

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 739 544**

51 Int. Cl.:

**G10L 19/025** (2013.01)

**G10L 19/002** (2013.01)

**H04N 19/115** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.02.2005 E 08021083 (4)**

97 Fecha y número de publicación de la concesión europea: **15.05.2019 EP 2034473**

54 Título: **Aparato y método de emisión de un valor estimado**

30 Prioridad:

**01.03.2004 DE 102004009949**

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

**31.01.2020**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.0%)  
Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**SCHUG, MICHAEL;  
HILPERT, JOHANNES;  
GEYERSBERGER, STEFAN y  
NEUENDORF, MAX**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

**ES 2 739 544 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

Aparato y método de emisión de un valor estimado

- 5 La presente invención se refiere a codificadores para codificar una señal incluyendo información de audio y/o de vídeo, y en particular a la estimación de una necesidad de unidades de información para codificar esta señal.

10 A continuación se presentará el codificador de la técnica anterior. En una entrada 1000 se alimenta una señal de audio que debe codificarse. Esta señal de audio se alimenta en primer lugar a una etapa 1002 de ajuste a escala, en la que se realiza un denominado control de amplificación AAC, para establecer el nivel de la señal de audio. La información secundaria del ajuste a escala se suministra a un formateador 1004 de flujo de bits, tal como se representa mediante la flecha ubicada entre el bloque 1002 y el bloque 1004. La señal de audio ajustada a escala se suministra después de esto a un banco 1006 de filtros MDCT. Con el codificador AAC, el banco de filtros implementa una transformación de coseno discreta modificada con un 50% de ventanas solapadas, determinándose la longitud de ventana mediante un bloque 1008.

15 Hablando en general, el bloque 1008 existe para que se aplique la función ventana a señales transitorias con ventanas relativamente cortas, y que se aplique la función ventana a señales más bien estacionarias con ventanas relativamente largas. Esto sirve para alcanzar un nivel mayor de resolución temporal (a costa de la resolución de frecuencia), para las señales transitorias debido a las ventanas relativamente cortas mientras que para señales más bien estacionarias se consiga una mayor resolución de frecuencia (a costa de la resolución temporal) por ventanas más largas, prefiriéndose según la tendencia ventanas más largas, dado que prometen una mayor ganancia de codificación. En la salida del banco 1006 de filtros se encuentran bloques sucesivos, considerado desde el punto de vista temporal, de valores espectrales, que según la realización del banco de filtros pueden ser coeficientes MDCT, coeficientes de Fourier o también señales de subbanda, teniendo cada señal de subbanda un determinado ancho de banda limitado, que se establece mediante el canal de subbandas correspondiente en el banco 1006 de filtros, y presentando cada señal de subbanda un determinado número de valores de muestreo de subbandas.

20 A continuación se explica a modo de ejemplo el caso, en el que el banco de filtros emite bloques sucesivos, considerado desde el punto de vista temporal, de coeficientes espectrales MDCT, que hablando en general, representan espectros de corta duración sucesivos de la señal de audio que debe codificarse en la entrada 1000. Un bloque de valores espectrales MDCT se alimenta entonces a un bloque 1010 de procesamiento TNS, en el que tiene lugar una conformación de ruido temporal (TNS = temporal noise shaping). La técnica TNS se usa para conformar la forma temporal del ruido de cuantificación dentro de cada ventana de la transformación. Esto se consigue aplicando un proceso de filtrado a partes de los datos espectrales de cada canal. La codificación se realiza en una base de ventana. En particular se realizan las siguientes etapas, para aplicar la herramienta de TNS a una ventana de datos espectrales, es decir a un bloque de valores espectrales.

25 En primer lugar se selecciona un intervalo de frecuencia para la herramienta de TNS. Una selección adecuada comprende cubrir un intervalo de frecuencia de desde 1,5 kHz hasta la máxima banda de factor de escala posible con un filtro. Se indica que este intervalo de frecuencia depende de la tasa de muestreo, tal como se especifica en la norma AAC (ISO/IEC 14496-3: 2001 (E)).

30 A continuación se realiza un cálculo LPC (LPC = linear predictive coding = codificación predictiva lineal), y concretamente con los coeficientes MDCT espectrales, que se encuentran en el intervalo de frecuencia objetivo seleccionado. Para una estabilidad aumentada se excluyen de este proceso los coeficientes que corresponden a frecuencias inferiores a 2,5 kHz. Los procedimientos LPC habituales, tal como se conocen del procesamiento de voz, pueden usarse para el cálculo LPC, por ejemplo, el algoritmo de Levinson-Durbin conocido. El cálculo se realiza para el orden máximo admisible del filtro de conformación de ruido.

35 Como resultado del cálculo LPC se obtiene la ganancia de predicción PG esperada. Además se obtienen los coeficientes de reflexión o coeficientes Parcor.

40 Si la ganancia de predicción no supera un determinado umbral, no se aplica la herramienta de TNS. En este caso se escribe una información de control en el flujo de bits, para que un decodificador sepa que no se ha realizado ningún procesamiento TNS.

45 Sin embargo, si la ganancia de predicción supera un umbral, se aplica el procesamiento TNS.

50 En una etapa siguiente se cuantifican los coeficientes de reflexión. El orden del filtro de conformación de ruido usado se determina mediante la eliminación de todos los coeficientes de reflexión que tienen un valor absoluto inferior a un umbral de la "cola" de la serie de coeficientes de reflexión. El número de los coeficientes de reflexión restantes se encuentra en el orden de magnitud del filtro de conformación de ruido. Un umbral adecuado es de 0,1.

Los coeficientes de reflexión restantes se convierten normalmente en coeficientes de predicción lineales, conociéndose esta técnica también como procedimiento "Step-Up".

Los coeficientes LPC calculados se usan entonces como coeficientes de filtro de conformación de ruido del codificador, es decir como coeficientes de filtro de predicción. Este filtro FIR se usa para filtrar en el intervalo de frecuencia objetivo especificado. En la decodificación se usa un filtro autorregresivo, mientras que durante la codificación se usa un denominado filtro promedio móvil. Finalmente se suministra la información secundaria para la herramienta de TNS al formateador de flujo de bits, tal como se representa mediante la flecha mostrada entre el bloque 1010 de procesamiento TNS y el formateador 1004 de flujo de bits en la figura 3.

Después de esto se hacen pasar a su través varias herramientas opcionales no mostradas en la figura 3, tales como, por ejemplo, una herramienta de predicción a largo plazo, una herramienta de intensidad/acoplamiento, una herramienta de predicción, una herramienta de sustitución de ruido, hasta que finalmente se llega a un codificador 1012 central/lateral. El codificador 1012 central/lateral está activo cuando la señal de audio que debe codificarse es una señal multicanal, es decir una señal estéreo con un canal izquierdo y un canal derecho. Hasta ahora, es decir en el sentido de procesamiento antes del bloque 1012 en la figura 3 se procesaron el canal estéreo izquierdo y el derecho independientemente entre sí, es decir se ajustaron a escala, se transformaron mediante el banco de filtros, se sometieron al procesamiento TNS o no, etc.

En el codificador central/lateral se comprueba entonces en primer lugar, si tiene sentido una codificación central/lateral, es decir si en realidad proporciona una ganancia de codificación. Una codificación central/lateral proporcionará entonces una ganancia de codificación, cuando el canal izquierdo y el derecho son más bien similares, dado que entonces el canal central, es decir la suma de los canales izquierdo y derecho es casi igual al canal izquierdo o al canal derecho, independientemente del ajuste a escala mediante un factor de 1/2, mientras que el canal lateral sólo tiene valores muy pequeños, dado que es igual a la diferencia entre los canales izquierdo y el derecho. Como consecuencia puede observarse que cuando los canales izquierdo y el derecho son prácticamente iguales, la diferencia es aproximadamente cero o incluye sólo valores muy pequeños, que se espera que se cuantifiquen como cero en un cuantificador 1014 posterior y por consiguiente puedan transmitirse de manera muy eficaz, dado que al cuantificador 1014 le sigue un codificador 1016 de entropía.

Al cuantificador 1014 se le suministra desde un modelo 1020 psicoacústico una interferencia permitida por banda de factor de escala. El cuantificador funciona de manera iterativa, es decir se llama en primer lugar a un bucle iterativo externo, que entonces llama a un bucle iterativo interno. Hablando en general, se realiza en primer lugar, partiendo de valores iniciales de anchos de etapa de cuantificación, una cuantificación de un bloque de valores en la entrada del cuantificador 1014. En particular, el bucle interno cuantifica los coeficientes MDCT, utilizándose un determinado número de bits en el proceso. El bucle externo calcula la distorsión y la energía modificada de los coeficientes usando el factor de escala, para llamar de nuevo a un bucle interno. Este proceso se itera, hasta que se cumple un determinado conjunto de condiciones. Para cada iteración en el bucle iterativo externo se reconstruye a este respecto la señal, para calcular la interferencia introducida por la cuantificación y compararla con la interferencia permitida proporcionada por el modelo 1020 psicoacústico. Además se aumentan los factores de escala de las bandas de frecuencia que tras esta comparación aún se consideran interferidas se amplían en una o más etapas desde interacción hasta interacción, y concretamente para cada iteración del bucle iterativo externo.

Entonces, cuando se alcanza una situación en la que la interferencia de cuantificación introducida por la cuantificación está por debajo de la interferencia permitida determinada mediante el modelo psicoacústico, y cuando al mismo tiempo se cumplen los requisitos de bits, concretamente, que no se supera una tasa de bits máxima, se finaliza la iteración, es decir el procedimiento de análisis por síntesis, y se codifican los factores de escala obtenidos, tal como se expone en el bloque 1014 y se suministra en forma codificada al formateador 1004 de flujo de bits, tal como se señala mediante la flecha, que está dibujada entre el bloque 1014 y el bloque 1004. Los valores cuantificados se suministran entonces al codificador 1016 de entropía, que realiza normalmente una codificación de entropía usando varias tablas de códigos Huffman para diferentes bandas de factor de escala, para transformar los valores cuantificados a un formato binario. Tal como se conoce, durante la codificación de entropía en forma de codificación de Huffman se recurre a tablas de códigos que se crean debido a una estadística de señales esperada, y en las que los valores que aparecen con mayor frecuencia reciben palabras de código más cortas que los valores que aparecen con menor frecuencia. Los valores codificados por entropía se suministran entonces también como información principal verdadera al formateador 1004 de flujo de bits, que emite entonces según una determinada sintaxis de flujo de bits en el lado de salida la señal de audio codificada.

La reducción de datos de señales de audio es entretanto una técnica conocida, que es el objetivo de una serie de normas internacionales (por ejemplo las normas ISO/MPEG-1, MPEG-2 AAC, MPEG-4).

Es común para los procedimientos mencionados anteriormente, que la señal de entrada se incorpore por medio de un denominado codificador aprovechando efectos relacionados con la percepción (psicoacústica, psicoóptica) en una reproducción compacta, con datos reducidos. Para ello se realiza habitualmente un análisis espectral de la señal

y las componentes de señal correspondientes se cuantifican teniendo en cuenta un modelo de percepción y a continuación se codifica de una manera lo más compacta posible como un denominado flujo de bits.

5 Para estimar antes de la verdadera cuantificación cuántos bits necesitará una determinada sección que debe codificarse de la señal, puede recurrirse a la denominada Perceptual Entropy (PE, entropía perceptual). La PE proporciona también una medida de cómo de difícil es para el codificador codificar una determinada señal o partes de la misma.

10 Para la calidad de la estimación es decisiva la desviación de la PE con respecto al número de bits necesarios en realidad.

15 Además puede recurrirse a la entropía perceptual o a cada valor estimado de una necesidad de unidades de información para codificar una señal para estimar si la señal es transitoria o estacionaria, dado que las señales transitorias también requieren más bits para la codificación que las señales más bien estacionarias. La estimación de una propiedad transitoria de una señal se usa, por ejemplo, para realizar una decisión de longitud de ventana, tal como se indica en el bloque 1008 en la figura 3.

20 En la figura 6 se ilustra la entropía perceptual calculada según la norma ISO/IEC IS 13818-7 (MPEG-2 advanced audio coding (AAC)). Para calcular esta entropía perceptual, es decir una entropía perceptual por bandas, se usa la ecuación representada en la figura 6. En esta ecuación el parámetro  $pe$  representa la entropía perceptual. Además  $width(b)$  representa el número de coeficientes espectrales en la respectiva banda  $b$ . Además  $e(b)$  es la energía de la señal en esta banda. Finalmente  $nb(b)$  es el umbral de enmascaramiento adecuado para ello o expresado de manera general, la interferencia permitida que puede introducirse en la señal, por ejemplo, mediante una cuantificación, para que aun así un oyente humano no oiga ninguna o sólo una interferencia reducida que se desvanece.

30 Las bandas pueden proceder de la clasificación de bandas del modelo psicoacústico (bloque 1020 en la figura 3), o pueden ser las denominadas bandas de factor de escala ( $safb$ ) usadas durante la cuantificación. El umbral de enmascaramiento psicoacústico es el valor de energía que no debe superar el error de cuantificación.

35 La ilustración mostrada en la figura 6 muestra por tanto cómo de bien funciona una entropía perceptual así determinada como estimación para el número de bits necesarios para la codificación. Para ese fin, en el ejemplo de un codificador AAC, se aplicó a diferentes tasas de bits para cada bloque individual la respectiva entropía perceptual en función de los bits utilizados. La pieza de prueba usada contiene una mezcla típica de música, voz e instrumentos individuales.

De manera ideal los puntos se acumularían a lo largo de una línea recta a través del punto cero. La extensión de la sucesión de puntos con las desviaciones con respecto a la línea ideal ilustra la estimación poco precisa.

40 Por tanto, en el concepto mostrado en la figura 6 es desventajosa la desviación que se manifiesta en el sentido de que, por ejemplo, resulte un valor demasiado elevado para la entropía perceptual, lo que significa a su vez que se señala al cuantificador, que se necesitan más bits que los requeridos en principio. Esto conduce al hecho de que el cuantificador realice una cuantificación demasiado fina, a que no aproveche por tanto la medida de interferencia permitida, lo que da como resultado una ganancia de codificación reducida. Por otro lado, cuando el valor para la entropía perceptual se determina como demasiado pequeño, entonces se señala al cuantificador, que se necesitan menos bits que los requeridos en principio, para la codificación de la señal. Esto tiene a su vez como consecuencia que el cuantificador realice una cuantificación demasiado aproximada, lo que conduciría directamente a una interferencia audible en la señal, siempre que no se tomen contramedidas. Las contramedidas pueden consistir en que el cuantificador necesite además uno o varios bucles iterativos, lo que hace aumentar el tiempo de cálculo del codificador.

55 Para mejorar el cálculo de la entropía perceptual podría introducirse, tal como se muestra en la figura 7, un término constante, tal como, por ejemplo, 1,5, en la expresión logarítmica. Entonces resulta ya un resultado mejorado, es decir una menor desviación hacia arriba o hacia abajo, aunque aún así puede observarse que teniendo en cuenta un término constante en la expresión logarítmica se reduce concretamente el caso de que la entropía perceptual señalice una necesidad demasiado optimista de bits. Sin embargo, por otro lado, a partir de la figura 7 puede reconocerse claramente que de manera significativa se señala un número demasiado elevado de bits, lo que conduce a que el cuantificador siempre cuantifique de manera demasiado fina, que por tanto se asuma la necesidad de bits mayor que lo que es en verdad, lo que a su vez da como resultado una ganancia de codificación reducida. La constante en la expresión logarítmica es una estimación aproximada de los bits necesarios para la información secundaria.

Así la inserción de un término en la expresión logarítmica proporciona una mejora de la entropía perceptual por bandas, tal como se ilustra en la figura 6, dado que las bandas con una separación muy reducida entre la energía y

el umbral de enmascaramiento se tienen en cuenta antes, dado que también para la transmisión de coeficientes espectrales que deben cuantificarse como cero es necesario un cierto número de bits.

5 En la figura 8 se ilustra un cálculo adicional, que no obstante requiere mucho tiempo de cálculo, de la entropía perceptual. En la figura 8 se muestra el caso en el que la entropía perceptual se calcula por líneas. Sin embargo, la desventaja radica en el mayor esfuerzo de cálculo del cálculo por líneas. En este caso en lugar de la energía se utilizan coeficientes espectrales  $X(k)$ , designando  $k_{\text{Offset}}(b)$  al primer índice de la banda  $b$ . Cuando la figura 8 se compara con la figura 7, puede observarse claramente en el intervalo entre 2000 y 3000 bits una reducción de las “desviaciones” hacia arriba. La estimación de PE será por tanto más precisa, es decir no se realizará una estimación demasiado pesimista, sino que se encontraría más bien en el óptimo, de modo que la ganancia de codificación puede aumentar en comparación con el procedimiento de cálculo mostrado en las figuras 6 y 7, o se reduce el número de iteraciones en el cuantificador.

10 Sin embargo, en el cálculo por líneas de la entropía perceptual es desventajoso el tiempo de cálculo necesario para obtener el valor de la ecuación mostrada en la figura 8.

15 Tales desventajas de tiempo de cálculo no desempeñan obligatoriamente ningún papel cuando el codificador se ejecuta en un PC potente o en una estación de trabajo potente. Por el contrario, resulta muy distinto cuando el codificador está ubicado en un dispositivo portátil, tal como, por ejemplo, un teléfono móvil UMTS, que por un lado tiene que ser pequeño y barato, que por otro lado debe tener un bajo consumo de energía, y que además tiene que trabajar rápidamente para permitir la codificación de una señal de vídeo o señal de audio transmitida a través de la conexión UMTS.

20 El documento US 2002/103637 A1 da a conocer un concepto para mejorar la capacidad de sistemas de codificación para utilizar los procedimientos de reconstrucción de alta frecuencia. Para ese fin se calcula por parte del codificador una dificultad de codificación o una medida para la carga de trabajo de un codificador, para controlar en función de esto la frecuencia de cruce, que determina hasta qué frecuencia se codifica una señal con un codificador de origen, codificándose la porción de la señal que se encuentra por encima de la frecuencia de cruce mediante un procedimiento de reconstrucción de alta frecuencia. Como medida de la dificultad para codificar una señal se calcula la entropía perceptual, que se basa en que se eleva al cuadrado un valor espectral y entonces se pondera con un número que es igual al número de líneas en la banda actual dividido entre el umbral psicoacústico para esta banda, para formar entonces a partir de dicho resultado un logaritmo. La suma de todos los logaritmos de este tipo en una banda da entonces como resultado la entropía perceptual en esa banda. Alternativamente a esto puede calcularse también una energía de distorsión al final del procedimiento de codificación de origen, sumándose la energía de distorsión en cada banda y ponderándose con una curva de intensidad del sonido.

25 El objetivo de la presente invención es proporcionar un concepto eficaz y aún así preciso para determinar un valor estimado de una necesidad de unidades de información para codificar una señal.

30 Este objetivo se logra mediante un aparato según la reivindicación 1, un procedimiento según la reivindicación 5 o un programa informático según la reivindicación 6.

35 La presente invención se basa en el hallazgo de que en un cálculo por bandas de frecuencia del valor estimado de una necesidad de unidades de información debe establecerse por motivos del tiempo de cálculo, que sin embargo, para obtener una determinación precisa del valor estimado, debe considerarse la distribución de la energía en la banda de frecuencia, que debe calcularse por bandas.

40 Con ello se “implica” en cierto modo de manera implícita al codificador de entropía que sigue al cuantificador en la determinación del valor estimado para la necesidad de unidades de información. La codificación de entropía permite concretamente, que para la transmisión de valores espectrales menores se necesita un menor número de bits que para la transmisión de valores espectrales mayores. El codificador de entropía es especialmente eficaz cuando pueden transmitirse valores espectrales cuantificados como cero. Dado que éstos aparecen normalmente con la mayor frecuencia, la palabra de código para transmitir una línea espectral cuantificada como cero es la palabra de código más corta, y la palabra de código para transmitir una línea espectral cuantificada cada vez mayor es cada vez más larga. Además para un concepto especialmente eficaz para transmitir una secuencia de valores espectrales cuantificados como cero puede recurrirse incluso a una codificación de longitud de recorrido, lo que tiene como consecuencia que en el caso de un recorrido de ceros por valor espectral cuantificado como cero no se necesita en promedio ni siquiera un solo bit.

45 Se ha encontrado que el cálculo de entropía perceptual por bandas usado en el estado de la técnica para determinar el valor estimado de la necesidad de unidades de información ignora completamente el modo de actuación del codificador de entropía dispuesto aguas abajo, cuando la distribución de la energía en la banda de frecuencia desvía de una distribución completamente uniforme.

Según la invención se tienen en cuenta por consiguiente para reducir las imprecisiones del cálculo por bandas cómo está distribuida la energía dentro de una banda.

5 Dependiendo la implementación puede determinarse la medida de la distribución de la energía en la banda de frecuencia basándose en amplitudes verdaderas, o mediante una estimación de las líneas de frecuencia, que no se cuantifican como cero mediante el cuantificador. Esta medida, que también se designa como "nl", representando nl "number of active lines", es decir el número de líneas activas, se prefiere por motivos de eficacia del tiempo de cálculo. Sin embargo, también puede tenerse en cuenta el número de líneas espectrales que deben cuantificarse como cero o una subdivisión más fina, siendo esta estimación más precisa cuanto más información del codificador de entropía dispuesto aguas abajo se tenga en cuenta. Si el codificador de entropía se basa en tablas de códigos de Huffman, entonces pueden integrarse de manera especialmente buena las propiedades de estas tablas de códigos, dado que las tablas de códigos no se calculan en cierto sentido en línea debido a la estadística de señales, sino dado que las tablas de códigos están fijadas de todas maneras independientemente de la verdadera señal.

15 Sin embargo, dependiendo de las limitaciones del tiempo de cálculo, en el caso de un cálculo especialmente eficaz se realiza la medida de la distribución de la energía en la banda de frecuencia mediante la determinación de las líneas aún restantes tras la cuantificación, es decir el número de líneas activas.

20 La presente invención es ventajosa en el sentido de que se determina un valor estimado de una necesidad de contenido de información, que por un lado es más preciso y por otro lado es más eficaz que el del estado de la técnica.

25 Además, la presente invención puede ajustarse a escala para diversas aplicaciones, dado que según la precisión deseada del valor estimado pueden incorporarse cada vez más propiedades del codificador de entropía, aunque a costa de un tiempo de cálculo aumentado, en la estimación de la necesidad de bits.

A continuación se explican detalladamente realizaciones preferidas de la presente invención haciendo referencia a los tiempos adjuntos, donde:

30 la figura 1 es un diagrama de circuito de bloques del aparato de la invención para determinar un valor estimado;

la figura 2 muestra una realización preferida de los medios para calcular una medida de la distribución de la energía en la banda de frecuencia;

35 la figura 2b muestra una realización preferida de los medios para calcular el valor estimado de la necesidad de bits;

la figura 3 es un diagrama de circuito de bloques de un codificador de audio conocido;

40 la figura 4 es una representación esquemática para explicar la influencia de la distribución de energía dentro de una banda sobre la determinación del valor estimado;

la figura 5 es un diagrama para el cálculo del valor estimado según la presente invención;

45 la figura 6 es un diagrama para el cálculo del valor estimado según la norma ISO/IEC IS 13818-7(AAC);

la figura 7 es un diagrama para el cálculo del valor estimado con término constante;

la figura 8 es un diagrama para el cálculo lineal del valor estimado con término constante.

50 A continuación con referencia a la figura 1 se ilustrará el aparato de la invención para determinar un valor estimado de una necesidad de unidades de información para codificar una señal. La señal, que puede ser una señal de audio y/o de vídeo, se alimenta a través de una entrada 100. Preferiblemente la señal ya se encuentra como representación espectral con valores espectrales. Sin embargo, esto no es obligatoriamente necesario, dado que mediante un filtrado correspondiente, por ejemplo paso banda, también pueden realizarse algunos cálculos con una señal de tiempo.

60 La señal se suministra a medios 102 para proporcionar una medida de una interferencia permitida para una banda de frecuencia de la señal. La interferencia permitida puede determinarse, por ejemplo, por medio de un modelo psicoacústico, tal como se ha explicado mediante la figura 3 (bloque 1020). Los medios 102 son además eficaces para proporcionar también una medida de la energía de la señal en la banda de frecuencia. La condición previa para un cálculo por bandas es que una banda de frecuencia, para la que se indica una interferencia permitida o una energía de señal, contenga al menos dos o más líneas espectrales de la representación espectral de la señal. En el caso de codificadores de audio normalizados típicos, la banda de frecuencia será preferiblemente una banda de factor de escala, dado que es necesaria la estimación de necesidad de bits directamente del cuantificador, para

establecer si una cuantificación realizada cumple o no un criterio de bits.

Los medios 102 están formados para suministrar tanto la interferencia permitida  $nb(b)$ , como la energía de señal  $e(b)$  de la señal en la banda a medios 104 para calcular el valor estimado de la necesidad de bits.

5 Según la invención los medios 104 para calcular el valor estimado de la necesidad de bits están formados para, además de la interferencia permitida y de la energía de señal, tener en cuenta una medida  $nl(b)$  de una distribución de la energía en la banda de frecuencia, desviando la distribución de la energía en la banda de frecuencia de una distribución completamente uniforme. La medida de la distribución de la energía se calcula en medios 106, necesitando los medios 106 al menos una banda, concretamente la banda de frecuencia considerada de la señal de audio o de vídeo o bien como señal paso banda o bien directamente como consecuencia de líneas espectrales, para poder realizar, por ejemplo, un análisis espectral de la banda, para obtener la medida de la distribución de las energías en la banda de frecuencia.

15 Naturalmente la señal de audio o de vídeo puede suministrarse a los medios 106 como señal de tiempo, realizando entonces los medios 106 un filtrado de banda así como un análisis en la banda. Como una alternativa la señal de audio o de vídeo suministrada a los medios 106, puede encontrarse ya en el intervalo de frecuencia, tal como, por ejemplo, como coeficiente MDCT, o si no también como señal paso banda en el banco de filtros con un número de filtros paso banda menor en comparación con un banco de filtros MDCT.

20 En una realización preferida los medios 106 para calcular están formados para tener en cuenta para calcular el valor estimado magnitudes actuales de valores espectrales en la banda de frecuencia.

25 Además los medios para calcular la medida de la distribución de la energía pueden estar formados para determinar como medida de la distribución de la energía varios valores espectrales, cuya magnitud es mayor o igual que un umbral de magnitud predeterminado, o cuya magnitud es menor o igual que el umbral de magnitud, siendo el umbral de magnitud preferiblemente una etapa de cuantificación estimada, que en un cuantificador hace que los valores menores o iguales a la etapa de cuantificación se cuantifiquen como cero. En este caso la medida de la energía es el número de líneas activas, es decir el número de líneas, que perduran tras la cuantificación o que no son iguales a cero.

30 La figura 2a muestra una realización preferida de los medios 106 para calcular la medida de la distribución de la energía en la banda de frecuencia. La medida de la distribución de la energía en la banda de frecuencia se designa en la figura 2a con  $nl(b)$ . El factor de forma  $ffac(b)$  es ya una medida de la distribución de la energía en la banda de frecuencia. Tal como puede observarse a partir del bloque 106, la medida de la distribución espectral  $nl$  se determina a partir del factor de forma  $ffac(b)$  mediante la ponderación con la raíz a la cuarta de la energía de señal  $e(b)$  dividido entre el ancho de banda  $width(b)$  o el número de líneas en la banda de factor de escala  $b$ . En este contexto debe especificarse que el factor de forma también es un ejemplo de una magnitud que indica una medida de la distribución de las energías, mientras que  $nl(b)$  es, a diferencia de esto, un ejemplo de una magnitud que representa un valor estimado del número de líneas relevantes para la cuantificación.

45 El factor de forma  $ffac(b)$  se calcula mediante la formación de magnitudes de una línea espectral y la formación de raíz siguiente de esta línea espectral y la suma siguiente de las magnitudes "a las que se ha aplicado la raíz" de las líneas espectrales en la banda.

La figura 2b muestra una realización preferida de los medios 104 para calcular el valor estimado  $pe$ , habiéndose introducido en la figura 2b además una diferenciación de caso, concretamente cuando el logaritmo de base 2 de la relación de la energía con respecto a la interferencia permitida es mayor que un factor constante  $c1$  o igual al factor constante. En este caso se toma la alternativa que se encuentra en la parte superior en el bloque 104, es decir la medida de la distribución espectral  $nl$  se multiplica por la expresión logarítmica.

50 Por otro lado se determina que el logaritmo de base 2 de la relación de las energías de señal con respecto a la interferencia permitida es menor que el valor  $c1$ , entonces se usa la alternativa inferior en el bloque 104 de la figura 2b, que presenta adicionalmente además una constante aditiva  $c2$  así como una constante multiplicativa  $c3$ , que se calculan a partir de las constantes  $c2$  y  $c1$ .

60 A continuación se explica mediante la figura 4a y la figura 4b el concepto de la invención. Así, la figura 4a muestra una banda, en la que hay cuatro líneas espectrales, que tienen todas el mismo tamaño. La energía en esta banda está distribuida por consiguiente de manera uniforme por la banda. Por el contrario, la figura 4b muestra una situación en la que la energía en la banda reside en una línea espectral, mientras que las otras tres líneas espectrales son iguales a cero. La banda mostrada en la figura 4b podría, por ejemplo, existir antes de la cuantificación, o podría obtenerse tras la cuantificación, cuando las líneas espectrales ajustadas a cero en la figura 4b antes de la cuantificación son menores que la primera etapa de cuantificación y por consiguiente se ajustan a cero por el cuantificador, es decir no "perduran".

5 El número de líneas activas en la figura 4b es por consiguiente igual a 1, calculándose el parámetro  $n_l$  en la figura 4b como la raíz cuadrada de 2. Por el contrario, el valor  $n_l$ , es decir la medida para la distribución espectral de la energía en la figura 4a se calcula como 4. Esto significa que la distribución espectral de la energía es más uniforme cuando la medida de la distribución de la energía espectral es mayor.

10 Debe indicarse que el cálculo por bandas de la entropía perceptual según el estado de la técnica no establece ninguna diferencia entre ambos casos. En particular no se establece ninguna diferencia cuando en ambas bandas, que se muestran en las figuras 4a y 4b, hay la misma energía.

15 Sin embargo, evidentemente el caso mostrado en la figura 4b puede codificarse con sólo una línea relevante con menos bits, dado que las tres líneas espectrales ajustadas a cero pueden transmitirse de manera muy eficaz. Hablando en general la capacidad de cuantificación más sencilla del caso mostrado en la figura 4b se basa en el hecho de que tras la cuantificación y la codificación sin pérdida los valores más pequeños y en particular los valores cuantificados como cero necesitan menos bits para su transmisión.

20 Según la invención se tiene por consiguiente en cuenta cómo la energía está distribuida dentro de la banda. Esto tiene lugar, tal como se ha expuesto, mediante la sustitución del número de líneas por banda en la ecuación conocida (figura 6) por una estimación del número de líneas que son distintas de cero tras la cuantificación. Esta estimación se muestra en la figura 2a.

25 Además debe indicarse que el factor de forma mostrado en la figura 2a también se necesita en otro punto en el codificador, por ejemplo, dentro del bloque 1014 de cuantificación para determinar el ancho de etapa de cuantificación. Entonces, cuando el factor de forma ya se calcula en otro punto, no tiene que volver a calcularse para la estimación de bits, de modo que el concepto según la invención para una estimación mejorada de la medida de los bits necesarios tiene suficiente con un mínimo de esfuerzo de cálculo adicional.

30 Tal como ya se ha expuesto, en el caso de  $X(k)$  se trata de los coeficientes espectrales que deben cuantificarse posteriormente, mientras que la variable  $k_{\text{Offset}}(b)$  designa el primer índice en la banda  $b$ .

Tal como puede observarse a partir de las figuras 4a y 4b, el espectro en la figura 4a da como resultado un valor  $n_l=4$ , mientras que el espectro en la figura 4b da como resultado un valor de 1,41. Por tanto con la ayuda del factor de forma se dispone una medida para la cuantificar la estructura de campo espectral dentro de la banda.

35 La nueva fórmula para calcular una entropía perceptual por bandas mejorada se basa por consiguiente en la multiplicación de la medida de la distribución espectral de la energía y de la expresión logarítmica, al aparecer la energía de señal  $e(b)$  en el numerador y la interferencia permitida en el denominador, pudiendo utilizarse según sea necesario un término dentro del logaritmo, tal como se ilustra ya en la figura 7. Este término puede ser, por ejemplo, también 1,5, sin embargo también puede ser igual a cero, tal como en el caso mostrado en la figura 2b, pudiendo determinarse éste por ejemplo empíricamente.

40 En este punto debe remitirse de nuevo a la figura 5, a partir de la que resulta evidente la entropía perceptual calculada según la invención, y concretamente indicada a través de los bits necesarios. Puede reconocerse claramente una mayor precisión de la estimación con respecto a los ejemplos comparativos en las figuras 6, 7 y 8. También con respecto al cálculo por líneas, el cálculo por bandas modificado según la invención tiene al menos un resultado equivalente.

45 Dependiendo de las circunstancias, el procedimiento según la invención puede implementarse en hardware o en software. La implementación puede tener lugar en un medio de almacenamiento digital, en particular en un disquete o CD con señales de control legibles electrónicamente, que pueden actuar conjuntamente con un sistema informático programable de manera que se realice el procedimiento. La invención consiste por tanto en general también en un producto de programa informático con un código de programa almacenado en un soporte legible por máquina para la realización del procedimiento según la invención, cuando el producto de programa informático se ejecuta en un ordenador. Dicho de otro modo, la invención puede realizarse por tanto como un programa informático con un código de programa para la realización del procedimiento, cuando el programa informático se ejecuta en un ordenador.



**REIVINDICACIONES**

1. Aparato para determinar un valor estimado (pe) de una necesidad de unidades de información para codificar una señal que presenta información de audio o de vídeo, presentando la señal varias bandas de frecuencia, que comprende:

medios (102) para proporcionar una medida (nb(b)) de una interferencia permitida para una banda de frecuencia (b) de la señal, incluyendo la banda de frecuencia (b) al menos dos valores espectrales de una representación espectral de la señal, y para proporcionar una medida de una energía (e(b)) de la señal en la banda de frecuencia (b);

caracterizado por

medios (106) para calcular una medida (nl(b)) de una distribución de la energía (e(b)) en la banda de frecuencia (b), desviando la distribución de la energía (e(b)) en la banda de frecuencia (b) de una distribución completamente uniforme,

donde los medios (106) para calcular la medida (nl(b)) para la distribución de la energía (e(b)) se forma para calcular la medida (nl(b)) de la distribución de la energía (e(b)) según las siguientes ecuaciones:

$$nl(b) = \frac{ffac(b)}{\left(\frac{e(b)}{width(b)}\right)^{0.25}}$$

$$ffac(b) = \sum_{k=kOffset(b)}^{kOffset(b+1)-1} \sqrt{|X(k)|},$$

donde X(k) es un valor espectral a un índice de frecuencia k, donde kOffset es un primer valor espectral en la frecuencia b, donde ffac(b) es un factor de forma, donde nl(b) representa la medida de la distribución de la energía (e(b)) en la banda de frecuencia b, donde e(b) es la energía en la banda de frecuencia b, y donde width(b) es un ancho de la banda de frecuencia; y

medios (104) para calcular el valor estimado (pe) mientras se usa la medida (nb(b)) para la interferencia permitida, la medida para la energía (e(b)), y la medida (nl(b)) para la distribución de la energía (e(b)).

2. Aparato según una de las reivindicaciones anteriores,

en el que los medios (104) para calcular el valor estimado se forma para usar un cociente de la energía (e(b)) en la banda de frecuencia (b) y la interferencia permitida en la banda de frecuencia (b).

3. Aparato según una de las reivindicaciones anteriores,

en el que los medios (104) para calcular el valor estimado se forma para calcular el valor estimado usando la siguiente expresión:

$$pe = \sum_b nl(b) \cdot \log_2 \left( \frac{e(b)}{nb(b)} + s \right)$$

donde pe es el valor estimado, donde nl(b) representa la medida de la distribución de la energía (e(b)) en la banda de frecuencia b, donde e(b) es la energía de la señal en la banda de frecuencia b, donde nb(b) es la interferencia permitida en la banda de frecuencia b, y donde s es un término aditivo, preferiblemente igual a 1,5.

4. Aparato según una de las reivindicaciones anteriores,

en el que la señal se facilita como representación espectral con valores espectrales.

5. Procedimiento para determinar un valor estimado de una necesidad de unidades de información para codificar una señal, que presenta información de audio o de vídeo, presentando la señal varias bandas de

frecuencia, comprendiendo las etapas de:

proporcionar (102) una medida (nb(b)) de una interferencia permitida para una banda de frecuencia (b) de la señal, incluyendo la banda de frecuencia (b) al menos dos valores espectrales de una representación espectral de la señal, y

proporcionar una medida de una energía (e(b)) de la señal en la banda de frecuencia (b);

caracterizado por

calcular (106) una medida (nl(b)) de una distribución de la energía (e(b)) en la banda de frecuencia (b), desviando la distribución de la energía (e(b)) en la banda de frecuencia (b) de una distribución completamente uniforme,

donde la medida (nl(b)) de la distribución de la energía (e(b)) se calcula según las siguientes ecuaciones:

$$nl(b) = \frac{ffac(b)}{\left(\frac{e(b)}{width(b)}\right)^{0.25}}$$

$$ffac(b) = \sum_{k=kOffset(b)}^{kOffset(b+1)-1} \sqrt{|X(k)|},$$

donde X(k) es un valor espectral a un índice de frecuencia k, donde kOffset es un primer valor espectral en la banda de frecuencia b, donde ffac(b) es un factor de forma, donde nl(b) representa la medida de la distribución de la energía (e(b)) en la banda de frecuencia b, donde e(b) es la energía en la banda de frecuencia b, y donde width(b) es un ancho de la banda de frecuencia; y

calcular el valor estimado (pe) mientras se usa la medida (nb(b)) para la interferencia permitida, la medida (e(b)) para la energía (e(b)), y la medida (nl(b)) para la distribución de la energía (e(b)).

6. Programa informático con un código de programa para realizar el procedimiento para determinar un valor estimado de una necesidad de unidades de información para codificar una señal según la reivindicación 5, cuando el programa se ejecuta en un ordenador.

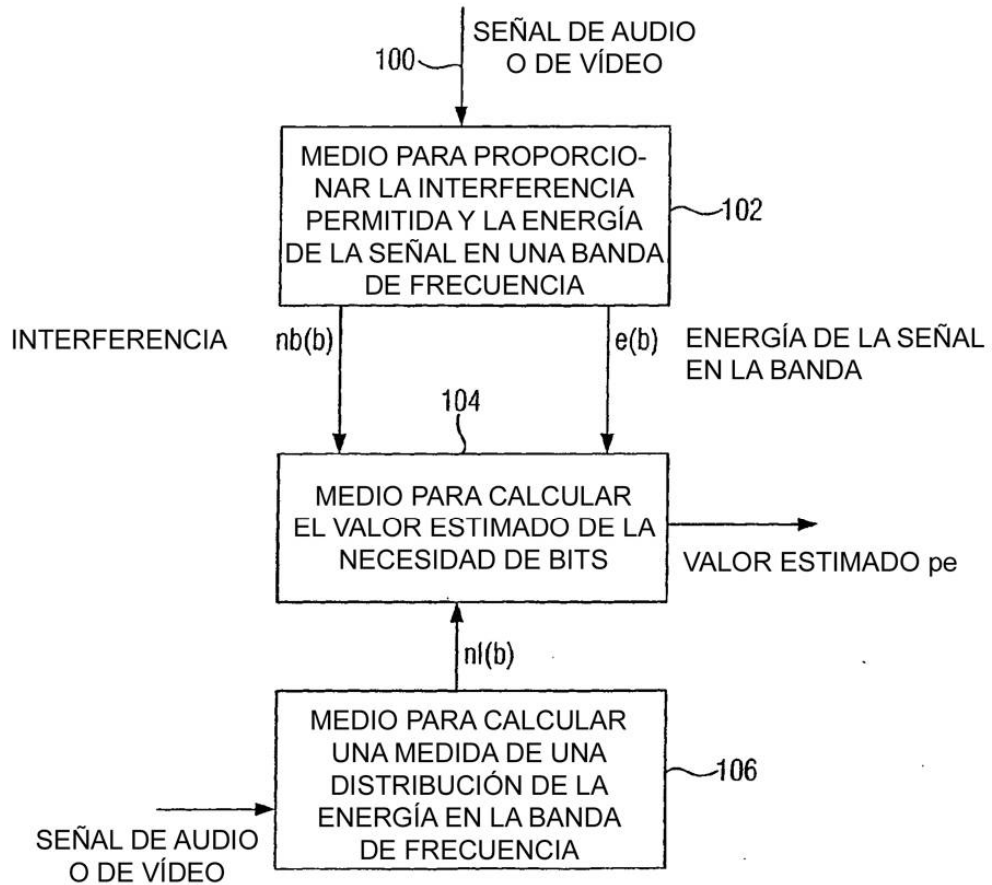


FIG. 1

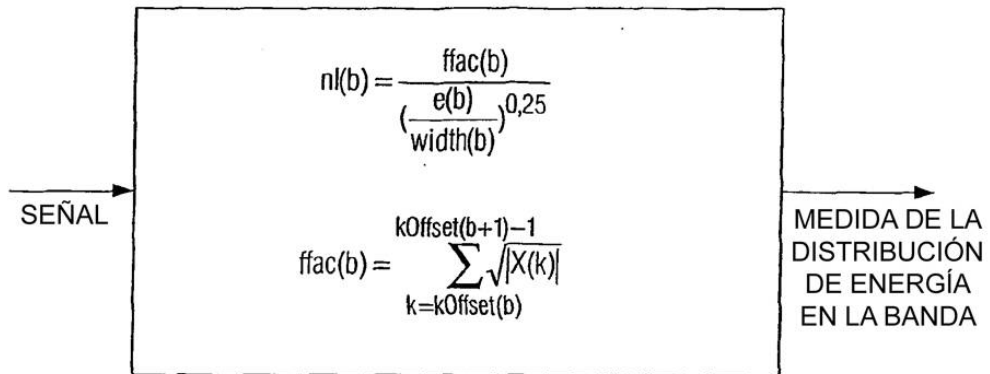


FIG. 2a

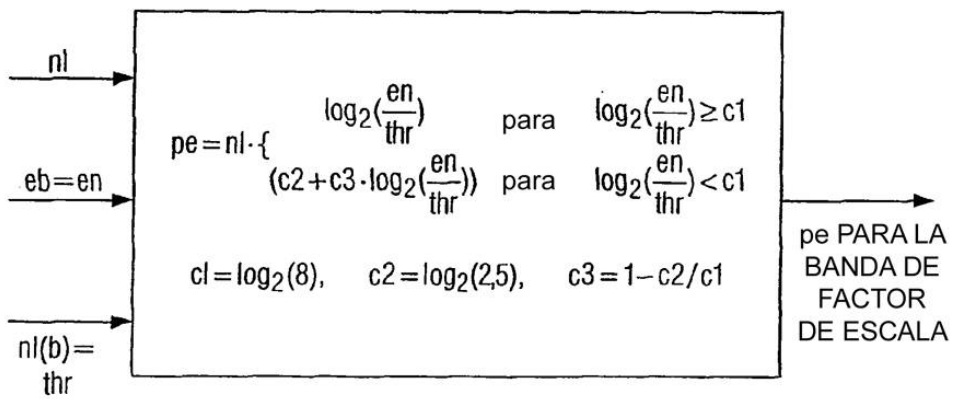


FIG. 2b

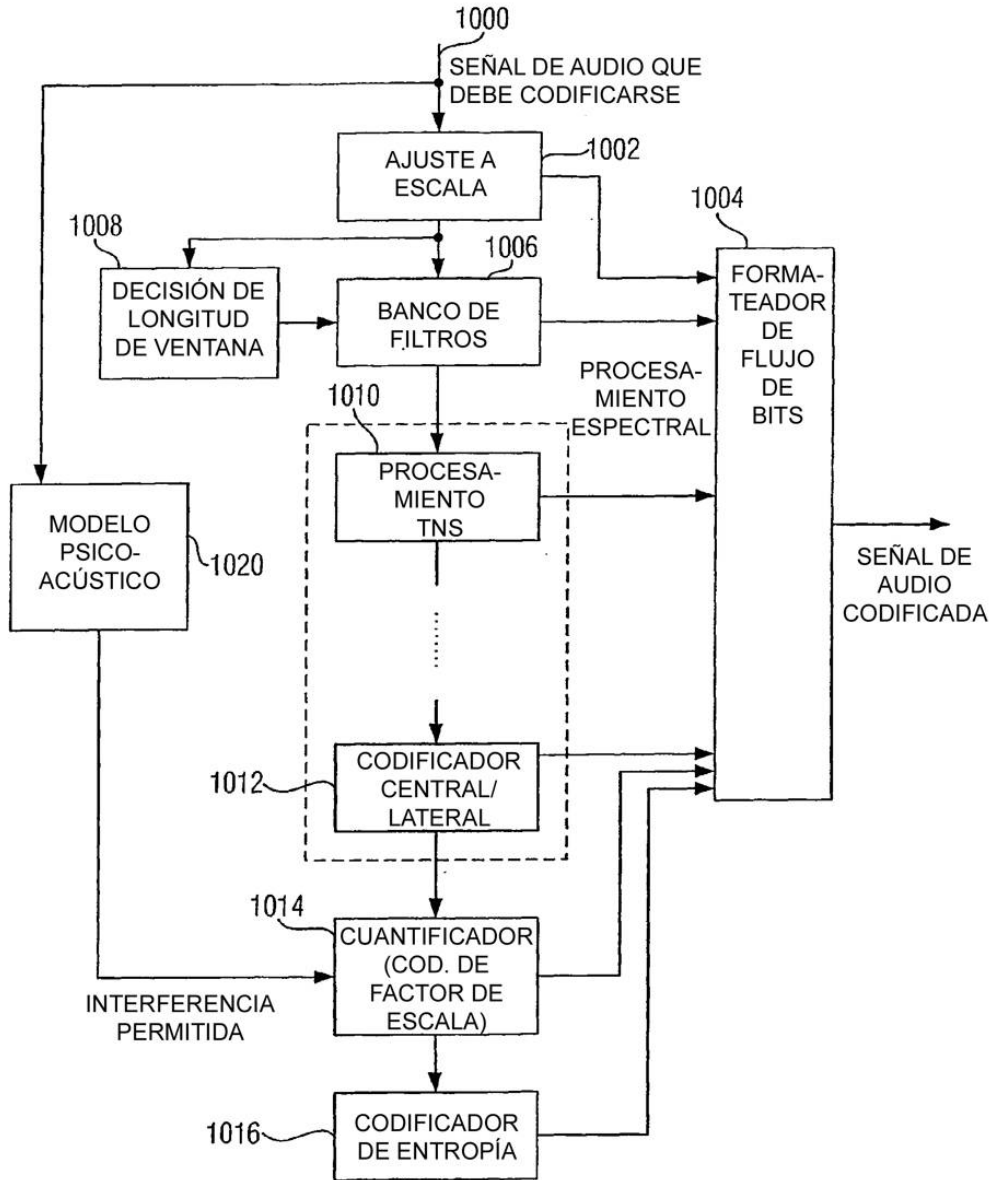
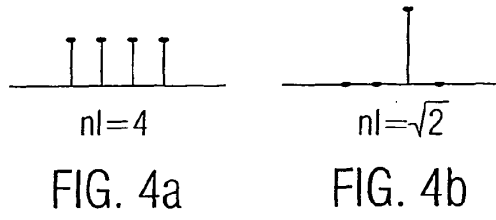


FIG. 3  
(ESTADO DE LA TÉCNICA)



$$pe = nl(b) \cdot \log_2 \left( \frac{e(b)}{nb(b)} + 1,5 \right)$$

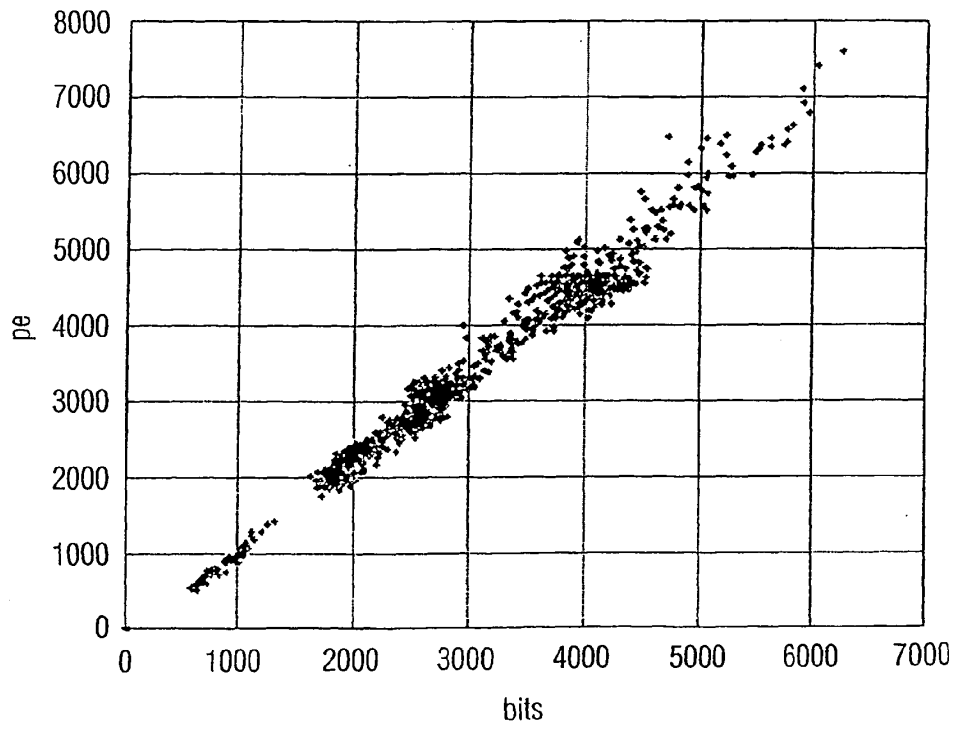
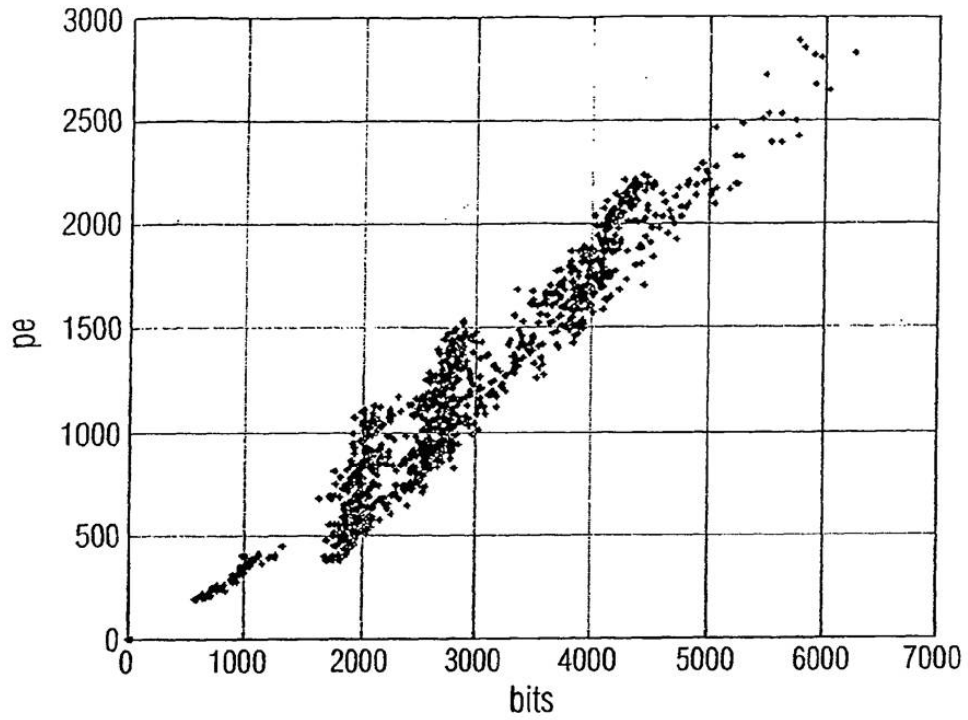


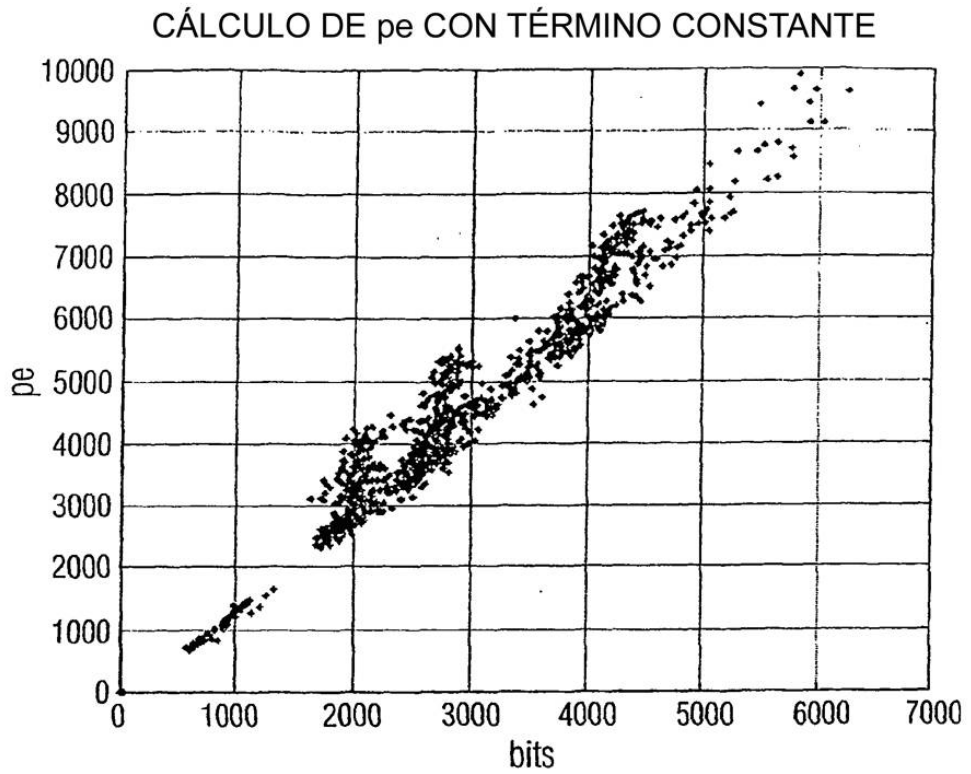
FIG. 5

pe CALCULADO SEGÚN LA NORMA ISO/IEC IS  
13818-7 MPEG-2 ADVANCED AUDIO CODING (AAC)



$$pe = -\sum_b \text{width}(b) \cdot \lg \left( \frac{nb(b)}{e(b)+1} \right)$$

FIG. 6  
(ESTADO DE LA TÉCNICA)

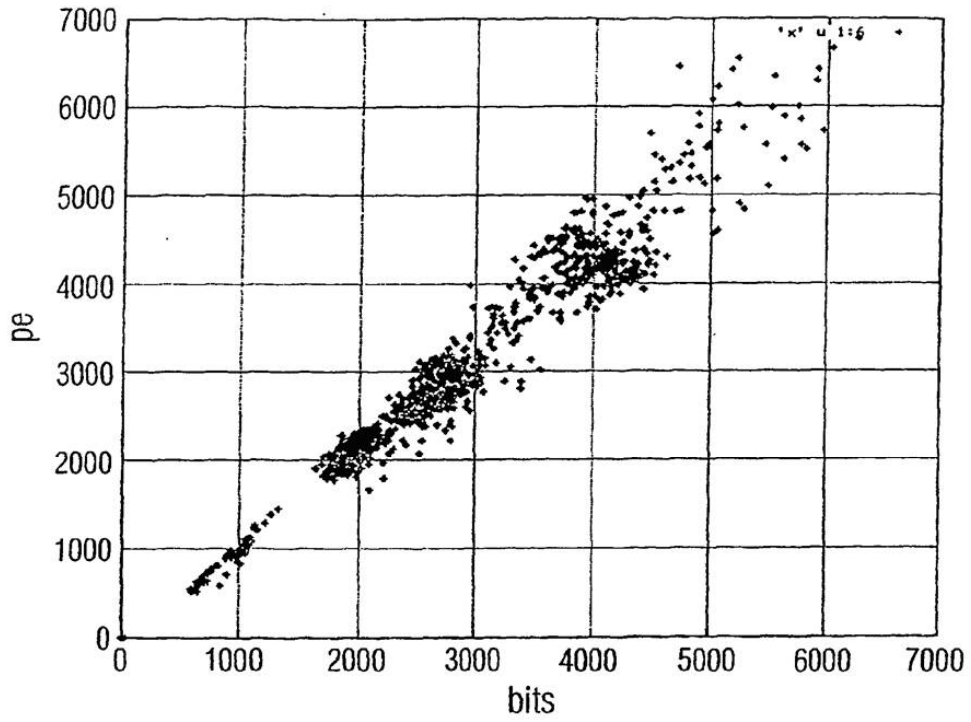


$$p_e = \sum_b \text{width}(b) \cdot \log_2 \left( \frac{e(b)}{nb(b)} + 1,5 \right)$$

FIG. 7



CÁLCULO DE  $pe$  POR LÍNEAS CON  
TÉRMINO CONSTANTE



$$pe = \sum_b \sum_{kOffset(b)}^{kOffset(b+1)-1} \log_2 \left( \frac{X^2(k) \cdot width(b)}{nb(b)} + 1,5 \right)$$

FIG. 8