Chemometrics and Intelligent Laboratory Systems 135 (2014) 126–132

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

The potential of multivariate analysis to phase identification based on X-ray diffraction patterns

G. Balcerowska-Czerniak [a,*], A. Wronkowski [a], A.J. Antończak [b], Ł. Skowroński [a], A.A. Wronkowska [a]

[a] Institute of Mathematics and Physics, University of Technology and Life Sciences, S. Kaliskiego 7, PL-85789 Bydgoszcz, Poland
[b] Department of Electronics, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, PL-50370 Wrocław, Poland

ARTICLE INFO

Article history:
Received 21 December 2013
Received in revised form 18 March 2014
Accepted 13 April 2014
Available online 21 April 2014

Keywords:
X-ray diffraction
Compositional analysis
PCA biplot
Au–Sn intermetallic compounds
Laser-induced titanium oxides

ABSTRACT

The possibility of using multivariate data analysis in the X-ray diffraction technique for chemical compound identification is presented. Two data sets consisting of X-ray patterns of the Au–Sn thin film system and laser-induced coloured oxide films on a titanium substrate (Ti–O) are analysed as examples. The proposed approach is based on the assumption that data has a bilinear mathematical structure and, therefore, it can be subjected to principal component analysis and a biplot can be created. The key information was gained when the PCA model that we have developed was employed to reference a database of X-ray patterns. The obtained biplot, in a simple way, allowed for a visual appraisal of the chemical compounds that constituted the analysed samples. Chemical information about the composition of the two sets of targets, Au–Sn and Ti–O, was found to be consistent with that reported in the literature and determined using other methods. Moreover, in the case of Ti–O, which is known as a very complex system, multivariate data analysis enabled the determination of new supplementary information concerning the role of the individual chemical compounds in terms of laser-induced colouring of the titanium surface.

1. Introduction

X-ray diffractometry (XRD) is widely used as an analytical technique for identification of crystalline materials. A diffractogram showing peak intensities against the glancing angle position (θ or 2θ) provides some chemical information about a sample that sets XRD apart from other analytical tools. The ability to accurately index the unit cell, i.e. find the assignment of Miller indices (hkl) to each observed peak, has become the important step in phase compositional analysis. In general, indexing is a routine task used in commercially available computer programmes and databases, which offer mathematical models applied for peak profile fitting (peak position, shape and intensity). It should be noted that a search among a huge number of results yields a list of suggestions, none of which result in a decisively correct profile fit. Having knowledge of the typical phases formed in the analysed target makes it relatively easy to identify all of the dominant phases. Problems can arise when the samples have a multicomponent structure and a lot of library profiles are potential candidates for their phase identification. In multiphase samples, a confusion can arise when two or more components have reflection peaks at the same (or close) diffraction angles, within the measurable limits. With a large number of closely spaced Bragg peaks, the developed profile fit methods may give unstable results. Moreover, statistically good curve fits are not always chemically meaningful and great care must be exercised when choosing the phase model to describe the data profile. In such cases any assistance is helpful for a proper modelling of experimental data and for understanding the material itself.

This paper provides a guide to rapid phase identification using multivariate data analysis (MVA) applied to X-ray diffraction patterns. Two MVA methods, namely principal component analysis (PCA) and multivariate curve resolution (MCR), are mainly used in analytical techniques for extracting the significant information from a large data set and have been reported on many times (e.g. Refs. [1–3] and references therein). Recently, Rodriguez et al. [4] reported the increasing importance of PCA application in the extraction of chemical information contained in micro-X-ray fluorescence (μ-XRF) data.

This study aims at expanding the utility of multivariate statistical analysis methods to phase identification based on X-ray diffractograms. In this case, PCA in combination with a biplot as a graphical display of the results could be an effective tool for data exploration. Application of a biplot in PCA analysis is a rather standard procedure in commercial computer programmes. For example such a biplot can be used as a diagnostic tool for choosing an appropriate model for data analysis, or to study a correlation between an analysed data set and an external data set.

The present work deals with the use of a reference database set as a new data set. This approach allowed us to find which of the reference diffraction profiles would be effective for phase identification. To the best of our knowledge this is a novel application of a PCA biplot in the XRD technique for data compositional analysis. In the commonly used

* Corresponding author. Tel.: +48 52 3408632; fax: +48 52 3408643.
E-mail address: grazyna.czerniak@utp.edu.pl (G. Balcerowska-Czerniak).

http://dx.doi.org/10.1016/j.chemolab.2014.04.010
0169-7439/

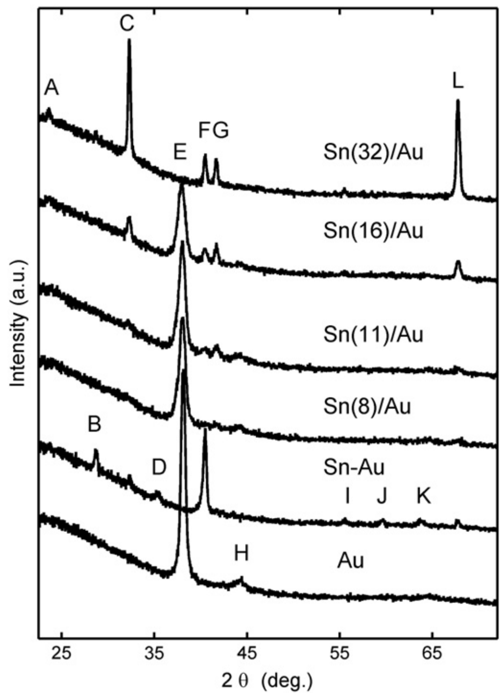

Fig. 1. X-ray diffractograms of Au, Sn–Au, and Sn/Au thin films deposited on glass substrate [6].

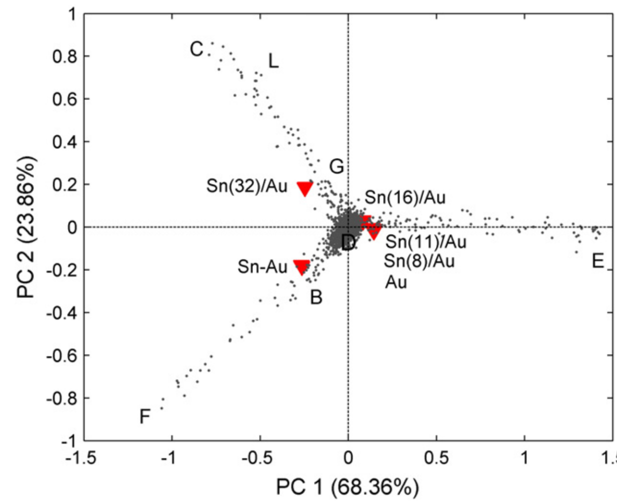

Fig. 2. The standard biplot of the first two PCs for the Sn/Au data set.

procedure the experimental XRD data set is qualitatively analysed using the Rietveld refinement method [5]. As a new concept, a PCA biplot can be an alternative tool for a rapid identification process that is useful in those cases where the qualitative information is satisfactory, without specifying the percentage of particular compounds.

To demonstrate the potential of the proposed approach in terms of chemical composition two experimental X-ray diffraction data sets were investigated by PCA. In order to test its correctness, the method has been applied to a set of X-ray diffraction patterns recorded for the Au–Sn thin film samples. In a previous study [6] the creation of AuSn, $AuSn_2$, $AuSn_5$ intermetallic phases and β-Sn phase was revealed; therefore the results given by MVA could be easily compared and checked.

The second type of data studied in this work is a set of X-ray diffraction patterns recorded for laser-induced coloured oxide films grown on a titanium substrate (Ti–O). In such a system compositional analyses revealed the formation of coatings of titanium oxides that consisted of a mixture of polycrystalline oxides such as Ti_2O, TiO, Ti_2O_3 and TiO_2 [7,8]. Apart from the mentioned phases, there are reports about the formation of β-Ti, Ti_6O and Ti_3O phases on laser-treated titanium surfaces [9,10]. The phases of titanium oxides $TiO_{1.04}$ and Ti_3O, titanium nitrides TiN and $TiN_{0.3}$ were identified and the appearance of titanium oxynitride TiN_xO_y was postulated for laser processing of Ti plates in air in Ref. [11]. The formation of titanium nitride phases was also reported by György et al. [12], but in their experiment the laser irradiation of the Ti surface took place in a nitrogen atmosphere. In addition to this, the formation of TiN, Ti_2N, $TiN_{0.6}O_{0.4}$, $Ti(NO_3)_4$ and TiO_2 was discussed in Ref. [13] for the films prepared by magnetron sputtering in controlled gas (argon, nitrogen and oxygen) flow rates.

In general, since there is a large number of reference database profiles (about 600 for the Ti–O system in the ICCD-JCPDS database) it is very difficult to decompose an experimental X-ray diffractogram into the crystalline components for a Ti–O multicomponent sample. Specifically, in terms of a compositional analysis, no detailed evidence of titanium inserted phases Ti_xX_y (X = N, O or N–O) being formed in the laser-induced colour marking films on titanium substrate has so far been shown. Therefore, here the capabilities of the PCA biplot for an examination of the titanium oxides, titanium nitrides and titanium oxynitride phases in the Ti–O system are discussed and compared with the results presented in the literature.

2. Materials and methods

2.1. Materials and experimental procedures

In a previous investigation of ours the Au–Sn X-ray data were recorded to study the reactive interdiffusion between Au and Sn nanolayers [6]. Two types of samples for the Au–Sn system were analysed. In bilayer samples (Sn/Au), the Au layer ($d = 20$ nm) was first thermally evaporated on a glass (Bk7 — Menzel) substrate, and subsequently, a thin Sn film of different thicknesses (Sn on Au: $d_{Sn} = 8$ nm, 11 nm, 16 nm, 32 nm) was deposited. The second type of binary system (containing 39 at. % Sn (Sn–Au)) was produced by co-deposition of metals from separate sources. Evaporation of particular metals (purity 99.999%, Johnson-Matthey, Co.) was performed from conical tungsten baskets in a vacuum of 10^{-3} Pa.

The second analysis was conducted for samples of titanium oxide films characterized by different colours obtained by laser-induced heating of the electropolished titanium substrate (purity >99.9%, made by MTI). The irradiation of the Ti surface was done using a Yb: glass fibre laser system (λ = 1.1 μm), operating in a pulse mode with a repetition rate in the range between 20 and 80 kHz and pulse duration of 230 ns at normal atmospheric pressure. The coating layers of various colours (yellow, orange, red, purple, blue and light blue) were formed on the target surface by varying the cumulated laser fluence (i.e. 158,

Table 1
Percentage of the variance captured by the PCA model for the Sn/Au thin film system.

Principal component number	Eigenvalue of Cov(X)	% Variance captured this PC	% Variance captured total
1	$9.43e + 002$	68.36	68.36
2	$3.29e + 002$	23.86	92.22
3	$7.09e + 001$	5.14	97.36
4	$3.51e + 001$	2.55	99.90