Measurement 47 (2014) 193–199

Contents lists available at ScienceDirect

Measurement

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

Feature extraction for identification of drug and explosive concealed by body packing based on positive matrix factorization

Yongcheng Jiang [a,b], Ping Liu [c,*]

[a] School of Electric Engineering and Automation, Anhui University, Hefei, Anhui 230601, PR China
[b] Key Laboratory of Intelligent Computing & Signal Processing, Anhui University, Hefei, Anhui 230601, PR China
[c] School of Electronic Science & Applied Physics, Hefei University of Technology, Hefei, Anhui 230009, PR China

ARTICLE INFO

Article history:
Received 5 September 2012
Received in revised form 11 August 2013
Accepted 22 August 2013
Available online 31 August 2013

Keywords:
Positive matrix factorization
Feature extraction
Body packing
Energy dispersive X-ray diffraction

ABSTRACT

In the analysis towards the energy dispersive X-ray diffraction (EDXRD) spectra of drug and explosive concealed by body packing, positive matrix factorization (PMF) was introduced to extract features from EDXRD spectra of samples in a set of drugs and explosive concealed in the anthropomorphic phantom, because PMF prevents the negative factors from occurring, avoids contradicting physical reality, and makes factors more easily interpretable. In order to compare with the features extracted by PMF, Principal Component Analysis (PCA) and robust PCA were investigated. Then, K-nearest neighbour (KNN) and support vector machine (SVM) were introduced to classify the samples according to the features extracted by PMF, PCA and robust PCA. It is shown that the recognition rates obtained by PMF are highest (above 99.5%) and insensitive to classifiers. This work demonstrates that PMF is effective in feature extraction for identification of drug and explosive concealed by body packing.

1. Introduction

With the spread of drug and explosive concealed by body packing, developing effective methods and instruments to detect and identify body packing has been widely focused on. Among various techniques, such as fluorescence of dual-emission quantum dots hybrid [1], fast neutron scattering [2], CT [3], terahertz imaging [4] and homogeneous phase protein-based assay [5], the potential of energy dispersive X-ray diffraction (EDXRD) for detection in complex background has been demonstrated due to its non-destructive, high selection and high efficiency [6–16]. However, during the analysis of EDXRD spectrum of drug and explosive concealed by body packing, feature extraction is still a challenge, mainly due to the interference of body, limit of detection time, and the poor signal-to-noise ratio of spectrum. Therefore, the position of diffraction peak of drug and explosive is not obvious [6,7]. Besides, dimensionality reduction is necessary to mitigate the so-called "curse of dimensionality" because the diffraction profile consists of high-dimensional data.

As methods of feature extraction, principal component analysis (PCA) [6,7,11–14] and robust PCA [13,14] have yielded satisfactory results in analyzing the EDXRD spectrum. However, their application is subject to the possible presence of negative factor loadings which are difficult to interpret in terms of non-negative definite physical parameter such as concentration, mass and spectral intensity. During the past few years, PMF has been developed [17–22], whose weighted least-squared problem resolution prevents the negative factors from occurring, avoids contradicting physical reality, and makes factors more eas-

* Corresponding author. Address: School of Electronic Science & Applied Physics, Hefei University of Technology, No. 193 Tunxi Road, Hefei, Anhui 230009, PR China. Tel.: +86 551 62901465; fax: +86 551 62919106.
E-mail address: 147213929@163.com (P. Liu).

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.08.056

ily interpretable. In addition, PMF makes each data point to be individually weighed, and allows the analyst to adjust the influence of each data point depending on the confidence in the measurement. For example, data below detection can be retained to be used in the model, with the associated uncertainty adjusted so these data points have less influence on the solution than measurements above the detection limit. It largely overcomes the limitations of PCA and robust PCA.

In recent years, PMF has been applied mainly and successfully to source apportionment of pollutants in aerosols, soils and sediments [21–26]. In this paper, PMF is introduced as a method of feature extraction to analyze EDXRD spectrum of drug and explosive concealed by body packing. Meanwhile, PCA and robust PCA are reviewed for comparison. Moreover, features extracted by above algorithms are evaluated by KNN [27] and SVM [28].

2. Materials and methods

2.1. Experimental system

Following the Bragg's law, momentum transfer q can be expressed as a function of the energy of the photon E and the angle θ according to $q = CE\sin(\theta/2)$, where $C = 1.01354\ \text{Å}^{-1}\ \text{keV}^{-1}$ [29]. The scheme of EDXRD spectrometer in our experiment is shown in Fig. 1, which is a laboratory-built energy-scanning diffractometer in the Institute of Intelligent Machines, Chinese Academy of Sciences [12]. The polychromatic source of X-ray comprised an X-ray tube and the related collimators to define the incident and scattered beams. The X-ray source has a tungsten target, and could be operated up to 75 kVp and up to 3 mA current. A Canberra planar high purity germanium detector with a resolution of 140 eV at 5.9 keV combined with a Canberra spectrum master InSpector 2000 was adopted. The spectrum master InSpector 2000 has 1024 channels, corresponding the value from 0 to 3 of the momentum transfer. The widths of the four slit collimator (from X-ray source to detector) are 0.5, 0.75, 0.75, and 2.0 mm, respectively. The distances between X-ray source and sample (anthropomorphic phantom) , between sample and the third collimator, and between the third collimator and the fourth collimator, are 35 mm, 130 mm and 50 mm respectively. The scatter angle is 4°. In order to minimize detection time and avoid the potential damage to human in the future practice, the radiation time of detector is limited to 30 s, and the radiation measured at the detector is 0.75 mSv in 30 s.

2.2. Materials

As a methodological study, an anthropomorphic phantom was developed. It is designed based on Chengdu dosimetric phantom which is certified by International Commission on Radiation Units and Measurements [30]. The anthropomorphic phantom is divided into three parts including of head, chest, abdomen, and simulates more than 20 kinds of organs. The anthropomorphic phantom is non-uniform and the main parameters are listed as follows: height is 170 cm, weight is 65 kg, sitting height is 91 cm, head circumference is 55 cm, bust is 75 cm, waistline is 75 cm, hipline is 85 cm, chest depth is 24 cm, chest breadth is 28 cm, shoulder breadth is 36 cm and hip breadth is 32 cm. In this experiment, drugs (heroin and methamphetamine), drug precursors (phenylacetic acid and piperonal), explosive (TNT) and disruptors (body) which the criminal commonly used were selected as samples. The samples were filled into centrifuges tube of 5 ml, which was fixed by a depth of 8 cm in the stomach of anthropomorphic phantom. All samples are shown in Table 1, and each sample was measured 40 times.

2.3. Positive matrix factorization model

2.3.1. Mathematical formulation

The mathematical development of PMF is described here based on Paatero and Tapper [17,18]. PMF is a multivariate factor analysis tool that decomposes a matrix of sample data into two matrices–factor contributions and factor profiles, which need to be interpreted by an analyst

Table 1
The labels of samples.

Label	Samples
Group 1	TNT + body
Group 2	Phenylacetic acid + body
Group 3	Methamphetamine + body
Group 4	Heroin + body
Group 5	Piperonal + body
Group 6	Body

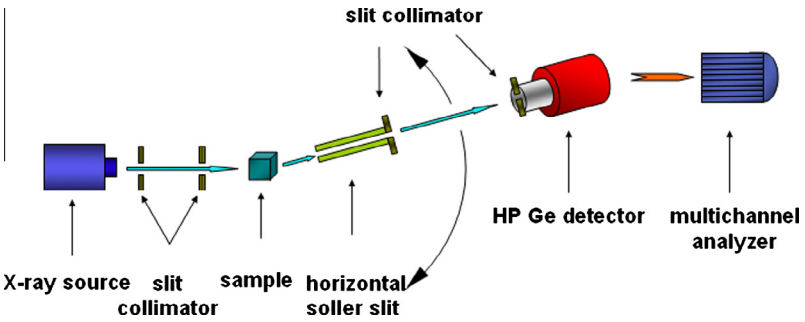

Fig. 1. The geometry of the energy dispersive X-ray diffraction system.