

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 605 779**

51 Int. Cl.:

**G10L 17/04** (2013.01)

**G10L 17/10** (2013.01)

**G10L 17/16** (2013.01)

**G10L 17/24** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.09.2012** **E 12006775 (6)**

97 Fecha y número de publicación de la concesión europea: **09.11.2016** **EP 2713367**

54 Título: **Reconocimiento de orador**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**16.03.2017**

73 Titular/es:

**AGNITIO S.L. (100.0%)**  
**C/ Gran Vía 39-8a planta**  
**28013 Madrid, ES**

72 Inventor/es:

**BUERA RODRÍGUEZ, LUIS;**  
**VAQUERO AVILÉS-CASCO, CARLOS;**  
**GARCÍA GOMAR, MARTA y**  
**ARTIAGA, ANTONIO MIGUEL**

74 Agente/Representante:

**MILTENYI, Peter**

ES 2 605 779 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

Reconocimiento de orador.

5 Los sistemas de reconocimiento de orador pueden utilizarse para confirmar o rechazar que una persona que está hablando es quien se ha indicado que es (verificación de quien habla) y también puede utilizarse para determinar  
 10 quién está hablando de entre una pluralidad de personas conocidas (identificación de quien habla). Tal sistema de identificación de orador puede ser de grupo abierto, ya que es posible que quien habla no sea una de las personas conocidas para el sistema, o de grupo cerrado, si quien habla se encuentra siempre en el conjunto del sistema.  
 Dichos sistemas pueden encontrar aplicación en banca telefónica, identificación de sospechosos y generalmente pueden utilizarse en un contexto relacionado con la seguridad.

15 Tales sistemas de reconocimiento de orador pueden exigir al usuario decir el mismo contenido léxico (por ejemplo, la misma frase clave), tanto para la inscripción como el reconocimiento. Dicho sistema es un sistema dependiente de texto, ofreciendo en algunos casos seguridad adicional, ya que requiere el reconocimiento de la identidad de quien habla, así como el contenido léxico de la expresión.

20 Dichos sistemas de reconocimiento también pueden ser independientes de texto, de modo que no se establece ninguna restricción respecto al contenido léxico de la inscripción y de las expresiones de reconocimiento. Tales sistemas pueden tener la ventaja de que las personas pueden identificarse, por ejemplo, a partir de conversaciones comunes, por ejemplo, conversaciones cotidianas o inscribirse con tales conversaciones comunes de las cuales ya existen archivos.

25 El documento US 2008/0312926 A1 describe la creación de huella vocal y reconocimiento de orador automático, dependiente de texto e independiente del idioma, basado en modelos ocultos de Markov (HMM) y sistemas de reconocimiento de voz automáticos (ASR).

30 El documento US 2007/0294083 A1 describe un procedimiento rápido, dependiente de texto e independiente del idioma, para la autenticación de usuario por voz basado en Deformación Dinámica de Tiempo (DTW).

El documento US 6.094.632 describe un dispositivo de reconocimiento de voz, donde se combinan el sistema ASR y salidas del sistema de identificación de voz (SID).

35 Patrick Kenny, en su artículo "*Joint Factor Analysis of Speaker Session Variability; Theory and Algorithms*", da una introducción a procedimientos relacionados con la verificación de orador, en particular, un algoritmo, que puede ser utilizado en sistemas de reconocimiento de orador.

El documento US 2007/198257 A1 se refiere a autenticación de orador.

40 El documento W 2007/131530 A1 se refiere a un procedimiento para compensar la variabilidad entre sesiones para la extracción automática de información de una señal de voz de entrada que representa una expresión de un orador.

45 El documento US 2010/0131273 describe un dispositivo, un sistema y un procedimiento para la detección de vida utilizando biometría de voz.

Otro documento de la técnica anterior es el documento "*Support Vector Machines versus Fast Scoring in the Low-Dimensional Total Variability Space for Speaker Verification*" de N. Dehak y otros en Interspeech, Brighton, Londres, septiembre de 2009.

50 Es conocido el uso de modelos ocultos de Markov (HMM) que consisten en un conjunto de estados que corresponden a un evento observable deterministamente y están conectados por arcos de probabilidad de transición. Los estados se definen en un vector de parámetros y se extraen de la señal de voz. Cada estado tiene asociada una función de densidad de probabilidad (pdf), que modela los vectores de características asociadas a ese estado. Tal función de densidad de probabilidad puede ser, por ejemplo, una mezcla de funciones Gaussianas (Modelos de Mezclas Gaussianas, GMM), en el espacio multidimensional de los vectores de características, pero pueden utilizarse también otras distribuciones.

60 El modelo oculto de Markov se define por las probabilidades de transición  $a_{ij}$  asociadas a los arcos que representan la probabilidad de pasar del estado  $i$  al estado  $j$ , las distribuciones de estado inicial  $p_i$ , que están asociadas a los estados y son las probabilidades iniciales de cada estado y la distribución de probabilidad de observación  $b_i$  que está asociada al estado  $i$  y puede ser, por ejemplo, un GMM. Esas distribuciones de probabilidad de observación están definidas por un conjunto de parámetros que dependen de la naturaleza de las distribuciones.

Las aproximaciones convencionales para el uso de modelos ocultos de Markov en marcos de reconocimiento de orador dependiente de texto generalmente requieren una transcripción de la frase utilizada que se requiere para construir el HMM de orador aplicando algún tipo de adaptación del HMM, por ejemplo, un máximo a posteriori (MAP) (tal como se describe, por ejemplo, en J. Gauvin y C. Lee "*Maximum Posteriori Estimation for Multivariate Gaussian Mixture Observations of Markov Chains*" *IEEE Transactions on Speech and Audio Processing*, 2(2): 291-298) o la regresión lineal de probabilidad máxima (MLLR) (tal como se describe, por ejemplo, en CJ Leggetter y P.C. Woodland en "*Maximum likelihood linear regression for speaker adaptation of the parameters of continuous density Hidden Markov Models*" u otras adaptaciones de un modelo de punto de partida como una concatenación de HMMs genéricos que representan unidades (por ejemplo, fonemas o palabras) de señales de audio, por ejemplo, la frase.

En este marco, los HMMs genéricos generalmente se denominan *Modelo de Fondo Universal*. A partir de esto puede calcularse una puntuación utilizando un algoritmo adecuado tal como, por ejemplo, el algoritmo de avance-retroceso o de Viterbi, tal como se describe en L. R. Rabiner "*a tutorial of Hidden Markov Models and selected applications in speech recognition*", Proc. de IEEE77 (2): 257-286, DOI: 10.1109/5. 18626. [1].

Tales HMMs genéricos normalmente requieren un entrenamiento supervisado ya que cada unidad (por ejemplo, fonema, palabra,...) tiene que asociarse a un determinado HMM. A partir de esto, el marco de reconocimiento de voz puede clasificarse en función de cómo se obtiene la transcripción. Las posibilidades de cómo puede obtenerse tal transcripción comprenden conocimiento previo, el uso de sistemas de reconocimiento de voz convencionales o el uso de sistemas de reconocimiento de voz universales tal como se describe, por ejemplo, en US 2008/0312926. Sin embargo, estos procedimientos generalmente requieren un entrenamiento supervisado y/o requieren cálculos intensos, requieren una gran cantidad de memoria, normalmente dependen del idioma y/o no son muy flexibles. Los enfoques clásicos para sistemas de reconocimiento de orador basados en HMM que dependen de texto pueden presentar el inconveniente, además, de que el modelo HMM de orador tiene una relación directa con la transcripción que puede robarse en por lo menos un punto del sistema.

En el reconocimiento de orador clásico utilizando técnicas de adaptación de HMM, toda la información de los vectores de características está incorporada en el modelo de orador, a pesar de que cierta información, tal como por ejemplo el canal, no es una característica típica de orador y, por lo tanto, no debería estar incluida en el modelo de orador.

Por estas razones, los procedimientos de reconocimiento de orador clásicos dependientes de texto tienen limitaciones considerables.

Algunos de sus problemas son los que se han descrito anteriormente; almacenamiento de la transcripción o una estimación de la transcripción de la frase del orador, el uso de un reconocimiento de orador o decodificador fonético que hace que el sistema utilice una gran cantidad de memoria y no sea apto para pequeños dispositivos tales como tabletas o teléfonos inteligentes, y el hecho de que no compensan el canal u otros efectos negativos de la señal de voz.

Preferiblemente, un sistema mejorado puede aprovecharse de la información de la secuencia temporal de los vectores de características y proporcionar un rendimiento y una precisión satisfactorios sin utilizar una transcripción de la expresión, por ejemplo, la frase del discurso.

La presente invención se refiere a un reconocimiento de orador automático y soluciona por lo menos uno o más de los problemas mencionados anteriormente. En particular, la invención se refiere a un reconocimiento de orador automático, independiente del idioma y dependiente de texto que se aprovecha de la correlación temporal de la información de orador de los vectores de características de la muestra de voz, pero sin incorporar su transcripción o estimación a cualquier punto del sistema. Por lo tanto, pueden conseguirse resultados satisfactorios con poca memoria y recursos de tiempo de cálculo, por lo que puede incorporarse incluso en dispositivos pequeños tales como, por ejemplo, teléfonos inteligentes.

La invención comprende una combinación de enfoques clásicos independientes de texto con un enfoque dependiente de texto con el fin de aprovechar la información dinámica. Efectivamente, los enfoques clásicos independientes de texto tales como análisis de factores conjuntos (AFC) tal como se describe, por ejemplo, en el artículo mencionado anteriormente "*Joint Factor Analysis of Speaker Session Variability: Theory and Algorithms*" de Patrick Kenny o paradigmas del vector-i tal como se explica, por ejemplo, en el documento "*Support Vector Machines versus Fast Scoring in the Low-Dimensional Total Variability Space for Speaker Verification*" de N. Dehak y otros en Interspeech, Brighton, Londres, septiembre de 2009, no utilizan correlación temporal de muestras de voz, pero compensan de manera muy eficiente los efectos del canal para extraer solamente información del orador. Por otra parte, los enfoques clásicos dependientes de texto aprovechan la correlación temporal teniendo en cuenta de este modo un punto clave en los problemas dependientes de texto, pero sin compensar efectos del canal de manera satisfactoria. Una combinación de estos paradigmas puede realizarse a nivel de sistema, por ejemplo, mediante la fusión de un sistema de reconocimiento de orador dependiente de texto y un sistema de reconocimiento de orador

independiente de texto o incorporando las virtudes de los independientes de texto en un marco de dependientes de texto.

5 La invención comprende un procedimiento de acuerdo con la reivindicación 1. En las reivindicaciones dependientes 2-13 se describen realizaciones favorables.

10 En particular, la invención comprende un procedimiento para el reconocimiento de orador dependiente de texto utilizando un modelo de orador obtenido por adaptación de un Modelo de Fondo Universal en el que el modelo de voz es un modelo oculto de Markov de orador adaptado que comprende corrección de canal. Dicho procedimiento puede comprender dos partes, a saber, inscripción, donde el modelo se adapta y la parte de reconocimiento en el que se verifica o se identifica el orador y se compara una prueba de audio con un conjunto de modelos o un modelo para la verificación.

15 Un Modelo de Fondo Universal genérico (UBM genérico) normalmente se adapta a un orador una vez que éste se ha inscrito.

20 Éste puede adaptarse en base a una o varias expresiones (expresiones de inscripción) del orador. A partir de tal expresión (por ejemplo, una frase, una palabra, un fragmento que normalmente esté presente como archivo de audio o información) pueden extraerse vectores de características (vectores de características de inscripción). Las expresiones comprenden información del orador e información del canal, en el que el canal es todo lo que hay entre el orador y el soporte de grabación, por ejemplo, comprendiendo un micrófono a través del cual se han transmitido las expresiones, y cualquier otra cosa a través de la cual se han transmitido las expresiones, por ejemplo, cables, altavoces, etc.

25 Esta adaptación puede realizarse de una manera no supervisada. En particular, puede no ser necesario ningún contenido léxico de las expresiones o una aproximación de los mismos. Por lo tanto, puede no ser necesario ningún reconocimiento de voz u otro procedimiento de transcripción.

30 Entonces puede construirse un modelo dependiente de texto para la(s) expresión(es) correspondiente(s) mediante la adaptación del Modelo de Fondo Universal genérico del sistema dependiente de texto con todos los vectores de características de inscripción de la(s) expresión(es) correspondiente(s) (en el que normalmente cada expresión comprende el mismo contenido léxico) y el orador. Los vectores de características de inscripción generalmente se extraen de las expresiones previstas para la inscripción. Puede aplicarse una corrección de canal de manera que el modelo de orador dependa sólo del orador, pero no del canal (por ejemplo, tal como se muestra en el documento de 35 Kenny, mencionado anteriormente).

40 Para cada orador pueden generarse varios modelos. Normalmente, un modelo de orador se adapta a partir del UBM para un contenido léxico, por ejemplo, una frase, una palabra o un fragmento. Cada modelo puede obtenerse con una, dos, o más expresiones, por ejemplo, cinco expresiones del mismo contenido léxico.

45 En particular, una vez que algunas expresiones de un orador están presentes como archivo de audio u otra información de audio, por ejemplo, una determinada frase, una palabra, un fragmento o algo similar, pueden extraerse vectores de características de tiempo corto. En tales vectores de características de tiempo puede utilizarse, por ejemplo, Coeficientes Cepstrales en las Frecuencias de Mel (MFCC) tal como muestra, por ejemplo, Davis, S. B. y Mermelstein, P. en "*Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences*", *IEEE Trans. on Acoustic, Speech and Signal Processing*, 28(4): 357-366, 1980. En los vectores de características de tiempo, toda la información relevante del orador, contenido léxico y otros aspectos (también efectos no deseados, por ejemplo, de canal) se comprimen en un pequeño vector en cada 50 intervalo de tiempo. Por ejemplo, cada 10 milisegundos puede crearse un vector de características 60 dimensional.

55 Con esta información, puede aplicarse un modelo de paradigma de análisis de factores conjuntos tal como se describe en el artículo mencionado anteriormente de Kenny para adaptar un UBM genérico, que es un HMM, que compensa al mismo tiempo aspectos no deseables de la señal de voz tal como, por ejemplo, el canal. Para un orador y un canal  $h$  dado, el modelo completo es:

$$M_h(s) = m + v y(s) + u x_h(s) + d z(s)$$

60 donde  $m$  es un supervector obtenido por la concatenación de todos los vectores medios del UBM genérico,  $M_h(s)$  es el supervector que modela el orador  $s$  y el canal  $h$ ,  $v$  es la matriz rectangular de voces propias,  $y(s)$  es la variable oculta que incluye la información de orador (factores de orador),  $u$  es la matriz rectangular de canales propios,  $x_h(s)$  es la variable oculta que incluye la información del canal (factores de canal),  $d$  es una matriz diagonal para modelar aquellos aspectos de la señal de voz que no están incluidos en  $u$  y  $v$ , y  $z(s)$  es una variable residual oculta asociada a  $d$ . El término  $d z(s)$  puede ser modelado como cero ( $y$ , por lo tanto, no tenerse en cuenta).

5 Para esta adaptación se requieren estadísticas cero y de primer orden para las expresiones que se utilizan para la inscripción para la adaptación del locutor del UMB (expresiones de inscripción). Puede llevarse a cabo algún algoritmo adecuado, tal como, por ejemplo, un algoritmo de avance-retroceso o un algoritmo de Viterbi para el UBM genérico y cada una de las correspondientes expresiones de la frase y el orador.

10 Para las expresiones de inscripción, el orador normalmente tiene que decir el mismo contenido léxico (por ejemplo, una frase, una palabra o un fragmento) más de una vez, en particular, más de 2, en particular, más de 3, o en particular más de 5 veces. Para la verificación (reconocimiento) en tales sistemas (explicado más adelante) puede ser suficiente una expresión.

15 Para cada orador, pueden crearse varios modelos (en particular, adaptados a partir de un UBM genérico) con diferentes contenidos léxicos. Para cada expresión con un contenido léxico para el reconocimiento de orador, en general se utiliza el modelo correspondiente a dicho contenido léxico.

Por ejemplo, utilizando un algoritmo de avance-retroceso, puede calcularse la probabilidad de que se encuentre en el Gaussiano  $k$  del estado  $i$  en el instante  $t$   $\gamma_{i,k}(t)$ . El Anexo muestra cómo pueden calcularse estas probabilidades.

20 A partir de esto, las estadísticas cero y de primer orden en el tiempo son:

$$N_{i,k} = \sum_t \gamma_{i,k}(t)$$

$$F_{i,k} = \sum_t \gamma_{i,k}(t) x_t$$

Aquí,  $x_t$  es el vector de características de inscripción en el intervalo de tiempo  $t$ .

25 A partir de esto, pueden extraerse factores de orador y canal y los vectores medios del HMM adaptado pueden darse como

$$\overline{M}_{l_c}(s) = m + v y(s).$$

30 Esta ecuación puede representar el medio del modelo adaptado de orador (modelo de orador) y también puede utilizarse en un proceso de reconocimiento más adelante. El índice  $l_c$  se refiere al contenido léxico, para el cual  $\overline{M}_{l_c}(s)$  está adaptado.  $y(s)$  puede calcularse a partir de la(s) expresión(es) de inscripción, tal como se muestra, por ejemplo, en el documento de Kenny, mientras que  $v$  (como  $u$  y  $d$ , si es necesario) pueden obtenerse en una fase de desarrollo previa (después de la formación sin supervisión del UBM genérico) tal como se muestra, por ejemplo, en el documento de Kenny.

35 Para fines prácticos, el sistema puede no guardar  $\overline{M}_{l_c}(s)$ , pero puede guardar solamente  $y(s)$ . Esto puede ser ventajoso ya que el vector  $y(s)$  puede tener unas dimensiones considerablemente menores que  $\overline{M}_{l_c}(s)$ . Debido a que  $v$  y  $m$  pueden tener que estar guardados de todos modos, el almacenamiento de  $y(s)$  en lugar de  $\overline{M}_{l_c}(s)$  puede reducir los recursos del sistema necesarios.

40 Las probabilidades de transición pueden adaptarse a partir del UBM genérico con el fin de completar el modelo de orador (modelo de orador adaptado), que pueden tener los mismos parámetros que el UBM excepto los medios, que pueden calcularse utilizando las ecuaciones que se han dado antes y las probabilidades de transición tal como se dan a continuación. Las probabilidades de transmisión del HMM genérico pueden adaptarse a partir de las probabilidades de transición conocidas del HMM genérico y pueden darse como

$$\hat{a}_{ij} = \frac{\sum_t \tau_t(i, j)}{\sum_j \sum_t \tau_t(i, j)}$$

$$\tau_t(i, j) = \frac{\alpha_i(t) a_{ij} b_j(x_{t+1}) \beta_j(t+1)}{\sum_j \alpha_i(t) a_{ij} b_j(x_{t+1}) \beta_j(t+1)}$$

El significado de las variables en estas ecuaciones puede ser, por ejemplo, tal como se explica en el anexo.

- 5 En algunas realizaciones de un procedimiento de acuerdo con la invención, solamente podrán adaptarse vectores medios y probabilidades de transición del UBM genérico en el modelo de orador, ya que han demostrado ser los parámetros más importantes. En otras realizaciones, todos los parámetros del UBM genérico pueden adaptarse al orador. En este punto, la inscripción puede completarse.
- 10 En un procedimiento de acuerdo con la invención, el Modelo de Fondo Universal genérico puede entrenarse en una formación sin supervisión antes de adaptarse a un orador y un contenido de texto determinado. En particular, puede entrenarse a partir de un conjunto de información de audio sin la información de las transcripciones utilizando un algoritmo adecuado tal como, por ejemplo, un algoritmo de Maximización de Expectativas (EM) (tal como se describe, por ejemplo, en la A.P. Dempster, N.M. Laird y D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society*, 39(1)) o un algoritmo variacional de Bayes (VB) (tal como se describe, por ejemplo, en C.M. Bishop, "Pattern Recognition and Machine Learning", Springer Verlag).

15 Para este entrenamiento no supervisado, pueden utilizarse expresiones de varios oradores, por ejemplo, de más de cinco, en particular, de más de diez, en particular, de más de veinte oradores que puedan hablar más de uno, en particular más de cinco, en particular más de diez, y en particular más de quince idiomas diferentes y/o expresiones de más de 1, en particular, más de 2, en particular más de 5, en particular más de 10 y, en particular más de 20 oradores húngaros para el entrenamiento sin supervisión del Modelo de Fondo Universal genérico. (Se considera que el húngaro abarca casi la totalidad de los posibles sonidos que pueden pronunciarse en la mayoría de los idiomas más hablados.) De uno, más o todos los oradores cuyas expresiones se utilizan pueden utilizarse más de una, en particular, más de 5 expresiones.

20 En una fase de dicho entrenamiento, se utilizan preferiblemente tantas conversaciones de libre expresión como sea posible para entrenar el Modelo de Fondo Universal genérico para estar preparado para cualquier expresión, por ejemplo, contraseña, frase o idioma que pueda utilizarse durante la operación.

30 Dicho entrenamiento no supervisado puede ser ventajoso, ya que la forma en que cada estado del Modelo de Fondo Universal genérico no tiene un significado físico (por ejemplo, no está limitado a un determinado fonema) aumentando así la seguridad y la reducción de los costes necesarios para crear dicho modelo de fondo universal genérico.

35 La topología del modelo de fondo universal genérico puede seleccionarse para comprender una probabilidad de transición de cada estado posible al mismo y cada otro estado posible asegurando de este modo que para cada expresión posible, por ejemplo, contraseña, frase o idioma pueda utilizarse en ese Modelo de Fondo Universal genérico.

40 Además, las distribuciones de probabilidad de observación pueden ser modelos de mezclas Gaussianas que opcionalmente comprendan matrices de covarianza diagonales y cuatro componentes.

45 En un procedimiento de acuerdo con la invención, el número de estados puede ser fijo.

En particular, el número de estados puede establecerse en menos de treinta estados, hasta treinta o más de treinta, o el número de estados puede establecerse al número de fonemas húngaros, o un número que pueda estimarse mediante un análisis de las propiedades espectrales de una señal de voz.

50 Esto puede realizarse, por ejemplo, cuando se necesita un UBM genérico preciso para un determinado contenido léxico, por ejemplo, una determinada frase, una determinada palabra o un determinado fragmento. Tal UBM genérico puede tener que entrenarse de nuevo para cada nuevo contenido léxico, por ejemplo, frase, palabra o fragmento.

En particular, puede suponerse que puede necesitarse tantos estados como cambios espectrales se detecten en una señal. Un algoritmo para encontrar el número de cambios espectrales puede estar basado en un análisis de predicción lineal suave de tiempo corto. Dados dos segmentos adyacentes, f y g, éstos pueden analizarse con el fin de determinar si hay un cambio espectral entre ellos o no. Una posible medida de divergencia puede calcularse mediante

$$D(f, g) = \log\left(\frac{a_g^H R_{ff} a_g}{a_f^H R_{ff} a_f}\right)$$

Si la medida de divergencia es superior a un cierto umbral, puede suponerse que el cambio espectral estaba presente entre segmentos f y g.  $R_{ff}$  puede ser la señal de autocorrelación del segmento f, y  $a^*$  (donde \* puede ser f o g) puede ser los coeficientes de filtro del filtro de predicción lineal óptimo para \* por ejemplo extraído de la señal de autocorrelación resolviendo la predicción de error problema de mínimos cuadrados. Esto puede ser similar al detector de actividad de voz (VAD) de GSM (ETSI EN 300 730).

El procedimiento puede comprender, además, adaptar uno o más parámetros, por ejemplo, una serie de estados y/o, por ejemplo, el Modelo de Fondo Universal genérico a un contenido léxico, por ejemplo, una frase de acceso o contraseña. Por lo tanto, si se utiliza un cierto contenido léxico y ya es conocido, puede no ser necesario que cubra todos los sonidos con el Modelo de Fondo Universal genérico. En particular, el resto de los parámetros del Modelo de Fondo Universal genérico pueden encontrarse entonces utilizando datos y algoritmos de adaptación de HMM clásico tal como por ejemplo MAP. Para ello, pueden utilizarse los datos de aproximadamente 10 oradores, por ejemplo, 5-15, que dicen el contenido léxico (por ejemplo, una frase (de contraseña), una palabra (contraseña) o un fragmento) unas pocas veces, por ejemplo, 2-5 veces, y puede ser ya suficiente.

Después de eso, puede entrenarse una matriz de canales propios u, y una matriz de voces propias v, en una sesión de desarrollo antes de que el UBM genérico pueda adaptarse para un orador. Teniendo estas dos matrices puede realizarse una corrección de canal ya que la información del canal (matriz de canales propios) se mantiene separada de la información de orador (matriz de voces propias), que puede ser procesada por separado. También puede entrenarse en esta sesión de desarrollo en la matriz d, que puede modelar los aspectos de la señal de voz no incluidos en u y v. En otras realizaciones, d puede modelarse como 0, no estando presente, no siendo necesario, por lo tanto, que se entrene.

Un procedimiento de acuerdo con la invención también puede comprender la etapa de verificar si la persona objetivo dijo una señal de prueba. Tal verificación puede realizarse de una manera sin supervisión. Por lo tanto, puede no ser necesaria ninguna transcripción o aproximación del contenido léxico para la etapa de verificación, y, en particular, por ejemplo, puede no ser necesario ningún reconocimiento de voz o procedimiento de transcripción.

Para ello, pueden crearse vectores de características de prueba  $\{x_1, x_2, x_3, \dots, x_T\}$  a partir de la señal de prueba que se ha dicho.

En particular, por ejemplo, puede utilizarse MFCC en el que en cada 10 milisegundos se crea un vector de características 60 dimensional.

Además, puede utilizarse para el cálculo el modelo de orador representado por los factores de orador y(s) y estadísticas cero y de primer orden calculadas con los vectores de características de prueba.

Las estadísticas pueden calcularse utilizando el UBM genérico y el modelo de orador con un algoritmo adecuado, por ejemplo, un algoritmo de avance-retroceso o de Viterbi. Entonces pueden calcularse las probabilidades de registro del modelo de orador y el modelo UBM genérico  $LL_{UBM}$ .

En particular, puede llevarse a cabo un algoritmo adecuado tal como, por ejemplo, el algoritmo de avance-retroceso o de Viterbi con los vectores de características de prueba y el HMM adaptado de orador, en el que los medios y las probabilidades de transición pueden ser los adaptados.

Dicho algoritmo adecuado tal como, por ejemplo, el algoritmo de avance-retroceso o de Viterbi, puede proporcionar entonces un camino más probable para los vectores de características (vectores de características de pruebas) en los estados de HMM adaptado de orador en este caso  $q = \{q_1, q_2, \dots, q_T\}$ .

Entonces, la probabilidad de encontrarse en el Gaussiano k del estado i del UBM genérico en el instante t puede calcularse como

$$\gamma_{i,k}^{vit}(t) = \begin{cases} \frac{w_{i,k} b_{i,k}(x_t; \mu_{i,k}, \Sigma_{i,k})}{\sum_k w_{i,k} b_{i,k}(x_t; \mu_{i,k}, \Sigma_{i,k})} & \text{si } i = q_t \\ 0 & \text{si no} \end{cases}$$

y las estadísticas cero y de primer orden pueden darse como

5

$$N_{i,k} = \sum_t \gamma_{i,k}^{vit}(t)$$

$$F_{i,k} = \sum_t \gamma_{i,k}^{vit}(t) x_t$$

Aquí,  $x_t$  es el vector de características en el intervalo de tiempo  $t$  y  $\mu_{i,k}$  y  $\Sigma_{i,k}$  son la media y la matriz de covarianza del Gaussiano  $k$  de estado  $i$  del UBM genérico.

10

A continuación, puede obtenerse una probabilidad de registro de los vectores de características relativas al HMM adaptado de orador, por ejemplo, descrito en el documento de Kenny (LL<sub>SPK</sub>). Aquí, el modelo adaptado de orador puede utilizarse solamente para determinar el camino (más probable), pero no para calcular las estadísticas.

15

Esto puede ser contrario al enfoque AFC clásico utilizando el modelo genérico para determinar un camino (más probable), produciendo diferente desalineación, especialmente en enfoques de dependientes de texto. Sin embargo, también puede adoptarse el enfoque AFC clásico.

20

A continuación, puede llevarse a cabo una etapa correspondiente con un UBM genérico y los vectores de características de prueba, lo que da lugar a un camino más probable  $Q^{UBM} = \{q_{UBM1}, q_{UBM2}, \dots, q_{UBMT}\}$ . Para esto, de nuevo, puede utilizarse un algoritmo adecuado tal como, por ejemplo, el algoritmo de avance-retroceso o de Viterbi.

La probabilidad de encontrarse en el estado Gaussiano  $k$  del UBM genérico puede calcularse como

$$\gamma_{i,k}^{UBM,vit}(t) = \begin{cases} \frac{w_{i,k} b_{i,k}(x_t; \mu_{i,k}, \Sigma_{i,k})}{\sum_k w_{i,k} b_{i,k}(x_t; \mu_{i,k}, \Sigma_{i,k})} & \text{si } i = q_{UBM,t} \\ 0 & \text{si no} \end{cases}$$

25

con correspondientes estadísticas cero y de primer orden:

$$N_{i,k}^{UBM} = \sum_t \gamma_{i,k}^{UBM,vit}(t)$$

$$F_{i,k}^{UBM} = \sum_t \gamma_{i,k}^{UBM,vit}(t) x_t$$

30

De nuevo,  $x_t$  es un vector de características en el intervalo de tiempo  $t$ . Entonces, puede calcularse la relación de probabilidad de registro de los vectores de características (por ejemplo, tal como se describe en el documento de Kenny (LL<sub>UBM</sub>)).

35

Como medida final para la comparación entre la señal (expresión) de prueba, por ejemplo, un archivo de audio con una expresión, por ejemplo, frase prueba o una contraseña hablada por un orador respecto a un modelo de orador seleccionado tiene la relación de probabilidad de registro (LLR<sub>(i)</sub>) calculada como

$$LLR_{td} = LL_{spk} - LL_{UBM} + \sum_{i=1}^{T-1} \log(\hat{a}_{q_i q_{i+1}}) - \sum_{i=1}^{T-1} \log(a_{q_{UBM,i} q_{UBM,i+1}})$$

5 En este caso, la relación de probabilidad de registro puede comprender también, además de los registros de la probabilidad ( $LL_{spk}$  y  $LL_{UBM}$ ), el logaritmo de términos que describen probabilidades de transición adaptada y probabilidades de transición genérica de orador en los caminos más probables de los modelos correspondientes que se han calculado. Esto puede ser ventajoso, ya que puede tener en cuenta la progresión temporal a lo largo de los HMMs.

10 Tal procedimiento puede comprender, además, una etapa de identificación de una persona objetivo mediante la identificación de orador de un modelo de orador adaptado del Modelo de Fondo Universal genérico con la puntuación de probabilidad más alta antes de verificar si la persona objetivo es precisamente la que ha dicho la señal de prueba, tal como se ha explicado antes.

15 Un ejemplo de otra implementación también puede comprender un procedimiento para el reconocimiento de orador dependiente de texto utilizando una combinación de un sistema dependiente de texto y uno independiente de texto, en el que el sistema dependiente de texto está adaptado y en el que, además, hay también un modelo del sistema independiente de texto para el orador y la frase.

20 La adaptación puede realizarse opcionalmente de manera no supervisada (modo no supervisado). En particular, puede no ser necesario ningún contenido léxico de las expresiones o aproximación de las mismas. Por lo tanto, puede no ser necesario ningún reconocimiento de voz u otro procedimiento de transcripción.

25 El modelo dependiente de texto puede utilizar un UBM genérico, por ejemplo, un HMM como punto de partida. En particular, puede ser un UBM que sea entrenado de manera no supervisada y que pueda presentar una topología tal como se ha explicado anteriormente. Por lo tanto, puede no ser necesaria ninguna transcripción del contenido léxico para el entreno y puede no ser necesaria ninguna transcripción del contenido léxico para que el siguiente se adapte a un orador. El UBM genérico puede ser por ejemplo un HMM y puede estar adaptado al orador, por ejemplo, mediante un algoritmo adecuado tal como MAP o un algoritmo Bayesiano (tal como se describe, por ejemplo, en C.M. Bishop "Pattern Recognition and Machine Learning", Springer Verlag).

30 Cuando se combina un sistema de reconocimiento de orador dependiente de texto e independiente de texto, el independiente de texto puede estar compuesto por un AFC clásico o un marco de vectores  $i$ , y puede dar una puntuación o una relación de probabilidad de registro con compensación de canal. Por otra parte, el sistema dependiente de texto puede no realizar ningún tipo de compensación de canal, y puede dar una puntuación o relación de probabilidad de registro.

El procedimiento puede ser de la siguiente manera:

40 Para el procedimiento pueden utilizarse, por ejemplo, vectores de características tal como se ha mencionado anteriormente, por ejemplo, con los parámetros de intervalos de tiempo de 10 ms y MFFC 60 dimensional.

45 Para el sistema dependiente de texto, el UBM genérico, que puede ser un HMM, puede adaptarse al orador con algún audio de inscripción por MAP, utilizando un algoritmo adecuado tal como, por ejemplo, utilizando un algoritmo de avance-retroceso o de Viterbi. Entonces, la inscripción puede terminarse. Para el reconocimiento, puede aplicarse un algoritmo adecuado tal como, por ejemplo, de Viterbi, con el audio de prueba y el modelo de orador, dando  $LLR_{spk}$ . El algoritmo adecuado tal como, por ejemplo, de Viterbi, puede aplicarse también con el audio de prueba y el UBM genérico, dando  $LLR_{UBM}$ . Puede no ser necesaria ninguna compensación de canal y, por lo tanto, la compensación de canal puede no realizarse. Finalmente, la  $LLR_{td}$  puede calcularse como

50  $LLR_{td} = LLR_{spk} - LLR_{UBM}$

55 Para el sistema independiente de texto, el UBM genérico, que puede ser un GMM, puede utilizarse para generar el modelo de orador utilizando AFC o un marco de vectores  $i$ . Entonces, la inscripción puede terminarse. Aquí, puede realizarse la compensación de canal. Para el reconocimiento, puede aplicarse AFC o un marco de vectores  $i$ , permitiendo así el cálculo de  $LLR_{ti}$ .

Finalmente, el registro final de relación de probabilidad de registro (LLR) puede obtenerse como:

60  $LLR = \alpha LLR_{td} + \beta LLR_{ti} + \delta.$

El audio de inscripción y el audio de prueba pueden ser el mismo para ambos sistemas, por lo que el usuario puede tener que decir el mismo contenido léxico, por ejemplo, frase (de contraseña) para inscripción y reconocimiento.

En tal procedimiento, las probabilidades de transición y los medios de la distribución de densidad de probabilidad, que pueden ser por ejemplo GMM, del UBM genérico del sistema dependiente de texto, que puede ser un HMM, pueden modificarse en tal enfoque. Se ha demostrado que esos parámetros son los más importantes. Si se utiliza una adaptación MAP, se requiere un etiquetado de cada vector de características de adaptación asociando un estado. Tal etiquetado puede realizarse mediante un algoritmo adecuado, tal como el algoritmo de avance-retroceso o algoritmo de Viterbi.

Para la construcción del modelo del sistema independiente de texto para la frase y orador correspondiente, puede elegirse también un enfoque adecuado tal como, por ejemplo, análisis de factores conjuntos o paradigmas de vectores  $i$  tal como describe Dehak y otros. En tal construcción, pueden utilizarse los mismos vectores de características como para adaptar el modelo del sistema dependiente de texto.

Como UBM genérico para el sistema dependiente de texto, puede utilizarse GMM, sin embargo, en otras realizaciones también es posible utilizar HMM.

Tal procedimiento también puede comprender la etapa de verificar si la persona objetivo dijo una señal de prueba. En particular, tal etapa de verificación puede comprender, por ejemplo, la extracción de vectores de características de tiempo corto de la señal de prueba, por ejemplo, de una frase de prueba o contraseña. Esos vectores de características pueden ser extraídos por MFCC, por ejemplo, donde los parámetros pueden ser, por ejemplo, una distancia temporal de 10 milisegundos con un vector de características 60 dimensional. Teniendo en cuenta esos vectores de características, puede calcularse una relación de probabilidad de registro de orador adaptado y el modelo no adaptado (UBM genérico) para el sistema independiente de texto ( $LLR_{ti}$ ). Tal verificación puede realizarse de una manera no supervisada, lo que puede tener las ventajas que se han descrito anteriormente respecto a una manera no supervisada en una etapa de verificación.

A continuación, el vector de características de orador también puede compararse con el modelo genérico del sistema dependiente de texto (con el cual se encuentra  $LL_{UBM}$ ), por ejemplo; un HMM y el modelo de orador del sistema dependiente de texto (con el cual se encuentra  $LL_{SPK}$ ) construido previamente utilizando un algoritmo adecuado tal como por ejemplo MAP o MLLR utilizando, por ejemplo, el algoritmo de avance-retroceso o algoritmo de Viterbi. A partir de las dos probabilidades de registro, una relación de probabilidad de registro combinado para los modelos dependientes de texto puede calcularse como:

$$LLR_{td} = LL_{SPK} - LL_{UBM}$$

La relación de probabilidad de registro final (LLR) puede obtenerse como una combinación lineal de las relaciones de probabilidad de registro del dependiente de texto y del independiente de texto al cual se añade un término independiente. Puede darse como:

$$LLR = \alpha LLR_{td} + \beta LLR_{ti} + \delta$$

donde  $LLR_{td}$  es la relación de probabilidad de registro del sistema dependiente de texto y  $LLR_{ti}$  es el cociente de probabilidad de registro del sistema independiente de texto y  $\delta$  es un término independiente. Los valores escalares  $\alpha$  y  $\beta$  que son los coeficientes para la relación de probabilidad de registro del sistema dependiente de texto y la relación de probabilidad de registro del sistema independiente de texto y un término independiente  $\delta$  puede entrenarse durante una sesión de desarrollo con datos externos, que suele ser independiente del orador. El término  $\delta$  puede utilizarse debido al significado físico de LLR.

La LLR se puede utilizar en escenarios forenses, por ejemplo, LLR puede representar la relación entre la hipótesis de que los dos audios que se comparan los ha dicho la misma persona y la hipótesis de que ambos audios no los ha dicho la misma persona.

Entonces, una LLR mayor que 0 puede significar que es más probable que los dos audios hayan sido dichos por la misma persona que no hayan sido dichos por la misma persona.

Sin embargo, puede que no sea simple y puede que el umbral no sea normalmente 0 dado que puede tenerse que considerar la información a priori que puede tener en cuenta alguna otra información que puede no ser la voz, y las penalizaciones de tener un error (falsa aceptación y falso rechazo).

Debido a eso, puede ser importante que la LLR pueda estar bien calibrada. En ese caso, puede fijarse un umbral teniendo en cuenta la información a priori y los costes de los errores muy fácilmente.

Debido a eso,  $\delta$  puede ser necesario y entrenarse. En algunas otras realizaciones,  $\delta$  puede establecerse en 0.

5 Tal combinación lineal de las dos relaciones de probabilidad de registro dependiente de texto y el sistema independiente de texto permite un entreno adecuado para lograr mejores resultados que cuando se utiliza un sistema dependiente de texto o independiente de texto sólo. De nuevo, para la verificación, la relación de probabilidad de registro final LLR puede compararse con un umbral.

10 Tal procedimiento también puede comprender la etapa de identificación de una persona objetivo mediante la identificación del modelo de orador adaptado con la puntuación de probabilidad más alta como posible orador objetivo antes de verificar que la señal de prueba la dijo ese orador.

La invención también comprende un medio legible por ordenador de acuerdo con la reivindicación 14.

15 La invención comprende, además, un sistema que comprende medios de acuerdo con la reivindicación 15.

La invención se explica adicionalmente con las figuras, en las cuales

20 La figura 1 muestra sistemas dependiente de texto de la técnica anterior; y  
La figura 2 muestra un sistema que comprende los aspectos de la invención.

En la figura 1, se explica un sistema de reconocimiento de orador de la técnica anterior.

25 A partir de una señal de audio se extrae una transcripción de texto. Después de eso, puede crearse un HMM de partida a partir del UBM genérico utilizando la transcripción o una estimación de la transcripción. Entonces, puede generarse un HMM adaptado de orador utilizando la señal de audio y el HMM de partida. Para la adaptación, puede utilizarse un algoritmo adecuado, tal como MAP o MLLR que pueda utilizar, además, uno o más algoritmos adecuados de modo que puede ser necesario también, por ejemplo, en algunos casos, un algoritmo de Viterbi o de avance-retroceso.

30 A continuación, durante la fase de prueba puede utilizarse un algoritmo adecuado tal como, por ejemplo, algoritmos de Viterbi y/o de avance-retroceso para decidir si la persona objetivo es de hecho quien ha dicho la señal de prueba. Tal decisión puede tomarse comparando un marcador frente a un umbral establecido.

35 La figura 2 muestra un sistema de acuerdo con una realización de la invención. En particular, partiendo de un UBM genérico, el UBM genérico, que puede ser un HMM, puede adaptarse utilizando una o más expresiones, por ejemplo, señales de audio. Así, puede generarse un HMM adaptado de orador. Tal como puede verse, en tal sistema, puede no ser necesaria ninguna transcripción del contenido del archivo de audio.

40 Después de la fase de inscripción, puede llevarse a cabo una prueba utilizando el HMM adaptado de orador en combinación con un archivo de audio de prueba y también utilizando el UBM genérico. Puede tomarse una decisión entonces acerca de la verificación de la comparación de la relación de probabilidad de registro del modelo de orador modificado frente al UBM no modificado genérico contra un umbral.

45 La figura 2 puede representar una solución dependiente de texto en la que el canal se compensa o una solución dependiente de texto en la que el canal no se compensa. Sin embargo, la figura 2 no muestra una solución independiente de texto que también puede utilizarse en algunas realizaciones de la invención.

50 En particular, en algunas realizaciones en las que el canal no se compensa en el sistema dependiente de texto, puede ser necesaria una fusión de la puntuación con la puntuación del sistema independiente de texto (no mostrado).

Anexo:

55 Dados los vectores de características para una expresión  $X = \{x_1, x_2, x_3, \dots\}$  y un HMM definido por

- $A = \{a_{ij}\}$ , las probabilidades de transición, que están asociadas a los arcos y representan la probabilidad de "pasar" del estado  $i$  al estado  $j$ .
  - $\Pi = \{p_i\}$ , la distribución de estado inicial, que se asocia a los estados y son las probabilidades iniciales de cada estado.
- 60

- $b_i$ , distribución de probabilidad de observación, que está asociada al estado  $i$  y es un GMM, definido por:

- $w_{i,k}$ : una probabilidad a priori del Gaussiano  $k$  del estado  $i$ .

$$b_{i,k}(x) = N\left(x; \mu_{i,k}, \Sigma_{i,k}\right) : \text{donde } \mu_{i,k} \text{ y } \Sigma_{i,k} \text{ son la media y la matriz de covarianza del Gaussiano } k \text{ de estado } i.$$

5

Puede definirse la probabilidad de producir la secuencia  $\{x_1, x_2, x_3, \dots, x_{t-1}\}$  mientras termina en el estado  $i$  en el instante  $t$  como  $\alpha_i(t)$ :

$$\alpha_i(t) = b_i(x_t) \sum_j a_{ji} \alpha_j(t-1)$$

10

$$b_i(x_t) = \sum_k w_{i,k} b_{i,k}(x_t)$$

donde es la probabilidad del estado  $i$  para  $x_t$ . Puede definirse la probabilidad de producir la secuencia  $\{x_t, x_{t-1}, \dots, x_1\}$ , donde  $T$  es el número total de vectores de características, dado que nos encontramos en el estado  $i$  en el instante  $t$  como  $\beta_i(t)$ :

15

$$\beta_i(t) = \sum_j a_{ij} \beta_j(t+1) b_j(x_{t+1})$$

Finalmente, la probabilidad de encontrarse en el Gaussiano  $k$  de estado  $i$  en el instante  $t$  es  $\gamma_{i,k}(t)$ :

$$\gamma_{i,k}(t) = \frac{\beta_i(t) w_{i,k} b_{i,k}(x_t) \sum_j a_{ji} \alpha_j(t-1)}{\sum_j \alpha_j(t) \beta_j(t)}$$

20

Los valores de inicialización para su uso en un algoritmo, por ejemplo, algoritmo de avance-retroceso, pueden ser  $\alpha_i(1) = p_i$  y  $\beta_i(T+1) = 1$ .

25

**REIVINDICACIONES**

- 5 1. Procedimiento para el reconocimiento de orador dependiente de texto utilizando un modelo de orador obtenido mediante la adaptación de un Modelo de Fondo Universal; en el que el modelo de orador es un modelo oculto de Markov adaptado de orador que comprende corrección de canal, en el que el procedimiento comprende adaptar un Modelo de Fondo Universal genérico a un contenido léxico y entrenar una matriz de canales propios y una matriz de voces propias en una sesión de desarrollo antes de adaptar el Modelo de Fondo Universal genérico para un orador.
- 10 2. Procedimiento para el reconocimiento de orador de acuerdo con la reivindicación 1, en el que el Modelo de Fondo Universal es no supervisado adaptado en base a expresiones de inscripción del orador, en el que se extraen estadísticas utilizando vectores de características de inscripción, el Modelo de Fondo Universal y un algoritmo de avance-retroceso o de Viterbi.
- 15 3. Procedimiento para el reconocimiento de orador de acuerdo con la reivindicación la reivindicación 1 ó 2, en el que solamente vectores medios y probabilidades de transición están adaptados en el modelo de orador o en el que todos los parámetros están adaptados en el modelo de orador.
- 20 4. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 3, en el que el Modelo de Fondo Universal del sistema dependiente de texto se entrena en un entrenamiento no supervisado antes de que se adapte.
- 25 5. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 4, en el que se utilizan expresiones de más de 10 oradores que hablan más de 5 idiomas diferentes o más de 5 oradores húngaros para un entrenamiento no supervisado del Modelo de Fondo Universal del sistema dependiente de texto.
- 30 6. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 5, en el que la topología del Modelo de Fondo Universal del sistema dependiente de texto se selecciona para comprender una posibilidad de transición de cada estado posible al mismo y cada otro estado posible.
- 35 7. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 6, en el que las distribuciones de probabilidad de observación son modelos de mezcla Gaussianos con matrices de covarianza diagonal y cuatro componentes;
- 40 8. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 7, en el que el número de estados es fijo.
- 45 9. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 8, en el que el número de estados se establece en menos de 30, 30, más de 30, el número de fonemas húngaros o un número estimado de un análisis de las propiedades espectrales de una señal.
- 50 10. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 9, que comprende, además, la etapa de verificar en un modo no supervisado si una señal de prueba la dijo una persona objetivo.
- 55 11. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 10, en el que el modelo adaptado de orador se utiliza sólo para determinar el camino más probable, pero no para calcular las estadísticas.
- 60 12. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 11, en el que la verificación de si la señal de prueba la dijo la persona objetivo comprende calcular la diferencia entre los dos términos de la probabilidad de registro del audio de prueba y el modelo de orador y el producto de registro de las probabilidades de transición del camino más probable obtenido con el modelo de orador y la probabilidad de registro del audio de prueba y el UBM genérico y el producto de registro de las probabilidades de transición del camino más probable obtenido con el UBM genérico.
13. Procedimiento para el reconocimiento de orador de acuerdo con una de las reivindicaciones 1 a 12, en el que el procedimiento comprende, además, una etapa de identificación de una persona objetivo mediante la identificación del modelo adaptado de orador con la puntuación de probabilidad más alta.
14. Medio legible por ordenador que comprende instrucciones legibles por ordenador para ejecutar un procedimiento de acuerdo con una de las reivindicaciones 1 a 13.

15. Sistema que comprende medios para ejecutar un procedimiento de acuerdo con una de las reivindicaciones 1 a 13.

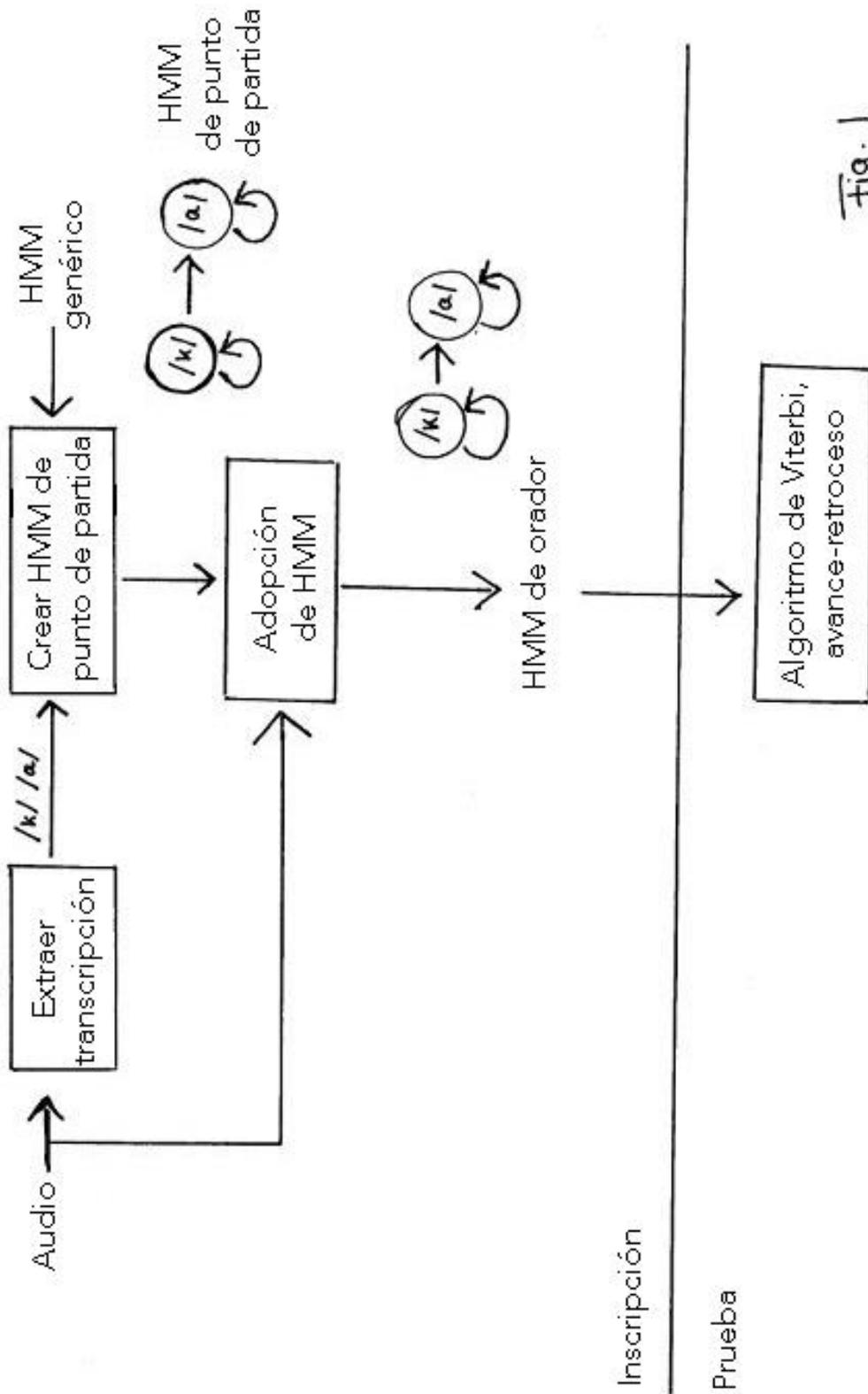


Fig. 1

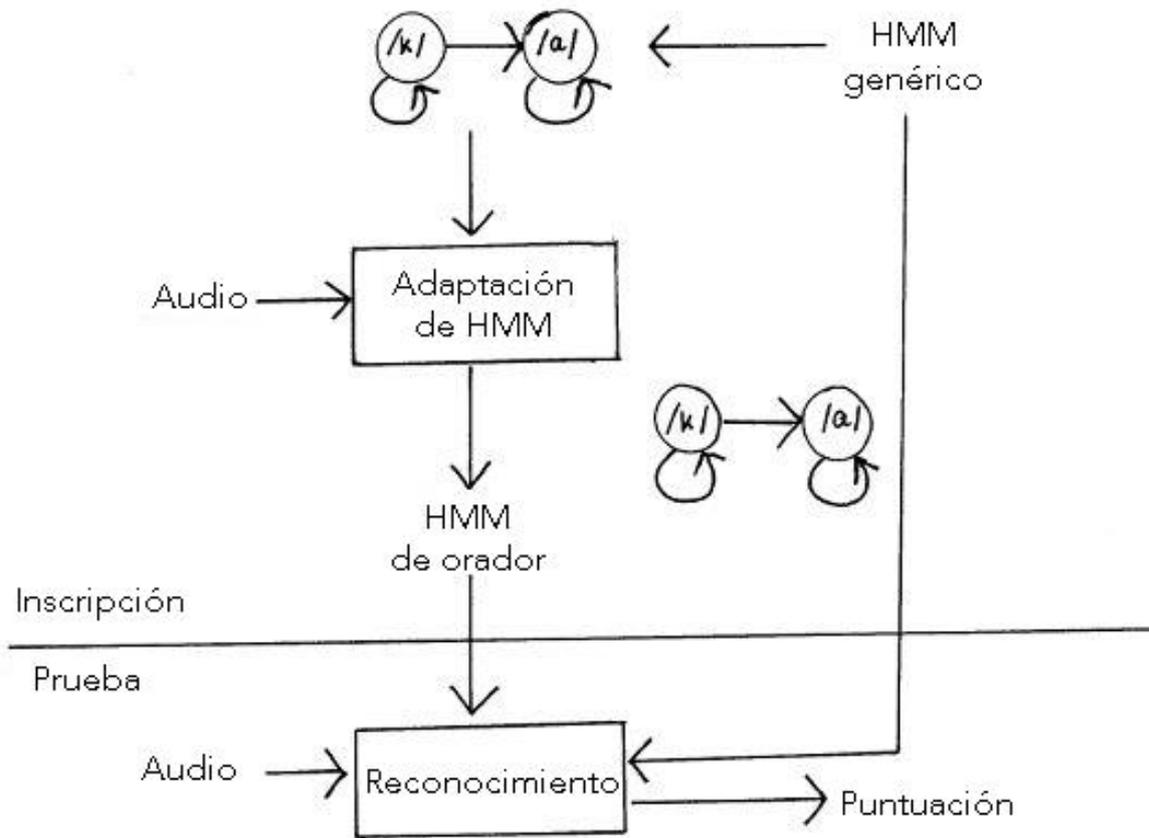


Fig. 2