

# Vergleich von Merkmalen in Flüssigchromatographie- Massenspektrometrie-Daten

Kristian Bredies

4. November 2005

# Inhalt

- 1 Einleitung
- 2 Flüssigchromatographie-Massenspektrometrie
  - Das Verfahren
  - Merkmale der Daten
  - Aufgabe
- 3 Erste Lösungsansätze
  - Ein Algorithmus
  - Ausblick

# Hintergrund

Kontakt mit dem Pharma-Unternehmen *Roche* durch Manfred Nölte seit August/September 2005

- Untersuchung von Substanzen auf bestimmte Stoffwechselprodukte

# Hintergrund

Kontakt mit dem Pharma-Unternehmen *Roche* durch Manfred Nölte seit August/September 2005

- Untersuchung von Substanzen auf bestimmte Stoffwechselprodukte
- Vorhandene Software unzureichend
- Bessere Algorithmen nötig, um “schwache Signale”, zu erkennen

# Allgemeine Aufgabe

## Gegeben:

Ein Gemisch von Substanzen in jeweils einer Blankprobe und einer Inkubationsprobe zu verschiedenen Zeitpunkten.

# Allgemeine Aufgabe

## Gegeben:

Ein Gemisch von Substanzen in jeweils einer Blankprobe und einer Inkubationsprobe zu verschiedenen Zeitpunkten.

## Gesucht:

Die Substanzen, um die sich die Proben unterscheiden und deren quantitative Entwicklung in den jeweiligen Zeiten.

# Das Verfahren

Die untersuchten Substanzen werden mit einem  
*Flüssigchromatographie-Massenspektrometer (LC-MS)* analysiert.

Blankprobe

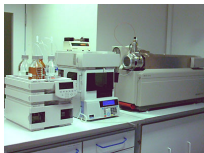
Inkubationsprobe

# Das Verfahren

Die untersuchten Substanzen werden mit einem  
*Flüssigchromatographie-Massenspektrometer (LC-MS)* analysiert.

Blankprobe

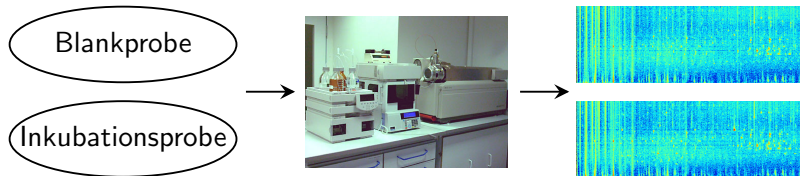
Inkubationsprobe





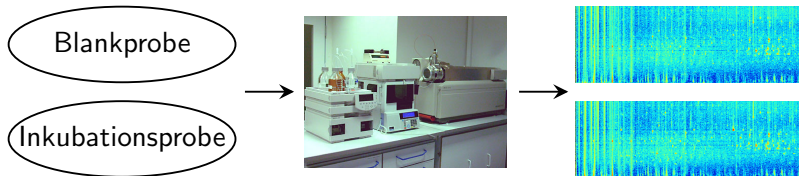
# Das Verfahren

Die untersuchten Substanzen werden mit einem  
*Flüssigchromatographie-Massenspektrometer (LC-MS)* analysiert.



# Das Verfahren

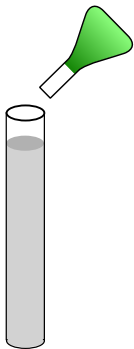
Die untersuchten Substanzen werden mit einem  
*Flüssigchromatographie-Massenspektrometer (LC-MS)* analysiert.



In diesen Geräten wird Chromatographie und Massenspektrometrie  
kombiniert.

# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



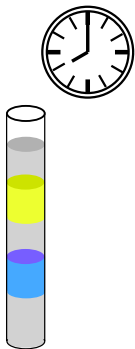
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



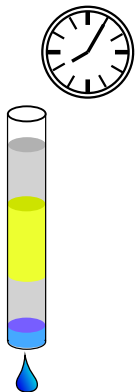
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



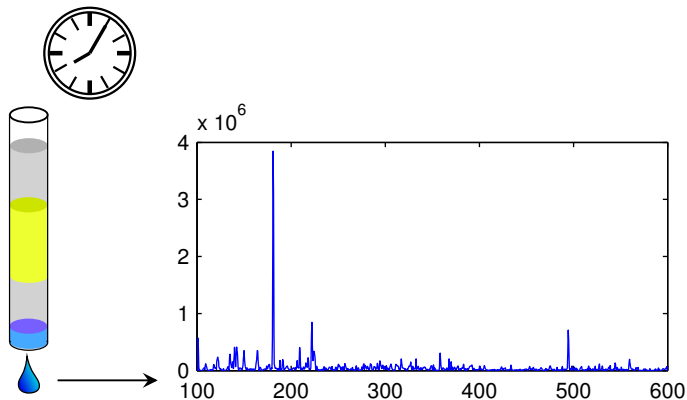
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



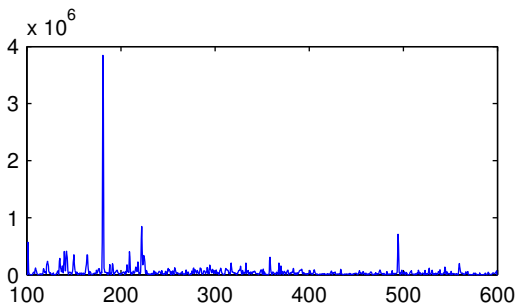
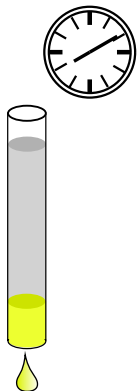
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



# Flüssigchromatographie-Massenspektrometrie

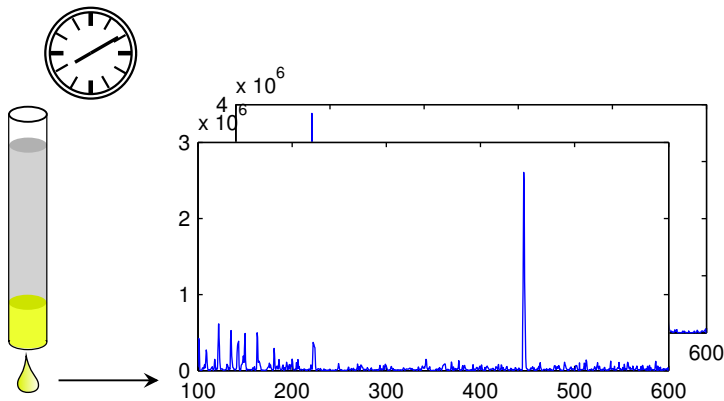
Schematische Funktionsweise:





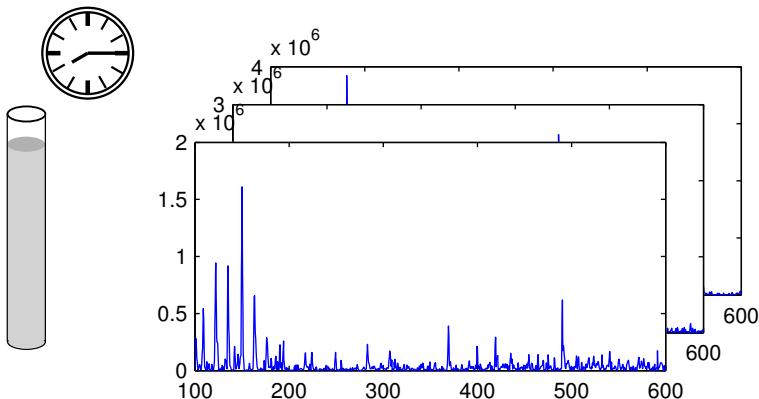
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



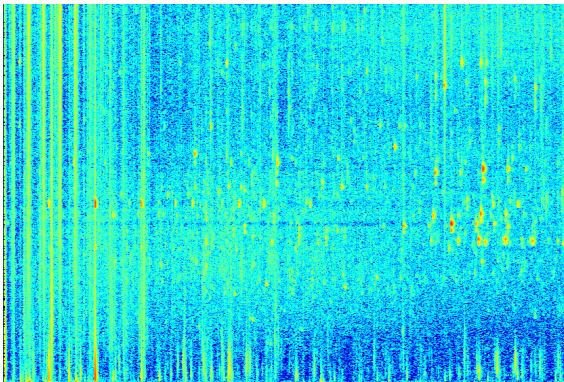
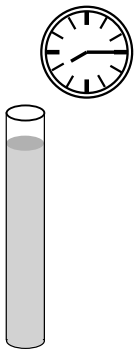
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



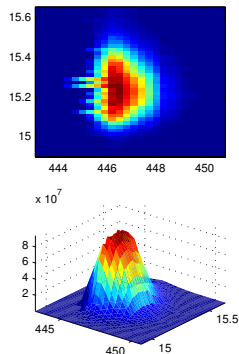
# Flüssigchromatographie-Massenspektrometrie

Schematische Funktionsweise:



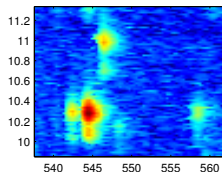
# Merkmale der Daten

- Interessant sind Peaks, die exponentiell in der Masse und wie eine Gaußkurve in der Zeit verlaufen.



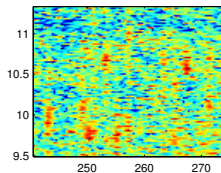
# Merkmale der Daten

- Interessant sind Peaks, die exponentiell in der Masse und wie eine Gaußkurve in der Zeit verlaufen.
- Die Peaks können sich überlappen.



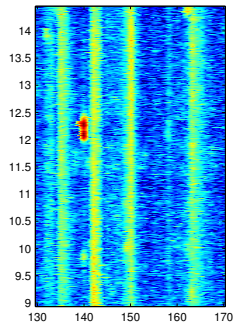
# Merkmale der Daten

- Interessant sind Peaks, die exponentiell in der Masse und wie eine Gaußkurve in der Zeit verlaufen.
- Die Peaks können sich überlappen.
- Die Datensätze sind teilweise sehr stark verrauscht, etwa in der Größenordnung der kleinsten interessanten Merkmale.



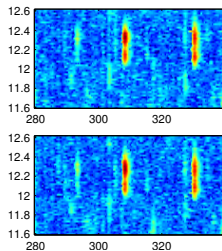
# Merkmale der Daten

- Die LC-MS-Daten sind durchsetzt von unerwünschten Dauerpeaks in zeitlicher Richtung.



# Merkmale der Daten

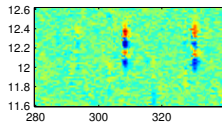
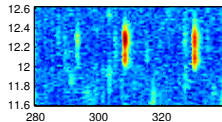
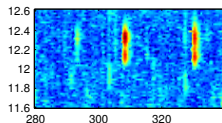
- Die LC-MS-Daten sind durchsetzt von unerwünschten Dauerpeaks in zeitlicher Richtung.
- Verschiedene Datensätze können zueinander zeitlich verzerrt und verschoben sein.





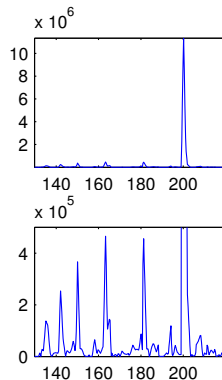
# Merkmale der Daten

- Die LC-MS-Daten sind durchsetzt von unerwünschten Dauerpeaks in zeitlicher Richtung.
- Verschiedene Datensätze können zueinander zeitlich verzerrt und verschoben sein.



# Merkmale der Daten

- Die LC-MS-Daten sind durchsetzt von unerwünschten Dauerpeaks in zeitlicher Richtung.
- Verschiedene Datensätze können zueinander zeitlich verzerrt und verschoben sein.
- Die Peakhöhen schwanken zwischen  $10^5$  und  $10^8$ .



# Aufgabe der Bildverarbeitung

## Gegeben:

Zu unterschiedlichen Zeitpunkten jeweils ein Paar  
LC-MS-Datensätze.

# Aufgabe der Bildverarbeitung

## Gegeben:

Zu unterschiedlichen Zeitpunkten jeweils ein Paar LC-MS-Datensätze.

## Gesucht:

- 1 Die Position und Höhe der interessanten Peaks in den jeweiligen Datensätzen.
- 2 Eine Zuordnung der gleichen Peaks in einem Datenpaar.
- 3 Eine Quantifizierung der zeitlichen Entwicklung der unterschiedlichen Peaks.

# Erste Lösungsansätze

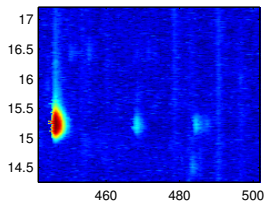
Mögliches Verfahren zur Erkennung der Peaks in einem Datensatz:

## Algorithmus

- 1 Entfernen von Rauschen.
- 2 Erkennen und Entfernen der Dauerpeaks in Zeitrichtung (vertikale Streifen).
- 3 Detektion von lokalen Maxima.

# Details zum Algorithmus

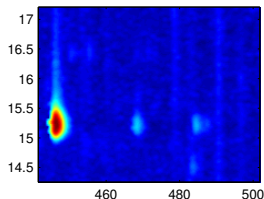
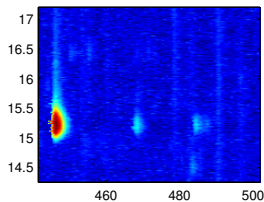
Entauschen



# Details zum Algorithmus

## Entrauschen

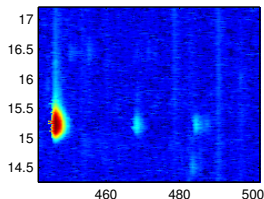
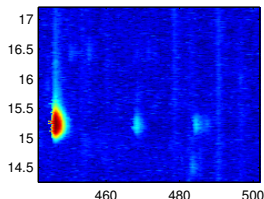
- Faltungs-Filtern entfernt kleine Peaks zu schnell.



# Details zum Algorithmus

## Entrauschen

- Faltungs-Filtern entfernt kleine Peaks zu schnell.
- Wavelet-Hardthresholding mit z.B.  $\text{db}2$  erhält kleine Peaks besser.
- Artefakte können durch Abschneiden negativer Werte verringert werden.

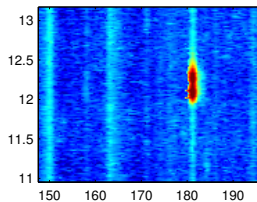




# Details zum Algorithmus

## Erkennen der Basislinie

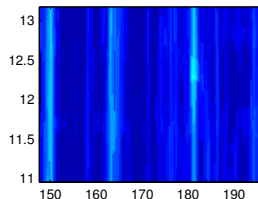
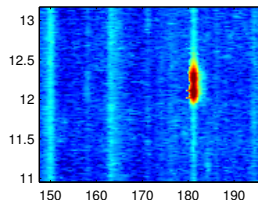
- Erosion mit vertikalem Strukturelement.



# Details zum Algorithmus

## Erkennen der Basislinie

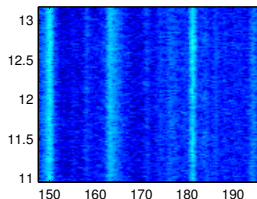
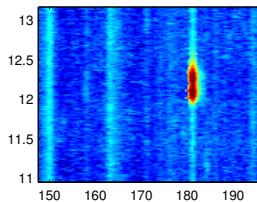
- Erosion mit vertikalem Strukturelement.
- Bei voller vertikaler Linie entstehen eigene unerwünschte Strukturen.



# Details zum Algorithmus

## Erkennen der Basislinie

- Erosion mit vertikalem Strukturelement.
- Bei voller vertikaler Linie entstehen eigene unerwünschte Strukturen.
- Drei-Punkt-Element bietet akzeptablen Kompromiss.



# Details zum Algorithmus

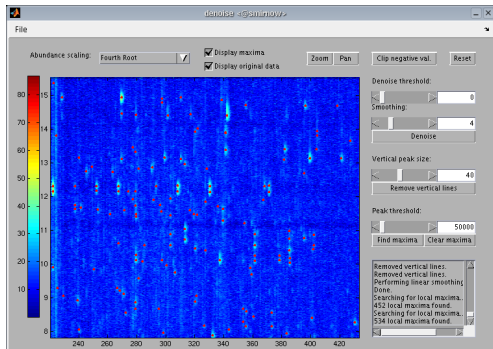
## Detektieren der Peaks

- Betrachte nur Peaks über einem gewissen Schwellwert.

# Details zum Algorithmus

## Detektieren der Peaks

- Betrachte nur Peaks über einem gewissen Schwellwert.
- Finde Peaks durch lokale Maxima.



# Ausblick

Folgende Schritte sind noch nötig:

- Exakte Bestimmung der Position und Höhe eines Peaks.
- Korrektur der Zeitverzerrung/-verschiebung zwischen zwei Datensätzen.
- Betrachtung und Vergleich mehrerer Datenpaare.