCs in ground coffee: a) pure arabica (Bozin, Brazil), b) pure robusta (Caffe Gourmet, India), c) Arabica/Robusta mixture (90% Arabica Bozin Brazil and 10% Robusta Caffe Gourmet).

4. Conclusion

MCC-GC-IMS method was used for ground coffee analysis. Due to the method features it was possible to perform the analysis with very little sample preparation, namely with only grinding of roasted beans followed by pouring of the powder to vials. The method is sensitive enough, so meaningful IMS spectra can be obtained by the sampling of VOCs from the vial headspace.

Furthermore, two-dimensional IMS spectra were successfully analyzed by the original software featured with ML algorithms. Application of random forest architecture allowed obtaining average accuracy of 0.96% in the determination of Arabica-Robusta composition in ground coffee mixes. Inclusion of this kind of software to IMS device software bundle could eliminate the need for the employment of skilled professional for the analysis results interpretation. Taking this into consideration, the IMS method may become easy-to-use and cost-effective thus attractive to small coffee-related enterprises. Moreover, it may become a key factor in guaranteed client satisfaction by uncompromised quality of coffee products.

5. References

[1] Oloruntoba A (ed.) 2020: https://www.statista.com/forecasts/758662/revenue-of-the-coffee-market-worldwide-by-country
[2] International Coffee Organization 2020: http://www.ico.org/trade_statistics.asp
[3] Manzo J 2014 *Journal of Arts and Humanities* **3** 8.
[4] Carvalho N B, Minim V P R *et al.* 2015 *Food Research International* **77** 400–407.
[5] Samoggia A and Riedel B 2018 *Appetite* **129** 70–81.
[6] Zhang C and Shen T 2017 *Sensors* **18** 1.
[7] Caporaso N, Whitworth M B *et al.* 2018 *Food Research International.* **108** 628–640
[8] Pauli E D and Barbieri F 2014 *Food Res. Int.* **61**
[9] Dias R C and Valderrama P 2018 *Food Chem.* **255**
[10] Preedy V R (ed.): Coffee in Health and Disease Prevention 2015 *Academic Press*.
[11] Konieczka P P, Aliaño-González M J *et al.* 2020 *Sensors.* **20** 3123.
[12] MCCGC-AIMS PeakMachine infopage; https://www.masatech.eu/peakmachine
[13] MaSaTECH control software; https://www.masatech.eu/control-software