

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 396 481**

51 Int. Cl.:

**G10L 19/14** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.04.2009 E 09730909 (0)**

97 Fecha y número de publicación de la concesión europea: **28.11.2012 EP 2272063**

54 Título: **Método y aparato para codificación selectiva de señales en base al rendimiento del codificador de núcleo**

30 Prioridad:

**09.04.2008 US 99842**

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

**21.02.2013**

73 Titular/es:

**MOTOROLA MOBILITY, INC. (100.0%)  
600 North US Highway 45  
Libertyville, IL 60048, US**

72 Inventor/es:

**ASHLEY, JAMES, P.;  
GIBBS, JONATHAN, A. y  
MITTAL, UDAR**

74 Agente/Representante:

**UNGRÍA LÓPEZ, Javier**

**ES 2 396 481 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

Método y aparato para codificación selectiva de señales en base al rendimiento del codificador de núcleo

5 **Antecedentes**

10 La transmisión de señales de texto, imágenes, voz y habla a través de canales de comunicación, incluyendo Internet, está aumentando rápidamente, como también la provisión de servicios multimedia capaces de acomodar varios tipos de información, tales como texto, imágenes y música. Las señales multimedia, incluyendo señales de voz y música, requieren una anchura de banda amplia al tiempo de la transmisión. Por lo tanto, para transmitir datos multimedia, incluyendo texto, imágenes y audio, es altamente deseable que los datos estén comprimidos.

15 La compresión de señales digitales de voz y audio es bien conocida. Generalmente se requiere compresión para transmitir eficientemente señales por un canal de comunicaciones, o para almacenar señales comprimidas en un dispositivo de medios digitales, tal como un dispositivo de memoria de estado sólido o disco duro de ordenador.

20 Un principio fundamental de la compresión de datos es la eliminación de datos redundantes. Los datos pueden ser comprimidos eliminando información temporal redundante tal como donde un sonido se repite, es predecible o es perceptivamente redundante. Esto tiene en cuenta la insensibilidad humana a altas frecuencias.

25 En general, la compresión da lugar a degradación de la señal, dando lugar a tasas de compresión más altas a mayor degradación. Se dice que un flujo de bits es escalable cuando se puede quitar partes del flujo de una forma que el flujo secundario resultante forme otro flujo de bits válido para algún decodificador deseado, y el flujo secundario representa el contenido de la fuente con una calidad de reconstrucción que es menor que la del flujo de bits completo original, pero es alta cuando se considera la cantidad inferior de los datos restantes. Los flujos de bits que no proporcionan esta propiedad se denominan flujos de bits monocapa. Los modos de escalabilidad usuales son escalabilidad temporal, espacial y de calidad. La escalabilidad permite que la señal comprimida sea ajustada para rendimiento óptimo por un canal de banda limitada.

30 La escalabilidad puede ser implementada de tal forma que se faciliten múltiples capas de codificación, incluyendo una capa base y al menos una capa de mejora, y las respectivas capas se forman de manera que tengan resoluciones diferentes.

35 Aunque muchos esquemas de codificación son genéricos, algunos esquemas de codificación incorporan modelos de la señal. En general, se logra una mejor compresión de la señal cuando el modelo es representativo de la señal que es codificada. Así, es conocido elegir el esquema de codificación en base a una clasificación del tipo de señal. Por ejemplo, una señal de voz puede ser modelada y codificada de una forma diferente a una señal de música. Sin embargo, por lo general, la clasificación de la señal es un problema difícil.

40 Un ejemplo de una técnica de compresión (o "codificación") que ha sido muy popular para codificación de voz digital es conocida como predicción lineal excitada por código (CELP), que es uno de una familia de algoritmos de codificación de "análisis por síntesis". Análisis por síntesis se refiere en general a un proceso de codificación por el que se usan múltiples parámetros de un modelo digital para sintetizar un conjunto de señales candidato que se comparan con una señal de entrada y analizan para ver si hay distorsión. Un conjunto de parámetros que produce la distorsión más baja es transmitido o almacenado posteriormente, y eventualmente se usa para reconstruir una estimación de la señal de entrada original. CELP es un método concreto de análisis por síntesis que usa uno o más libros de código, cada uno de los cuales incluye esencialmente conjuntos de vectores de código que son recuperados del libro de código en respuesta a un índice de libro de código.

50 En los codificadores CELP modernos, existe el problema de mantener reproducción de voz y audio de alta calidad a tasas de datos razonablemente bajas. Esto es especialmente verdadero con respecto a la música u otras señales audio genéricas que no encajan muy bien en el modelo de voz CELP. En este caso, la incongruencia del modelo puede producir una calidad audio severamente degradada que puede ser inaceptable para un usuario final del equipo que emplea tales métodos.

55 La Publicación de la Solicitud de Patente PCT número WO 97/15983 describe un método y aparato que permite la codificación, manipulación y decodificación de señales audio independientemente del contenido de señal específico. Un codificador (AC) lleva a cabo una codificación orientada a objeto y genera una señal codificada, con tasa de bits y anchura de banda variables, incluyendo una capa base y una o más capas de mejora para cada objeto. Las capas de mejora se seleccionan en base a la relación de señal a ruido SNR. El codificador (AC) es capaz de operar según algoritmos ad hoc y cualquier algoritmo estandarizado y selecciona el algoritmo más conveniente dependiendo del objeto a codificar.

65 La Publicación de la Solicitud de Patente PCT número WO 03/073741 describe cuantificadores escalables para audio caracterizados por una métrica de distorsión basada en percepción, no uniforme, que operan en un dominio comprimido-expandido común que incluye tanto capa base como una o más capas de mejora. El dominio

comprimido-expandido común está diseñado para permitir el uso de la misma métrica MSE no ponderada para selección óptima del parámetro de cuantificación en múltiples capas, explotando la dependencia estadística de la señal de capa de mejora en los parámetros de cuantificación usados en la capa precedente.

5 Un artículo de Ramprashad, S. A., titulado "Embedded coding using a mixed speech and audio coding paradigm", International Journal of Speech Technology, Kluwer, Dordrecht, NL, vol. 2, nº 4, 1 Mayo 1999, páginas 359-372, propone un algoritmo y estructura de codificación de voz/audio incorporados híbridos de dos etapas. La primera etapa de la estructura consta de un codificador de voz de núcleo que proporciona una tasa de bits de salida mínima y un rendimiento aceptable en entradas de voz limpias. La segunda etapa es un codificador basado en percepción/transformada que proporciona un flujo de bits opcional separado para la mejora de la salida de etapa de núcleo. La estructura de dos etapas puede ser usada para mejorar la calidad de un codec existente sin modificación del algoritmo de codificación original.

### Breve descripción de las figuras

15 Las figuras acompañantes, en las que números de referencia análogos se refieren a elementos idénticos o funcionalmente similares en todas las vistas separadas y que conjuntamente con la descripción detallada siguiente se incorporan y forman parte de la memoria descriptiva, sirven para ilustrar mejor varias realizaciones y para explicar varios principios y ventajas según la presente invención.

20 La figura 1 es un diagrama de bloques de un sistema de codificación y un sistema de decodificación de la técnica anterior.

25 La figura 2 es un diagrama de bloques de un sistema de codificación y un sistema de decodificación según algunas realizaciones de la invención.

La figura 3 es un diagrama de flujo del método para seleccionar un sistema de codificación según algunas realizaciones de la invención.

30 Las figuras 4-6 son una serie de gráficos que representan señales ejemplares en un comparador/selector según algunas realizaciones de la invención cuando se introduce una señal de voz.

35 Las figuras 7-9 son una serie de gráficos que representan señales ejemplares en un comparador/selector según algunas realizaciones de la invención cuando se introduce una señal de música.

La figura 10 es un diagrama de flujo de un método para codificación selectiva de señales según algunas realizaciones de la invención.

40 Los expertos apreciarán que los elementos de las figuras se ilustran a efectos de simplicidad y claridad y no se representan necesariamente a escala. Por ejemplo, las dimensiones de algunos elementos de las figuras pueden haberse exagerado con relación a otros elementos para ayudar a mejorar la comprensión de realizaciones de la presente invención.

### Descripción detallada

45 Antes de describir en detalle realizaciones según la presente invención, se deberá observar que las realizaciones residen primariamente en combinaciones de pasos del método y componentes del aparato relacionados con la codificación selectiva de señales en base al encaje de modelo. Consiguientemente, los componentes de aparato y los pasos del método se han representado, donde sea apropiado, por símbolos convencionales en los dibujos, que representan solamente los detalles específicos que son pertinentes para la comprensión de las realizaciones de la presente invención definida en las reivindicaciones independientes así como para no oscurecer la descripción con detalles que serán fácilmente evidentes para los expertos en la técnica que conozcan la presente descripción.

55 En este documento, términos relacionales tal como primero y segundo, superior e inferior, y análogos pueden ser usados únicamente para distinguir una entidad o acción de otra entidad o acción sin requerir o implicar necesariamente ninguna relación u orden real entre tales entidades o acciones. Se pretende que los términos "incluye", "incluyendo" o cualquier otra variación de los mismos, cubra una inclusión no exclusiva, de tal manera que un proceso, método, artículo o aparato que incluya una lista de elementos no incluya solamente dichos elementos, sino que pueda incluir otros elementos no expresamente enumerados o inherentes a tal proceso, método, artículo o aparato. Un elemento precedido por "incluye ... un" no excluye, sin más limitaciones, la existencia de elementos idénticos adicionales en el proceso, método, elemento o aparato que incluya el elemento.

65 Se apreciará que las realizaciones de la invención aquí descrita pueden incluir uno o más procesadores convencionales e instrucciones de programa almacenadas únicas que controlen el único o más procesadores para implementar, en unión con algunos circuitos no de procesador, algunas, la mayoría o todas las funciones de codificación selectiva de señales en base al encaje de modelo aquí descrito. Alternativamente, algunas o todas las

funciones podrían ser implementadas por una máquina de estado que no tenga instrucciones de programa almacenadas, o en uno o más circuitos integrados específicos de aplicación (ASICs), en los que cada función o algunas combinaciones de algunas de las funciones se implementan como lógica personalizada. Naturalmente, se podría usar una combinación de los dos métodos. Así, aquí se describen los métodos y los medios para estas funciones. Además, se espera que los expertos, a pesar del esfuerzo posiblemente significativo y muchas opciones de diseño motivadas, por ejemplo, por el tiempo disponible, la tecnología actual, y consideraciones económicas, cuando se dejen guiar por los conceptos y principios aquí descritos, serán fácilmente capaces de generar tales instrucciones de software y programas y CIs con experimentación mínima.

La figura 1 es un diagrama de bloques de un sistema de codificación y decodificación incorporado 100 de la técnica anterior. En la figura 1, se introduce una señal original  $s(n)$  102 en un codificador de capa de núcleo 104 de un sistema de codificación. El codificador de capa de núcleo 104 codifica la señal 102 y produce una señal codificada de capa de núcleo 106. Además, se introduce una señal original 102 en un codificador de capa de mejora 108 del sistema de codificación. El codificador de capa de mejora 108 también recibe una primera señal reconstruida  $s_c(n)$  110 como una entrada. La primera señal reconstruida 110 es producida pasando la señal codificada de capa de núcleo 106 a través de un primer decodificador de capa de núcleo 112. El codificador de capa de mejora 108 se usa para codificar información adicional en base a alguna comparación de las señales  $s(n)$  (102) y  $s_c(n)$  (110), y puede usar opcionalmente parámetros del codificador de capa de núcleo 104. En una realización, el codificador de capa de mejora 108 codifica una señal de error que es la diferencia entre la señal reconstruida 110 y la señal de entrada 102. El codificador de capa de mejora 108 produce una señal codificada de capa de mejora 114. Tanto la señal codificada de capa de núcleo 106 como la señal codificada de capa de mejora 114 son pasadas a un canal 116. El canal representa un medio, tal como un canal de comunicación y/o medio de almacenamiento.

Después de pasar a través del canal, se produce una segunda señal reconstruida 118 pasando la señal codificada de capa de núcleo recibida 106' a través de un segundo decodificador de capa de núcleo 120. El segundo decodificador de capa de núcleo 120 realiza la misma función que el primer decodificador de capa de núcleo 112. Si la señal codificada de capa de mejora 114 también se pasa a través del canal 116 y es recibida como señal 114', se puede pasar a un decodificador de capa de mejora 122. El decodificador de capa de mejora 122 también recibe la segunda señal reconstruida 118 como una entrada y produce una tercera señal reconstruida 124 como salida. La tercera señal reconstruida 124 concuerda con la señal original 102 más estrechamente que la segunda señal reconstruida 118.

La señal codificada de capa de mejora 114 incluye información adicional que permite que la señal 102 sea reconstruida más exactamente que la segunda señal reconstruida 118. Es decir, es una reconstrucción mejorada.

Una ventaja de dicho sistema de codificación incorporado es que un canal particular 116 puede no ser capaz de soportar sistemáticamente el requisito de anchura de banda asociado con algoritmos de codificación audio de alta calidad. Sin embargo, un codificador incorporado permite recibir un flujo de bits parcial (por ejemplo, solamente el flujo de bits de la capa de núcleo) del canal 116 para producir, por ejemplo, solamente la salida audio de núcleo cuando el flujo de bits de capa de mejora se pierde o corrompe. Sin embargo, hay compromisos de calidad entre codificadores incorporados frente a no incorporados, y también entre diferentes objetivos de optimización de codificación incorporada. Es decir, la codificación de capa de mejora de calidad más alta puede ayudar a lograr un mejor equilibrio entre capas de núcleo y de mejora, y también reduce la tasa general de datos para mejores características de transmisión (por ejemplo, congestión reducida), que puede dar lugar a menores tasas de error de paquete para las capas de mejora.

Aunque muchos esquemas de codificación son genéricos, algunos esquemas de codificación incorporan modelos de la señal. En general, se logra una mejor compresión de señal cuando el modelo es representativo de la señal que se codifica. Así, es conocido elegir el esquema de codificación en base a una clasificación del tipo de señal. Por ejemplo, una señal de voz puede ser modelada y codificada de forma diferente a una señal de música. Sin embargo, la clasificación de la señal es un problema difícil en general.

La figura 2 es un diagrama de bloques de un sistema de codificación y decodificación 200 según algunas realizaciones de la invención. Con referencia a la figura 2, una señal original 102 es introducida en un codificador de capa de núcleo 104 de un sistema de codificación. La señal original 102 puede ser una señal de voz/audio u otro tipo de señal. El codificador de capa de núcleo 104 codifica la señal 102 y produce una señal codificada de capa de núcleo 106. Se produce una primera señal reconstruida 110 pasando la señal codificada de capa de núcleo 106 a través de un primer decodificador de capa de núcleo 112. La señal original 102 y la primera señal reconstruida 110 son comparadas en un módulo comparador/selector 202. El módulo comparador/selector 202 compara la señal original 102 con la primera señal reconstruida 110 y, en base a la comparación, produce una señal de selección 204 que selecciona cuál de los codificadores de capa de mejora 206 usar. Aunque solamente se representan en la figura dos codificadores de capa de mejora, se deberá entender que se puede usar múltiples codificadores de capa de mejora. El módulo comparador/selector 202 puede seleccionar el codificador de capa de mejora que más probablemente genere la mejor señal reconstruida.

Aunque se representa el decodificador de capa de núcleo 112 para recibir la señal codificada de capa de núcleo 106

que es enviada correspondientemente al canal 116, la conexión física entre elementos 104 y 106 puede permitir una implementación más eficiente de tal manera que los elementos y/o estados de procesado comunes puedan ser compartidos y así no requerirían regeneración o duplicación.

5 Cada codificador de capa de mejora 206 recibe la señal original 102 y la primera señal reconstruida como entradas (o una señal, tal como una señal de diferencia, derivada de estas señales), y el codificador seleccionado produce una señal codificada de capa de mejora 208. En una realización, el codificador de capa de mejora 206 codifica una  
10 señal de error que es la diferencia entre la señal reconstruida 110 y la señal de entrada 102. La señal codificada de capa de mejora 208 contiene información adicional en base a una comparación de las señales  $s(n)$  (102) y  $s_c(n)$  (110). Opcionalmente, puede usar parámetros del decodificador de capa de núcleo 104. La señal codificada de capa de núcleo 106, la señal codificada de capa de mejora 208 y la señal de selección 204 son pasadas al canal 116. El canal representa un medio, tal como un canal de comunicación y/o un medio de almacenamiento.

15 Después de pasar a través del canal, se produce una segunda señal reconstruida 118 pasando la señal codificada de capa de núcleo recibida 106' a través de un segundo decodificador de capa de núcleo 120. El segundo decodificador de capa de núcleo 120 realiza la misma función que el primer decodificador de capa de núcleo 112. Si la señal codificada de capa de mejora 208 también se pasa a través del canal 116 y es recibida como señal 208', se puede pasar a un decodificador de capa de mejora 210. El decodificador de capa de mejora 210 también recibe la  
20 segunda señal reconstruida 118 y la señal de selección recibida 204' como entradas y produce una tercera señal reconstruida 212 como salida. La operación del decodificador de capa de mejora 210 depende de la señal de selección recibida 204'. La tercera señal reconstruida 212 concuerda con la señal original 102 más estrechamente que la segunda señal reconstruida 118.

25 La señal codificada de capa de mejora 208 incluye información adicional, así la tercera señal reconstruida 212 concuerda con la señal 102 más exactamente que la segunda señal reconstruida 118.

La figura 3 es un diagrama de flujo del método para seleccionar un sistema de codificación según algunas realizaciones de la invención. En particular, la figura 3 describe la operación de un módulo comparador/selector en una realización de la invención. Después del bloque de inicio 302, la señal de entrada (102 en la figura 2) y la señal  
30 reconstruida (110 en la figura 2) son transformadas, si se desea, a un dominio de señal seleccionado. Las señales de dominio de tiempo pueden ser usadas sin transformación o, en el bloque 304, las señales pueden ser transformadas a un dominio espectral, tal como el dominio de frecuencia, un dominio de transformada de coseno discreta modificada (MDCT), o un dominio de ondita, por ejemplo, y también pueden ser procesadas por otros elementos opcionales, tales como ponderación perceptible de algunas características de frecuencia o temporales de  
35 las señales. La señal de entrada transformada (o dominio de tiempo) se denota como  $S(k)$  para el componente espectral  $k$ , y la señal reconstruida transformada (o dominio de tiempo) se denota como  $S_c(k)$  para el componente espectral  $k$ . Para cada componente  $k$  en un conjunto seleccionado de componentes (que pueden ser todos o solamente algunos de los componentes), la energía,  $E_{tot}$ , en todos los componentes  $S_c(k)$  de la señal reconstruida es comparada con la energía,  $E_{err}$ , en los componentes que son mayores (en algún factor, por ejemplo) que el  
40 componente  $S(k)$  correspondiente de la señal de entrada original.

Aunque los componentes de señal de entrada y reconstruida pueden diferir de forma significativa en amplitud, un aumento significativo de la amplitud de un componente de señal reconstruida es indicativo de una señal de entrada  
45 pobremente modelada. Como tal, un componente de señal reconstruida de amplitud más baja puede ser compensado por un método dado de codificación de capa de mejora, mientras que un componente de señal reconstruida de amplitud más alta (es decir, pobremente modelado) puede ser más adecuado para un método alternativo de codificación de capa de mejora. Dicho método alternativo de codificación de capa de mejora puede implicar reducir la energía de algunos componentes de la señal reconstruida antes de la codificación de capa de mejora, de tal manera que se reduzca el ruido audible o la distorsión producidos como resultado de la incongruencia  
50 del modelo de señal de capa de núcleo.

Con referencia de nuevo a la figura 3, un bucle de componentes es inicializado en el bloque 306, donde el componente  $k$  es inicializado y las medidas de energía  $E_{tot}$  y  $E_{err}$  son inicializadas a cero. En el bloque de decisión 308, se realiza una comprobación para determinar si el valor absoluto del componente de la señal  
55 reconstruida es significativamente mayor que el componente correspondiente de la señal de entrada. Si es significativamente mayor, como ilustra la bifurcación positiva a partir del bloque de decisión 308, el componente es añadido a la energía de error  $E_{err}$  en el bloque 310 y el flujo continúa al bloque 312. En el bloque 312, el componente de las señales reconstruidas es añadido al valor de energía total,  $E_{tot}$ . En el bloque de decisión 314, el valor de componente se incrementa y se lleva a cabo una comprobación para determinar si todos los componentes han sido procesados. En caso negativo, como ilustra la bifurcación negativa a partir del bloque de decisión 314, el  
60 flujo vuelve al bloque 308. De otro modo, como ilustra la bifurcación positiva a partir del bloque de decisión 316, el bucle es completado y las energías acumuladas totales son comparadas en el bloque de decisión 316. Si la energía de error  $E_{err}$  es muy inferior al error total  $E_{tot}$ , como ilustra la bifurcación negativa a partir del bloque de decisión 316, se selecciona la capa de mejora tipo 1 en el bloque 318. De otro modo, como ilustra la bifurcación positiva a partir del bloque de decisión 316, se selecciona la capa de mejora tipo 2 en el bloque 320. El procesado de este  
65 bloque de señal de entrada termina en el bloque 322.

5 Será evidente a los expertos en la técnica que se puede usar otras medidas de energía de señal, tal como el valor absoluto del componente elevado a alguna potencia. Por ejemplo, la energía de un componente  $Sc(k)$  puede ser estimada como  $|Sc(k)|^P$ , y la energía de un componente  $S(k)$  puede ser estimada como  $|S(k)|^P$ , donde  $P$  es un número mayor que cero.

Será evidente a los expertos en la técnica que la energía de error  $E_{err}$  puede ser comparada con la energía total en la señal de entrada más bien que la energía total en la señal reconstruida.

10 El codificador puede ser implementado en un procesador programado. A continuación se ofrece un listado de código ejemplar correspondiente a la figura 3. Las variables  $energy\_tot$  y  $energy\_err$  son denotadas por  $E_{tot}$  y  $E_{err}$ , respectivamente, en la figura.

```

Thresh1 = 0.49;
Thresh2 = 0.264;
energy_tot = 0;
energy_err = 0;
for (k = kStart; k < kMax; k++)
{
    if (Thresh1*abs(Sc[k]) > abs(S[k])) {
        energy_err += abs(Sc[k]);
    }
    energy_tot += abs(Sc[k]);
}
if (energy_err < Thresh2*energy_tot)
    type = 1;
else
    type = 2;

```

15 En este ejemplo los valores umbral  $Thresh1$  y  $Thresh2$  se ponen a 0,49 y 0,264, respectivamente. Se puede usar otros valores dependiendo de los tipos de codificadores de capa de mejora usados y dependiendo también de qué dominio de transformada se use.

20 Se puede añadir una etapa de histéresis, de modo que el tipo de capa de mejora solamente se cambie si un número especificado de bloques de señal es del mismo tipo. Por ejemplo, si se está usando un codificador de tipo 1, no se seleccionará el tipo 2 a no ser que dos bloques consecutivos indiquen el uso del tipo 2.

25 Las figuras 4-6 son una serie de gráficos que muestran resultados ejemplares para una señal de voz. La gráfica 402 en la figura 4 representa la energía  $E_{tot}$  de la señal reconstruida. La energía es calculada en tramas de 20 milisegundos, de modo que la gráfica representa la variación en la energía de señal en un intervalo de 10 segundos. La gráfica 502 en la figura 5 representa la relación de la energía de error  $E_{err}$  a la energía total  $E_{tot}$  en el mismo período de tiempo. El valor umbral  $Thresh2$  se representa como la línea discontinua 504. La señal de voz en tramas donde la relación excede del umbral no está bien modelada por el codificador. Sin embargo, en la mayoría de las tramas no se supera el umbral. La gráfica 602 en la figura 6 representa la señal de selección o decisión en el mismo período de tiempo. En este ejemplo, el valor 0 indica que se selecciona el codificador de capa de mejora tipo 1 y un valor 1 indica que se selecciona el codificador de capa de mejora tipo 2. Se ignoran las tramas aisladas donde la relación es más alta que el umbral y la selección solamente se cambia cuando dos tramas consecutivas indican la misma selección. Así, por ejemplo, el codificador de capa de mejora tipo 1 se selecciona para la trama 141 aunque la relación exceda del umbral.

35 Las figuras 7-9 muestran una serie correspondiente de gráficos de una señal de música. La gráfica 702 en la figura 7 representa la energía  $E_{tot}$  de la señal de entrada. De nuevo, la energía es calculada en tramas de 20 milisegundos, así la gráfica representa la variación en energía de entrada en un intervalo de 10 segundos. La gráfica 802 en la figura 8 representa la relación de la energía de error  $E_{err}$  a la energía total  $E_{tot}$  en el mismo período de tiempo. El valor umbral  $Thresh2$  se representa como la línea discontinua 504. La señal de música en las tramas donde la relación excede del umbral no está bien modelada por el codificador. Éste es el caso de la mayoría de las tramas, dado que el codificador de núcleo está diseñado para señales de voz. La gráfica 902 en la figura 9 representa la señal de selección o decisión en el mismo período de tiempo. De nuevo, el valor 0 indica que el codificador de capa de mejora de tipo 1 se selecciona y un valor 1 indica que se selecciona el codificador de capa de mejora de tipo 2. Así, el codificador de capa de mejora de tipo 2 se selecciona la mayor parte del tiempo. Sin embargo, en las tramas donde el codificador de núcleo opera bien para música, se selecciona el codificador de capa de mejora de tipo 1.

En una prueba con 22.803 tramas de una señal de voz, se seleccionó el codificador de capa de mejora de tipo 2 solamente en 227 tramas, es decir, solamente 1% del tiempo. En una prueba con 29.644 tramas de música, se

seleccionó el codificador de capa de mejora de tipo 2 en 16.145 tramas, es decir, 54% del tiempo. En las otras tramas el núcleo codificador opera bien para la música y se seleccionó el codificador de capa de mejora para voz. Así, el comparador/selector no es un clasificador de voz/música. Esto contrasta con los esquemas anteriores que pretenden clasificar la señal de entrada como voz o música y luego seleccionar consiguientemente el esquema de codificación. El acercamiento de la invención es seleccionar el codificador de capa de mejora dependiendo del rendimiento del codificador de capa de núcleo.

La figura 10 es un diagrama de flujo que representa la operación de un codificador incorporado según algunas realizaciones de la invención. El diagrama de flujo representa un método usado para codificar una trama de datos de señal. La longitud de la trama se selecciona en base a una característica temporal de la señal. Por ejemplo, una trama de 20 ms puede ser usada para señales de voz. Después del bloque de inicio 1002 en la figura 10, la señal de entrada es codificada en el bloque 1004 usando un codificador de capa de núcleo para producir una señal codificada de capa de núcleo. En el bloque 1006 la señal codificada de capa de núcleo es decodificada para producir una señal reconstruida. En esta realización, se genera una señal de error, en el bloque 1008, como la diferencia entre la señal reconstruida y la señal de entrada. La señal reconstruida es comparada con la señal de entrada en el bloque 1010 y en el bloque de decisión 1012 se determina si la señal reconstruida es una buena concordancia para la señal de entrada. Si la concordancia es buena, como ilustra la bifurcación positiva a partir del bloque de decisión 1012, se usa el codificador de capa de mejora de tipo 1 para codificar la señal de error en el bloque 1014. Si la concordancia no es buena, como ilustra la bifurcación negativa a partir del bloque de decisión 1012, se usa el codificador de capa de mejora de tipo 2 para codificar la señal de error en el bloque 1016. En el bloque 1018, la señal codificada de capa de núcleo, la señal codificada de capa de mejora y el indicador de selección son enviadas al canal (para transmisión o almacenamiento, por ejemplo). El procesado de la trama termina en el bloque 1020.

En esta realización, el codificador de capa de mejora es sensible a una señal de error; sin embargo, en una realización alternativa, el codificador de capa de mejora es sensible a la señal de entrada y, opcionalmente, una o más señales del codificador de capa de núcleo y/o el decodificador de capa de núcleo. En otra realización, se usa una señal de error alternativa, tal como una diferencia ponderada entre la señal de entrada y la señal reconstruida. Por ejemplo, algunas frecuencias de la señal reconstruida pueden ser atenuadas antes de la formación de la señal de error. La señal de error resultante se puede denominar una señal de error ponderada.

En otra realización alternativa, el codificador y el decodificador de capa de núcleo también pueden incluir otras capas de mejora, y el comparador de la presente invención puede recibir como entrada la salida de una de las capas de mejora previas como la señal reconstruida. Adicionalmente, puede haber capas de mejora posteriores a dichas capas de mejora que pueden ser o no conmutadas como resultado de la comparación. Por ejemplo, un sistema de codificación incorporada puede incluir cinco capas. La capa de núcleo (L1) y la segunda capa (L2) pueden producir la señal reconstruida  $S_c(k)$ . La señal reconstruida  $S_c(k)$  y la señal de entrada  $S(k)$  se pueden usar luego para seleccionar los métodos de codificación de capa de mejora en las capas tres y cuatro (L3, L4). Finalmente, la capa cinco (L5) puede incluir solamente un solo método de codificación de capa de mejora.

El codificador puede seleccionar entre dos o más codificadores de capa de mejora dependiendo de la comparación entre la señal reconstruida y la señal de entrada.

El codificador y el decodificador pueden ser implementados en un procesador programado, en un procesador reconfigurable o en un circuito integrado específico de aplicación, por ejemplo.

En la memoria descriptiva anterior se ha descrito realizaciones específicas de la presente invención. Sin embargo, los expertos en la técnica observarán que se puede hacer varias modificaciones y cambios sin apartarse del alcance de la presente invención expuesto en las reivindicaciones siguientes. Consiguientemente, la memoria descriptiva y las figuras se han de considerar como ilustrativas más bien que en sentido restrictivo, y se ha previsto que tales modificaciones queden incluidas dentro del alcance de la presente invención. Los beneficios, las ventajas, las soluciones a problemas y cualquier elemento(s) que puedan producir cualquier beneficio, ventaja o solución o que sea más pronunciado, no se han de interpretar como unas características o elementos críticos, requeridos o esenciales de alguna o de todas las reivindicaciones. La invención se define únicamente por las reivindicaciones anexas incluyendo cualesquiera modificaciones realizadas durante la tramitación de esta solicitud y todos los equivalentes de las reivindicaciones concedidas.

**REIVINDICACIONES**

1. Un método (300) para codificar una señal de entrada incluyendo una señal audio, incluyendo el método:
  - 5 codificar la señal de entrada (102) usando un codificador de capa de núcleo (104) para producir una señal codificada de capa de núcleo (106);
  - decodificar la señal codificada de capa de núcleo para producir una señal reconstruida (110);
  - 10 comparar (308, 316) la señal reconstruida con la señal de entrada donde la comparación incluye estimar una energía  $E_{tot}$  como una suma de energías en un conjunto seleccionado de componentes de la señal reconstruida o la señal de entrada y estimar una energía  $E_{err}$  como una suma de energías de los componentes  $S_c(k)$  de la señal reconstruida para los que la relación  $S(k)/S_c(k)$  del componente  $S(k)$  de la señal de entrada al componente  $S_c(k)$  de la señal reconstruida excede de un valor umbral y donde la comparación incluye además comparar la energía  $E_{tot}$  con la energía  $E_{err}$ ;
  - 15 seleccionar (318, 320) un codificador de capa de mejora (206) de una pluralidad de codificadores de capa de mejora dependiendo de la comparación entre el componente de la señal reconstruida y el componente correspondiente de la señal de entrada; y
  - 20 generar una señal codificada de capa de mejora (208) usando el codificador de capa de mejora seleccionado, dependiendo la señal codificada de capa de mejora de la señal de entrada.
2. Un método según la reivindicación 1, incluyendo además:
  - 25 generar una señal de error como la diferencia entre la señal reconstruida y la señal de entrada, donde generar la señal codificada de capa de mejora incluye codificar la señal de error.
3. Un método según la reivindicación 2, donde la señal de error incluye una diferencia ponderada entre la señal reconstruida y la señal de entrada.
4. Un método según la reivindicación 1, incluyendo además:
  - 35 transformar la señal reconstruida para producir los componentes de la señal reconstruida; y
  - transformar la señal de entrada para producir los componentes de la señal de entrada,
  - 40 donde la transformada se selecciona del grupo de transformadas que consta de una transformada Fourier, una transformada de coseno discreta modificada (MDCT) y una transformada de ondita.
5. Un método según la reivindicación 1, donde la energía de un componente  $S_c(k)$  se estima como  $|S_c(k)|^P$ , y donde la energía de un componente  $S(k)$  se estima como  $|S(k)|^P$  donde P es un número mayor que cero.
6. Un método según la reivindicación 1, donde comparar la energía  $E_{tot}$  con la energía  $E_{err}$  incluye: comparar la relación de energías  $E_{err}/E_{tot}$  con un valor umbral.
7. Un método según la reivindicación 1, donde la capa de núcleo codificada incluye un codificador de voz.
8. Un método según la reivindicación 1, incluyendo además enviar la señal codificada de capa de núcleo, la señal codificada de capa de mejora y un indicador del codificador de capa de mejora seleccionado a un canal.
9. Un codificador de señal selectivo (200) incluyendo:
  - 55 un codificador de capa de núcleo (104) para recibir una señal de entrada (102) incluyendo una señal audio a codificar y para producir una señal codificada de capa de núcleo (106);
  - un decodificador de capa de núcleo (112) para recibir la señal codificada de capa de núcleo como entrada y para producir una señal reconstruida (110);
  - 60 estando dispuesto cada uno de una pluralidad de codificadores de capa de mejora (206) de manera que sea seleccionable para codificar una señal de error para producir una señal codificada de capa mejorada (208), incluyendo la señal de error una diferencia entre la señal de entrada y la señal reconstruida; y
  - 65 un módulo comparador/selector (202) para seleccionar un codificador de capa de mejora de la pluralidad de codificadores de capa de mejora dependiendo de una comparación de la señal reconstruida con la señal de entrada,

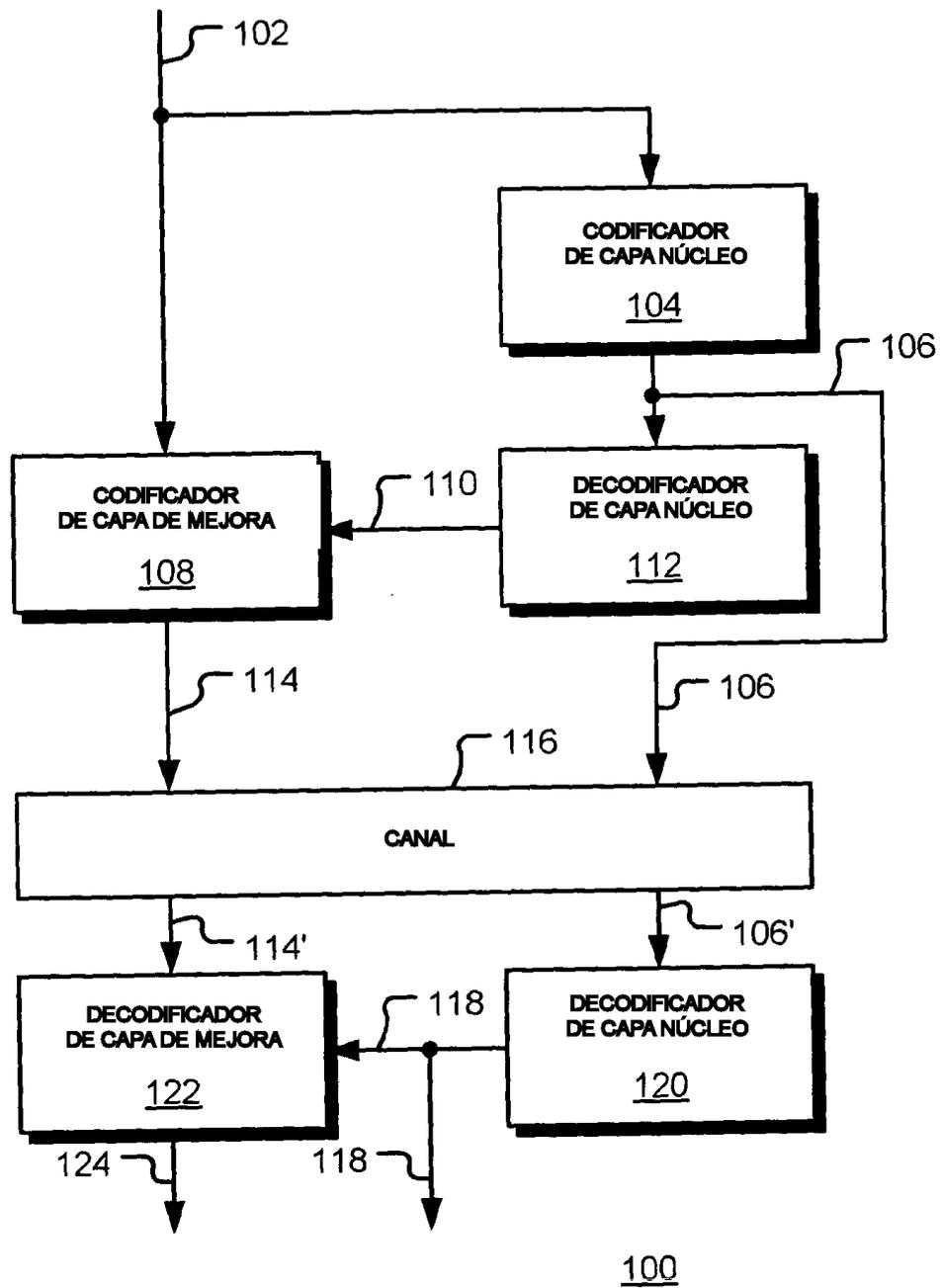
5 donde el módulo comparador/selector está configurado para estimar una energía  $E_{tot}$  como una suma de energías en un conjunto seleccionado de componentes de la señal reconstruida o la señal de entrada y para estimar una energía  $E_{err}$  como una suma de energías en componentes de la señal reconstruida sumando las energías de dichos componentes  $Sc(k)$  de la señal reconstruida para que la relación  $S(k)/Sc(k)$  del componente  $S(k)$  de la señal de entrada al componente  $Sc(k)$  de la señal reconstruida exceda de un valor umbral y donde el módulo comparador/selector está configurado además para comparar la energía  $E_{tot}$  con la energía  $E_{err}$ , y además,

10 donde la señal de entrada es codificada como la señal codificada de capa de núcleo, la señal codificada de capa mejorada y un indicador del codificador de capa mejorada seleccionado.

10 10. Un codificador de señal selectivo según la reivindicación 9, donde el codificador de capa de núcleo incluye un codificador de voz.

15 11. Un codificador de señal selectivo según la reivindicación 9, donde el módulo comparador/selector está dispuesto para comparar la energía  $E_{tot}$  con la energía  $E_{err}$  comparando la relación de energías  $E_{err}/E_{tot}$  con un valor umbral.

20 12. Un codificador de señal selectivo según la reivindicación 9, donde los componentes de la señal reconstruida y los componentes de la señal de entrada son calculados mediante una transformada seleccionada del grupo de transformadas que consta de una transformada Fourier, una transformada de coseno discreta modificada (MDCT) y una transformada de ondita.



**FIG. 1**

TÉCNICA ANTERIOR

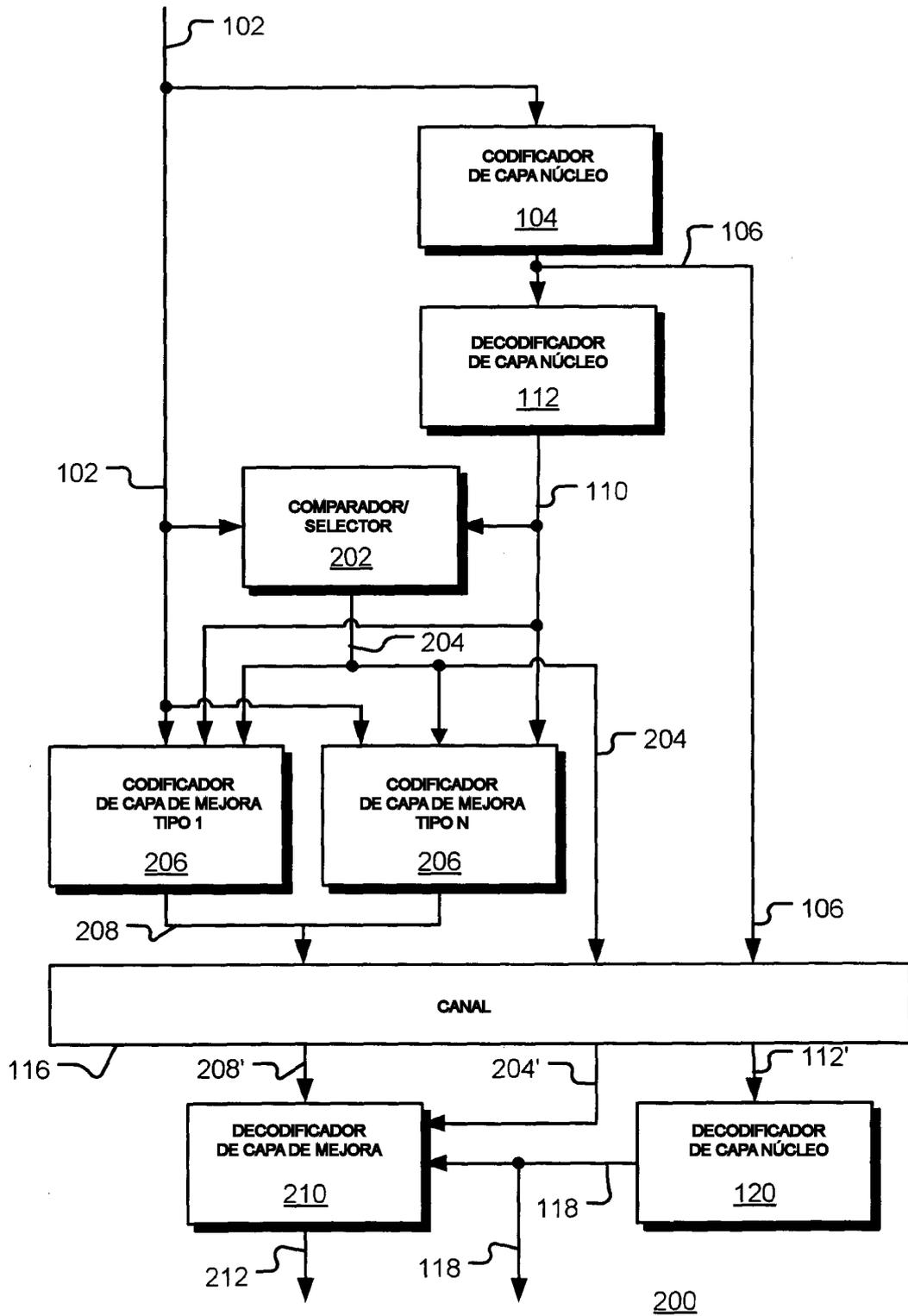


FIG. 2

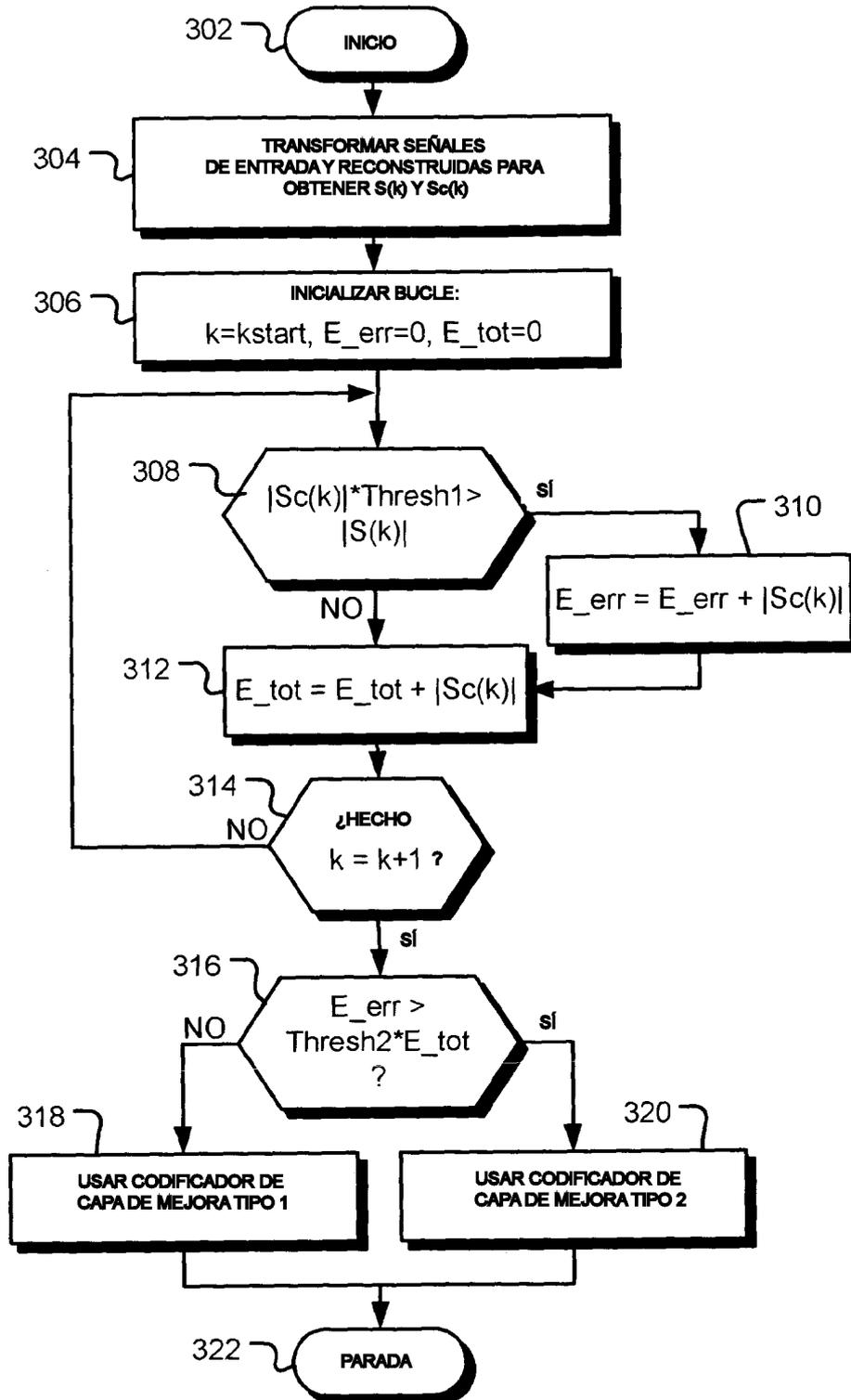
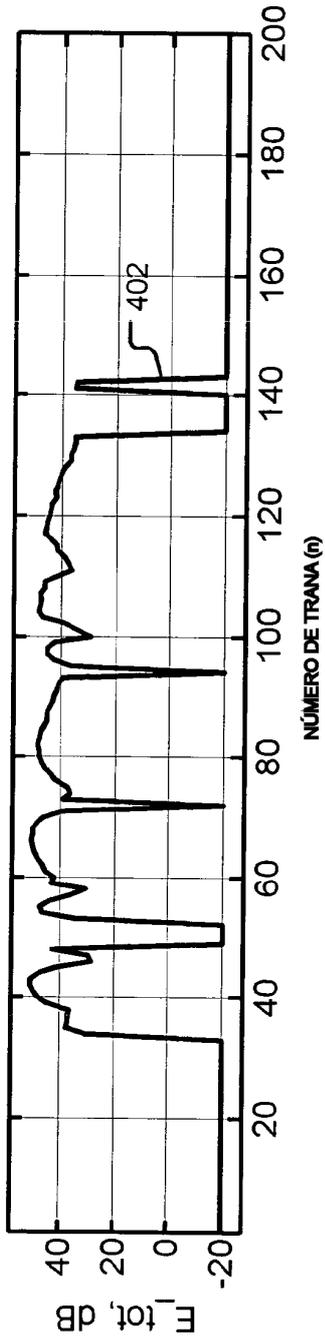
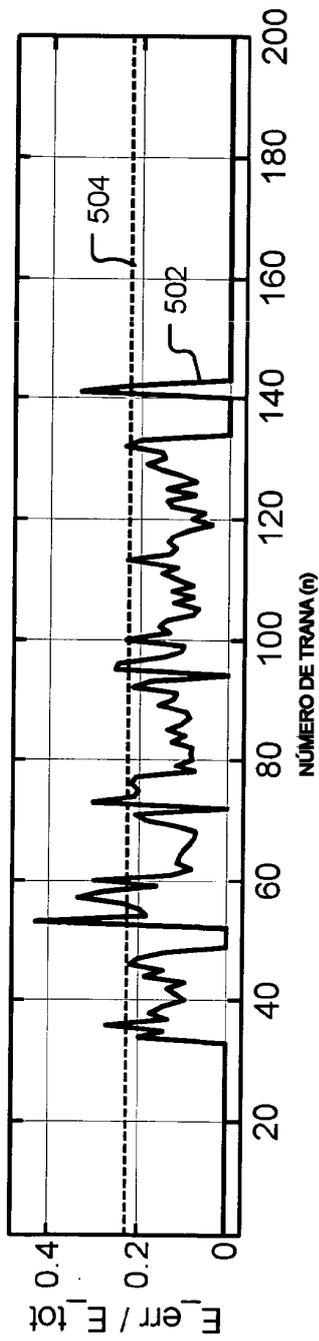


FIG. 3

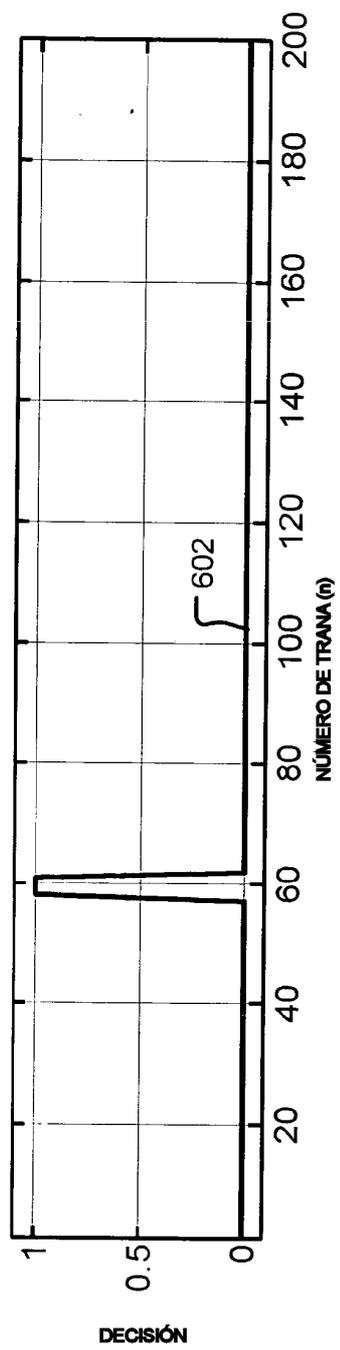
**FIG. 4**

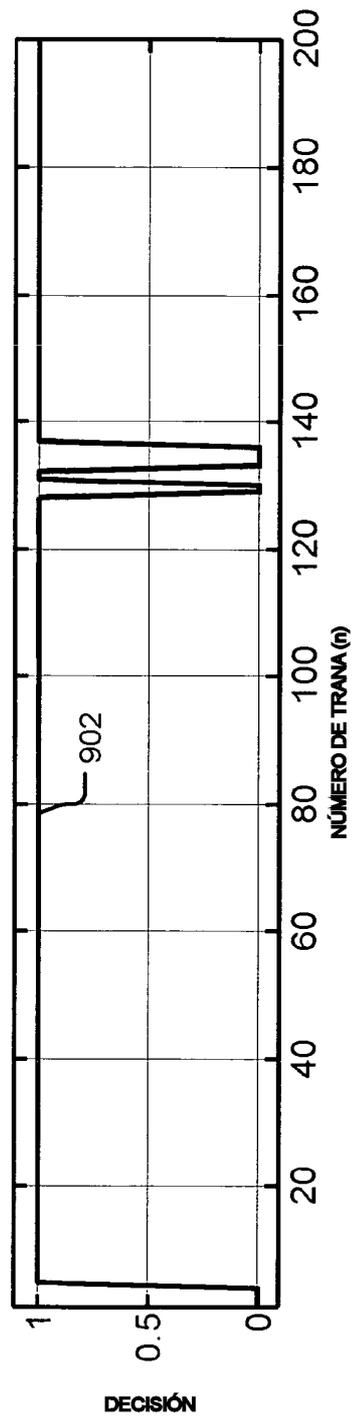
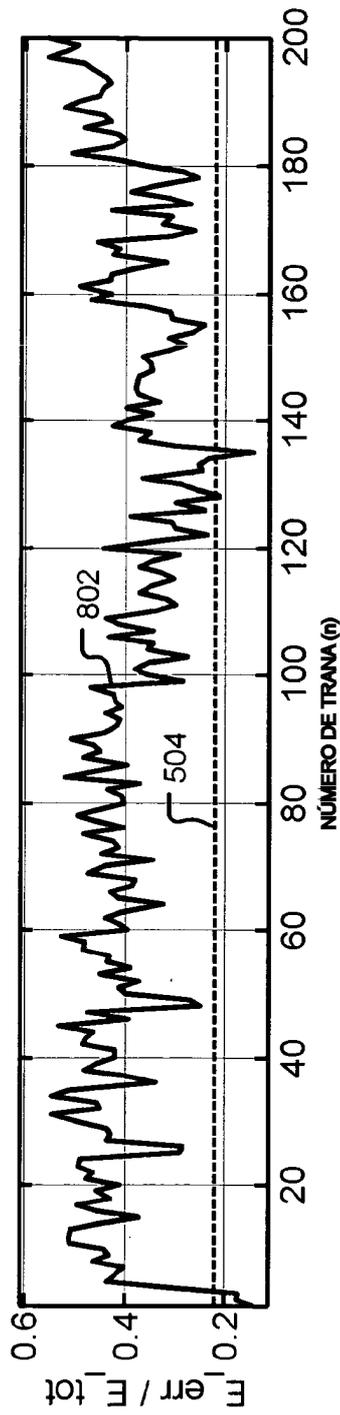
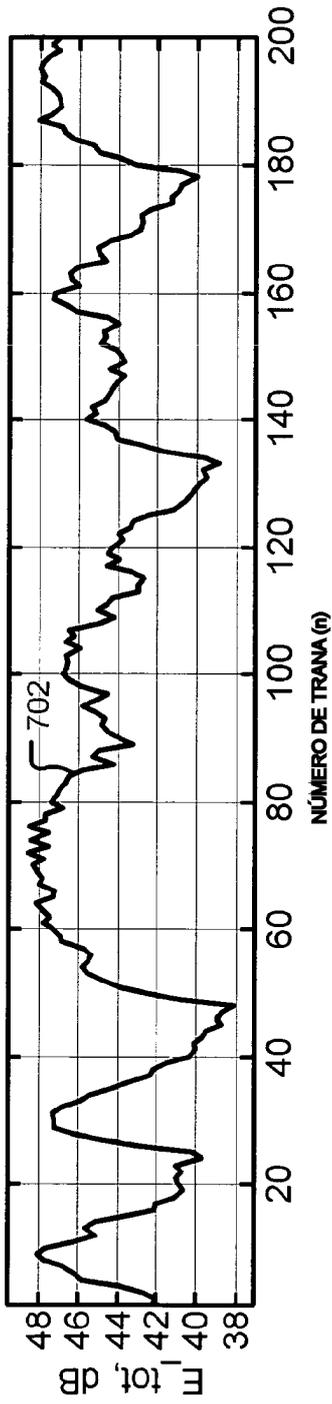


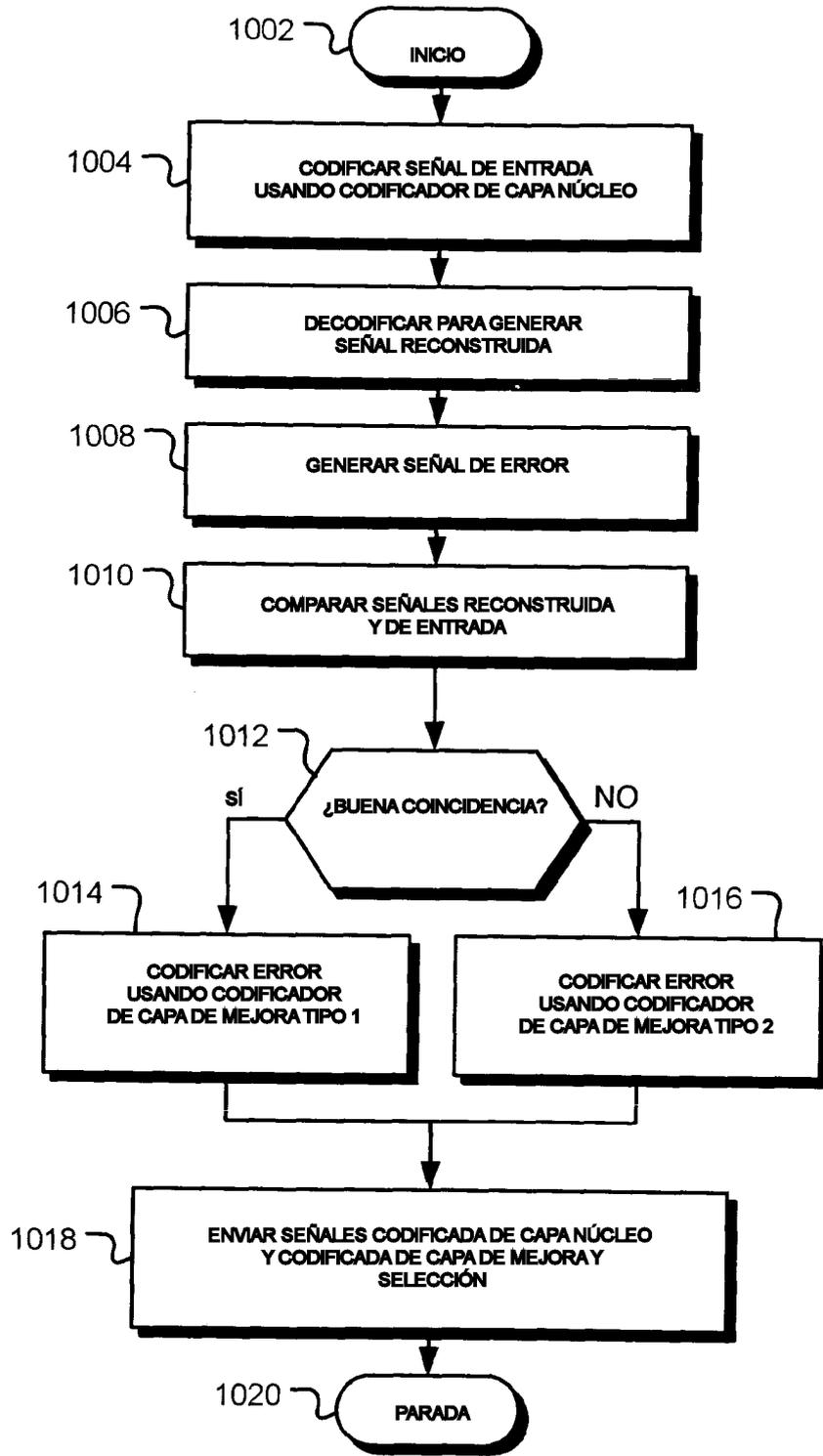
**FIG. 5**



**FIG. 6**







**FIG. 10**