



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 269 137**

(51) Int. Cl.:

H04M 9/08 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **00927602 .3**

(86) Fecha de presentación : **18.05.2000**

(87) Número de publicación de la solicitud: **1180300**

(87) Fecha de publicación de la solicitud: **20.02.2002**

(54) Título: **Cancelación del eco acústico.**

(30) Prioridad: **25.05.1999 EP 99304040**

(73) Titular/es: **BRITISH TELECOMMUNICATIONS
public limited company
81 Newgate Street
London EC1A 7AJ, GB**

(45) Fecha de publicación de la mención BOPI:
01.04.2007

(72) Inventor/es: **Reed, Martin y
Hawksford, Malcolm John**

(45) Fecha de la publicación del folleto de la patente:
01.04.2007

(74) Agente: **Curell Suñol, Marcelino**

ES 2 269 137 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

Cancelación del eco acústico.

5 La presente invención se refiere a un procedimiento para la cancelación del eco acústico de canal múltiple (AEC) aplicable a sistemas que derivan una señal multicanal espacializada a partir de una señal monofónica, cada canal de la cual se aplica a un elemento respectivo de una serie de altavoces a diferentes ganancias para proporcionar la percepción o ilusión audible de direccionalidad. Esta clase de señal espacializada se denominará en este documento mono orientado. Un sistema mono orientado utiliza dos o más elementos de ganancia para representar la espacialización, 10 que se representa en un procesador panorámico para generar las salidas de altavoz correspondientes. En las formas de realización que se describirán se utiliza una señal estereofónica de dos canales, con dos altavoces - un sistema conocido como "estéreo a partir de mono orientado" (SSM), pero los principios de la invención pueden aplicarse a sistemas con más de dos canales. La invención tiene aplicación en sistemas de teleconferencia en los cuales se proporciona a la voz de cada interlocutor posicionamiento espacial en provecho del oyente.

15 En el caso de una comunicación de voz cómoda en un sistema de teleconferencia que utiliza un altavoz y un micrófono, a diferencia del casco con auriculares, se necesita un procedimiento de cancelación del eco acústico (AEC). Para sistemas monofónicos puede utilizarse la topología que muestra la figura 1, con diversos procesos adaptativos diferentes como por ejemplo el de mínimos cuadráticos medios (LMS), mínimos cuadrados recursivos (RLS) o proyección 20 rápida afín (FAP). No obstante, en el caso de sistema estereofónicos y de canal múltiple, las soluciones existentes están mucho menos avanzadas y todavía presentan obstáculos importantes que deben superarse. El ejemplo de la figura 2 muestra que en el caso de un sistema estereofónico hay dos trayectos de eco, h_1 y h_2 (que comprenden las respuestas de impulso del micrófono y el altavoz), en comparación con el trayecto único en el caso del sistema monofónico. (Se supone que se utiliza un solo micrófono, como ocurre generalmente cuando la espacialización debe crearse artificialmente. Más generalmente, el número de trayectos de eco es el producto del número de altavoces por el número de micrófonos).

30 Las soluciones existentes para el problema de cancelación del eco acústico generalmente presentan la disposición del sistema que muestra la figura 2, en la cual las respuestas del trayecto de la persona que habla al micrófono son desconocidas. El objetivo del proceso adaptativo en el cancelador de eco es utilizar las señales $x_1(t)$, $x_2(t)$ y $e(t)$ para entrenar los filtros adaptativos \hat{h}_1 y \hat{h}_2 para que

$$e(t) \rightarrow 0 \quad (1)$$

35 40 Con los procesos existentes de filtro adaptativo no es posible alcanzar un ajuste convergente de los filtros como por ejemplo

$$h_1 = \hat{h}_1 \text{ y } h_2 = \hat{h}_2 \quad (2)$$

En lugar de ello, se obtiene una solución convergente como la siguiente

$$h_1 * g_1 + h_2 * g_2 = \hat{h}_1 * g_1 + \hat{h}_2 * g_2 \quad (3)$$

45 en la que $*$ es el operador de convolución. Debe tenerse en cuenta que la ecuación (3) satisface la ecuación (1), pero la ecuación (2) no es la única solución para la ecuación (3), por lo tanto los valores de h_1 y h_2 no pueden derivarse de este resultado.

50 Si los filtros g_1 o g_2 cambian, posiblemente debido al movimiento de la persona que habla, la igualdad de la ecuación (3) ya no se mantiene (a no ser que también se cumpla la ecuación (2)). Por lo tanto, el cancelador de eco ya no produce una solución convergente y el eco que oye la persona que habla aumenta de nivel.

55 Se han propuesto diversas soluciones para este problema, que consisten en manipular las señales del altavoz x_1 y x_2 , o en utilizar las propiedades de las señales x_1 y x_2 . El propósito de estas soluciones es utilizar las propiedades de correlación mutua de las dos señales, ya que puede demostrarse que existe una solución a la ecuación (2) cuando las dos señales se encuentran suficientemente descorrelacionadas. No obstante, al estar, en un sistema de teleconferencia las señales x_1 y x_2 , inherentemente correlacionadas, las técnicas que aprovechan las pequeñas características descorrelacionadas de las señales presentan un bajo rendimiento, excepto en condiciones ideales.

60 Se ha propuesto añadir una pequeña cantidad de ruido blanco independiente a las señales x_1 y x_2 . Introduciendo alguna señal descorrelacionada, se ha comprobado que esto contribuye significativamente a la convergencia de la solución hacia la presentada en la Ecuación (2). No obstante, aunque añadir ruido de este modo mejorar la convergencia, el ruido debe añadirse a un nivel tal que resulta indeseablemente audible.

65 Un artículo titulado "A Multichannel Affine projection Algorithm With Applications To Multichannel Acoustic Echo Cancellation" de Benesty *et al* en "IEEE Signal Processing Letters", vol. 3, nº 2, páginas 36-37, publicado en

1996 por la IEEE Signal Processing Society describe un sistema de supresión de eco multicanal que alimenta los respectivos canales con filtros adaptativos correspondientes, pero no tiene en cuenta los factores de ganancia de los respectivos canales.

5 Según la presente invención, se dispone un procedimiento de cancelación de eco acústico para una señal espacializada orientada de canal múltiple generada a partir de una entrada de señal modificada según las funciones de ganancia de espacialización respectivas para generar una pluralidad de canales de audio, utilizando el proceso de cancelación de eco una espacialización y una estimación de los trayectos de eco combinadas, calculándose la estimación a partir de las funciones de ganancia aplicadas a los canales respectivos, con lo cual, cuando cambian las funciones de ganancia aplicadas a los canales respectivos, se genera una estimación de los trayectos de eco, basándose la estimación en una estimación previa del trayecto de eco y en las funciones de ganancia, utilizándose la estimación de los trayectos de eco para generar una señal de cancelación de eco.

10 Según otro aspecto de la presente invención, se dispone un aparato para cancelación del eco acústico en un sistema de audio espacializado orientado de canal múltiple, en el que el sistema de audio espacializado comprende:

15 unos medios de entrada de señal para recibir una señal de audio;

20 una pluralidad de medios de salida de audio para generar señales acústicas derivadas de la señal de audio;

25 unos medios de control asociados con los medios de salida de audio para generar entradas de control de ganancia que controlan los medios de salida de audio de manera que dicha pluralidad de medios de salida de audio generan una versión espacializada de la señal de audio;

30 unos medios de entrada de audio para detectar señales acústicas;

35 unos medios de salida de señales para transmitir una señal derivada de las señales acústicas detectadas por los medios de entrada de audio;

40 unos medios de estimación de trayectos de eco que comprenden unos medios de detección para identificar cambios en las funciones de ganancia de los medios de control respectivos, y medios de estimación para generar una estimación del trayecto de eco entre los medios de salida acústica y los medios de entrada acústica, basándose la estimación en una estimación previa del trayecto de eco y en las funciones de ganancia detectadas por los medios de detección;

45 unos medios de generación de señales de cancelación de eco para generar una señal de cancelación de eco derivada de las señales de audio espacializadas generadas por los medios de control y la estimación derivada por los medios de estimación del trayecto de eco, y

50 unos medios de combinación de señales para aplicar la señal de cancelación de eco a la señal generada por los medios de entrada de audio.

La presente invención es una adaptación del proceso LMS monofónico. La invención evita actualizaciones múltiples de dos o más estimaciones de trayectos de eco, tales como \hat{h}_1 y \hat{h}_2 , y reduce el número de operaciones de filtración requeridas en comparación con los procesos de cancelación de eco estéreo existentes, tales como $\hat{h}_1 * x_1$ y $\hat{h}_2 * x_2$. Adicionalmente, la invención utiliza los parámetros de espacialización en el proceso adaptativo, a diferencia de los procesos de cancelación de eco existentes. La actualización del LMS se modifica para tomar en consideración los parámetros de espacialización que se utilizan para actualizar la estimación del trayecto de eco agregado cada vez que cambia la espacialización. Después de cambiar la espacialización N, (donde N es el número de canales del sistema), la estimación del trayecto de eco agregado converge con el trayecto de eco agregado para futuros cambios en las espacializaciones. Antes del cambio N de la espacialización, el proceso converge hacia una solución local para la estimación del trayecto de eco agregado, de modo que sigue existiendo alguna reducción de la señal de eco en el estadio de aprendizaje del proceso.

55 El estadio de aprendizaje puede ser parte de una fase de preparación del sistema antes de la utilización del mismo en funcionamiento operativo. Por ejemplo, el número requerido de cambios de espacialización puede alcanzarse operando el proceso LMS monofónico para cada canal sucesivamente, ajustando las ganancias de los demás canales a cero.

60 A continuación se describe, a título de ejemplo, una forma de realización de la invención considerada conjuntamente con las figuras en las cuales

65 la figura 1 ilustra los principios de la cancelación de eco monofónico,

la figura 2 ilustra los principios básicos de la cancelación de eco estereofónico, (sólo se muestra la mitad del trayecto de retorno)

65 la figura 3 ilustra una forma de realización de la invención,

ES 2 269 137 T3

la figura 4 ilustra un sistema teórico equivalente al de la figura 3,

la figura 5 es una ilustración gráfica del error medio del sistema que muestra la figura 3 comparando sistemas anteriores con el de la invención,

la figura 6 ilustra el descenso a lo largo del tiempo del desajuste de las estimaciones de trayecto erróneas \hat{h}_1 y \hat{h}_2 utilizando el procedimiento según la invención.

la figura 7 es un diagrama de flujo que muestra el proceso de actualización de la forma de realización preferida.

El sistema monofónico ilustrado en la figura 1 comprende un trayecto de entrada 12 conectado a la fuente (micrófono 11 y altavoz 10) de la señal original $s(t)$. Esta señal $s(t)$ se modifica mediante el trayecto de transmisión 12 para generar una señal de altavoz $x(t)$ que se suministra al altavoz 13. El trayecto de retorno consiste en un micrófono 21, un trayecto de transmisión de retorno 22 y un altavoz 23. Entre el altavoz 13 de un trayecto y el micrófono 21 del otro trayecto existe un trayecto acústico h_1 . Por lo tanto, algunos sonidos originados por la persona que habla 10 serán devueltos al altavoz 23 y oídos por la persona que habla 10 como eco. Este efecto puede resultar intrusivo, especialmente cuando los trayectos de transmisión 12, 22 pueden introducir retardos, de modo que se instala un procesador de cancelación de eco 24, 25, 26 para eliminar esta señal. Un procesador adaptativo 24 compara la señal $e(t)$ que debe transmitirse por la vía de retorno 22 con la señal $x(t)$ que llega por el trayecto de entrada 12 e identifica las correlaciones entre ellas. Esto se utiliza para controlar un banco de filtros adaptativos 25 para generar un vector \hat{h}_1 . Este vector es una estimación del trayecto acústico h_1 . Este vector se aplica a la señal de entrada $x(t)$ mediante el filtro y el resultado se sustrae de la entrada $y(t)$ del micrófono 21 en un combinador 26 para generar una señal de retorno $e(t)$. Si el filtro adaptativo 25 genera una estimación \hat{h}_1 exacta de la función h_1 , la señal de eco $y(t)$ debería ser anulada por la salida del filtro adaptativo 25 y $e(t)$ debería ser cero. Por lo tanto, las únicas señales transmitidas por el trayecto de retorno 22 serán los sonidos generados cerca del micrófono 21 y no recibidos del altavoz 13.

La figura 2 ilustra un sistema estereofónico generalizado. La situación es similar a la de la figura 1, excepto porque en este caso se han dispuesto dos altavoces y dos micrófonos (indicados por los sufijos "L" y "R" en la figura 2) en cada extremo. Sólo se muestra un canal 22L del trayecto de retorno. Para cada trayecto de retorno 22L, (22R) existen dos estimaciones de trayecto de eco \hat{h}_{1L} , \hat{h}_{2L} , una para cada trayecto de ida 12_L, 12_R, porque cada micrófono 21L, (21R) es capaz de recibir retroalimentación acústica de ambos altavoces 13L, 13R. Debido a que los dos trayectos de ida 12_L, 12_R transportan señales de la misma fuente 10, existe una correlación significativa entre las dos estimaciones \hat{h}_{1L} , \hat{h}_{2L} y, por lo tanto, los procesos que generan los valores de los filtros adaptativos 25LL, 25RL no son independientes.

En el caso general que muestra la figura 2, las funciones originales g_1 , g_2 son desconocidas en el extremo receptor, siendo dependientes de las posiciones relativas de los micrófonos 11L, 11R y la persona que habla 10. Por lo tanto, no es posible calcular una solución única para \hat{h}_{1L} , \hat{h}_{2L} . Para una espacialización determinada, el procesador adaptativo 24L generará dos funciones adecuadas \hat{h}_{1L} , \hat{h}_{2L} para las cuales resulta válida la ecuación (3) anteriormente citada:

$$h_1 * g_1 + h_2 * g_2 = \hat{h}_1 * g_1 + \hat{h}_2 * g_2 \quad (3)$$

pero esto no implica necesariamente que también pueda aplicarse la ecuación (2):

$$h_1 = \hat{h}_1 \text{ y } h_2 = \hat{h}_2 \quad (2)$$

Si la espacialización, y por lo tanto las funciones g_1 , g_2 cambian, se apreciará que los filtros adaptativos deben reiniciarse para corresponder a la nueva espacialización.

En un sistema orientado como el que ilustran las figuras 3 y 4, se utiliza una fuente monofónica y las funciones g_1 , g_2 se convierten en valores escalares g_1 , g_2 que representan ganancia. Por consiguiente, sólo hay un trayecto de transmisión 12, 22 en cada dirección. La información de control, generada tanto en el extremo receptor como en el transmisor con la señal monofónica $s(t)$, se utiliza para controlar los amplificadores de ganancia variable 14L, 14R para variar las funciones de ganancia g_1 , g_2 definiendo la espacialización en el extremo receptor, para emular una fuente estereofónica. El hecho de que las funciones de ganancia g_1 , g_2 sean conocidas en el extremo receptor permite utilizarlas en el procesador adaptativo 24, 25.

A continuación se describirá el funcionamiento del procesador adaptativo 24, 25 con referencia a la figura 7. La teoría matemática subyacente se describirá a continuación con referencia a las figuras 3 y 4.

Después de la etapa de inicialización 101, en la cual se ajustan valores nacionales para las funciones de ganancia g_1 , g_2 , el proceso se ejecuta en un bucle iterativo para cada periodo muestreado n del modo siguiente.

En primer lugar, se ajustan los valores k_1 , k_2 , k_3 (etapa 102). Estos valores identifican los tres últimos períodos muestreados en los cuales han cambiado los valores de espacialización g_1 , g_2 . Si los valores de ganancia de espacialización g_1 , g_2 no han cambiado desde la muestra previa n-1 los valores de k_1 , k_2 , k_3 son los mismos que para la muestra

ES 2 269 137 T3

anterior. No obstante, si los valores han cambiado, k_3 se ajusta al valor previo de k_2 , k_2 se ajusta al valor previo de k_1 , y k_1 se ajusta a $n-1$.

5 A continuación se determina la función de la ganancia estimada (etapa 103). Esta función es la matriz

$$\begin{bmatrix} g_1(k_3) & g_2(k_3) \\ g_1(k_2) & g_2(k_2) \end{bmatrix} = \mathbf{G}(k_2, k_3)$$
. Si los valores de espacialización g_1, g_2 no han cambiado, esta matriz tampoco cambia y no es necesario volver a calcularla. A continuación se determina la inversa de esta matriz.

10

$$\begin{bmatrix} y_{00} & y_{01} \\ y_{10} & y_{11} \end{bmatrix} = \mathbf{G}^{-1}(k_2, k_3)$$

15 De nuevo, si los valores de espacialización g_1, g_2 no han cambiado, esta matriz tampoco cambia y por lo tanto no es necesario volver a calcularla.

20 A continuación, (etapa 104), si $r = n - k_1 + 1$ es inferior al número de términos L en el vector de trayecto de eco
 \hat{h} estimado (en otras palabras, si el número de muestras r transcurrido desde el último cambio de espacialización es inferior a L), se corrige un término del vector de trayecto de eco estimado \hat{h}_{n-1} como sigue

25

$$\begin{aligned} h_{n-1}(r) &= \left(y_{00} \hat{h}_{k3}(r) + y_{01} \hat{h}_{k2}(r) \right) g_1 \\ &\quad + \left(y_{10} \hat{h}_{k3}(r) + y_{11} \hat{h}_{k2}(r) \right) g_2 \end{aligned}$$

30 (siendo \hat{h}_{n-1} el caso específico de la función de trayecto de eco estimado \hat{h} de la iteración previa). Todos los demás términos $\hat{h}_{n-1}(0) \dots \hat{h}_{n-1}(r-1)$ y $\hat{h}_{n-1}(r+1) \dots \hat{h}_{n-1}(L-1)$ permanecen sin cambios.

35 A continuación se genera la señal de cancelación de error $s_n^T \hat{h}_{n-1}$ (en la que s_n es el vector que representa las últimas muestras L de la señal de entrada $s(n)$), utilizando el vector de trayecto de eco estimado \hat{h}_{n-1} (etapa 105), y se sustrae de la señal $y(n)$ para generar la señal de salida $e(n)$.

40 El vector de trayecto de eco estimado \hat{h}_{n-1} se adapta, a continuación, como respuesta a la señal de eco $e(n)$ (etapa 106) para utilizarlo en la siguiente iteración.

$$\hat{h}_n = \hat{h}_{n-1} + \mu s_n e(n)$$

45 en la que $e(n) = s_n^T s_n e(n)$ y μ es el tamaño de la etapa.

El progreso es estable a no ser que la espacialización cambie en una escala de tiempo superior al período L , y que $0 \leq \mu < 2$.

50 La complejidad de cálculo de las etapas 105 y 106 del proceso anterior es la misma que la del proceso LMS normalizado, que es del orden de $2L$. El número de cálculos es del orden de dos multiplicaciones y una división para la inversión de matriz utilizada en la etapa 103. Como que se realiza solamente una vez después de cada cambio de las espacialización, aumenta poco la complejidad del proceso para L grandes. La etapa 104 sólo se calcula en las L primeras muestras después de un cambio de la espacialización y es insignificante para las L grandes. Por lo tanto, 55 cuando se utiliza el proceso mostrado anteriormente para la cancelación del eco acústico con un sistema orientado mono, para el cual es probable que $L > 100$, el proceso presenta una complejidad de aproximadamente $2L$.

Sigue una descripción matemática, que se iniciará a partir del sistema mostrado en la figura 4 y a continuación procederá a mostrar que es equivalente al sistema según la invención mostrado en la figura 3 que utiliza una única estimación de trayecto de eco agregado.

60 Como muestra la figura 4, los dos trayectos de respuesta de impulso de micrófono 12L, 12R (figura 2) se sustituyen por dos respuestas de impulso generadas artificialmente g_1 y g_2 que actúan sobre una única señal de micrófono $s(t)$ recibida a través del trayecto monofónico 12 para generar la percepción o ilusión de espacialización. En el caso de la solución presentada aquí, las funciones g_1 y g_2 se han simplificado todavía más representándolas como simples funciones de ganancia antes que como respuestas de impulso. Este sistema trabaja de un modo completamente diferente al del sistema anterior de la figura 2 ya que, antes que intentar descorrelacionar las señales x_1 y x_2 para ayudar a la convergencia, el procedimiento presentado aquí depende de la naturaleza inherentemente altamente correlacionada de

las señales x_1 y x_2 . El procedimiento utiliza el conocimiento de las funciones g_1 y g_2 (que son desconocidas en los sistemas de la técnica anterior) para conseguir una cancelación de eco convergente. Antes de describir una solución de filtro adaptativo al problema del eco, se demostrará que existe una solución al problema.

5 Haciendo que la entrada al bloque de espacialización en el tiempo de muestra n esté representada por un vector de columna $s_n = [s(n) \ s(n-1) \ \dots \ s(n-(L-1))]^T$, la entrada al micrófono del extremo del oyente esté representada por $y_n = [y(n) \ y(n-1) \ \dots \ y(n-(L-1))]^T$ y que los dos trayectos de eco del altavoz al micrófono sean los vectores L de columna h_1 y h_2 (que incorporan las respuestas de impulso del altavoz y el micrófono) de longitud L , entonces

10
$$y_n = g_1(n)S_nh_1 + g_2(n)S_nh_2 \quad (4)$$

en la que la espacialización está representada como valores de ganancia $g_1(n)$ y $g_2(n)$ que son constantes a lo largo de los períodos de muestra $n-(L-1) \ \dots \ n$, y $S_n = [s_n \ \dots \ s_{n-(L-1)}]^T$ (es una matriz de “Toeplitz”, es decir, una matriz simétrica de orden $L \times L$, que presenta los términos de s_n en la primera fila y la primera columna, los términos de s_{n-1} en la segunda fila y columna, y así sucesivamente). Puede apreciarse que h_1 y h_2 no pueden resolverse a partir de la ecuación (4).

20 No obstante, considerando ahora la utilización de un segundo conjunto de observaciones de entradas y salidas en un tiempo de muestra $n+a$, en la que $a \geq L$

25
$$\begin{aligned} g_1(n+a) &= g_1(n+a-1) = \dots = g_1(n+1) \neq g_1(n) \\ g_2(n+a) &= g_2(n+a-1) = \dots = g_2(n+1) \neq g_2(n) \end{aligned} \quad (5)$$

en otras palabras, las funciones g_1 y g_2 han cambiado entre el tiempo de muestra n y el tiempo de muestra $n+1$, pero a continuación se han mantenido sin cambios entre el tiempo $n+1$ y el tiempo de muestra $n+a$.

30 Por lo tanto,

35
$$\begin{bmatrix} y_n \\ y_{n+a} \end{bmatrix} = (\mathbf{G}_{n,n+a} \otimes \mathbf{I}_L) \begin{bmatrix} S_n \\ S_{n+a} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (6)$$

en la que \mathbf{I}_L es la matriz de identidad $L \times L$,

40
$$\mathbf{G}_{n,n+a} = \begin{bmatrix} g_1(n) & g_2(n) \\ g_1(n+a) & g_2(n+a) \end{bmatrix} \quad (7)$$

45 y \otimes es el producto de Kronecker. (El producto de Kronecker de dos matrices A y B se obtienen multiplicando la matriz B independientemente por cada término individual de la matriz A y formando una nueva matriz (cuyo orden es el producto de las dos matrices originales) con los términos resultantes).

La solución a la ecuación (6) es

50
$$\begin{bmatrix} S_n \\ S_{n+a} \end{bmatrix}^{-1} (\mathbf{G}_{n,n+a} \otimes \mathbf{I}_L)^{-1} \begin{bmatrix} y_n \\ y_{n+a} \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (8)$$

55 y utilizando las identidades del producto de Kronecker

60
$$\begin{bmatrix} S_n \\ S_{n+a} \end{bmatrix}^{-1} (\mathbf{G}_{n,n+a}^{-1} \otimes \mathbf{I}_L) \begin{bmatrix} y_n \\ y_{n+a} \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad (9)$$

Por lo tanto existe una solución para h_1 y h_2 si la señal s excita persistentemente (es decir, presenta un contenido totalmente espectral) y la matriz $\mathbf{G}_{n,n+a}$ es no singular, es decir, tiene una matriz inversa. La condición de no singular de la matriz $\mathbf{G}_{n,n+a}$ se cumple si los valores de espacialización en los tiempos de muestra n y $n+a$ son diferentes y no son múltiplos escalares uno de otro (es decir, $g_1(n)/g_2(n) \neq g_1(n+a)/g_2(n+a)$). Idealmente, los valores deberían ser suficientemente diferentes para que la solución de la ecuación (9) presente las condiciones adecuadas.

ES 2 269 137 T3

Una vez establecido que existe una solución, a continuación se deriva del proceso LMS el proceso adaptativo para la solución. El proceso LMS normalizado (NLMS) se utiliza para efectuar la cancelación del eco monofónico, como se ha mencionado con referencia a la figura 1, utilizando las actualizaciones siguientes

$$5 \quad e(n) = y(n) - s_n^T \hat{h}_{n-1} \quad (10)$$

$$10 \quad \varepsilon(n) = s_n^T S_n e(n) \quad (11)$$

$$15 \quad \hat{h}_n = \hat{h}_{n-1} + \mu s_n \varepsilon(n) \quad (12)$$

en las que $e(n)$ es la señal de eco, μ es el parámetro de tamaño de la etapa y \hat{h}_n es la estimación del trayecto de eco en el caso de muestra n . Las ecuaciones anteriores del LMS normalizado pueden modificarse para el caso mono orientado utilizando una estimación de trayecto de eco agregado único y redefiniendo \hat{h}_{n-1} como

$$20 \quad \hat{h}_{n-1}^T = [\hat{h}_1(0)g_1(n) + h_2(0)g_2(n), \dots, \hat{h}_1(L-1)g_1(n-(L-1)) + h_2(L-1)g_2(n-(L-1))] \quad (13)$$

en la que $\hat{h}_1(t)$ y $\hat{h}_2(t)$ son funciones que representan las dos estimaciones de eco en el intervalo de muestra n .
25 Similarmente se define h como la combinación de $h_1(t)$ y $h_2(t)$ en una forma equivalente a la que muestra la ecuación 13

$$30 \quad h^T = [h_1(0)g_1(n) + h_2(0)g_2(n), \dots, h_1(L-1)g_1(n-(L-1)) + h_2(L-1)g_2(n-(L-1))] \quad (14)$$

35 La tarea consiste entonces en utilizar y actualizar \hat{h} de modo que las actualizaciones del LMS normalizado de las ecuaciones 10, 11 y 12 se utilizan para la cancelación de eco en vez de utilizar explícitamente dos estimaciones de trayecto de eco. Si los valores de $g_1(n)$ y $g_2(n)$ son constantes para todos los n , las actualizaciones de la ecuación 10, 11 y 12 pueden utilizarse sin cambios para determinar una estimación de h , cuando h es constante a lo largo del tiempo. No obstante, si $g_1(n)$ y $g_2(n)$ cambian a lo largo del tiempo, esta solución no puede utilizarse, ya que un cambio en h no se tiene en cuenta en las actualizaciones del LMS de las ecuaciones 10, 11 y 12.

40 Considerando tres épocas de muestra i , $i-a$ e $i-b$ donde

$$b >> L > a \quad (15)$$

45

y

50

$$g_1(i) = g_1(i-1) = \dots = g_1(i-a) \neq g_1(i-a-1) = g_1(i-a-2) = \dots = g_1(i-b) \\ g_1(i-b) \neq g_1(i-b-1) = g_1(i-b-2) = \dots = g_1(i-\infty)$$

55

$$(16)$$

y similarmente para $g_2(n)$, es decir, los valores de $g_1(n)$ y $g_2(n)$ cambian sólo en las épocas $i-a$ e $i-b$.

60 Considerando que el valor del coeficiente j en el trayecto de eco combinado en las épocas $i-a-1$ e $i-b-1$ (es decir, justo antes de los cambios en la espacialización) que forman las ecuaciones 14 y 16 viene dado por

$$65 \quad h_{i-b-1}(j) = h_1(j)g_1(i-b-1) + h_2(j)g_2(i-b-1) \quad (17)$$

$$h_{i-a-1}(j) = h_1(j)g_1(i-a-1) + h_2(j)g_2(i-a-1) \quad (18)$$

ES 2 269 137 T3

Las ecuaciones 17 y 18 pueden expresarse como

$$5 \quad \begin{bmatrix} h_{i-b-1}(j) \\ h_{i-a-1}(j) \end{bmatrix} = \begin{bmatrix} g_1(i-b-1) & g_2(i-b-1) \\ g_1(i-a-1) & g_2(i-a-1) \end{bmatrix} \begin{bmatrix} h_1(j) \\ h_2(j) \end{bmatrix} \quad (19)$$

y por lo tanto

$$10 \quad \mathbf{G}_{i-b-1, i-a-1}^{-1} \begin{bmatrix} h_{i-b-1}(j) \\ h_{i-a-1}(j) \end{bmatrix} = \begin{bmatrix} h_1(j) \\ h_2(j) \end{bmatrix} \quad (20)$$

15 utilizando la definición de \mathbf{G} de (7).

Además, considerando el valor del coeficiente j en el trayecto de eco combinado en la época i que, a partir de (14) y (16), viene dada por

$$20 \quad 25 \quad h_i(j) = \begin{cases} h_{i-1}(j) & j < i-a \\ h_1(j)g_1(i) + h_2(j)g_2(i) & j = i-a \\ h_{i-1}(j) & j > i-a \end{cases} \quad (21)$$

30 Si los elementos de \mathbf{G}^{-1} se definen mediante una variable γ tal como

$$35 \quad \mathbf{G}_i^{-1} = \begin{bmatrix} \gamma_{00} & \gamma_{01} \\ \gamma_{10} & \gamma_{11} \end{bmatrix} \quad (22)$$

40 a partir de (19), (21) y (22)

$$45 \quad h_i(j) = (\gamma_{00} h_{i-b-1}(j) + \gamma_{01} h_{i-a-1}(j))g_1(i) + \\ (\gamma_{10} h_{i-b-1}(j) + \gamma_{11} h_{i-a-1}(j))g_2(i)$$

(Ecuación 23)

50 Esta ecuación es la actualización adicional requerida para la actualización del LMS normalizado de las ecuaciones (10), (11) y (12). Debe tenerse en cuenta que a partir de (21) sólo necesita actualizarse un coeficiente en \mathbf{h} de cada período de muestra para tomar en consideración un cambio en la espacialización.

55 El proceso puede extenderse a un sistema que presenta más de dos canales efectuando una pequeña modificación para el proceso. Específicamente, para un sistema de N canales, los cambios previos $N+1$ en la posición de espacialización se registran en variables k_{N+1}, \dots, k_1 desde el menor al más reciente respectivamente. La matriz \mathbf{G} se generaliza como

$$60 \quad 65 \quad \mathbf{G} = \begin{bmatrix} g_1(k_{N+1}) & \dots & g_N(k_{N+1}) \\ \vdots & \ddots & \vdots \\ g_1(k_2) & \dots & g_N(k_2) \end{bmatrix} \quad (24)$$

La etapa 104 puede generalizarse como

$$5 \quad \hat{h}(r) = [g_1(n) \quad \dots \quad g_N(n)] \mathbf{G}^{-1} \begin{bmatrix} \hat{h}(r) \\ \vdots \\ \hat{h}(r) \end{bmatrix} \quad (25)$$

10 Las etapas 105 y 106 permanecen sin cambios.

Para demostrar el proceso descrito, tanto el proceso de mínimos cuadráticos medios normalizados estéreo según la presente invención como el de mínimos cuadráticos medios normalizados se simularon utilizando la configuración mostrada en la figura 3 y con trayectos de eco tomados de los trayectos micrófono/persona que habla reales medidos. 15 La posición de espacialización se cambió cinco veces durante la simulación para comprobar la capacidad del proceso para adaptarse a la espacialización cambiante. La figura 5 muestra el error medio de conjunto en la cancelación del eco para una entrada de ruido blanco gausiano (GWN). Para la simulación, el parámetro de tamaño de etapa μ era 0,1, la longitud del filtro era de $L = 200$, la velocidad de muestra era de 8 KHz y se añadió una señal de ruido de 40 dB debajo de la señal de entrada. La figura 5 muestra que inicialmente el proceso LMS normalizado (figura 5a) y el proceso LMS normalizado estéreo según la invención (figura 5b) presentan el mismo rendimiento hasta 40.000 muestras. Debe 20 tenerse en cuenta que el error en ambos procesos se produce inicialmente cuando los filtros adaptativos convergen con el valor estático del vector de trayecto agregado h . No obstante, en 19.000 períodos de muestra la espacialización cambia, de modo que h cambia y por consiguiente el error aumenta de repente en ambos procesos. El proceso LMS normalizado (figura 5a) todavía no dispone de suficiente información para formar el vector de trayecto agregado 25 correcto h , ya que necesita dos posiciones de espacialización para determinar los vectores de trayecto individuales h_1 y h_2 . No obstante, después de la segunda posición de espacialización, el procedimiento según la invención puede efectuar una estimación del valor del h cambiado. Esto puede observarse en la figura 5 cuando el error no aumenta después de la muestra 40.000 para la curva de error en b), a diferencia de lo que ocurre en el proceso LMS normalizado en a) que demuestra incrementos del error cada vez que cambia la espacialización.

30 También puede observarse el rendimiento del procedimiento según la invención para señales de voz en la figura 6, que muestra el desajuste en los dos vectores de trayecto de eco estimados \hat{h}_1 y \hat{h}_2 . Estos vectores no se calculan realmente de forma explícita en este proceso: no obstante, se utiliza una solución para ellos como parte de la espacialización. La ecuación 20 se utilizó para formar los gráficos que muestra la figura 6. Puede apreciarse que el desajuste desciende en uno de los filtros y aumenta en el otro a lo largo de la trayectoria de la primera posición de espacialización, que está situada entre las muestra 0 y 19.000. Esto se debe al hecho de que $g_1(0) = 0,2$ y $g_2(0) = 0,8$, lo cual 35 significa que la mayor parte de la energía de señal se deriva del segundo trayecto de filtro, de modo que la estimación del filtro agregado \hat{h} tiende al segundo trayecto de filtro h_2 . En este punto, no existe una información *a priori* que 40 permita la solución verdadera de h_1 y h_2 . No obstante, tan pronto como se inicia la segunda posición de espacialización en la muestra 19.000, el desajuste en \hat{h}_1 y \hat{h}_2 empieza a descender y el proceso LMS normalizado estéreo permite que se forme una solución para h_1 y h_2 .

45 El procedimiento descrito utiliza el filtro adaptativo de mínimos cuadráticos medios normalizado para formar la actualización de la estimación del trayecto de eco combinado. No obstante, puede utilizarse cualquier procedimiento adaptativo actual o futuro que actualice una estimación de un filtro desconocido sobre la base muestra a muestra en lugar del algoritmo de mínimos cuadráticos medios normalizados descrito. La única modificación requerida es sustituir la etapa del proceso 106 por otra actualización de filtro. Son ejemplos adecuados existentes la proyección afín rápida, y los filtros adaptativos de mínimos cuadráticos medios o de mínimos cuadráticos medios recursivos.

50

55

60

65

REIVINDICACIONES

1. Procedimiento de cancelación del eco acústico para una señal espacializada orientada de canal múltiple, estando generada la señal espacializada orientada a partir de una entrada de señal (12) modificada según las funciones de ganancia de espacialización (g_1, g_2) respectivas para generar una pluralidad de canales de audio ($x_1(t), x_2(t)$), utilizando el proceso de cancelación de eco una espacialización y una estimación de los trayectos de eco combinadas (\hat{h}_1, \hat{h}_2), desviándose la estimación a partir de las funciones de ganancia aplicadas a los canales respectivos, con lo cual, cuando cambian las funciones de ganancia aplicadas a los canales respectivos, se genera una estimación (\hat{h}_1, \hat{h}_2) de los trayectos de eco (h_1, h_2), basándose la estimación en una estimación previa del trayecto de eco y en las funciones de ganancia (g_1, g_2), utilizándose las estimaciones de los trayectos de eco (\hat{h}_1, \hat{h}_2) para generar una señal de cancelación de eco $e(t)$.

2. Procedimiento según la reivindicación 1, en el que las funciones de ganancia aplicadas a los canales respectivos son factores de ganancia escalares.

3. Procedimiento según la reivindicación 1 ó 2, en el que se utiliza un filtro adaptativo de mínimos cuadráticos medios normalizado para formar la actualización de la estimación de los trayectos de eco combinada.

4. Procedimiento según la reivindicación 1 ó 2, en el que se utiliza un filtro adaptativo de proyección afín rápida para formar la actualización de la estimación de los trayectos de eco combinada.

5. Procedimiento según la reivindicación 1 ó 2, en el que se utiliza un filtro adaptativo de mínimos cuadráticos medios recursivo para formar la actualización de la estimación de los trayectos de eco combinada.

6. Aparato para la cancelación del eco acústico en un sistema de audio espacializado orientado de canal múltiple, comprendiendo el sistema de audio espacializado:

- unos medios de entrada de señal (12) para recibir una señal de audio $s(t)$,
- una pluralidad de medios de salida de audio (13L, 13R) para generar señales acústicas ($x_1(t), x_2(t)$) derivadas de la señal de audio $s(t)$;
- unos medios de control (14L, 14R) asociados con los medios de salida de audio (13L, 13R) para generar entradas de control de ganancia g_1, g_2 que controlan los medios de salida de audio de manera que dicha pluralidad de medios de salida de audio (13L, 13R) generan una versión espacializada h_1, h_2 de la señal de audio $s(t)$;
- unos medios de entrada de audio (21) para detectar señales acústicas;
- unos medios de salida de señales (22) para transmitir una señal $y(t)$ derivada de las señales acústicas detectadas por los medios de entrada de audio (21);
- unos medios de estimación de trayectos de eco (25) que comprenden unos medios de detección para identificar cambios en las funciones de ganancia g_1, g_2 de los medios de control (14L, 14R) respectivos, y unos medios de estimación para generar una estimación del trayecto de eco entre los medios de salida acústica 13L, 13R y los medios de entrada acústica (21), basándose la estimación en una estimación previa del trayecto de eco y en las funciones de ganancia g_1, g_2 detectadas por los medios de detección;
- unos medios de generación de señales de cancelación de eco (24L, 24R) para generar una señal de cancelación de eco derivada de las señales de audio espacializadas generadas por los medios de control (14L, 14R) y las estimaciones derivadas por los medios de estimación del trayecto de eco (25), y
- unos medios de combinación de señales (26) para aplicar la señal de cancelación de eco a la señal generada por los medios de entrada de audio.

7. Aparato según la reivindicación 6, en el que las funciones de ganancia aplicadas a los canales respectivos son factores escalares de ganancia.

8. Aparato según la reivindicación 6 ó 7, en el que se utiliza un filtro adaptativo de mínimos cuadráticos medios normalizado para formar la actualización de las estimaciones de trayecto de eco.

9. Aparato según la reivindicación 6 ó 7, en el que se utiliza un filtro adaptativo de proyección afín rápida para formar la actualización de las estimaciones de trayecto de eco.

10. Aparato según la reivindicación 6 ó 7, en el que se utiliza un filtro adaptativo de mínimos cuadráticos medios recursivo para formar la actualización de las estimaciones de trayecto de eco.

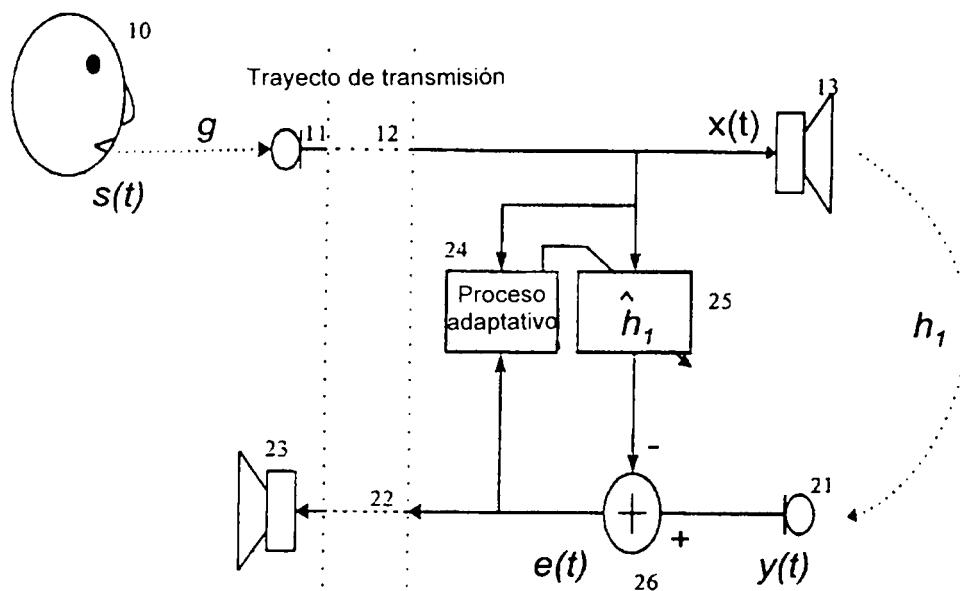


FIGURA 1

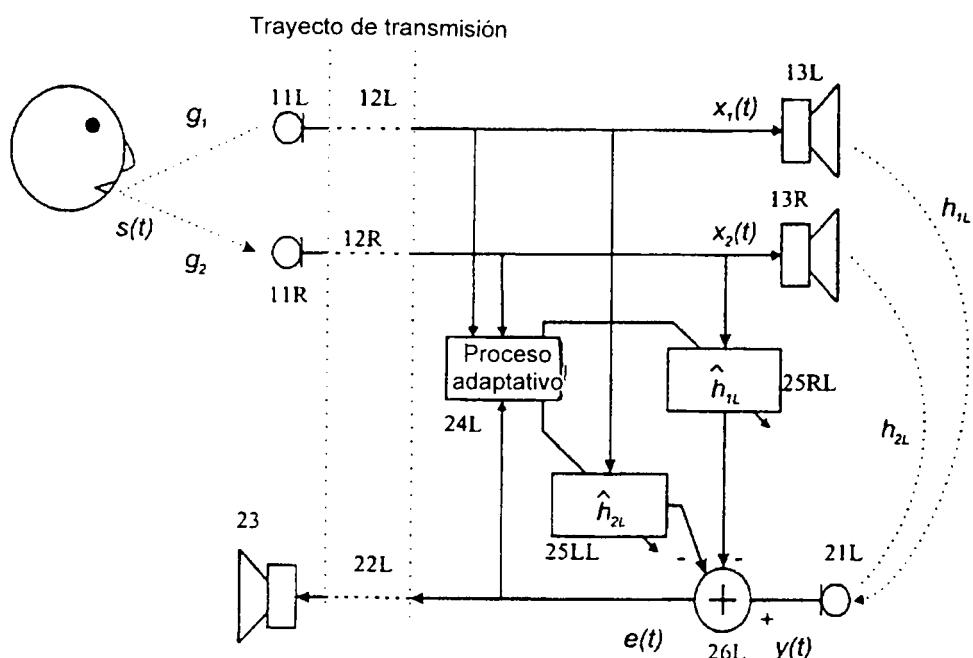


FIGURA 2

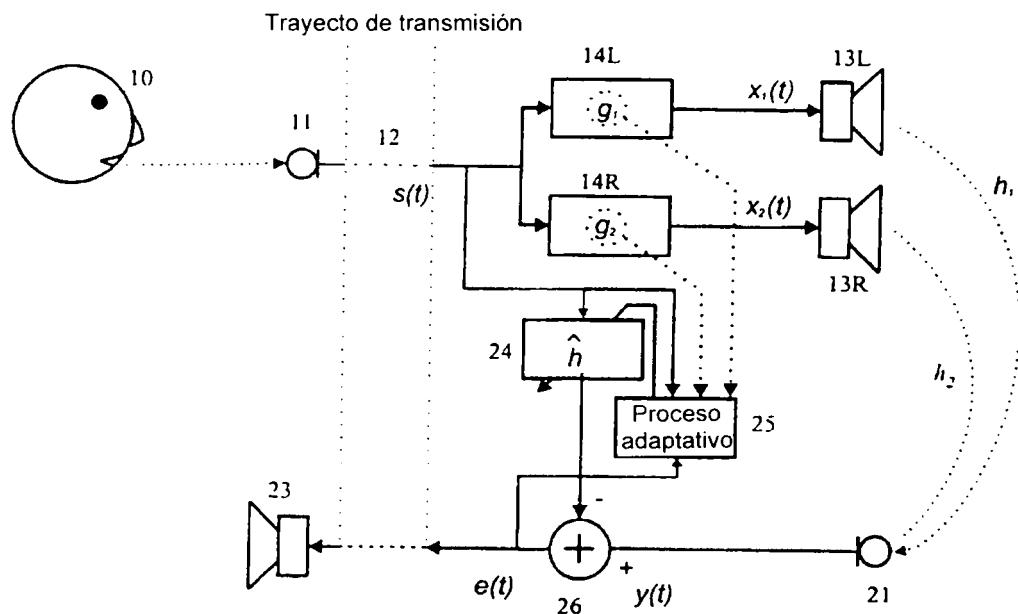


FIGURA 3

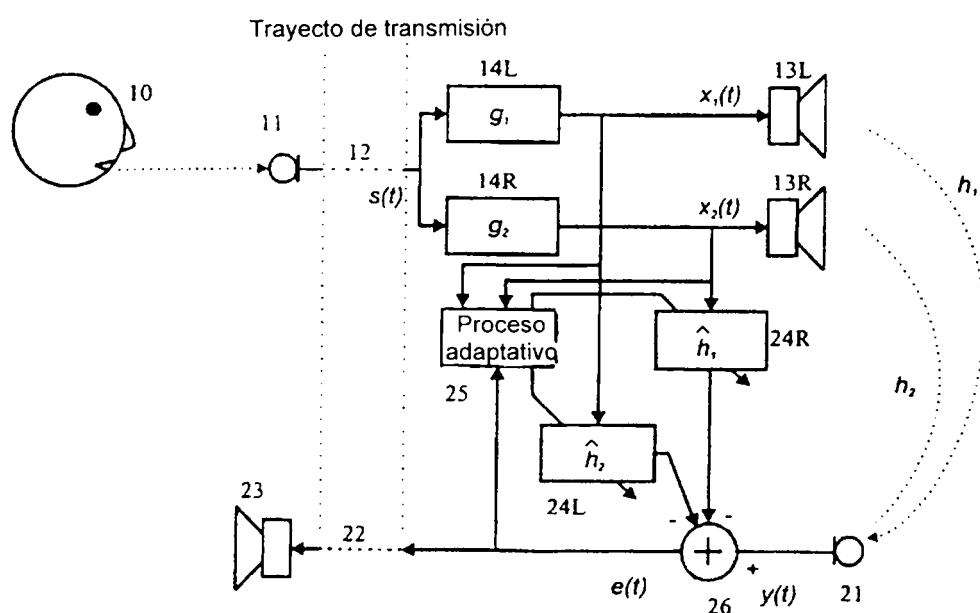


FIGURA 4

ES 2 269 137 T3

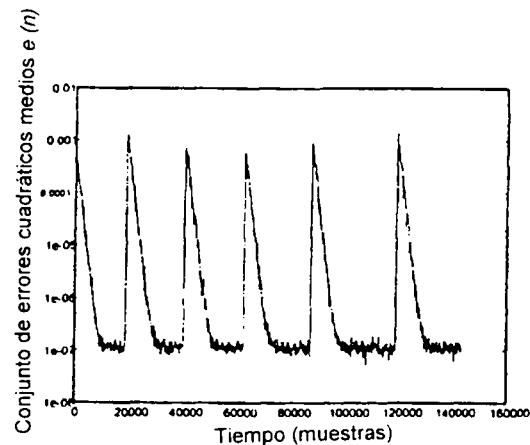


FIGURA 5a

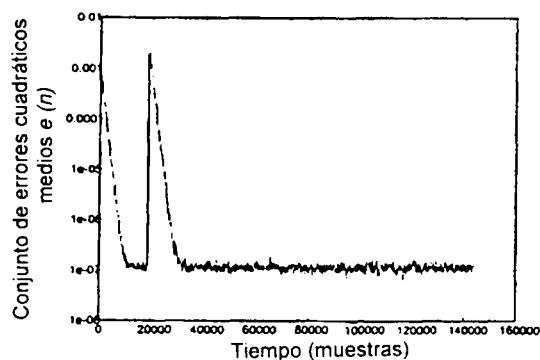


FIGURA 5b

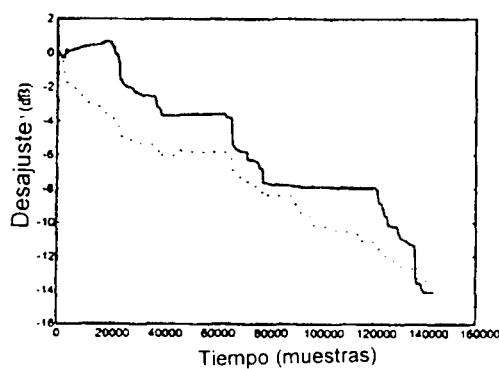


FIGURA 6

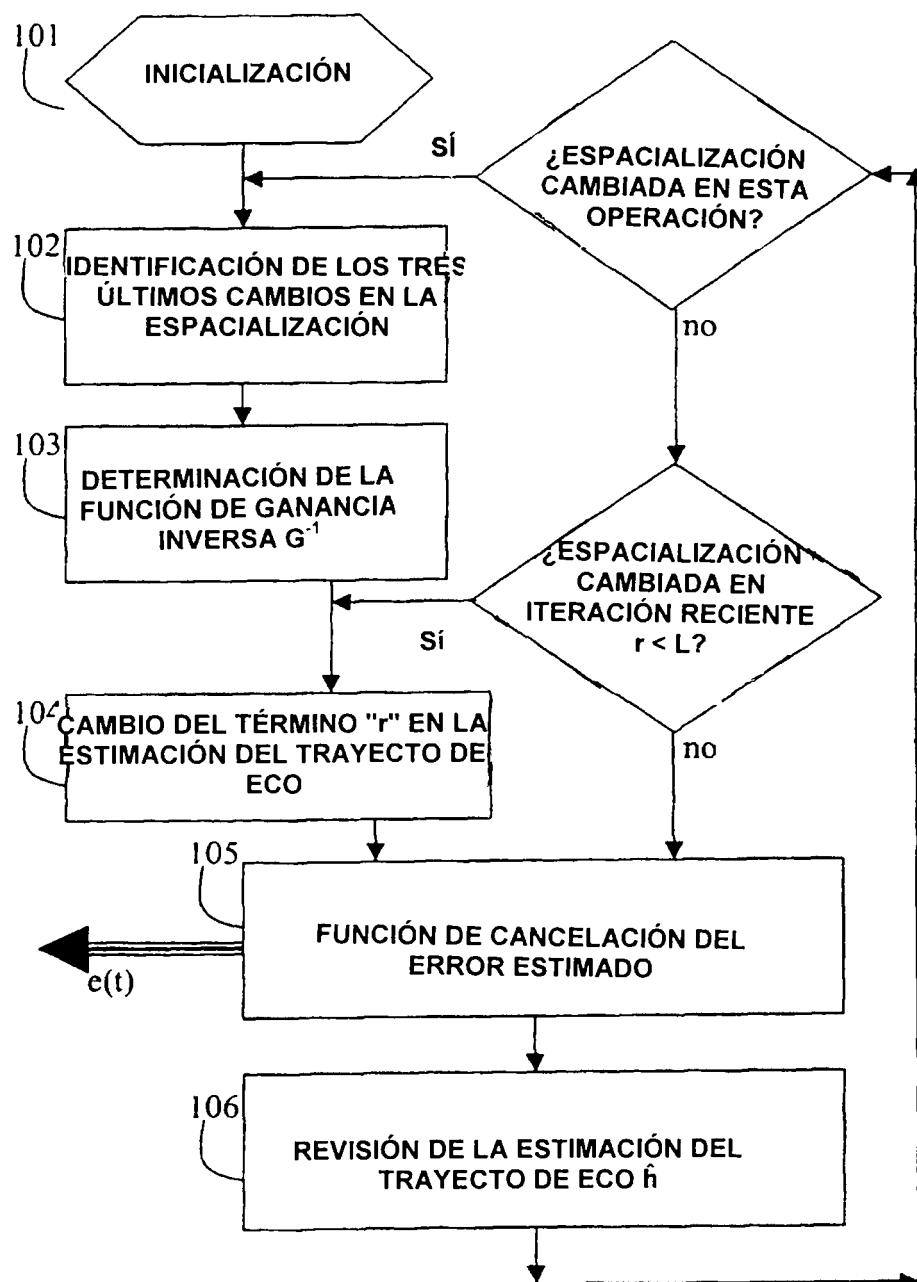


FIGURA 7