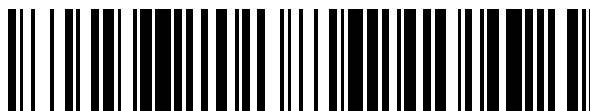


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 578 353**

51 Int. Cl.:

G01J 3/42 (2006.01)

G01N 21/35 (2014.01)

G01J 3/28 (2006.01)

G01N 21/27 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.10.2011 E 11782047 (2)**

97 Fecha y número de publicación de la concesión europea: **30.03.2016 EP 2635882**

54 Título: **Procedimiento para la determinación de componentes químicos de sustancias sólidas o líquidas con ayuda de la espectroscopía THz**

30 Prioridad:

04.11.2010 DE 102010050198

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

26.07.2016

73 Titular/es:

**HÜBNER GMBH & CO. KG (100.0%)
Heinrich-Hertz-Strasse 2
34123 Kassel, DE**

72 Inventor/es:

**HEISE, MICHAEL;
LAMPEN, PETER;
NALPANTIDIS, KONSTANTINOS y
PLATTE, FRANK**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 578 353 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Procedimiento para la determinación de componentes químicos de sustancias sólidas o líquidas con ayuda de la espectroscopía THz

5 La invención se refiere a un procedimiento para la determinación de componentes químicos de sustancias sólidas o líquidas con ayuda de la espectroscopía THz, en la que las sustancias son puestas en oscilación mediante la radiación THz y al menos se detectan las partes de radiación reflejadas y se evalúa el espectro de reflexión respectivo.

10 La espectroscopía Terahertz ha adquirido en los últimos años una importancia particular, dado que para este intervalo espectral se desarrollaron nuevas fuentes de radiación y detectores. Para el intervalo espectral conocido desde hace muchos decenios como infrarrojo lejano se posibilitaron nuevas aplicaciones en la analítica química basada en espectroscopía molecular y para la caracterización de los materiales más diversos ("Terahertz Spectroscopy; Henry M. Mantsch, Dieter Naumann, The renaissance of far infrared spectroscopy, Journal of Molecular Structure 964 (2010) 1-4; S. L. Dexheimer (Comp.), Terahertz Spectroscopy: Principles and Applications; CRS Press, Nueva York, 2007"). En particular, desde hace algunos años está disponible la denominada
15 espectroscopía en el dominio del tiempo (TDS) con la que, a diferencia de la espectroscopía infrarroja convencional, junto a la amplitud también se mide la fase del pulso de terahercios, con lo que junto a los coeficientes de absorción espectrales se puede determinar también, mediante procedimientos matemáticos adecuados, el índice de refracción espectral de la muestra.

20 Si se irradia una muestra mediante la radiación de una fuente de irradiación THz (o fuente de radiación IR lejana), entonces se produce una interacción entre la muestra y la radiación. La radiación que se produce de determinadas frecuencias puede ser transmitida, reflejada o absorbida, en donde, en función de las propiedades químicas y ópticas, las porciones pueden ser diferentes. Si se registran las porciones de radiación que interactúan con la muestra a lo largo de un intervalo de frecuencias mayor (también conocido como intervalo espectral), entonces se obtienen para la sustancia atenuaciones dependientes de la frecuencia características de las intensidades de
25 radiación espectral irradiadas, que están a disposición como la denominada huella digital espectral. Si en las muestras se han de realizar mediciones de transmisión con el fin de analizar éstas, entonces las muestras deben presentar propiedades de transmisión suficientes, con el fin de que puedan detectarse todavía porciones de radiación suficientemente elevadas para una medición de los datos espectrales. En el caso de sustancias no transparentes, se pueden añadir por mezcla a las muestras aditivos para la dilución, con el fin de reducir la absorción en un intervalo de frecuencias determinado. Otra posibilidad para la caracterización química de sustancias es el análisis de espectros de reflexión de la muestra irradiada. Una ventaja de este procedimiento es que ya no es necesaria una preparación de la muestra. Si este tipo de muestras se miden a una distancia de metros bajo condiciones atmosféricas normales, entonces se produce una transposición perturbadora de las señales de medición por parte de la absorción de radiación del vapor de agua atmosférico que, en muchos intervalos espectrales
30 estrechos, puede conducir a una absorción completa de la radiación y reducir considerablemente la calidad de los espectros de la sustancia a identificar.

35 Para la medición de espectros de reflexión de sustancias peligrosas tales como, p. ej., explosivos, se conocen diferentes soluciones. En la revista "Appl. Phys. Lett. 93, 081906 (2008), M. Ortolani, J.S. Lee, U. Schade y H.-W. Hübers, Surface roughness effects on the terahertz reflectance of pure explosive materials") se describen espectros de reflexión de elevada calidad, pero para ello son necesarias fuentes de radiación de un sincrotrón que, bajo condiciones en la práctica, no está disponible. En la publicación "Appl. Phys. Lett. 86, 241116 (2005), Y.C. Shen, T. Lo, P.F. Taday, B.E. Cole, W.R. Tribe y M.C. Kemp, Detection and identification of explosives using terahertz pulsed spectroscopic imaging") se describen espectros de reflexión de RDIX que, sin embargo, no se midieron bajo condiciones de medición atmosféricas difíciles.

45 En el documento US 2007/0114419 A1 se representa un escenario de medición en el que incluso se representan, como posibles, distancias de 50 m y superiores. La viabilidad se investigó con ayuda de cuatro sustancias explosivas diferentes (los espectros medidos en la reflexión y transmisión) y otras sustancias tales como jabón, harina, sal y azúcar. Se llevó a cabo una clasificación de las sustancias explosivas con ayuda de una reducción de las propiedades (a saber reducción de variables) y subsiguiente análisis de los componentes principales (PCA) de los espectros. Se pretendieron representaciones de componentes principales bidimensionales o bien
50 tridimensionales para el análisis de la agrupación en relación con el grupo de sustancias explosivas. En este caso, se ha de recurrir a criterios de probabilidad estadísticos con el fin de estimar la pertenencia a las sustancias críticas buscadas. La quimiometría presentada en este documento es estado conocido de la técnica. Clasificadores sencillos

no se adecúan, sin embargo, para la difícil misión del reconocimiento del modelo con ayuda de los espectros THz de las clases de sustancias mencionadas, registrados bajo condiciones límites malas. Mediante diferentes espectros superpuestos, p. ej., mediante porciones de espectros de prendas de vestir superpuestas por encima de las sustancias críticas, los espectros THz de las muestras, después del tratamiento previo de los espectros y de una
5 reducción de las propiedades, no forman obligatoriamente zonas continuas en el espacio dimensional bajo. Más bien, existen muchas zonas que han de asociarse a la misma clase de sustancias. Además, este procedimiento tampoco tiene en cuenta la dispersión en el caso de mediciones repetitivas.

En el documento US 2009/0101843 A1 se representa un modo de proceder matemático para el tratamiento previo de datos espectrales. En este caso, junto a otros espectros se mencionan, entre otros, también espectros de THz.
10 Como aplicación particular se menciona la espectroscopía de fluorescencia y de por vida.

El tratamiento previo se dirige a un análisis de los componentes principales (PCA) de los espectros de fondo, cuyos vectores se utilizan para la proyección del espectro de muestras en el espacio complementario de los espectros de fondo. Los espectros de muestras purificados de esta forma de componentes perturbadores pueden compararse entonces con un espectro de biblioteca tratado de igual manera de la sustancia buscada, en base a medidas de
15 similitud. Por ejemplo, se utiliza el producto a escala vectorial, cuyo valor se compara con un valor umbral, oscilando los valores para los vectores normalizados en longitud entre 0 y 1 (colinealidad completa en el caso de este último).

Correcciones de señales ortogonales en el caso de calibraciones multivariantes no son particularmente buenas, dado que, en última instancia, la corrección de la dimensión no conduce a mejoras en la calidad de predicción, dado que los datos en bruto sólo requieren, por ejemplo, otros factores PLS que implícitamente conducen a una
20 proyección con radiación de las sensibilidades transversales. Esto se puede también comprender con el tratado de la teoría del analito de la red ("C.D. Brown, Discordance between Net Analyte Signal Theory and Practical Multivariate Calibration, Anal. Chem. 2004, 76, 4364-4373"). Otro modo de proceder en el que, a priori, se utiliza el conocimiento sobre la forma del espectro de muestra buscado, es la calibración basada en la ciencia (SBC) que tiene en cuenta una modelación implícita de las señales perturbadoras de fondo de un modo estadístico ("H.H. Heise, P. Lampen, R. Marbach, Near-infrared reflection spectroscopy for non-invasive monitoring of glucose –
25 established and novel strategies for multivariate calibration, Handbook of Optical Sensing of Glucose in Biological fluids and Tissues, V.V. Tuchin (Comp.) CRS Press, Capítulo 5 (2008)"). Un estado de la técnica adicional se da a conocer en Baker et al., Proc. SPIE Vol. 5790 (2005), páginas 1-10.

Hasta la fecha, no se conoce una solución satisfactoria para la caracterización automática y la clasificación de
30 sustancias sólidas o líquidas con ayuda de la espectroscopía THz, bajo condiciones límite difíciles, a las concentraciones de vapor de agua atmosféricas más diversas. En este caso, las sustancias a detectar pueden ser, p. ej., cubiertas por una o varias capas de materiales desconocidos en donde la atmósfera en el paso de la radiación, entre la fuente de radiación, el objeto de medición y el detector, presenta una humedad del aire y una temperatura variables, y también la longitud del paso de la radiación puede ser diferente. Condiciones límite difíciles
35 de este tipo se manifiestan, p. ej., en el sector del control de seguridad, en donde personas pueden llevar en el cuerpo sustancias peligrosas por debajo de una prenda de vestir. Como escenario adicional, una sustancia no permitida puede ser escondida en el equipaje entre prendas de vestir.

Misión de la invención es proporcionar un procedimiento para la caracterización automatizada y la clasificación de
40 sustancias sólidas o líquidas con ayuda de la espectroscopía THz bajo condiciones límite difíciles a las concentraciones de vapor de agua atmosféricas más diversas.

Este problema se resuelve en el caso de un procedimiento del tipo mencionado al comienzo, de acuerdo con la invención, mediante las características especificadas en la reivindicación 1. Antes de la evaluación del espectro de reflexión respectivo se somete al espectro bruto en cada caso detectado a un tratamiento previo, eliminándose en
45 una primera etapa de tratamiento previo la influencia de las bandas de agua en el espectro bruto respectivo mediante filtros de pico, en una segunda etapa de tratamiento previo se somete al espectro bruto respectivo a una corrección de la línea base, y en una tercera etapa de tratamiento previo se eliminan mediante filtración porciones de alta frecuencia del espectro bruto.

De este modo, es posible eliminar informaciones no necesarias que están contenidas en el espectro bruto y obtener, en la medida de lo posible, sólo las informaciones que se basan en las propiedades (químicas) de las muestras
50 medidas. En este caso, se tiene en cuenta que los materiales tejidos de las prendas de vestir reducen muy fuertemente las intensidades de radiación que se utilizan para la medición de la muestra a investigar, de modo que las intensidades de señal que llegan al detector pueden ser muy pequeñas y, por consiguiente, pueden presentar una razón señal-ruido muy desfavorable. Señales perturbadoras adicionales pueden proceder de espectros

- 5 superpuestos a base de las partes de reflexión de materiales tejidos desconocidos con una composición de productos tales como algodón, viscosa, cuero y muchos más, También la naturaleza física de la superficie (superficie de la textura, p. ej. tejida, lisa, etc.) puede dificultar fuertemente o imposibilitar la caracterización espectral de las sustancias propiamente dichas. Otra influencia, no a subestimar, la representa la humedad relativa del entorno. Las bandas de agua atmosféricas perturbadoras que, se producen mediante absorción, pueden ser en el espacio de frecuencias tan elevadas que la calidad de los espectros de reflexión se vuelve asimismo peor y, en parte, intervalos de frecuencia completos se vuelven inservibles (mediante el denominado recorte de líneas). Además, los espectros de reflexión de los sólidos dependen también de la granulometría del sólido respectivo, y también influye el ángulo de la radiación incidente.
- 10 En una ejecución ventajosa, está previsto que la señal bruta detectada en cada caso antes del tratamiento previo sea transformada en el intervalo de frecuencias. Naturalmente, esto sólo se requiere en el caso de utilizar espectrómetros que miden el intervalo de tiempo (p. ej., en el caso de la THz-TDS). Entonces, los interferogramas o bien las formas de ondas deben ser transformados primeramente en el intervalo de frecuencias, lo cual puede llevarse a cabo preferiblemente con ayuda de una transformación de Fourier que proporciona los espectros brutos.
- 15 Con el fin de eliminar lo más ampliamente posible en una etapa previa informaciones espectrales perturbadoras y no necesarias, p. ej., condicionadas por la característica del espectrómetro de THz y por la influencia de la humedad del aire, está previsto, además, preferiblemente, que el espectro bruto respectivo sea dividido antes del tratamiento previo mediante un espectro de referencia del espectrómetro utilizado.
- 20 Con el fin de llevar a cabo la corrección de la línea base en el tratamiento previo, está previsto preferiblemente que para la corrección de la línea base del espectro bruto respectivo se utilice un filtro pasabajos o que el espectro bruto respectivo sea dividido por el espectro obtenido con el valor medio deslizante, o del espectro bruto respectivo, o sea sustraída la función tratada de igual manera.
- 25 Además, puede estar previsto que el espectro previamente tratado respectivo sea transformado con ayuda de una transformación de Kramers-Kronig. La ventaja de esta transformación estriba en que las informaciones espectrales en los espectros de reflexión se convierten en bandas de absorción. En el caso ideal, se produce una mayor densidad de información en los espectros. Signaturas de dispersión presentes en los espectros de reflexión se pueden convertir en bandas de absorción más estrechas.
- 30 Los espectros se componen, también después del tratamiento previo, todavía de varios cientos de pares de valores de grado de frecuencia y reflexión. Para la clasificación, no se requiere, por una parte, un espacio dimensional tan elevado y, por otra parte, la cantidad de datos ralentiza el análisis. Por lo tanto, se prefiere, además, prever que con el espectro previamente tratado respectivo se lleve a cabo una reducción de las propiedades, p. ej., mediante análisis de los componentes principales en un espacio dimensional bajo. La representación de los datos originales en el espacio transformado, que es abarcado por los primeros vectores del componente principal, muestra a menudo ya una separación visible de los datos. El clasificador propiamente dicho define zonas (geométricas) en un espacio
- 35 la mayoría de las veces dimensional bajo al que se deben de asociar sustancias de diferentes clases. En este caso, se utilizan preferiblemente para la clasificación, clasificadores QDC (clasificador discriminante cuadrático). También se pueden utilizar clasificadores de modelo mixto de Parzen o Gauss.
- 40 El procedimiento de acuerdo con la invención es particularmente adecuado para uso en la vigilancia y la defensa de peligros en aeropuertos, estaciones ferroviarias, administraciones, sectores militares, en eventos multitudinarios o similares.
- Puede utilizarse para la detección o bien identificación de sustancias o bien materiales, en particular de sustancias peligrosas tales como explosivos o drogas. Además, pueden detectarse, reconocerse o bien identificarse de este modo básicamente materiales que se encuentran detrás de o bien en otros materiales.
- 45 Además, el procedimiento puede utilizarse para el control de personas, control de bultos, control de paquetes postales o control de embalaje.
- La invención se explica con más detalle a modo de ejemplo en lo que sigue con ayuda de los dibujos. Estos muestran en
- la Fig. 1 un diagrama de operaciones de una forma de realización preferida del procedimiento de acuerdo con la invención,

- la Fig. 2a un espectro bruto de amplitud de una medición de referencia con y sin humedad del aire,
- la Fig. 2b un espectro bruto de amplitud de una medición de la muestra con y sin humedad del aire,
- la Fig. 3a un espectro de reflexión de una muestra sin materiales tejidos,
- la Fig. 3b un espectro de reflexión de una muestra con material tejido cubierto,
- 5 la Fig. 4a el espectro conforme a la Fig. 3a, después de la eliminación de los picos condicionados por las bandas de agua y establecimiento del intervalo espectral,
- la Fig. 4b el espectro correspondiente al de la Fig. 3b para materiales puros con vestimenta después de un tratamiento previo correspondiente,
- las Figs. 5a
- 10 y 5b la normalización de los espectros conforme a las Figs. 4a y 4b,
- las Figs. 6a
- y 6b el resultado de la eliminación de niveles de ruido de alta frecuencia mediante filtración en los espectros según las Figs. 5a y 5b,
- la Fig. 7 la optimización del parámetro para la corrección de la línea base,
- 15 la Fig. 8 la optimización del parámetro para la filtración de los espectros y en la
- la Fig. 9 la elección del número óptimo de componentes principales.

En la Fig. 1 se representa el transcurso fundamental de una forma de realización preferida del procedimiento de acuerdo con la invención desde el registro de un espectro bruto a través del tratamiento previo de los espectros hasta la clasificación de las sustancias.

- 20 El objetivo del tratamiento previo de los espectros es eliminar todas las informaciones no necesarias en el espectro y contener en la medida de lo posible sólo las informaciones que se basen en las propiedades (químicas) de la muestra medida.

- 25 En el caso de espectrómetros que miden en el intervalo de tiempo, como en el caso de los THz-TDS, interferogramas o bien las formas de ondas deben transformarse primeramente en el espacio de la frecuencia. Esto se lleva a cabo habitualmente con ayuda de una transformación de Fourier.

En la Fig. 2a está representado el transcurso cualitativo de dos espectros de referencia (sin muestra) con y sin humedad del aire. Frente a ello, en la Fig. 2b están recogidos dos espectros de una sustancia de ensayo, medidos de nuevo con humedad del aire presente y con una atmósfera seca. Las bandas de agua inferiores grises relativamente anchas se pueden ver tanto en la medición de referencia como también en la medición de la muestra.

- 30 Mediante la división del espectro de la muestra mediante un espectro de referencia se eliminan ampliamente de forma matemática las informaciones espectrales perturbadoras y no necesarias – condicionadas por la característica del espectrómetro THz y por la influencia de la humedad del aire -. Mediante precisiones de medición y condiciones variables del entorno (humedad del aire variable), las bandas de agua atmosféricas se compensan y se separan por
- 35 filtración de una manera cada vez peor, en particular en el caso de frecuencias elevadas. En las Figs. 3a y 3b se pueden ver en todos los espectros las puntas (“picos”) resultantes mediante la formación de cocientes del espectro y perturbadoras. Están representados tres espectros de la sustancia de ensayo no cubierta ASS (ácido acetilsalicílico). La capacidad de reproducción de los espectros de la sustancia investigada es extremadamente mala en virtud de la mala razón señal-ruido, del fondo espectral fuertemente variable y de los picos muy agudos. Dificultades todavía mayores las ofrecen mediciones de sustancias que están cubiertas por diferentes materiales. El fondo espectral
- 40 varía aquí todavía más intensamente que en el caso de las sustancias puras, dado que los materiales absorben con diferente intensidad en función de la frecuencia (Fig. 3b para la sustancia de ensayo PABA (abreviatura inglesa para p-Aminobenzoic Acid, traducido ácido p-aminobenzoico). Además, los materiales examinados al trasluz pueden

conducir también a bandas adicionales en el espectro que dificultan adicionalmente una caracterización y clasificación de la sustancia investigada.

La eliminación de los picos condicionados por las bandas de agua tiene lugar bien sólo en el intervalo de hasta 2,1 THz, y esto también sólo cuando el espectro de referencia fue registrado en el caso de humedades del aire iguales. Las informaciones espectrales por encima de 2,1 THz pueden ser despreciadas o son extrapoladas de forma constante. En las Figs. 4a y 4b se representa el resultado de las dos medidas de tratamiento previo para los espectros de sustancias puras (Fig. 4a) y de sustancias puras con prenda de vestir (Fig. 4b).

Los picos pueden definirse como criterio matemático. En un pico, dos puntos consecutivos tienen una pendiente claramente mayor que la modificación media de puntos consecutivos. Un vector y contiene, por ejemplo, un espectro. Cuando se cumple

$$|y(i+1)-y(i)| / \Delta > \text{umbral} \quad (\text{PICO } 1)$$

entonces, en el caso de $(i+1)$ está presente un pico. El valor "umbral" se ha de elegir de manera específica para el problema, pero habitualmente siempre es igual en el caso de la misma instalación de medición. El "delta" puede, pero no tiene que ser constante. En particular, en el caso de una distribución homogénea en las abscisas, se puede trabajar con un "delta" constante. El criterio (PICO 1) se aplica a lo largo de todos los puntos. Un acceso pragmático alternativo consiste en almacenar todas las pendientes en un vector y , a continuación, clasificarlas en función del tamaño. El experto en la materia reconocerá mediante picos qué porción del espectro no es aprovechable. Cuando se determinó esta porción, pueden eliminarse a continuación simplemente, p. ej., las pendientes mayores del 30%.

El objetivo es elegir el porcentaje de modo que desaparezcan en la medida de lo posible todos los picos, pero no se eliminen innecesariamente muchos puntos en los que no existan picos. Por eliminación se quiere dar a entender en el caso de los filtros de picos que los puntos en los que están presentes picos queden desatendidos simplemente. No se interpolan nuevos valores. Otros métodos se describen en Phillips, G. R.; Harris, J. M. *Anal. Chem.* 1990, 62, 2351-2357 y Hill, W. y Rogalia, D.: Spike-correction of weak signals from charge-coupled devices and its application to Raman spectroscopy *Anal. Chem.*, 1992, 64 (21), págs. 2575-2579, en los que los puntos de pico son rellenados. Cuando los puntos en los que se forman picos son siempre los mismos en todos los espectros, también se puede trabajar con listas.

La información espectral (buscada) (Figs. 4a, 4b) se encuentra en un intervalo de frecuencias medio. La variación de las líneas base puede interpretarse como perturbación de baja frecuencia, mientras que las perturbaciones de alta frecuencia en el espectro se plasman como una razón señal-ruido mala. Con el fin de hacer equiparables entre sí los espectros, sin borrar la información de sustancia espectral buscada, se determina primeramente el transcurso de la línea base a través de un filtro pasabajos. La corrección tiene lugar sustrayendo la línea base del espectro de partida o dividiendo el espectro de partida por la línea base. Ambos modos de proceder normalizan los espectros presentes. Las Figs. 5a y 5b muestran la elevada calidad de reproducibilidad de sustancias iguales, sin (Fig. 5a) y con (Fig. 5b) materiales tejidos en el caso de utilizar el filtro pasabajos especialmente adaptado.

Básicamente, existen diferentes métodos para la corrección de la línea base. En el caso más sencillo, los espectros son transformados mediante adición o sustracción de líneas base en la zona positiva. Alternativamente, todos los espectros pueden elaborarse también de manera que los espectros con sus mínimos locales sean desplazados en cada caso a las abscisas, es decir, éstos representen correspondientes valores cero. Muchos programas ofrecen también la posibilidad de elegir varios puntos en el espectro que luego se describen mediante un transcurso individual lineal o deslizante (p. ej., mediante funciones segmentarias) y representan la línea base. En este caso, los puntos elegidos son la mayoría de las veces los mínimos locales de los espectros. El método está ligado, sin embargo, a demasiado trabajo manual y dificulta una automatización. Planteamientos más productivos utilizan, p. ej., un filtro pasabajos para calcular la tendencia del espectro. Esta tendencia es la línea base. Para la determinación de la tendencia se adecúan de nuevo el método del valor medio deslizante, la regresión de una función polinómica a través del espectro o también la regresión local según Savitzky-Golay (Savitzky A. Golay, M.J.E., *Smoothing and Differentiation of Data by Simplified Least Squares Procedures*, Analytical Chemistry, Volumen 36, 1627-1639 (1964)). Como última etapa del tratamiento previo de los espectros se lleva a cabo todavía una filtración de los espectros que elimina las porciones de fluido de alta frecuencia y perturbadoras.

El método matemáticamente más sencillo de la filtración es la aplicación del valor medio deslizante (moving average). Esto conduce ya a menudo a resultados muy buenos. El valor medio deslizante puede considerarse como un tipo de filtro pasabajos. Los otros métodos de la regresión se basan en una reacción local de los datos. Ésta se puede basar en funciones de forma de primer orden (lowess) Cleveland, W.S. Lowess: a program for smoothing

scatterplots by robust locally weighted regression, American Statistician, Volumen 35, Ejemplar 1, 54 (1981) o de segundo orden (loess). Derivadas de estos dos métodos se encuentran también a disposición todavía variantes que proveen a datos atípicos de la serie de datos de un peso menor y, con ello, son más robustos (rlowess, rloess). Otra variante generalizada es aquella según Savitzky-Golay (sgolay) que se basa asimismo en una regresión local con una función de forma la mayoría de las veces cuadrática (Orfanidis, S.J., Introduction to Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, (1996), ISBN y Savitzky, A. Golay, M.J.E., Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry, Volumen 36, 1627-1639, (1964)). Una propiedad ventajosa de este método es que los valores extremos locales se mantienen mejor. Nuevos métodos de filtración en la técnica de filtración trabajan según el método wavelet (Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D., Wavelet Shrinkage: asymptopia, Jour. Roy. Stat. Soc., Serie B, Volumen 57, Ejemplar 2, 301-369 (1995); DeVore, R.A., Jawerth, B., Lucier, B.J., Image compression through wavelet transform coding, IEEE Trans. on Inf. Theory, Volumen 38, Ejemplar 2, 719-746, (1992) y Donoho D.L., Johnstone, I.M. Ideal spatial adaptation by wavelet shrinkage, Biometrika, Volumen 81, 425-455, (1994). Este método se utilizó originalmente para la compresión de señales. La representación jerárquica de la señal de partida en diferentes escalas se puede utilizar, sin embargo, también para la filtración. Lo ventajoso de los métodos wavelet es una libertad muy amplia de parámetros, con lo que el método gana en validez general. No obstante, es de gran importancia para el procedimiento wavelet la elección de la función de forma.

Los espectros elaborados de esta manera pueden ser aportados directamente a la clasificación. Alternativamente, existe la posibilidad de transformar los espectros de reflexión con ayuda de la transformación de Kramers-Kronig (KKT). La ventaja estriba en que las informaciones espectrales en los espectros de reflexión se convierten en características de absorción. En el caso ideal se produce una densidad de información mayor en los espectros. Características de dispersión presentes en los espectros de reflexión se pueden convertir en bandas de absorción más estrechas. Las Figs. 6a y 6b muestran el resultado de la filtración a continuación de los espectros conforme a la Fig. 5a o bien 5b.

Los espectros se componen, también después del tratamiento previo, todavía de varios cientos de pares de valores de frecuencia y grado de reflexión. Para la clasificación no es necesario, por una parte, un espacio dimensional tan elevado y, por otra parte, la cantidad de datos ralentiza el análisis. Por lo tanto, tiene lugar una reducción de la dimensión adicional o también una reducción de las propiedades o bien selección de propiedades (en inglés Feature reduction, feature selection). La reducción de propiedades tiene lugar con ayuda del análisis de componentes principales (en inglés PCA: principal component analysis). La representación de los datos originales en el espacio transformado, que es abarcado por los primeros vectores de componentes principales, muestra a menudo ya una separación visible de los datos. Otra alternativa para la reducción de la dimensión lo representa el denominado mapeo de Fisher (o análisis de discriminación lineal LDA), (véase, p. ej., de Ridder, D.; Loog, M.; Reinders, M.J.T.: Local Fisher embedding, Proceedings of the 17th International Conference on Pattern Recognition, 2004. (Vol.2) págs. 295 – 298 (2004)). El clasificador propiamente dicho define zonas (geométricas) en un espacio la mayoría de las veces de baja dimensión, al que se han de asociar sustancias de diferentes clases. Clasificadores utilizados a menudo son LDC (clasificador discriminante lineal), QDC (clasificador discriminante cuadrático) o Parzen, que se basan todos en distancias o bien probabilidades de permanencia de los puntos de datos. Durante el ensayo, una medición (espectro) recorre los mismos pasos del tratamiento previo de los espectros y la reducción de propiedades. Con ayuda de la posición o bien de una medida de distancia, el clasificador decide qué sustancia se ha de asociar a la medición del ensayo. Esto sucede, por ejemplo, con ayuda de la distancia mínima a la muestra siguiente (conocida).

Las etapas de tratamiento del tratamiento previo de los espectros no están libres de parámetros. Seguidamente se describe la forma en que pueden determinarse los parámetros óptimos de las etapas de tratamiento individuales. El modo de proceder se describe con ayuda de siete series de medición a razón de 10 mediciones con marcadores conocidos. Marcadores conocidos significa a este respecto que se conoce la asociación espectros/sustancia.

La calidad de la predicción se sometió a ensayo con ayuda de dos indicadores de errores.

En el caso del ensayo 1 (eps1) se utilizaron para la clasificación seis de las siete series de medición. La mitad de los espectros se utilizan para la constitución del clasificador (p. ej., mediante análisis de los componentes principales y Parzen) y la otra mitad para el ensayo. El valor eps1 indica qué porción porcentual de los espectros utilizados para los ensayos es asociada de forma falsa positiva o falsa negativa.

El ensayo 2 (eps2) toma para el ensayo espectros de una nueva serie de medición que no se utilizaron en absoluto para el entrenamiento. También aquí, el valor de eps2 indica qué porción porcentual de los espectros utilizados para el ensayo es asociada de forma falsa positiva o falsa negativa.

5 La ventaja de este modo de proceder estriba en que la dispersión de las mediciones dentro de una serie de medición y la dispersión entre las series de medición (otra persona de medición, otras condiciones del entorno) pueden ser examinadas independientemente entre sí. Si se observan sólo los valores de eps1, entonces existe el riesgo del sobre-ajuste del clasificador. El valor de eps1 es un error aparente mientras que eps2 permite una estimación de errores realista.

10 La Fig. 7 muestra la calidad de predicción del clasificador en función de los parámetros para la corrección de la línea base. Se consiguen buenos resultados para el parámetro 1 de 50 hasta 150, lo cual corresponde a la eliminación de frecuencias bajas (véase la etapa de tratamiento previo de las Figs. 3a, 3b a Figs. 4a, 4b). El sobre-ajuste es claramente visible para parámetros que aumentan. El aparente error eps1 permanece siendo pequeño, mientras que los valores de eps2 aumentan.

El segundo parámetro modifica la filtración de los espectros (véanse las Figs. 6a, 6b). La Fig. 8 muestra, con ayuda del transcurso de eps1 y eps2, que en absoluto ninguna filtración y demasiada filtración aumenta la tasa de errores. Se alcanzan buenos resultados para parámetros 2 de 2 a 5, lo cual corresponde a una filtración moderada.

15 Como último parámetro importante se investigó la influencia que tiene la reducción de propiedades sobre la calidad de predicción (parámetro 3, número de los componentes principales utilizados según la PCA). En principio, se puede recurrir también a los componentes principales más elevados para la clasificación. Sin embargo, precisamente en este caso existe el riesgo de un sobre-ajuste, dado que los componentes más elevados no contienen necesariamente informaciones específicas para la sustancia. La Fig. 9 muestra que el error (aparente) eps1 para un número cada vez mayor de componentes principales utilizados se desplaza hacia el 0. El error real eps2 muestra, sin embargo, un intervalo en donde es mínimo. Se alcanzan buenos resultados para el parámetro 3 de 10. En el presente caso los primeros 10 componentes principales han cubierto aprox. el 90% de la varianza espectral total.

25 Los valores de parámetros óptimos antes mencionados no son válidos en general. Más bien, dependen de la instalación de medición, del número de clases consideradas y de espectros por clase y del procedimiento para la filtración y la corrección de la línea base propiamente dicha. La idea fundamental en el modo de proceder en la elección de la clasificación es, sin embargo, válida en general.

Básicamente, el procedimiento se adecúa no sólo para el análisis de espectros de reflexión, sino también de espectros de transmisión.

REIVINDICACIONES

- 5 1. Procedimiento para la determinación de componentes químicos de sustancias sólidas o líquidas con ayuda de la espectroscopía THz en el caso de concentraciones de vapor de agua atmosféricas diferentes, en el que las sustancias son puestas en oscilación mediante la radiación THz y al menos son detectadas las porciones de radiación reflejadas y se evalúa el espectro de reflexión respectivo, en donde primero se divide el espectro bruto detectado en cada caso por un espectro de referencia sin muestra y, a continuación, se eliminan las bandas de agua atmosféricas provocadas por la humedad del aire en el conjunto de radiación entre la fuente de radiación de THz, el objeto de medición y el detector mediante absorción, caracterizado por que después de la división por el espectro de referencia sin muestra se eliminan, en una primera etapa de tratamiento previo, las puntas condicionadas por las
- 10 bandas de agua en el espectro bruto respectivo mediante filtros de picos, en una segunda etapa de tratamiento previo se somete al espectro bruto respectivo a una corrección de la línea base y en una tercera etapa de tratamiento previo las porciones de alta frecuencia del espectro bruto son eliminadas mediante filtración y, a continuación, el espectro, así tratado previamente, se aporta para la evaluación de una clasificación.
- 15 2. Procedimiento según la reivindicación 1, caracterizado por que el espectro bruto detectado en cada caso es transformado antes del tratamiento previo en el intervalo de frecuencias.
3. Procedimiento según una o varias de las reivindicaciones 1 ó 2, caracterizado por que para la corrección de la línea base del espectro bruto respectivo se utiliza un filtro pasabajos, y el espectro bruto respectivo se divide por el espectro obtenido con el valor medio deslizante o se sustrae del espectro bruto respectivo el espectro antes mencionado.
- 20 4. Procedimiento según una o varias de las reivindicaciones 1 a 3, caracterizado por que el espectro pre-tratado respectivo se transforma con ayuda de una transformación de Kramer-Kronig.
5. Procedimiento según una o varias de las reivindicaciones 1 a 4, caracterizado por que con el espectro pretratado respectivo se lleva a cabo una reducción de las propiedades, mediante análisis de los componentes principales a un espacio dimensional más bajo.
- 25 6. Procedimiento según una o varias de las reivindicaciones 1 a 5, caracterizado por que el espectro pretratado respectivo se clasifica mediante clasificadores QDC .
7. Procedimiento según una o varias de las reivindicaciones 1 a 6, caracterizado por que se utiliza para la vigilancia y la defensa de riesgo en aeropuertos, estaciones ferroviarias, administraciones, sectores militares o en eventos multitudinarios.

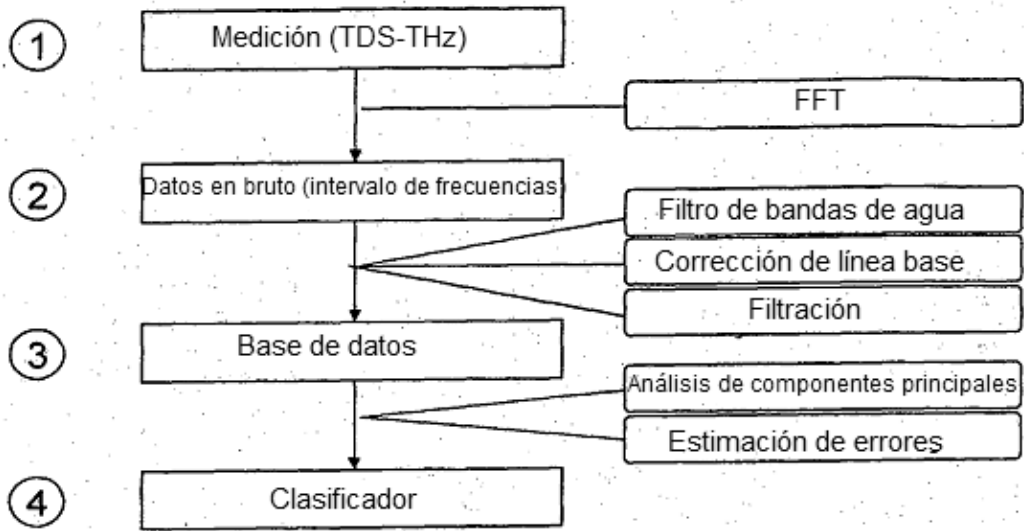


Fig. 1

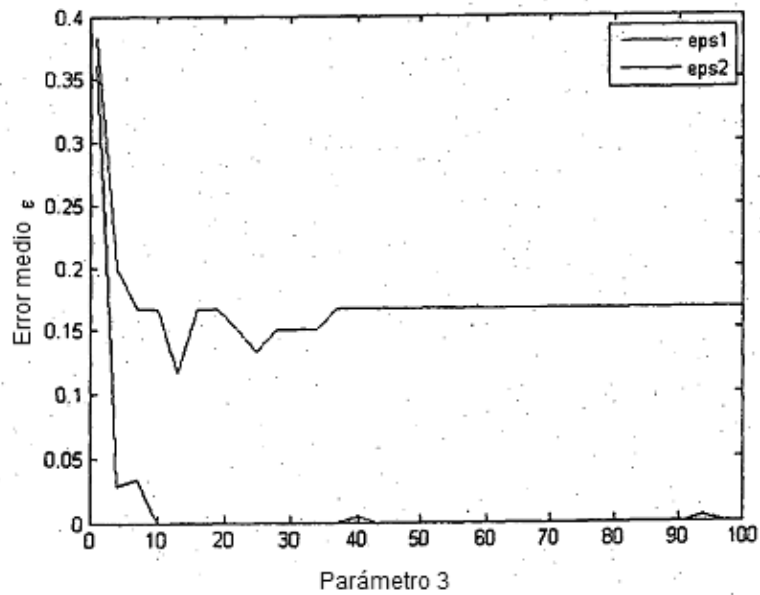


Fig. 9

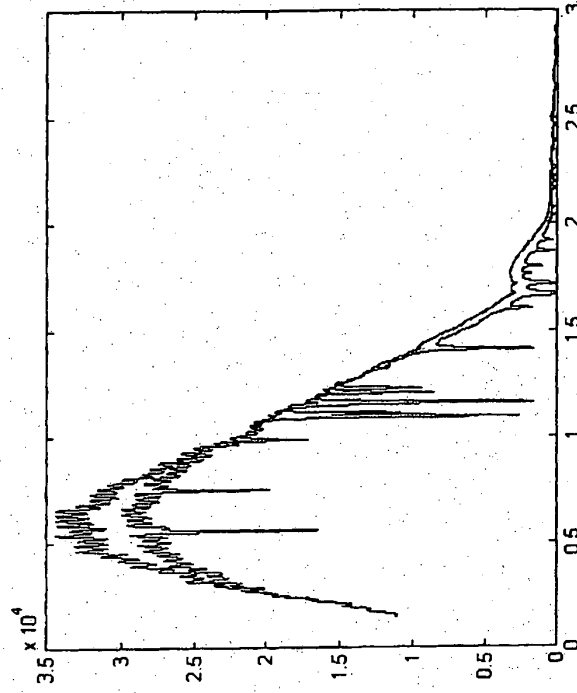


Fig. 2b

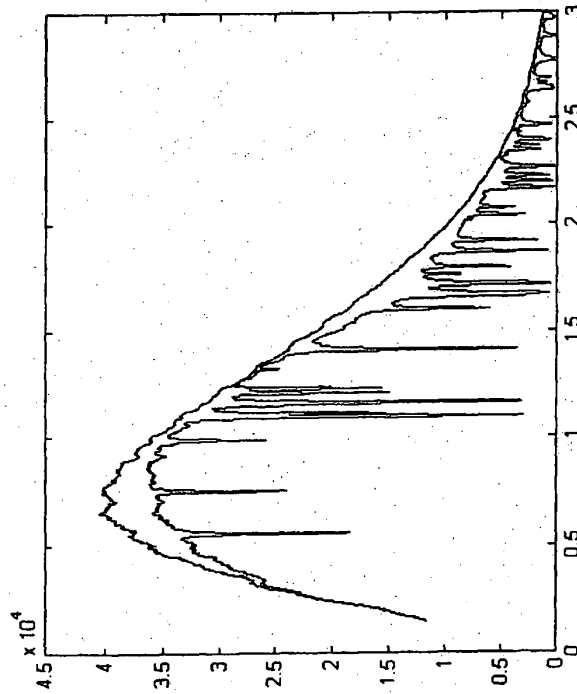


Fig. 2a

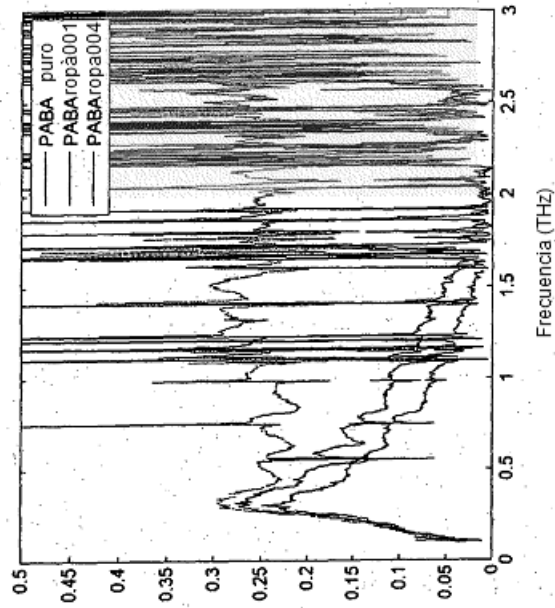


Fig. 3b

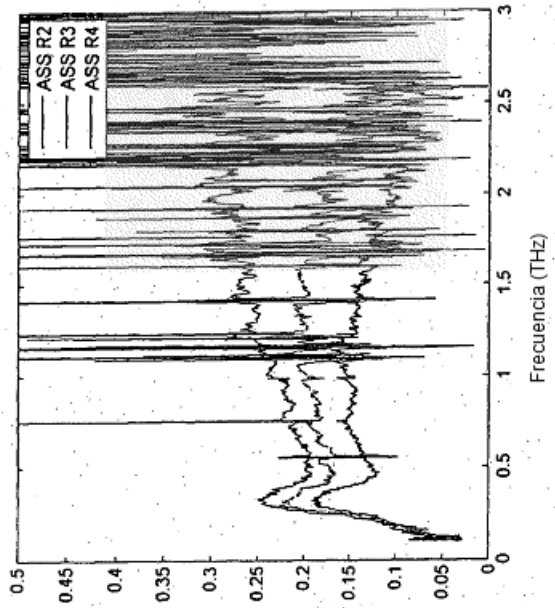


Fig. 3a

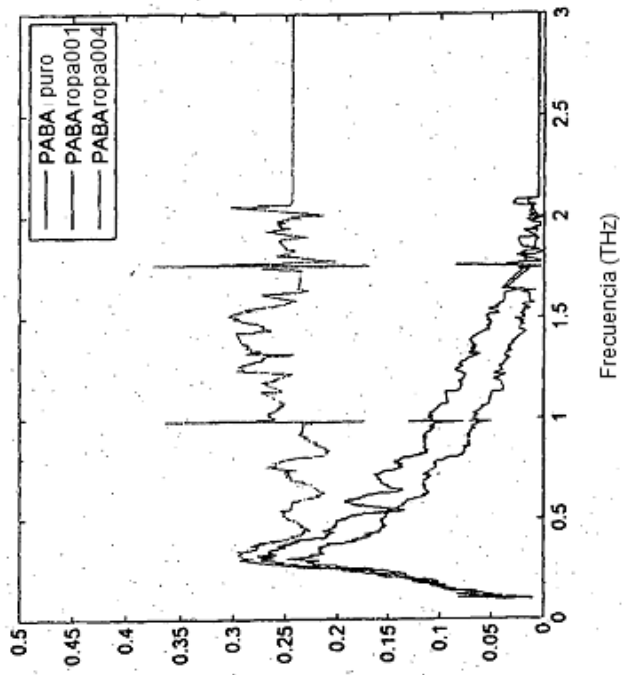


Fig. 4b

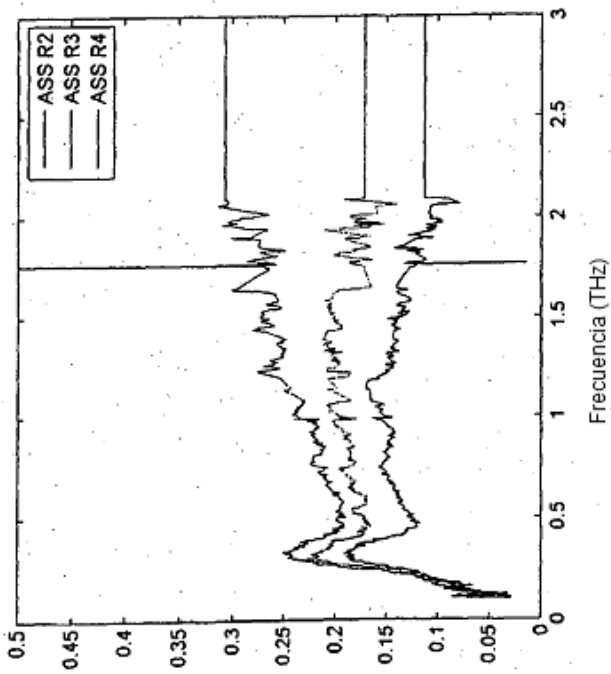


Fig. 4a

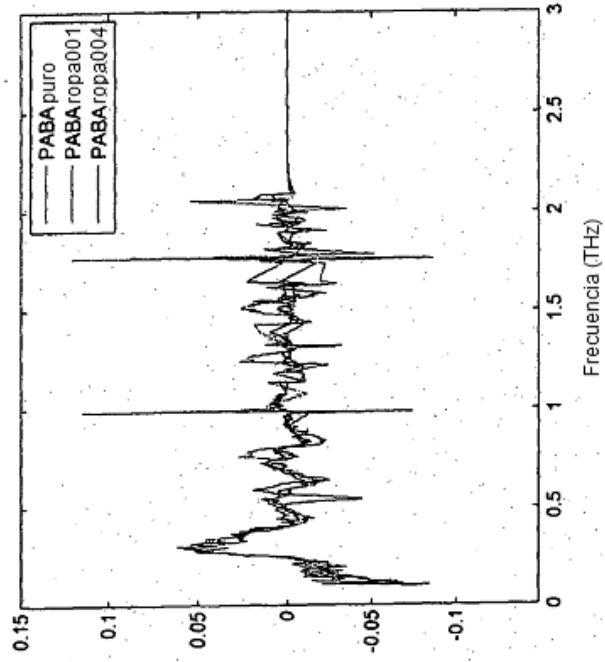


Fig. 5b

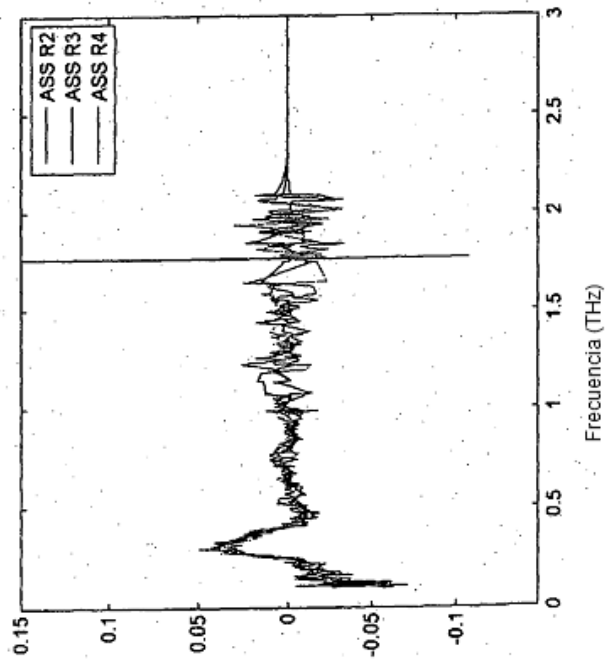


Fig. 5a

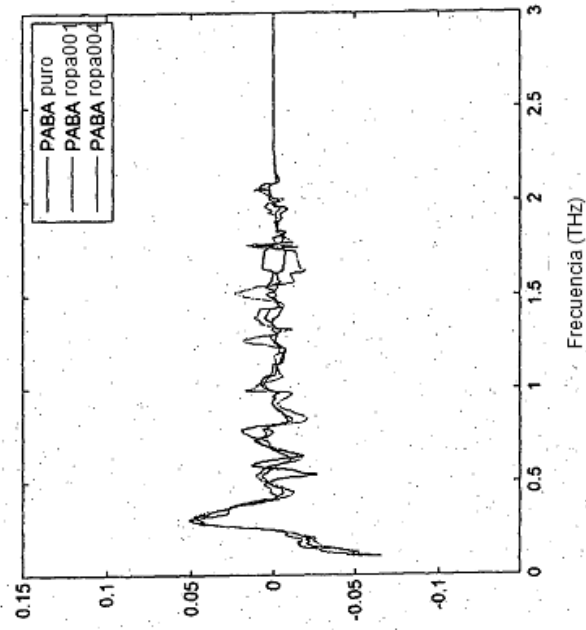


Fig. 6b

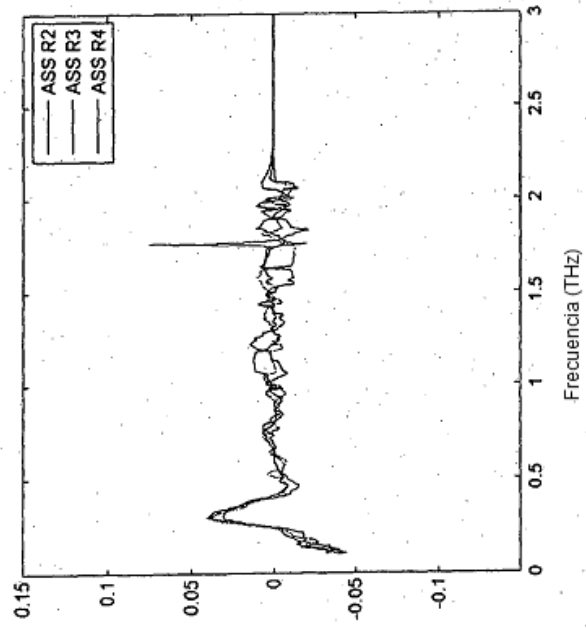


Fig. 6a

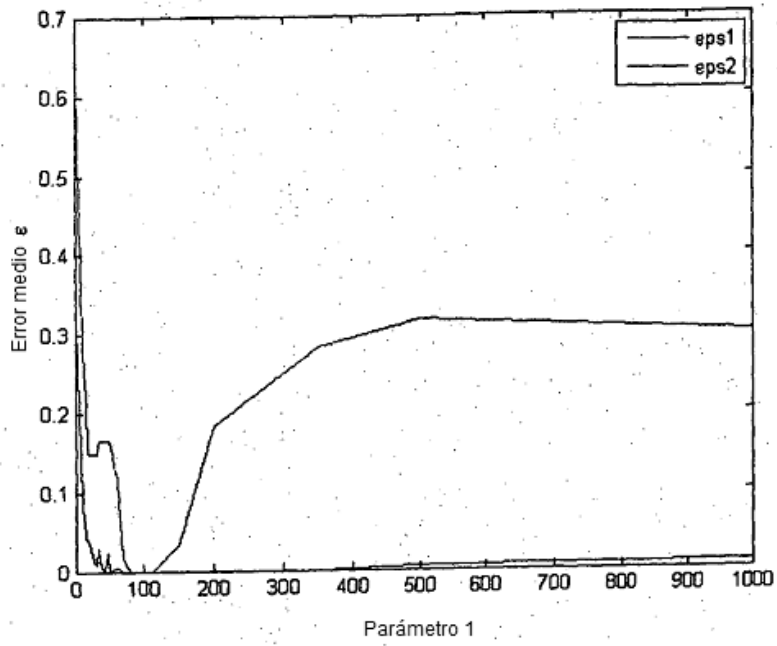


Fig. 7

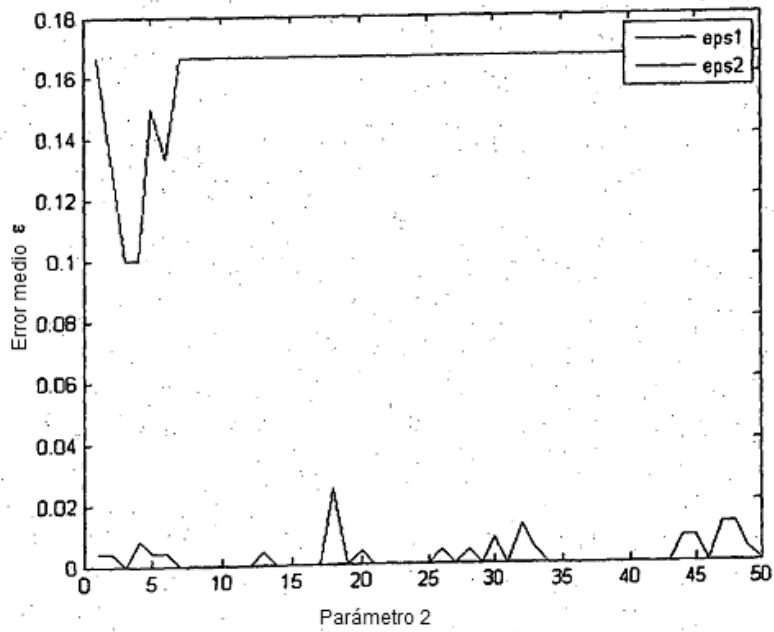


Fig. 8