



Analytical Methods

Continuous statistical modelling for rapid detection of adulteration of extra virgin olive oil using mid infrared and Raman spectroscopic data

Konstantia Georgouli^a, Jesus Martinez Del Rincon^b, Anastasios Koidis^{a,*}^a Queens University Belfast, Institute for Global Food Security, Belfast, Northern Ireland, UK^b Queens University Belfast, Institute of Electronics, Communications and Information Technology, Belfast, Northern Ireland, UK

ARTICLE INFO

Article history:

Received 1 March 2016

Received in revised form 11 August 2016

Accepted 3 September 2016

Available online 08 September 2016

Keywords:

Continuous statistical modelling

Dimensionality reduction

Rapid detection

Adulteration

Extra virgin olive oil

FT-IR

RAMAN

Spectroscopy

ABSTRACT

The main objective of this work was to develop a novel dimensionality reduction technique as a part of an integrated pattern recognition solution capable of identifying adulterants such as hazelnut oil in extra virgin olive oil at low percentages based on spectroscopic chemical fingerprints. A novel Continuous Locality Preserving Projections (CLPP) technique is proposed which allows the modelling of the continuous nature of the produced in-house admixtures as data series instead of discrete points. The maintenance of the continuous structure of the data manifold enables the better visualisation of this examined classification problem and facilitates the more accurate utilisation of the manifold for detecting the adulterants. The performance of the proposed technique is validated with two different spectroscopic techniques (Raman and Fourier transform infrared, FT-IR). In all cases studied, CLPP accompanied by k-Nearest Neighbors (kNN) algorithm was found to outperform any other state-of-the-art pattern recognition techniques.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The interdisciplinary collaborations between engineering, computer science and analytical science have led to the development of contemporary analytical instruments that allow the extraction of great amount of chemical information for a large number of samples relatively quickly and effortless. However, the produced analytical data (spectroscopic, chromatographic, isotopic, sensorial, etc.) are often multivariate data matrices which demand appropriate chemometric analysis. In chemometrics, mathematical and statistical methods are used for processing and capturing the most important and relevant content within the multivariate data. Despite the fact that a few multivariate methods are used in the area of food analysis either alone or in combination with other methods (Berrueta, Alonso-Salces, & Héberger, 2007), there is an increasing demand for the introduction of novel and more intelligent pattern recognition methods for tackling more complex food analysis challenges such as food adulteration issues observed worldwide (Lohumi, Lee, Lee, & Cho, 2015).

One of the most common adulterations occurring is mixing one commodity product or ingredient with another one in small

percentages where the two ingredients are of a very similar chemical nature. In these cases, current chemometric techniques somehow fail to identify the fraudulent sample accurately (Ozen & Mauer, 2002; Šmejkalová & Piccolo, 2010) or use the same samples for both calibration and validation steps of the model (Christy, Kasemsumran, Du, & OZAKI, 2004; López-Díez, Bianchi, & Goodacre, 2003), which biases the results. An indicative example of on-going food fraud is the adulteration of extra virgin olive oil, a premium and high value commodity with renowned health properties (Zhang, Qi, Zou, & Liu, 2011). Despite the establishment of a strict legislation framework, including specific analytical parameters defining the purity of the oil (International Olive Council (a), 2016; Agriculture & Rural Development, European Commission, 2013), the extra virgin olive oil adulteration with other lower value vegetable oils still remains an important issue for the consumers and the olive oil sector alike (European Commission, 2013; Frankel, 2010).

One of these adulterants is hazelnut oil, which has very similar triacylglycerol, total sterol and fatty acid composition with extra virgin olive oil and has concerned numerous researchers (Koidis & Osorio, 2013; Parker et al., 2014; Pena, Cárdenas, Gallego, & Valcárcel, 2005). Extra virgin olive oil can be adulterated with hazelnut oil in two different ways: adulteration with crude hazelnut oil and adulteration with refined hazelnut oil. The identification of the adulteration with refined hazelnut oil is increasingly

* Corresponding author.

E-mail addresses: kgeorgouli01@qub.ac.uk (K. Georgouli), j.martinez-del-rincon@qub.ac.uk (J. Martinez Del Rincon), t.koidis@qub.ac.uk (A. Koidis).

difficult due to the removal of markers like filberstone, a volatile compound unique to hazelnut oil, and other minor components through the refining process in addition to the similarity of the triacylglycerol profile of both oils (Flores, Del Castillo, Blanch, & Herraiz, 2006).

Most research efforts aiming to address this adulteration problem have made use of chromatographic analytical methods. Despite providing satisfactory results by analysing the triacylglycerol content (International Olive Council (b), 2013), polar components (Zabaras & Gordon, 2004) and using sterol fractions, 4,4'-Dimethylsterols (Damirchi, Savage, & Dutta, 2005), n-alkanes (Webster, Simpson, & Shanks, 2001) and filberstone (Flores et al., 2006) as possible markers, chromatographic methods involve complicated process steps, demand a large amount of time and financial resources and require access to laboratory facilities. Therefore, it is urgent to develop simple, inexpensive, rapid and accurate alternative methods to determine adulterants in extra virgin olive oil in environments that time and fast decisions are important (ports, control points, market surveys and other rapid testing environments).

Apart from chromatographic, several spectroscopic techniques in combination with chemometric methods have been proposed as rapid screening techniques for the authentication of extra virgin olive oil and the detection and quantification of its adulteration with hazelnut oil. Adulteration of olive oil with hazelnut oil at levels of 25% and higher was detected using Fourier transform infrared (FT-IR) coupled with partial least squares (PLS) analysis (Ozen & Mauer, 2002). Moreover, the same combination has been used for developing a method for the estimation of extra virgin olive oil adulteration with edible oils including hazelnut oil. The produced PLS models for the case of the hazelnut oil showed a relatively good performance (relative error of prediction, REP = 20.8 and correlation factor $R^2 = 0.9351$) (Maggio, Cerretani, Chiavaro, Kaufman, & Bendini, 2010). Multiple linear regression (MLR) models constructed using FT-IR data for extra virgin olive oil-hazelnut oil admixtures claim to be capable of detecting hazelnut oil content in olive oil with a 5% limit of detection (Lerma-García, Ramis-Ramos, Herrero-Martínez, & Simó-Alfonso, 2010). In another study, high gradient diffusion NMR spectroscopy coupled with discriminant analysis (DA) was used for detecting rapidly the adulteration of extra virgin olive oils with seed and nut oils. The lower limit of detection for the case of hazelnut oil was 30% (Šmejkalová & Piccolo, 2010). The development of an artificial neural network in 600 MHz ^1H NMR and ^{13}C NMR data achieved a limit of 8% (García-González, Mannina, DImperio, Segre, & Aparicio, 2004). In a recent study, 60 MHz ^1H NMR spectral data in combination with PLS regression achieved a limit of detection at the level of 11.2% w/w (Parker et al., 2014). However, it has to be highlighted that the aforementioned studies tackling this adulteration of extra virgin olive oil with little or great success do not claim explicitly if

the hazelnut oil is refined or crude and they are not often validated adequately and correctly which might produce overestimated and/or overfitted results.

The detection of adulterants at low levels (5–20%) is still quite challenging even for high end methods such as chromatography (Osorio, Haughey, Elliott, & Koidis, 2014a; Zhang et al., 2011). There is a need for more research in the field of data analysis of complex chemical data, especially spectroscopic data which are by nature multivariate. More accurate statistical methods are required to be used on top of existing analytical methods that would not necessarily demand a large number of samples and are independent of statistical interpretations (Frankel, 2010).

The present work introduces a novel continuous statistical modelling technique which extends the Locality Preserving Projections (LPP) dimensionality reduction technique to the cases where data are considered as a continuous variable. Data are modelled as data series and the continuity is preserved during the learning and dimensionality reduction by building two graphs incorporating neighbourhood information of the data set. In this way, the proposed technique has been designed, developed and tested coupled with k-Nearest Neighbors (kNN) classifier on the adulteration of extra virgin olive oil with hazelnut oil using spectra from two different spectroscopic techniques. Preliminary results obtained are compared with the performance of state-of-the-art supervised pattern recognition techniques.

2. Theory and algorithm

2.1. The proposed method: Continuous Locality Preserving Projections (CLPP)

Continuous Locality Preserving Projections technique is a semi-supervised linear method that enables the dimensionality reduction for learning manifolds characterised by continuous data. It extends the linear dimensionality reduction technique LPP (He & Niyogi, 2003) preserving continuity as in previous non-linear techniques such as Temporal Laplacian Eigenmaps (TLE) (Lewandowski, Martinez-del Rincon, Makris, & Nebel, 2010). LPP was chosen as the base method due to its properties and advantages against other dimensionality reduction techniques such as principal component analysis (PCA) (Wold, Esbensen, & Geladi, 1987) or linear discriminant analysis (LDA) (Fisher, 1938), especially when the input data show linear properties (He & Niyogi, 2003). Given a set of $Y = y_1, y_2, \dots, y_n$ data points in high dimensional space ($y_k \in R^D$) (see Fig. 1a), CLPP is able to transform this into its low dimensional space by mapping it to a set of points $Z = m_1, m_2, \dots, m_n$ ($m_k \in R^d$) with $d \ll D$ (see Fig. 1b), while preserving the continuity of the data.

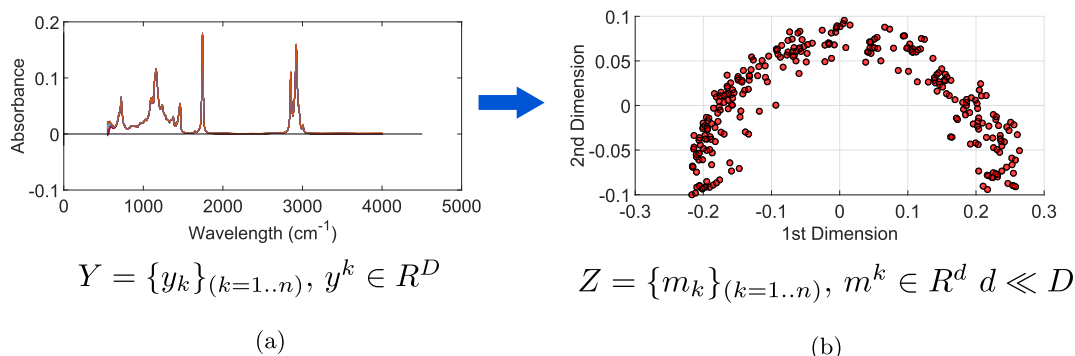


Fig. 1. Definition and application of CLPP: (a) Data points in high dimensional space; (b) Data points in low dimensional space.

Download English Version:

<https://daneshyari.com/en/article/7586970>

Download Persian Version:

<https://daneshyari.com/article/7586970>

[Daneshyari.com](https://daneshyari.com)