

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 536 560**

51 Int. Cl.:

G06K 9/68 (2006.01)

G06K 9/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.12.2009** **E 09799660 (7)**

97 Fecha y número de publicación de la concesión europea: **18.02.2015** **EP 2382587**

54 Título: **Método para descubrir y reconocer patrones**

30 Prioridad:

31.12.2008 FI 20086260

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

26.05.2015

73 Titular/es:

AALTO-KORKEAKOULUSÄÄTIÖ (100.0%)

PL 11000

00076 Aalto, FI

72 Inventor/es:

**LAINÉ, UNTO y
RÄSÄNEN, OKKO**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 536 560 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

Método para descubrir y reconocer patrones

Campo técnico de la invención

La invención se refiere en general a un método para descubrir y reconocer patrones.

5 Antecedentes de la invención

Una idea básica en todos los métodos para descubrir patrones es modelar regularidades estadísticas y comparar un modelo con una representación real con el fin de medir una similitud o similitudes entre el modelo creado (aprendido) y un patrón actual que está analizándose.

10 Uno de los métodos y técnicas conocidos utilizado para modelar y reconocer patrones en secuencias es el modelo de Markov, que supone que una secuencia que va a modelarse tiene la propiedad de Markov. Tener la propiedad de Markov significa que, dado un estado actual, los estados futuros son independientes de los estados pasados. Dicho de otro modo, la descripción del estado actual en sí misma capta completamente toda la información que podría influir en la evolución futura del procedimiento. Los estados futuros se alcanzarán a través de un procedimiento probabilístico en lugar de un procedimiento determinístico.

15 En cada etapa un sistema puede cambiar su estado del estado actual a otro estado, o permanecer en el mismo estado, según una cierta distribución de probabilidad. Los cambios del estado se denominan transiciones y las probabilidades asociadas con diversos cambios de estado se denominan probabilidades de transición.

20 Muchos procedimientos físicos y correspondientes secuencias observables, que se crean mediante los procedimientos físicos, tienen estructuras fuertes tales como estructuras temporales que pueden medirse mediante coeficientes de correlación de orden superior. Por tanto, dependiendo de la resolución temporal usada la secuencia creada puede tener estructuras amplias (por el tiempo o espacio), que no pueden modelarse con precisión mediante una cadena de Markov en la que los estados futuros son independientes de los estados pasados.

Un objeto de la invención es proporcionar un método para descubrir y reconocer patrones, un producto de programa informático para descubrir y reconocer patrones y un aparato para descubrir y reconocer patrones.

25 Un objeto de la invención se alcanza proporcionando un método según la reivindicación 1, un producto de programa informático según la reivindicación 8 y un aparato según la reivindicación 9.

Una realización de la presente invención se refiere a un método según la reivindicación 1, una realización a un producto de programa informático según la reivindicación 8 y una realización a un aparato según la reivindicación 9.

En las reivindicaciones dependientes se definen realizaciones adicionales.

30 Según una realización de la invención un método comprende recibir una primera secuencia que comprende símbolos de primera secuencia en relación con el concepto y una etiqueta asociada a la primera secuencia, obtener matrices de probabilidad de transición a partir de matrices de frecuencia de transición que representan datos de frecuencia de incidencia de transiciones entre los símbolos de primera secuencia a diferentes distancias en la primera secuencia y aprender las matrices de probabilidad de transición para cada etiqueta y cada distancia para
35 obtener una función de activación que determina el concepto que se produce en una segunda secuencia.

40 El método según las realizaciones de la invención puede realizarse en un aparato, que puede ser por ejemplo un sistema informático, un portátil o un teléfono móvil, en el que un software ejecuta el procedimiento para descubrir y reconocer patrones. En segundo lugar, el procedimiento para descubrir y reconocer patrones puede ejecutarse mediante un procesador (por ejemplo un chip de procesador de señal digital (DSP)), que se ha adaptado para realizar el procedimiento para descubrir y reconocer patrones.

El término "concepto" (patrón) se refiere a un cierto tipo de regularidad estadística en una secuencia (o entre dos secuencias diferentes), que difiere de otros conceptos (otras regularidades) y forma un claro contraste con todas las irregularidades de tipo ruido.

45 El término "etiqueta" se refiere a una indicación, que está asociada con una secuencia o varias etiquetas están asociadas a la misma secuencia y estas asociaciones se denominan conceptos.

El término "transiciones" se refiere a pares de símbolos de secuencia (índice, indicador). Los símbolos de secuencia

son los elementos básicos de las secuencias y los pares de símbolos de secuencia tienen distancias entre los símbolos de secuencia dentro de un par (Dicho de otro modo, cada par de símbolos tiene una cierta ubicación y distancia entre sí (diferencia en las ubicaciones) en la secuencia).

5 Según una realización de la invención el método, que se da a conocer en una realización anterior, comprende recibir la primera información que comprende el concepto. La primera información puede ser por ejemplo voz o imagen.

Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, comprende asociar la etiqueta a la primera secuencia.

10 Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, comprende discretizar y cuantificar primera información que comprende el concepto y representar la primera información en forma de la primera secuencia unidimensional (1D). Además, puede muestrearse información 2D visual (espacial) de diferentes maneras para formar secuencias 1D. Por este motivo la descripción detallada se limita solamente a la información secuencial (1D). En un caso más complicado es posible tratar con un conjunto de secuencias 1D y también asociaciones entre estas secuencias (por ejemplo asociaciones entre representaciones de audio y visuales).

15 Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, recopila los datos de frecuencia de la incidencia de las transiciones entre los símbolos de secuencia en la primera secuencia. Se formula un grupo de matrices que comprende al menos una matriz, que consiste en los datos de frecuencia, para cada etiqueta. El número de las matrices en el grupo depende del número de las diferentes distancias de los pares de símbolos recopilados a partir de la secuencia y el número de etiquetas asociadas con
20 conceptos (por ejemplo matriz $M(a,b|distancia, etiqueta)$, donde $a*b$ define el tamaño de la matriz y el número de las matrices es igual a distancias por etiquetas).

Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, comprende almacenar los datos de frecuencia de la incidencia de las transiciones en la primera secuencia en las matrices de frecuencia de transición.

25 Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, comprende normalizar las matrices de frecuencia de transición a las matrices de probabilidad de transición según ecuaciones representadas posteriormente (1)-(3).

30 Según una realización de la invención en el método, que se da a conocer en cualquiera de las realizaciones anteriores, la(s) segunda(s) secuencia(s) se procesa(n) a partir de segunda información real que comprende el concepto discretizando y cuantificando la segunda información para representar la segunda información en forma de la(s) segunda(s) secuencia(s).

35 Según una realización de la invención el método, que se da a conocer en cualquiera de las realizaciones anteriores, se usa para un reconocimiento de patrones en relación con por ejemplo un reconocimiento de voz o un reconocimiento de imagen. El algoritmo usado que ejecuta el método toma un flujo de información dado en forma discreta, cuantificada. La información puede crearse mediante cualquier procedimiento con variación en el tiempo, por ejemplo, señales audiovisuales, o puede ser variable sólo espacialmente, por ejemplo una imagen fija.

40 Según una realización de la invención un producto de programa informático, que recibe una primera secuencia que comprende símbolos de primera secuencia en relación con el concepto y una etiqueta asociada con la primera secuencia, obtiene matrices de probabilidad de transición a partir de matrices de frecuencia de transición que representan datos de frecuencia de incidencia de transiciones entre los símbolos de primera secuencia a diferentes distancias en la primera secuencia y aprende las matrices de probabilidad de transición para cada etiqueta y cada distancia para obtener una función de activación que determina el concepto que se produce en una segunda secuencia.

45 Según una realización de la invención un aparato, que recibe una primera secuencia que comprende símbolos de primera secuencia en relación con el concepto y una etiqueta asociada con la primera secuencia, obtiene matrices de probabilidad de transición a partir de matrices de frecuencia de transición que representan datos de frecuencia de incidencia de transiciones entre los símbolos de primera secuencia a diferentes distancias en la primera secuencia y aprende las matrices de probabilidad de transición para cada etiqueta y cada distancia para obtener una función de activación que determina el concepto que se produce en una segunda secuencia.

50 El método según las realizaciones de la invención puede realizarse en un aparato, por ejemplo un sistema informático, un portátil o un teléfono móvil, en el que un software ejecuta el procedimiento para descubrir y reconocer patrones.

En segundo lugar, el procedimiento para descubrir y reconocer patrones puede ejecutarse mediante un procesador programado (por ejemplo un chip de procesador de señal digital (DSP)), que está adaptado para realizar el procedimiento para descubrir y reconocer patrones.

5 El algoritmo que ejecuta el método según las realizaciones de la invención aprende estadísticas recopilando evidencias de secuencias reales, construye modelos estadísticos para esas secuencias y los aplica en el descubrimiento y la clasificación de patrones.

10 Las estadísticas se basan en la frecuencia de los pares de símbolos de secuencia hallados en diferentes distancias espaciales en casos espaciales o retardos en casos temporales. En el caso de secuencias de tiempo los pares de símbolos de secuencia pueden interpretarse como transiciones de estado y sus frecuencias relativas correspondientes pueden normalizarse para producir correspondientes probabilidades transicionales. El número de incidencias de los diferentes pares de símbolos de secuencia pueden recopilarse para formar una matriz de frecuencia o una tabla. Cada distancia temporal (retardo) o espacial creará una matriz o una tabla propia. De manera correspondiente, pueden crearse estadísticas asociativas entre dos secuencias diferentes.

15 Las estadísticas de transición están vinculadas a la presencia de una entrada multimodal (una etiqueta) de modo que cada etiqueta indica un concepto (patrón) en la serie de tiempo que se asocia con la misma. Después del entrenamiento del modelo, puede reconocerse un concepto desconocido que es uno de los aprendidos anteriormente, a partir de una nueva secuencia. El algoritmo está adaptado especialmente para el reconocimiento de patrones en condiciones en las que las secuencias están corrompidas con altos niveles de distorsión y ruido, y condiciones en las que las dependencias de eventos se distribuyen por el tiempo o espacio, de modo que los eventos adyacentes no pueden considerarse de manera fiable como dependientes entre sí.

20 Una diferencia entre el método según las realizaciones de la invención y la cadena de Markov es que el método según las realizaciones de la invención recopila evidencia estadística a partir de una secuencia contando en primer lugar el número de todos los pares de indicadores hallados a diferentes distancias y a continuación realiza dos normalizaciones diferentes (ecuaciones (1) y (2) en la descripción detallada), combina estas dos representaciones estadísticas antes de la tercera normalización realizada para derivar el conjunto final de matrices de activación.

Sólo una de las representaciones derivadas, la matriz de transición de estado del retardo uno, es equivalente a la representación conocida por el método de la cadena de Markov. En otras dos representaciones usadas en la derivación de las matrices de activación el tratamiento de la evidencia estadística difiere considerablemente de los conocidos anteriormente.

30 Otra diferencia es que la cadena de Markov normalmente no utiliza estadísticas por retardos mayores de uno y cuando las usa, las estadísticas pueden derivarse de la matriz de transición de retardo uno mediante una simple multiplicación de matrices (ecuación de Chapman-Kolmogorov). En el método según las realizaciones de la invención las estadísticas por retardos mayores se toman directamente de la secuencia y no se aproximan (estiman) mediante una multiplicación de matrices.

35 La tercera diferencia entre el método según las realizaciones de la invención y la cadena de Markov es que el método según las realizaciones de la invención utiliza en paralelo dos representaciones diferentes para cada concepto (ecuación (2)) y todas las matrices de activación obtenidas para pares de indicadores de distancias diferentes (ecuación (4)). Por tanto, el patrón de activación combinado no sólo depende del estado actual o anterior sino de todos los eventos anteriores en la secuencia descrita por el conjunto de matrices de activación.

40 Las ventajas del método según las realizaciones de la invención se obtienen cuando se trata de datos en los que una perturbación ha destruido por completo o corrompido gravemente la información original en una ventana temporal o espacial limitada. La información original puede recuperarse (corregirse) sólo cuando un modelo puede utilizar la información disponible en las proximidades, y aún más, cuando el modelo puede realizar asociaciones apropiadas basándose en la información no corrompida en las proximidades.

45 Breve descripción de los dibujos

A continuación, se describirán los aspectos de la invención en más detalle con referencia a realizaciones a modo de ejemplo según los dibujos adjuntos, en los que

la figura 1 ilustra un diagrama de flujo general a modo de ejemplo del método para descubrir y reconocer patrones según una realización ventajosa de la invención,

50 la figura 2A ilustra un diagrama de flujo a modo de ejemplo del método para un descubrimiento de patrones según una realización ventajosa de la invención,

la figura 2B ilustra un diagrama de flujo a modo de ejemplo del método para un reconocimiento de patrones según una realización ventajosa de la invención,

la figura 3 ilustra una vista de la activación de las representaciones de palabras en una expresión vocal en función del tiempo sin el filtrado de mediana de la activación,

5 la figura 4 ilustra una vista de la activación de las representaciones de palabras en una expresión vocal en función del tiempo con el filtrado de mediana de la activación y

la figura 5 ilustra una vista de la curva de aprendizaje característica del algoritmo.

Descripción detallada

10 La figura 1 representa un diagrama de flujo general que describe un método 100 para descubrir y reconocer patrones según la realización de la invención.

En primer lugar en la etapa 110 se recopilan datos, por ejemplo expresiones vocales que comprenden varias palabras en un procedimiento de reconocimiento de voz.

15 Los datos recopilados se procesan durante la etapa 120 de modo que a partir de los datos procesados pueden hallarse regularidades estadísticas (patrones) y crearse modelos estadísticos para los patrones hallados. En la fase de aprendizaje una etiqueta simultánea, externa indica cuál del conjunto de patrones (que se reconocerán posteriormente) está presente en algún lugar en la secuencia real.

Finalmente, los modelos estadísticos de los patrones se comparan con datos reales nuevos, no vistos anteriormente, para descubrir similitudes entre los modelos estadísticos creados de los patrones y un patrón real en la etapa 130.

20 La figura 2A da a conocer, mediante un ejemplo solamente, un diagrama de flujo que describe un método 200 de descubrimiento según la realización de la invención en más detalle.

Durante el inicio del método en la etapa 205, se enciende un ordenador y/o una aplicación que ejecuta el método y se proporcionan las fases necesarias antes de un procedimiento de aprendizaje de patrones, descubrimiento y reconocimiento, tales como la definición de la configuración de aplicación y la inicialización de diferentes variables y parámetros.

25 En este caso, un usuario define la configuración, las variables y los parámetros en vista del aprendizaje, descubrimiento y reconocimiento de patrones.

30 A continuación, en la etapa 210, la información usada en el procedimiento de aprendizaje (entrenamiento), tal como expresiones vocales que comprenden una o más palabras pronunciadas por un ser humano, se introduce en un software de descubrimiento y reconocimiento de patrones a través de un receptor, por ejemplo un micrófono, y una unidad de procesamiento de señales que puede modificar la señal de audio recibida para obtener una forma apropiada para el software de descubrimiento y reconocimiento de patrones.

En la etapa 215 se discretiza y cuantifica la información recibida (muestra), y después, en la etapa 220, se representa la información discreta y cuantificada en forma de secuencia o secuencias 1D.

35 Entonces, en la etapa 225, algún evento (patrón), que se asocia con la(s) secuencia(s) se especifica mediante una etiqueta, que puede ser por ejemplo un número o letras.

40 Así, la entrada en el sistema consiste en una serie de tiempo de elementos discretos o información espacial muestreada para formar secuencias 1D, y en la fase de entrenamiento, etiquetas que especifican algún evento o patrón asociado con las secuencias. En algunos casos una modalidad de información puede proporcionar una etiqueta para otra modalidad. Los elementos básicos de las secuencias se denominan indicadores y en el caso más sencillo pueden hacer referencia a elementos en un libro de códigos de cuantificación vectorial o pueden producirse mediante cualquier tipo de discretización de series de tiempo o imágenes. En un caso más complejo pueden hacer referencia a una representación de nivel superior de información, por ejemplo eventos o elementos posibles que reflejan propiedades cualitativas claras. Un ejemplo podría ser modelar mercados bursátiles y procedimientos econométricos complejos. La otra fuente de información (posiblemente otra fuente de modalidad) se representa mediante un conjunto de denominadas etiquetas de concepto c.

45 Las etiquetas son normalmente valores de número entero que representan salidas invariables de otro procedimiento que se asocian a la entrada de serie de tiempo (por ejemplo un procedimiento de categorización realizado en otra

modalidad como la percepción visual o táctil en caso de reconocimiento de voz, o algún otro grupo de eventos definidos manualmente que pretenden asociarse con la serie de tiempo).

5 El mecanismo puede funcionar también en sentido inverso; un evento acústico puede servir como etiqueta para aprender patrones visuales. Una modalidad puede formar etiquetas para otras modalidades para ayudar en el aprendizaje. Más generalmente, el método permite la construcción de asociaciones estadísticas entre modalidades diferentes. Éste es uno de los temas clave para modelar y entender la formación y el aprendizaje de los significados (por parte de medios y seres humanos).

10 En la etapa 230, cuando el concepto (etiqueta) se ha activado y se ha representado la secuencia, el algoritmo empieza a recopilar datos de frecuencia de la incidencia de pares de indicadores (transiciones) en la secuencia a una distancia l .

El algoritmo almacena los datos de frecuencia recopilados de la incidencia de pares de indicadores en una tabla de histograma o una matriz T (matriz de frecuencia de transición) en la etapa 235. Los indicadores originales pueden usarse como referencias para T , cuando es necesario el número de incidencias del correspondiente par de indicadores.

15 Durante la siguiente etapa 240 el histograma recopilado en T se usa entonces para producir otra representación P .

20 La estructura principal del algoritmo es una matriz $P_{l,c}$ de tamaño $N_q \times N_q$, donde N_q es el tamaño del libro de códigos, que se asemeja a las matrices de probabilidad de transición pero no contiene probabilidades bien definidas sino en su lugar una especie de sumas de probabilidades acumulativas. Mantiene un registro de probabilidades de transición normalizadas del indicador $a[t-l]$ al indicador $a[t]$ con la presencia simultánea del concepto c , donde $l, c, t \in Z$, e l es un elemento del conjunto $l = \{l_1, l_2, l_3, \dots, l_n\}$ y c es un elemento del conjunto $c = \{1, 2, 3, \dots, N_c\}$. Dicho de otro modo, N_c es el número total de conceptos introducidos en el sistema. Si se define $N_l = |l|$, hay un total de $N_P = N_l \cdot N_c$ instancias de P matrices, una para cada concepto con un retardo específico. Una matriz $T_{l,c}$ es de otro modo similar a $P_{l,c}$ excepto porque mantiene un registro de las frecuencias de transición en lugar de probabilidades normalizadas del indicador $a[t-l]$ al indicador $a[t]$ en presencia del concepto c .

25 Como los valores de P no son probabilidades clásicas en el intervalo entre 0 y 1 debido a un procedimiento de normalización de tres fases, los valores de P se denominarán valores de activación y P se denominará matriz de activación. Los valores de activación almacenados en P se calcularán usando la información de frecuencia almacenada en T .

30 A continuación, se representa cómo se proporciona el entrenamiento en el método. Para simplificar la notación, los elementos de las matrices $P_{l,c}$ y $T_{l,c}$ se designan en forma de $P(a_i, a_j | l, c)$ y $T(a_i, a_j | l, c)$, donde las dos primeras variables a_i y a_j definen los índices de elemento de matriz de los indicadores (transición de a_i a a_j o aparición conjunta de a_i y a_j), mientras que l define el retardo y c define el concepto.

35 La entrada consiste en secuencias de entrenamiento $S = \{s_1, s_2, \dots, s_n\}$ y conceptos relacionados con secuencia $V = \{v_1, v_2, \dots, v_n\}$, donde cada $v_i = \{c_1, c_2, \dots, c_n\}$, $v \in c$. Todas las transiciones en la secuencia s_i que se producen en los retardos l se actualizan a las matrices de frecuencia de transición $T_{l,c}$, donde c es un elemento de v_i asociado con s_i . Este procedimiento se repite para todas las S en el material de entrenamiento.

El siguiente ejemplo de pseudocódigo ilustra el procedimiento de recopilación de las frecuencias de transición:

para $i = 1:\text{longitud}\{S\}$

$s = S(i)$;

40 $v = V(i)$

 para retardo = $1:\text{longitud}(l)$

 para $t = 1:\text{longitud}(s)$

 para $c = 1:\text{longitud}(v)$

$T(s[t-\text{retardo}], s[t] | \text{retardo}, c) = T(s[t-\text{retardo}], s[t] | \text{retardo}, c) + 1$;

45 fin

fin

fin

fin

5 En la etapa 240, como todas las transiciones que se producen en el material de entrenamiento se suman a y almacenan en las matrices de frecuencia de transición T, las matrices se normalizan a matrices de probabilidad de transición P' normalizando la probabilidad de transición de cada indicador a todos los demás indicadores $\sum_x Pr(a_i, a_x) = 1$ porque

$$P'(a_i, a_j | l_d, c_k) = \frac{T(a_i, a_j | l_d, c_k)}{\sum_{x=1}^{N_q} T(a_i, a_x | l_d, c_k)} \quad (1)$$

donde Nq es el tamaño de libro de códigos, es decir, el número de elementos únicos, en la serie de tiempo.

10 La probabilidad de que se produzca una transición específica durante la presencia de una etiqueta en lugar de todas las demás transiciones se suma de manera acumulativa a P'_{l,c}:

$$P''(a_i, a_j | l_d, c_k) = P'(a_i, a_j | l_d, c_k) + \frac{T(a_i, a_j | l_d, c_k)}{\sum_{x=1}^{N_q} \sum_{y=1}^{N_q} T(a_x, a_y | l_d, c_k)} \quad (2)$$

15 Esto mejora el valor de aquellas transiciones que son muy comunes en presencia del concepto. Se observará que ahora la matriz ya no es una matriz de probabilidad de transición bien definida en el sentido de que las probabilidades del siguiente estado no se suman para dar uno. Por tanto, los valores de P se denominan a partir de ahora valores de activación (específicos del concepto) y los resultados del procedimiento de reconocimiento se denominan activaciones de concepto.

Finalmente, se incorpora una probabilidad de que se produzca una transición durante la presencia de un concepto c_k en lugar de cualquier otro concepto a la matriz de activación final P porque

$$P(a_i, a_j | l_d, c_k) = \frac{P''(a_i, a_j | l_d, c_k)}{\sum_{z=1}^{N_c} P''(a_i, a_j | l_d, c_z)} - \frac{1}{N_c} \quad (3)$$

20 Dicho de otro modo, la probabilidad acumulativa de una transición de a_i a a_j en caso de que la etiqueta c se divide entre la suma de probabilidades de la misma transición que se produce durante todas las posibles etiquetas c. Si una transición pasa a ser igualmente probable para todos los conceptos, no conteniendo por tanto valor de información, tendría una probabilidad de 1/Nc. Por tanto, en cada elemento en todas las matrices se ha restado 1/Nc de su valor para tener una activación de cero para un caso completamente aleatorio y un valor negativo para transiciones que se producen con más frecuencia durante otros conceptos. La resta de 1/Nc mencionada anteriormente no es una etapa necesaria, sino que hace que el procedimiento sea más conveniente.

Cuando ha finalizado el entrenamiento, el método termina en la etapa 245.

30 La figura 2B ilustra, mediante un ejemplo solamente, un diagrama de flujo que describe un método 260 para reconocer patrones según la realización de la invención

El inicio del procedimiento de reconocimiento en la etapa 265, es similar al inicio del procedimiento de descubrimiento.

35 En la etapa 270, se introduce segunda información, datos de entrada reales, en el sistema y la segunda información se procesa discretizando y cuantificando para presentar los datos de entrada reales en forma de secuencia o secuencias que tienen patrones durante las etapas 275 y 280.

A continuación, en la etapa 285, se determina un concepto a partir de los datos de entrada reales estudiando las transiciones de la(s) secuencia(s). Las transiciones tienen probabilidades de transición P en relación con cada distancia y cada etiqueta. Estas probabilidades se estiman durante la fase de aprendizaje. Las probabilidades se

combinan para crear una función de activación A y se reconocerá el concepto que tiene el nivel de activación más alto.

Por tanto, el nivel de activación del concepto c_i en el momento t dadas las secuencias de entrada reales puede expresarse como

$$A(c_i, t) = \sum_{d=1}^{N_t} P(s[t - l_d], s[t] | l_d, c_i) \quad (4)$$

5

cuando sólo se incluye el historial hacia atrás de la secuencia de entrada.

Cuando se completa el reconocimiento de patrones, el método finaliza en la etapa 290.

10 También es posible tener un procedimiento de reconocimiento bidireccional incluyendo $P(s[t], s[t+1] | l_d, c_i)$ valores de activación en la suma en la ecuación (4) si se conocen de antemano los siguientes indicadores hasta el mayor retardo $\max(l)$ en la secuencia. Esto mejora la localización del evento reconocido, puesto que el valor pico de la curva de activación se centra en un punto en el que hay un mayor soporte estadístico para el concepto específico, que se distribuye simétricamente alrededor de ese punto en cuanto a probabilidades transicionales.

15 La ecuación (4) proporciona una estimación de activación local para cada candidato de concepto aunque en muchas aplicaciones resulta útil examinar la salida de activación en una ventana temporal mayor puesto que los eventos que están reconociéndose se dispersan por varias tramas de tiempo posteriores. Una posibilidad para hacer esto es, en primer lugar, aplicar un filtro de mediana o paso bajo a las curvas de activación en una ventana temporal mayor. Entonces, en cada una de estas curvas de activación temporales relacionadas con conceptos se busca una subsecuencia de longitud $L_i \in [L_{\min}, L_{\max}]$ que tiene una suma acumulativa máxima de valores de activación. Después de hallar estas subsecuencias para cada modelo de concepto c , la subsecuencia i con la suma acumulativa
20 más alta define la hipótesis de concepto c_i .

25 L_{\min} establece un límite temporal mínimo para la información que se incluye en el procedimiento de decisión de reconocimiento y debería ser al menos tan larga como el evento más corto posible que está reconociéndose. De manera similar, L_{\max} define un límite superior temporal para la integración de información y debería ser al menos tan larga como el evento más largo posible que está reconociéndose. Sin embargo, tener valores incluso más grandes para L_{\max} puede ser beneficioso en varias situaciones, puesto que el contexto de un evento contiene a menudo pistas para el propio evento y las estadísticas introducidas en las matrices de probabilidad de transición tienen en cuenta esta información.

30 Se realizan sumas (combinación lineal) en las ecuaciones presentadas anteriormente (2) y (4) ponderando factores con un valor uno. Sin embargo, es posible usar factores de ponderación α que tienen valores diferentes de uno. Estos valores pueden determinarse basándose en algún criterio adicional, por ejemplo basándose de manera iterativa en un éxito de reconocimiento de patrones.

35 Además, también es posible ejecutar todo el algoritmo en paralelo para varios flujos de entrada cuantificados de manera sincronizada con el fin de incorporar varias fuentes de información. Esto transforma las matrices de frecuencia y activación en la forma $T_{\psi}(a_i, a_j | l, c)$ y $P_{\psi}(a_i, a_j | l, c)$, donde ψ designa el número del flujo de entrada que está procesándose. El entrenamiento se realiza de manera similar a la condición de flujo único con el fin de construir matrices de concepto separadas para cada concepto en cada retardo y para cada flujo. En la fase de prueba la salida de probabilidad de todos los flujos se combina para tener una probabilidad de un concepto c_i en el momento t de

$$A(c, t) = \sum_{\psi=1}^{||\Psi||} \left(\sum_{d=1}^{N_t} P_{\psi}(s[t - l_d], s[t] | l_d, c_i) \right) * \omega_{\psi} \quad (5)$$

40 donde ω_{ψ} es un factor de ponderación definido para cada flujo de entrada.

En las figuras 3 y 4 se muestra un ejemplo de resultado en relación con el método descrito para reconocer patrones.

45 El algoritmo de matriz de concepto se aplicó a un experimento de aprendizaje de palabras no supervisado. El objetivo fue aprender 11 palabras clave diferentes a partir de un corpus que contenía 4000 expresiones vocales pronunciadas por cuatro hablantes en inglés británico (dos hombres y dos mujeres, 1000 expresiones vocales cada uno). Cada expresión vocal contiene de una a dos palabras clave y viene con una metaetiqueta que simula la presencia de estas palabras clave en otra modalidad. Se consigue una precisión de reconocimiento del 100%

cuando se dedican aproximadamente 3000 expresiones vocales para el entrenamiento y las 1000 restantes para las pruebas.

5 El material de voz se cuantificó obteniendo un flujo de índices de cuantificación vectorial (VQ) (longitud de trama de 10 ms, tamaño de libro de códigos $N_c = 150$) con un algoritmo de agrupamiento de k-medias. La distancia euclidiana de vectores MFCC estáticos se usó como medida de distancia. El entrenamiento se realizó con una expresión vocal cada vez con su etiqueta de concepto asociada. En la fase de prueba sólo se usó el flujo VQ de una expresión vocal como entrada y el sistema tenía que reconocer qué palabra clave (etiqueta de concepto) se ha introducido en la expresión vocal. Las curvas de activación se sometieron a un filtro de mediana con formación de ventanas de 150 ms y L_{\min} se estableció en 350 ms y L_{\max} en 450 ms.

10 La figura 3 muestra la activación de las representaciones de concepto (o palabra) interior en la expresión vocal "Papá se acerca" ("*Daddy comes close*") en función del tiempo sin el filtrado de mediana de la activación. En la parte superior de la figura 3 se muestra la actividad de cada concepto en función del tiempo mostrando curvas separadas con valores de probabilidad y en la parte inferior están las mejores sumas acumulativas de las subsecuencias halladas para cada concepto. Los límites de la subsecuencia del concepto ganador se indican con flechas.

15 La figura 4 muestra el mismo procedimiento con el filtrado de mediana. Puede verse a partir de la figura que la palabra clave correcta "papá" ("*daddy*") se reconoce con un margen claro respecto a otros candidatos de palabras. Además, el filtrado de mediana integra información por una ventana temporal mayor y por tanto mejora la diferencia entre conceptos que reciben un soporte continuo de la entrada y aquéllos que sólo hallan aleatoriamente alguna estructura familiar en la misma. En la parte superior se encuentra la actividad de cada concepto en función del tiempo mostrando curvas separadas con valores de probabilidad y en la parte inferior se muestran las mejores sumas acumulativas de las subsecuencias halladas para cada concepto. Los límites de la subsecuencia del concepto ganador se indican con flechas.

20 La figura 5 muestra la curva de aprendizaje característica del algoritmo en función del número de expresiones vocales entrenadas. La curva de aprendizaje revela que la tasa de reconocimiento aumenta muy rápidamente después de sólo algunas muestras y consigue una precisión del 100% a las 3000 expresiones vocales. Las primeras 500 expresiones vocales se muestran en detalle en la parte derecha inferior.

25 La invención se explicó anteriormente con referencia a las realizaciones mencionadas anteriormente y se han mostrado las diversas ventajas de la invención. Resulta evidente que la invención no está limitada solamente a estas realizaciones, sino que comprende todas las posibles realizaciones dentro del alcance de las siguientes reivindicaciones de patente.

30

REIVINDICACIONES

1. Método (200) para reconocer un concepto en una señal, por ejemplo una señal de voz, mediante un aparato, comprendiendo el método:

recibir (210), mediante un receptor del aparato, una primera señal,

5 muestrear (215, 220, 225), mediante el aparato, la primera señal para formar una primera secuencia de símbolos unidimensional a partir de la primera señal recibida, por ejemplo usando cuantificación vectorial,

especificar la presencia de un concepto en la primera secuencia mediante una etiqueta de concepto, para cada etiqueta de concepto c_k , variando k desde 1 hasta N_c , y para cada retardo l_d dado, variando d desde 1 hasta N_l , donde un retardo l_d indica una distancia de d elementos en la secuencia de símbolos:

10 obtener (230, 235), mediante el aparato, una matriz de frecuencia de transición $T(a_i, a_j|l_d, c_k)$ de tamaño $N_q \times N_q$, donde N_q es el número de posibles símbolos, incrementando el elemento de matriz en la posición i, j siempre que los símbolos a_i y a_j aparezcan en conjunto en la primera secuencia con el retardo l_d en presencia de la etiqueta de concepto c_k ,

15 obtener (240), mediante el aparato, una matriz de probabilidad de transición $P'(a_i, a_j|l_d, c_k)$ normalizando la matriz de frecuencia de transición T mediante todas las posibles transiciones

$$P'(a_i, a_j|l_d, c_k) = \frac{T(a_i, a_j|l_d, c_k)}{\sum_{x=1}^{N_q} T(a_i, a_x|l_d, c_k)},$$

obtener, mediante el aparato, una matriz de probabilidad de etiqueta $P(a_i, a_j|l_d, c_k)$ mediante

$$P(a_i, a_j|l_d, c_k) = \frac{P''(a_i, a_j|l_d, c_k)}{\sum_{z=1}^{N_c} P''(a_i, a_j|l_d, c_z)} - \frac{1}{N_c},$$

donde

20
$$P''(a_i, a_j|l_d, c_k) = P'(a_i, a_j|l_d, c_k) + \frac{T(a_i, a_j|l_d, c_k)}{\sum_{x=1}^{N_q} \sum_{y=1}^{N_q} T(a_x, a_y|l_d, c_k)},$$

recibir (270), mediante el receptor del aparato, una segunda señal, y

reconocer (280, 285), mediante el aparato, la presencia de un concepto en la posición t en una segunda secuencia de símbolos, que se forma a partir de la segunda señal recibida, como la etiqueta de concepto c_k que maximiza el valor de

25
$$A(c_i, t) = \sum_{d=1}^{N_l} P(s[t - l_d], s[t]|l_d, c_i)$$

siendo i desde 1 hasta N_c .

2. Método según la reivindicación 1, que comprende además sumar las probabilidades de suma obtenidas en una ventana temporal que tiene una longitud que corresponde a una duración esperada del patrón que va a reconocerse para detectar la etiqueta de valor máximo.

30 3. Método según cualquier reivindicación anterior, en el que el resultado del reconocimiento se procesa mediante un filtrado de mediana o de paso bajo.

4. Método según cualquier reivindicación anterior, en el que una longitud de un filtro de mediana o paso bajo se estima mediante una longitud del patrón.

35 5. Método según cualquier reivindicación anterior, que comprende además almacenar (235) los datos de frecuencia recopilados de la incidencia de las transiciones en la primera secuencia en la matriz de frecuencia de transición T .

6. Método según cualquier reivindicación anterior, en el que la segunda señal se discretiza y cuantifica (275) para representarla en forma de segunda secuencia.
7. Método según cualquier reivindicación anterior, que se usa para un reconocimiento de voz o un reconocimiento de imagen.
- 5 8. Producto de programa informático configurado para ejecutar el método según cualquiera de las reivindicaciones 1-7, cuando el producto de programa informático se ejecuta mediante un procesador.
9. Aparato configurado para ejecutar el método según cualquiera de las reivindicaciones 1-7.

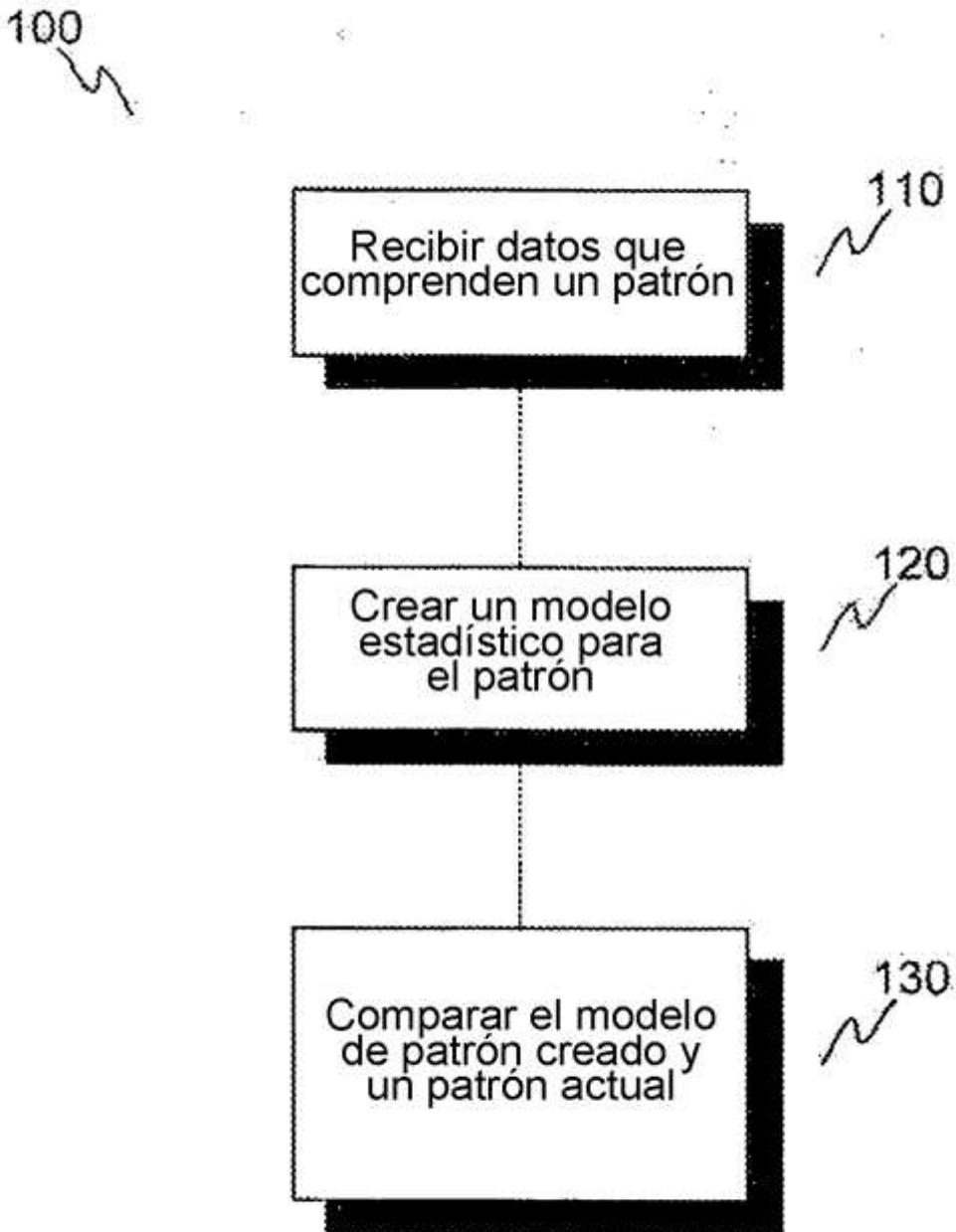


Figura 1

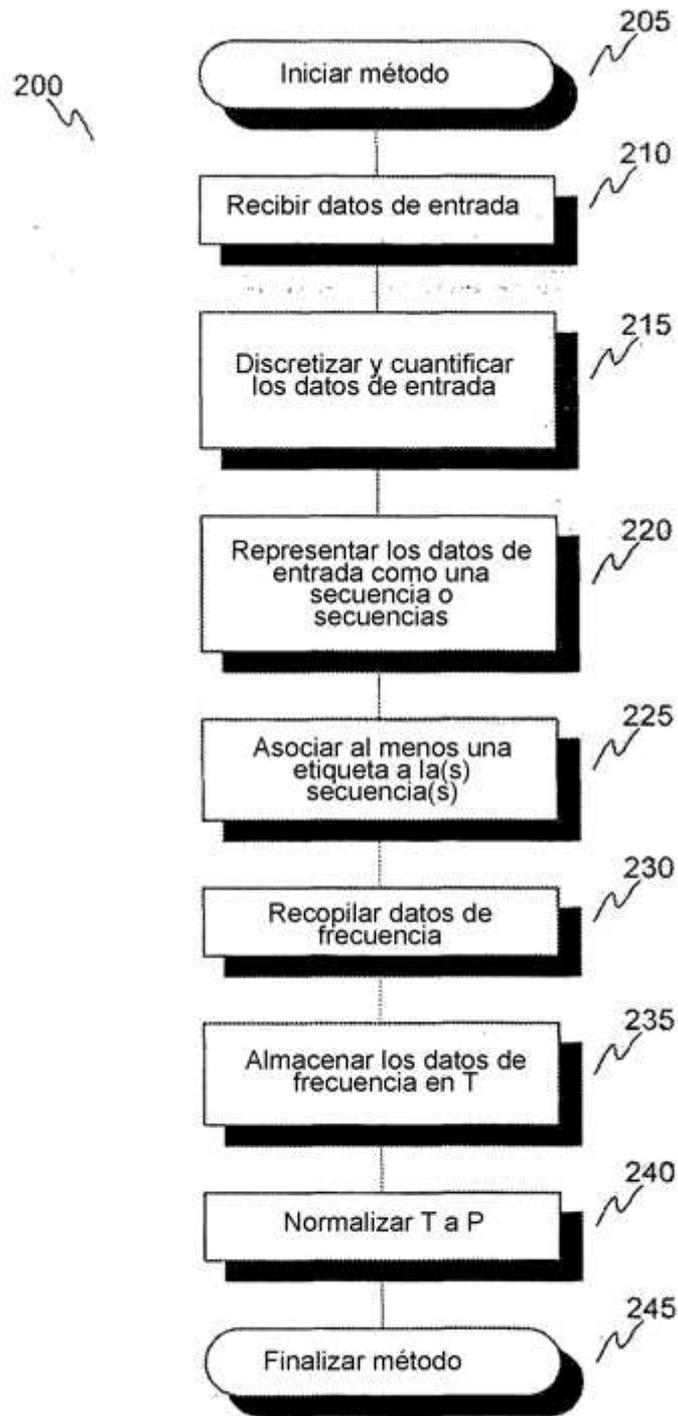


Figura 2A

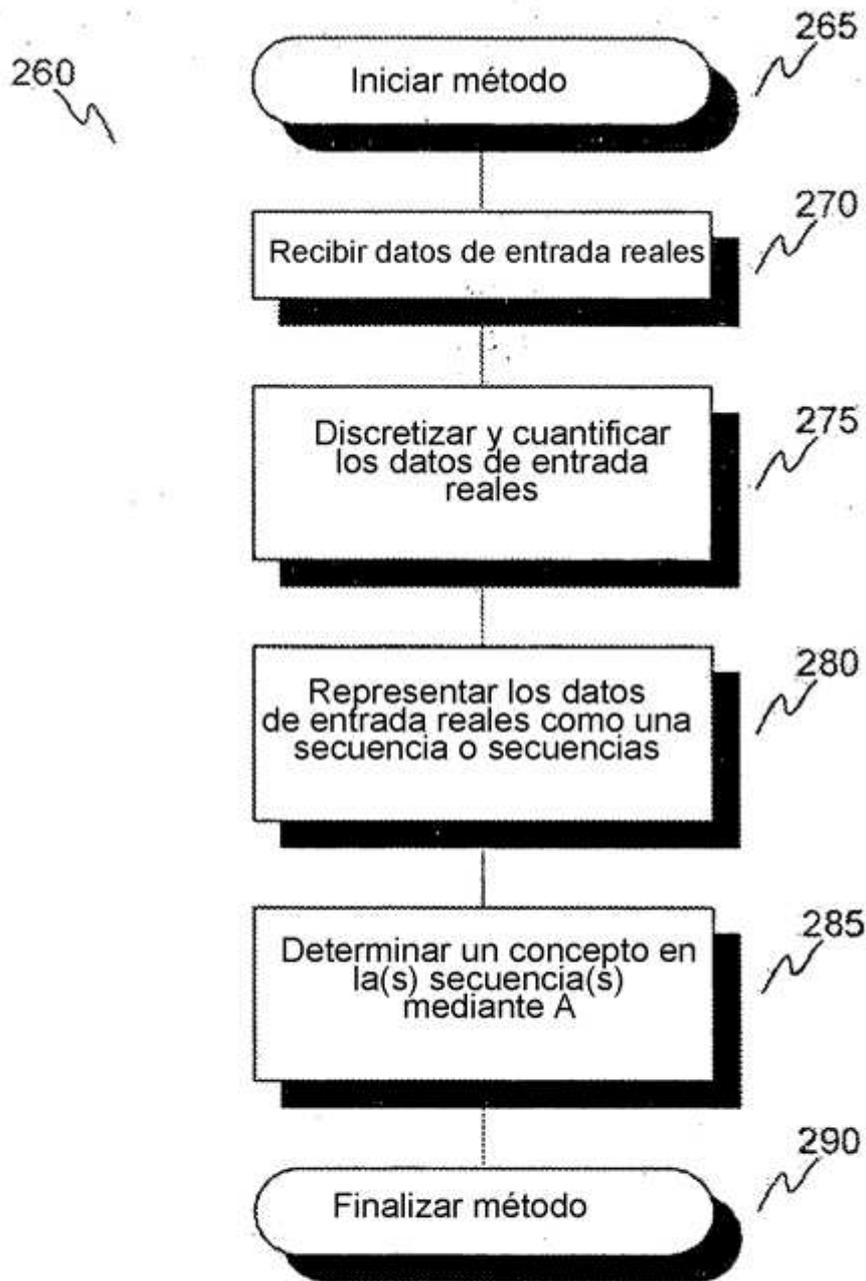


Figura 2B

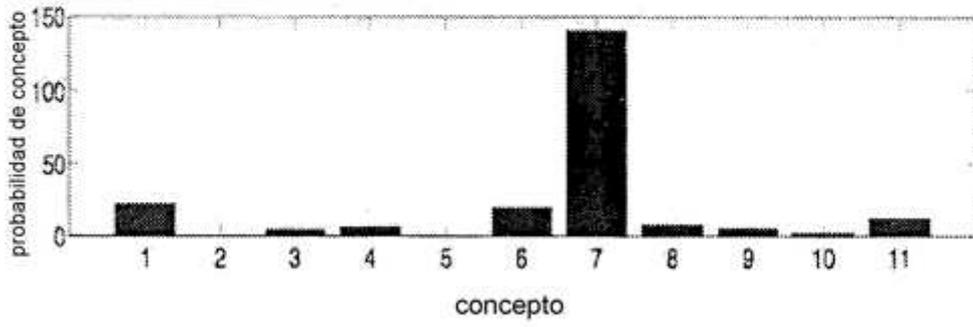
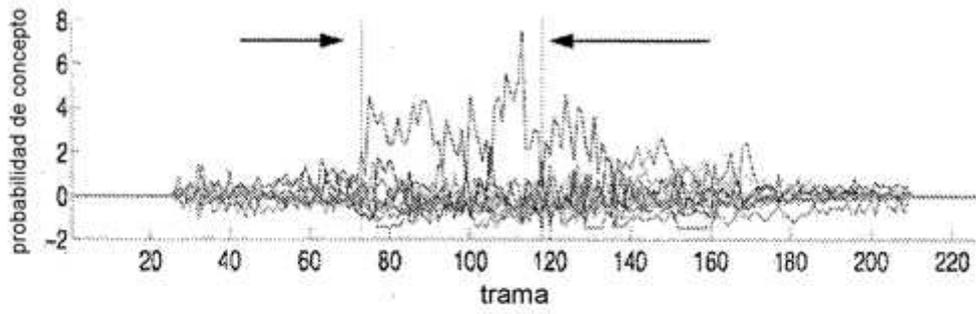


Figura 3

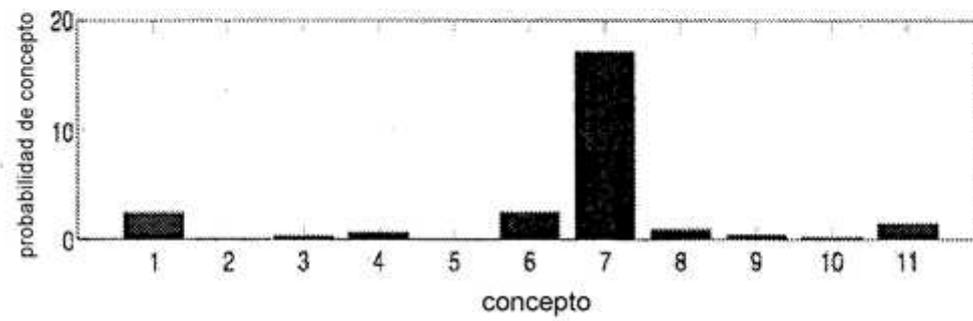
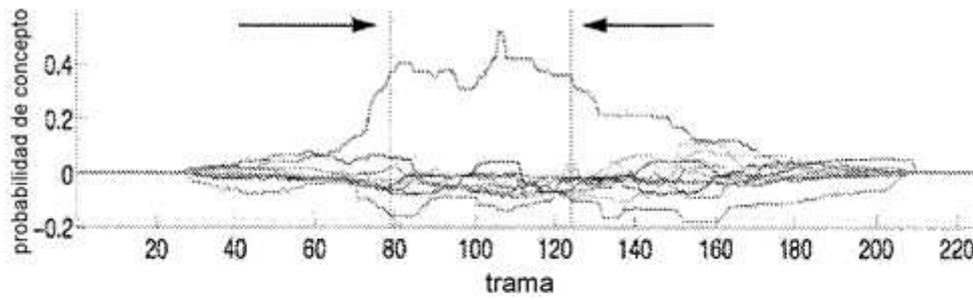


Figura 4

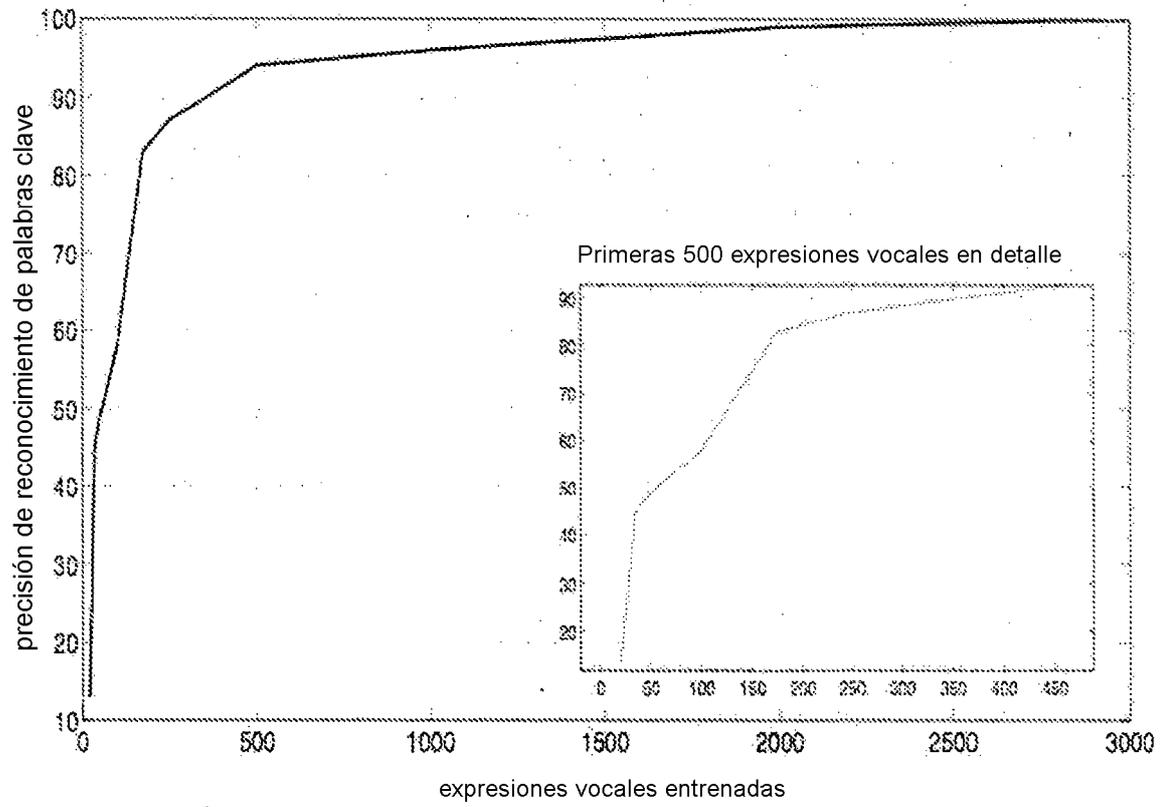


Figura 5