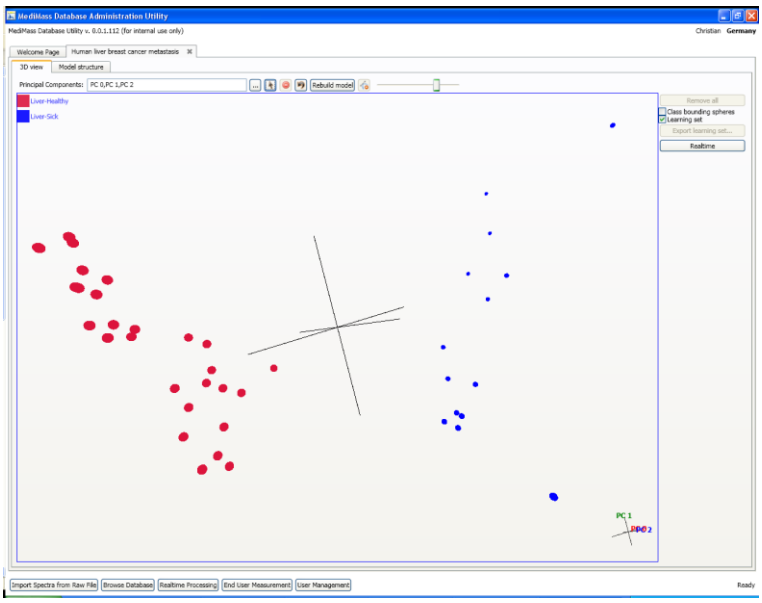
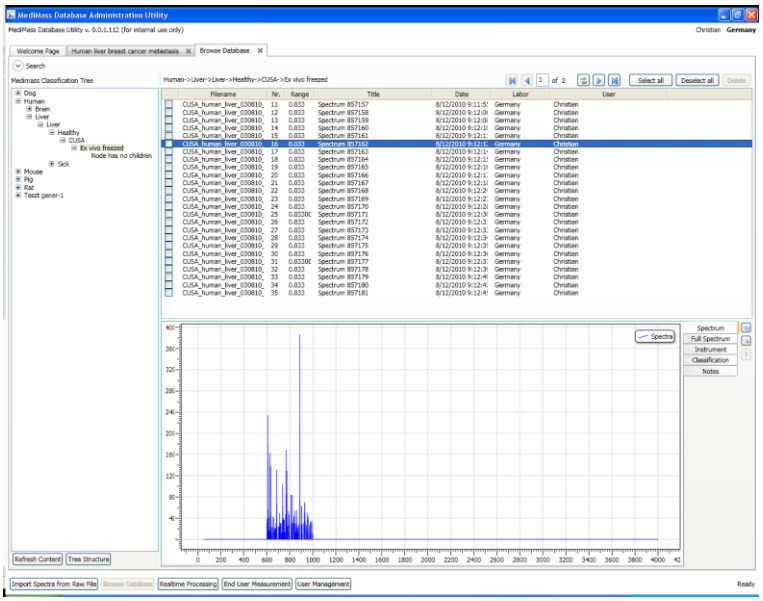
Chemometrics – data analysis method development for real-time tissue identification

Although the acquisition of real-time mass spectrometric data is fairly simple, the processing and analysis of spectra within 50 – 150 ms raises a number of challenges. This timeframe is defined by the surgical environment, since feedback delay should not exceed 1 s, and the physical processes of ion formation, transfer, mass analysis, detection, signal conversion and signal transmission leave ~ 100 ms for the data analysis and visual feedback.

Our present data analysis approach is based on comparison with entries in a spectral library, and classification of data point into the group with highest similarity. Raw mass spectra are noise-filtered, normalized, aligned, binned, and reduced to 60 dimensional data vectors via principal component analysis. Noise filtering is based on peak shape and also the presence or absence of isotope peaks. (Note, that all of the corresponding ionization methods produce singly charged ions under the given conditions.) Alignment of spectra is a simple, multi-point recalibration, which is critical due to subsequent binning of data. Bin size depends on analyzer technique utilized, e.g. in case of Orbitrap data we use 0.01 Da bins, which means that PCA reduces number of dimensions from 30,000 to 60. Spectral library entry groups (corresponding to individual histology groups) are separated further using linear discriminant analysis (LDA). Since all the PCA and LDA is performed prior to experiment, during the surgical intervention the data points are localized in the 60 dimensional LDA space and Mahalanobis distances are measured from the centers of tissue specific data groups. Data point is classified into the closest group, with a well-defined threshold value. In those cases where all distances exceed the threshold, the data point qualifies as “unknown”. The outlined data analysis algorithm gives good identification specificity (95-97%) which is close to that of histology. Our current research is focused on alternative, heuristic methods of data analysis, and also the identification of mixed spectra. This latter application is extremely critical in case of detection of tumor infiltrations.

Similar data analysis algorithm is being developed for the processing and identification of imaging data, using higher number of dimensions, due to fewer restrictions on analysis time.




MediMass Database Administration Utility
MediMass Database Utility v. 0.0.1.112 (for internal use only)
Christian Germany
Welcome Page
Human liver breast cancer metastasis
Browse Database
Search
Medimass Classification Tree
Dog
Human
Brain
Liver
Liver
Healthy
CUSA
Ex vivo freezed
Node has no children
Sick
Mouse
Pig
Rat
Testt gener-1
Refresh Content
Tree Structure
Human->Liver->Liver->Healthy->CUSA->Ex vivo freezed
Select all
Deselect all
Delete
Filename
Nr.
Range
Title
Date
Labor
User
Spectra
Spectrum
Full Spectrum
Instrument
Classification
Notes
Import Spectra from Raw File
Browse Database
Realtime Processing
End User Measurement
User Management
Ready