



(10) **DE 10 2014 224 916 A1** 2016.06.09

(12)

## Offenlegungsschrift

(21) Aktenzeichen: **10 2014 224 916.3**

(22) Anmeldetag: **04.12.2014**

(43) Offenlegungstag: **09.06.2016**

(51) Int Cl.: **G01N 27/62** (2006.01)

**H01J 49/02** (2006.01)

(71) Anmelder:

**Universität Bremen, 28359 Bremen, DE**

(74) Vertreter:

**Fink Numrich Patentanwälte, 80634 München, DE**

(72) Erfinder:

**Maaß, Peter, Prof. Dr., 28211 Bremen, DE; Trede,  
Dennis, Dr., 28359 Bremen, DE; Aleksandrov,  
Fedor, Dr., 28209 Bremen, DE; Oetjen, Janina, Dr.,  
28865 Lilienthal, DE; Hauberg-Lotte, Lena, Dr.,  
28201 Bremen, DE**

(56) Ermittelter Stand der Technik:

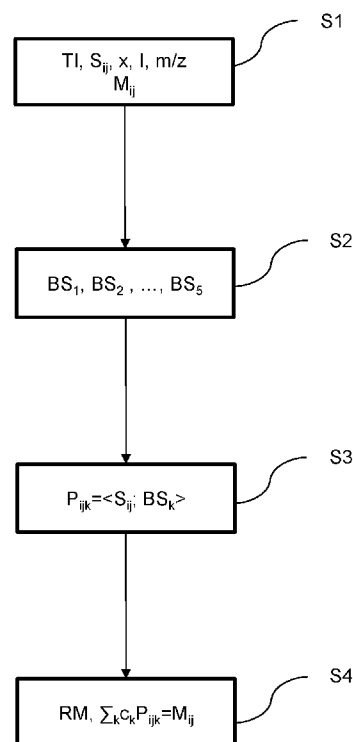
|           |                         |           |
|-----------|-------------------------|-----------|
| <b>DE</b> | <b>10 2013 207 402</b>  | <b>A1</b> |
| <b>US</b> | <b>2009 / 0 154 781</b> | <b>A1</b> |
| <b>US</b> | <b>2013 / 0 060 550</b> | <b>A1</b> |
| <b>WO</b> | <b>02/ 010 740</b>      | <b>A2</b> |
| <b>WO</b> | <b>2008/ 110 581</b>    | <b>A2</b> |
| <b>WO</b> | <b>2013/ 086 516</b>    | <b>A1</b> |
| <b>WO</b> | <b>2013/ 096 845</b>    | <b>A2</b> |

Prüfungsantrag gemäß § 44 PatG ist gestellt.

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

(54) Bezeichnung: **Verfahren zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte des menschlichen oder tierischen Körpers**

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte (TI) des menschlichen oder tierischen Körpers zur Präparation eines digitalen Färbetests, wobei in dem digitalen Färbetest rechnergestützt in einem getesteten Gewebeschnitt (TI') Gewebebereiche bestimmt werden, die ein vorbestimmtes Krankheitsbild aufweisen. Im Rahmen des Verfahrens werden räumlich aufgelöste Hyperspektral-daten und insbesondere Massenspektrometriedaten für eine Vielzahl von Messpositionen (x) verarbeitet. Dabei werden aus den Hyperspektral-daten mittels einer Methode zur Bestimmung von Basisvektoren mehrere Basisspektren (BS<sub>1</sub>, BS<sub>2</sub>, ..., BS<sub>5</sub>) bestimmt, die mit den Hyperspektral-daten korreliert werden, wodurch für jede Messposition (x) eine Mehrzahl von Korrelationswerten (P<sub>ijk</sub>) für jeweilige Basisspektren (BS<sub>1</sub>, BS<sub>2</sub>, ..., BS<sub>5</sub>) erhalten wird. Anschließend wird mittels eines Regressionsverfahrens ein Regressionsmodell (RM) bestimmt, welches eine Maske (M<sub>ij</sub>) des oder der Gewebeschnitte, in der krankhafte Gewebebereiche markiert sind, basierend auf den ermittelten Korrelationswerten beschreibt. Das Regressionsmodell (RM) und die Basisspektren (BS<sub>1</sub>, BS<sub>2</sub>, ..., BS<sub>5</sub>), die im Rahmen des Verfahrens gewonnen werden, können anschließend zur Durchführung des digitalen Färbetests genutzt werden.



## Beschreibung

**[0001]** Die Erfindung betrifft ein Verfahren zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte des menschlichen oder tierischen Körpers sowie ein Verfahren zur rechnergestützten Durchführung eines digitalen Färbetests.

**[0002]** In der klassischen Pathologie werden Gewebeproben mit dem Ziel der Diagnostik bzw. Prognose bestimmter medizinischer Indikationen analysiert. Dazu werden mittels biochemischer Verfahren Gewebeschnitte bearbeitet, wie z.B. mittels HE-Färbung (Hämatoxylin-Eosin-Färbung), immunhistochemischen Färbetests (auch als IHC-Tests abgekürzt) oder molekularen Testverfahren. Mit diesen Verfahren wird krankhaftes Gewebe in dem Gewebeschnitt markiert, so dass die visuelle Inspektion und die Diagnostik unterstützt werden. Beispielsweise werden im Rahmen von IHC-Tests Antikörper zur visuellen Markierung von Biomarkern, die für bestimmte Krankheitsbilder spezifisch sind, auf einem Gewebeschnitt aufgebracht.

**[0003]** Aus dem Stand der Technik sind ferner massenspektrometrische Methoden zur Identifikation von veränderten Gewebestrukturen bekannt. Häufig werden hierzu bildgebende Verfahren der Massenspektrometrie, z.B. DESI oder MALDI, eingesetzt. Dabei wird für jeden Bildpunkt eines Gewebeschnitts ein volles Spektrum von Masse-zu-Ladungs-Verhältnissen im Bereich von wenigen 100 bis ca. 10 kDa erfasst. Man erhält somit ein Massenspektrum, das jeweils eine räumliche Koordinate und eine spektrale Masse-zu-Ladungs-Koordinate aufweist. Aus diesen massenspektrometrischen Daten lassen sich bis dato die darin enthaltenen Informationen nicht derart extrahieren, dass die massenspektrometrische Analyse herkömmliche histologische Färbetests ersetzen kann.

**[0004]** Aufgabe der Erfindung ist es, die aus einem oder mehreren Gewebeschnitten gewonnenen Hyperspektraldaten derart zu analysieren, dass diese Daten zur Durchführung von digitalen bzw. rechnergestützten Färbetests genutzt werden können. Ferner ist es Aufgabe der Erfindung ein Verfahren zur Durchführung eines digitalen Färbetests zu schaffen.

**[0005]** Diese Aufgabe wird durch das Verfahren gemäß Patentanspruch 1 bzw. Patentanspruch 10 bzw. die Vorrichtung gemäß Patentanspruch 13 bzw. Patentanspruch 14 gelöst. Weiterbildungen der Erfindung sind in den abhängigen Ansprüchen definiert.

**[0006]** Das erfindungsgemäße Verfahren dient zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte des menschlichen oder tierischen Körpers. Ein jeweiliger Gewebeschnitt wurde mittels Biopsie aus dem Körper entnommen und wird außer-

halb des Körpers bearbeitet, ohne dass das entnommene Gewebe dem Körper zurückgeführt wird. Der Begriff des Gewebeschnitts ist hier und im Folgenden weit zu verstehen. Ein Gewebeschnitt kann insbesondere auch ein Tissue-Microarray darstellen, bei dem eine Vielzahl von Gewebeproben (sog. Tissue-Cores), die vorzugsweise von unterschiedlichen Patienten stammen, auf einem Probenträger aufgebracht sind. Mittels der erfindungsgemäßen Analyse wird ein digitaler Färbetest präpariert bzw. vorbereitet, bei dem rechnergestützt in einem getesteten Gewebeschnitt Gewebebereiche bestimmt werden, die ein vorbestimmtes Krankheitsbild aufweisen.

**[0007]** Im Folgenden werden die einzelnen Merkmale des Verfahrens anhand der Schritte a) bis e) erläutert. Diese Schritte dienen lediglich zur besseren Referenzierung der Merkmale des Verfahrens und mit ihnen ist nicht zwangsläufig eine zeitliche Reihenfolge der Durchführung der Schritte korreliert.

**[0008]** In einem Schritt a) werden räumlich aufgelöste (digitale) Hyperspektraldaten von einem oder mehreren vorbestimmten zu analysierenden Gewebeschnitten eingelesen, wobei der oder die Gewebeschnitte Gewebebereiche mit einem vorbestimmten Krankheitsbild aufweisen. Die Hyperspektraldaten enthalten für eine Vielzahl von räumlichen (insbesondere zweidimensionalen) Messpositionen in dem oder den vorbestimmten Gewebeschnitten jeweilige Spektren, wobei ein jeweiliges Spektrum die Zusammensetzung des oder der vorbestimmten Gewebeschnitte an einer Messposition in der Form eines Vektors aus Intensitätswerten für eine Vielzahl von Kanälen repräsentiert, wobei ein Intensitätswert eine für den entsprechenden Kanal detektierte physikalische Größe ist.

**[0009]** Im Falle der Analyse mehrerer vorbestimmter Gewebeschnitte sind die Messpositionen über die Gewebeschnitte hinweg unterscheidbar. Mit anderen Worten entspricht in diesem Fall die Gesamtanzahl der Messpositionen der Summe von allen in den Gewebeschnitten enthaltenen Messpositionen, selbst wenn Messpositionen in unterschiedlichen Gewebeschnitten die gleichen räumlichen Koordinaten aufweisen. Zur Unterscheidung von Messpositionen aus unterschiedlichen Gewebeschnitten kann ein Index verwendet werden, welcher den jeweiligen Gewebeschnitt spezifiziert. Die Gewebeschnitte stammen vorzugsweise von unterschiedlichen Patienten (d.h. unterschiedlichen Menschen oder Tieren). In einer bevorzugten Variante der Erfindung werden mit dem erfindungsgemäßen Verfahren zwischen 50 und 100 Gewebeschnitte analysiert. Je größer die Anzahl der Gewebeschnitte ist, desto verlässlicher ist der digitale Färbetest, der basierend auf den Ergebnissen des Analyseverfahrens durchgeführt wird.

**[0010]** Die obigen digitalen Hyperspektraldaten können vorab erfasst worden sein und in einem digitalen Speicher hinterlegt sein, aus dem sie in Schritt a) ausgelesen werden. Nichtsdestotrotz kann das erfindungsgemäße Verfahren auch den Schritt der Erfassung der Hyperspektraldaten mittels eines geeigneten Spektrometers beinhalten.

**[0011]** Vorzugsweise handelt es sich bei den Hyperspektraldaten um Massenspektrometriedaten, insbesondere sog. MALDI-Massenspektrometriedaten (MALDI = Matrix-Assisted Laser Desorption/Ionization). Bei Massenspektrometriedaten sind die Kanäle Masse-zu-Ladungs-Verhältnisse und die für den entsprechenden Kanal detektierte physikalische Größe repräsentiert die Anzahl der für das entsprechende Masse-zu-Ladungs-Verhältnis detektierten Moleküle. Nichtsdestotrotz können die Hyperspektraldaten ggf. auch anderen Daten darstellen, wie z.B. aus der Raman-Spektroskopie gewonnene Daten. Im Falle der Raman-Spektroskopie stellen die Kanäle Frequenzen von an der Gewebeprobe reflektiertem bzw. gestreutem Licht dar und die physikalische Größe ist die Intensität dieses reflektierten bzw. gestreuten Lichts.

**[0012]** In einem Schritt b) des erfindungsgemäßen Verfahrens wird eine (digitale) Maske des oder der vorbestimmten Gewebeschnitte eingelesen, welche für die Vielzahl von Messpositionen jeweilige Skalenwerte enthält, wobei Bereiche des oder der vorbestimmten Gewebeschnitte mit gegenüber dem restlichen Bereich des oder der Gewebeschnitte abweichenden (z.B. erhöhten) Skalenwerten Gewebebereiche mit dem vorbestimmten Krankheitsbild sind. Die digitale Maske kann aus einem herkömmlichen histologischen Färbetest und insbesondere einem immunhistochemischen Färbetest des oder der zu analysierenden Gewebeschnitte gewonnen sein. Im Falle der Analyse von mehreren Gewebeschnitten umfasst die Maske jeweilige Teilmasken für jeden Gewebeschnitt. Die Messpositionen in den Teilmasken korrelieren mit entsprechenden Messpositionen in den jeweiligen Gewebeschnitten. Wie oben erwähnt, sind die Messpositionen dabei über die Gewebeschnitte und damit auch über die Teilmasken hinweg unterscheidbar. Die digitale Maske ist in der Regel vorab in einem digitalen Speicher hinterlegt, aus dem die Daten in der Form der entsprechenden Skalenwerte ausgelesen werden. Je nach Ausgestaltung können die Skalenwerte aus einem diskreten oder kontinuierlichen Wertebereich stammen. Gegebenenfalls können die Skalenwerte auch nur die Werte 0 und 1 annehmen, wobei der Wert 0 für einen Gewebebereich ohne Krankheitsbild und der Wert 1 für einen Gewebebereich mit Krankheitsbild oder umgekehrt steht.

**[0013]** In einem Schritt c) des erfindungsgemäßen Verfahrens werden mittels einer Methode zur Bestimmung von Basisvektoren mehrere Basisspektren im

Raum der Vektoren der Hyperspektraldaten oder in einem Raum von abgeleiteten Vektoren ermittelt. Dabei werden die abgeleiteten Vektoren aus den Hyperspektraldaten für jede Messposition berechnet. Spezielle Varianten der Berechnung der abgeleiteten Vektoren werden weiter unten bzw. in Anspruch 3 beschrieben. Die Basisspektren stellen die Basisvektoren dar. Die Anzahl der Basisspektren liegt vorzugsweise zwischen 5 und 12. Unter einer Methode zur Bestimmung von Basisvektoren ist ein Verfahren zu verstehen, welches Basisvektoren derart bestimmt, dass die Vektoren der Hyperspektraldaten bzw. die abgeleiteten Vektoren durch Kombinationen (insbesondere Linearkombinationen) der Basisvektoren darstellbar sind. Bevorzugte Varianten von Methoden zur Bestimmung von Basisvektoren werden weiter unten genannt.

**[0014]** In einem Schritt d) des erfindungsgemäßen Verfahrens werden die Basisspektren mit den Vektoren der Hyperspektraldaten oder den abgeleiteten Vektoren an jeder Messposition korreliert, wodurch für jede Messposition eine Mehrzahl von Korrelationswerten für jeweilige Basisspektren erhalten wird. Es existiert somit für jede Messposition und jedes Basisspektrum ein Korrelationswert. Verfahren zur Korrelation von Vektoren sind hinlänglich bekannt. Vorzugsweise ist der Korrelationswert das Skalarprodukt zwischen normierten Vektoren, d.h. zwischen Vektoren, die jeweils mit dem Kehrwert ihres Betrags multipliziert werden. Der Korrelationswert liegt in diesem Fall zwischen  $-1$  und  $1$ . Gegebenenfalls kann der Korrelationswert auch der Betrag des Skalarprodukts sein. In diesem Fall liegt der Korrelationswert zwischen  $0$  und  $1$  und kann als Ähnlichkeitsmaß zwischen den Vektoren interpretiert werden, wobei die Vektoren umso ähnlicher sind, je größer der Korrelationswert ist.

**[0015]** In einem Schritt e) wird schließlich mittels eines Regressionsverfahrens ein Regressionsmodell (d.h. entsprechende Parameter des Regressionsmodells) bestimmt. Das Regressionsmodell beschreibt die Maske basierend auf der Mehrzahl von Korrelationswerten für die jeweiligen Messpositionen. Das Regressionsmodell mit seinen Parametern und die Basisvektoren werden schließlich für die spätere Durchführung des digitalen Färbetests digital gespeichert.

**[0016]** Mit den Ergebnissen des erfindungsgemäßen Verfahrens in der Form der Basisvektoren und des Regressionsmodells lassen sich auf einfache Weise rechnergestützte Färbetests von neuen Gewebeschnitten durchführen, um zu ermitteln, ob in diesen Gewebeschnitten Gewebebereiche mit dem vorbestimmten Krankheitsbild enthalten sind. Der Begriff des Färbetests ist weit zu verstehen. Insbesondere muss der Färbetest Gewebebereiche mit dem vorbestimmten Krankheitsbild nicht zwangsläufig

fig farblich codieren, sondern die Codierung kann allgemein über die oben beschriebenen Skalenwerte erfolgen.

**[0017]** Wie im obigen Schritt c) beschrieben, kann in einer Variante des erfindungsgemäßen Verfahrens die Methode zur Bestimmung von Basisvektoren direkt im Raum der Vektoren der Hyperspektraldaten angewendet werden. In diesem Fall können auf einfache Weise ohne spezifische Informationen über die in krankhaften Gewebereichen vorhandenen Biomarker Ergebnisse gewonnen werden, welche die Durchführung von digitalen Färbetests ermöglichen. Nichtsdestotrotz können die Basispektren ggf. auch in einem Raum von abgeleiteten Vektoren bestimmt werden, die aus den Hyperspektraldaten für jede Messposition ermittelt bzw. berechnet werden. Vorzugsweise werden hierzu Informationen über die im krankhaften Gewebe vorhandenen Molekülzusammensetzungen genutzt, wie dies in der nachfolgend erläuterten Variante des erfindungsgemäßen Verfahrens der Fall ist.

**[0018]** In einer speziellen Variante des erfindungsgemäßen Verfahrens, bei der in Schritt c) die Basispektren in einem Raum von abgeleiteten Vektoren bestimmt werden, wird ein abgeleiteter Vektor für jede Messposition wie nachfolgend erläutert ermittelt.

**[0019]** Es sind für das vorbestimmte Krankheitsbild (theoretisch oder experimentell bestimmte) Hyperspektraldaten für eine Mehrzahl von Molekülzusammensetzungen vorgegeben. Eine jeweilige Molekülzusammensetzung stellt in einer bevorzugten Ausführungsform ein vorgegebenes Peptid mit seinen Isotopen dar. Im Falle von Massenspektrometriedaten werden die Hyperspektraldaten eines Peptids durch das Massenspektrum des Peptids für die unterschiedlichen Masse-zu-Ladungs-Verhältnisse seiner Isotope repräsentiert. Gegebenenfalls kann eine Molekülzusammensetzung auch mehrere Peptide bzw. andere Moleküle betreffen, die durch Hyperspektraldaten und insbesondere durch ein entsprechendes Massenspektrum dargestellt werden. Die Molekülzusammensetzungen und deren Hyperspektraldaten wurden vorab mit geeigneten Verfahren bestimmt, wobei diese Verfahren nicht Bestandteil der Erfindung sind und deshalb nicht näher erläutert werden.

**[0020]** Die Hyperspektraldaten für eine Molekülzusammensetzung enthalten allgemein einen Vektor in der Form eines für die Molekülzusammensetzung charakteristischen Spektrums aus Intensitätswerten für eine vorgegebene Anzahl von Kanälen aus der Vielzahl von Kanälen. Die Anzahl der Kanäle stellt in der Regel einen Teil der Kanäle der im obigen Schritt a) eingelesenen Hyperspektraldaten dar.

**[0021]** Die für die Molekülzusammensetzungen charakteristischen Spektren werden in einem nächsten Schritt mit den Spektren der eingelesenen Hyperspektraldaten an den jeweiligen Messpositionen korreliert, wobei die Korrelation, analog wie oben beschrieben, über das Skalarprodukt der normierten Vektoren erfolgen kann. Bei der Korrelation werden nicht alle Kanäle, sondern nur die vorgegebene Anzahl von Kanälen berücksichtigt. Als Ergebnis der Korrelation wird als abgeleiteter Vektor für jede Messposition ein Vektor aus mehreren Korrelationskoeffizienten ermittelt, wobei für eine jeweilige Messposition ein Korrelationskoeffizient für jede Molekülzusammensetzung vorliegt. Vorzugsweise existieren für eine jeweilige Messposition zwischen 10 und 200 Korrelationskoeffizienten.

**[0022]** In der soeben beschriebenen Variante des erfindungsgemäßen Verfahrens werden somit aus den ursprünglichen Hyperspektraldaten Vektoren aus Korrelationskoeffizienten gewonnen, welche die Korrelation mit vorab bekannten Molekülzusammensetzungen beschreiben. Anstatt der Hyperspektraldaten werden dann diese abgeleiteten Vektoren in den Schritten c) bis e) weiterverarbeitet.

**[0023]** In einer besonders bevorzugten Variante sind der oder die zu analysierenden Gewebeschnitte jeweilige Gewebeschnitte mit Verdau. Der Begriff des Verdau ist aus der Histologie hinlänglich bekannt. Dabei wird der Gewebeschnitt über biochemische Verfahrensschritte vorverarbeitet. Insbesondere werden im Rahmen des Verdau größere Moleküle durch Enzyme in Peptide zerlegt.

**[0024]** Als Methode zur Bestimmung von Basisvektoren können in Schritt c) aus dem Stand der Technik bekannte Verfahren eingesetzt werden. In einer bevorzugten Ausführungsform wird eine nicht-negative Matrixfaktorisierung verwendet. Ebenso kann eine PCA-Analyse (PCA = Principal Component Analysis, deutsch: Hauptkomponentenanalyse) eingesetzt werden. In einer besonders bevorzugten Variante wird zur Bestimmung der Basisvektoren eine nicht-negative Matrixfaktorisierung mit der Nebenbedingung der spärlichen Besetzung der durch die nicht-negative Matrixfaktorisierung bestimmten Matrizen verwendet. Insbesondere wird die in den Druckschriften [1] und [2] beschriebene Matrixfaktorisierung genutzt. Die gesamte Offenbarung dieser Druckschriften [1] und [2] wird durch Verweis zum Inhalt der vorliegenden Anmeldung gemacht.

**[0025]** Auch in Schritt d) können an sich bekannte Regressionsverfahren zur Bestimmung des Regressionsmodells verwendet werden. Zum Beispiel kann ein lineares Regressionsverfahren zum Einsatz kommen, welches entsprechende Gewichtungsfaktoren zur Linearkombination der Mehrzahl von Korrelationswerten für die jeweiligen Messpositionen festlegt.

Anstatt eines linearen Regressionsverfahrens wird in einer weiteren Variante der Erfindung eine an sich bekannte SVM-Regression (SVM = Support Vector Machine) verwendet.

**[0026]** Das vorbestimmte Krankheitsbild, welches im Rahmen des oben beschriebenen Verfahrens betrachtet wird, stellt insbesondere das Vorhandensein von bestimmten Tumorzellen dar. Das Krankheitsbild betrifft somit eine Tumorerkrankung. Ebenso kann das Krankheitsbild ggf. das Vorhandensein von Insulin-Inseln darstellen. Das Krankheitsbild betrifft in diesem Fall eine Diabetes-Erkrankung.

**[0027]** Neben dem oben beschriebenen Verfahren zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte betrifft die Erfindung ferner ein Verfahren zur rechnergestützten Durchführung eines digitalen Färbetests unter Verwendung von Basisspektren und eines Regressionsmodells, die mit dem obigen Analyseverfahren für ein vorbestimmtes Krankheitsbild bestimmt und digital gespeichert wurden. Bei diesem digitalen Färbetest werden die nachfolgend erläuterten Schritte i) bis iii) durchgeführt, wobei die Schritte wiederum nicht zwangsläufig eine Reihenfolge für deren Durchführung implizieren.

**[0028]** In einem Schritt i) werden räumlich aufgelöste Hyperspektraldaten eines vorgegebenen zu testenden Gewebeschnitts eingelesen, wobei die Hyperspektraldaten für eine Vielzahl von Messpositionen in dem vorgegebenen Gewebeschnitt jeweilige Spektren enthalten, wobei ein jeweiliges Spektrum die Zusammensetzung des vorgegebenen Gewebeschnitts an einer Messposition in der Form eines Vektors aus Intensitätswerten für Kanäle aus einer Vielzahl von Kanälen repräsentiert, wobei ein Intensitätswert eine für den entsprechenden Kanal detektierte physikalische Größe ist. Vorzugsweise sind die Hyperspektraldaten wiederum die oben definierten Massenspektrometriedaten bzw. MALDI-Massenspektrometriedaten.

**[0029]** In einem Schritt ii) werden die Vektoren der Hyperspektraldaten oder abgeleitete Vektoren, die aus den Hyperspektraldaten berechnet werden, an jeder Messposition mit den Basisspektren korreliert, wobei die Basisspektren mit dem oben beschriebenen Analyseverfahren bzw. einer bevorzugten Variante dieses Verfahrens ermittelt wurden. Hierdurch wird für jede Messposition eine Mehrzahl von Korrelationswerten für jeweilige Basisspektren erhalten. Die Korrelation erfolgt vorzugsweise wieder über ein Skalarprodukt aus normierten Vektoren.

**[0030]** Schließlich wird in einem Schritt iii) das Regressionsmodell, das mit dem obigen Analyseverfahren bzw. einer bevorzugten Variante dieses Verfahrens bestimmt wurde, auf die Korrelationswerte angewendet. Im Falle eines linearen Regressionsmo-

dells werden die Korrelationswerte für eine jeweilige Messposition basierend auf den Gewichtungsfaktoren des Modells linear miteinander kombiniert. Mittels des Regressionsmodells wird eine Maske erhalten, welche für die Vielzahl von Messpositionen jeweilige Skalenwerte enthält, wobei Bereiche des vorgegebenen Gewebeschnitts mit gegenüber dem restlichen Gewebeschnitt abweichenden Skalenwerten Gewebebereiche mit dem vorbestimmten Krankheitsbild sind.

**[0031]** Mit dem soeben erläuterten Verfahren können somit digitale Masken generiert werden, welche Informationen nach Art eines Färbetests enthalten. Vorzugsweise wird dabei eine Visualisierung auf einem Bildschirm bzw. Display derart generiert, dass ein Bild aus den Messpositionen repräsentierenden Bildpunkten erzeugt wird, wobei ein jeweiliger Bildpunkt den Skalenwert der in Schritt iii) ermittelten Maske für die durch den jeweiligen Bildpunkt repräsentierte Messposition codiert. Auf diese Weise wird eine digitale Darstellung geschaffen, die einem herkömmlichen Färbetest entspricht.

**[0032]** In einer Variante des soeben beschriebenen digitalen Färbetests werden Ergebnisse derjenigen Ausführungsform des erfindungsgemäßen Analyseverfahrens genutzt, welche auf abgeleiteten Vektoren basiert. In dieser Variante des Färbetests werden in Schritt ii) abgeleitete Vektoren mit den Basisspektren korreliert, wobei die abgeleiteten Vektoren wie nachfolgend erläutert bestimmt werden.

**[0033]** Für das vorbestimmte Krankheitsbild sind Hyperspektraldaten für eine Mehrzahl von Molekülzusammensetzungen vorgegeben, wobei die Molekülzusammensetzungen in gleicher Weise wie im obigen Analyseverfahren definiert sind. Die Hyperspektraldaten für eine Molekülzusammensetzung enthalten somit einen Vektor in der Form eines für die Molekülzusammensetzung charakteristischen Spektrums aus Intensitätswerten für eine vorgegebene Anzahl von Kanälen aus der Vielzahl von Kanälen. Die für die Molekülzusammensetzung charakteristischen Spektren werden mit den Spektren der in Schritt i) eingelesenen Hyperspektraldaten an den jeweiligen Messpositionen korreliert, wodurch als abgeleiteter Vektor für jede Messposition ein Vektor aus mehreren Korrelationskoeffizienten ermittelt wird, wobei für eine jeweilige Messposition ein Korrelationskoeffizient für jede Molekülzusammensetzung vorliegt.

**[0034]** Neben den oben beschriebenen Verfahren betrifft die Erfindung ferner einer Vorrichtung zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte des menschlichen oder tierischen Körpers zur Präparation eines digitalen Färbetests, wobei die Vorrichtung zur Durchführung des oben beschriebenen Analyseverfahrens bzw. einer oder mehrerer bevorzugter Varianten dieses Analysever-

fahrens eingerichtet ist. Mit anderen Worten umfasst die Vorrichtung Mittel zur Durchführung der Schritte a) bis e) des erfindungsgemäßen Analyseverfahrens.

**[0035]** Die Erfindung betrifft darüber hinaus eine Vorrichtung zur rechnergestützten Durchführung eines digitalen Färbetests unter Verwendung von Basisspektren und eines Regressionsmodells, die mit dem obigen Analyseverfahren bzw. einer oder mehrerer bevorzugter Varianten dieses Verfahrens für ein vorbestimmtes Krankheitsbild bestimmt und digital gespeichert wurden. Dabei ist die Vorrichtung zur Durchführung des obigen digitalen Färbetests bzw. einer oder mehrerer bevorzugter Varianten dieses Färbetests eingerichtet. Mit anderen Worten umfasst die Vorrichtung Mittel zur Durchführung der Schritte i) bis iii) des erfindungsgemäßen Färbetests.

**[0036]** Die Erfindung betrifft darüber hinaus ein Computerprogrammprodukt mit einem auf einem maschinenlesbaren Träger gespeicherten Programmcode zur Durchführung eines Verfahrens zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte bzw. zur rechnergestützten Durchführung eines digitalen Färbetests gemäß der Erfindung bzw. bevorzugter Varianten der Erfindung, wenn der Programmcode auf einem Computer ausgeführt wird.

**[0037]** Ferner betrifft die Erfindung ein Computerprogramm mit einem Programmcode zur Durchführung eines Verfahrens zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte bzw. zur rechnergestützten Durchführung eines digitalen Färbetests gemäß der Erfindung bzw. bevorzugter Ausführungsformen der Erfindung, wenn der Programmcode auf einem Computer ausgeführt wird.

**[0038]** Ausführungsbeispiele der Erfindung werden nachfolgend anhand der beigegeführten Figuren detailliert beschrieben.

**[0039]** Es zeigen:

**[0040]** Fig. 1 ein Diagramm, welches die in einer Ausführungsform des erfindungsgemäßen Analyseverfahrens durchgeführten Schritte verdeutlicht;

**[0041]** Fig. 2 ein Diagramm, welches die in einer Ausführungsform des erfindungsgemäßen digitalen Färbetests durchgeführten Schritte verdeutlicht;

**[0042]** Fig. 3 eine Darstellung, welche die in einer Ausführungsform des erfindungsgemäßen Analyseverfahrens ermittelten Basisspektren wiedergibt;

**[0043]** Fig. 4 eine Darstellung, welche die Bestimmung eines Regressionsmodells basierend auf den Basisspektren der Fig. 3 wiedergibt; und

**[0044]** Fig. 5 eine Visualisierung, welche gemäß einer Ausführungsform des erfindungsgemäßen digitalen Färbetests generiert wird.

**[0045]** Nachfolgend werden Ausführungsformen der Erfindung basierend auf Hyperspektraldaten in der Form der bereits eingangs erwähnten MALDI-Massenspektrometriedaten erläutert. Im Rahmen dieser Spektrometrie werden die Moleküle des zu analysierenden Gewebeschnitts mit einer Matrix-Lösung überzogen, und ein Laserstrahl (z.B. eines Stickstofflasers) wird auf verschiedene Messpositionen des entsprechenden Probenträgers gerichtet. Über die Laserpulse des Laserstrahls erfolgt eine Teilchenablösung von ionisierten Molekülen des Gewebes, deren Masse-zu-Ladungs-Verhältnis in an sich bekannter Weise mit einem Massenspektrometer erfasst wird. Hierzu kann z.B. ein TOF-Massenspektrometer (TOF = Time of Flight) oder FTICR-Massenspektrometer (FTICR = Fourier Transform Ion Cyclotron Resonance) genutzt werden. Die hier beschriebenen Verfahren der Erfindung sind jedoch ggf. auch auf andere Massenspektrometriedaten anwendbar, sofern die Daten orts aufgelöst an verschiedenen Messpositionen Spektren enthalten, die in Abhängigkeit von dem Masse-zu-Ladungs-Verhältnis die Anzahl von mit dem Massenspektrometer detektierten Ionen wiedergeben. Ferner kann das Verfahren der Erfindung auch allgemein auf Hyperspektraldaten angewendet werden (z.B. Raman-Spektroskopie), sofern die Daten orts aufgelöst an verschiedenen Messpositionen Intensitätsvektoren mit Einträgen für verschiedene Kanäle enthalten. Eine Intensität kann z.B. den Anteil von reflektiertem bzw. gestreutem Licht einer bestimmten Wellenlänge umfassen, wobei die Wellenlänge den Kanal darstellt.

**[0046]** Im Folgenden wird zunächst anhand von Fig. 1 ein Verfahren zur Analyse eines Gewebeschnitts erläutert, dessen Ergebnisse dann im Rahmen eines digitalen Färbetests verwendet werden, wie anschließend anhand von Fig. 2 beschrieben wird. Fig. 3 bis Fig. 5 zeigen Ergebnisse des Analyseverfahrens bzw. des Färbetests, die anhand eines Gewebeschnitts der Bauchspeicheldrüse einer Ratte ermittelt wurden. Die entsprechenden Gewebeschnitte umfassen Insulin-Inseln in der Bauchspeicheldrüse, so dass die Gewebeschnitte dem Krankheitsbild von Diabetes entsprechen.

**[0047]** Im Rahmen der hier beschriebenen Ausführungsform des Analyseverfahrens wird gemäß Fig. 1 ein vorbestimmter Gewebeschnitt TI betrachtet, der mittels der obigen MALDI-Spektrometrie analysiert wurde. Demzufolge liegen an einer Vielzahl von zweidimensionalen Messpositionen  $x$  des Gewebeschnitts jeweilige Spektren  $S_{ij}$  vor, wobei  $i, j$  die kartesischen Koordinaten der jeweiligen Messposition indizieren. Ein einzelnes Spektrum enthält für eine Vielzahl von Kanälen in der Form von  $m/z$ -Werten (d.h.

Masse-zu-Ladungs-Verhältnissen) Intensitätswerte  $I$ , welche jeweils die Anzahl der für den entsprechenden  $m/z$ -Wert detektierten Moleküle darstellen. Das nachfolgend beschriebene Verfahren ist auf die Analyse von mehreren Gewebeschnitten übertragbar, wobei sich in diesem Fall die Messpositionen über die Gewebeschnitte hinweg fortsetzen. Die Messpositionen aus unterschiedlichen Gewebeschnitten können z.B. über einen Index für den jeweiligen Gewebeschnitt voneinander unterschieden werden.

**[0048]** In einer bevorzugten Variante wurde der Gewebeschnitt TI vor der Durchführung der MALDI-Spektrometrie einem Verdau unterzogen, bei dem die Moleküle des Gewebeschnitts über biochemische Prozesse geeignet zersetzt wurden. Hierzu werden Enzyme auf den Gewebeschnitt aufgebracht, welche die Moleküle des Gewebeschnitts in Peptide zerschneiden. Über einen Verdau lassen sich in einem entsprechenden Gewebeschnitt leichter Krankheitsbilder anhand von Massenspektrometriedaten ermitteln. Neben den Massenspektrometriedaten  $S_{ij}$  wird im Rahmen der Analyse des Gewebeschnitts TI auch eine (digitale) Maske  $M_{ij}$  des entsprechenden Gewebeschnitts benötigt. Diese Maske gibt für die einzelnen Messpositionen  $x$  des Gewebeschnitts einen jeweiligen Skalenwert an, der z.B. zwischen 0 und 1 liegt. Die Größe des Skalenwerts zeigt an, wie groß der Anteil der dem vorbestimmten Krankheitsbild entsprechenden Moleküle an der jeweiligen Messposition ist. Gegebenenfalls können die Skalenwerte auch nur die Werte 0 und 1 annehmen, wobei der Wert von 0 für einen Gewebebereich ohne Krankheitsbild und der Wert von 1 für einen Gewebebereich mit Krankheitsbild steht. Die Maske  $M_{ij}$  kann z.B. aus einem herkömmlichen immunhistochemischen Färbetest gewonnen worden sein, bei dem über den Zusatz von Antikörpern in dem Gewebeschnitt Biomarker, die für das vorbestimmte Krankheitsbild spezifisch sind, visuell über Gewebeverfärbung hervorgehoben werden, wobei die Intensität der Gewebeverfärbung den Anteil der Biomarker an der entsprechenden Gewebeposition anzeigt. Die Ergebnisse des immunhistochemischen Färbetests werden dann in digitale Daten der Maske  $M_{ij}$  gewandelt.

**[0049]** Wie soeben erläutert, sind die Eingangsdaten des hier beschriebenen Analyseverfahrens der Massenspektrometriedaten  $S_{ij}$  sowie die Maske  $M_{ij}$ , was in Schritt S1 in **Fig. 1** verdeutlicht ist. Diese Daten werden in digitaler Form eingelesen, wobei der Schritt der Ermittlung der Daten und insbesondere die massenspektrometrische Erfassung der Spektren  $S_{ij}$  ggf. auch Teil des Analyseverfahrens sein kann.

**[0050]** In einem Schritt S2 des Analyseverfahrens werden nunmehr aus den Spektren  $S_{ij}$  spektrale Signaturen in der Form von Basisspektren zur Unterscheidung metabolisch relevanter Strukturen bestimmt. Beispielhaft sind in **Fig. 1** die fünf Basisspek-

tren  $BS_1, BS_2, \dots, BS_5$  angegeben, wobei je nach Ausführungsform mehr oder weniger Basisspektren ermittelt werden können. Die Basisspektren werden mit einem an sich bekannten Verfahren zur Bestimmung von Basisvektoren in dem Vektorraum der Spektraldaten berechnet. Dabei repräsentiert ein jeweiliger Vektor in dem Vektorraum ein einzelnes Spektrum  $S_{ij}$  an einer entsprechenden Messposition und die Einträge des Vektors sind die Intensitätswerte an den  $m/z$ -Werten. Die Basisspektren stellen die mit der Methode bestimmten Basisvektoren dar. Insbesondere approximieren die Basisspektren dabei eine Basis, in der die Spektren  $S_{ij}$  als Linearkombination der Basisspektren beschrieben werden. In der hier beschriebenen Ausführungsform wird eine nicht-negative Matrixfaktorisierung mit der Nebenbedingung einer spärlichen Besetzung der daraus abgeleiteten Matrizen zur Bestimmung der Basisvektoren eingesetzt. Dabei wird die nicht-negative Matrixfaktorisierung aus der Druckschrift [1] bzw. der Druckschrift [2] verwendet. Die Komponenten in der Terminologie der Druckschrift [1] entsprechen den Basisvektoren.

**[0051]** In einem nächsten Schritt S3 werden die ursprünglichen Spektren  $S_{ij}$  mit den in Schritt S2 ermittelten Basisspektren  $BS_k$  basierend auf einem Skalarprodukt korreliert. Das heißt, entsprechende Korrelationswerte  $P_{ijk}$  für jedes Spektrum  $S_{ij}$  und jeden Basisvektor  $BS_k$  werden wie folgt erhalten:

$$P_{ijk} = \langle S_{ij}, BS_k \rangle$$

**[0052]** Der Korrelationswert  $P_{ijk}$  ist auf 1 normiert und liegt zwischen  $-1$  und  $1$ . Er entspricht bei zwei identischen Vektoren dem Wert 1. Dies wird dadurch erreicht, dass die Vektoren  $S_{ij}$  und  $BS_k$  vor Bildung des Skalarprodukts durch ihren Betrag geteilt und damit normiert werden.

**[0053]** In einem Schritt S4 wird schließlich mittels eines Regressionsverfahrens ein Regressionsmodell RM bestimmt, welches die ursprüngliche Maske  $M_{ij}$  basierend auf den Korrelationswerten  $P_{ijk}$  beschreibt. In der hier beschriebenen Ausführungsform wird ein lineares Regressionsverfahren verwendet, bei dem die Maske  $M_{ij}$  durch eine Linearkombination der Korrelationswerte approximiert wird, d.h. es gilt:

$$\sum_k c_k P_{ijk} = M_{ij}$$

**[0054]** Die mittels des linearen Regressionsverfahrens bestimmten Gewichtungsfaktoren  $c_k$  sind die Parameter des ermittelten Regressionsmodells RM. Anstatt eines linearen Regressionsverfahrens können auch beliebige andere Regressionsverfahren zur Bestimmung eines Regressionsmodells verwendet werden, das die Maske in Abhängigkeit von den Korrelationswerten beschreibt. In einer besonders bevor-

zugten Ausführungsform wird ein an sich bekanntes SVM-Regressionsverfahren (SVM = Support Vector Machines) eingesetzt.

$$MA_{ij} = \sum_k c_k N_{ijk}$$

**[0055]** Als Endergebnis des Verfahrens der **Fig. 1** erhält man schließlich das Regressionsmodell RM mit entsprechenden Parametern  $c_k$  sowie eine Mehrzahl von Basisspektren  $BS_k$ . Diese Informationen werden digital gespeichert, um anschließend in einem Verfahren zur Durchführung eines digitalen Färbetests eingesetzt zu werden. Dieses Verfahren wird im Folgenden anhand von **Fig. 2** erläutert wird.

**[0056]** Ausgangspunkt des Verfahrens der **Fig. 2** ist der Schritt S1', in dem für eine Vielzahl von zweidimensionalen Messpositionen  $y$  eines zu testenden Gewebeschnitts Tl' Massenspektrometriedaten  $SP_{ij}$  mit entsprechenden Intensitätswerten  $I$  an jeweiligen  $m/z$ -Werten eingelesen werden. Gegebenenfalls kann die massenspektrometrische Messung zur Bestimmung der Spektren  $SP_{ij}$  selbst Teil des Verfahrens sein. Nichtsdestotrotz können die Daten auch schon vorgegeben sein und aus einem digitalen Speicher ausgelesen werden. Für den Gewebeschnitt Tl' soll rein anhand der Massenspektrometriedaten eine Maske  $MA_{ij}$  erzeugt werden, die analog zu der Maske  $M_{ij}$  aus **Fig. 1** über entsprechende Skalenwerte das Vorhandensein eines vorbestimmten Krankheitsbilds an entsprechenden Messpositionen codiert.

**[0057]** In dem Schritt S2' des digitalen Färbetests werden nunmehr in gleicher Weise wie in Schritt S3 der **Fig. 1** Korrelationswerte bestimmt, die nunmehr jedoch nicht auf dem Skalarprodukt zwischen den Spektren  $S_{ij}$  und den Basisvektoren  $BS_k$ , sondern auf dem Skalarprodukt zwischen den Spektren  $SP_{ij}$  und den Basisvektoren  $BS_k$  basieren. Mit anderen Worten werden entsprechende Korrelationswerte  $N_{ijk}$  wie folgt erhalten:

$$N_{ijk} = \langle SP_{ij}; BS_k \rangle$$

**[0058]** In Analogie zu den Korrelationswerten  $P_{ijk}$  aus **Fig. 1** liegen auch die Korrelationswerte  $N_{ijk}$  zwischen  $-1$  und  $1$ . Die im Schritt S2' verarbeiteten Basisvektoren  $BS_k$  und auch das im folgenden Schritt S3' verarbeitete Regressionsmodell RM werden aus einem digitalen Speicher ausgelesen und wurden mit dem Verfahren der **Fig. 1** ermittelt.

**[0059]** In dem Schritt S3' der **Fig. 2** wird nunmehr basierend auf den Korrelationswerten  $N_{ijk}$  unter Verwendung des Regressionsmodells RM bzw. dessen Parametern  $c_k$  die Maske  $MA_{ij}$  mit deren Skalenwerten ermittelt, d.h. es wird bestimmt, wie sich die Korrelationswerte der Spektren der zu testenden Gewebeprobe Tl' gemäß dem Regressionsmodell überlagern und hierdurch die Maskendaten  $MA_{ij}$  bilden. Mit anderen Worten werden die Maskendaten wie folgt ermittelt:

**[0060]** Auf diese Weise werden rechnergestützt durch Rückgriff auf die Ergebnisse des Analyseverfahrens der **Fig. 1** Gewebebereiche mit einem vorbestimmten Krankheitsbild entsprechend markiert. Vorzugsweise wird die Maske  $MA_{ij}$  in der Form eines digitalen Bildes auf einem Display visualisiert. Dieses Bild zeigt den Gewebeschnitt Tl' in der Form einer Vielzahl von Pixeln, die jeweils einer Messposition des Gewebeschnitts entsprechen. Jedes Bildpixel codiert dabei über seinen Grauwert oder ggf. Farbwert den entsprechenden Skalenwert der Maske  $MA_{ij}$ , wodurch ein Färbesbild entsteht, wie es herkömmlicherweise mittels immunhistochemischer Färbetests bestimmt wird. Erfindungsgemäß müssen jedoch keine biochemischen Verfahrensschritte durchgeführt werden, sondern das Färbesbildwert wird rein rechnergestützt aus massenspektrometrischen Daten abgeleitet.

**[0061]** Die anhand von **Fig. 1** und **Fig. 2** beschriebene Verfahren stellen einen sog. Bottom-Up-Ansatz dar, bei dem ohne Wissen über die relevanten Molekülzusammensetzungen in der Gewebeprobe die Basisspektren rein über ein rechnergestütztes Verfahren ermittelt werden. In einer abgewandelten Variante (Top-Down-Ansatz) ist es auch möglich, die Basisvektoren mit dem Vorwissen über die Biomarker bzw. entsprechender Peptide, in welche sich die Biomarker durch Verdau zersetzen, zu bestimmen. Dabei wurde vorab bestimmt bzw. simuliert, wie sich die bekannten Biomarker unter dem verwendeten Verdau (z.B. Trypsin) in Peptide zersetzen. Für die chemische Zusammensetzung eines jeweiligen Peptids wurde ferner ein Massenspektrum bestimmt, welches für die  $m/z$ -Werte der Isotope des Peptids jeweils einen Intensitätswert in der Form der Anzahl der detektierten Moleküle angibt. Verfahren zur Bestimmung der entsprechenden Peptide und deren Spektren sind an sich bekannt. Da diese Verfahren nicht Bestandteil der Erfindung sind, sondern lediglich deren Ergebnisse in der Form von Peptid-Spektren genutzt werden, werden sie nicht im Detail beschrieben.

**[0062]** Die einzelnen Spektren der Peptide werden gemäß der abgewandelten Variante in einem Vorverarbeitungsschritt mit den Spektren  $S_{ij}$  der Massenspektrometriedaten aus **Fig. 1** korreliert. Die Korrelation erfolgt wiederum mittels eines Skalarprodukts, wobei zur Bildung des Skalarprodukts die (normierten)  $m/z$ -Werte der Spektren der Peptide mit den entsprechenden (normierten)  $m/z$ -Werten der Spektren  $S_{ij}$  multipliziert werden und diese Produkte summiert werden. Die sich ergebenden Korrelationswerte, die zur Unterscheidung gegenüber den obigen Korrelationswerten als Korrelationskoeffizienten bezeichnet werden, liegen wiederum zwischen  $0$  und  $1$ . Als Er-



gebnis erhält man schließlich für jede Messposition eine der Anzahl von Peptiden entsprechende Anzahl an Korrelationskoeffizienten, welche abgeleitete Vektoren im Sinne der Patentansprüche darstellen. Auf diese abgeleiteten Vektoren aus Korrelationskoeffizienten wird in gleicher Weise wie in **Fig. 1** ein Verfahren zur Bestimmung von Basisvektoren und insbesondere eine nicht-negative Matrixfaktorisierung angewendet, wobei die Basisvektoren nunmehr in dem Vektorraum der abgeleiteten Vektoren ermittelt werden. Diese Basisvektoren werden dann mit den abgeleiteten Vektoren anstatt den ursprünglichen Spektren korreliert, woraus sich wiederum Korrelationswerte ergeben, welche – analog wie oben beschrieben – so bestimmt werden, dass sie zwischen 0 und 1 liegen. In Analogie zu Schritt S4 der **Fig. 1** wird dann ein Regressionsmodell basierend auf diesen Korrelationswerten bestimmt, welches die Maske  $M_{ij}$  beschreibt.

**[0063]** Werden die Ergebnisse der soeben beschriebenen abgewandelten Ausführungsform des Analyseverfahrens nunmehr im Rahmen des digitalen Färbetests eingesetzt, werden vor Durchführung des Schritts S2' aus **Fig. 2** zunächst abgeleitete Vektoren mit entsprechenden Korrelationskoeffizienten durch Korrelation die Spektren der Peptide mit den ursprünglichen Spektren  $SP_{ij}$  ermittelt. Anschließend erfolgt in Analogie zu Schritt S2' die Ermittlung der Korrelationswerte, wobei nunmehr jedoch die abgeleiteten Vektoren verwendet werden. Schließlich wird genauso wie im Schritt S3' der **Fig. 2** die Maske  $MA_{ij}$  mit dem Regressionsmodell bestimmt.

**[0064]** Wie bereits eingangs erwähnt, zeigen **Fig. 3** bis **Fig. 5** Ergebnisse einer Ausführungsform des erfindungsgemäßen Analyseverfahrens bzw. des erfindungsgemäßen Färbetests für den Gewebeschnitt der Bauchspeicheldrüse einer Ratte, wobei die Bauchspeicheldrüse Insulin-Inseln enthält. Der Gewebeschnitt wurde dabei basierend auf einem Bottom-Up-Ansatz analysiert. Es wurden insgesamt zehn Basisspektren aus den ursprünglichen Spektren  $S_{ij}$  ermittelt, wobei in **Fig. 3** beispielhaft fünf dieser Basisspektren  $BS_1$  bis  $BS_5$  wiedergegeben sind. Für die einzelnen Basisspektren wurden die entsprechenden Korrelationswerte und basierend darauf die Koeffizienten  $c_k$  eines linearen Regressionsmodells bestimmt. Dies wird aus **Fig. 4** ersichtlich.

**[0065]** In **Fig. 4** sind im oberen rechten Diagramm die zehn Gewichtungsfaktoren  $c_k$  des Regressionsmodells für die Basisvektoren  $k = 1, \dots, 10$  wiedergegeben. Ferner ist im unteren rechten Diagramm die Überlagerung der Basisspektren entsprechend den Gewichtungsfaktoren  $c_k$  angedeutet, d.h.  $\sum c_k \cdot BS_k$ . Darüber hinaus ist in dem digitalen Bild auf der linken Seite ein Ausschnitt des Gewebeschnitts mit entsprechenden Pixeln als Bildpositionen  $x$  dargestellt, wobei die weißen Umrandungen, die aus Über-

sichtlichkeitsgründen nur teilweise mit dem Bezugszeichen  $L$  bezeichnet sind, Bereiche abgrenzen, bei denen in dem Gewebeschnitt Insulin-Inseln vorliegen. Über die Umrandungen  $L$  werden somit die ursprünglichen Maskendaten angedeutet, die im vorliegenden Fall nur die Skalenwerte 0 und 1 einnehmen, wobei der Skalenwert 0 einen Bereich ohne Insulin und der Skalenwert 1 einen Bereich einer Insulin-Insel darstellt. Die einzelnen Pixelwerte in dem dargestellten Bild stellen die aus dem ermittelten Regressionsmodell rekonstruierten Maskendaten  $M_{ij}$  dar, wobei die Skalenwerte der einzelnen Pixel farbcodiert sind, was aus der Schwarz-Weiß-Darstellung der **Fig. 4** nicht ersichtlich wird. Nichtsdestotrotz stellen dunkle Pixelwerte große Skalenwerte im Bereich von 1 dar. Man kann somit erkennen, dass die Rekonstruktion die ursprünglichen Maskendaten gut wiedergibt.

**[0066]** Das Regressionsmodell mit den entsprechenden Gewichtungsfaktoren  $c_k$  aus **Fig. 4** wurde schließlich auf einen neuen Gewebeschnitt der Bauchspeicheldrüse der Ratte angewendet. Mit anderen Worten wurden für den neuen Gewebeschnitt MALDI-Massenspektrometriedaten ermittelt und anschließend basierend auf den Schritten S2' und S3' der **Fig. 2** die Maskendaten  $MA_{ij}$  abgeleitet. Diese Maskendaten sind für den neuen Gewebeschnitt in **Fig. 5** wiedergegeben, wobei die Messpositionen  $y$  wieder einzelne farbcodierte Pixel sind, was aus der Schwarz-Weiß-Darstellung der **Fig. 5** jedoch nicht ersichtlich ist. Schwarze Pixelwerte stellen jedoch große Skalenwerte dar. In **Fig. 5** sind wiederum durch entsprechende weiße Umrandungen  $L$  (nur teilweise mit diesem Bezugszeichen bezeichnet) Insulin-Inseln angedeutet, wie sie tatsächlich in dem Gewebeschnitt identifiziert werden konnten. Wie man erkennt, repräsentieren die dargestellten Maskendaten  $MA_{ij}$  sehr gut die tatsächlich im Gewebeschnitt vorhandenen Insulin-Inseln.

**[0067]** Die im Vorangegangenen beschriebenen Ausführungsformen der Erfindung weisen eine Reihe von Vorteilen auf. Insbesondere können erstmalig Gewebeschnitte rein basierend auf massenspektrometrischen Methoden analysiert werden, um hierdurch pathologische Gewebereiche abzuleiten. Das Verfahren kann somit zumindest zum Teil aufwändige immunhistochemische Tests oder andere molekulare Testverfahren ersetzen. Es müssen hierzu lediglich einmal basierend auf einem Analyseverfahren ein Regressionsmodell und Basisspektren ermittelt werden, die dann bei der Auswertung der Massenspektrometriedaten eingesetzt werden. Auf diese Weise wird ein digitaler Färbetest geschaffen, der krankhafte Gewebereiche in einem Gewebeschnitt identifiziert. Vorzugsweise wird einem Pathologen der Gewebeschnitt in der Form eines Bilds repräsentiert, in dem krankhafte Bereiche basierend auf einer Farbcodierung wiedergegeben sind.

## Literaturverzeichnis:

- [1] DE 10 2013 207 402 A1  
 [2] J. H. Kobarg et al., "Numerical experiments with MALDI Imaging data", Adv Comput Math (2014) 40, Seiten 667 bis 682, November 2013

## Patentansprüche

1. Verfahren zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte (TI) des menschlichen oder tierischen Körpers zur Präparation eines digitalen Färbetests, wobei in dem digitalen Färbetest rechnergestützt in einem getesteten Gewebeschnitt (TI) Gewebebereiche bestimmt werden, die ein vorbestimmtes Krankheitsbild aufweisen, wobei:

a) räumlich aufgelöste Hyperspektraldaten von einem oder mehreren vorbestimmten zu analysierenden Gewebeschnitten (TI), die Gewebebereiche mit dem vorbestimmten Krankheitsbild aufweisen, eingelesen werden, wobei die Hyperspektraldaten für eine Vielzahl von Messpositionen ( $x$ ) in dem oder den vorbestimmten Gewebeschnitten (TI) jeweilige Spektren ( $S_{ij}$ ) enthalten, wobei ein jeweiliges Spektrum ( $S_{ij}$ ) die Zusammensetzung des oder der vorbestimmten Gewebeschnitte (TI) an einer Messposition ( $x$ ) in der Form eines Vektors aus Intensitätswerten ( $I$ ) für eine Vielzahl von Kanälen ( $m/z$ ) repräsentiert, wobei ein Intensitätswert ( $I$ ) eine für den entsprechenden Kanal ( $m/z$ ) detektierte physikalische Größe ist;

b) eine Maske ( $M_{ij}$ ) des oder der vorbestimmten Gewebeschnitte (TI) eingelesen wird, welche für die Vielzahl von Messpositionen ( $x$ ) jeweilige Skalenwerte enthält, wobei Bereiche des oder der vorbestimmten Gewebeschnitte (TI) mit gegenüber dem restlichen Bereich des oder der Gewebeschnitte (TI) abweichenden Skalenwerten Gewebebereiche mit dem vorbestimmten Krankheitsbild sind;

c) mittels einer Methode zur Bestimmung von Basisvektoren mehrere Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) im Raum der Vektoren der Hyperspektraldaten oder in einem Raum von abgeleiteten Vektoren, die aus den Hyperspektraldaten für jede Messposition ( $x$ ) berechnet werden, ermittelt werden;

d) die Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) mit den Vektoren der Hyperspektraldaten oder den abgeleiteten Vektoren an jeder Messposition ( $x$ ) korreliert werden, wodurch für jede Messposition ( $x$ ) eine Mehrzahl von Korrelationswerten ( $P_{ijk}$ ) für jeweilige Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) erhalten wird;

e) mittels eines Regressionsverfahrens ein Regressionsmodell (RM) bestimmt wird, welches die Maske ( $M_{ij}$ ) basierend auf der Mehrzahl von Korrelationswerten ( $P_{ijk}$ ) für die jeweiligen Messpositionen ( $x$ ) beschreibt, wobei das Regressionsmodell (RM) und die Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) für die Durchführung des digitalen Färbetests digital gespeichert werden.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, dass die Hyperspektraldaten Massenspektrometriedaten, insbesondere MALDI-Massenspektrometriedaten (MALDI = Matrix-Assisted Laser Desorption/Ionization), sind, in denen die Kanäle ( $m/z$ ) Masse-zu-Ladungs-Verhältnisse darstellen und die für den entsprechenden Kanal ( $m/z$ ) detektierte physikalische Größe die Anzahl der für das entsprechende Masse-zu-Ladungs-Verhältnis detektierten Moleküle repräsentiert.

3. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass in Schritt c) die Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) im Raum der abgeleiteten Vektoren bestimmt werden, wobei ein abgeleiteter Vektor für jede Messposition ( $x$ ) derart berechnet wird, dass

– für das vorbestimmte Krankheitsbild Hyperspektraldaten für eine Mehrzahl von Molekülzusammensetzungen vorgegeben sind, wobei die Hyperspektraldaten für eine Molekülzusammensetzung einen Vektor in der Form eines für die Molekülzusammensetzung charakteristischen Spektrums aus Intensitätswerten ( $I$ ) für eine vorgegebene Anzahl von Kanälen ( $m/z$ ) aus der Vielzahl von Kanälen ( $m/z$ ) enthalten;

– die für die Molekülzusammensetzungen charakteristischen Spektren mit den Spektren ( $S_{ij}$ ) der eingelesenen Hyperspektraldaten an den jeweiligen Messpositionen ( $x$ ) korreliert werden, wodurch als abgeleiteter Vektor für jede Messposition ( $x$ ) ein Vektor aus mehreren Korrelationskoeffizienten ermittelt wird.

4. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass der oder die zu analysierenden Gewebeschnitte jeweilige Gewebeschnitte mit Verdau sind.

5. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass die in Schritt c) verwendete Methode zur Bestimmung von Basisvektoren eine nicht-negative Matrixfaktorisierung oder eine PCA-Analyse ist.

6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet**, dass die nicht-negative Matrixfaktorisierung die Nebenbedingung der spärlichen Besetzung der durch die nicht-negative Matrixfaktorisierung bestimmten Matrizen umfasst.

7. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass das in Schritt e) verwendete Regressionsverfahren eine lineare Regressionsverfahren oder ein SVM-Regressionsverfahren ist.

8. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass die Maske ( $M_{ij}$ ) in Schritt b) aus einem histologischen und insbesondere immunhistochemischen Färbetest des oder der vorbestimmten Gewebeschnitte ermittelt ist.

9. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet**, dass das vorbestimmte Krankheitsbild das Vorhandensein vorbestimmter Tumorzellen oder das Vorhandensein von Insulin-Inseln in dem oder den vorbestimmten Gewebeschnitten (TI) darstellt.

10. Verfahren zur rechnergestützten Durchführung eines digitalen Färbetests unter Verwendung von Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) und eines Regressionsmodells (RM), die mit einem Verfahren nach einem der vorhergehenden Ansprüche für ein vorbestimmtes Krankheitsbild bestimmt und digital gespeichert wurden, bei dem:

i) räumlich aufgelöste Hyperspektraldaten eines vorgegebenen zu testenden Gewebeschnitts (TI') eingelesen werden, wobei die Hyperspektraldaten für eine Vielzahl von Messpositionen (y) in dem vorgegebenen Gewebeschnitt (TI') jeweilige Spektren ( $SP_{ij}$ ) enthalten, wobei ein jeweiliges Spektrum ( $SP_{ij}$ ) die Zusammensetzung des vorgegebenen Gewebeschnitts (TI') an einer Messposition (y) in der Form eines Vektors aus Intensitätswerten (I) für Kanäle (m/z) aus einer Vielzahl von Kanälen (m/z) repräsentiert, wobei ein Intensitätswert (I) eine für den entsprechenden Kanal (m/z) detektierte physikalische Größe ist;

ii) die Vektoren der Hyperspektraldaten oder abgeleitete Vektoren, die aus den Hyperspektraldaten berechnet werden, an jeder Messposition (x) mit den Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) korreliert werden, welche mit einem Verfahren nach einem der vorhergehenden Ansprüche bestimmt wurden, wodurch für jede Messposition (y) eine Mehrzahl von Korrelationswerten ( $N_{ijk}$ ) für jeweilige Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) erhalten wird;

iii) das Regressionsmodell (RM), das mit einem Verfahren nach einem der vorhergehenden Ansprüche bestimmt wurde, auf die Korrelationswerte ( $N_{ijk}$ ) angewendet wird und hierdurch eine Maske ( $MA_{ij}$ ) erhalten wird, welche für die Vielzahl von Messpositionen (y) jeweilige Skalenwerte enthält, wobei Bereiche des vorgegebenen Gewebeschnitts (TI') mit gegenüber dem restlichen Gewebeschnitt (TI') abweichenden Skalenwerten Gewebereiche mit dem vorbestimmte Krankheitsbild sind.

11. Verfahren nach Anspruch 10, **dadurch gekennzeichnet**, dass eine Visualisierung auf einem Display derart generiert wird, dass ein Bild aus die Messpositionen (y) repräsentierenden Bildpunkten erzeugt wird, wobei ein jeweiliger Bildpunkt den Skalenwert der in Schritt iii) ermittelten Maske ( $MA_{ij}$ ) für die durch den jeweiligen Bildpunkt (BP) repräsentierte Messposition (y) codiert.

12. Verfahren nach Anspruch 10 oder 11, **dadurch gekennzeichnet**, dass in Schritt ii) abgeleitete Vektoren mit den Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) korreliert werden, wobei die abgeleiteten Vektoren derart berechnet werden, dass

– für das vorbestimmte Krankheitsbild Hyperspektraldaten für eine Mehrzahl von Molekülzusammensetzungen vorgegeben sind, wobei die Hyperspektraldaten für eine Molekülzusammensetzung einen Vektor in der Form eines für die Molekülzusammensetzung charakteristischen Spektrums aus Intensitätswerten (I) für eine vorgegebene Anzahl von Kanälen (m/z) aus der Vielzahl von Kanälen (m/z) enthalten;

– die für die Molekülzusammensetzungen charakteristischen Spektren mit den Spektren ( $S_{ij}$ ) der in Schritt i) eingelesenen Hyperspektraldaten an den jeweiligen Messpositionen (y) korreliert werden, wodurch als abgeleiteter Vektor für jede Messposition (y) ein Vektor aus mehreren Korrelationskoeffizienten ermittelt wird.

13. Vorrichtung zur rechnergestützten Analyse eines oder mehrerer Gewebeschnitte (TI) des menschlichen oder tierischen Körpers zur Präparation eines digitalen Färbetests, wobei in dem digitalen Färbetest rechnergestützt in einem getesteten Gewebeschnitt (TI') Gewebereiche bestimmt werden, die ein vorbestimmtes Krankheitsbild aufweisen, **dadurch gekennzeichnet**, dass die Vorrichtung zur Durchführung eines Verfahrens nach einem der Ansprüche 1 bis 9 eingerichtet ist.

14. Vorrichtung zur rechnergestützten Durchführung eines digitalen Färbetests unter Verwendung von Basisspektren ( $BS_1, BS_2, \dots, BS_5$ ) und eines Regressionsmodells (RM), die mit einem Verfahren nach einem der Ansprüche 1 bis 9 für ein vorbestimmtes Krankheitsbild bestimmt und digital gespeichert wurden, **dadurch gekennzeichnet**, dass die Vorrichtung zur Durchführung eines Verfahrens nach einem der Ansprüche 10 bis 12 eingerichtet ist.

15. Computerprogrammprodukt mit einem auf einem maschinenlesbaren Träger gespeicherten Programmcode zur Durchführung eines Verfahrens nach einem der Ansprüche 1 bis 12, wenn der Programmcode auf einem Computer ausgeführt wird.

16. Computerprogramm mit einem Programmcode zur Durchführung eines Verfahrens nach einem der Ansprüche 1 bis 12, wenn der Programmcode auf einem Computer ausgeführt wird.

Es folgen 5 Seiten Zeichnungen

## Anhängende Zeichnungen

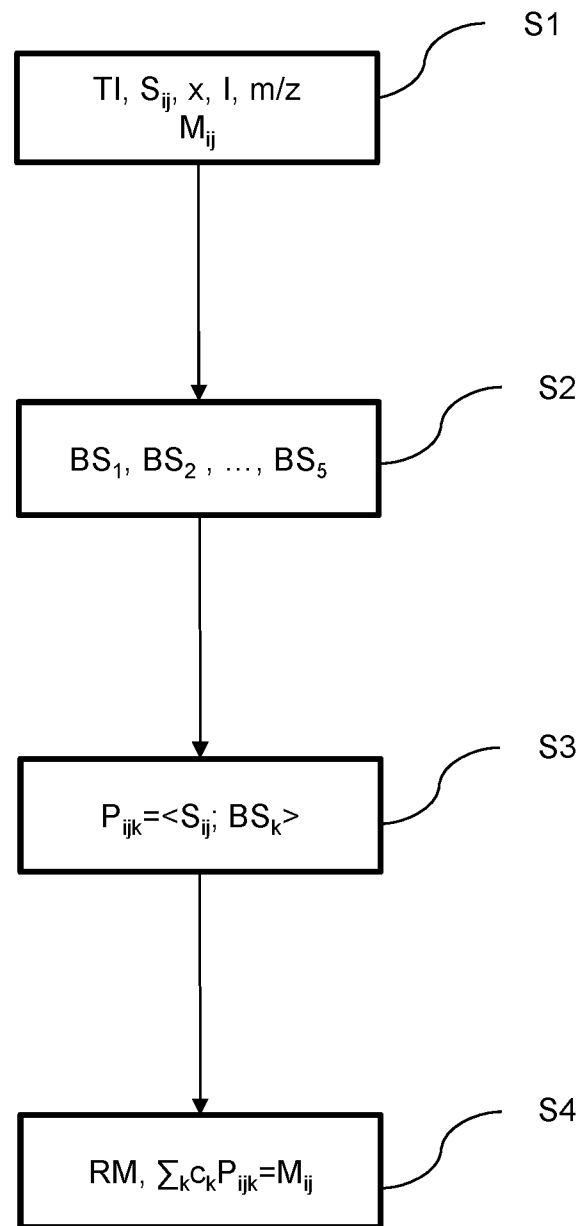


Fig. 1

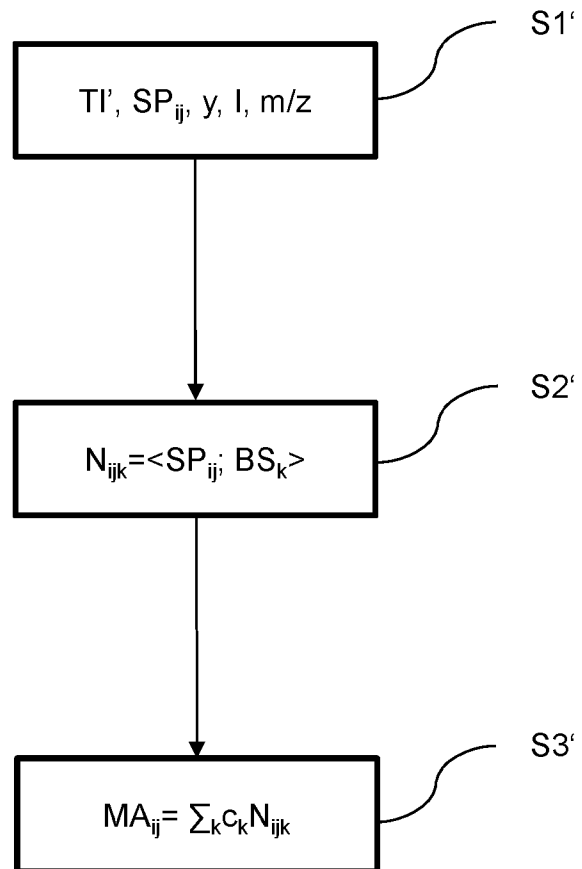


Fig. 2

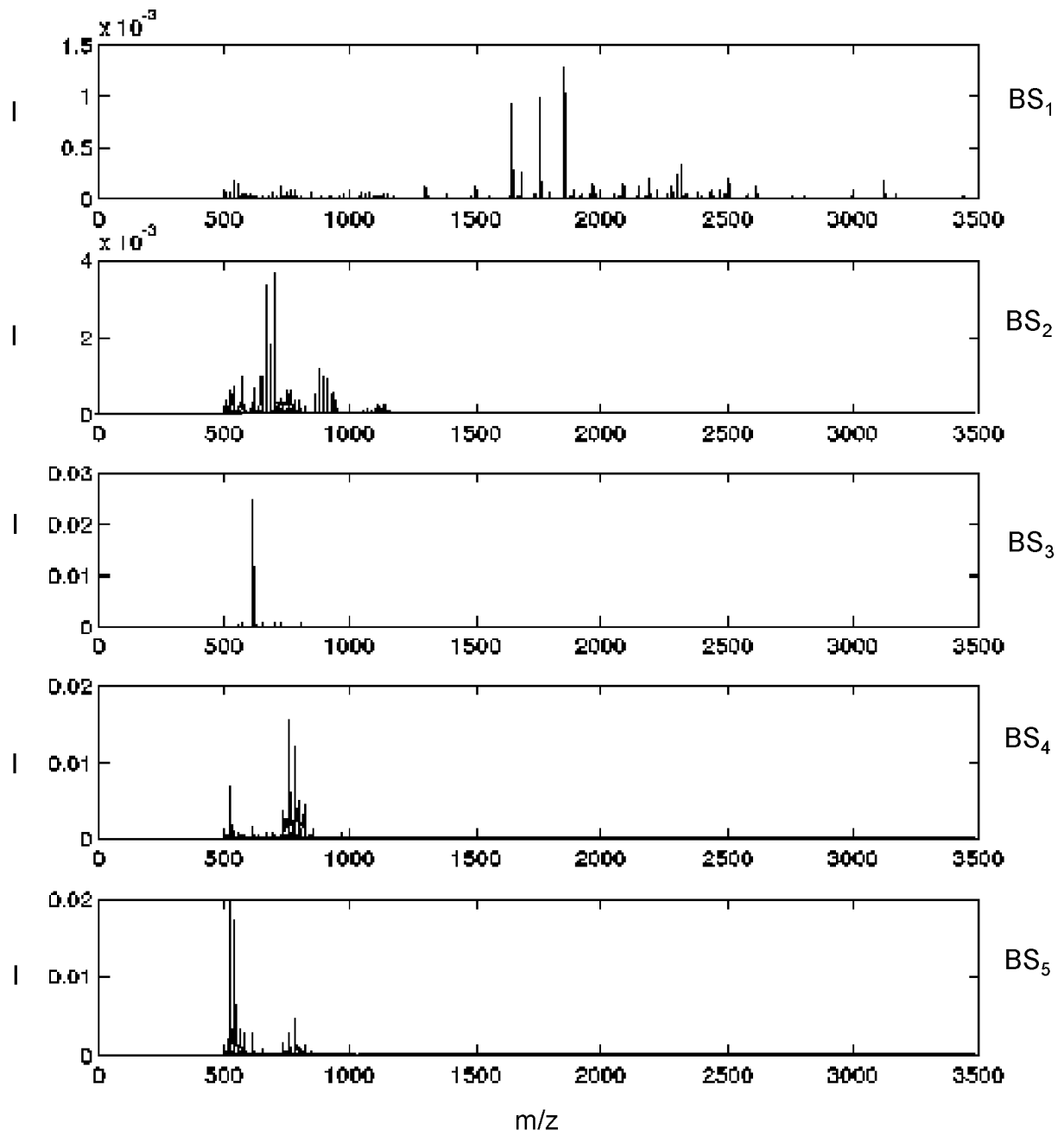


Fig. 3

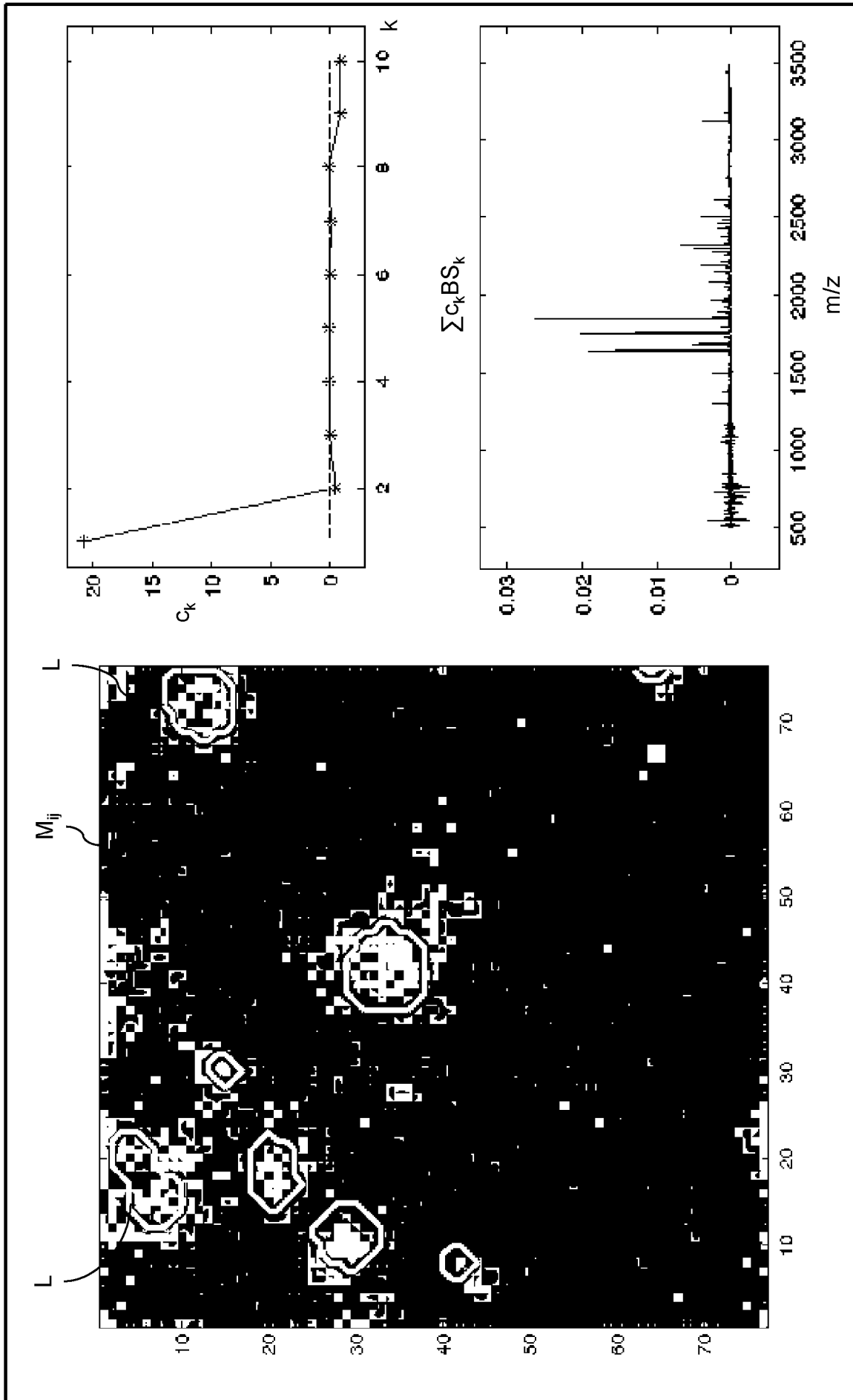


Fig. 4

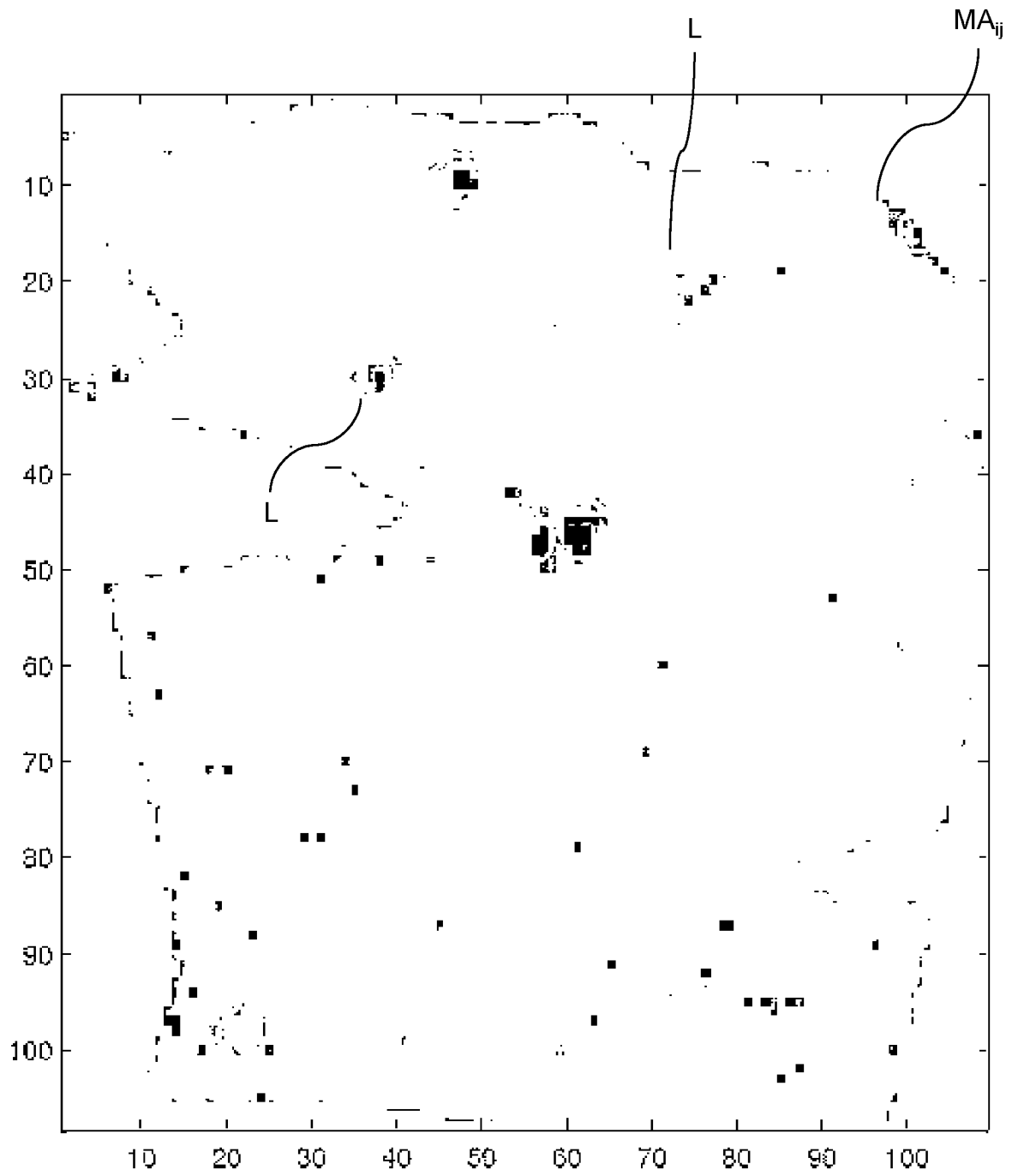


Fig. 5