



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 281 854**

51 Int. Cl.:
G10L 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **05001938 .9**

86 Fecha de presentación : **01.08.1995**

87 Número de publicación de la solicitud: **1530201**

87 Fecha de publicación de la solicitud: **11.05.2005**

54 Título: **Procedimiento y aparato para seleccionar una velocidad de codificación en un vocodificador de velocidad variable.**

30 Prioridad: **10.08.1994 US 288413**

45 Fecha de publicación de la mención BOPI:
01.10.2007

45 Fecha de la publicación del folleto de la patente:
01.10.2007

73 Titular/es: **QUALCOMM INCORPORATED**
5775 Morehouse Drive
San Diego, California 92121-1714, US

72 Inventor/es: **Dejaco, Andrew P. y**
Garnder, William R.

74 Agente: **Carpintero López, Francisco**

ES 2 281 854 T3

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

DESCRIPCIÓN

Procedimiento y aparato para seleccionar una velocidad de codificación en un vocodificador de velocidad variable.

5 Antecedentes de la invención**I. Campo de la invención**

10 La presente invención se refiere a los vocodificadores. Más particularmente, la presente invención se refiere a un procedimiento nuevo y mejorado para determinar la velocidad de codificación de la voz en un vocodificador de velocidad variable.

II. Descripción de la técnica relacionada

15 Los sistemas de compresión de voz de velocidad variable suelen utilizar alguna forma de algoritmo de determinación de velocidad antes de que empiece la codificación. El algoritmo de determinación de velocidad asigna un esquema de codificación de velocidad de transferencia de bits más alta a los segmentos de la señal de audio en los que está presente la voz y un sistema de codificación de velocidad mas baja a los segmentos de silencio. De esta forma, se obtiene una velocidad de transferencia de bits media inferior, mientras que la calidad de la voz reconstruida sigue
20 siendo alta. Por lo tanto, para operar con eficacia un codificador de voz de velocidad variable se requiere un algoritmo de determinación de velocidad sólido que pueda diferenciar entre voz y silencio en una diversidad de entornos de ruido de fondo.

25 Uno de dichos sistemas de compresión de voz de velocidad variable o vocodificadores de velocidad variable se da a conocer en el documento WO-A1-92/22891 transferido al cesionario de la presente invención. En esta ejecución particular de vocodificador de velocidad variable, la voz de entrada se codifica utilizando técnicas de predicción lineal con excitación por código (CELP) a una de las diversas velocidades determinada por el nivel de actividad oral. El nivel de actividad oral se determina a partir de la energía de las muestras de audio de entrada, que pueden contener ruido de fondo además de voz. Para que el vocodificador proporcione codificación de voz de alta calidad con niveles variables
30 de ruido de fondo, es necesario emplear una técnica de ajuste de umbral de manera adaptativa para compensar el efecto del ruido de fondo sobre el algoritmo de decisión de velocidad.

35 Los vocodificadores se suelen utilizar en dispositivos de comunicación tales como teléfonos celulares o dispositivos de comunicación personal para proporcionar la compresión digital de señal de una señal de audio analógica que se convierte a forma digital para la transmisión. En un entorno móvil en el que puede utilizarse un teléfono celular o un dispositivo de comunicación personal, los niveles altos de energía de ruido de fondo dificultan la diferenciación mediante el algoritmo de determinación de velocidad entre los sonidos sordos de baja energía y el silencio con ruido de fondo, utilizando un algoritmo de determinación de velocidad basado en la energía de la señal. Por lo tanto, los sonidos sordos a menudo se codifican a velocidades de transferencia de bits inferiores y se degrada la calidad de la
40 voz, ya que las consonantes como "s", "x", "ch", "sh", "t", etc. se pierden en la voz reconstruida.

45 Los vocodificadores que basan sus decisiones de velocidad únicamente en la energía del ruido de fondo fracasan en tener en cuenta la intensidad de la señal en relación con el ruido de fondo al establecer los valores umbral. Un vocodificador que basa sus niveles umbral únicamente en el ruido de fondo tiende a comprimir entre sí los niveles umbral cuando el ruido de fondo aumenta. Si se desea que el nivel de la señal se mantenga fijo, ésta es la forma correcta de establecer los niveles umbral; sin embargo, si se desea que el nivel de la señal aumente con el nivel de ruido de fondo, entonces la compresión de los niveles umbral no es una solución óptima. Se necesita un procedimiento alternativo para establecer niveles umbral que tenga en cuenta la intensidad de la señal en los vocodificadores de velocidad variable.

50 Un problema final que continúa surge durante la reproducción de música mediante vocodificadores de decisión de velocidad basada en la energía de ruido de fondo. Cuando la gente habla, deben efectuar pausas para respirar, lo que permite que los niveles umbral se restablezcan en el nivel de ruido de fondo adecuado. Sin embargo, en la transmisión de música a través de un vocodificador, como surge en condiciones de música en espera, no se producen pausas y los niveles umbral continúan aumentando hasta que la música empieza a codificarse a una velocidad inferior a la velocidad
55 completa. En tal condición, el codificador de velocidad variable ha confundido música con ruido de fondo.

60 Merece atención el artículo de Paksoy *et al* "Variable Rate Speech Coding with Phonectic Segmentation", ICASSP 1993, páginas II-155-158. El artículo da a conocer un filtro de eliminación de ruido adaptativo utilizado para distinguir entre ruido y voz. Cada trama de la señal de entrada se pasa a través del filtro y la potencia en la salida del filtro se compara con un umbral adaptativo para detectar la presencia de voz. La capacidad de detección de actividad de voz en identificar voz en un entorno de SNR baja se fortalece introduciendo un esquema de umbral adaptativo diferente, en el que se realizan las comparaciones de nivel de energía en subbandas de frecuencia individuales. Un criterio de energía dependiente de la banda utiliza cuatro subbandas de frecuencia con el fin de detección de voz. Se obtiene un umbral adaptativo para cada una de estas cuatro bandas basado en la energía de bandas correspondientes de ruido estacionario.
65 La energía de la señal de entrada para cada una de estas cuatro bandas se calcula y si cualquiera de éstas sobrepasa el umbral adaptativo correspondiente, entonces se indica sonido.

Sumario de la invención

Según la presente invención se proporcionan un aparato para determinar una velocidad de codificación, tal como se establece en la reivindicación 1, y un procedimiento para determinar una velocidad de codificación, tal como se establece en la reivindicación 17. Las realizaciones de la presente invención se reivindican en las reivindicaciones dependientes.

La presente invención es un procedimiento y aparato nuevos y mejorados para determinar una velocidad de codificación en un vocodificador de velocidad variable. Un primer objetivo de la presente invención es proporcionar un procedimiento mediante el cual reducir la probabilidad de codificar sonidos sordos de baja energía como ruido con fondo. En la presente invención, la señal de entrada se filtra para dar una componente de alta frecuencia y una componente de baja frecuencia. Las componentes filtradas de la señal de entrada se analizan a continuación por separado para detectar la presencia de voz. Debido a que los sonidos sordos tienen una componente de alta frecuencia su intensidad relativa a una banda de frecuencia alta se distingue más del ruido de fondo en esa banda que cuando se compara con el ruido de fondo sobre toda la banda de frecuencias.

Un segundo objetivo de la presente invención de la presente invención es proporcionar medios para establecer los niveles umbral que tengan en cuenta la energía de la señal así como la energía del ruido de fondo. En la presente invención, el establecimiento de umbrales de detección de voz se basa en una estimación de la relación señal-ruido (SNR) de la señal de entrada. En la realización ejemplar, la energía de la señal se estima como la energía máxima de la señal durante periodos de actividad oral y la energía de ruido de fondo se estima como la energía mínima de la señal durante periodos de silencio.

Un tercer objetivo de la presente invención es proporcionar un procedimiento para codificar música que pasa a través de un vocodificador de velocidad variable. En la realización ejemplar, el aparato de selección de velocidad detecta un grupo de tramas consecutivas durante las cuales los niveles umbral han aumentado y comprueba la periodicidad relativa a dicho grupo de tramas. Si la señal de entrada es periódica indicaría la presencia de música. Si se detecta la presencia de música, entonces se establecen los umbrales a niveles tales que la señal se codifica a velocidad completa.

Breve descripción de los dibujos

Las características, los objetivos y las ventajas de la presente invención resultarán más evidentes a partir de la descripción detallada expuesta a continuación cuando se toman junto con los dibujos, en los que se emplean en toda la memoria los mismos números de referencia para las partes equivalentes, y en los que:

la figura 1 es un diagrama de bloques de la presente invención.

Descripción detallada de las realizaciones preferidas

En relación con la figura 1, la señal de entrada, $S(n)$, se proporciona al elemento de cálculo de energía de subbanda 4 y al elemento de cálculo de energía de subbanda 6. La señal de entrada $S(n)$ se compone de una señal de audio y ruido de fondo. La señal de audio suele ser voz, pero también puede tratarse de música. En la realización ejemplar, $S(n)$ se proporciona en tramas de veinte milisegundos de 160 muestras cada una. En la realización ejemplar, la señal de entrada $S(n)$ tiene componentes de frecuencia entre 0 kHz y 4 kHz, que es aproximadamente el ancho de banda de una señal de voz humana.

En la realización ejemplar, la señal de entrada de 4 kHz, $S(n)$, se filtra para obtener dos subbandas separadas. Las dos subbandas separadas se hallan entre 0 y 2 kHz y 2 kHz y 4kHz, respectivamente. En una realización ejemplar, la señal de entrada puede dividirse en subbandas mediante filtros de subbandas, cuyo diseño se conoce ampliamente en la técnica y se detalla en la patente US-A-5.644.596, transferida al cesionario de la presente invención.

Las respuestas impulsivas de los filtros de subbandas se indican por $h_L(n)$ para el filtro paso bajo, y $h_H(n)$ para el filtro paso alto. La energía de las componentes de subbandas resultantes de la señal pueden calcularse para obtener los valores $R_L(0)$ Y $R_H(0)$ simplemente sumando los cuadrados de las muestras de salida del filtro, de una forma ampliamente conocida en la técnica.

En una realización preferida, cuando la señal de entrada $S(n)$ se proporciona al elemento de cálculo de energía de subbanda 4, el valor de energía de la componente de baja frecuencia de la trama de entrada, $R_L(0)$ se calcula como:

$$R_L(0) = R_S(0) \cdot R_{h_L}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{h_L}(i) \quad (1)$$

donde L es el número de tomas en un filtro paso bajo con respuesta de impulso $h_L(n)$,

ES 2 281 854 T3

donde $R_S(i)$ es la función de autocorrelación de la señal de entrada, $S(n)$, dada por la ecuación:

$$R_S(i) = \sum_{n=1}^N S(n) \cdot S(n-1) \quad \text{para } i \in [0, L-1] \quad (2)$$

donde N es el número de muestras de la trama,

y donde R_{h_L} es la función de autocorrelación del filtro paso bajo $h_L(n)$ dada por:

$$R_{h_L}(i) = \sum_{n=0}^{L-1} h_L(n) \cdot h_L(n-i) \quad \text{para } i \in [0, L-1] \quad (3)$$

$=0$ en otro caso

La energía de alta frecuencia, $R_H(0)$, se calcula de forma similar en el elemento de cálculo de energía de subbanda

Los valores de la función de autocorrelación de los filtros subbanda pueden calcularse por adelantado para reducir la cantidad de cálculos. Además, algunos de los valores calculados de $R_S(i)$ se utilizan en otros cálculos en la codificación de la señal de entrada, $S(n)$, que reduce adicionalmente la carga de cálculo neta del procedimiento de selección de velocidad de codificación de la presente invención. Por ejemplo, la obtención de los valores de tomas del filtro LPC requiere el cálculo de un grupo de coeficientes de autocorrelación de la señal de entrada.

El cálculo de los valores de tomas del filtro LPC se conoce ampliamente en la técnica y se detalla en el documento WO-A1-92/22891. Si va a codificarse la voz con un procedimiento que requiere un filtro LPC de diez tomas sólo necesitan calcularse los valores de $R_S(i)$ para valores de i desde 11 a $L-1$, además de aquellos que se utilizan en la codificación de la señal, porque en el cálculo de los valores de tomas del filtro LPC se utiliza $R_S(i)$ para valores de i desde 0 a 10. En la realización ejemplar, los filtros subbanda tienen 17 tomas, $L=17$.

El elemento de cálculo de energía de subbanda 4 proporciona el valor calculado de $R_L(0)$ al elemento de decisión de velocidad de subbanda 12, y el elemento de cálculo de energía de subbanda 6 proporciona el valor calculado de $R_H(0)$ al elemento de decisión de velocidad de subbanda 14. El elemento de decisión de velocidad 12 compara el valor de $R_L(0)$ con dos valores umbral predeterminados $T_{L1/2}$ y T_{Lfull} y asigna una velocidad de codificación recomendada, $RATE_L$ según la comparación. La asignación de velocidad se lleva a cabo de la siguiente forma:

$$RATE_L = \text{octavo de velocidad} \quad R_L(0) \leq T_{L1/2} \quad (4)$$

$$RATE_L = \text{media velocidad} \quad T_{L1/2} < R_L(0) \leq T_{Lfull} \quad (5)$$

$$RATE_L = \text{velocidad completa} \quad R_L(0) > T_{Lfull} \quad (6)$$

El elemento de decisión de velocidad de subbanda 14 opera de forma similar y selecciona una velocidad de codificación recomendada, $RATE_H$, según el valor de energía de alta frecuencia $R_H(0)$ y basándose en un grupo diferente de valores umbral $T_{H1/2}$ y T_{Hfull} . El elemento de decisión de velocidad de subbanda 12 proporciona su velocidad de codificación recomendada, $RATE_L$, al elemento de selección de velocidad de codificación 16, y el elemento de decisión de velocidad de subbanda 14 proporciona su velocidad de codificación recomendada, $RATE_H$, al elemento de selección de velocidad de codificación 16. En la realización ejemplar el elemento de selección de velocidad de codificación 16 selecciona la más alta de las dos velocidades recomendadas y proporciona la velocidad más alta como la VELOCIDAD DE CODIFICACIÓN seleccionada.

El elemento de cálculo de energía de subbanda 4 también proporciona el valor de energía de baja frecuencia, $R_L(0)$, al elemento de adaptación de umbral 8, donde se calculan los valores umbral $T_{L1/2}$ y T_{Lfull} para la siguiente trama de entrada. De modo similar, el elemento de cálculo de energía de subbanda 6 proporciona el valor de energía de alta frecuencia, $R_H(0)$, al elemento de adaptación de umbral 10, donde se calculan los valores umbral $T_{H1/2}$ y T_{Hfull} de la siguiente trama de entrada.

El elemento de adaptación de umbral 8 recibe el valor de energía de baja frecuencia, $R_L(0)$, y determina si $S(n)$ contiene ruido de fondo o señal de audio. En un ejemplo de ejecución, el procedimiento mediante el cual el elemento de adaptación de umbral 8 determina si está presente una señal de audio examinando la función de autocorrelación normalizada NACF, que viene dada por la ecuación siguiente:

$$NACF = \max_T \frac{\sum_{n=0}^{N-1} e(n) \cdot e(n-T)}{\frac{1}{2} \cdot \left[\sum_{n=0}^{N-1} e^2(n) + \sum_{n=0}^{N-1} e^2(n-T) \right]} \quad (7)$$

donde e(n) es la señal residual formante obtenida tras el filtrado de la señal de entrada, S(n), mediante un filtro PLC.

El diseño y la filtración de una señal mediante un filtro LPC son ampliamente conocidos en la técnica y se detallan en el documento WO-A1-92/22891 mencionado anteriormente. La señal de entrada, S(n), se filtra mediante el filtro LPC para eliminar la interacción de los formantes. La función NACF se compara con un valor umbral para determinar si está presente una señal de audio. Si la función NACF es superior a un valor umbral predeterminado, indica que la trama de entrada tiene una característica periódica indicativa de la presencia de una señal de audio tal como voz o música. Obsérvese que aunque las partes de voz y música no son periódicas y presentarán valores bajos de la función NACF, el ruido de fondo normalmente nunca muestra ninguna periodicidad y casi siempre presenta valores bajos de la función NACF.

Si se determina que S(n) contiene ruido de fondo, el valor de la función NACF es inferior a un valor umbral TH1, entonces se utiliza el valor $R_L(0)$ para actualizar el valor de la estimación de ruido de fondo actual BGN_L . En la realización a modo de ejemplo, TH1 es 0,35. $R_L(0)$ se compara con el valor actual de la estimación de ruido de fondo BGN_L . Si $R_L(0)$ es inferior que BGN_L , entonces la estimación de ruido de fondo BGN_L se establece igual a $R_L(0)$ independientemente del valor de la función NACF.

La estimación de ruido de fondo BGN_L sólo se aumenta cuando la función NACF es inferior al valor umbral TH1. Si $R_L(0)$ es superior a BGN_L , y la función NACF es inferior a TH1, entonces la energía de ruido de fondo BGN_L se establece en $\alpha_1 \cdot BGN_L$, siendo α_1 un número superior a 1. En la realización ejemplar, α_1 es igual a 1,03. BGN_L continuará aumentando mientras la función NACF sea inferior al valor umbral TH1 y $R_L(0)$ sea superior al valor actual de BGN_L , hasta que BGN_L llegue a un valor máximo predeterminado BGN_{max} , momento en el que la estimación de ruido de fondo se establece en BGN_{max} .

Si se detecta una señal de audio, expresado por el valor de la función NACF que sobrepasa un segundo valor umbral TH2, entonces la estimación de la energía de señal, S_L , se actualiza. En la realización ejemplar, TH2 se establece en 0,5. El valor de $R_L(0)$ se compara con una estimación actual de la energía de señal de paso bajo, S_L . Si $R_L(0)$ es superior al valor actual de S_L , entonces S_L se establece en $R_L(0)$. Si $R_L(0)$ es inferior al valor actual de S_L , entonces S_L se establece en $\alpha_2 \cdot S_L$, de nuevo sólo si la función NACF es superior a TH2. En la realización ejemplo, α_2 se establece en 0,96.

A continuación el elemento de adaptación de umbral 8 calcula una estimación de la relación señal-ruido según la ecuación 8 siguiente:

$$SNR_L = 10 \cdot \log \left[\frac{S_L}{BGN_L} \right] \quad (8)$$

A continuación el elemento de adaptación de umbral 8 determina un índice de la relación señal-ruido cuantificada I_{SNRL} , según las ecuaciones 9 a 12 siguientes:

$$I_{SNRL} = \text{nint} \left[\frac{SNR_L - 20}{5} \right] \quad \text{para } 20 < SNR_L < 55 \quad (9)$$

$$\begin{aligned} &= 0 && \text{para } SNR_L \leq 20, && (10) \\ &= 7 && \text{para } SNR_L \geq 55. \end{aligned}$$

donde nint es una función que redondea el valor fraccional al entero más cercano.

A continuación el elemento de adaptación de umbral 8 selecciona o calcula dos factores de escala, $k_{L1/2}$ y k_{Lfull} , según el índice de la relación señal-ruido, I_{SNRL} . Se proporciona en la tabla 1 siguiente una de tabla de consulta ejemplar de valores de escala:

ES 2 281 854 T3

TABLA 1

	ISNRL	$K_{L1/2}$	K_{Lfull}
5	0	7, 0	9, 0
	1	7, 0	12, 6
10	2	8, 0	17, 0
	3	8, 6	18, 5
	4	8, 9	19, 4
15	5	9, 4	20, 9
	6	11, 0	25, 5
20	7	15, 8	39, 8

Estos dos valores se utilizan para calcular los valores umbral para la selección de velocidad según las ecuaciones siguientes:

$$T_{L1/2} = K_{L1/2} \cdot BGN_L \quad (11)$$

y

$$T_{Lfull} = K_{Lfull} \cdot BGN_L \quad (12)$$

donde $T_{L1/2}$ es el valor umbral de media velocidad de baja frecuencia y T_{Lfull} es el valor umbral de velocidad completa de baja frecuencia.

El elemento de adaptación de umbral 8 proporciona los valores umbral adaptados $T_{L1/2}$ y T_{Lfull} al elemento de decisión de velocidad 12. El elemento de adaptación de umbral 10 opera de forma similar y proporciona los valores umbral $T_{H1/2}$ y T_{Hfull} al elemento de decisión de velocidad de subbanda 14.

El valor inicial de la estimación de la energía de señal de audio S , donde S puede ser S_L o S_H , se establece según sigue. La estimación de energía de señal inicial, S_{INIT} , se establece en $-18,0$ dBm0, donde $3,17$ dBm0 denota la intensidad de la señal de una onda sinusoidal completa que, en la realización ejemplar, es una onda sinusoidal digital con un rango de amplitudes entre -8031 y 8031 . S_{INIT} se utiliza hasta que se determina que está presente una señal acústica.

El procedimiento mediante el que se detecta inicialmente una señal acústica es comparar el valor de la función NACF con un umbral, cuando la función NACF sobrepasa el umbral durante un número predeterminado de tramas consecutivas, entonces se determina si está presente una señal acústica. En la realización ejemplar, la función NACF debe sobrepasar el umbral durante diez tramas consecutivas. Una vez que se cumple esta condición, la estimación de la energía de la señal, S , se establece en la energía máxima de señal en las diez tramas precedentes.

El valor inicial de la estimación de ruido de fondo BGN_L se establece inicialmente en BGN_{max} . Tan pronto como se recibe una energía de trama de subbanda que es inferior a BGN_{max} , la estimación de ruido de fondo se restablece en el valor del nivel de energía de subbanda recibido, y se procede a la generación de la estimación de ruido de fondo BGN_L de la forma descrita anteriormente.

En una realización preferida, se acciona una condición de bloqueo cuando se detecta una trama de una velocidad baja que sigue a una serie de tramas de voz de velocidad completa. En la realización ejemplar, cuando se codifican a velocidad completa cuatro tramas de voz consecutivas seguidas de una trama en la que la VELOCIDAD DE CODIFICACIÓN se establece en una velocidad inferior a la velocidad completa y las relaciones señal-ruido calculadas son inferiores a una SNR mínima predeterminada, la VELOCIDAD DE CODIFICACIÓN para esa trama se establece en la velocidad completa. En la realización ejemplar la SNR mínima predeterminada es $27,5$ dB como se define en la ecuación 8.

En una realización preferida, el número de tramas del periodo de bloqueo es una función de la relación señal-ruido. En la realización ejemplar, el número de tramas del periodo de bloqueo se determina según sigue:

ES 2 281 854 T3

número de tramas de periodo de bloqueo=1 $22,5 < \text{SNR} < 27,5$ (13)

número de tramas de periodo de bloqueo=2 $\text{SNR} \leq 22,5$ (14)

5 número de tramas de periodo de bloqueo=0 $\text{SNR} \geq 27,5$ (15)

La presente invención también proporciona un procedimiento con el que detectar la presencia de música, que como se ha descrito anteriormente carece de las pausas que permiten restablecer las medidas de ruido de fondo. El procedimiento para detectar la presencia de música presupone que música no está presente la música al principio de la llamada. Esto permite al aparato de selección de velocidad de codificación de la presente invención estimar correctamente la energía de ruido de fondo inicial, BGN_{INIT} . Debido a que la música, a diferencia del ruido de fondo, tiene una característica periódica, la presente invención examina el valor de la función NACF para diferenciar la música del ruido de fondo. El procedimiento de detección de música de la presente invención calcula una función NACF media según la ecuación siguiente:

$$\text{NACF}_{\text{AVE}} = \frac{1}{T} \sum_{i=1}^T \text{NACF}(i) \quad (16)$$

donde NACF se define en la ecuación 7, y

donde T es el número de tramas consecutivas en las que el valor estimado del ruido de fondo ha ido aumentando a partir de la estimación de ruido de fondo inicial BGN_{INIT} .

Si el ruido de fondo BGN ha ido aumentando durante el número de tramas T predeterminadas y la función NACF_{AVE} sobrepasa un umbral predeterminado, entonces se detecta música y el ruido de fondo BGN se restablece en BGN_{init} . Debe observarse que, para ser eficaz, el valor T debe establecerse suficientemente bajo para que la velocidad de codificación no descienda por debajo de la velocidad completa. Por lo tanto, el valor de T debe establecerse como una función de la señal acústica y de BGN_{init} .

La descripción anterior de las realizaciones preferidas se proporciona para permitir que cualquier experto en la técnica pueda crear o utilizar la presente invención. Las diversas modificaciones de estas realizaciones resultarán evidentes fácilmente para los expertos en la técnica, y los principios genéricos definidos en el presente documento pueden aplicarse a otras realizaciones sin el uso de la actividad inventiva. Por lo tanto, la presente invención no está prevista para limitarse a las realizaciones mostradas en el presente documento, sino que debe de estar de acuerdo con el alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

5 1. Un aparato para determinar una velocidad de codificación para un vocodificador de velocidad variable que comprende:

medios de cálculo de energía de subbanda (4, 6) para recibir una señal de entrada (S(n)) y determinar una pluralidad de valores de energía de subbanda según un formato de cálculo de energía de subbanda predeterminado;

10 medios de cálculo de umbral (8, 10) para determinar una estimación de energía de señal y una estimación de ruido de fondo, y para determinar una pluralidad de valores umbral de velocidad de codificación en cada subbanda, basándose cada valor umbral de velocidad de codificación en una relación de dicha estimación de energía de señal con respecto a dicha estimación de ruido de fondo; y

15 medios de determinación de velocidad (12, 14, 16) para recibir dicha pluralidad de valores de energía de subbanda y dicha pluralidad de valores umbral de velocidad de codificación, y para determinar dicha velocidad de codificación para dicha señal de entrada (S(n)) con dicha pluralidad de valores de energía de subbanda y dicha pluralidad de valores umbral de velocidad de codificación.

20 2. Aparato según la reivindicación 1, en el que dichos medios (4, 6) de cálculo de energía de subbanda están adaptados para determinar cada una de dicha pluralidad de valores de energía de subbanda según la ecuación:

$$25 \text{ energía de subbanda} = R_S(0) \cdot R_{\text{hbp}}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{\text{hbp}}(i),$$

donde L es el número de tomas en un filtro pasa bandas hbp (n), donde $R_S(i)$ es la función de autocorrelación de la señal de entrada, S(n), y donde R_{hbp} es la función de autocorrelación del filtro pasa bandas hbp (n).

30 3. Aparato según la reivindicación 1, en el que dichos medios de cálculo de umbral (8, 10) están adaptados para determinar un valor de escala según dicho valor de relación señal-ruido.

35 4. Aparato según la reivindicación 3, en el que los medios (8, 10) de cálculo de umbral están adaptados para determinar al menos un valor umbral multiplicando una estimación de ruido de fondo por dicho valor de escala.

5. Aparato según la reivindicación 1, en el que dichos medios de determinación de velocidad están adaptados para comparar al menos uno de dicha pluralidad de valores de energía de subbanda con al menos un valor umbral para determinar dicha velocidad de codificación.

40 6. Aparato según la reivindicación 4, en el que dichos medios de determinación de velocidad están adaptados para comparar al menos uno de dicha pluralidad de valores de energía de subbanda con dicho al menos un valor umbral para determinar dicha velocidad de codificación.

45 7. Aparato según la reivindicación 1, en el que dichos medios de determinación de velocidad (12, 14, 16) están adaptados para determinar a una pluralidad de velocidades de codificación recomendadas en el que cada velocidad de codificación recomendada corresponde a cada uno de entre dicha pluralidad de valores de energía de subbanda, y en el que dichos medios de determinación de velocidad están adaptados para determinar dicha velocidad de codificación según dicha pluralidad de velocidades de codificación recomendadas.

50 8. Aparato según la reivindicación 1, en el que dichos medios de cálculo de energía de subbanda (4, 6) comprenden un calculador de energía de subbanda, y en el que dichos medios de determinación de velocidad (12, 14, 16) comprenden un selector de velocidad que está adaptado para recibir dicha pluralidad de valores de energía de subbanda y para seleccionar dicha velocidad de codificación según dicha pluralidad de valores de energía de subbanda.

55 9. Aparato según la reivindicación 8, en el que dicho calculador de energía de subbanda está adaptado para determinar cada uno de entre dicha pluralidad de valores de energía de subbanda según la ecuación:

$$60 \text{ energía de subbanda} = R_S(0) \cdot R_{\text{hbp}}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{\text{hbp}}(i)$$

65 donde L es el número de tomas en un filtro pasa bandas hbp (n), donde $R_S(i)$ es la función de autocorrelación de la señal de entrada, S(n), y donde R_{hbp} es la función de autocorrelación del filtro pasa bandas hbp (n).

10. Aparato según la reivindicación 8, que comprende además un calculador de umbral dispuesto entre dicho calculador de energía de subbanda y dicho selector de velocidad, estando adaptado dicho calculador de umbral para

ES 2 281 854 T3

recibir dichos valores de energía de subbanda y determinar una serie de valores umbral de velocidad de codificación según la pluralidad de valores de energía de subbanda.

5 11. Aparato según la reivindicación 10, en el que dicho calculador de umbral está adaptado para determinar un valor de relación señal-ruido según dicha pluralidad de valores de energía de subbanda.

12. Aparato según la reivindicación 11, en el que dicho calculador de umbral está adaptado para determinar un valor de escala según dicho valor de relación señal-ruido.

10 13. Aparato según la reivindicación 12, en el que el calculador de umbral está adaptado para determinar al menos un valor umbral multiplicando una estimación de ruido de fondo por dicho valor de escala.

14. Aparato según la reivindicación 8, en el que dicho selector de velocidad está adaptado para comparar al menos uno de dicha pluralidad de valores de energía de subbanda con al menos un valor umbral para determinar dicha velocidad de codificación.

15 15. Aparato según la reivindicación 13, en el que dicho selector de está adaptado para comparar al menos uno de dicha pluralidad de valores de energía de subbanda con dicho al menos un valor umbral para determinar dicha velocidad de codificación.

20 16. Aparato según la reivindicación 8, en el que dicho selector de velocidad está adaptado para determinar una pluralidad de velocidades de codificación recomendadas, en el que dicha velocidad de codificación recomendada corresponde a cada uno de entre dicha pluralidad de valores de energía de subbanda, y en el que dicho selector de velocidad está adaptado para determinar dicha velocidad de codificación según dicha pluralidad de velocidades de codificación recomendadas

25 17. Un procedimiento para determinar una velocidad de codificación para un vocodificador de velocidad variable que comprende las etapas de:

30 recibir una señal de entrada (S(n));

determinar una pluralidad de valores de energía de subbanda según un formato de cálculo de energía de subbanda predeterminado;

35 determinar un valor de relación señal-ruido basado en una relación de una estimación de energía de señal respecto a una estimación de ruido de fondo;

40 determinar una pluralidad de valores umbral de velocidad de codificación en cada subbanda basado en dicho valor de relación señal-ruido; y

determinar dicha velocidad de codificación para dicha señal de entrada (S(n)) según dicha pluralidad de valores de energía de subbanda y dicha pluralidad de valores umbral de velocidad de codificación.

45 18. Procedimiento según la reivindicación 17, el que dicha etapa de determinar una pluralidad de valores de energía de subbanda se realiza según la ecuación:

50
$$\text{energía de subbanda} = R_s(0) \cdot R_{\text{hbp}}(0) + 2 \cdot \sum_{i=1}^{L-1} R_s(i) \cdot R_{\text{hbp}}(i)$$

55 donde L es el número de tomas en un filtro pasa bandas hbp (n), donde $R_s(i)$ es la función de autocorrelación de la señal de entrada, S(n), y donde R_{hbp} es la función de autocorrelación del filtro pasa bandas hbp (n).

19. Procedimiento según la reivindicación 17, en el que dicha etapa de determinar una serie de valores umbral de velocidad de codificación determina un valor de escala según dicho valor de relación señal-ruido.

60 20. Procedimiento según la reivindicación 19, en el que dicha etapa de determinar una serie de valores umbral de velocidad de codificación determina dicho valor umbral de velocidad multiplicando una estimación de ruido de fondo por dicho valor de escala.

65 21. Procedimiento según la reivindicación 17, en el que dicha determinación de dicha velocidad de codificación compara al menos uno de dicha pluralidad de valores de energía de subbanda con al menos un valor umbral para determinar dicha velocidad de codificación.

ES 2 281 854 T3

22. Procedimiento según la reivindicación 20, en el que dicha etapa de dicha determinación de dicha velocidad de codificación compara al menos uno de dicha pluralidad de valores de energía de subbanda con dicho al menos un valor umbral para determinar dicha velocidad de codificación.

5 23. Procedimiento según la reivindicación 17, que comprende además la etapa de generar una velocidad de codificación recomendada según cada uno de entre dicha pluralidad de valores de energía de subbanda, y en el que dicha etapa de determinar una velocidad de codificación selecciona una de dichas velocidades de codificación recomendadas.

10 24. Aparato según la reivindicación 1, en el que dichos medios de cálculo de energía de subbanda comprenden un subsistema de filtro de subbanda para determinar una energía de señal para cada subbanda de frecuencia de la señal de entrada; y en el que dichos medios de determinación de velocidad comprenden un subsistema de selección de velocidad para seleccionar la velocidad de codificación de la señal de entrada basada en las energías de señal de cada subbanda de frecuencia de la señal de entrada (S(n)).

15 25. Aparato según la reivindicación 24, en el que el subsistema de filtro de subbanda comprende una pluralidad de elementos de cálculo de energía de subbanda (4, 6), y cada uno de entre la pluralidad de elementos de cálculo de energía de subbanda está adaptado para determinar una energía de señal de subbanda de frecuencia.

20 26. Aparato según la reivindicación 25, en el que el subsistema de selección de velocidad comprende una pluralidad de elementos de adaptación de umbral (8, 10), y cada uno de entre la pluralidad de elementos de adaptación de umbral está adaptado para utilizar la energía de señal de subbanda de frecuencia a partir de un elemento de cálculo de energía de subbanda correspondiente (4, 6) adaptado para determinar si está presente una señal de audio en la subbanda de frecuencia.

25 27. Aparato según la reivindicación 26, en el que cada elemento de adaptación de umbral (8, 10) está configurado para determinar un valor umbral basado en la energía de señal y una estimación de ruido de la subbanda de frecuencia correspondiente, en el que el valor umbral se utiliza para determinar si la señal de audio está presente en la subbanda de frecuencia.

30 28. Aparato según la reivindicación 26, en el que la pluralidad de elementos de adaptación de umbral (8, 10) están configurados para determinar un valor umbral basado en las energías de señal combinadas para cada una de entre las subbandas de frecuencia de la señal de entrada (S(n)), en el que el valor umbral se utiliza para determinar si la señal de audio está presente en la subbanda de frecuencia.

35

40

45

50

55

60

65

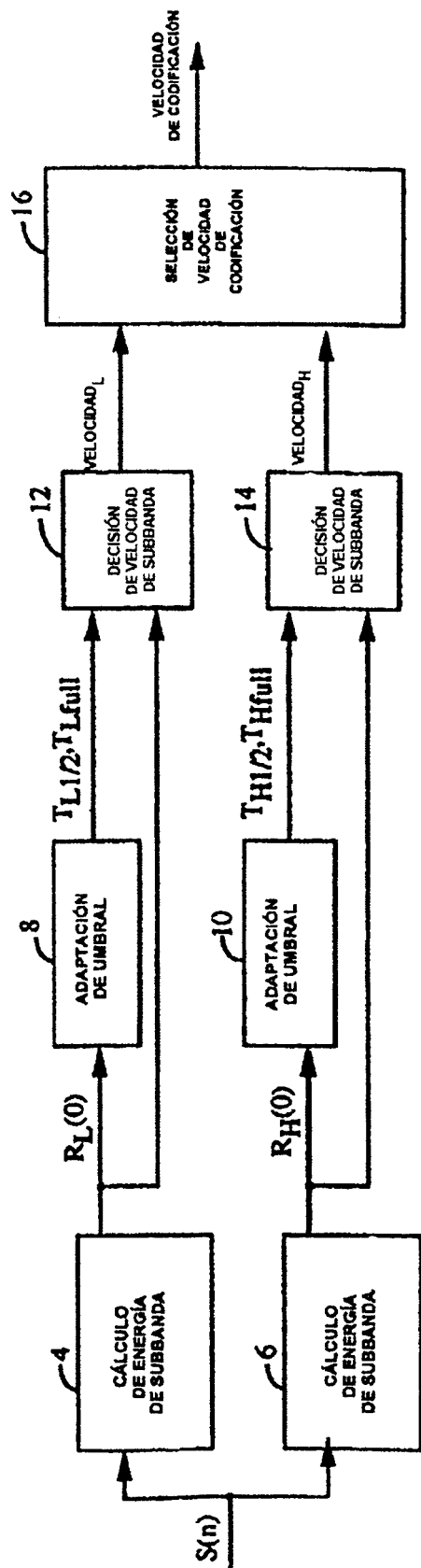


FIG. 1