



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 312 852**

51 Int. Cl.:  
**G10L 19/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03808458 .8**

96 Fecha de presentación : **16.12.2003**

97 Número de publicación de la solicitud: **1581928**

97 Fecha de publicación de la solicitud: **05.10.2005**

54 Título: **Reducción del coste de transmisión del factor de escala para MPEG-2 AAC mediante una retícula.**

30 Prioridad: **02.01.2003 US 336637**

45 Fecha de publicación de la mención BOPI:  
**01.03.2009**

45 Fecha de la publicación del folleto de la patente:  
**01.03.2009**

73 Titular/es: **DOLBY LABORATORIES LICENSING CORPORATION**  
**100 Potrero Avenue**  
**San Francisco, California 94103-4813, US**

72 Inventor/es: **Vinton, Mark, Stuart**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 312 852 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).

# ES 2 312 852 T3

## DESCRIPCIÓN

Reducción del coste de transmisión del factor de escala para MPEG-2 AAC mediante una retícula.

### 5 Técnica anterior

Las técnicas de codificación de audio por banco de filtros y transformación típica tales como MPEG-1 capas 1 a 3, Dolby AC-3 (conocido también como Dolby Digital) (Dolby, Dolby Digital y Dolby AC-3 son marcas comerciales de Dolby Laboratories Licensing Corporation) y Codificación de Audio Avanzada MPEG-2 (AAC) reducen los regímenes de transmisión de datos al asignar dinámicamente bits tanto en tiempo como en frecuencia para eliminar redundancias inaudibles en la señal de audio. La asignación dinámica de bits está basada típicamente en principios psicoacústicos dependientes de la señal. Detalles adicionales del Dolby AC-3 se pueden encontrar en el Estándar de Compresión Digital de Audio (AC-3). Aprobado el 10 de noviembre de 1994. (Rev 1) Anexo A añadido el 12 de abril de 1995. (Rev 2) 13 fe de erratas añadida el 24 de mayo de 1995. (Rev 3) Anexo B y C añadidos el 20 de diciembre de 1995. Detalles adicionales del AAC se pueden encontrar en el documento “Codificación de Audio MPEG-2 ISO/IEC” por Bosi y otros, presentada en la convención 101, del 8 al 11 de noviembre de 1996, Los Ángeles, Borrador de la Sociedad de Ingeniería de Audio 4382).

En AAC, la asignación de bits se consigue usando factores de escala y parámetros de ganancia global contenidos en el flujo de bits. El espectro de audio, transformado usando una bien conocida transformada directa de coseno modificada (MDCT: modified discrete cosine transform) conocida como cancelación de irregularidades de dominio en el tiempo (TDAC: time domain alias cancellation) (véase el documento de Princen y otros, “diseño de banco de filtros por análisis/síntesis basado en cancelación de irregularidades de dominio en el tiempo”, IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-34, páginas. 1153-1161, octubre de 1986), es repartido en bandas de aproximadamente la mitad de un ancho de banda crítico y se aplican factores de escala de manera multiplicativa. Los factores de escala y la ganancia global representan de forma unida representan asignación de bits en etapas de 1,5 dB o aproximadamente incrementos de cuartos de bits (la asignación exacta de bits conseguida es dependiente de las características estocásticas de la señal de audio y se complica además por el cuantificador no lineal incorporado en AAC). Aumentar el factor de escala en una banda reduce efectivamente el ruido de cuantificación en esa banda al asignar más bits a esa banda. Inversamente, disminuir un factor de escala aumenta el ruido de cuantificación en una banda particular al reducir los bits asignados a ella.

Como AAC es un sistema de codificación de audio adaptable hacia delante, los factores de escala son comunicados al decodificador. Esto se consigue codificando diferencialmente los factores de escala y después codificando con Huffman las diferencias. Los códigos Huffman definidos en el estándar AAC, son tales que variaciones grandes en los parámetros del factor de escala de banda a banda llevan al consumo excesivo de los bits disponibles en forma de información lateral, lo que complica las derivaciones de factores de escala como se explica en la siguiente sección.

### 40 *Cálculo del factor de escala*

Calcular los factores de escala en un codificador AAC es un problema muy difícil debido a la incertidumbre en la asignación de ruido conseguida al alterar los factores de escala y el uso de una etapa de cuantificador no lineal. Comúnmente se usan dos técnicas en AAC para calcular factores de escala, a saber análisis por síntesis y estimación directamente desde el modelo de enmascarado, que se describen después. Mientras que la selección de los factores de escala puede ser arbitraria, dentro de algunas limitaciones impuestas por el estándar, estas dos técnicas son las más conocidas.

### *Cálculo de factor de escala usando análisis por síntesis*

El cálculo de factor de escala usando análisis por síntesis se consigue usando dos bucles anidados, un bucle interno responsable de la cuantificación y conteo de bits y un bucle externo, que analiza el resultado del bucle interno y altera los factores de escala correspondientemente.

El bucle interno altera el parámetro de ganancia global contenido en el flujo de bits AAC para asegurar que el número de bits usado para codificar el espectro de audio no es más del número de bits disponible. La ganancia global se establece en un valor inicial y el espectro es cuantificado. Los números de bits usados son contados después. Si el número de bits usado es más grande del número de bits disponible, entonces la ganancia global es aumentada y el espectro es cuantificado de nuevo y el número de bits usados es recontado. Este procedimiento se repite hasta que el número de bits usados es menor que el número de bits disponible. El bucle interno es referido a menudo como un “bucle de régimen” ya que controla el régimen de codificación de bits.

El bucle externo analiza el resultado conseguido por el bucle interno y altera los factores de escala de tal manera que el ruido de cuantificación en cada banda cumple los requisitos psicoacústicos tan estrechamente como sea posible. El bucle externo empieza con todos los factores de escala establecidos a cero y el bucle interno es llamado para cuantificar el espectro. La distorsión (ruido cuantificador) en cada banda es calculada entonces y comparada con los requisitos de ruido para cada banda según se calcula por el modelo psicoacústico. Si la distorsión en alguna banda es mayor que la distorsión permisible calculada por el modelo psicoacústico, entonces el factor de escala para la banda es aumentado. El bucle interno es llamado de nuevo con los factores de escala ajustados y el procedimiento se repite hasta

que (1) la distorsión en todas bandas es menor que el nivel de enmascarado calculado por el modelo psicoacústico o (2) todos factores de escala han sido aumentados.

La técnica de análisis por síntesis padece varios problemas; primero, la técnica es extremadamente compleja y, consecuentemente, no es apropiada para aplicaciones de complejidad limitada. Además de esto, el procedimiento de bucle dual descrito antes no garantiza convergencia en una solución óptima; sin embargo, a mayores regímenes de datos se ha mostrado que produce resultados excelentes.

#### *Estimación del factor de escala para el nivel de enmascaramiento*

Asumiendo que aumentando el factor de escala en una unidad en una banda lleva a una reducción de 1,5 dB de distorsión de cuantificación en esa banda (un aumento del ratio señal a ruido) (tanto la ganancia global como los factores de escala son cuantificados en etapas de 1,5 dB), los factores de escala pueden ser derivados directamente desde el modelo de enmascaramiento como se describe en "Codificación MPEG-2 AAC de eficiencia aumentada" por Smithers y otros, Audio Engineering Society Convention Paper, presentado en la 111 convención, del 21 al 24 de septiembre de 2001, en Nueva York. Para esta técnica, los factores de escala son calculados primero directamente del modelo de enmascaramiento, por ejemplo, usando la expresión establecida abajo en EQN. 1, en la que  $s_i$  es el factor de escala para la banda  $i$  y  $m_i$  es el nivel de enmascaramiento para la banda  $i$  calculado por el modelo psicoacústico.

$$s_i = \frac{-2}{\log_{10}(2)} \cdot \log_{10}(m_i) \quad (\text{EQN. 1})$$

El espectro es cuantificado entonces usando el bucle interno (o bucle de régimen) descrito en la sección previa, eliminando así la necesidad de un bucle externo de alta complejidad. Aunque esta técnica es mucho más simple que la técnica de análisis por síntesis descrita en la sección previa y por tanto es apropiada para sistemas de complejidad limitada, el cálculo de los factores de escala del modelo de enmascarado genera factores de escala que exhiben mayor variación de banda a banda que los generados por la técnica análisis por síntesis de dos bucles. Como los factores de escala están codificados diferencialmente y después codificados por Huffman (mayores diferencias implican palabras de código Huffman más largas), alta variación en los factores de escala significa que el coste de bits de transmitir los factores de escala es muy alto, lo que degrada el rendimiento de la estimación del factor de escala de la técnica de nivel de enmascarado.

El documento de Aggarwal A; Regunathan S L; Rose K: "OPTIMIZACIÓN BASADA EN TRELIS DE CODIFICACIÓN AVANZADA DE AUDIO CON MPEG-4", Seminario IEEE en Codificación de Lenguaje. Procedimientos. Centrarse en los retos del nuevo Milenio, Septiembre de 2000, páginas 142 a 144 ("TRELIS-BASED OPTIMIZATION OF MPEG-4 ADVANCED AUDIO CODING", IEEE Workshop on Speech Coding. Proceeding. Meeting the Challenges of the new Millenium, September 2000, pages 142 to 144) está dirigido a mejorar el rendimiento del régimen bajo de bits del MPEG-4 AAC al tener en cuenta el régimen total de bits, incluyendo información lateral. La optimización de la distorsión de régimen se realiza empleando una búsqueda Viterbi basada en entramado de todas combinaciones de posibles factores de escala (SF) y estados de valor de libro de código Huffman (HCB). El documento proporciona una visión general del procedimiento convencional de cuantificación AAC y formula un problema - el objeto de un codificador AAC es elegir los parámetros de cuantificación (valores de libro de código Huffman y factores de escala) para todas bandas de frecuencia para minimizar la relación de ruido a máscara (NMR) sometido al límite dado de régimen de bits, pero el procedimiento de optimización del parámetro de cuantificación AAC ignora dependencias entre bandas introducidas por codificación AAC y falla también al tener en cuenta el número total de bits requeridos, incluyendo bits para la información lateral. Un procedimiento de búsqueda, que emplea un procedimiento de programación dinámica (Viterbi) que busca todas combinaciones de posibles de estados de valor de libro de código Huffman y factores de escala sin emplear valores preliminares SF y/o HCB, encuentra los parámetros de cuantificación que minimizan las funciones de coste.

#### **Descripción del invento**

El presente invento está dirigido a un método para reducir el coste total de bits de un codificador perceptivo de audio que emplea asignación adaptable de bits de acuerdo a la reclamación 1.

Ninguna de las técnicas descritas antes para calcular factores de escala en AAC tiene en cuenta explícitamente el coste de transmitir los factores de escala al descodificador. En particular, la técnica de derivación directa más simple puede permitir que el coste de transmisión del factor de escala supere el 10% (a 128 kbps para material estéreo) del régimen total de datos disponible para transmisión de audio, degradando así la representación descodificada. Para afrontar el problema, el presente invento emplea una técnica dinámica de optimización de programación, que incluye, por ejemplo, un algoritmo de búsqueda Viterbi y por entramado, para reducir el coste de bits de transmitir información del factor de escala en AAC (MPEG-2/4 Advanced Audio Coding: Codificación Avanzada de Audio). El invento minimiza una función de coste que equilibra el coste de transmitir los factores de escala contra el coste de cambiar los factores de escala los factores de escala de los valores preliminares derivados por una técnica preliminar de cálculo

de factor de escala. En particular, factores de escala que tienen valores menores que otros pueden ser desplazados a valores mayores para reducir la extensión de variaciones en valor de factor de escala de una banda de factor de escala a la siguiente. Aunque un aumento en el valor de factor de escala provoca que más bits sean asignados a una banda de factor de escala, hay un ahorro general de bits en reducir el grado de variaciones banda a banda en valores de factor de escala ya que diferencias de banda a banda son codificadas en Huffman de tal manera que la longitud de código aumenta con variaciones banda a banda en aumento. El ahorro general de bits hace más bits disponibles para que el cuantificador asigne a bandas de factor de escala otros distintos de los que el valor del factor de escala es aumentado con la finalidad de reducir variaciones banda a banda, dando lugar, por tanto, a una mejora de la calidad de audio percibida.

Aunque el invento es aplicable a formas de AAC que emplea dos bucles anidados en el cuantificador para derivar factores de escala preliminares, tanto un bucle de iteración interno como un bucle de iteración externo (como se describe en el documento mencionado antes de Bosi y otros), el invento es particularmente beneficioso cuando se emplea en una forma de AAC en la que el bucle externo, que calcula el error de cuantificador y deriva factores de escala usando análisis por síntesis, es omitido y los factores de escala preliminares son estimados usando el umbral de enmascarado derivado por la parte porcentual de modelo del codificador AAC. Tal forma modificada de AAC se describe en el documento de convención identificado antes de Smithers y otros. La técnica de programación dinámica acorde con el presente invento es substancialmente menos compleja en el aspecto computacional que el bucle externo omitido, pero da lugar a señal codificada que tiene substancialmente la misma calidad que la producida por un codificador AAC que emplea dos bucles anidados.

### Descripción de los dibujos

La figura 1 es un diagrama de bloques esquemático funcional de un procedimiento de codificación que incorpora optimización dinámica de factor de escala de programación acorde con el presente invento.

La figura 2 es un diagrama de flujo simplificado que muestra la aplicación de un algoritmo de búsqueda de Viterbi a una ecuación de coste de bits del tipo empleado preferiblemente en el presente invento.

La figura 3 son representaciones de valores de factores de escala de ejemplo frente a bandas de factor de escala para el caso de factores de escala preliminares que provienen de una técnica directa de estimación de factor de escala y para factores de escala ajustados resultantes de la optimización del coste de bits de acuerdo con el presente invento.

La figura 4 son representaciones de formas de onda de ejemplo que indican el coste de bits de factores de escala por estructura resultantes de una técnica de estimación directa de factor de escala y para factores de escala ajustados resultantes de optimización de coste de bits de acuerdo con el presente invento.

### Mejor modo de realizar el invento

La figura 1 muestra un esquema simple de alto nivel de un procedimiento de codificación AAC que incorpora optimización dinámica de factor de escala de programación de acuerdo con el presente invento. La figura anterior muestra la optimización de factor de escala de acuerdo con el presente invento junto con la estimación directa de factor de escala desde la información de modelo descrita antes. Aunque otras técnicas de derivación de factor de escala pueden ser mejoradas usando las enseñanzas de este invento, el invento es particularmente adecuado para usar con esta técnica de estimación directa.

En la figura 1, el audio de entrada es transformado usando un MDCT2, seguido por procesamiento previo 4 (por ejemplo, modelación temporal de ruido (TNS: temporal noise shaping), predicción y codificación de medio lado (MS: mid-side) para aplicaciones estéreo). La entrada es pasada también a un modelo psicoacústico 6, que calcula el nivel de enmascarado. Como se ha explicado antes, el modelo de enmascarado se usa directamente para computar los factores de escala para cada banda ("cálculo del factor de escala" 8). Mientras que los factores de escala preliminares derivados por esta técnica se aproximan al requisito psicoacústico muy estrechamente, la alta variación banda a banda en los valores del factor de escala llevan a un alto coste de transmisión. Para minimizar este coste, la optimización 10 del factor de escala de acuerdo con el presente invento procesa los factores de escala preliminares antes de su aplicación al espectro MDCT en el bucle 12 de régimen y codificación sin ruido (codificación diferencial Huffman) 14.

Se asume que aumentando el valor de un factor de escala por una unidad en una banda aumenta el número de bits usados en esa banda por un quarter bit por coeficiente MDCT. Aunque esto no siempre es preciso debido a la naturaleza estocástica desconocida de la señal y el cuantificador no uniforme usado en AAC, en promedio es una suposición razonable. Se asume adicionalmente que factores de escala preliminares ya han sido determinados para un rendimiento psicoacústico apropiado, bien por análisis por síntesis o por técnicas de estimación directa de enmascarado. La siguiente fórmula de costes equilibra el coste de la transmisión de factor de escala con el coste de aplicar más bits a una banda particular. La función de coste se da más abajo en EQN. 2.

$$C = \sum_i \left( \alpha_i \frac{(\tilde{s}_i - s_i)}{4} B_i + D(\tilde{s}_i - s_{i-1}) - D(s_i - s_{i-1}) \right) \quad (\text{EQN. 2})$$

5

En la ecuación EQN. 2, C es el coste total de cambiar los factores de escala, que se debería hacer tan negativo como sea posible para reducir los costes relativos de transmisión de factores de escala. El símbolo  $s_i$  representa los factores de escala preliminares derivados, por ejemplo, para consideraciones psicoacústicas por cualquiera de las técnicas descritas antes. Además de esto,  $\tilde{s}_i$  es el nuevo grupo de factores de escala en EQN. 2 y  $B_i$  es el número de coeficientes en la banda  $i$  de factor de escala. La función  $D()$  es la búsqueda Huffman de los factores de escala diferenciales codificados. La escala  $\alpha_i$  de banda previa es un valor entre 0 y 1 que estima el número de coeficientes MDCT que serán cuantificados a valores distintos de cero. El parámetro  $\alpha_i$ , que es una función del valor del factor de escala, es opcional (si se omite, es reemplazado por un valor constante igual a 1) pero mejora enormemente el rendimiento del algoritmo si se estima con precisión. En esta ecuación, se asume que  $\alpha_i$  sea constante si los factores de escala solo son modificados ligeramente de su valor preliminar. Por simplicidad, esto puede conseguirse contando el número de coeficientes MDCT en una banda que tiene un valor absoluto mayor que algún umbral predefinido.

10

15

20

Para la EQN. 2 de coste de bits del factor de escala, los nuevos factores de escala solo se les permite tomar valores mayores o iguales a los valores preliminares, por tanto el sistema no puede disminuir los bits asignados a una banda pero solo puede aumentar el número de bits si los bits adicionales resultantes de un factor de escala aumentado es más barato que el coste diferencial codificado de los factores de escala. La función  $D(s_i - s_{i-1})$ , la búsqueda Huffman de los factores de escala diferenciales codificados aplicados al grupo original de factores de escala, es una constante en la EQN. 2 y puede ser eliminado en la práctica.

25

30

35

40

Es deseable optimizar el valor de factor de escala en cada banda de factor de escala para minimizar el número total de bits requeridos. Una optimización adecuada puede ser conseguida poblando un entramado (a veces se le hace referencia como una "retícula") de tal manera que sus nodos en cada etapa o nivel consecutivo (bandas de factor de escala "i") sean los estados posibles (valores "k" de factor de escala) para esa etapa y aplicando un algoritmo adecuado de búsqueda, tal como un algoritmo de búsqueda Viterbi, que es una técnica de búsqueda de mínimo coste particularmente adecuada para un entramado. En este contexto, el algoritmo Viterbi determina el mínimo recorrido de bits a través del entramado, optimizando por tanto el valor del factor de escala en cada banda de factor de escala. El algoritmo Viterbi computa el mejor recorrido (más barato) para cada nodo (valor del factor de escala) en cada etapa (banda de factor de escala) al encontrar la mejor extensión (menor régimen de bits) de los nodos previos (valores de factor de escala). Tales computaciones son realizadas para cada etapa (banda de factor de escala) hasta la última. En cada etapa (banda de factor de escala), el algoritmo lleva la cuenta de: (1) el mejor recorrido en cada nodo (valor del factor de escala), y (2) el coste acumulado hasta ese nodo (valor de factor de escala). Conociendo el mejor recorrido en un nodo es equivalente dentro de un nodo es equivalente a conocer en cada valor de nodo (factor de escala) el mejor valor de nodo predecesor (factor de escala), determinando así el mejor recorrido a través del entramado y minimizando el número total de bits requeridos. El valor del factor de escala en cada banda de factor de escala es optimizado para cada estructura sucesiva (bloque) de audio digital. El algoritmo de búsqueda Viterbi se conoce bien. Véase, por ejemplo, el capítulo 15 ("Codificación en entramado y Árbol": "Tree an Trellis Encoding") de Compresión de Señal y Cuantificación de Vector por Allen Gersho y Robert M. Gray, Kluwer Academic Publishers, Boston, 1992, páginas 555-586.

45

50

Más específicamente, para minimizar la función de coste en EQN. 2, se puede emplear una técnica dinámica de optimización de programación, tal como un algoritmo de búsqueda Viterbi, como sigue. Una retícula o entramado está construida con el estado número  $k$  en la etapa número  $i$  denominada  $s_{ki}$  y el coste acumulado en cualquier estado  $k$  y etapa  $i$  se denomina como  $C_{ki}$ . Cada estado en la retícula representa los valores posibles del nuevo grupo de factor de escala después de la optimización. El algoritmo se calcula después usando las siguientes etapas:

1) Iniciar  $i = 0$  y  $C_{ki} = 0$

2) para todos  $k$  de tal manera que  $s_{ki} > s_i$  ( $s_i$  son el grupo de factores de escala preliminares) encontrar

55

$$C = \min \left( \alpha_i \frac{(\tilde{s}_i - s_i)}{4} B_i + D(\tilde{s}_i - s_{i-1}) - D(s_i - s_{i-1}) \right) \forall I \quad (\text{EQN. 3})$$

60

Si  $i <$  número de bandas de factor de escala  $i = i + 1$ , volver a la etapa 2

65

El nuevo grupo de factores de escala,  $\tilde{s}_i$ , son el recorrido a través del entramado de tal manera que  $C_{ki}$  es minimizado en la etapa final. El algoritmo de búsqueda Viterbi se entiende bien y hay ampliamente disponibles técnicas de implementación eficientes. Se pueden emplear alternativas a un algoritmo de búsqueda Viterbi de tal como, por ejemplo, otras técnicas de optimización de retícula.

## ES 2 312 852 T3

Un ejemplo de la aplicación de un algoritmo de búsqueda Viterbi en la EQN. 3 se describe ahora con relación al diagrama de flujo de la figura 2.

5 La figura 2 muestra un diagrama de flujo de un procedimiento que emplea un algoritmo de búsqueda Viterbi para minimizar la función de coste de la EQN. 3 para cada estructura de audio digital. Como se indica en el bloque 102, primero, el factor de escala para cada banda de factor de escala es estimado, teniendo en cuenta requisitos psicoacústicos. Esto puede conseguirse, por ejemplo, en la manera descrita en el documento de Smithers y otros, mencionado antes.

10 Los factores de escala para cada banda de factor de escala se representan por una matriz,  $SF[i]$ , en la que la variable “ $i$ ” puede estar comprendida entre entre cero y  $N-1$ , en el que  $N$  es el número de bandas de factor de escala en una estructura de audio. Una segunda serie,  $Cost[k]$ , representa el coste acumulado de un recorrido a través del entramado. Una serie  $Historia[i][k]$ , almacena el recorrido más barato para cada nodo (valor del factor de escala) en una etapa (banda de factor de escala) en el entramado. La variable “ $k$ ” (el valor de factor de escala) puede estar comprendida  
15 entre cero y  $MAX-1$ , donde  $MAX$  es el número de valores de factor de escala.

Un contador de etapa (banda de factor de escala) “ $i$ ” es iniciado a cero en el bloque iniciador 104, que, además de iniciar la banda “ $i$ ” de factor de escala a cero, inicia también  $Historia[i][k]$  a cero y  $Cost[k]$  a cero. El contador de etapa es aumentado en el bloque 116 hasta que todas bandas  $i$  de factor de escala son procesadas según se determina  
20 en el bloque de decisión 114.

Para cada etapa (banda de factor de escala)  $i$  en el entramado, se determina la ruta más barata hacia cada nodo (valor de factor de escala)  $k$  en esa etapa. Esto se hace usando los dos bucles anidados, un bucle 108 y un bucle 110.

25 La variable  $k$  en el bloque de decisión 118 es iniciada a cero por el bloque 116 y es aumentada por el bloque 128 del primer bucle anidado 118, el bucle “ $k$ ”, hasta que todos valores posibles de factor de escala, representados por los nodos en la etapa  $i$  (banda de factor de escala  $i$ ) son comprobados respecto al coste usando el segundo bucle anidado 110, el bucle “ $m$ ”. En el bloque 130, el segundo bucle anidado 110 calcula el coste acumulado de recorrido desde la etapa  $i-1$  (banda de factor de escala  $i-1$ ) a la etapa  $i$  (banda de factor de escala  $i$ ) del entramado de acuerdo con EQN. 3 si el valor de factor de escala para la banda de factor de escala  $i$  es mayor o igual que la estimación preliminar de factor de escala (bloque 102). Si el factor de escala no es mayor o igual que el factor de escala preliminar para esa  
30 banda de factor de escala, entonces el coste acumulado para esa banda de factor de escala es establecido, por ejemplo, en un valor grande arbitrario para asegurar que este recorrido a través del entramado no es posible. La variable  $m$  en el bloque de decisión 124 es iniciada a cero por el bloque 122 y es aumentada por el bloque 132 del segundo bucle anidado 110. La variable “ $m$ ” (el número de nodos de recorrido pasados) puede estar comprendido entre cero y  $MAX-1$ , donde  $MAX$  es el número de nodos de recorrido pasados.  
35

El coste acumulado para cada grupo de nodos de recorrido pasados es almacenado en una matriz temporal,  $TempCost[m]$ , cuyo valor se da por:

$$40 \quad TemCost[m] = Cost[m] + Alpha[i]*(k-SF[i])*B[i]/4+D(k-m),$$

45 donde  $Alpha[i]$  es una escala para banda de factor de escala para compensar coeficientes MDCT cuantificados a cero (véase  $\alpha_i$ , en EQN. 3),  $B[i]$  es el ancho de banda de factor de escala (véase  $B$ , en EQN. 3) y  $D()$  es la tabla de búsqueda Huffman del coste de transmisión de factor de escala (véase EQN. 3). El coste acumulado temporal es calculado y almacenado para todos posibles valores de los nodos de nodo de recorrido pasados,  $m$ , al presente nodo,  $k$ , son calculados, como se determina por el bloque de decisión 124, el coste mínimo se encuentra y se almacena en la matriz  $Cost2[k]$  en el bloque 126. También, el recorrido más barato en la etapa  $i$  y el nodo  $k$  es almacenado en la matriz  
50  $Historia[i][k]$  en el bloque 126.

Una vez que todos nodos presentes  $k$  en la etapa  $i$ , han sido procesados, como se determina por el bloque de decisión 118, la matriz  $Cost2[k]$  es copiada en la matriz  $Cost[k]$  en el bloque 120 en un bucle anidado  $i$  106 y el procedimiento se repite hasta que todas bandas de factor de escala han sido procesadas.  
55

Una vez que todas bandas han sido procesadas, como se determina por el bloque de decisión 114, la matriz  $Cost[k]$  contiene el coste acumulado para cada recorrido a través del entramado. El valor mínimo en la matriz  $Cost[k]$  es determinado por el bloque 134 y el indexto a ese valor ( $L$ ) identifica el nuevo valor de factor de escala ajustado para la última banda de factor de escala ( $i = N-1$ ). Un contador “ $i$ ” es disminuido repetidamente por un segundo bucle  $i$   
60 (no anidado), que empieza desde  $i = N-1$  por el bloque 140. La matriz  $Historia[i][k]$  se usa para rastrear hacia atrás a través del entramado para encontrar cada nodo anterior a lo largo del recorrido más barato conforme las bandas  $i$  de factor de escala van hacia atrás desde  $N-1$  a cero, identificando por tanto el valor óptimo de factor de escala de coste de bit para cada banda de factor de escala, que está dispuesto en la salida 146. Esto se consigue en el bucle 112 disminuyendo repetidamente  $i$  en el bloque 140 y determinando el valor óptimo  $k$  histórico de factor de escala para cada banda  $i$  de factor de escala en el bloque 142. El bloque 144 identifica el nuevo valor de factor de escala ajustado  
65 para cada banda de factor de escala sucesivo hacia atrás conforme  $i$  es disminuido desde  $N-1$  a cero.

## ES 2 312 852 T3

La figura 3 muestra el efecto de aplicar la optimización del factor de escala del presente invento a factores de escala preliminares derivados por medio de la técnica de estimación directa para una sola estructura de audio AAC. Los círculos representados en la figura 3 representan los factores de escala no ajustados; mientras que los puntos más dibujados representan factores de escala ajustados de acuerdo con una aplicación del presente invento. La técnica de optimización del factor de escala acorde con el presente invento reduce en gran medida la variación en los factores de escala. También los factores de escala ajustados son aumentados siempre, no solo ahorrando bits en general sino disminuyendo el ruido de cuantificación no solo en las bandas en la que los factores de escala son aumentados, sino también en otras bandas como un resultado de ahorros generales de bits (permitiendo así que más bits sean asignados a otras bandas). Los ahorros de bits conseguidos por esta técnica se muestran en la figura 4, que representa el coste de transmitir los factores de escala por estructura de un solo segmento de audio, tanto con como sin el uso de optimización acorde con el presente invento. La línea superior en la figura 4 es el coste de transmisión sin el uso del presente invento, mientras que la línea inferior muestra el coste de bits de transmisión con el uso del presente invento. De la figura 4, se verá que el coste de bits por estructura para la transmisión de los factores de escala es reducida en gran medida por el presente invento.

Debería entenderse que esta implementación de otras variaciones y modificaciones del invento y sus diversos aspectos serán claros para los expertos en la técnica, y que el invento no está limitado por estas realizaciones específicas descritas.

El presente invento y su diversos aspectos puede ser implementado como funciones de software realizadas en procesadores de señales digitales, computadoras digitales de uso general programadas, y/o computadores digitales de uso especial. Interfases entre flujos de señal digital y analógico pueden ser realizadas en equipo físico apropiado y/o como funciones en software yo firmware.

REIVINDICACIONES

5 1. Un método para reducir el coste total de bits de un codificador de audio perceptivo que emplea asignación adaptable de bits en la que una representación en el dominio del tiempo de una señal de audio es dividida en bloques de tiempo sucesivos, cada bloque de tiempo está dividido en bandas de frecuencia, y un factor de escala es asignado a cada una de las bandas de frecuencia, en la que el número de bits requerido para representar cada bloque aumenta con aumentos de los valores de factor de escala y con aumentos en variaciones banda a banda en valores de factor de escala, que comprende

10 determinar un factor de escala preliminar para dichos cada una de las bandas de frecuencia, y

15 optimizar los factores de escala para cada una de dichas bandas de frecuencia al aumentar el factor de escala a un valor mayor que el valor de factor de escala preliminar para una o más de las bandas de frecuencia si el aumento en coste de bits de dicho aumento es el mismo o menor que la reducción en coste de bits resultante de la disminución en variaciones banda a banda en valores de factor de escala resultante del aumento del factor de escala para una o más bandas de frecuencia, en las que para cada una de las bandas de frecuencia los factores de escala optimizados solo se les permite tomar valores mayores o iguales que los valores de factor de escala preliminares.

20 2. Un método acorde con la reivindicación 1, en el que dicha optimización incluye disminuir una función de coste de bits.

25 3. Un método acorde con la reivindicación 2, en el que dicha disminución disminuye el coste de bits de un recorrido a través de un entramado en el que sus nodos son los posibles valores de factor de escala en cada banda consecutiva de factor de escala.

4. Un método acorde con la reivindicación 3, en el que la disminución se realiza por un algoritmo de búsqueda Viterbi.

30 5. Un método acorde con cualquiera de las reivindicaciones 1-4, en el que el codificador de audio perceptivo Huffman codifica las diferencias entre los valores de factores de escala de bandas de frecuencia vecinas, en el que un aumento en variaciones banda a banda en valores de factor de escala aumenta el número de bits requeridos para la codificación Huffman.

35 6. Un método acorde con cualquiera de las reivindicaciones 1-5, en el que dicha derivación de un factor de escala preliminar para dichas bandas de frecuencia emplea al menos un bucle iterativo.

40 7. Un método acorde con la reivindicación 6, en el que dicho codificador de audio perceptivo genera un modelo de enmascarado, y dicha derivación emplea un bucle iterativo y calcula factores de escala basados en el modelo de enmascarado.

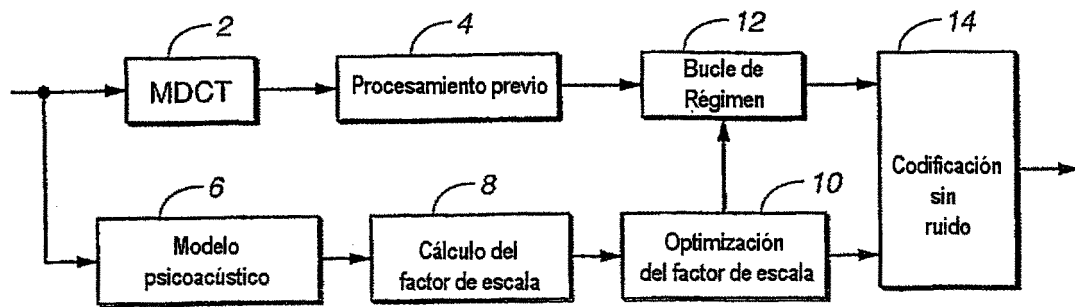
45

50

55

60

65



**FIG.\_1**

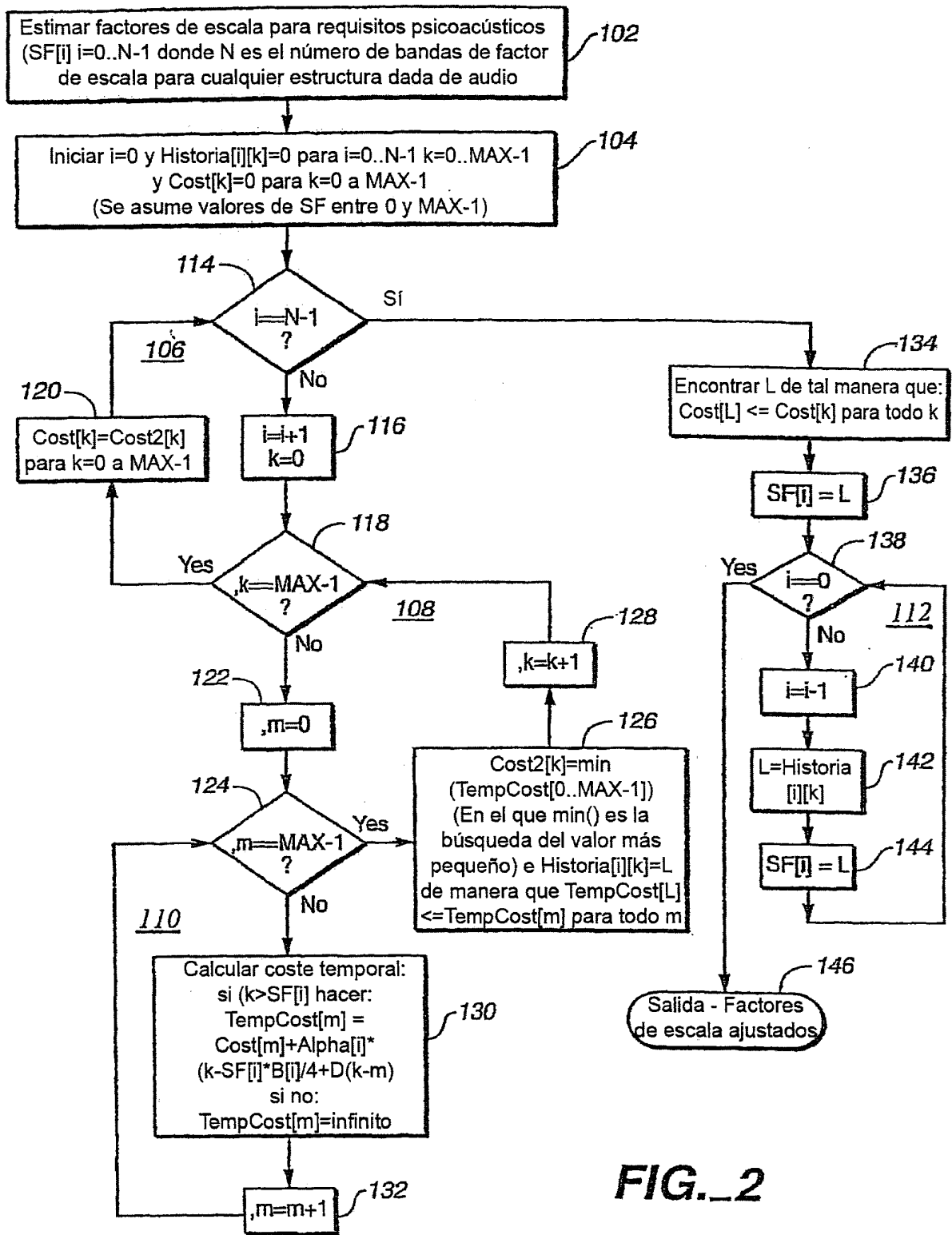
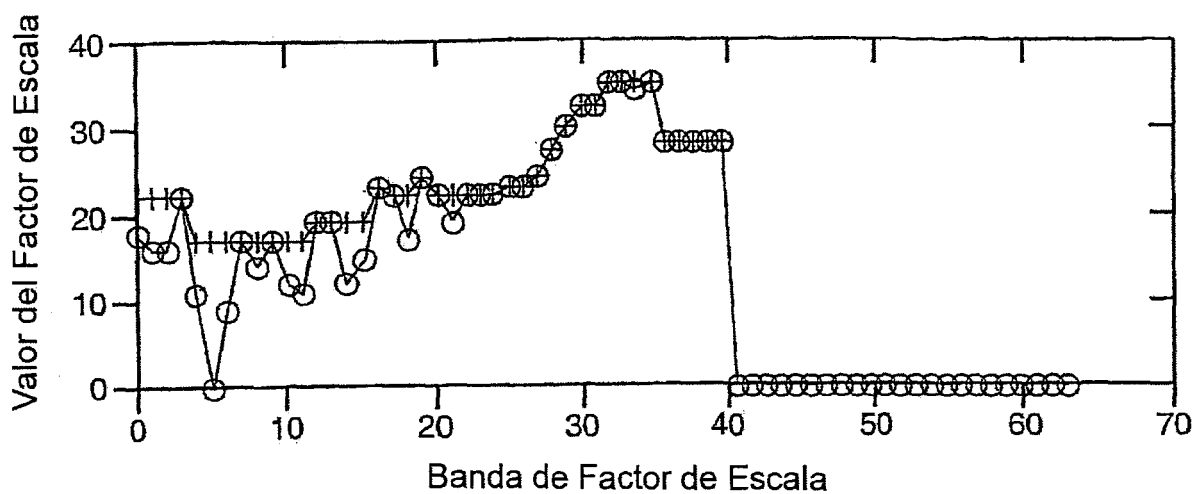
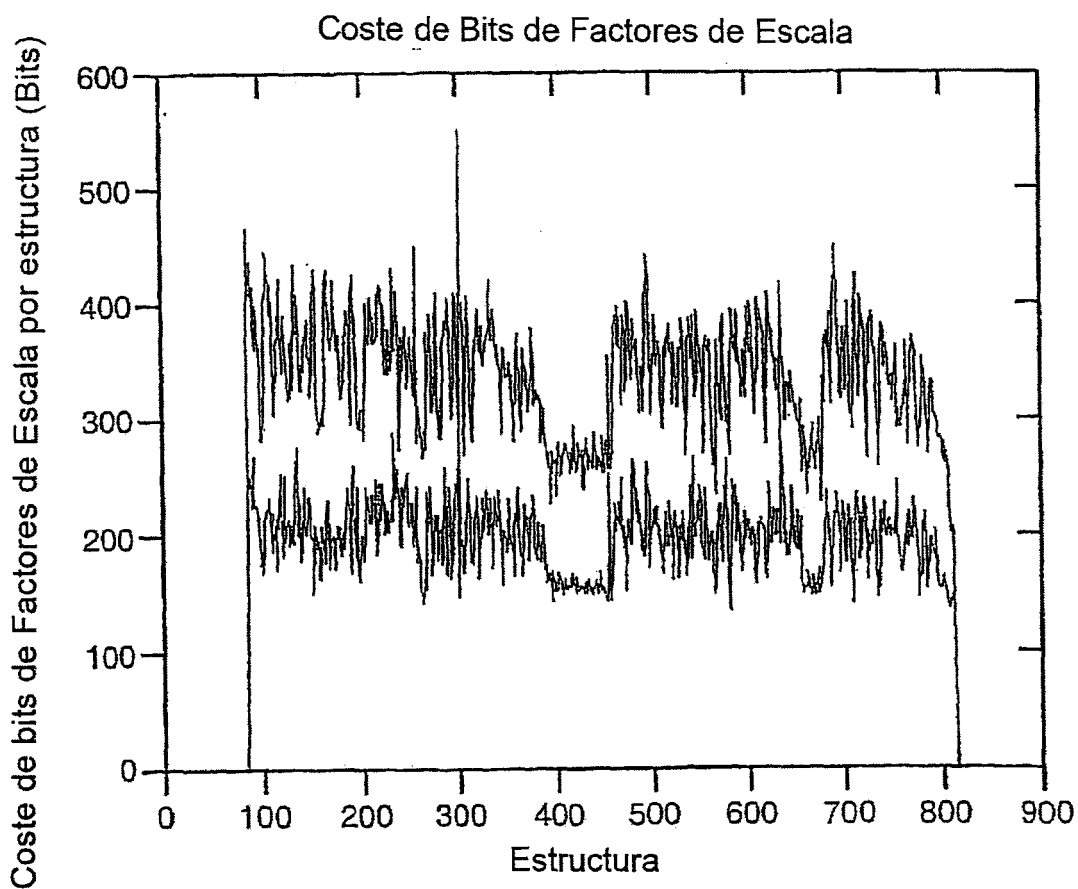


FIG. 2



**FIG.\_3**



**FIG.\_4**