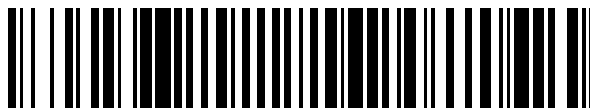


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 401 143**

21 Número de solicitud: 201100407

51 Int. Cl.:

A61B 5/117 (2006.01)

G01N 21/71 (2006.01)

G01J 3/28 (2006.01)

G01J 3/443 (2006.01)

12

SOLICITUD DE PATENTE

A2

22 Fecha de presentación:

07.04.2011

43 Fecha de publicación de la solicitud:

17.04.2013

71 Solicitantes:

**UNIVERSIDAD COMPLUTENSE DE MADRID
(100.0%)**

**AVDA. SENECA 2
28040 MADRID ES**

72 Inventor/es:

**CACERES GIANNI, Jorge;
GILABERT BEGUERIA, Alejandra;
MARCOS MARTINEZ, Daniel;
FERNANDEZ GARCIA, Guadalupe;
MANUEL DE VILLENA RUEDA, Francisco Javier;
ROSALES MARTINEZ, Daniel y
IZQUIERDO HORNILLOS, Roberto**

74 Agente/Representante:

ACEBAL SARABIA, Carmen

54 Título: **DISCRIMINACION DE INDIVIDUOS A PARTIR DE TEJIDO OSEO UTILIZANDO
ESPECTROSCOPIA DE ABLACION LASER Y PROCESAMIENTOS MATEMATICOS.**

57 Resumen:

Discriminación de individuos a partir de tejido óseo utilizando espectroscopía de ablación láser y procesamientos matemáticos.

El método se basa en el análisis de tejido óseo empleando la técnica de espectroscopia de ablación láser combinada con una red neuronal que permite manejar un gran número de datos en un tiempo muy corto.

La metodología empleada proporciona un método adecuado de alta sensibilidad y especificidad en la discriminación de tejido óseo, independientemente de su antigüedad, que permite, de forma rápida, agrupar un conjunto de restos óseos pertenecientes a un individuo y diferenciarlos claramente de otros individuos o restos óseos que pudieran encontrarse junto a los primeros. El método desarrollado tiene aplicación en diversos campos como, por ejemplo, en antropología, arqueología y, principalmente, en determinaciones forenses.

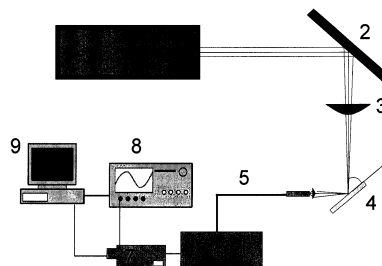


FIGURA 1

DESCRIPCIÓN

Discriminación de individuos a partir de tejido óseo utilizando espectroscopía de ablación láser y procesamientos matemáticos.

5 Objeto y campo de la invención

Esta invención se enmarca en el campo discriminación de individuos a partir de restos óseos mediante análisis espectral de material biológico, concretamente para analizar tejido óseo empleando la técnica de espectroscopía de ablación láser. La metodología empleada proporciona un método adecuado de alta sensibilidad y especificidad en la
10 discriminación de tejido óseo, independientemente de su antigüedad, que permiten agrupar un conjunto de restos óseos pertenecientes a un individuo y diferenciarlos claramente de otros individuos o restos óseos que pudieran encontrarse junto a los primeros. El método desarrollado tiene aplicación en diversos campos como, por ejemplo, en antropología, arqueología y, principalmente, en determinaciones forenses.

15

Estado de la Técnica

Las técnicas clásicas utilizadas en la identificación positiva de restos humanos consisten en el análisis y comparación de huellas dactilares, radiografías esqueléticas o dentales y, en ocasiones, por la presencia de marcas únicas, tales como tatuajes o
20 cicatrices.

Actualmente, en el entorno forense se utilizan análisis genéticos (análisis de ADN) sobre muestras biológicas debido a su objetividad ya que se basan en fundamentos científicos. Sin embargo, este tipo de análisis para determinar a qué individuo
25 pertenece una muestra biológica presenta, fundamentalmente, la limitación del factor tiempo pues los protocolos exigen unos tiempos mínimos de tratamiento de las muestras; además, las muestras a analizar pueden haber sufrido una serie de modificaciones debido al tiempo transcurrido desde su recogida hasta el momento de su estudio, a las condiciones ambientales a las cuales se vieron sometidas, a la
30 cantidad de muestra, al soporte donde se encuentran, al lugar de procedencia o al tipo de muestra.

En relación al tipo de muestra de donde se extrae el ADN, para los cadáveres ya esqueletizados en los que la única muestra disponible son los huesos y los dientes la extracción es más larga y costosa debido a su estado de degradación y contaminación.

- 5 Antes de realizar un estudio de ADN existen pasos previos que permiten discriminar el tipo de resto biológico. Así, se aplican pruebas orientativas (de bajo coste, sencillas de realizar y rápidas que revelan la posible naturaleza de la muestra sin ser concluyentes), pruebas de certeza (que sirven para determinar la naturaleza de la muestra) y pruebas específicas (que sirven para determinar a qué tipo de organismo
10 pertenece la muestra). En algunas ocasiones no es posible realizar este tipo de pruebas preliminares debido a la escasa cantidad de muestra disponible.

Por todo ello, en los casos en que el número de restos es grande, los procedimientos existentes basados en el análisis de ADN para agruparlos pueden representar un alto
15 costo tanto temporal como económico.

En el caso de restos humanos, el análisis elemental de huesos y dientes puede utilizarse para discriminar entre diferentes individuos y para conocer determinados hábitos del individuo y su posible exposición a diferentes condiciones ambientales.

20

En los últimos años, se han planteado otras pruebas alternativas o complementarias al análisis de ADN basadas en técnicas espectrales que realizan un análisis elemental de muestras con aplicación forense. Así, se ha utilizado la emisión de rayos X inducida por partículas (PIXE) para el análisis elemental de huesos en casos de cremación
25 donde se cuestiona la naturaleza de los restos y casos de muerte por bala como herramienta diagnóstica donde la identificación de componentes elementales permite una interpretación más exacta de la evidencia (Warren et al. *Forensic Science Internacional* 125 (2002)). La espectrometría de masas con fuente de plasma de acoplamiento inductivo (LA-ICP-MS) se ha utilizado también con fines forenses para
30 análisis de tierras raras en esqueletos de vertebrados para estudiar su diagénesis (Rogers et al. *PALAIOS*, 2010, v. 25), para el análisis de porcentajes de isótopos de Sr en huesos y dientes enterrados diagenéticamente alterados con el objetivo de

reconstruir procesos de migración humana (Prohaska et al. *J. Anal. At. Spectrom.*, 2002, **17**) o para el análisis de trazas de metales encontrados en fragmentos de huesos o dientes de restos enterrados con el fin de conocer su exposición a fenómenos medioambientales y poder distinguir entre diferentes individuos (Castro et al. *Forensic Science International* **195** (2010)). Aunque estas técnicas permiten obtener una imagen espectral característica de cada resto, parece haber diferencias en la composición elemental de diferentes huesos pertenecientes a un mismo individuo, lo cual hace que la técnica no consiga discriminar totalmente entre diferentes individuos. Por tanto, sería deseable disponer de un método analítico que tenga en cuenta la heterogeneidad que puede darse entre los restos de un mismo individuo para distinguirlo de otro individuo.

En la presente invención se plantea la técnica de espectroscopia de plasma inducido por láser (LIBS) para análisis elemental de restos óseos.

15

La técnica de LIBS (Laser Induced Breakdown Spectroscopy) es un método rápido y versátil y más económico que otras técnicas espectrales, como LA-ICP-MS, para efectuar el análisis de distintos tipos de muestras que pueden ser inaccesibles o tediosos por técnicas analíticas convencionales, siendo particularmente útil en el análisis de muestras con una matriz compleja, (J.O. Cáceres et al. *Spectrochimica Acta B: Atomic Spectroscopy* **56**(6) 831 (2001), WB Lee, et al, *Appl Spectrosc. Rev.* **39**(1) 27 (2004)).

En el análisis cualitativo por LIBS, no se busca una composición química detallada, más bien una identificación instantánea de la muestra utilizando una característica única de LIBS, la cual es capaz de generar una “huella digital” de la muestra. Esta “huella digital” es un espectro LIBS de la muestra, el cual proviene exclusivamente de ésta y por lo tanto depende de su matriz. Además, debido a la naturaleza del espectro de emisión, el cual está dominado por líneas iónicas, que a menudo se ven inhibidas en su relación directa entre la concentración elemental en la muestra y sus intensidades, proporciona un espectro único y solo perteneciente a la muestra bajo análisis.

La técnica LIBS ya ha sido utilizada para el análisis de muestras biológicas de **tejidos blandos de organismos vivos o muertos** (por ejemplo, el documento US 7,092,087 describe el análisis por LIBS muestras de vejiga, colon, endometrio, pulmón, ovario, próstata o recto con el fin de detectar cáncer y Yueh et al. en *Spectrochemica Acta Part B* **64** (2009) describen el uso de LIBS para diferenciar hígado, riñón, cerebro, pulmón, Músculo o bazo en pollos muertos), de **tejidos calcificados de organismos vivos** (por ejemplo, análisis de minerales en pelo de individuos expuestos a contaminantes ambientales según describen Corsi et al. en *Applied Optics* **42**(30), 2003) y Smek et al. en *Spectrochemica Acta Part B* **56**(2001), que describe la aplicación de análisis por LIBS al análisis de minerales y elementos tóxicos en huesos o dientes de niños y adultos con el fin de conocer el efecto de su exposición a distintos ambientes), de **tejidos calcificados de restos fósiles** (por ejemplo, Alvira et al. en *Applied Spectroscopy* **64**(3), 2010 describen el análisis de trazas en dientes de *Homo sapiens* con el fin de conocer sus hábitos o Hrdlickka et al. en *Applied Optics* **49**(13), (2010) describen el análisis remoto de restos óseos en excavaciones arqueológicas o Martin et al. en *Bulletin of Laser and Spectroscopy of India* **14** (2005-2006) describen la aplicación de LIBS al análisis de huesos procedentes de restos de diferentes mamíferos) o de **restos humanos incinerados** (Martin et al. en *Spectrochemica Acta Part B* **62**(2007) donde describe la aplicación de LIBS al análisis de restos humanos incinerados para obtener información sobre un crimen).

Para que los resultados obtenidos del análisis LIBS permitan la discriminación entre diferentes muestras analizadas, es necesario realizar una correlación correcta de los resultados cuantificando el grado de similitud entre dos muestras. Así, utilizando un procedimiento de correlación adecuado el sistema LIBS puede ser entrenado para reconocer espectros procedentes de diferentes muestras, lo cual significa evaluar la similitud de espectros desconocidos con un banco espectral de muestras clasificadas.

Actualmente, se han desarrollado métodos de análisis basados en LIBS combinados con redes neuronales para la clasificación de muestras como, por ejemplo, aleaciones de aluminio (Koujelev et al. *Planetary and Space Science* **58** (2010)) o cerámicas

arqueológicas (Ramil et al. Applied Physics A 92 (2008)). Sin embargo, el número de trabajos encaminados a la aplicación de redes neuronales en LIBS sigue siendo escaso y no correcto para todas las muestras utilizadas o el error asociado a la identificación es mayor al 30%. Por otra parte el campo de las redes neuronales puede mejorarse en
5 muchas direcciones para conseguir la aplicación de LIBS en numerosos campos donde resultaría de gran utilidad la identificación y clasificación eficaz de muestras.

Descripción de la invención

La presente invención consiste en un método rápido y directo para la identificación y
10 discriminación de tejido óseo y/o análisis espectral completo de elementos químicos presentes en estos mediante la combinación de la técnica LIBS con redes neuronales.

El método de análisis se basa en un análisis de un único pulso láser que produce un proceso de vaporización y posterior formación de un plasma de la superficie de la
15 muestra, la obtención del espectro de emisión de este plasma en el orden de unos pocos microsegundos sin ningún tipo de tratamiento previo y la posterior comparación espectral con una base de datos espectral dinámica. Para ello se utilizan algoritmos matemáticos entrenados y la comparación de espectros adecuados para el análisis con los de las matrices espectrales que pueden tener 2048x100 puntos como mínimo y con
20 un tiempo de análisis total menor a 60 (sesenta) segundos.

En esta invención se utiliza un número muy alto de puntos espectrales obtenidos de la emisión del plasma de restos óseos, combinado con un sistema de algoritmos adecuados para manejar un gran número de datos en un tiempo muy corto y de forma
25 eficiente y efectiva, asegurando un 100% de sensibilidad y un 100% de especificidad que permite la caracterización, identificación y discriminación de las muestras óseas.

El dispositivo utilizado para la identificación y caracterización de muestras, mostrado en la figura 1, se compone de un láser pulsado (1), en este caso particular se ha
30 utilizado un láser de Nd:YAG, aunque en la practica podría utilizarse cualquier otro tipo de láser tanto de estado sólido o gaseosos (láser de nitrógeno, láser de dióxido de carbono láser de exímeros, láser OPO, etc.) que permita obtener condiciones

energéticas suficientes para poder producir un plasma. Por este motivo pueden utilizarse láseres tanto en el ultra violeta, visible o infrarrojo.

El láser de Nd:YAG utilizado trabaja a una frecuencia de 1 a 20 Hertz a una longitud
5 de onda fija de 1064 nm respectivamente. Esta longitud de onda no es limitativa para
el uso de otras longitudes de onda, que también puede emitir este láser como 266, 355
ó 532 nm, u otras longitudes de onda producidas por cualquier otro tipo de láser que
permita obtener condiciones energéticas suficientes para poder producir un plasma.
La duración del pulso es de 4 nanosegundos. Dicho láser puede proporcionar hasta
10 180 mJ/pulso de energía de salida. Se han utilizado espejos (2) y lentes (3) adecuados
a fin de focalizar el haz del láser sobre la muestra (4). La emisión del plasma es
colectada utilizando una fibra óptica de 1 metro acoplada al detector, que a su vez es
activado por el pulso láser. El detector es un sensor óptico CCD que proporciona 2048
puntos espectrales en un rango de 200 a 1100 nm. La señal del detector es
15 posteriormente comparada con una base espectral almacenada utilizando un sistema
de algoritmos matemáticos. Una interfaz gráfica permite al usuario el control del
láser, la observación del espectro de la muestra obtenido y la representación de los
resultados para su evaluación.

20 A pesar de que la calidad del espectro y su reproducibilidad pueden variar
considerablemente, no afectan a la capacidad del sistema para la identificación de la
muestra. Esto es así debido a que los espectros utilizados como base para la
comparación reflejan todas las condiciones de medición; por lo tanto, la precisión y
exactitud de la medición del espectro LIBS no es relevante. El sistema es adiestrado
25 para evaluar la similitud de espectros desconocidos con su banco de datos de muestras
clasificadas. Esta característica, en donde los parámetros de ablación, por ejemplo la
distancia lente-muestra, hace menos crítica la calidad del espectro y por lo tanto,
factible la identificación correcta con un único disparo láser.

30 Con el objeto de crear una base de datos espectral dinámica que recoja las “huellas
dactilares” de cada muestra, se utiliza una red neuronal (NN) con conexiones hacia
delante, concretamente se trata de un modelo de propagación *perceptron (back-*

propagating perceptron model). La NN empleada consta de tres capas denominadas de entrada, de salida y ocultas, formadas por neuronas (único elemento operativo). La capa de entrada se utiliza sólo para la entrada de la matriz de datos en la NN. En las otras capas, se realizan cálculos no lineares. En la capa oculta, cada neurona recibe
 5 señales de otras neuronas de entrada, sumándose éstas mediante la función activación, ecuación 1. Después, el resultado es transformado por la función de transferencia, ecuación 2. Finalmente, el resultado es enviado a las neuronas de salida (neuronas en la capa de salida).

$$10 \quad x_k = \sum_{j=1} w_{jk} y_j \quad (1)$$

$$y_k = f(x) = \left(\frac{2}{1 + \exp(-2x_k)} - 1 \right) \quad (2)$$

En las ecuaciones 1 y 2, y_j e y_k representan la salida de las neuronas ocultas (j) y de
 15 las neuronas de salida (k) respectivamente, w_{jk} representa el peso entre el j^{th} de las neuronas ocultas y el k^{th} de las neuronas de salida. Como función de transferencia se ha utilizado el algoritmo tangente hiperbólico sigmoidal, ecuación 2. Los pesos fueron parámetros ajustables de las NN asociados con cada una de las conexiones entre neuronas y estas pueden modificar la señal de comunicación entre éstas. El proceso
 20 de optimización de la matriz de pesos de la NN se llevó a cabo mediante el algoritmo de entrenamiento de gradiente escalado conjugado SCG. Este algoritmo fue seleccionado con el fin de prevenir el sobreajuste (violation of Occam's razor) y el sobreentrenamiento (I.V. Tetko, et al J.Chem. Inf. Coput. Sci. 35, 826 (1995)). El problema del sobreentrenamiento se refiere al hecho de que la red sólo memoriza el
 25 conjunto de aprendizaje y pierde su habilidad para generalizar. Por otra parte, el algoritmo SCG, fue desarrollado para entrenamientos rápidos de sistemas de aprendizaje de redes neuronales (H. Demuth et al, Neural Network Toolbox for use with Matlab User's guide version 4.0.6. ninth printing revised; The math Works: Natick, MA, 2005). Una importante característica de este tipo de algoritmos de alto

rendimiento, es que pueden converger cientos de veces mas rápido que otros tipos de algoritmos, aunque a su vez, consume una gran cantidad de memoria. El SCG minimiza el error de predicción empleando una combinación lineal del error cuadrático medio (MSE), ecuación 3. Esto determina la combinación correcta para producir una red que generalice nuevos datos de entrada dentro del rango de los datos de aprendizaje.

$$MSE = \frac{1}{N} \sum_k^N (r_k - y_k)^2 \quad (3)$$

- 10 En la ecuación 3, N corresponde al número de observaciones, y_k el valor real, r_k las estimaciones en el modelo de NN y k capa de salida de NN, respectivamente.

Los algoritmos básicos de backpropagation ajustan los pesos en una fuerte pendiente de dirección descendente (valores negativos del gradiente), esto es la dirección en la cual la función de rendimiento disminuye más rápidamente. Resulta que, aunque la función disminuye más rápidamente a lo largo de los valores negativos de la pendiente, ésta no produce necesariamente la convergencia más rápida. En el algoritmo de gradiente conjugado, realiza una búsqueda a lo largo de direcciones conjugadas, lo que produce generalmente una convergencia más rápida. (Ver Hagan, M.T., H.B. Demuth, and M.H. Beale, Neural Network Design , Boston, MA: PWS Publishing, 1996. para una discusión más amplia de los algoritmos de gradiente conjugado.)

Cada uno de los algoritmos de gradiente conjugado requiere de una línea de búsqueda en cada interacción. Esta línea de búsqueda es computacionalmente costosa, porque requiere que la red dé respuesta a todas las entradas de entrenamiento varias veces para cada búsqueda. El algoritmo de gradiente conjugado desarrollado por Moller [(M.F. Moller, Neural Networks, 6 525 (1993)), fue diseñado para prevenir el consumo de tiempo de la línea de búsqueda.

- 30 Todos los algoritmos de gradiente conjugado comienzan con la búsqueda de la dirección descendente del gradiente (gradiente negativo) en la primera interacción.

$$p_0 = -g_0 \quad (4)$$

Posteriormente, se realiza una línea de búsqueda, para determinar la distancia optima de movimiento a lo largo de esta dirección de búsqueda:

5

$$x_{k+1} = x_k + \alpha_k p_k \quad (5)$$

El procedimiento general para determinar la nueva dirección de búsqueda es la combinación de la nueva pendiente de dirección con la búsqueda anterior:

10

$$p_k = -g_k + \beta_k p_k - 1 \quad (6)$$

Las distintas versiones de los algoritmos de gradiente conjugado se distinguen por la manera en que la constante β_k es calculada, para el algoritmo SCG el procedimiento de actualización es:

15

$$\beta_k = (|r_{k+1}|^2 - r_{k+1}^r - r_k) / \mu_k \quad (7)$$

Este algoritmo combina el enfoque de región verdadera (usada en el algoritmo de Levenberg-Marquardt (D. Marquardt, J. Appl. Math. 11, 431-441, (1963))), con la aproximación del gradiente conjugado. (ver Moller, para una explicación detallada de este algoritmo).

La topología de la NN empleada en este trabajo consistió en una capa de entrada con 2048 neuronas para las variables de entrada (espectro LIBS). Una capa oculta (optimizada después) y una capa de salida con dos neuronas para determinar si el reconocimiento es positivo o negativo. Esta topología con una sola capa oculta fue adecuada para resolver los problemas de todas las muestras estudiadas. Cada modelo de NN empleado en este trabajo se diseñó utilizando el programa Matlab, seleccionado por su eficacia en el cálculo vectorial.

Dado que la NN utilizada se basa en el algoritmo supervisado, para optimizar la matriz de pesos es necesario emplear datos de entrada y salida que caractericen adecuadamente el proceso que será modelado. Estos datos han sido tomados directamente de 50 espectros LIBS obtenidos a partir de las muestras. Cada espectro
 5 corresponde a un único pulso láser. Tanto las muestras como los datos espectrales obtenidos no tuvieron ningún tipo de tratamiento previo. El conjunto de datos (espectros LIBS de las muestras) fue distribuido al azar en el aprendizaje (80%) y en la verificación (20%) de las muestras, teniendo en cuenta que ningún conjunto de datos o cualquiera de sus réplicas deben ser presentados en la verificación de la
 10 muestra.

La precisión es la principal característica de un procedimiento de reconocimiento como recurso para la toma de decisión, motivo por el cual, las métricas para evaluar procesos de detección tienen una importancia significativa e involucran la frecuencia
 15 relativa de los reconocimientos correctos e incorrectos que hace un observador a partir de los resultados obtenidos. Las medidas básicas son el número de positivos y negativos, (verdaderos y falsos, VP, VN, FP, FN), a partir de los cuales se calculan la sensibilidad (S) y la especificidad (E) y exactitud (Ex) de los procesos de detección.

20 Un verdadero positivo (VP) corresponde a la detección correcta de una sustancia, compuesto o característica en una muestra, cuando esta realmente existe. Un verdadero negativo (VN) corresponde a la detección negativa de una sustancia, compuesto o característica en una muestra cuando efectivamente esta no existe. Las detecciones falsas (FP, FN) corresponden a los casos en los que la detección no
 25 corresponde con la realidad de la muestra.

S y E son dos métricas del desempeño de un proceso de detección que se construyen a partir del número de VP, FP, VN y FN en una muestra de validación. La sensibilidad de un proceso de detección se refiere a la probabilidad de que una sustancia,
 30 compuesto o característica sea detectada cuando realmente existe. La sensibilidad se especifica como una fracción entre 0 y 1, o un porcentaje entre 0 y 100.

La suma de VP y FN corresponde al total de positivos en el proceso de detección así S de un sistema de detección se puede calcular como:

$$S = VP / (VP + FN) \quad (8)$$

5

Una S=1 indica que todas las sustancias, compuestos o características son detectados. S también se denomina Fracción de Verdaderos Positivos (FVP).

La métrica que complementa a la sensibilidad es la especificidad la cual mide la probabilidad de que un proceso de detección reporte correctamente la no existencia de una sustancia, compuesto o característica cuando efectivamente no existe. La suma de VN y FP corresponde al total de falsos en el proceso de detección así E en un sistema de detección se puede calcular como:

$$E = VN / (VN + FP) \quad (10)$$

15

Una E=1 indica que nunca se reporta la existencia de una sustancia, compuesto o característica cuando éste no existe. La Fracción de Falsos Positivos (FFP) esta definida como (1-E) y es la fracción de muestras que se reportan equivocadamente.

Para evaluar un proceso de detección es necesario tener los valores de sus dos métricas: (S y E), ya que una sola métrica no puede evaluar correctamente el proceso. Esto debido a que se puede forzar a S=1 si nuestro sistema de detección reporta todos los casos como positivos (a lo que corresponde a E=0) y también se puede forzar a E=1 si nuestro sistema reporta todos los casos como negativos (a esto corresponde S=0).

20

La exactitud (Ex) es el principal parámetro de reconocimiento en un proceso de toma de decisión, y la razón por la cual las métricas para evaluar el proceso de detección son tan impotentes. *Esta incluye la frecuencia relativa de la identificación correcta e incorrecta de los resultados obtenidos y se puede calcular de acuerdo como:*

25

$$A = (VP + VN) / (VP + VN + FP + FN) \quad (11)$$

El análisis de la Curva Característica de Operación del Receptor (ROC) (del ingles Receiver Operating Chacteristic) desarrollada en el contexto de la detección de

señales electrónicas en los inicios de la década de los 50, encontró una importante aplicación en el proceso de toma de decisión medicas. La curva ROC es la herramienta estándar para mostrar todas las posibles combinaciones de sensibilidad y especificidad de un proceso de detección. Usualmente se gráfica la Fracción de Verdaderos Positivos (FVP) o S en el eje Y, vs la Fracción de Falsos Positivos (FFP) o (1-E) en el eje X. En una curva ROC el punto de operación ideal es la esquina superior izquierda donde FVP=1, FFP=0

Como se muestra en la Figura 3, el procedimiento de identificación permite la completa separación y por lo tanto su identificación para muestras con matrices muy similares. Las curvas ROC de estas muestras muestran que el procedimiento de identificación, funciona perfectamente, y a todos los efectos prácticos puede considerarse que se resuelven con un 100% de sensibilidad y un 100% especificidad.

La nueva metodología permite la identificación del material óseo sin ningún tipo de preparación previa, con un solo disparo láser, en un tiempo menor de 1 segundo. No obstante, esta metodología no reemplaza el procedimiento habitual de identificación de restos óseos por odontología y/o estomatología forense, antropología forense, genética forense o ADN sino que es una herramienta útil que permite agrupar un conjunto de restos óseos facilitando la tarea de reconstrucción de un esqueleto o del material óseo encontrado, con un coste considerablemente inferior al análisis de ADN y que puede ser realizada por personal no cualificado.

DESCRIPCIÓN DE LAS FIGURAS

La figura 1 muestra una vista esquematizada del dispositivo comprendiendo tanto el equipo láser (1), los espejos (2), lentes (3), muestra (4), fibra óptica (5), monocromador (6), controlador de retrasos de pulso (7), osciloscopio (8) y ordenador personal (9).

La figura 2 muestra los espectros LIBS de las muestra óseas de tres individuos, clasificadas como H11 (a), H21 (b) y H31 (c), donde se ha utilizado una nomenclatura basada la letra H y dos números (x e y) de la forma: Hxy donde x

identifica al individuo e y identifica al hueso utilizado. En todos los casos el hueso 1 corresponde a la primera costilla derecha de cada individuo (ver tabla 1).

La figura 3 muestra los resultados obtenidos de la red neuronal. Dado que la repuesta de la red es un número, la repuesta de ésta se ha correlacionado con los números que identifican a cada muestra. De esta forma una identificación perfecta corresponde cuando la salida de la red neuronal coincide con el valor asignado a la muestra ósea. El cero siempre se utiliza para indicar que no hay coincidencia. Se muestra los resultados correspondientes a las muestras H11, H21 y H31 a las que se ha asignado los números 11, 21 y 31 respectivamente.

La figura 4 muestra la curva de repuesta ROC de las muestras (a) H11, (b) H21 y (c) H31. La confianza de la predicción puede ser expresada por una probabilidad condicional, es decir, la tasa de clasificación correcta de los espectros clasificados (exactitud) expresada en porcentajes y es representado por el índice de correlación (IC). Para los casos mostrados en la figura el índice de correlación fue 98,5, 100 y 99,8 respectivamente.

La figura 5a muestra los resultados obtenidos de la red neuronal para un mismo individuo (1) y tres huesos diferentes (H11, H12 y H13) correspondientes al torso=primera costilla, brazo=cubito y pierna=fémur, respectivamente, mientras que la figura 5b muestra los resultados obtenidos de la red neuronal para dos individuos, el primero clasificado como 2 (muestras H21 y H22) y el segundo clasificado como 3 (muestra H31). Esta última muestra no ha sido introducida en la base de datos. Como puede observarse todas las muestras han sido correctamente identificadas. Incluso la muestra H31 fue correctamente identificada como desconocida y no como otra muestra presente en la base de datos utilizada como referencia, lo cual demuestra la robustez del método utilizado.

La figura 6 muestra el efecto del número de espectros utilizados en el entrenamiento de la red neuronal, sobre la robustez del método para lograr una identificación correcta. La robustez se define como la capacidad del método para identificar

correctamente una muestra desconocida como desconocida y no como alguna de las que están contenidas en su base de datos.

La figura 7 muestra una imagen de la interfase optimizada para mostrar de forma simple al usuario los resultados de la red neuronal durante el análisis binario de muestras. En este caso se utilizan dos muestras óseas de dos individuos como referencia (H11 y H21). El test se realiza sobre las restantes muestras óseas. El resultado se expresa indicando el índice de correlación porcentual para cada individuo o como muestra desconocida. El ejemplo muestra el resultado obtenido para la librería de espectros de la muestra H25 indicada como 25LH50 (individuo 2 hueso 5, donde 50 indica la cantidad de espectros contenidos en la librería), con un índice de correlación de 100, lo que indica que todos los espectros analizados de esta muestra han sido correctamente adjudicados al individuo 2.

15 **MODO DE REALIZACIÓN DE LA INVENCION**

La presente invención se ilustra mediante los siguientes ejemplos que no son en ningún caso limitativo de su alcance, el cual viene definido exclusivamente por la nota reivindicatoria.

20 **Ejemplo 1: Discriminación y clasificación de individuos.**

Como ejemplo, a fin de identificar una muestra ósea, se procede a la obtención de espectros LIBS obtenidos de un único disparo láser, procedente de uno de los huesos a analizar. Un total de 50 pulsos láser de cada muestra son almacenados en una base de datos como huellas digitales de las muestras. A fin de prevenir la radiación de cuerpo negro generada en los primeros momentos del plasma se utilizó como tiempo óptimo un retraso de 4 μ s entre el pulso láser y la obtención del espectro en la evolución del plasma. Como un ejemplo, los espectros LIBS correspondientes a tres individuos, clasificadas como H11, H21 y H31, donde se ha utilizado una nomenclatura basada en la letra H y dos números (x e y) de la forma Hxy (ver tabla 1), donde x identifica al individuo e y identifica al hueso utilizado, se muestra en la figura 2. Dado que la repuesta de la red es un número, la repuesta de ésta se ha correlacionado con los números que identifican a cada muestra. De esta forma una identificación perfecta

corresponde cuando la salida de la red neuronal coincide con el valor asignado a la muestra ósea. El cero siempre se utiliza para indicar que no hay coincidencia. Los resultados obtenidos por la red neuronal se muestran en la figura 3. Los primeros 50 puntos corresponden a la muestra del individuo 3 (H31), los siguientes 50 corresponden a la muestra del individuo 2 (H21) y los restantes al individuo 1 (H11). En todos los casos la muestra de hueso corresponde a la primera costilla derecha de cada individuo. Como puede observarse en la figura 3, con el primer pulso la red reconoce el espectro como perteneciente a uno del conjunto de datos de referencia y la salida de la red es 31. Por lo tanto el modelo de NN no puede “ver” diferencias entre la muestra analizada y la muestra de referencia con un número de identificación igual a 31. El mismo resultado se obtuvo hasta el espectro 50. Luego, cuando la muestra usada en el test es remplazada por la muestra H21 la salida de la red neuronal es 21. Lo mismo ocurre cuando se remplace la muestra H21 por la H11 produciendo una salida de la red de 11. En este punto, el modelo de NNs ha asignado correctamente los espectros para las tres muestras analizadas. La mayoría de los espectros para estas muestras son asignados correctamente y sólo 1 en la muestra H31 y dos en la muestra H11 se desvían más de un 1% del comportamiento esperado. Por este motivo son tomados como valores erróneos y afectan de manera considerable tanto a la curva ROC como el Índice de Correlación, los cuales fueron 98,7, 100 y 97,4 respectivamente y son mostrados en la figura 4. Dado que los espectros analizados provienen de un pulso láser simple, las perturbaciones observadas en solo 3 de 150 es más que aceptable, pero es esencial tener en cuenta los resultados del Índice de Correlación. La evaluación del procedimiento, curva ROC (figura 4), muestra que el 100% de los espectros han sido clasificados correctamente, aproximando la evaluación realizada a un ideal 100 % de sensibilidad y 100 % especificidad. Se obtuvieron exactamente los mismos resultados para todos los huesos analizados, mostrados en la Tabla 2.

Tabla 1: Nomenclatura utilizada para las muestras óseas.

Sección del Cuerpo	Hueso	Clasificación (y)
Torso	Primera Costilla derecha	1
Brazo	Cúbito izquierdo	2
Pierna	Fémur derecho	3
Huesos (Hxy)*		
Individuo 1	Individuo 2	Individuo 3
H11 H12 H13	H21 H22 H23	H31 H32 H33

* Todos los huesos fueron medidos con un mínimo de 3 réplicas.

Tabla 2. Resultados de clasificación del método de identificación

<i>Resultado de clasificación</i>						
<i>Test de identificación</i>					<i>Test de Robustez</i>	
Individuo/Hueso	Índice de correlación	No identificado	Mal identificados	Identificación correcta	Mal identificados	Correcto como desconocido
<i>Individuo 1/Hueso 11</i>	99.8	0	0	100	0	100
<i>Individuo 1/Hueso 12</i>	98.2	0	0	100	0	100
<i>Individuo 1/Hueso 13</i>	100	0	0	100	0	100
<i>Individuo 2/Hueso 21</i>	100	0	0	100	0	100
<i>Individuo 2/Hueso 22</i>	100	0	0	100	0	100
<i>Individuo 2/Hueso 23</i>	100	0	0	100	0	100
<i>Individuo 3/Hueso 31</i>	98.5	0	0	100	0	100
<i>Individuo 3/Hueso 32</i>	98.0	0	0	100	0	100
<i>Individuo 3/Hueso 33</i>	100	0	0	100	0	100

Ejemplo 2: Discriminación y clasificación de huesos de un mismo individuo.

Como ejemplo, a fin de identificar diferentes muestras óseas de un individuo se han utilizado tres muestras correspondientes a un mismo individuo en este caso clasificado como 1, dos correspondientes a otro individuo, en este caso clasificado como 2 y una tercera muestra correspondiente al individuo clasificado como 3. Esta última muestra no ha sido introducida en la base de datos y por lo tanto tendrá que ser observada como desconocida y la salida de la NN debe ser cero. Los resultados de las muestras del individuo 1 (H11, H12 y H13) se muestran en la figura 5a mientras que los resultados de las muestras del individuo 2 (H21 y H22) y 3 (H31) se muestran en la figura 5b. Como puede observarse, todas las muestras han sido correctamente identificadas. Incluso la muestra H31 fue correctamente identificada como desconocida y no como otra muestra presente en la base de datos utilizada como referencia, lo cual demuestra la robustez del método utilizado.

15

Ejemplo 3: Entrenamiento de la red neuronal

Como ejemplo, a fin de evaluar el número óptimo de espectros usados en el proceso de entrenamiento, se estudio la variación de la robustez como una función del número de espectros utilizado en la matriz de entrenamiento. La figura 6 muestra un gráfico de estos resultados. Como puede observarse, la robustez aumenta rápidamente con el número de espectros. Incluso para un valor muy bajo de espectros utilizados (menor a 10) la robustez es aceptable.

25

El tiempo requerido para la obtención de los espectros es muy bajo (menor a 1 minuto). Una vez almacenados, éstos pueden ser seleccionados para un análisis posterior en tiempo real. Los análisis llevados a cabo para diferentes muestras óseas analizadas en diferentes días, muestran que las librerías son adecuadas para una correcta identificación de las muestra óseas, incluso cuando se producen pequeñas variaciones en el sistema experimental, tales como cambios en la energía del láser, la temperatura de la habitación o la distancia entre la lente y la muestra.

30

- Las medidas con un pulso láser simple fueron suficientes para la identificación de las muestras óseas estudiadas. A la luz de estos resultados, el modelo optimizado de LIBS y NN desarrollado, provee resultados confiables para todas las muestras analizadas.
- 5 Estos resultados son el mejor indicador de la capacidad de la metodología presentada en esta patente.

Ejemplo 4: Presentación de resultados al usuario

- 10 Como ejemplo, la figura 7 muestra una imagen de la interfase optimizada para mostrar de forma simple al usuario los resultados de la red neuronal durante el análisis de la muestra H23. En este caso se utiliza como referencia la muestra H21 (Cúbito-brazo) mientras que el test se realiza sobre la muestra H23. Aunque la matriz de datos puede ser considerablemente grande, el tiempo de cálculo para el entrenamiento de la NNs
- 15 fue siempre menor a 10 segundos, mientras que el tiempo de cálculo para el análisis nunca superó los 5 segundos, en todas las muestras analizadas.

REIVINDICACIONES

1. Método de discriminación de individuos mediante la identificación y caracterización instantánea de tejido óseo caracterizado porque comprende: (a) irradiar una muestra empleando un láser focalizado sobre la superficie de una muestra, (b) detectar la radiación producida por los elementos químicos del plasma que se forma en la superficie sobre la que se focaliza el haz láser, obteniendo un espectro único de la muestra, utilizando un analizador óptico (c) comparar la señal del detector con una base espectral dinámica utilizando un sistema de redes neuronales artificiales que incluyen algoritmos matemáticos supervisados.
2. Método de discriminación de individuos mediante la identificación y caracterización instantánea de tejido óseo, según reivindicación 1, donde el láser es un láser en estado sólido o gaseoso que emite en el rango ultravioleta, visible o infrarrojo.
3. Método de discriminación de individuos mediante la identificación y caracterización instantánea de tejido óseo, según reivindicaciones 1 y 2, donde el láser es del tipo Nd:YAG trabajando a una frecuencia de 1 a 20 Hz, a una longitud de onda fija de 1064 nm, con una duración de pulso de 4 nanosegundos
4. Método de discriminación de individuos mediante la identificación y caracterización instantánea de tejido óseo, según reivindicación 1, donde la detección de la radiación producida por los elementos químicos del plasma se realiza a través de un sensor óptico CCD que proporciona 2048 puntos espectrales en un rango de 200 a 1100 nm.
5. Método de discriminación de individuos mediante la identificación y caracterización instantánea de tejido óseo, según reivindicación 1, donde la red neuronal posee conexiones hacia delante basadas en un modelo de propagación *perceptron*, tres capas (entrada, salida y oculta), una función de transferencia entre capas que utiliza un algoritmo tangente hiperbólico sigmoidal y su matriz de pesos se optimiza mediante el algoritmo de entrenamiento de gradiente escalado conjugado SCG previniendo el sobreajuste y el sobreentrenamiento.

6. Dispositivo para la identificación y caracterización instantánea de muestras de tejido óseo que comprende: (a) un equipo láser capaz de producir un proceso de vaporización y posterior formación de un plasma en la superficie de una muestra, (b) un analizador óptico que detecta la radiación producida por los elementos químicos del plasma, (c) un ordenador personal con interfaz gráfica que permite al usuario el control del láser, la observación del espectro de la muestra obtenido y la representación de los resultados para su análisis y que incluye un base de datos espectral dinámica que utiliza un sistema de redes neuronales artificiales que incluyen algoritmos matemáticos supervisados.
7. Dispositivo para la identificación y caracterización instantánea de muestras de tejido óseo, según reivindicación 6, donde el láser es cualquier tipo de láser es un láser en estado sólido o gaseoso que emite en el rango de ultravioleta, visible o infrarrojo.
8. Dispositivo para la identificación y caracterización instantánea de muestras óseas, según reivindicaciones 6 y 7, donde el láser es del tipo Nd:YAG trabajando a una frecuencia de 1 a 20 Hz, a una longitud de onda fija de 1064 nm, con una duración de pulso de 4 nanosegundos.
9. Dispositivo para la identificación y caracterización instantánea de muestras de tejido óseo, según reivindicación 6, donde la detección de la radiación producida por los elementos químicos del plasma se realiza a través de un sensor óptico CCD que proporciona 2048 puntos espectrales en un rango de 200 a 1100 nm.
10. Dispositivo para la identificación y caracterización instantánea de muestras de tejido óseo, según reivindicación 6, donde la red neuronal posee conexiones hacia delante basadas en un modelo de propagación *perceptron*, tres capas (entrada, salida y oculta), una función de transferencia entre capas que utiliza una algoritmo tangente hiperbólico sigmoideal y su matriz de pesos se optimiza mediante el algoritmo de entrenamiento de gradiente escalado conjugado SCG previniendo el sobreajuste y el sobreentrenamiento.

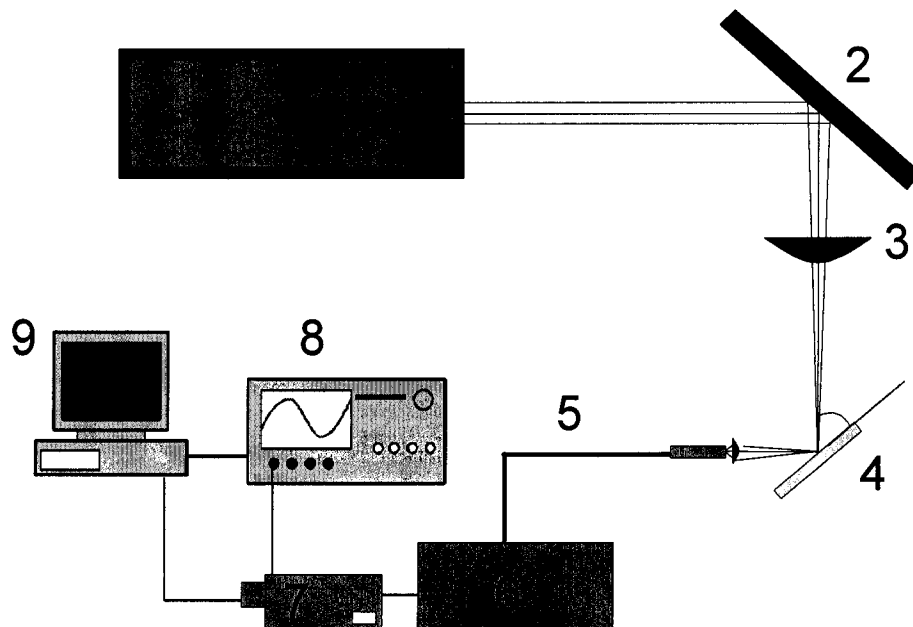


FIGURA 1

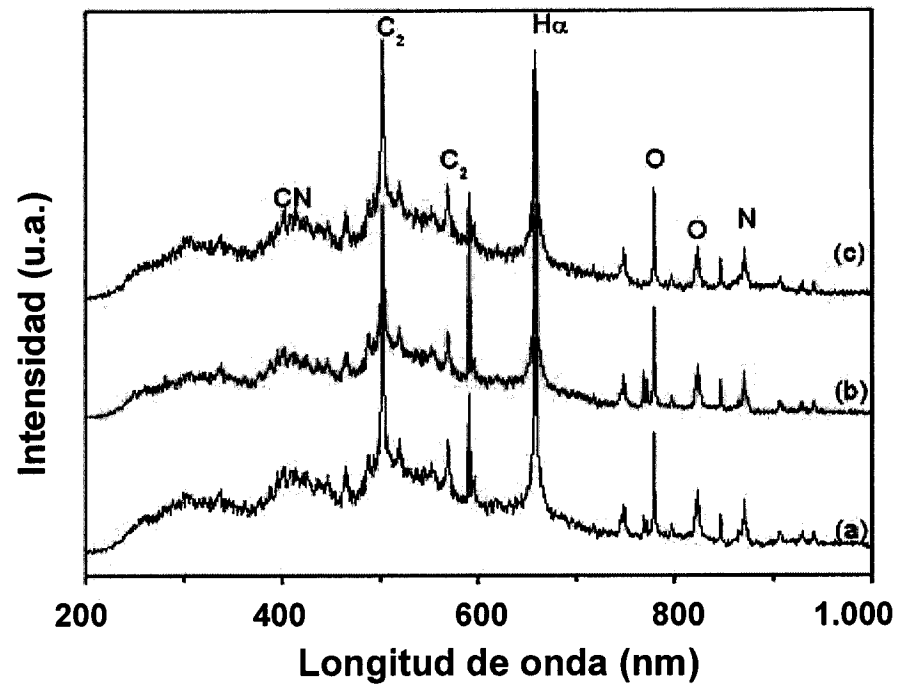


FIGURA 2

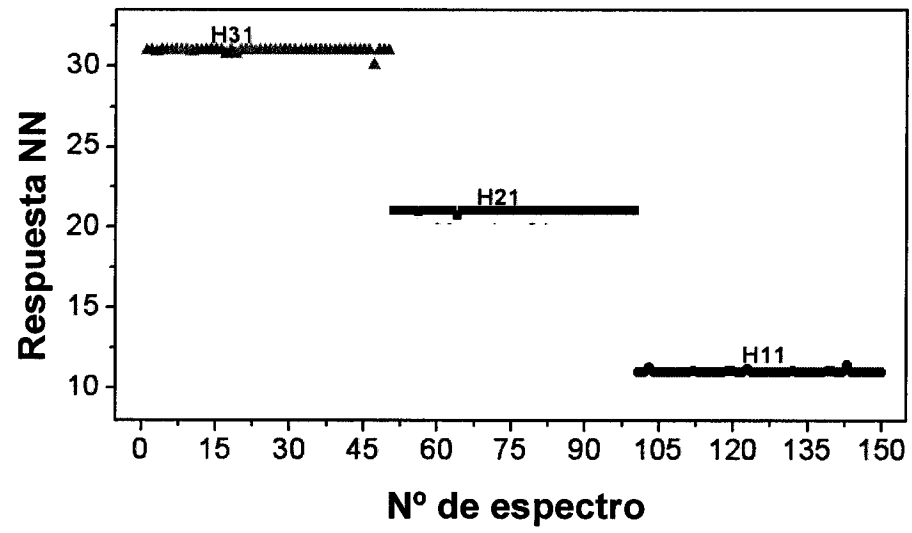


FIGURA 3

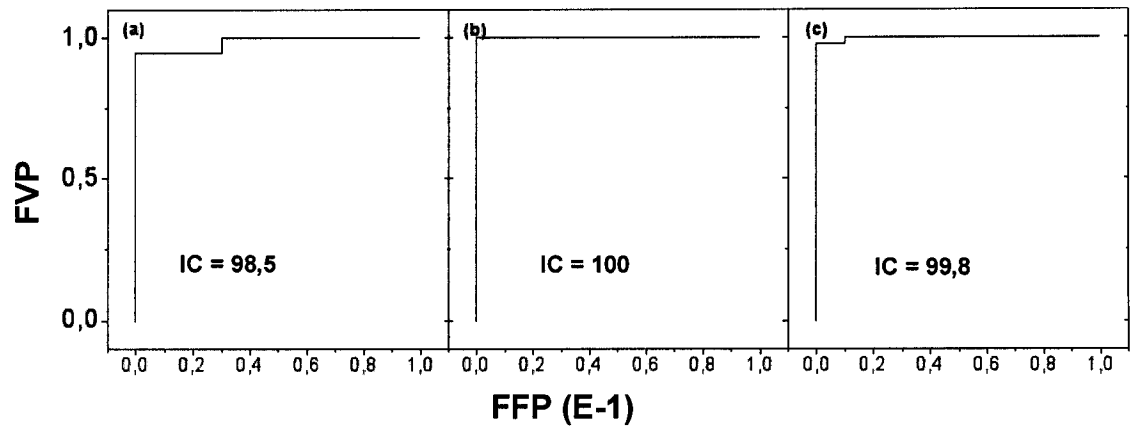


FIGURA 4

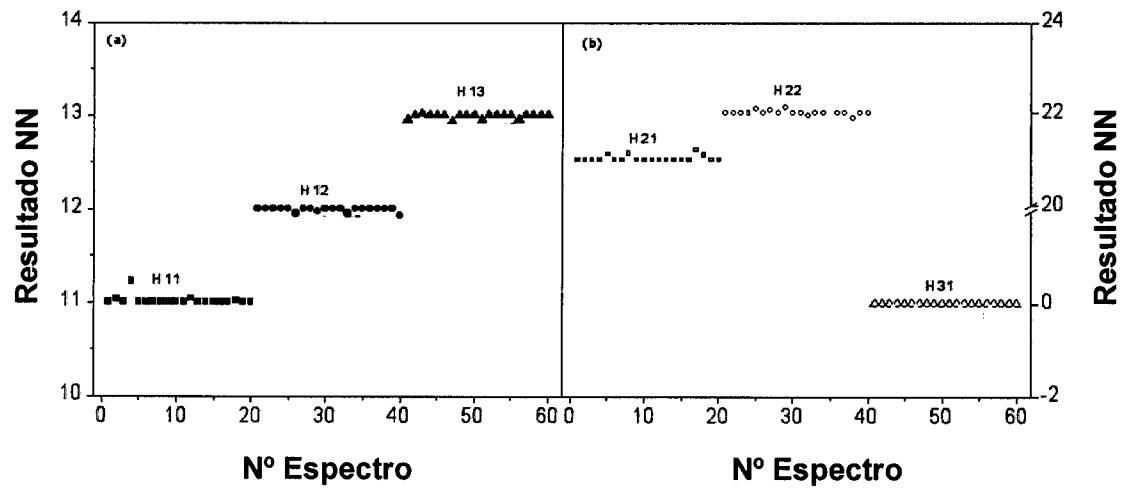


FIGURA 5

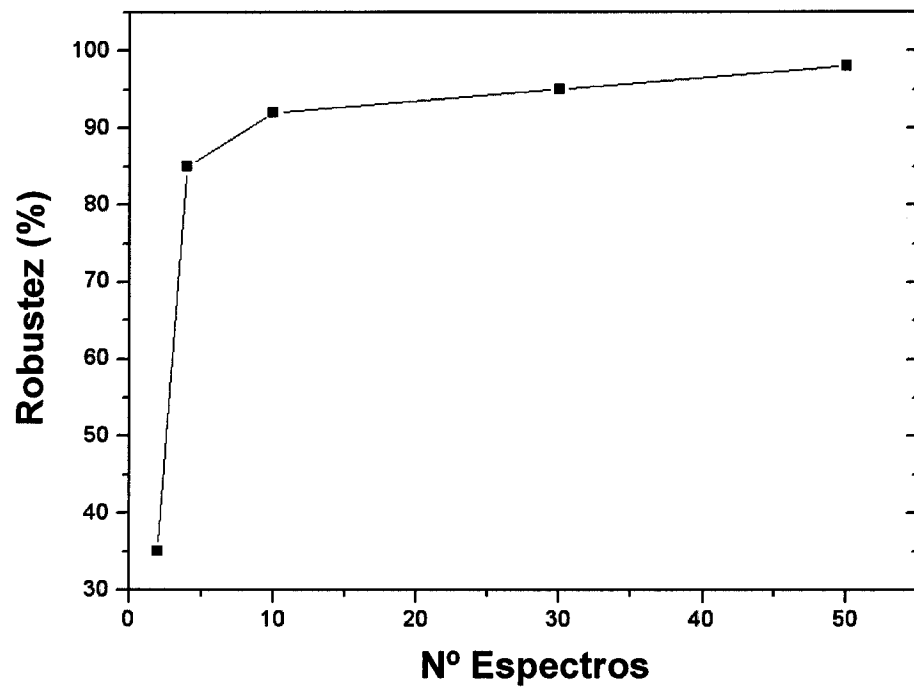


FIGURA 6

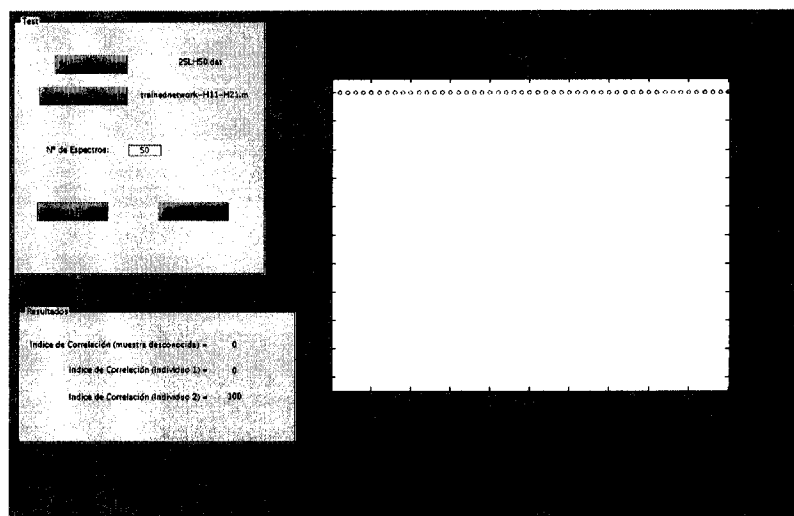


FIGURA 7