Improved Peak Detection and Quantification for Mass Spectrometry Analysis using CWT based Method

Tzu-Ching Wu and Kun-Pin Wu

Institute of Biomedical Informatics, National Yang Ming University, Taipei, Taiwan

Introduction

In general, peak detection procedure can decompose into three consequent steps: smoothing, baseline correction and peak finding. Nevertheless, difficulties arise to detect true peaks for low Signal to Noise Ratio (SNR) and complexity of MS signals. Baseline removal and smoothing steps are unrecoverable after applied on MS data inadequately and results are inconsistent among common peak detection algorithms.

Due to increasing practicability in the proteomics field though high computational challenges label-free quantification techniques are, we present a novel and concise algorithm for MS data peak detection and especially peak quantification by utilize 2-D Continuous Wavelet Transform (CWT) coefficients matrix information, which derived from applying CWT over MS raw data. No further baseline removal or peak smoothing preprocessing steps are required before peak detection.

Contribution

This algorithm improves the accuracy to detect true peaks in a robust manner by investigate the ridge lines on the CWT coefficients matrix respect to peak positions. Furthermore, peak quantification and resolution of overlapping peaks, which are crucial process when neighboring peaks overlap strongly enough, are presented in this approach by utilize specific scales CWT coefficients information. Our approach has been validated by using both simulated Gaussian overlapping MS data and physics-based computer model of mass spectrometry.
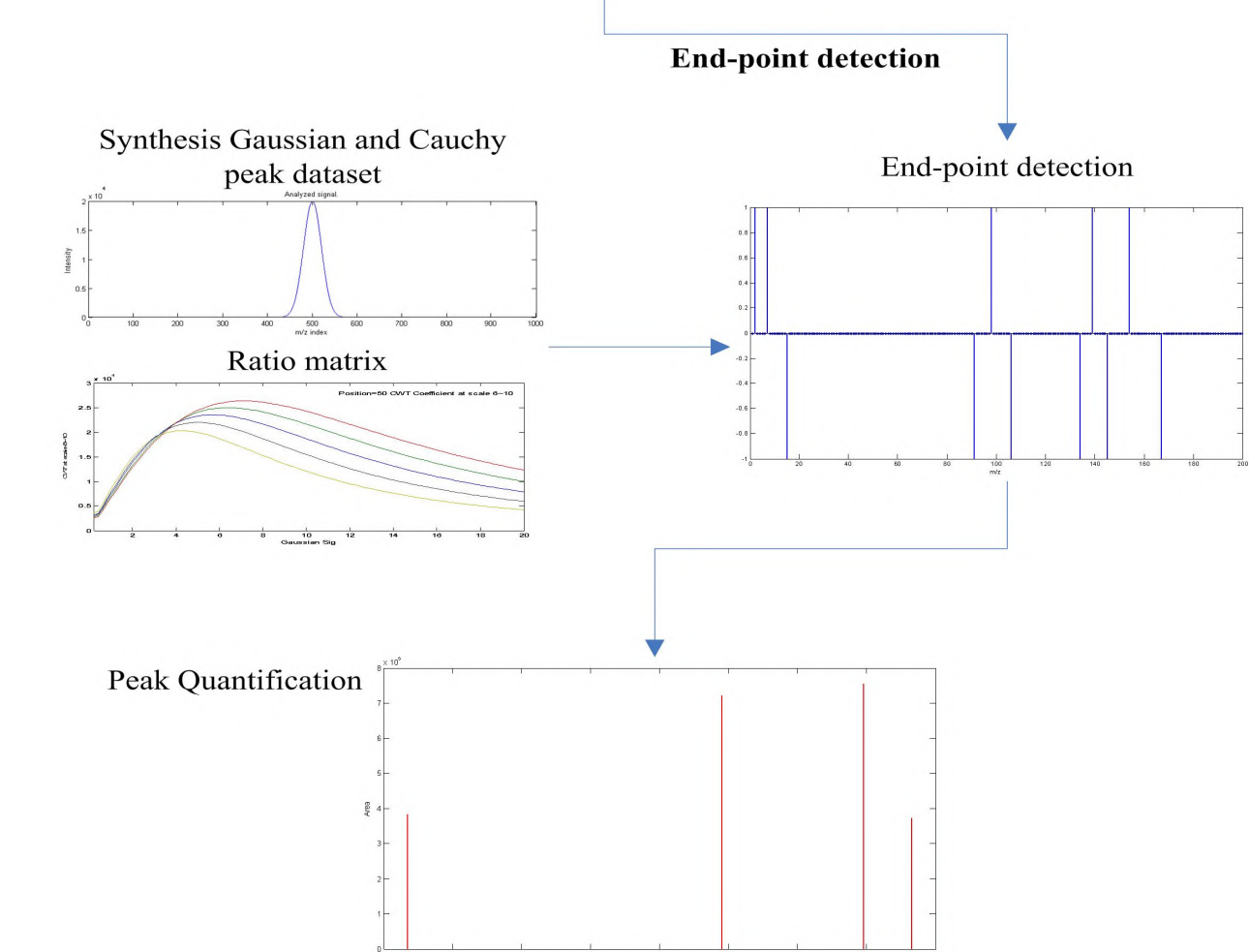
Flowchart
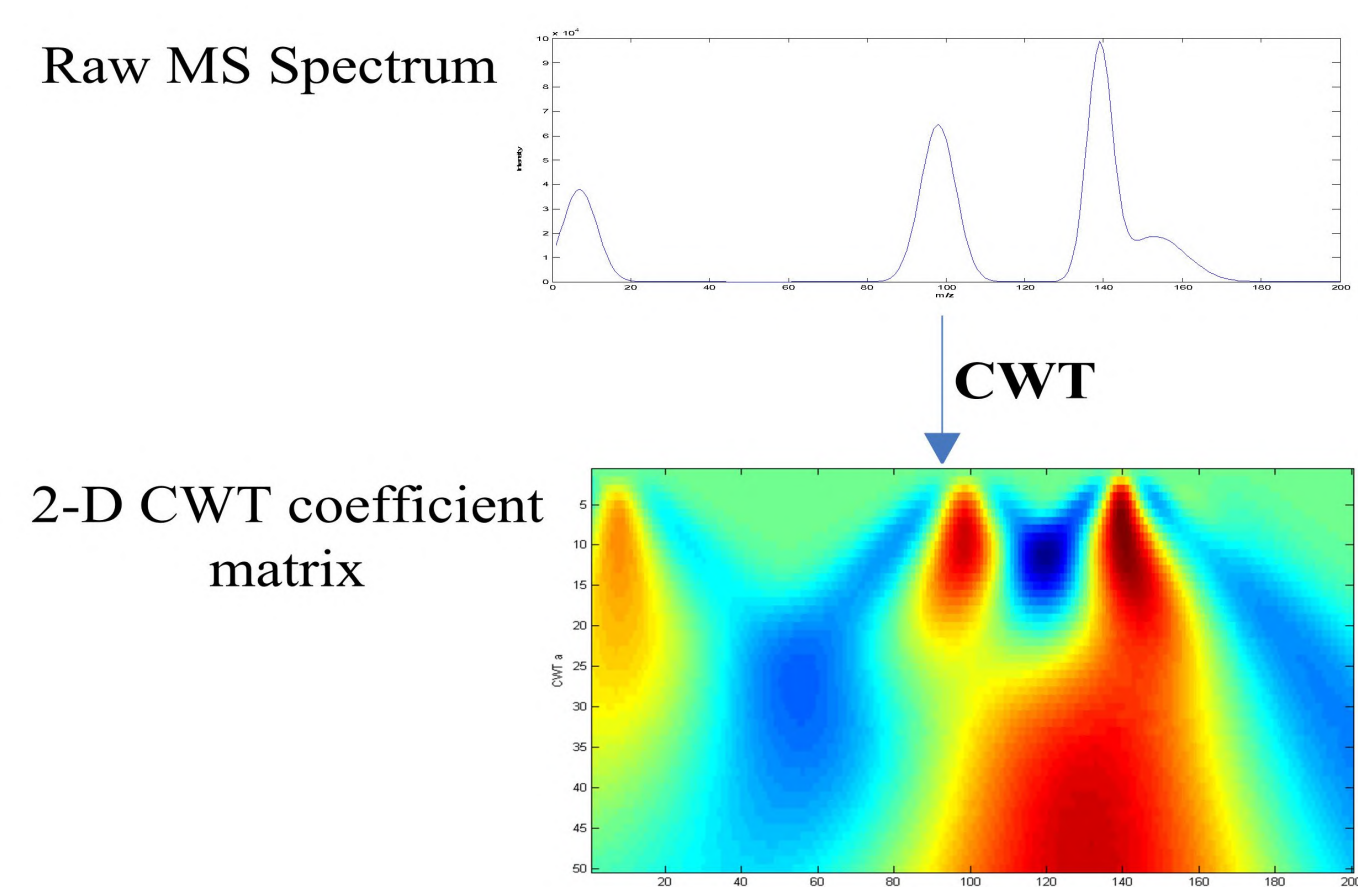
Continuous wavelet transform
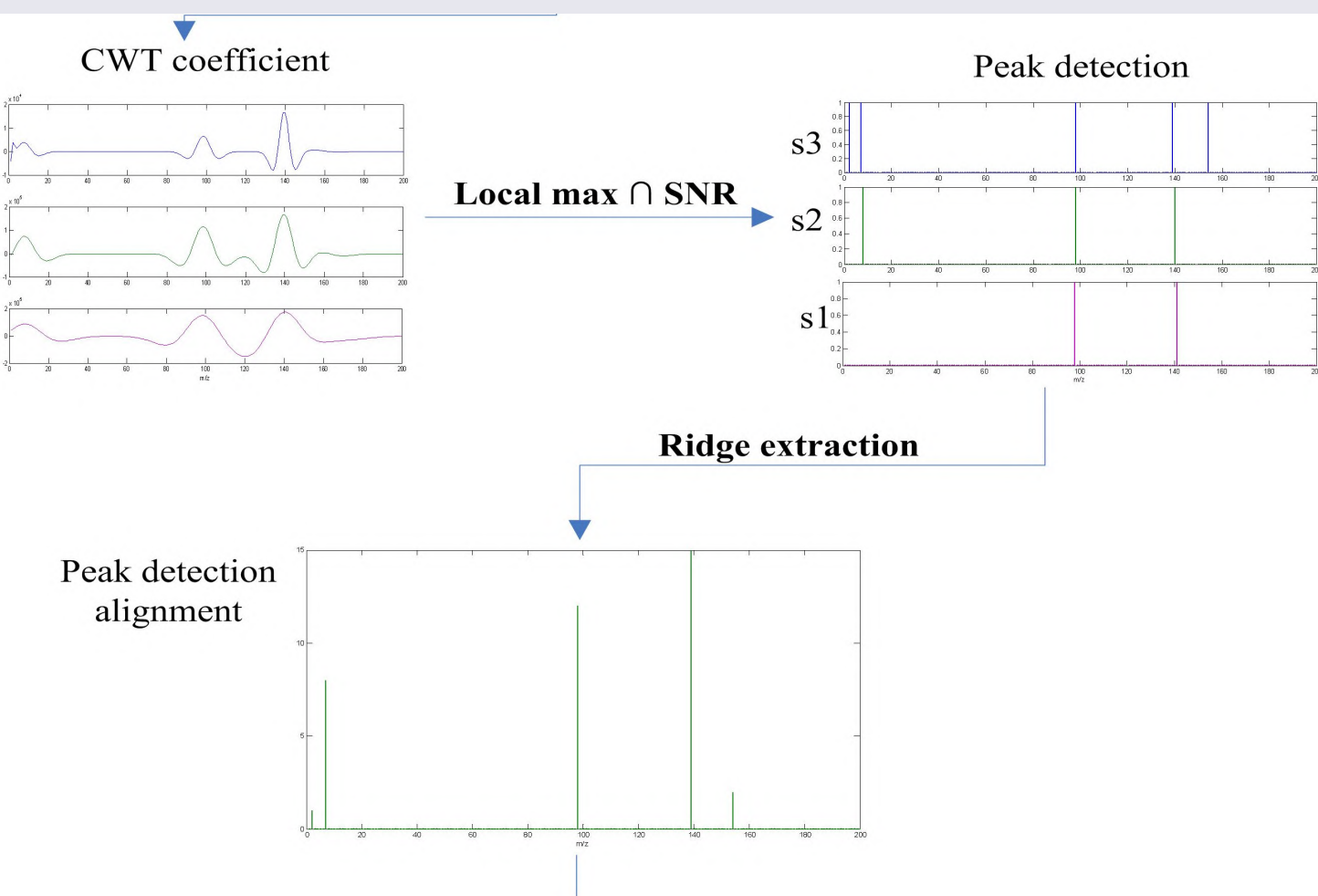
Peak Detection processing

1. Peak detection in every scale
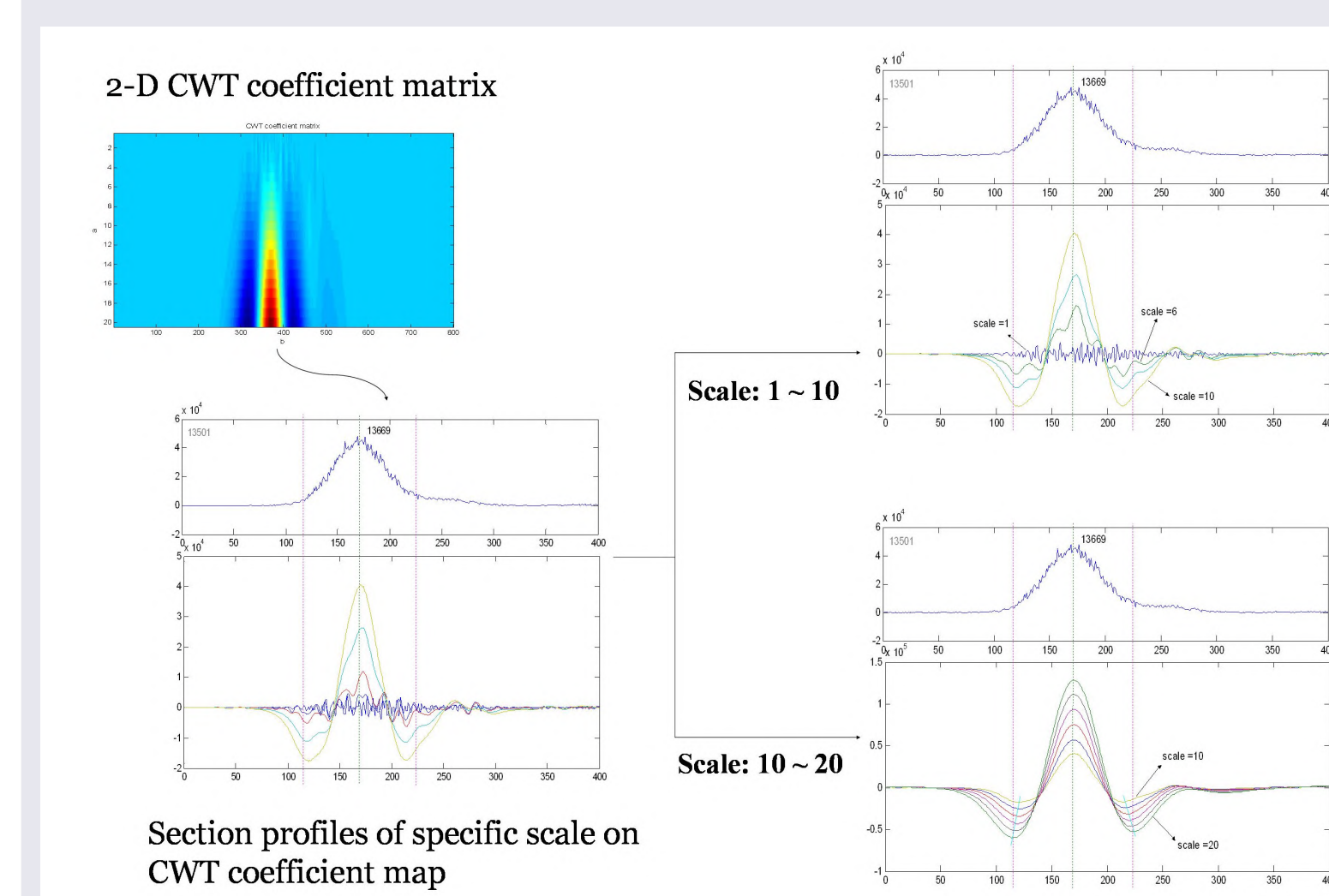2. Ridge extraction
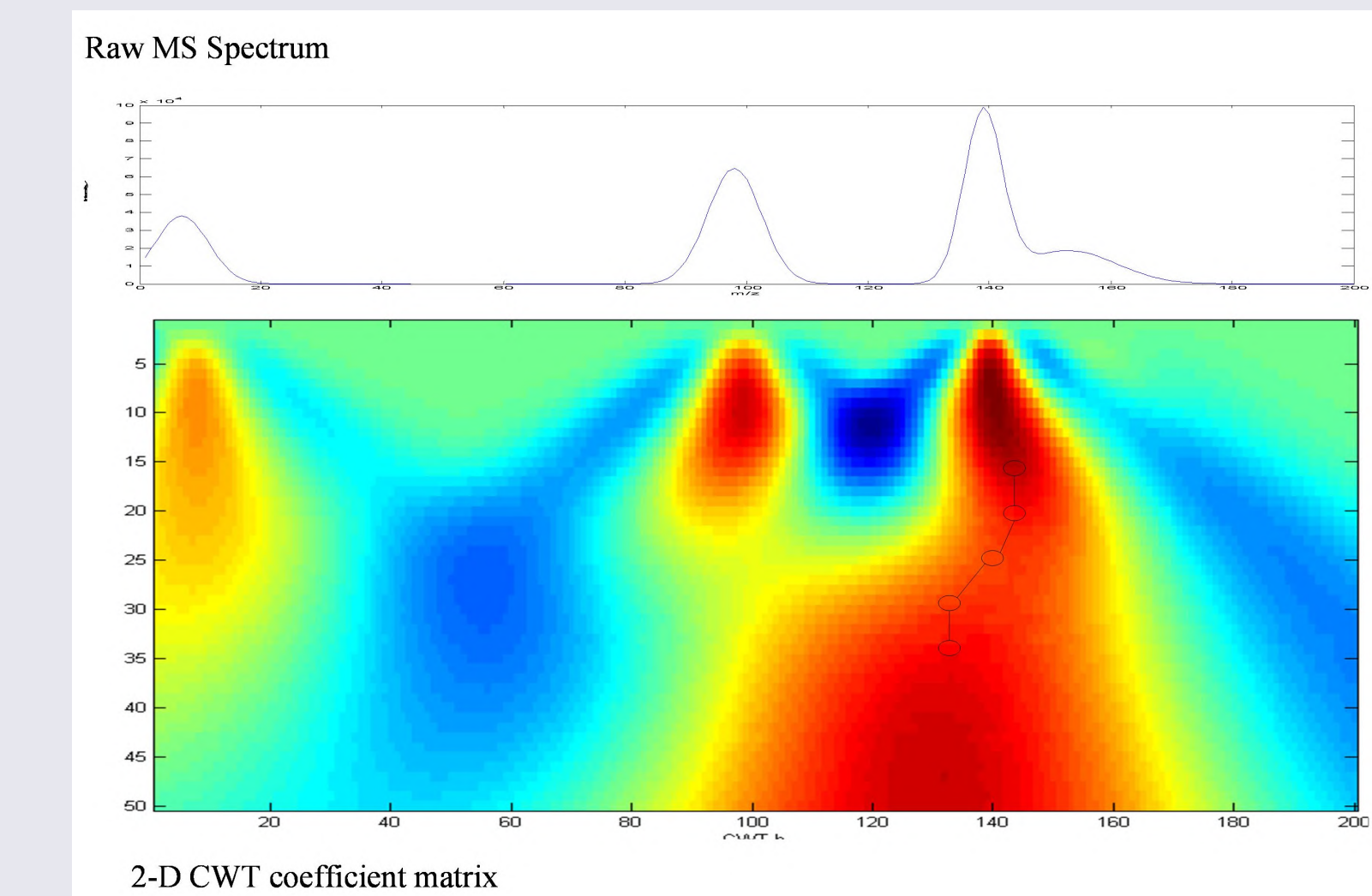
Peak quantification processing

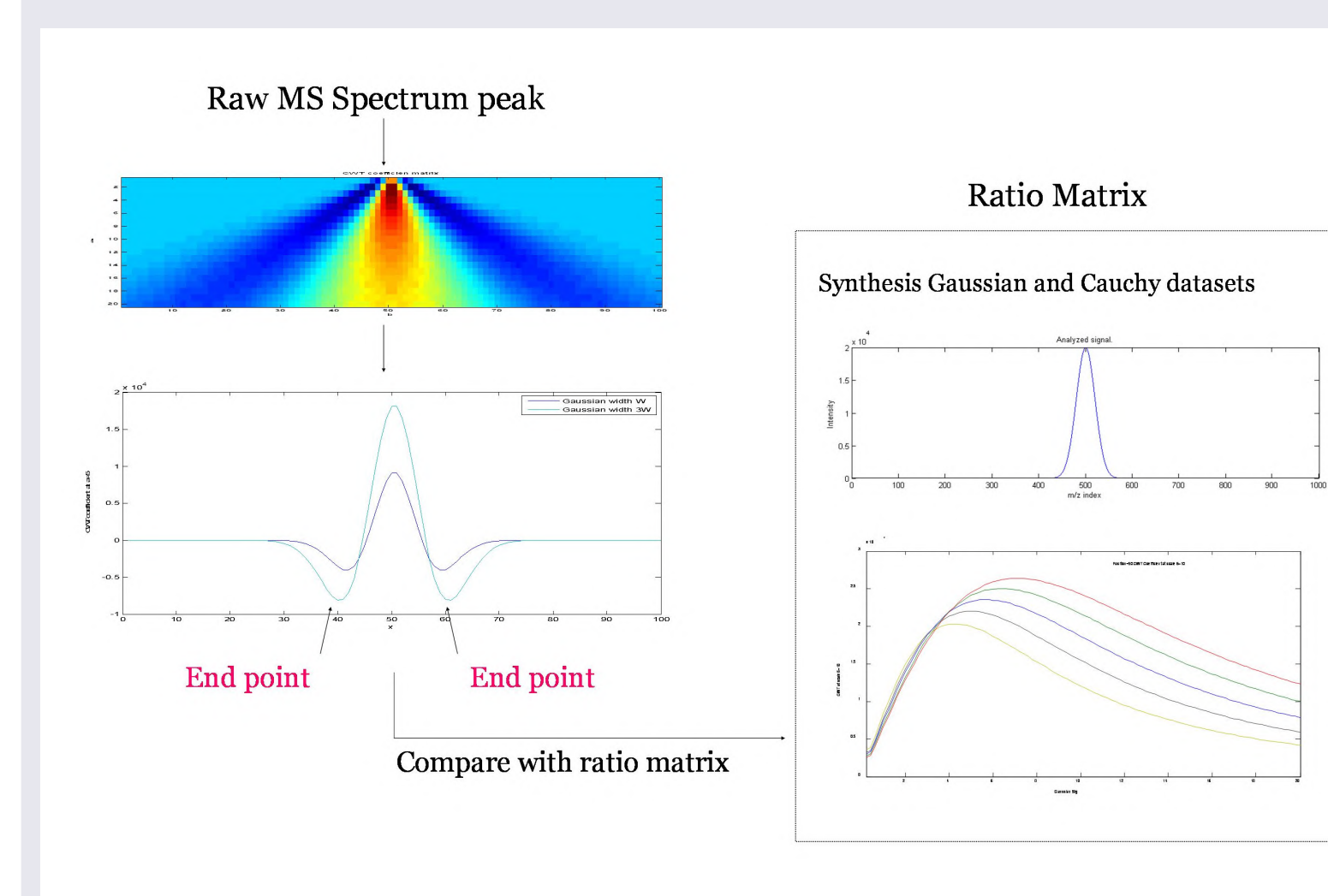1. Establish ratio matrix
2. End-Point detection
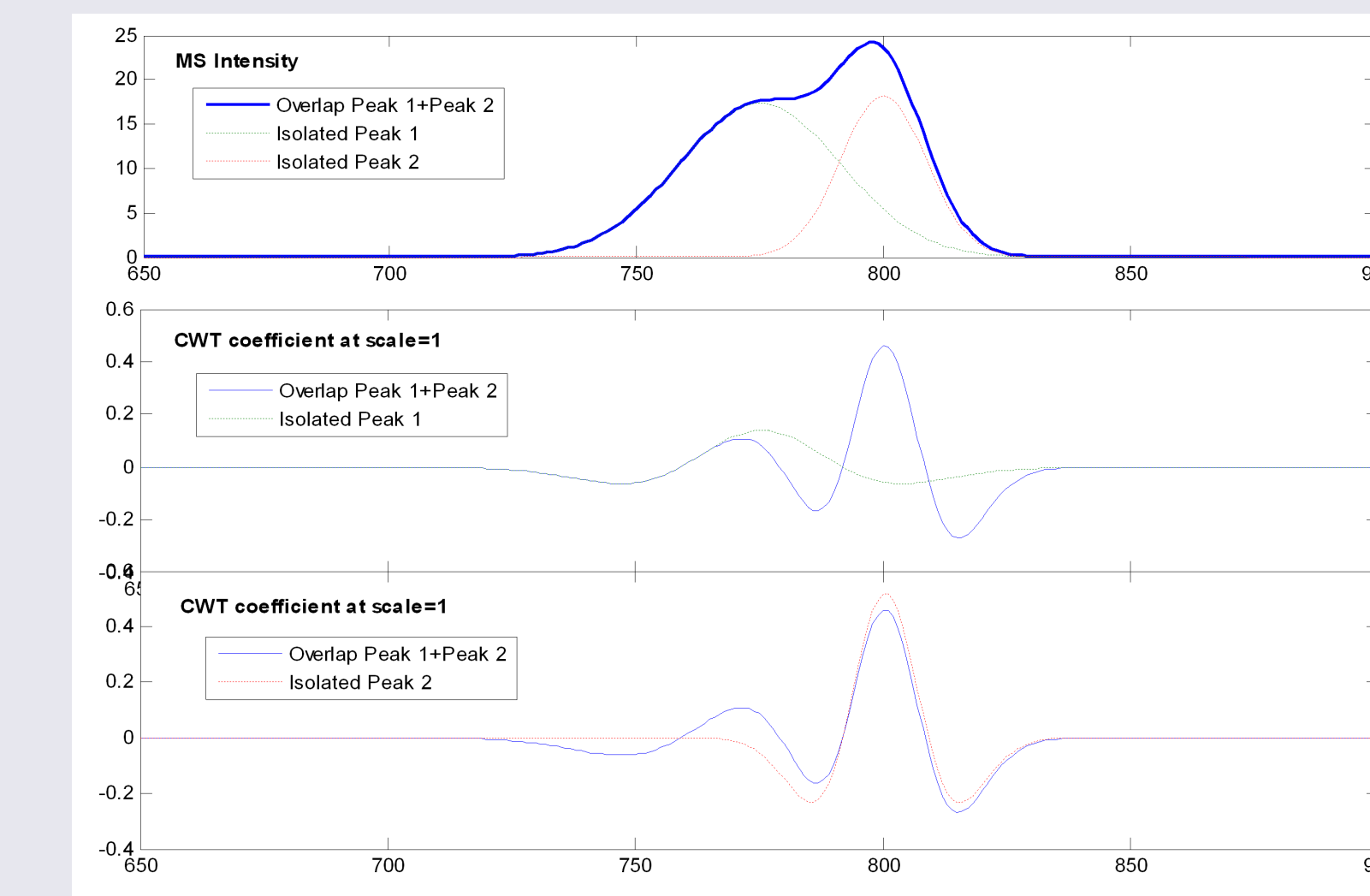3. Quantification

Methods

Multi-scale profile analysis

Section profiles of specific scale on CWT coefficient map

Peak detection

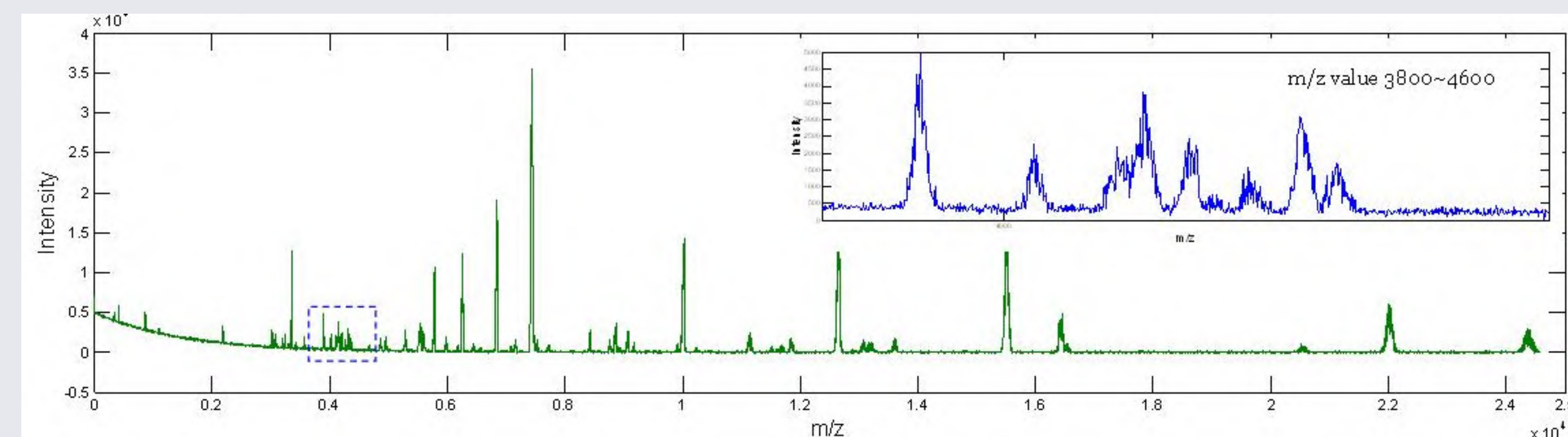### Peak quantification

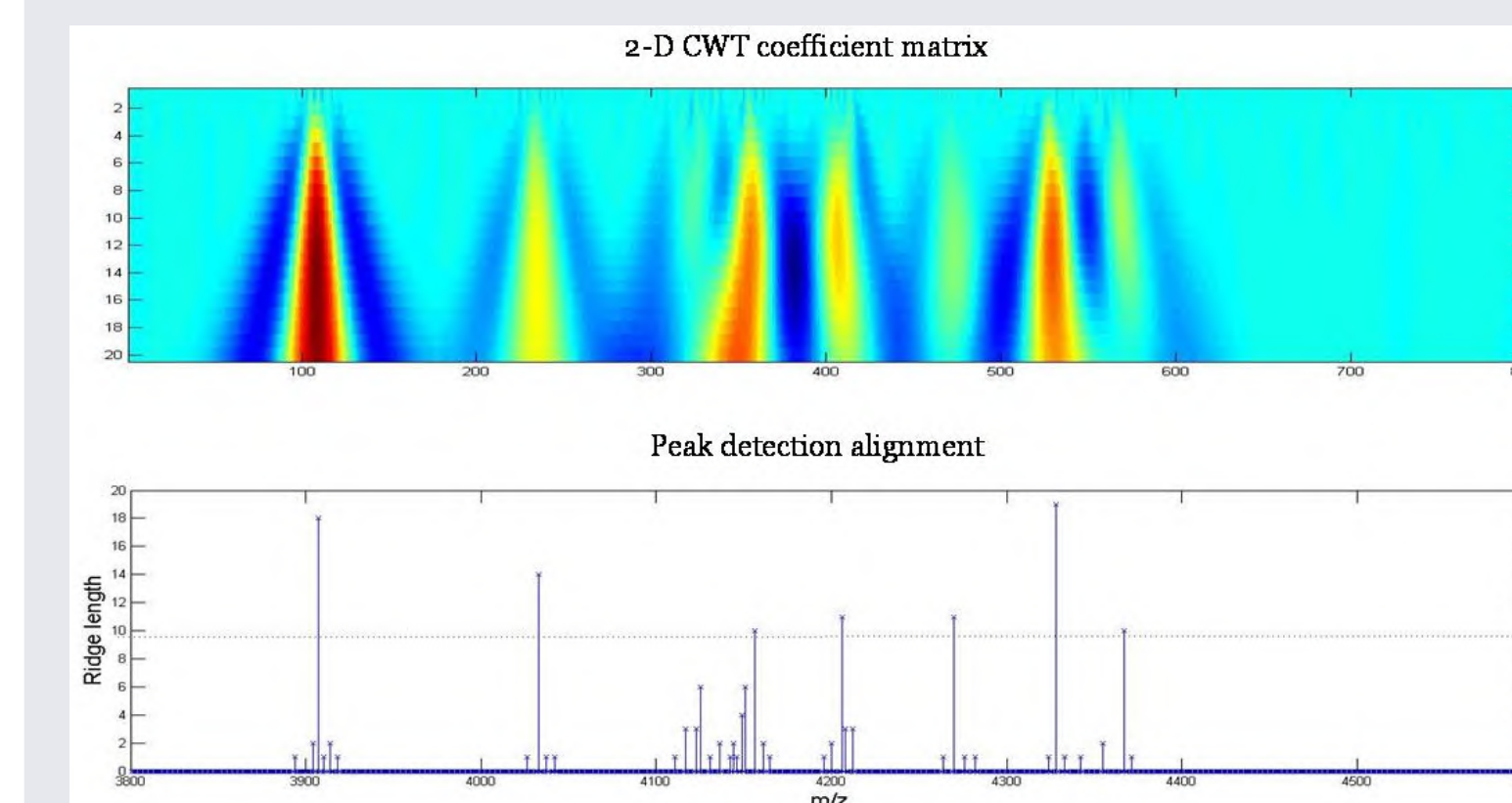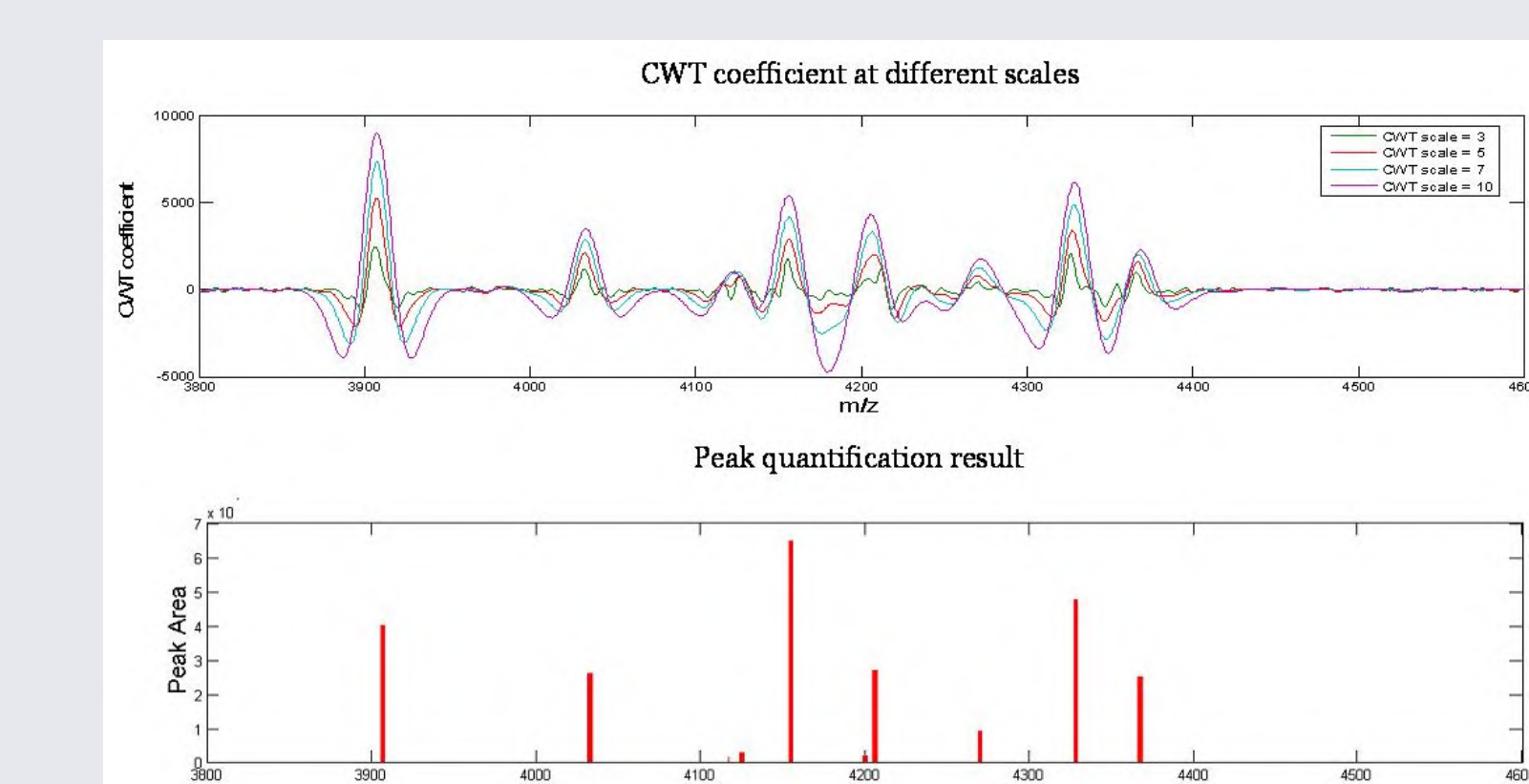### Resolution overlapping peaks

Results

Example: Simulated Proteomics Spectra made by Coombes et al.(2005)

Peak detection result

Peak quantification result

Reference

1. Du *Bioinformatics* **2006**; 22, 2059-2065.
2. Chao *BMC Bioinformatics* **2009**; 10(4).
3. Lange *Pac. Symp. Biocomput.* **2006**, 243-254.
4. Wee *Electrophoresis* **2008**; 29, 4215-4225.
5. Morris *Bioinformatics* **2005**; 21, 1764-1775.
6. Coombes *Proteomics* **2005**; 5(16), 4107-4117.