Mining Raman Big Data with Independent Component Analysis Based Blind Source Separation Algorithm

(*Collage of Science and Engineering, Aoyama Gakuin University*) ○Rintaro Shimada
Keywords: Raman Spectroscopy; Multivariate Analysis; Big Data

Raman spectroscopy is a powerful method for the non-invasive analysis of complex chemical systems such as living cells. However, due to the comprehensive nature of Raman scattering, in which virtually every molecule under laser irradiation contributes to the signal, the analysis of complex spectral data obtained from biological samples has been a challenge to researchers, especially who were not experts in spectral analysis. Therefore, the development of easy-to-use methodologies to extract and separate biochemical information from the observed spectral dataset is an important challenge that must be overcome.

In this study, we present a blind source separation algorithm for a large Raman spectral dataset based on independent component analysis (ICA) for separating mixed Raman spectra obtained from heterogeneous mixtures into component spectra. Systematic study of artificially prepared model Raman data sets revealed that a simple preprocess through numerical differentiation could enhance intrinsically low statistical independence of typical Raman spectral patterns. Hence, we propose a new spectral component decomposition method based on the scheme shown in Figure 1. Here, the principal component analysis serves as preprocessing for dimensionality and noise reductions. The numerically differentiated principal component loading spectra were analyzed with ICA for deducing the unmixing matrix (**W**) of the observed dataset (**A**) into chemically interpretable components spectra ($\mathbf{P}_{chem}$) and its quantity vectors ($\mathbf{T}_{chem}$). Application to experimentally obtained Raman datasets from the biological specimen will be discussed in the presentation.

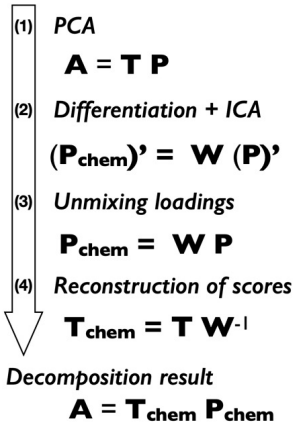

Figure 1. Schematic of the blind source separation algorithm.
PCA: Principal component analysis.