Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 199 (2018) 170–178

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

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

Adaptive handling of Rayleigh and Raman scatter of fluorescence data based on evaluation of the degree of spectral overlap

Yingtian Hu [a], Chao Liu [a], Xiaoping Wang [a,b,*], Dongdong Zhao [c]

[a] State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China
[b] Key Laboratory of Ocean Observation-Imaging Testbed of Zhejiang University, Zhoushan 316021, China
[c] Institute of Advanced Digital Technology and Instrumentation, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 25 August 2017
Received in revised form 19 March 2018
Accepted 23 March 2018
Available online 26 March 2018

Keywords:
Excitation emission matrix
Scatter handling
Degree of spectral overlap
Adaptability
Fitting and interpolation

ABSTRACT

At present the general scatter handling methods are unsatisfactory when scatter and fluorescence seriously overlap in excitation emission matrix. In this study, an adaptive method for scatter handling of fluorescence data is proposed. Firstly, the Raman scatter was corrected by subtracting the baseline of deionized water which was collected in each experiment to adapt to the intensity fluctuations. Then, the degrees of spectral overlap between Rayleigh scatter and fluorescence were classified into three categories based on the distance between the spectral peaks. The corresponding algorithms, including setting to zero, fitting on single or both sides, were implemented after the evaluation of the degree of overlap for individual emission spectra. The proposed method minimized the number of fitting and interpolation processes, which reduced complexity, saved time, avoided overfitting, and most importantly assured the authenticity of data. Furthermore, the effectiveness of this procedure on the subsequent PARAFAC analysis was assessed and compared to Delaunay interpolation by conducting experiments with four typical organic chemicals and real water samples. Using this method, we conducted long-term monitoring of tap water and river water near a dyeing and printing plant. This method can be used for improving adaptability and accuracy in the scatter handling of fluorescence data.

1. Introduction

Fluorescence spectroscopy is receiving increasing attention for water quality and pollution monitoring, as it is a rapid, highly sensitive, and reagentless technique that requires no sample pretreatment [1]. Fluorescence spectrophotometry can discriminate between dissolved organic matter (DOM) fractions; it can, in particular, differentiate the labile fraction into environmentally significant components [2] . In recent years, a wide range of both excitation and emission wavelengths have been simultaneously scanned to generate fluorescence excitation emission matrix (EEM), which represents a detailed map of the fluorescence properties of the mixture. Some common analysis methods for EEM include observational methods of fluorescence peaks [3], parallel factor analysis (PARAFAC) [4]. However, the intrinsic presence of scatter effects in the EEM measurements poses a practical problem. Moreover, a technical difficulty has to be faced that there is no way to evaluate the real contribution of scatter in the data while handling it because the Rayleigh and Raman scatter always exists at the same time with fluorescence signals.

Until now, handling methods of scatter can largely be divided into two categories. The first category involves techniques where the spectral data are selectively used in computation and modeling, but the values remain unchanged, such as introducing weight factors [5], independent component analysis [6,7], and two-direction resection [8]. The second category involves techniques through which the spectral data are corrected using missing values [9], interpolated data [10–12], or subtracting the standard baseline [3]. However, the first category does not eliminate the impact completely and only involves mathematical statistics methods. In regard to the second category, Elcoroaristizabal et al. [9] proposed that it is more robust to use interpolated data than missing values. Recently, an algorithm has been widely adopted for replacing the scatter areas by three-dimensional Delaunay interpolation [13]. Unfortunately, the correction results of present methods are always unsatisfactory when the scatter seriously overlaps with the signal of analytes. The main reason for this is that it is too blind to utilize the same algorithm in two-dimensional interpolation without considering the degree of overlap. In addition, two-dimensional interpolation is too complex and often causes overfitting. Therefore, one-dimensional interpolation on individual emission spectra is used in this study.

* Corresponding author at: State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China.
E-mail address: xpwang@zju.edu.cn (X. Wang).

https://doi.org/10.1016/j.saa.2018.03.058
1386-1425/

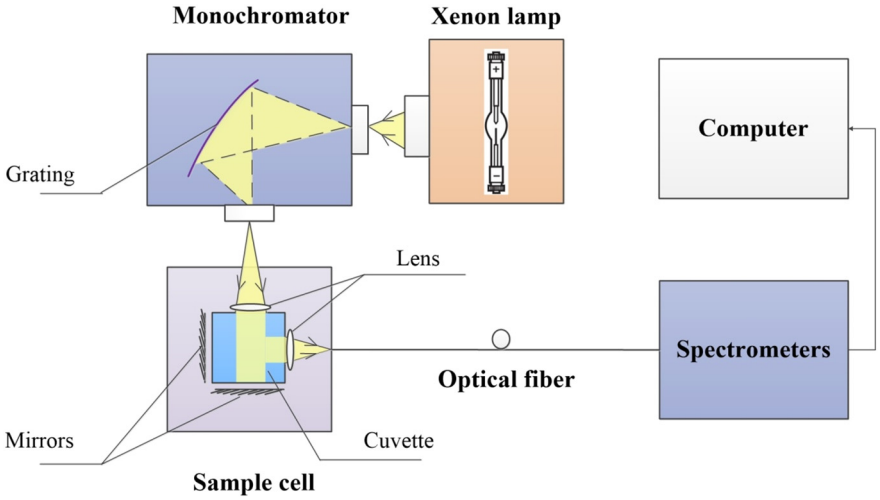


Fig. 1. Experimental setup.

In the present work, an adaptive method is proposed to minimize the influence of scatter based on evaluation of the degree of spectral overlap. Because interpolations are only the estimation of true spectra and sometimes they are not real, very few fitted values are used in this method based on their individual degrees of spectral overlap after the evaluation. The effect of this procedure on the subsequent PARAFAC analysis is assessed and compared to Delaunay interpolation. Data demonstrating better correction results were obtained by performing experiments using four typical organic chemicals and by monitoring the quality of real tap water and river water samples. The results suggest that the proposed method enables adaptive handling of Rayleigh and Raman scatter, prevents overfitting, and guarantees better authenticity of the fluorescence data.

2. Materials and Methods

2.1. Experimental Setup

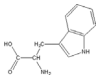
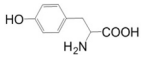
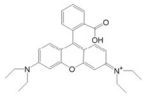
An experimental system was designed and constructed with four basic components: a light source, a monochromator, a sample cell, and an optical receiver, to measure the three-dimensional fluorescence spectrum (Fig. 1). The light source was a 150 W xenon lamp (XBO 150 W/4, OSRAM). The monochromator (7IMS3011B, Saifan Photoelectric Instrument, China) with 1200 grooves mm-1 gratings blazed at 300 nm was used to produce a single exciting light. Both the entrance and exit slit widths were 5 mm. A spectrometer (MAYA2000PRO, Ocean Optics, USA) was used as the optical receiver. A quartz cuvette was held in a 4-way holder (CUV-ALL, Ocean Optics, USA) with two collimating lenses and two mirrors that couple to optical fibers. The fluorescence spectra were obtained with excitations from 240 to 600 nm and emission wavelengths from 240 to 700 nm with step sizes of 5 and 2 nm, respectively. In addition, the effects related to wavelength-dependent efficiencies of the instrumental components were corrected on both excitation and emission intensities following Coble et al. [14].

All computations were performed using Matlab (The MathWorks Inc., Natick, MA, USA) R2013a. The PARAFAC models were constructed with the N-way toolbox (www.models.kvl.dk) [15]. The comparison method of Delaunay interpolation was also obtained from www.models.kvl.dk.

2.2. Water Samples

In this study, we chose four typical chemicals, tryptophan, tyrosine, humic acid, and rhodamine B, and prepared their aqueous solutions to collect their three-dimensional fluorescence spectra. The first three are natural organic matter, occurring commonly in natural water. Rhodamine B is a common and carcinogenic dye and is often found in rivers near printing and dyeing mills. The aqueous solutions of each chemical were prepared by dissolving a certain amount of AnalaR™ or AristAR™ grade substances in deionized water from a Milli-Q water purification system (Millipore, Billerica, MA, USA). Appropriate concentration ranges were selected to ensure that the fluorescence intensity was

Table 1
Tap water, river water, and water samples containing four organic chemicals.

Water sample	Structural formula	Concentrations	Number of samples
Tryptophan		50–500 μg/L	5
Tyrosine		50–500 μg/L	5
Humic acid	Complex mixture	2–50 mg/L	5
Rhodamine B		5–100 μg/L	5
Tap water	–	–	25
River water	–	–	25