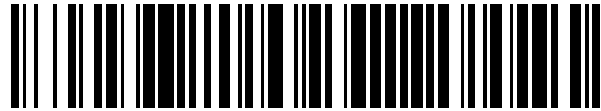


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 596 457**

51 Int. Cl.:

G10L 21/0388 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.07.2013 PCT/EP2013/063827**

87 Fecha y número de publicación internacional: **09.01.2014 WO14005992**

96 Fecha de presentación y número de la solicitud europea: **01.07.2013 E 13737170 (4)**

97 Fecha y número de publicación de la concesión europea: **24.08.2016 EP 2867894**

54 Título: **Dispositivo, método y programa informático para desplazamientos de frecuencia libremente seleccionables en el dominio de subbanda**

30 Prioridad:

02.07.2012 US 201261667119 P
10.07.2012 EP 12175808

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.01.2017

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)**
Hansastraße 27c
80686 München, DE y
TECHNISCHE UNIVERSITÄT ILMENAU (50.0%)

72 Inventor/es:

NAGEL, FREDERIK;
SCHNABEL, MICHAEL;
NEUKAM, CHRISTIAN y
SCHULLER, GERALD

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 596 457 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).

Dispositivo, método y programa informático para desplazamientos de frecuencia libremente seleccionables en el dominio de subbanda

DESCRIPCIÓN

5

La presente invención se refiere al procesamiento de señales de audio y, en particular, a un dispositivo, un método y un programa informático para desplazamientos de frecuencia arbitrarios en el dominio de subbanda.

10 Los sistemas de procesamiento de datos ayudados por computadora son una parte integral de la vida diaria en la sociedad actual la cual está caracterizada por nuevos medios. Desde hace ya algún tiempo se encuentran sistemas para consumir nuevos medios en casi cualquier hogar. Ejemplos de tales sistemas, que transmiten y reproducen datos en forma digital, son reproductores de datos de vídeo y audio, como, por ejemplo, para DVD y BluRay, para CD y el formato de archivo mp3. Estos sistemas de reproducción están caracterizados por una reproducción casi sin
15 pérdida de contenidos de medios. Aparte de las telecomunicaciones clásicas, internet es un portal importante para comunicaciones, por ejemplo por medio de VoIP. El procesamiento digital de señales subyacente es común en todas las tecnologías mencionadas. Es de importancia decisiva para la calidad de la reproducción y eficacia de las tecnologías digitales.

20 El procesamiento de señales de audio es cada vez más importante a este respecto. En la actualidad hay una pluralidad de codificadores de audio en el mercado, implementados, por ejemplo, mediante algoritmos para tratar digitalmente el material de audio para su almacenamiento o transmisión. El objetivo de cada método de codificación es comprimir el contenido de información de una señal de modo que se requiera un espacio de almacenamiento mínimo al tiempo que se mantiene la mejor calidad de reproducción posible. La eficacia de los codificadores de
25 audio modernos depende principalmente del almacenamiento requerido y, entre otras cosas, de la complejidad de cálculo requerida del algoritmo.

Básicamente, un codificador de audio digital es un instrumento para transferir señales de audio a un formato apropiado para almacenamiento o transmisión. Esto tiene lugar en el lado del transmisor del codificador de audio
30 (*encoder*). Los datos generados de esta forma se devuelven a la forma original en el receptor (*decoder*) y, en el caso ideal, corresponden a los datos originales, excepto por un retardo constante (*delay*). El objetivo general de los codificadores de audio es minimizar la cantidad de datos requeridos para representar la señal de audio y al mismo tiempo maximizar la calidad de reproducción percibida. Cuando se desarrollan codificadores de audio, una serie de factores han de tenerse en cuenta, como, por ejemplo, la fidelidad de la reproducción, la tasa de transmisión de
35 datos y la complejidad. Aparte de eso, el retardo añadido al procesar la señal (el retardo añadido) también desempeña un papel importante (Bosi y Goldberg, 2003).

En particular en los comienzos de la codificación de audio, la eficacia de los métodos era muy importante ya que solo había disponible un almacenamiento y una potencia computacional muy limitados. Hoy en día, esta demanda
40 parece ser menos importante. Incluso los ordenadores personales domésticos o portátiles son capaces de calcular sin dificultad algoritmos complicados en tiempo real, y las conexiones de internet de banda ancha proporcionan suficiente ancho de banda para transmitir material de audio codificado. Sin embargo, perfeccionar los métodos de codificación de audio es de importancia particular. En el campo de las comunicaciones móviles y la transmisión por satélite, el ancho de banda es muy limitado. Reducir las cantidades de datos que han de transmitirse es importante.
45 Adicionalmente, en estos campos se le da mucha importancia a la eficacia de la tecnología de codificación usada. Los algoritmos subyacentes deben presentar una estructura simple a fin de minimizar la potencia computacional y el consumo de corriente.

Otro aspecto es la calidad de las señales de audio codificadas reproducidas. Muchos codificadores de audio reducen
50 la cantidad de datos al usar una reducción de irrelevancia. A este respecto se pierden fracciones de señal en función de la tasa de transmisión de datos. Con tasas de transmisión de datos bajas, la calidad de las señales de audio reproducidas disminuye por tanto.

Generalmente, se puede diferenciar entre dos tipos de codificación de audio, codificación de audio sin pérdida y con
55 pérdida. La codificación de audio sin pérdida permite una reconstrucción exacta de la señal original del lado del receptor. El método con pérdida en cambio causa desviaciones de la señal original irreversibles por medio de un modelo de percepción subjetiva (Zölzer, 2005).

La codificación de audio sin pérdida está basada en reducir la redundancia contenida en la señal que ha de
60 codificarse. Un método común para ello es, por ejemplo, la predicción lineal (*linear predictive coding*, LPC) en asociación con codificación por entropía posterior. Tales métodos de codificación de audio permiten reconstruir la señal de entrada de manera precisa bit por bit a partir de la corriente de bits codificada.

La predicción lineal usa dependencias estadísticas entre muestras (*samples*) consecutivas de la señal a fin de poder

predecir valores futuros. Esto está basado en el hecho de que las muestras consecutivas son más similares de lo que lo son las muestras que están a una mayor distancia. La predicción se implementa mediante un filtro de predicción lineal que estima la muestra actual mediante una serie de muestras previas. Sin embargo, no es la estimación en sí misma la que es procesada adicionalmente, sino la diferencia entre este valor y la verdadera muestra en este punto. El objetivo de la predicción lineal es minimizar la energía de esta señal de error mediante filtros optimizados y transmitir esta señal de error, que requiere únicamente un pequeño ancho de banda (Weinzierl, 2008).

A continuación se codifica la señal de error por entropía. La entropía es una medida del contenido de información promedio e indica el mínimo teórico de los bits requeridos para la codificación. Un método típico para ello es la codificación Huffman. Ciertas palabras de código son asociadas aquí a muestras individuales, en función de su probabilidad estadística de ocurrencia. Se asocian símbolos cortos a muestras que ocurren frecuentemente y los valores de señal que ocurren raramente son representados por palabras de código largas. En promedio, la señal codificada se representa entonces por el menor número posible de bits (Bosi y Goldberg, 2003).

Tanto la predicción lineal como la codificación por entropía son reversibles y por consiguiente no eliminan información de la señal. Cuando se combinan los dos métodos, solo se eliminan las redundancias de la señal que va a codificarse. Dado que tales enfoques sin pérdida dependen en gran medida de la característica de señal, la ganancia de codificación es relativamente pequeña. La tasa de compresión lograda, es decir la relación de tasa de transmisión de bits de entrada y tasa de transmisión de bits de la señal codificada, se sitúa en un intervalo de 1,5 : 1 a 3 : 1 (Weinzierl, 2008).

La codificación de audio con pérdida está basada en el principio de la reducción de irrelevancia. Este método requiere un modelo de percepción humana, el cual describe fenómenos psicoacústicos del sentido del oído con respecto a la resolución de tiempo y frecuencia. Así, la codificación de audio con pérdida también se denomina codificación adaptada a la percepción o codificación psicoacústica. En el campo de la codificación de audio, todas las fracciones de señal que no pueden ser percibidas por el ser humano y son por consiguiente inaudibles se denominan irrelevantes (Zölzer, 2005). Para entender mejor el modo de funcionamiento de un codificador de audio adaptado a la percepción, es de gran importancia un conocimiento profundo de la psicoacústica.

El oído humano analiza un evento de sonido descomponiéndolo en grupos de frecuencias. Estos grupos de frecuencias se representan en la escala Bark y se denominan, según la literatura inglesa, como bandas críticas. Cada uno de estos grupos de frecuencias agrupa un intervalo de frecuencias el cual es evaluado conjuntamente por el oído humano. Así, un intervalo de frecuencias corresponde a un área espacial en la membrana basilar. En total hay 24 bandas críticas asociadas a la membrana basilar, cuyo ancho de banda aumenta a medida que aumenta la frecuencia (Fastl y Zwicker, 2007). Los codificadores de audio con pérdida adoptan este modelo de grupos de frecuencias para descomponer señales de banda ancha en subbandas y codificar cada banda individualmente (Zölzer, 2005). Este modelo se adapta frecuentemente y se usa una división de frecuencias lineal de más de 24 bandas en lugar de la escala Bark.

Otra característica importante de la percepción auditiva es la sensación de intensidad de sonido, dependiente de la frecuencia, de sonidos con el mismo nivel de presión acústica. Dos características del oído resultan de esto. Por una parte, sonidos de diferente frecuencia pero con el mismo nivel de presión acústica son percibidos como con diferente intensidad del sonido, por otra parte hay un umbral dependiente de la frecuencia a partir del cual aún pueden percibirse los sonidos (Fastl y Zwicker, 2007). Este umbral también se denomina umbral auditivo absoluto o umbral auditivo en silencio y está ilustrado en la Fig. 22. A partir de esto se pueden extraer dos conclusiones para la codificación de audio. Las señales cuyos niveles están por debajo del umbral auditivo absoluto no necesitan ser procesadas ya que no pueden ser percibidas de todos modos. Aparte de ello, el número de pasos de cuantificación requeridos por banda de frecuencias puede determinarse también a partir la distancia entre el umbral auditivo en silencio y el nivel de señal (Zölzer, 2005).

Los efectos de tapado o enmascaramiento tienen una gran influencia sobre la codificación de audio. Puede diferenciarse aquí entre enmascaramiento temporal y dependiente de la frecuencia. En ambos casos se denomina enmascarador a un evento de sonido mediante el cual se tapa otro evento de sonido. Así, el evento enmascarado es inaudible. Con el enmascaramiento temporal se tapa un evento antes o después del enmascarador. El preenmascaramiento (*premasking*) es independiente de la duración del enmascarador y tapa eventos de sonido hasta 50 ms antes de percibirse el propio enmascarador (Yost, 1994). El postenmascaramiento (*postmasking*), en cambio, depende de la duración del enmascarador. Los eventos de sonido se tapan en este caso después de que termine el enmascarador. Dependiendo de la duración del enmascarador pueden pasar hasta 200 ms hasta que el oído sea receptivo de nuevo a señales en el rango del umbral auditivo en silencio (Fastl y Zwicker, 2007).

La Fig. 21 muestra una ilustración esquemática de enmascaramiento temporal. En particular, la Fig. 21 muestra de manera esquemática las regiones de pre y postenmascaramiento y el nivel respectivo por debajo del cual se tapan las señales. El enmascaramiento temporal puede usarse en codificación de audio a fin de ocultar el ruido espurio

causado por el proceso de codificación, tal como, por ejemplo, ruido de cuantificación, delante de secuencias de señal de nivel alto (transitorios).

Los efectos de enmascaramiento en el dominio de frecuencia desempeñan un papel mucho más importante que los efectos de tapado temporal. El enmascaramiento dependiente de la frecuencia describe el cambio en el umbral auditivo en silencio para sonidos individuales y sonido de banda estrecha. Estas señales distorsionan el umbral auditivo en silencio considerablemente debido a su umbral de enmascarado específico. Las señales cuyo nivel es inferior al umbral de enmascarado del enmascarador y que se encuentran en el rango efectivo de dicho umbral, no pueden percibirse (Fastl y Zwicker, 2007). Este contexto está ilustrado en la Fig. 22.

La Fig. 22 muestra una ilustración esquemática del enmascaramiento dependiente de la frecuencia en el oído humano. Como puede verse, el sonido enmascarado se sitúa por debajo del umbral de enmascarado del enmascarador y es, así, inaudible. Este efecto es aprovechado por los métodos de codificación de audio con pérdida. Las fracciones de señal por debajo del umbral de enmascarado dependiente de la frecuencia son eliminadas de la señal y no son procesadas adicionalmente (Zölzer, 2005).

La estructura general de un codificador típico adaptado a la percepción se ilustra en la Fig. 23. La Fig. 23 muestra un diagrama de bloques de un codificador de audio psicoacústico. En primer lugar, la señal PCM que va a codificarse se descompone en bandas de frecuencia mediante el banco de filtros de análisis y se alimenta al modelo psicoacústico. Aquí, un umbral de enmascarado dependiente del tiempo, el cual regula la precisión de cuantificación para las diferentes bandas de frecuencia, es determinado por las características psicoacústicas del oído descritas. Así, las bandas de frecuencia importantes, es decir las bandas de frecuencia fáciles de percibir, son cuantificadas con una muy alta resolución y las no importantes son representadas con una resolución con un menor número de bits. A continuación, al igual que en la codificación de audio sin pérdida, se realiza una codificación por entropía para la reducción de datos. Dado que tienen que transmitirse parámetros de control adicionales debido al banco de filtros de análisis y al modelo psicoacústico, la corriente de bits propiamente dicha la compone el multiplexor de corriente de bits. La ganancia de codificación en codificadores de audio con pérdida se consigue aquí al combinar cuantificación y codificación por entropía (Zölzer, 2005). Dependiendo de la calidad que deba conseguirse, la tasa de compresión se sitúa entre 4 : 1 y 50 : 1 (Weinzierl, 2008).

El decodificador está construido de manera relativamente simple. En primer lugar, la corriente de bits recibida se divide de nuevo en datos de señal y parámetros de control por un multiplexor. Después de eso, se realizan una decodificación por entropía y una cuantificación inversa. Los parámetros de control controlan a este respecto la cuantificación inversa de los datos útiles. Las señales de subbanda obtenidas de esta forma son entonces alimentadas al banco de filtros de síntesis para la reconstrucción de la señal PCM de banda ancha (Zölzer, 2005). El diagrama de bloques respectivo de un codificador de audio psicoacústico está ilustrado en la Fig. 24.

A continuación se comentarán una serie de transformaciones de señal del estado de la técnica. Dado que la cuantificación en muchos codificadores de audio está basada en un modelo de percepción que describe la percepción del ser humano en el dominio de frecuencia, es necesario transferir la señal que va a codificarse al dominio de frecuencia también. Existen para ello un gran número de transformaciones con diferentes características y campos de aplicación. Las transformaciones relevantes para codificación de audio se exponen a continuación y se muestra la estructura de un banco de filtros.

La transformada de Fourier es el método más importante para analizar la estructura armónica de una señal. Es parte del análisis de Fourier y recibe su nombre por el matemático y físico francés Jean-Baptiste-Joseph Fourier (1768 a 1830) quien fue el primero en introducirla. La transformada de Fourier es una función para transferir una señal de tiempo a su representación en el dominio de frecuencia. Es usada, entre otras cosas, para describir el comportamiento de sistemas lineales invariantes en el tiempo (LTI, por sus siglas en inglés) y para poder predecirlos (Burrus y Parks, 1985). Así, es, por ejemplo, de gran importancia en acústica y en la caracterización del oído humano. El procedimiento básico de la transformada de Fourier es descomponer una señal de tiempo en una suma ponderada de oscilaciones de seno y coseno. Para señales continuas aperiódicas se calcula de la siguiente manera (Bosi y Goldberg, 2003):

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (2.1)$$

Aquí, $x(t)$ es la señal que va a analizarse en el dominio de tiempo y $X(f)$ el respectivo espectro de Fourier en el dominio de frecuencia. Debe tenerse en cuenta que el resultado es complejo aunque se transforme una señal real. Mediante la relación de Euler en la ecuación 2.2 puede demostrarse que la parte real de $X(f)$ corresponde a los términos de coseno de $x(t)$ y que la parte imaginaria corresponde a las componentes de seno. Mediante

$$e^{-j2\pi ft} = \cos(2\pi ft) - j \operatorname{sen}(2\pi ft) \quad (2.2),$$

el resultado de la ecuación 2.1 es:

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot (\cos(2\pi ft) - j \operatorname{sen}(2\pi ft)) dt \quad (2.3)$$

5

$$= \int_{-\infty}^{\infty} x(t) \cdot \cos(2\pi ft) dt - j \int_{-\infty}^{\infty} x(t) \cdot \operatorname{sen}(2\pi ft) dt \quad (2.4)$$

que resulta en:

$$X(f) = \operatorname{Re} \{X(f)\} + j \operatorname{Im} \{X(f)\} \quad (2.5)$$

10

Dado que el seno y el coseno difieren entre sí únicamente en su fase, se puede concluir la fase de la señal a partir de la relación de los términos correspondientes. Es válido que:

$$X(f) = |X(f)| \cdot e^{j\varphi(f)} \quad (2.6)$$

15

con:

$$|X(f)| = \sqrt{(\operatorname{Re} \{X(f)\})^2 + (\operatorname{Im} \{X(f)\})^2} \quad (2.7)$$

$$\varphi(f) = \arctan \left(\frac{\operatorname{Im} \{X(f)\}}{\operatorname{Re} \{X(f)\}} \right) \quad (2.8)$$

20

Así, $|X(f)|$ se denomina respuesta en frecuencia de valor absoluto y $\varphi(f)$ se denomina respuesta en frecuencia de fase o simplemente fase.

25

Mediante la transformada de Fourier inversa (ecuación 2.9), la señal transformada es transferida de nuevo a su representación original en el dominio de tiempo. Deberá tenerse en cuenta que la transformada de Fourier y su inversa difieren por un prefactor constante y el signo de la función exponencial (Burrus y Parks, 1985).

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df \quad (2.9)$$

30

A continuación se comentará la transformada discreta de Fourier.

35

En la práctica surgen problemas en calculadoras digitales al usar la transformada de Fourier. Por un lado, esto es debido al hecho de que únicamente pueden ser procesados un número finito de valores de tiempo, y por otra parte, además de la variable de tiempo, también tiene que ser muestreada discretamente la variable de frecuencia. La solución de estos problemas es la transformada discreta de Fourier (DFT, por sus siglas en inglés). Mediante la DFT, una señal finita, discreta en el tiempo, es transferida a un espectro periódico discreto. Esto significa que es una de las transformadas más importantes en el procesamiento digital de señales. El origen de la DFT se encuentra en la transformada de Fourier y puede encontrarse una derivación precisa en (Lochmann, 1990). La DFT de una señal discreta en el tiempo $x[n]$ de longitud N se define como (Burrus y Parks, 1985):

$$X[k] = \sum_{n=0}^{N-1} x[n]W^{kn}, \quad \forall k \in [0, N-1] \quad (2.10)$$

y, de manera análoga a ello, la transformada discreta de Fourier inversa (IDFT, por sus siglas en inglés):

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k]W^{-kn}, \quad \forall k \in [0, N-1] \quad (2.11)$$

5 con el fasor de rotación complejo W :

$$W = e^{-j\frac{2\pi}{N}} \quad (2.12)$$

10 Así, $X[k]$ es el espectro periódico discreto de $x[n]$ con $\forall k, n \in \mathbb{N}$. La longitud de periodo del espectro corresponde a la longitud de transformada N y las frecuencias normalizadas son mapeadas en el intervalo $[0, 2\pi]$.

15 Para señales de entrada reales, la DFT tiene una característica importante. Aquí no se calculan N coeficientes de frecuencia independientes, como en el caso general, sino únicamente la mitad. Esta característica puede aprovecharse por ejemplo para el almacenamiento o transmisión de los datos. Para la transformada inversa, los segundos $N/2$ valores son calculados entonces mediante la siguiente correlación (Rao y Yip, 2001):

$$X[N-k] = X[k]^* \quad (2.13)$$

20 El operador $*$ en la fórmula 2.13 caracteriza a este respecto la conjugada compleja. Así, $X(k)^*$ es la secuencia de valores conjugados complejos para $X(k)$.

25 La complejidad de cálculo de la DFT y la IDFT es de N^2 multiplicaciones y adiciones complejas. Si se usan simetrías en el cálculo, el número de pasos de cálculo requeridos se reduce a $N \log N$ y la complejidad corresponde a $O(\log N)$. Sin embargo, en los métodos rápidos, la longitud de transformada N tiene que corresponder a una potencia del número dos. La transformada rápida de Fourier se denomina FFT (por sus siglas en inglés) (Kiencke y Jäkel, 2005).

30 La transformada discreta de Fourier no ha ganado aceptación en el campo de la compresión de datos. Las grandes desventajas de la DFT son una complejidad de cálculo alta y la redundancia contenida en el espectro. Aunque existen métodos eficaces para cálculo de la DFT en forma de FFT, el resultado será siempre un espectro complejo. Esto significa que a partir de N valores de transformada se calculan N pares de valores complejos. Además, únicamente los primeros $N/2$ valores espectrales contienen información nueva.

A continuación se comentan las transformadas de coseno y de seno discretas.

35 La transformada de coseno discreta (DCT, por sus siglas en inglés) es una solución para los problemas de la DFT mencionados antes. La DCT es una transformada real, discreta, lineal y ortogonal. Debido a estas mismas características, es la transformada usada más frecuentemente en compresión de datos digital (Britanak *et al.*, 2007).

40 La DCT es una transformada trigonométrica discreta. En total se diferencian ocho formas de DCT. Dependiendo de su extensión se dividen en transformadas pares e impares, así como en los tipos I, II, III y IV. Sin embargo, para procesamiento digital de señales, únicamente los tipos pares de DCT son de importancia. Estos se enumeran a continuación (Rao y Yip, 2001):

$$X^I[k] = \varepsilon[k] \sum_{n=0}^N \varepsilon[n]x[n] \cos\left(\frac{\pi nk}{N}\right), \quad \forall k \in [0, N] \quad (2.14a)$$

$$X^{II}[k] = \varepsilon[k] \sum_{n=0}^{N-1} x[n] \cos\left(\frac{\pi(n+0,5)k}{N}\right), \quad \forall k \in [0, N-1] \quad (2.14b)$$

$$X^{III}[k] = \sum_{n=0}^{N-1} \varepsilon[n]x[n] \cos\left(\frac{\pi(k+0,5)n}{N}\right), \quad \forall k \in [0, N-1] \quad (2.14c)$$

$$X^{IV}[k] = \sum_{n=0}^{N-1} x[n] \cos\left(\frac{\pi(n+0,5)(k+0,5)}{N}\right), \quad \forall k \in [0, N-1] \quad (2.14d)$$

con:

$$\varepsilon[p] = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } p = 0 \vee p = N \\ 1 & \text{si no} \end{cases} \quad (2.15)$$

10 Cada una de estas fórmulas tiene su propia aplicación en codificación. La DCT-II se usa principalmente como transformada para datos de imagen. En la literatura es el primer tipo de DCT descrito. Esto significa que generalmente el término "DCT" se refiere a DCT-II (Ahmed *et al.*, 1974). Excepto por un prefactor, la DCT-III es la transformada inversa de la DCT-II y viceversa. Para codificación de audio, la DCT-IV es de especial importancia. Es la base de la transformada de coseno discreta modificada.

15 A fin de poder mostrar características importantes de la DCT, a continuación se expone una correlación entre DFT y DCT. Como se ha ilustrado antes, la DFT calcula, a partir de una señal de valor real de longitud N , únicamente $N/2$ coeficientes de frecuencia independientes. A la inversa, esto significa que se necesitan $2N$ valores en el dominio de tiempo para obtener N valores espectrales. Sin embargo, si únicamente hay N valores de tiempo disponibles, la señal tiene que extenderse de manera apropiada. La ampliación simétrica mediante reflejo de la señal completa parece ser apropiado aquí. La señal ampliada parece repetirse por tanto con una longitud de periodo de $2N$. Esto tiene la ventaja de que se suprime el molesto efecto de fuga de la DFT con señales cortadas (Kiencke y Jäkel, 2005).

Cualquier señal real $x[n]$ de longitud N se amplía simétricamente, siendo el resultado:

$$\tilde{x}[n] = [x[0], \dots, x[N-1], x[N-1], \dots, x[0]] \quad (2.16)$$

con $0 \leq n \leq 2N-1$. La longitud de $\tilde{x}[n]$ es por tanto $2N$. La DFT de la ecuación 2.10 se aplica ahora sobre esta señal con la ecuación 2.12 y se convierte (Rao y Yip, 2001). Una derivación detallada puede encontrarse en el anexo A.1. Es válido que:

$$\tilde{X}[k] = \sum_{n=0}^{2N-1} \tilde{x}[n] e^{-j\frac{2\pi}{2N}kn} \quad (2.17a)$$

$$= 2e^{j\frac{\pi}{2N}k} \sum_{n=0}^{N-1} \tilde{x}[n] \cos\left(\frac{\pi(n+0,5)k}{N}\right), \quad \forall k \in [0, N-1] \quad (2.17b)$$

35 Cuando se compara este resultado con la DCT-II en la ecuación 2.14b, se puede observar que estas dos ecuaciones difieren únicamente en el término de fase $2e^{j\frac{\pi}{2N}k}$. Dado que éste es independiente de la señal y no contiene ninguna información, puede ignorarse para el cálculo de la DCT (Rao y Yip, 2001). Para la DCT-I puede mostrarse una correlación similar, pero con una extensión de señal diferente de $x[n]$. La DCT-IV se obtiene entonces a partir de una rotación de fase de la función base de la DCT-II. Una derivación detallada de esto puede encontrarse en (Rao y Yip, 2001).

Pueden sacarse algunas conclusiones de este resultado. En primer lugar, se observa que la DCT, a diferencia de la DFT, es una transformada puramente real. Dos ventajas resultan de esto. Primero, no se tienen que realizar multiplicaciones y adiciones complejas para el cálculo y, segundo, únicamente se necesita la mitad del espacio de almacenamiento para almacenar los datos ya que no hay pares de valores complejos. Asimismo, se observa que la DCT, para calcular N coeficientes de frecuencia independientes, requiere exactamente N valores para la transformación. Las frecuencias se sitúan todas en el intervalo $[0, \pi]$. A diferencia de la DFT, la redundancia contenida en el espectro para señales de entrada de valor real se ha desvanecido y así la resolución de frecuencia es el doble de alta. Sin embargo, es desventajoso que el espectro de la DCT no pueda convertirse en valor absoluto y fase. Además, puede suceder que en la señal estén contenidas frecuencias que si bien corresponden a las funciones base de la DCT (véanse las ecuaciones 2.14a a 2.14d), están rotadas sin embargo 90° en relación con las mismas en su fase. Estas frecuencias no son representadas entonces por la DCT, es decir el respectivo coeficiente DCT es cero. Por estas razones, la DCT es apropiada para una compresión de datos eficaz y rápida, pero menos para análisis de señal (Malvar, 1992).

Aparte de la transformada de coseno discreta, existe también la transformada de seno discreta (DST, por sus siglas en inglés). En total se diferencian ocho formas de DST. Únicamente la DST-IV es de importancia aquí. Con respecto a su forma y características, corresponde a la DCT-IV (Rao y Yip, 2001):

$$X_s^{IV}[k] = \sum_{n=0}^{N-1} x[n] \sin\left(\frac{\pi(n+0,5)(k+0,5)}{N}\right), \quad \forall k \in [0, N-1] \quad (2.18)$$

Cuando una señal se transforma tanto con la DCT-IV como con la DST-IV, el espectro complejo formado por la combinación de los dos espectros reales de nuevo contiene información sobre valor absoluto y fase. La resolución en frecuencia aquí es aun el doble de grande que en la DFT, lo cual significa que se mapean N frecuencias en el intervalo $[0, \pi]$ (Malvar, 1992).

Para el procesamiento de señales de audio largas, no es posible transformar la señal como un todo. Por una parte, la complejidad de cálculo se incrementa con ello enormemente ya que se requieren N^2 operaciones de cálculo también para calcular la DCT. Por otra parte, no es posible procesar la señal en tiempo real ya que tiene que esperarse siempre a la transmisión de la corriente de datos completa antes de que la señal pueda reconstruirse. Por consiguiente, es necesario dividir la señal en bloques. En este caso, la DCT se aplica como denominada transformada de bloques (Rao y Yip, 2001). Con el índice de bloque $b \in \mathbb{N}$ se obtiene para la DCT-IV a partir de la ecuación 2.14d:

$$X_b^{IV}[k] = \sum_{n=0}^{N-1} x[n + bN] \cos\left(\frac{\pi(n+0,5)(k+0,5)}{N}\right), \quad \forall k \in [0, N-1] \quad (2.19)$$

La longitud de señal de $x[n]$ corresponde aquí a bN . Con transformadas de bloque aparecen artefactos de bloque debido a la cuantificación. Un ejemplo conocido en el que pueden observarse artefactos de este tipo es el método de compresión JPEG. Los artefactos de bloque se originan por las extensiones que han de realizarse para la periodización. No se corresponden con las extensiones de señal asumidas originalmente (véase la ecuación 2.16). El resultado son saltos en los límites de bloque que en el dominio de frecuencia desplazan la energía hacia frecuencias altas (Malvar, 1992). Los saltos en una señal de audio pueden percibirse como crujidos. El oído humano es muy sensible a tales artefactos. Así, estos deben evitarse absolutamente.

La transformada de coseno discreta modificada se explicará a continuación.

La transformada de coseno discreta modificada (MDCT, por sus siglas en inglés) es la transformada central para compresión de audio. Es usada entre otros, en mp3, codificación de audio avanzada o AAC (por sus siglas en inglés) y Dolby Digital (ac-3). La MDCT es una transformada real, discreta y ortogonal y una modificación de la DCT-IV. Se define de la siguiente manera (Rao y Yip, 2001):

$$X_b[k] = \sum_{n=0}^{2N-1} x[n + bN] \cos\left(\frac{\pi(n+0,5-\frac{N}{2})(k+0,5)}{N}\right), \quad \forall k \in [0, N-1] \quad (2.20)$$

Una ventaja de la MDCT con respecto a la DCT-IV es que se evitan los artefactos de bloque. Esto puede

conseguirse principalmente mediante la superposición de varios bloques sucesivos. Este tipo de transformada también es conocida como transformada ortogonal superpuesta (LOT, por sus siglas en inglés) (Malvar y Staelin, 1989).

5 La redundancia puede eliminarse de nuevo mediante el método de superposición-adición (OLA, por sus siglas en inglés). Así, los bloques que se forman en la transformada inversa se solapan hasta un 50 % y se suman, este procedimiento se denomina superposición-adición.

10 La resolución en frecuencia de la MDCT puede mejorarse de manera adicional mediante la ponderación de la secuencia de entrada $x[n + bN]$ con una función ventana. En la ecuación 2.20, la ventana corresponde a una función rectangular que recorta el bloque actual b a partir de la señal global. En el dominio de frecuencia, esto corresponde a una convolución con la función sinc . La atenuación de banda de rechazo pobre de la función sinc puede mejorarse mediante adaptación de esta función ventana y así lograrse una selectividad de frecuencia incrementada. Para que la MDCT sea capaz de reconstruir perfectamente, la función ventana $w[n]$ de longitud $2N$ tiene que cumplir las condiciones Princen-Bradley (PR) (Princen *et al.*, 1987):

$$w[n] = w[2N - 1 - n] \quad (2.21a)$$

$$w^2[n] + w^2[n + N] = 1 \quad (2.21b)$$

20 Una ventana simple que cumpla estas condiciones y exhiba suficiente atenuación de banda de rechazo es la ventana de mitad de onda sinusoidal. Esta se usa, entre otros, en mp3 y AAC y se define de la siguiente manera (Malvar, 1992):

$$w[n] = \text{sen} \left(\frac{\pi(n + 0,5)}{2N} \right), \quad \forall n \in [0, 2N - 1] \quad (2.22)$$

25 Al insertar la función ventana $w[2N - 1 - n]$ en la ecuación 2.20, puede observarse otra característica importante de la MDCT. El resultado corresponde a la convolución discreta de $x[n + bN]$ con la función ventana modulada $w[n]$. Así, para $\forall k \in [0, N - 1]$, es válido que (Schuller y Smith, 1996):

$$X_k[b] = \sum_{n=0}^{2N-1} x[n + bN] w[2N - 1 - n] \cos \left(\frac{\pi(n + 0,5 - \frac{N}{2})(k + 0,5)}{N} \right) \quad (2.23)$$

30 Así, la MDCT no puede verse únicamente como transformada de bloque, sino también como banco de filtros modulado (Malvar, 1992). Así, la función ventana corresponde al filtro FIR de prototipo de paso bajo el cual se modula con el núcleo de coseno y así representa las bandas de frecuencia del banco de filtros. El resultado de esto es que la secuencia de entrada $x[n + bN]$ se descompone en exactamente N subbandas. En relación con la característica de TDA, la MDCT cumple las condiciones de un denominado "banco de filtros muestreado de manera crítica".

35 Tal banco de filtros muestreado de manera crítica está ilustrado en la Fig. 25. En particular, la Fig. 25 muestra un banco de filtros PR muestreado críticamente de N bandas con un retardo de sistema de n_d muestras. Tales bancos de filtros son de importancia particular para la codificación de audio ya que describen una señal de la manera más precisa y completa posible con el menor número de muestras (Rao y Yip, 2001).

40 El símbolo $\downarrow N$ corresponde a una reducción de la tasa de muestreo por el factor $1/N$ y $\uparrow N$ a un incremento por el factor N . la señal después del banco de filtros de síntesis $\hat{x}[n] = x[n - n_d]$ es idéntica a la señal de entrada $x[n]$ antes del banco de filtros de análisis, excepto por un retardo constante de n_d muestras. En el caso de la MDCT, $hk[n]$ es la función de ventana modulada $wk[n]$. Dado que $w[n]$ cumple las condiciones PR, los filtros de análisis h_k son idénticos a los filtros de síntesis g_k .

45 Desde un punto de vista matemático, es apropiado formular sistemas de ecuaciones lineales, entre los que se incluyen todas las transformadas mencionadas hasta ahora, en la notación vector-matriz. Una señal $x[n]$ de longitud

bN se representa como un vector de columna $\underline{\mathbf{x}} = [x[0], x[1], \dots, x[bN - 1]]^T$. El operador T identifica la transposición. La formación de bloque puede representarse como una matriz en la cual cada columna de la matriz contiene un bloque de $x[n]$:

$$\underline{\mathbf{x}} = \begin{bmatrix} x[0] & x[N] & \dots & x[(b-1)N] \\ x[1] & x[N+1] & \dots & x[(b-1)N+1] \\ \vdots & \ddots & \dots & \vdots \\ x[N-1] & x[2N-1] & \dots & x[bN-1] \end{bmatrix} \quad (2.24)$$

La regla de transformación también puede representarse como una matriz. Las funciones ventana moduladas forman aquí las líneas de la matriz. Para $\forall k \in [0, N - 1], \forall n \in [0, 2N - 1]$, es válido lo siguiente:

$$\underline{\mathbf{T}}_{MDCT}(k, n) := w[n] \cos\left(\frac{\pi(n + 0,5 - \frac{N}{2})(k + 0,5)}{N}\right) \in \mathbb{R}^{N \times 2N} \quad (2.25)$$

Para poder calcular la MDCT de $\underline{\mathbf{x}}$, la estructura de bloques de $\underline{\mathbf{X}}$ debe ampliarse en un 50% de superposición para TDA. Así, la MDCT puede escribirse de la siguiente forma:

$$\underline{\mathbf{X}} = \underline{\mathbf{T}}_{MDCT} \cdot \underline{\mathbf{X}}_{TDA} \quad (2.26)$$

con:

$$\underline{\mathbf{X}}_{TDA} = \begin{bmatrix} x[0] & x[N] & \dots & x[(b-2)N] \\ x[1] & x[N+1] & \dots & x[(b-1)N+1] \\ \vdots & \vdots & \ddots & \vdots \\ x[2N-1] & x[3N-1] & \dots & x[bN-1] \end{bmatrix} \quad (2.27)$$

Cada columna de $\underline{\mathbf{X}}$ forma el espectro de la MDCT del bloque en cuestión con el índice b en $\underline{\mathbf{x}}$.

Para calcular un bloque, esta fórmula de la MDCT requiere $2N^2$ multiplicaciones y adiciones. Sin embargo, la complejidad de cálculo puede reducirse considerablemente.

Para ello es necesario transferir el banco de filtros en la Fig. 25 a un banco de filtros polifásico equivalente (véase la Fig. 26). Mediante la representación polifásica y la transformada z , los sistemas multitasa, como el banco de filtros de la MDCT, pueden analizarse más ampliamente.

Un filtro FIR $h[n]$ puede dividirse siempre en $M \in \mathbb{N}$ fases cuando la longitud del filtro corresponde a un múltiplo entero de M . La fase m -ésima $p_m[n]$ de $h[n]$ se produce al retardar $h[n]$ con z^m y reducir la tasa de muestreo por el factor M (Malvar, 1992). Es válido que:

$$p_m[n] = h[nM + m] \quad (2.28)$$

Con esta descomposición y la transformada z , el filtro $h[n]$ puede representarse de la siguiente manera (Malvar, 1992):

$$H[z] = \sum_{n=0}^{MN-1} h[n]z^{-n} \quad (2.29)$$

$$= \sum_{m=0}^{M-1} z^{-m} \sum_{n=0}^{N-1} h[nM + m]z^{-nM} \quad (2.30)$$

5 En vez de notación de suma, la notación de vector es preferida aquí también. La ecuación 2.30 puede representarse por tanto como un vector de dimensión N:

$$\underline{\mathbf{H}} = \begin{bmatrix} H_0[z] \\ H_1[z] \\ \vdots \\ H_{N-1}[z] \end{bmatrix} \quad (2.31)$$

con:

10

$$H_n[z] = \sum_{m=0}^{M-1} h[nM + m]z^{-nM-m} \quad (2.32)$$

Esta descomposición polifásica puede ahora aplicarse a cada filtro del banco de filtros de la MDCT. El resultado es la representación polifásica equivalente del banco de filtros ya mencionada, en la Fig. 26 (Schuller y Smith, 1996). Así, la Fig. 26 representa un banco de filtros polifásico PR muestreado de manera crítica de N bandas equivalente.

15

Al hacer uso de las simetrías en el núcleo de la MDCT y la característica TDA, las matrices de filtros polifásicas de análisis y de síntesis $\underline{\mathbf{P}}_a$ y $\underline{\mathbf{P}}_s$ pueden dividirse cada una en una matriz de convolución (*folding*) ocupada escasamente y una matriz de transformación (Schuller y Smith, 1996). Las matrices de convolución $\underline{\mathbf{F}}_a$ y $\underline{\mathbf{F}}_s$ tienen una estructura de diamante con los coeficientes de la función ventana $w[n]$ como polinomios en el dominio z. Pueden descomponerse adicionalmente en una matriz de ventana y una matriz de retardo:

20

$$\underline{\mathbf{F}}_a = \underline{\mathbf{D}} \cdot \underline{\mathbf{F}} \quad (2.33a)$$

$$\underline{\mathbf{F}}_s = \underline{\mathbf{F}}^T \cdot \underline{\mathbf{D}}^{-1} \quad (2.33b)$$

25

La forma precisa y la división de las matrices de convolución se muestran más adelante. Las matrices de transformación corresponden a la matriz DCT-IV:

30

$$\underline{\mathbf{T}}(k, n) := \cos\left(\frac{\pi(n+0,5)(k+0,5)}{N}\right) \in \mathbb{R}^{N \times N} \quad (2.34a)$$

$$\underline{\mathbf{T}}^{-1} = \frac{2}{N} \cdot \underline{\mathbf{T}} \quad (2.34b)$$

Con estas matrices, el espectro de la MDCT de la señal de entrada dividida en bloques \underline{x} se calcula de la siguiente manera (Schuller y Smith, 1996):

$$\underline{X} = \underline{T} \cdot \underline{D} \cdot \underline{F} \cdot \underline{x} \quad (2.35)$$

5 y para la transformada inversa es válido que:

$$\underline{\hat{x}} = \underline{F}^T \cdot \underline{D}^{-1} \cdot \underline{T}^{-1} \cdot \underline{X} \quad (2.36)$$

10 Esta solución ofrece varias ventajas con respecto a calcular la MDCT según la ecuación 2.26. En primer lugar, el solapamiento de dominio de tiempo (*time domain aliasing*) que se forma puede reconocerse más fácilmente. Con la representación polifásica de la matriz de convolución en la ecuación 2.33a, el proceso puede interpretarse como entregar fracciones de señal ponderadas del bloque $(b - 1)$ al bloque actual b . Mediante la adición de estas fracciones de señal se forma TDA. La mayor ventaja de calcular la MDCT usando polifases es la complejidad de cálculo reducida considerablemente. Al usar la matriz de DCT-IV cuadrática y la matriz de convolución ocupada escasamente, la complejidad de cálculo se reduce a $N(N + 2)$ multiplicaciones y adiciones. Al usar implementaciones rápidas de la DCT, de manera similar a la FFT, el número de operaciones requerido puede disminuirse hasta $N(\log N + 2)$ y así la complejidad se reduce a $O(N \log N)$ (Rao y Yip, 2001). Por estas razones, la MDCT se considera aquí como implementada de acuerdo con el enfoque polifásico.

20 En el procesamiento de señales de audio puede ser necesario desplazar una señal de una frecuencia baja a frecuencias más altas, debiendo ser este desplazamiento en frecuencia seleccionable libremente y preciso. Los codificadores de audio que quieran restaurar las frecuencias más altas de una señal tienen que encarar este problema. Las tecnologías de codificación de audio modernas usan métodos de ampliación de ancho de banda para una compresión más eficaz de datos de audio. Aparte de las características psicoacústicas del oído humano, se aprovecha la correlación de fracciones de señal de baja frecuencia con respecto a las fracciones de alta frecuencia para la reducción de datos.

30 Un ejemplo de implementación de una ampliación de ancho de banda se describe en el documento US2012065983 A1.

A pesar de la existencia de numerosas formas de reducir la tasa de transmisión de datos mediante codificación de audio, los codificadores de audio actuales encuentran limitaciones cuando se desean tasas de transmisión de bits bajas. En particular los métodos psicoacústicos en este caso producen falseamientos de señal indeseados. Esto puede verse en artefactos de interferencia, como alturas perdidas, transitorios borrosos o siseo artificial de la señal de audio reproducida. En muchos casos de aplicación, sin embargo, únicamente está disponible un ancho de banda de transmisión limitado. La ampliación de ancho de banda (*bandwidth extension*, BWE) ofrece una solución para estos problemas. Generalmente, la ampliación de ancho de banda aún una serie de métodos con los que puede ensancharse espectralmente una señal de audio de banda limitada a fin de obtener de nuevo el ancho de banda original. Se diferencian en total cuatro categorías de métodos para ampliación de ancho de banda (Larsen y Aarts, 2004). Estos se ilustran de manera gráfica en la Fig. 27.

45 La Fig. 27 muestra categorías de ampliación de ancho de banda (Larsen y Aarts, 2004). En la Fig. 27 se muestra el BWE psicoacústico de baja frecuencia en la esquina superior izquierda. En la Fig. 27 puede verse el BWE psicoacústico de alta frecuencia en la esquina superior derecha. En la Fig. 27 se ilustra el BWE de baja frecuencia en la esquina inferior izquierda. Además, en la Fig. 27 se ilustra el BWE de alta frecuencia en la esquina inferior derecha. La energía de la banda 'a' (línea discontinua) es desplazada a la banda 'b' (línea de puntos).

50 Solo la categoría III (esquina inferior derecha en la Fig. 27) es útil para grabación de audio. En el denominado "BWE de alta frecuencia" se usan las frecuencias presentes en la señal de banda limitada a fin de reconstruir el rango de alta frecuencia del espectro. La idea de usar un método de este tipo para la ampliación de ancho de banda de señales de audio está basada en el hecho de que hay una fuerte correlación entre la fracción de alta frecuencia y de baja frecuencia de la señal. Así, es posible reconstruir las frecuencias altas que faltan a partir de las fracciones de señal baja presentes (Larsen y Aarts, 2004). A continuación se exponen técnicas y métodos actuales con los que es posible ampliar una señal de banda limitada a su ancho de banda original por medio de BWE de alta frecuencia.

La replicación de banda espectral (*spectral band replication*, SBR) es conocida por el estado de la técnica, tal como se usa, entre otras cosas, en HE-AAC. En la replicación de banda espectral con SBR se aprovechan las

correlaciones entre fracciones de señal de baja frecuencia y alta frecuencia a fin de ensanchar espectralmente la señal de paso bajo proporcionada por el codificador. Las bandas de baja frecuencia del banco de filtros subyacente son copiadas a las bandas altas que faltan y se adapta la envolvente espectral. Este proceso de copiado causa, en particular con frecuencias de corte bajas, artefactos capaces de percibirse tales como rugosidad acústica y cambios indeseados en timbre. Estos son causados principalmente por la extensión armónica del espectro que falta en el límite entre la banda base y las bandas de alta frecuencia producidas de manera algorítmica.

Un codificador de audio SBR del estado de la técnica usa descomposición en subbandas pQMF de la señal y de esta forma asegura una alta eficacia de codificación [Eckstrand, 2002]. Esto se consigue al transmitir únicamente las bandas de frecuencia más bajas, mientras que las fracciones de frecuencia más altas son reconstruidas usando información lateral y el desplazamiento en frecuencia de las bandas más bajas mencionado anteriormente.

La replicación de banda espectral es en la actualidad el método más generalizado para ampliación de ancho de banda. Se emplea, entre otras cosas, en HE-AAC y mp3PRO. SBR fue desarrollado por la empresa Coding Technologies, con el objetivo de incrementar la eficacia de los codificadores de audio existentes. Esto se consigue al procesar, mediante un codificador, únicamente las frecuencias por debajo de una determinada frecuencia de corte f_0 . En los ejemplos mencionados se utilizan codificadores de mp3 y AAC como codificadores de núcleo. Las frecuencias por encima de la frecuencia de corte se describen únicamente mediante unos pocos parámetros. En función de la calidad que haya de alcanzarse, esta se sitúa entre 5 kHz y 13 kHz. Las fracciones de alta frecuencia son entonces reconstruidas en el receptor con ayuda de esta información lateral y la señal de banda limitada decodificada (Eckstrand, 2002).

La Fig. 28 muestra el diagrama de bloques de un codificador SBR ampliado. La tasa de muestreo de la señal de entrada se disminuye y a continuación se alimenta al codificador propiamente dicho. En paralelo, la señal es analizada por un banco de filtros de espejo en cuadratura complejo (*quadrature-mirror-filterbank*, QMF) y se realiza un cálculo de energía. El QMF usado consiste en 64 subbandas. Los parámetros requeridos para estimar las envolventes espectrales pueden derivarse de las mismas. Parámetros adicionales permiten reaccionar a las características especiales de la señal de entrada. Al conocer el codificador SBR la producción de la banda de alta frecuencia puede reconocer diferencias fuertes entre la fracción de alta frecuencia (HF, por sus siglas en inglés) original y la sintetizada.

Cuando, por ejemplo, sonidos individuales perceptibles fuertemente por encima de la frecuencia de corte están presentes en la señal, estos se describen mediante parámetros adicionales y pueden alimentarse de nuevo a la señal reconstruida. La información lateral producida es insertada en la corriente de bits de salida, junto a los datos de audio propiamente dichos (Larsen y Aarts, 2004).

La Fig. 29 muestra el diagrama de bloques del decodificador respectivo ampliado por SBR. Los datos de audio de banda limitada son decodificados por el decodificador y los parámetros de control son extraídos de la corriente de bits. A continuación, los datos de audio son alimentados de nuevo a un banco de filtros QMF para reconstruir las fracciones de alta frecuencia. La banda base es copiada dentro de este banco de filtros e insertada por encima de la frecuencia de corte (véase la Fig. 30, a la izquierda).

La Fig. 30 es una ilustración esquemática de la respuesta en frecuencia de valor absoluto. Así, la Fig. 30 es una ilustración esquemática de la reconstrucción de HF por SBR. La Fig. 30 muestra, a la izquierda, un copiado y desplazamiento de la banda base. La Fig. 30 ilustra, a la derecha, un espectro después de adaptar la envolvente espectral.

La información, producida en el codificador SBR, sobre la envolvente espectral se usa ahora para ajustar la curva envolvente del espectro copiado a la original. Esta adaptación se hace usando el parámetro de control transmitido y la energía de la banda QMF respectiva. Si las características del espectro reconstruido difieren de las originales, se añaden adicionalmente componentes tonales o ruido a la señal (Larsen y Aarts, 2004). La Fig. 30 muestra a la derecha el espectro reconstruido adaptado.

Finalmente, la señal de banda limitada y la señal de alta frecuencia reconstruida son mezcladas y transferidas al dominio de tiempo mediante el banco de filtros de síntesis. De esta forma, se ha formado una señal de ancho de banda ampliado que está lista ahora para la reproducción.

En esta forma de ampliación de ancho de banda surgen problemas con las señales altamente tonales con una estructura armónica altamente perceptible. Incluso aunque el método SBR prevé técnicas para la adaptación tonal del espectro, estas no son suficientes para restaurar una estructura armónica destruida. El resultado es una rugosidad acústica perceptible en la señal (Wilde, 2009). Estos artefactos son muy desagradables para el oyente. El origen se encuentra en el proceso de copiado del codificador SBR. No se tiene en cuenta la estructura fina armónica de la señal y simplemente se replica la banda base. El resultado puede verse en la Fig. 31.

La Fig. 31 muestra una destrucción de la estructura armónica con SBR. La Fig. 31 muestra a la izquierda un espectro de banda ancha original. La Fig. 31 muestra a la derecha un espectro después de la reconstrucción de HF por SBR.

5 Es claramente reconocible que los armónicos están desplazados con respecto al espectro original en el rango por encima de la frecuencia de corte. El espectro HF reconstruido es armónico, pero la estructura armónica está afectada por una desviación de frecuencia de cresta adicional f_{lag} a la frecuencia de corte. Adicionalmente, las relaciones de amplitud de los subtonos armónicos se distorsionan al reconstruir la curva envolvente. Este efecto ocurrirá con todas las señales armónicas, tales como las generadas por ejemplo por instrumentos musicales.

10 Para señales armónicas, tales como, por ejemplo, las de un diapasón, SBR y métodos de ampliación de ancho de banda equivalentes producen artefactos indeseados, tales como, por ejemplo, rugosidad acústica y timbres desagradables, ya que la estructura armónica de la señal no se mantiene completamente. Para señales que presentan una estructura armónica pronunciada se producen artefactos indeseados, como rugosidad acústica y cambios de timbre, cuando se aplica SBR.

15 Se han desarrollado por tanto dos métodos de ampliación de ancho de banda en el dominio tiempo, los cuales obtienen estas estructuras: la ampliación de ancho de banda armónica (*harmonic bandwidth extension*, HBE) controlada mediante vocoder de fase y la BWE por modulación continua (*continuous modulation*, CM) que usa modulación de banda lateral individual [Nagel y Disch 2009], [Nagel *et al.*, 2010]. Debido a la modulación continua con frecuencias seleccionables libremente, en particular la BWE por CM logra una buena restauración armónica.

20 Existen métodos de ampliación de ancho de banda alternativos que evitan el problema de extensión espectral no armónica. Dos de estos métodos se presentaran a continuación. Básicamente, estos métodos reemplazan el generador de HF del decodificador SBR en la Fig. 29 y representan por tanto una alternativa al simple proceso de copiado. La adaptación de la envolvente espectral y de la tonalidad permanece sin cambios. Dado que la señal de entrada debe estar en el dominio de tiempo, este método también se denomina método en el dominio de tiempo para la ampliación de ancho de banda.

25 La ampliación de ancho de banda armónica puede mencionarse primero. La ampliación de ancho de banda armónica (*harmonic bandwidth extension*, HBE) usa un vocoder de fase para producir el rango agudo. El espectro se estira al aplicar un vocoder de fase. Como se muestra a la izquierda en la Fig. 32, la banda base se expande hasta la frecuencia de señal máxima f_{max} y se recorta el intervalo de frecuencias entre la frecuencia de corte y f_{max} . El espectro se compone entonces de esta fracción y de la banda base (véase la Fig. 32, a la derecha). La envolvente se adapta al igual que en SBR (Nagel y Disch, 2009).

30 La Fig. 32 es una ilustración esquemática de la reconstrucción de HF por HBE. La Fig. 32 muestra a la izquierda el estiramiento de la banda base por el factor 2. La Fig. 32 muestra a la derecha un espectro después de haber adaptado la envolvente espectral.

35 Usando factores de estirado de número entero $\sigma \in \mathbb{N}^+$ se asegura que la frecuencia de corte f_g no cambia la estructura armónica. Es válido que:

$$f_{max} = \sigma \cdot f_g \quad (3.1)$$

40 Resulta desventajoso el hecho de que la distancia entre los subtonos en el rango de HF cambia con el factor de estirado debido a la expansión del espectro, como puede verse en la Fig. 33. Además, se requieren cálculos complicados para expandir el espectro. Entre estos están la DFT de alta resolución, la adaptación de la fase y la conversión de la tasa de muestreo (Dolson, 1986). Cuando la señal de audio está subdividida en bloques, es necesaria adicionalmente una estructura de superposición-adición a fin de poder extender la fase de bloques colindantes de manera continua. Para señales altamente tonales, se pueden conseguir muy buenos resultados al usar la técnica de vocoder de fase, sin embargo en señales de percusión los transitorios se difuminan y se vuelve necesario realizar un tratamiento transitorio separado (Wilde, 2009).

45 La Fig. 33 muestra una estructura armónica con HBE. La Fig. 33 muestra a la izquierda un espectro de banda ancha original. La Fig. 33 ilustra a la derecha un espectro después de la reconstrucción de HF por HBE.

A continuación se expone la modulación de banda lateral única continua.

50 La modulación de banda lateral única continua (*continuously modulated bandwidth extension*, CM-BWE) es otro método en el dominio de tiempo para la ampliación de ancho de banda. En este método, la banda base es modulada con la frecuencia f_{mod} por medio de modulación de banda lateral única y así es desplazada a otra ubicación

espectral, como se ilustra en la Fig. 34. Una frecuencia de modulación variable asegura el mantenimiento de la estructura armónica de la señal de ancho de banda ampliado. Con frecuencias de modulación mayores que la frecuencia de corte f_g , la brecha que se forma en el espectro tiene que llenarse con ruido (Nagel *et al.*, 2010).

5 La Fig. 34 muestra una ilustración esquemática de la reconstrucción de HF por CM-BWE. La Fig. 34 muestra a la izquierda una modulación de la banda base por la frecuencia f_{mod} . La Fig. 34 muestra a la derecha un espectro después de adaptar la envolvente espectral.

10 Aparte del caso ilustrado en la Fig. 34, también puede ser necesario modular la banda base varias veces. En tal caso, la frecuencia de modulación tiene que ser adaptada para cada modulación en la cual se selecciona su respectivo siguiente múltiplo entero (Nagel *et al.*, 2010). Antes de la modulación, la banda base tiene que filtrarse paso bajo conforme a la frecuencia de modulación, a fin de que la frecuencia de señal permitida máxima f_{max} no se exceda después de la modulación. De manera similar a los métodos ya expuestos, a continuación se conforma la envolvente espectral y se adapta la tonalidad.

15 La Fig. 35 muestra la estructura armónica tal como se forma en una señal ampliada por medio de CM-BWE. La Fig. 35 muestra a la izquierda un espectro de banda ancha original. La Fig. 35 muestra a la derecha un espectro después de la reconstrucción de HF por CM-BWE. Como en el método HBE, en CM-BWE no hay un tono armónico en el espectro. Sin embargo, esto no atrae la atención en una forma negativa, ya que la estructura armónica en sí misma se mantiene.

20 Las desventajas de este método radican en tener que calcular la modulación de banda lateral única. Se requiere una señal analítica para un cálculo correcto, es decir una señal que contenga únicamente frecuencias positivas. Un transformador Hilbert es necesario para calcular esta señal. Se trata básicamente de un filtro no causal de respuesta de impulso infinito. No es posible implementar un filtro de este tipo y tiene que simplificarse. Sin embargo a fin de conseguir una atenuación de banda de rechazo lo más alta posible con un orden de filtro mínimo, se añade un retardo no insignificante a la señal mediante causalización del filtro (Wilde, 2009).

25 Sin embargo, cuando se realiza el desplazamiento de frecuencia en el dominio de tiempo, esto puede ser muy complejo. Realizar el desplazamiento en el dominio de subbanda de un codificador de audio de subbanda en cambio puede hacer que la resolución de frecuencia sea demasiado basta para el desplazamiento en frecuencia requerido.

30 Lo que se desea es minimizar el espacio de memoria requerido por datos digitales o el ancho de banda requerido para transmitir estos datos mediante la codificación de señales de audio. Al mismo tiempo, la calidad percibida de la señal de audio reproducida debe ser comparable al estándar CD (frecuencia de muestreo 44100 Hz con una profundidad de cuantificación de 16 bits). Así, la calidad ha de maximizarse con una tasa de transmisión de datos decreciente.

35 Por tanto, un objetivo de la presente invención es proporcionar diseños mejorados para desplazamientos de frecuencia seleccionables libremente en el dominio de subbanda. El objetivo de la presente invención se consigue mediante un dispositivo según la reivindicación 1, mediante un método según la reivindicación 23 y mediante un programa informático según la reivindicación 24.

40 Se proporciona un dispositivo para producir una señal de audio desplazada en frecuencia basándose en una señal de entrada de audio. La señal de entrada de audio puede representarse a este respecto, para una pluralidad de primeras subbandas, mediante uno o varios primeros valores de subbanda. El dispositivo incluye una interfaz y una unidad de desplazamiento en frecuencia. La interfaz está configurada para recibir la señal de entrada de audio. La unidad de desplazamiento en frecuencia está configurada para producir la señal de audio desplazada en frecuencia, presentando la señal de audio desplazada en frecuencia, para una pluralidad de segundas subbandas, en cada caso uno o varios segundos valores de subbanda. Además, cada uno de los primeros y los segundos valores de subbanda presentan información de un ángulo de fase respectivo.

45 La unidad de desplazamiento en frecuencia está configurada adicionalmente para producir uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de modo que el segundo ángulo de fase de este segundo valor de subbanda puede diferir del primer ángulo de fase de este primer valor de subbanda por una diferencia de ángulo de fase, dependiendo la diferencia de ángulo de fase de una información de frecuencia que indica por qué diferencia de frecuencia se va a desplazar la señal de entrada de audio a fin de obtener la señal de audio desplazada en frecuencia, y dependiendo la diferencia de ángulo de fase de un ancho de banda de frecuencia de una de las primeras subbandas.

50 Las formas de realización proporcionan conceptos mejorados para ampliación de ancho de banda, denominándose estos conceptos mejorados en lo sucesivo como "ampliación de ancho de banda espectral armónica" o "HSBE". Esta ampliación de ancho de banda armónica desarrollada en el dominio de frecuencia permite suprimir artefactos indeseados. El espectro replicado aquí es modulado de modo que la estructura armónica original se mantiene. En contraste con otros métodos, la HSBE puede basarse a este respecto en la representación de señal en el dominio

MDCT y así permite una implementación eficaz. La ampliación de ancho de banda correcta armónicamente es conseguida mediante un proceso de copiado de los valores espectrales con modulación posterior.

5 De acuerdo con una forma de realización, se usa el dominio de subbanda de la MDCT el cual se implementa usualmente ya en codificadores de audio. De esta forma, la transformada no resulta en una complejidad adicional o retardo.

10 En formas de realización, cuando la señal es desplazada por un múltiplo entero par del ancho de banda de subbanda, las señales de subbanda de frecuencias inferiores son desplazadas a las bandas de frecuencia superiores correspondientes.

15 Cuando la señal va a desplazarse por un múltiplo entero impar, en formas de realización cada segunda muestra de las señales de subbanda que han de copiarse se dota de signo inverso (índice del bloque creciente, en la dirección de tiempo). De esta forma, la característica de compensación de solapamiento (*alias cancellation property*) del banco de filtros MDCT aún funciona para la señal desplazada en frecuencia y copiada.

20 Cuando las señales de frecuencias inferiores son desplazadas por un múltiplo no entero del ancho de banda de subbanda, en formas de realización se producen versiones de valor complejo de las subbandas de frecuencia inferior, se toma una copia de estas y se modulan (multiplican) por una función exponencial compleja, donde esta función exponencial compleja tiene una frecuencia que corresponde a la diferencia de la frecuencia del siguiente múltiplo entero del ancho de banda de subbanda (hay un desplazamiento de ancho de banda de subbanda fraccionario). Sin embargo, la característica de compensación de solapamiento del banco de filtros MDCT se ve afectada o es destruida por ello.

25 A fin de evitar o reducir la distorsión de frecuencia resultante, se realiza un procesamiento similar a una convolución entre señales de subbanda adyacentes, donde una versión ponderada de una de las señales de su banda se suma a la señal de subbanda de una banda en su cercanía, de modo que exhibe el signo inverso de la componente de solapamiento, y así el solapamiento es compensado o reducido. En una forma de realización, los pesos aquí son elegidos de modo que correspondan al desplazamiento en frecuencia fraccionario deseado.

30 Mediante la modulación del espectro se infringe la característica TDAC de la MDCT, de modo que se producen componentes de solapamiento. A fin de eliminarlas, se proporciona en formas de realización una estructura de filtro FIR para compensación de solapamiento. Las respuestas de impulso de filtro requeridas para esto son optimizadas por medio de aproximación sucesiva y, por ejemplo, almacenadas como una tabla de consulta.

35 Para la aplicabilidad de los conceptos presentados aquí no es necesario adaptar el proceso de codificación, es decir adaptar la compresión de los datos, por ejemplo aplicando un modelo psicoacústico.

40 Los conceptos proporcionados se basan en métodos de ampliación de ancho de banda ya existentes y los mejoran. Con este nuevo método es posible incrementar la calidad del material de audio reproducido, manteniendo los requerimientos de almacenamiento. No se influye aquí en el proceso de codificación, sino que se desarrolla el decodificador. El método desarrollado implementa una ampliación de ancho de banda armónica. Está basado en replicación de banda espectral (SBR), tal como se usa en la tecnología HE-AAC.

45 Los eficaces conceptos de replicación de banda espectral inventivos proporcionados mantienen la estructura armónica del espectro original y así reducen los artefactos descritos de la tecnología SBR conocida.

50 Se proporcionan conceptos para la ampliación de banda espectral armónica (HSBE). Dado que la eficacia desempeña un papel importante, se describirá el procesamiento de señales subyacente por el banco de filtros usado hasta la supresión de artefactos por los filtros FIR.

La ampliación de banda espectral armónica presentada aquí proporciona una posibilidad potente y eficaz de ampliar el espectro de banda limitada de una señal de audio continuando con su estructura armónica.

55 Se ha mostrado que en la ampliación de banda espectral armónica han de tenerse en cuenta numerosos problemas para que el espectro pueda continuarse de una manera correcta armónicamente. Aunque se conocen algunos métodos para reconstruir la estructura armónica del espectro, es útil aun así desarrollar técnicas novedosas más eficientes. En particular para la integración en estructuras de codificador de audio modernas, como AAC o USAC, es indispensable la adaptación a las condiciones generales de las mismas.

60 Un componente esencial para una ampliación de banda espectral armónica satisfactoria es suprimir las componentes espurias que se producen. Es por ello que un punto esencial radica en el modo de funcionamiento y la optimización de la cancelación de solapamiento destinada a ello. Se ha mostrado que las fracciones de señal indeseadas pueden reducirse enormemente al seleccionar filtros FIR apropiados. Así, en contraste con el filtrado

clásico, no se convoluciona una señal de tiempo con la respuesta de impulso de filtro, sino los coeficientes MDCT/MDST complejos. Aunque mediante este método no es posible cancelar todas las componentes de solapamiento que se producen, es suficiente para eliminar algunas fracciones espurias dominantes.

5 Aparte de eso, la transformación de los coeficientes MDCT en valores espectrales MDST constituye otro elemento importante de la ampliación de banda espectral armónica. Los codificadores de audio modernos operan exclusivamente en el dominio MDCT. Si bien la señal es descrita de manera precisa suficientemente en su representación espectral, sin embargo esta información no es suficiente para replicar el espectro con HSBE. La evolución de fase requerida solo puede ser modificada mediante coeficientes MDST adicionales. Para ello se introduce aquí una transformación, la cual permite con un retardo constante calcular los coeficientes MDST que faltan lo más eficazmente posible a partir de los valores MDCT conocidos. Aparte de una solución exacta, se presentará una alternativa propensa a errores, pero con ahorro de recursos.

15 La modulación del espectro es importante en la HSBE. Se ha mostrado aquí que para una replicación eficaz del espectro son útiles dos pasos. Por una parte, el espectro es desplazado por subbandas MDCT de número entero y, por otra parte, se realiza una modulación dentro del ancho de banda de una subbanda MDCT para una resolución fina. Para señales en calidad CD, la resolución conseguida con esta técnica se sitúa alrededor de 0,5 Hz. Esto significa que la estructura armónica del espectro puede ser replicada de manera muy precisa.

20 La determinación de la frecuencia de retraso (*lag*) requerida para la modulación puede proporcionarse por ejemplo por el codificador.

25 En formas de realización, se proporcionan un sistema, un dispositivo o un método o un programa informático para producir una señal desplazada en frecuencia, usándose descomposición en subbandas, donde, para desplazamientos de ancho de banda de subbanda fraccionarios, las bandas son multiplicadas por medio de multiplicación por una función exponencial compleja.

30 En formas de realización, las componentes de solapamiento se compensan o al menos se reducen al realizar un procesamiento de mariposa entre señales de subbanda adyacentes

De acuerdo con formas de realización adicionales, el desplazamiento en frecuencia es realizado en el dominio de subbanda de un sistema de codificación de audio.

35 En formas de realización, el desplazamiento en frecuencia es usado para completar fracciones de frecuencia que faltan y/o huecos espectrales de una representación en frecuencia de una señal en un sistema de codificación de audio.

40 De acuerdo con formas de realización, el desplazamiento en frecuencia es usado en combinación con la conversión de tasa de muestreo para cambiar la velocidad de reproducción, mientras que la altura tonal (*pitch*) permanece igual.

45 Como se incrementa por ejemplo en primer lugar la frecuencia por medio de desplazamiento en frecuencia y entonces se reduce la velocidad de reproducción, el tiempo de reproducción de una cierta cantidad de datos de audio se volverá más largo con una altura tonal constante. Cuando, a la inversa, se reduce por ejemplo en primer lugar la frecuencia mediante desplazamiento en frecuencia y entonces se incrementa el tiempo de reproducción de una cierta cantidad de datos de audio, el tiempo de reproducción se vuelve más corto con una altura tonal constante.

50 En formas de realización adicionales, los conceptos son usados para ajustar finamente una señal musical. Los conceptos proporcionados pueden usarse por ejemplo en una forma particularmente ventajosa para Auto Tune. Cuando, por ejemplo, se van a realizar únicamente pequeños cambios de altura tonal de una señal musical digital, es decir, por ejemplo, cambios de frecuencia más pequeños que el ancho de banda de una subbanda, por ejemplo más pequeños que una subbanda MDCT o QMF, los conceptos proporcionados son particularmente ventajosos.

55 De acuerdo con formas de realización, los conceptos son usados para producir frecuencias más altas de un espectro al copiar o desplazar en frecuencia fracciones de frecuencia más pequeñas de un espectro.

En formas de realización, la descomposición en subbandas es una transformada de coseno discreta modificada (*modified discrete cosine transform*, MDCT).

60 En formas de realización adicionales, la descomposición en subbandas es un banco de filtros de espejo en cuadratura (*quadrature mirror filterbank*, QMF) polifásico.

Los conceptos proporcionados en las formas de realización anteriores pueden implementarse, entre otras cosas, como sistema, como dispositivo o como método o como programa informático.

Una implementación aceptable de la ampliación de banda espectral armónica se ha formado sobre la base de los algoritmos y principios funcionales proporcionados y desarrollados. Es posible una implementación en tiempo real en Matlab o en el lenguaje de programación C o en otro lenguaje de programación. Esto significa que los conceptos proporcionados pueden ser aplicados en sistemas en tiempo real. Se espera que la calidad de la señal reproducida se incremente al usar estos métodos, como también es el caso con SBR.

Analizar la complejidad de la modulación del espectro de acuerdo con los conceptos proporcionados proporciona muy buenos valores. La complejidad de cálculo depende a este respecto en gran medida de la transformación MDCT-MDST.

Se proporciona además un método para producir una señal de audio desplazada en frecuencia basándose en una señal de entrada de audio, pudiendo representarse la señal de entrada de audio, para una pluralidad de primeras subbandas, por uno o varios primeros valores de subbanda.

El método comprende:

Recibir la señal de entrada de audio, y procesar la señal de audio desplazada en frecuencia, presentando la señal de audio desplazada en frecuencia, para una pluralidad de segundas subbandas, en cada caso uno o varios segundos valores de subbanda, presentando cada uno de los primeros y los segundos valores de subbanda información de un ángulo de fase respectivo, y produciéndose uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de tal manera que el segundo ángulo de fase de este segundo valor de subbanda puede diferir del primer ángulo de fase de este primer valor de subbanda por una diferencia de ángulo de fase, dependiendo la diferencia de ángulo de fase de una información de frecuencia que indica por qué diferencia de frecuencia se va a desplazar la señal de entrada de audio a fin de obtener la señal de audio desplazada en frecuencia, y dependiendo la diferencia de ángulo de fase de un ancho de banda de frecuencia de una de las primeras subbandas.

Se proporciona además un programa informático para realizar el método anterior, cuando el programa informático es ejecutado en un ordenador o un procesador de señales.

Las formas de realización preferidas se encuentran en las reivindicaciones dependientes.

Las formas de realización preferidas se describirán a continuación con referencia a los dibujos, en los cuales:

- La Fig. 1A muestra un dispositivo para la ampliación de ancho de banda de acuerdo con una forma de realización,
- la Fig. 1B muestra un dispositivo de acuerdo con un ejemplo de realización, donde el dispositivo está configurado para producir una señal de audio ampliada en frecuencia,
- la Fig. 2 es una ilustración esquemática de la reconstrucción de HF por HSBE de acuerdo con una forma de realización,
- la Fig. 3 muestra un dispositivo 300 para producir una señal de audio desplazada en frecuencia de acuerdo con una forma de realización,
- la Fig. 4A-4B muestra una estimación de la matriz de transformación MDCT-MDST de acuerdo con una forma de realización,
- la Fig. 5 ilustra respuestas de impulso de las matrices de transformación MDCT-MDST de acuerdo con una forma de realización,
- la Fig. 6 muestra una estimación del espectro MDST para ruido blanco,
- la Fig. 7 muestra una estructura armónica con HSBE de acuerdo con una forma de realización,
- la Fig. 8 ilustra un diagrama de la reconstrucción de HF por HSBE ampliada de acuerdo con una forma de realización,
- la Fig. 9 ilustra componentes de solapamiento para $\phi = 45^\circ$,
- la Fig. 10 ilustra un filtro respuesta de impulso de un filtro antisolapamiento de acuerdo con un ejemplo de realización para $\phi = 90^\circ$,
- la Fig. 11 muestra la influencia de la filtración antisolapamiento en una señal de seno para $\phi = 90^\circ$ de acuerdo con una forma de realización,
- la Fig. 12 muestra una estructura de mariposa para reducción de solapamiento de acuerdo con una forma de realización,
- la Fig. 13 muestra adaptación de curva envolvente por HSBE-LPC de acuerdo con una forma de realización,
- la Fig. 14 ilustra la complejidad de la modulación y filtración antisolapamiento de HSBE,
- la Fig. 15 ilustra la complejidad de la transformación MDCT/MDST rápida,
- la Fig. 16 es una ilustración de la complejidad de la transformación MDCT/MDST,
- la Fig. 17 es una interferencia residual en HSBE en función de la longitud de transformada,
- la Fig. 18 muestra una comparación de HSBE y SBR,

- la Fig. 19 muestra una estructura DCT-III/DST-III rápida universal,
 la Fig. 20 muestra una estructura DCT-IV rápida,
 la Fig. 21 muestra una ilustración esquemática de un enmascaramiento temporal,
 la Fig. 22 muestra una ilustración esquemática del enmascaramiento dependiente de la frecuencia en el oído humano,
- 5 la Fig. 23 muestra un diagrama de bloques de un codificador de audio psicoacústico,
 la Fig. 24 muestra un diagrama de bloques de un decodificador de audio psicoacústico,
 la Fig. 25 muestra un banco de filtros con un retardo de sistema de n_d muestras,
 la Fig. 26 muestra un banco de filtros polifásico,
- 10 la Fig. 27 muestra categorías de la ampliación de ancho de banda,
 la Fig. 28 muestra un diagrama de bloques de un codificador SBR ampliado,
 la Fig. 29 muestra un diagrama de bloques de un decodificador ampliado por SBR,
 la Fig. 30 muestra una ilustración esquemática de la reconstrucción de HF por SBR,
 la Fig. 31 muestra una destrucción de la estructura armónica con SBR,
- 15 la Fig. 32 es una ilustración esquemática de la reconstrucción de HF por HBE,
 la Fig. 33 ilustra una estructura armónica con HBE,
 la Fig. 34 muestra una ilustración esquemática de la reconstrucción de HF por CM-BWE, y
 la Fig. 35 ilustra una estructura armónica con CM-BWE.

20 La Fig. 1A muestra un dispositivo 100 para producir una señal de audio desplazada en frecuencia basada en una señal de entrada de audio. La señal de entrada de audio puede representarse a este respecto, para una pluralidad de primeras subbandas, por uno o varios primeros valores de subbanda. El dispositivo incluye una interfaz 110 y una unidad de desplazamiento en frecuencia 120. La interfaz 110 está configurada para recibir la señal de entrada de audio. La unidad de desplazamiento en frecuencia 120 está configurada para producir la señal de audio desplazada en frecuencia, presentando la señal de audio desplazada en frecuencia, para una pluralidad de segundas subbandas, en cada caso uno o varios segundos valores de subbanda. Además, cada uno de los primeros y los segundos valores de subbanda presentan información de un ángulo de fase respectivo. La unidad de desplazamiento en frecuencia 120 está configurada adicionalmente para producir uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de tal manera que el segundo ángulo de fase de este segundo valor de subbanda puede diferir del primer ángulo de fase de este primer valor de subbanda por una diferencia de ángulo de fase, dependiendo la diferencia de ángulo de fase de una información de frecuencia que indica por qué diferencia de frecuencia se va a desplazar la señal de entrada de audio, es decir, por ejemplo, por qué diferencia de frecuencia se van a desplazar los primeros valores de subbanda de las subbandas de la señal de entrada de audio a fin de obtener la señal de audio desplazada en frecuencia, y dependiendo la diferencia de ángulo de fase de un ancho de banda de frecuencia de una de las primeras subbandas.

25
30
35

En algunas formas de realización, la interfaz puede estar configurada para recibir la información de frecuencia que indica por qué diferencia de frecuencia se van a desplazar los primeros valores de subbanda de las subbandas de la señal de entrada de audio.

40

La Fig. 1b muestra un dispositivo 150 de acuerdo con un ejemplo de realización. El dispositivo 150 está configurado para producir una señal de audio de frecuencia ampliada. El dispositivo 150 está configurado a este respecto para producir la señal de audio de frecuencia ampliada al producir el dispositivo 150 los segundos valores de subbanda de la señal de audio desplazada en frecuencia, comprendiendo la señal de audio de frecuencia ampliada los primeros valores de subbanda de la señal de entrada de audio y los segundos valores de subbanda de la señal de audio desplazada en frecuencia.

45

A continuación se introducen los conceptos para la ampliación de ancho de banda de acuerdo con formas de realización, los cuales se denominan ampliación de banda espectral armónica (HSBE). Se trata de conceptos que combinan las ventajas de SBR y las de la modulación de banda lateral única continua. Se basa en representación de señales en el dominio MDCT. Así, puede integrarse HSBE directamente en codificadores de audio actuales, como HE-AAC o USAC, sin usar un banco de filtros QMF adicional, como se hace en SBR. A diferencia de los métodos de dominio de tiempo, ha de calcularse una DFT de alta resolución y no se requiere ninguna señal analítica.

50

A continuación se explicará el principio de funcionamiento de la ampliación de banda espectral armónica. La ampliación de banda espectral armónica usa una copia de banda base para generar la fracción de HF. La banda base es replicada a través de un proceso de copiado en el dominio de alta frecuencia. A diferencia de CM-BWE, donde debido al copiado se produce un hueco en el que faltan ciertos subtonos armónicos, el desplazamiento de la banda base se amplía en HSBE. La banda base es primero copiada también hacia arriba, de modo que la frecuencia de 0 Hz se sitúa entonces en f_g . El hueco así formado entre el último armónico de la frecuencia $f < f_g$ en la banda base y la frecuencia f_g es compensado al desplazar la banda base copiada de nuevo hacia abajo de modo que la estructura armónica será continua de nuevo. Así, se evita un hueco causado por la omisión de un subtono armónico, como en los métodos de dominio de tiempo. El proceso de ampliación de ancho de banda consiste aquí en dos partes. Una parte se realiza mediante un proceso de copiado en el dominio MDCT. Los coeficientes MDCT de baja

55
60

frecuencia son replicados mediante un copiado simple. La otra parte de la ampliación de ancho de banda, es decir el mantenimiento de la estructura armónica, se obtiene mediante una manipulación de la fase. Así, tiene que estar presente información de fase para este paso. La ampliación de banda espectral armónica básicamente opera con coeficientes MDCT puramente reales. Esto significa que una transferencia a un espectro complejo tiene lugar a fin de cambiar la información de fase. Esto es logrado mediante la transformación MDCT-MDST proporcionada aquí.

A fin de que las altas frecuencias de la banda base no se superpongan a las frecuencias de la banda replicada durante la adaptación, la banda HF es sometida a filtración de paso alto. Debido a la representación de la señal como coeficientes MDCT, esta filtración es muy simple ya que los coeficientes indeseados pueden ajustarse a cero. Sin embargo, este tipo de desplazamiento causa una limitación de banda de la señal sintetizada. Esto significa que, después de reconstruir la banda HF, no se alcanza la frecuencia de señal máxima original f_{max} , sino únicamente la frecuencia f_{syn} . El hueco que aparece entre f_{max} y f_{syn} puede, si es necesario, llenarse con ruido.

La Fig. 2 es una ilustración esquemática del proceso de copiado, que incluye adaptación armónica. Así, la Fig. 2 es una ilustración esquemática de la reconstrucción de HF por HSBE. La Fig. 2 muestra a la izquierda un copiado y desplazamiento de la banda base. La Fig. 2 muestra a la derecha un espectro después de la adaptación de la envolvente espectral.

La adaptación necesaria de la fase causa componentes espurios adicionales en la señal. Estas son suprimidas mediante la filtración antisolapamiento desarrollada de los valores espectrales MDCT/MDST complejos. Finalmente, se adapta la envolvente espectral de nuevo a su curso original mediante un método apropiado.

La Fig. 3 muestra un decodificador HSBE, es decir un decodificador ampliado por HSBE, que se obtiene con el procedimiento mencionado.

La Fig. 3 muestra un dispositivo 300 para producir una señal de audio desplazada en frecuencia de acuerdo con una forma de realización. En una forma de realización, puede tratarse de un decodificador HSBE, es decir un decodificador ampliado por HSBE.

El dispositivo 300 comprende una interfaz 310 y una unidad de desplazamiento en frecuencia 320. Una unidad de transformación MDCT/MDST 315 está ubicada entre la interfaz 310 y la unidad de desplazamiento en frecuencia 320. De manera adicional, el dispositivo 300 comprende una unidad de filtro 330. Además, el dispositivo 300 presenta una unidad de transformación de síntesis 340, por ejemplo en forma de un banco de filtros, y una unidad de adaptación de curva envolvente 350. De manera adicional, el dispositivo 300 en el ejemplo de realización de la Fig. 3 incluye una unidad para calcular τ y φ (318).

La unidad de transformación MDCT/MDST 315 puede estar configurada para obtener uno o varios primeros coeficientes MDCT de la señal de entrada de audio, los cuales son coeficientes de una transformada de coseno discreta modificada de la señal de entrada de audio. La unidad de transformación MDCT/MDST 315 puede obtener estos primeros coeficientes MDCT por ejemplo desde la interfaz 310.

La unidad de transformación MDCT/MDST 315 está configurada para determinar, basándose en uno o varios de los primeros coeficientes MDCT de la señal de entrada de audio, uno o varios primeros coeficientes MDST de la señal de entrada de audio, los cuales son coeficientes de una transformada de seno discreta modificada.

La unidad de desplazamiento en frecuencia 320 puede estar configurada a este respecto para producir segundos valores de subbanda basándose en cada caso en uno de los primeros valores de subbanda, basándose cada uno de los primeros valores de subbanda en uno de los primeros coeficientes MDCT y uno de los primeros coeficientes MDST que se ha determinado basándose en este primer coeficiente MDCT.

La estructura mostrada del dispositivo 300, realizado por ejemplo como decodificador HSBE, depende de los algoritmos implementados. Para usar este decodificador en otros entornos, puede ser necesario realizar la reconstrucción de la curva envolvente en el dominio de frecuencia. El bloque correspondiente en este caso estará directamente antes del banco de filtros de síntesis MDCT/MDST. Componentes adicionales pueden también estar insertados aquí, como adaptación de tonalidad usada en SBR. Sin embargo, estos métodos no tienen un efecto en el principio general de funcionamiento de la ampliación de banda espectral armónica.

De la Fig. 3 también se desprende el proceso de decodificación de una señal, de acuerdo con una forma de realización, que se codificó en el dominio MDCT. Para desplazar algunas partes del espectro en la medida deseada, los coeficientes MDCT decodificados son transformados primero en una representación MDCT/MDST combinada. Esto es útil ya que la modulación de un espectro complejo solo producirá componentes de solapamiento mayores en cada segunda subbanda, realizándose esta compensación mediante el método de compensación de solapamiento propuesto.

El generador de HF desplaza las entradas de frecuencia complejas de la representación de transformada MDCT/MDST de acuerdo con el desplazamiento deseado, ya sea de una manera decodificada a partir de la corriente de bits o determinada en el decodificador o mediante procesos externos. El término de modulación usado es:

5

$$e^{-jb\varphi} \frac{\pi}{180^\circ}$$

donde b es el índice de bloque, φ es el desplazamiento en frecuencia en grados (un desplazamiento en frecuencia de 180° corresponde un desplazamiento al centro de la siguiente subbanda).

10

Después de eso, se realiza la reducción de solapamiento, y entonces el espectro complejo es transformado de vuelta de manera inversa al dominio de tiempo y reproducido.

15

El término de modulación usado es una función exponencial compleja. φ es un ángulo en grados, el cual es dependiente en la diferencia de frecuencia por la cual se van a desplazar los primeros valores de subbanda de las subbandas.

La transferencia de MDCT a MDST se describirá a continuación.

20

La modulación de banda lateral única para mantener la estructura armónica es realizada parcialmente a través de una manipulación de la fase. Para la ampliación de banda espectral armónica, la respuesta de fase es de importancia esencial. Como ya se ha comentado, HSBE generalmente opera en el dominio MDCT real.

25

El decodificador pone a disposición únicamente coeficientes MDCT de modo que para la respuesta de fase se requieren adicionalmente los coeficientes MDST. La transferencia de los coeficientes MDCT a los coeficientes MDST correspondientes es posible y se comentará más adelante.

El cálculo exacto de la MDST se describirá a continuación.

30

Como en la DCT, en la MDCT existe una función correspondiente para calcular las fracciones de seno en la señal: la transformada de seno discreta modificada (MDST). Para la MDST, son válidas las mismas características que para la MDCT, aunque apenas se usa en codificación de audio.

Para algunas aplicaciones, como, por ejemplo, HSBE, es aun así útil calcular el espectro MDST de una señal. Puede obtenerse un espectro en valor absoluto y fase preciso al combinar los dos aspectos (Cheng, 2004).

35

La MDST se calcula de manera similar a la MDCT en la ecuación 2.35. La matriz de transformación $\underline{\mathbf{T}}$ y la matriz de ventana $\underline{\mathbf{F}}$ exhiben diferencias. La matriz de transformación MDST se calcula usando el núcleo de modulación de la DST-IV (véase la ecuación 2.18):

40

$$\underline{\mathbf{T}}_{DST}(k, n) := \text{sen} \left(\frac{\pi(n + 0,5)(k + 0,5)}{N} \right) \in \mathbb{R}^{N \times N} \quad (4.1a)$$

$$\underline{\mathbf{T}}_{DST}^{-1} = \frac{2}{N} \cdot \underline{\mathbf{T}}_{DST} \quad (4.1b)$$

45

Debido a otras características simétricas y otras consideraciones secundarias de la DST-IV comparado con la DCT-IV, las matrices de convolución deberán ser adaptadas de manera correspondiente. La modificación es un cambio de signo del segundo y cuarto cuadrantes de la matriz de ventana $\underline{\mathbf{F}}$:

$$\underline{\mathbf{F}}_{sin} = \begin{bmatrix} 0 & -w[\frac{N}{2} - 1] & w[\frac{N}{2}] & 0 \\ \vdots & & & \vdots \\ -w[0] & 0 & & w[N - 1] \\ w[N] & & & w[2N - 1] \\ \vdots & & & \vdots \\ 0 & w[N + \frac{N}{2} - 1] & w[N + \frac{N}{2}] & 0 \end{bmatrix} \quad (4.2)$$

Con estas adaptaciones puede calcularse la MDST de una señal $\underline{\mathbf{X}}$ dividida en bloques de la siguiente manera:

$$\underline{\mathbf{X}}_{MDST} = \underline{\mathbf{T}}_{DST} \cdot \underline{\mathbf{D}} \cdot \underline{\mathbf{F}}_{sen} \cdot \underline{\mathbf{X}} \quad (4.3)$$

Para la transformada inversa, es válido entonces que:

$$\underline{\hat{\mathbf{X}}} = \underline{\mathbf{F}}_{sin}^T \cdot \underline{\mathbf{D}}^{-1} \cdot \underline{\mathbf{T}}_{DST}^{-1} \cdot \underline{\mathbf{X}} \quad (4.4)$$

Una función de transferencia compleja calculada a partir de la combinación de espectros MDCT y MDST es necesaria a fin de manipular la respuesta de fase, por ejemplo. El método implementado para transferir el espectro de MDCT a coeficientes MDST se presenta a continuación.

Un método trivial y costoso desde el punto de vista computacional es transferir la señal en el dominio MDCT devuelta al dominio de tiempo, con MDST posterior:

$$\underline{\mathbf{X}}_{MDST} = \underline{\mathbf{T}}_{DST} \cdot \underline{\mathbf{D}} \cdot \underline{\mathbf{F}}_{sen} \cdot \underline{\mathbf{F}}^T \cdot \underline{\mathbf{D}}^{-1} \cdot \underline{\mathbf{T}}^{-1} \cdot \underline{\mathbf{X}} \quad (4.5)$$

Este cálculo puede simplificarse a fin de reducir el esfuerzo requerido. Así, primero la matriz polifásica $\underline{\mathbf{H}}$ se define como:

$$\underline{\mathbf{H}} = \underline{\mathbf{T}}_{DST} \cdot \underline{\mathbf{D}} \cdot \underline{\mathbf{F}}_{sen} \cdot \underline{\mathbf{F}}^T \cdot \underline{\mathbf{D}}^{-1} \cdot \underline{\mathbf{T}}^{-1} \quad (4.6)$$

Los elementos de la matriz $\underline{\mathbf{H}}$ consisten en respectivamente en polinomios de tercer orden en z. Esta característica puede ser usada para representar $\underline{\mathbf{H}}$ como una adición de tres matrices:

$$\underline{\mathbf{H}} = \underline{\mathbf{H}}_0 z^0 + \underline{\mathbf{H}}_1 z^{-1} + \underline{\mathbf{H}}_2 z^{-2} \quad (4.7)$$

Las tres sub matrices $\underline{\mathbf{H}}_0$, $\underline{\mathbf{H}}_1$ y $\underline{\mathbf{H}}_2$ muestran propiedades características que resultan en un cálculo eficaz. La matriz $\underline{\mathbf{H}}_1$ es una matriz ocupada escasamente con elementos 0,5 y -0,5. Existe una conexión directa entre las matrices $\underline{\mathbf{H}}_0$ y $\underline{\mathbf{H}}_2$ de modo que la matriz $\underline{\mathbf{H}}_2$ puede producirse como reflejo de los elementos de $\underline{\mathbf{H}}_0$ en su diagonal secundaria. La forma precisa y un cálculo detallado de estas matrices se presentarán más adelante. El espectro MDST del bloque $(b - 1)$ puede entonces calcularse de la siguiente manera:

$$\underline{\mathbf{X}}_{MDST}(b - 1) = \underline{\mathbf{H}}_0 \cdot \underline{\mathbf{X}}(b) + \underline{\mathbf{H}}_1 \cdot \underline{\mathbf{X}}(b - 1) + \underline{\mathbf{H}}_2 \cdot \underline{\mathbf{X}}(b - 2) \quad (4.8)$$

Así, $\underline{\mathbf{X}}(b)$ es la b -ésima columna de la matriz $\underline{\mathbf{X}}$. Puede observarse también en esta ecuación que se introduce un retardo de un bloque para calcular el espectro MDST. Cuando el espectro MDST del bloque b está presente, el espectro MDST del bloque previo está disponible. Con ayuda de los coeficientes MDST conseguidos esta forma, puede calcularse una respuesta de fase a partir del espectro MDCT/MDST complejo y manipularse usando una rotación de fase a fin de mantener la estructura armónica, como se ha comentado anteriormente.

A continuación se explicará un cálculo simplificado de la MDST.

Incluso aunque, de acuerdo con el método derivado, el cálculo de la MDST se simplifica enormemente, el cálculo de esta transformada es muy costoso desde el punto de vista computacional. Además, se requiere mucho espacio de memoria para almacenar la matriz \underline{H}_0 . Esto significa que se desea una simplificación adicional de la transformada.

Quando se analizan las matrices \underline{H}_0 y \underline{H}_2 más precisamente, es notable que contienen un número muy alto de valores los cuales tienden a cero. Los coeficientes con los valores absolutos más grandes se concentran en una región estrecha cercana a la diagonal principal de las matrices. Así, parece obvio poner los demás coeficientes a cero a fin de ahorrar de esta forma tanto potencia de cálculo como requisitos de almacenamiento. Además, los valores en las diagonales son muy similares. Básicamente difieren únicamente unos de otros por sus signos. Únicamente en las regiones cercanas a las esquinas adoptan los coeficientes valores más altos.

Con vistas a la simplificación se asume además que los valores por debajo y por encima de la diagonal principal son iguales, lo que significa que existe simetría axial de la matriz en relación con la diagonal principal. Se determina ahora una matriz simplificada, cuyos valores son tomados de la columna central de la matriz \underline{H}_0 . Así, de la columna central se recorta una región que incluye el elemento de la diagonal principal y cualquier número de elementos adicionales por debajo de la diagonal principal. Este fragmento se denomina $\underline{h}[n]$. La columna central de la nueva matriz se forma ahora a partir de $\underline{h}[n]$ y un reflejo puntual de $\underline{h}[n]$ en el elemento axial principal h_j , siendo el resto de la columna cero. Las otras columnas de la matriz simplificada se forman entonces mediante desplazamiento cíclico de esta columna. Se adapta signo de cada segunda columna. Al usar estos métodos para simplificar la matriz ocupada completamente \underline{H}_0 , la matriz ocupada escasamente \underline{H}'_0 puede determinarse ahora usando un número muy pequeño de coeficientes. Una particularidad de esta matriz es su estructura simétrica similar a una matriz de Toeplitz. Está formada por el desplazamiento cíclico de la respuesta de impulso recortada $h[n]$ y reflejada en la diagonal principal con:

$$h[n] = [h_{i,j}, h_{i+1,j}, \dots, h_{i+\sigma-1,j}, h_{i+\sigma,j}], \quad i, j = \frac{N}{2} - 1 \quad (4.9)$$

Así, i es el índice de línea y j del índice de columna de la matriz \underline{H}_0 y σ es el índice de ventana que determina la longitud del fragmento. La longitud de este fragmento siempre será $2\sigma + 1$. Para $\sigma = 1$ y $N = 6$, lo siguiente es válido para la estructura de la matriz \underline{H}'_0 :

$$\underline{H}'_0 = \begin{bmatrix} -h[0] & h[1] & 0 & 0 & 0 & 0 \\ -h[1] & h[0] & -h[1] & 0 & 0 & 0 \\ 0 & h[1] & -h[0] & h[1] & 0 & 0 \\ 0 & 0 & -h[1] & h[0] & -h[1] & 0 \\ 0 & 0 & 0 & h[1] & -h[0] & h[1] \\ 0 & 0 & 0 & 0 & -h[1] & h[0] \end{bmatrix} \quad (4.10)$$

Deberá tenerse en cuenta que, comenzando con la primera columna, cada segunda columna es multiplicada por -1. En las siguientes consideraciones, el índice de ventana σ corresponde a alrededor del 10% de la longitud de transformada N , es decir $\sigma = [0, 1 \cdot N]$. Esto significa que los requisitos de memoria para la matriz de transformación han disminuido al 20% debido al reflejo de los valores de $h[n]$.

Las Fig. 4A-4B representan una estimación de la matriz de transformación MDCT-MDST. La Fig. 4A muestra a la izquierda una matriz de transformación completamente ocupada \underline{H}_0 para $N = 64$. La Fig. 4B muestra a la derecha una matriz de transformación simplificada \underline{H}'_0 para $N = 64$.

En la Fig. 4A puede verse la matriz ocupada completamente \underline{H}_0 y, en comparación, en la Fig. 4B la matriz simplificada \underline{H}'_0 en una estructura simétrica similar a una matriz de Toeplitz. Como puede verse, una gran parte de los coeficientes alejados de la diagonal principal de \underline{H}'_0 son iguales a cero gracias a la simplificación. La Fig. 5 representa las respuestas de impulso de las matrices de transformación MDCT-MDST. En particular, la Fig.

5 ilustra la respuesta de impulso original de la columna 33-ésima de la matriz $\tilde{\mathbf{H}}_0$ (línea continua). En comparación con la misma puede verse, adicionalmente, la respuesta de impulso correspondiente, que se ha formado por el proceso de copiado y reflejo, de la nueva matriz $\tilde{\mathbf{H}}'_0$. El recorte de los valores tiene lugar usando una ventana rectangular con el índice de ventana $\sigma = 6$.

5 Este tipo de simplificación de la transformación MDCT-MDST no proporciona el espectro MDST exacto, tal como se produce por medio del cálculo según la ecuación 4.8. Un error es añadido al espectro por la simplificación realizada de las matrices $\tilde{\mathbf{H}}_0$ y $\tilde{\mathbf{H}}_2$. Esto causa una reducción de la distancia señal a ruido a aproximadamente -70 dB, como se muestra en la Fig. 6. La Fig. 6 muestra una estimación del espectro MDST para ruido blanco. El error de estimación del espectro MDST se incrementa en las regiones de borde del espectro. Este efecto se debe a la estimación imprecisa de los coeficientes de matriz próximos a los extremos de la diagonal principal. El error producido es reducido mediante la filtración de paso alto del algoritmo HSBE y está presente únicamente de manera consecutiva en las frecuencias altas.

A continuación se describirá la adaptación de la estructura armónica del espectro de frecuencia.

Una ventaja del método HSBE es mantener la estructura armónica después de la ampliación de ancho de banda.

15 Como ya se ha mencionado, esto tiene lugar mediante una manipulación de fase en el dominio MDCT/MDST complejo. La banda espectral copiada con el ancho de banda $B = f_{max} - f_g$ se considerara aquí. El objetivo es desplazar el espectro hacia abajo de modo que el primer armónico en esta banda (por ejemplo con la frecuencia $f_{H,n} > f_g$) después del desplazamiento se sitúe en la frecuencia del armónico más alto en la banda base de frecuencia $f_{H,\alpha} < f_g$. La distancia entre las frecuencias $f_{H,n}$ y $f_{H,\alpha}$ se denomina frecuencia de retraso f_{lag} .

20 La adaptación de la estructura armónica es regulada a través de esta frecuencia. Esta frecuencia puede también ser representada como el correspondiente múltiplo entero y no entero de las subbandas MDCT, por los que se va a desplazar la banda de frecuencia hacia abajo. Esto permite una máxima flexibilidad en el método desarrollado. Después de haber cumplido la condición mencionada anteriormente, todos los coeficientes MDCT con una frecuencia discreta inferior a f_g se ponen a cero a fin de que la banda base y la banda desplazada no se superpongan.

La Fig. 7 es una ilustración esquemática del resultado deseado del método HSBE para una señal tonal. Así, la Fig. 7 muestra la estructura armónica en HSBE. La Fig. 7 muestra el espectro de banda ancha original a la izquierda. La Fig. 7 muestra el espectro después de la reconstrucción de HF por HSBE a la derecha.

30 La estructura armónica original se mantiene así. No se deja fuera ningún subtono armónico por el desplazamiento descrito de la banda de alta frecuencia replicada a frecuencias más pequeñas que la frecuencia de corte f_g . El desplazamiento del espectro puede entonces ser interpretado como una modulación de banda lateral única de la señal de banda base filtrada por paso alto con la frecuencia de modulación f_{mod} . Lo siguiente es válido:

$$35 \quad f_{mod} = f_g - f_{lag} \quad (4.11)$$

Así, deberá mantenerse mente que, a frecuencias f_{lag} más grandes que la mitad del ancho de banda de una banda MDCT, los coeficientes MDCT cercanos a $f = f_{max}$ se vuelven cero por el desplazamiento. Estos pueden rellenarse con ruido. Esto no es necesario cuando la frecuencia f_{lag} es más pequeña que la mitad del ancho de banda de una banda MDCT, ya que ningún coeficiente MDCT se fija a cero.

A continuación se describirá la adaptación de las subbandas no de número entero

45 Es complicado mantener la estructura armónica cuando una banda MDCT comprende un alto ancho de banda comparado con la diferencia frecuencia de subtonos consecutivos de la estructura armónica. Cuando se realiza la modulación solo con frecuencias que representan un múltiplo entero del ancho de banda de una banda MDCT, la resolución de la reconstrucción armónica está limitada fuertemente y consecuentemente una estructura armónica fina no puede ser restaurada. Así, es necesario posibilitar una alta precisión de modulación de modo que el espectro de la banda base no solo pueda modularse por múltiplos enteros del ancho de banda de la banda MDCT, sino también por fracciones de los mismos.

Usando el siguiente enfoque es posible desplazar el espectro dentro del ancho de banda de una subbanda MDCT. El método está basado en modificar la fase del espectro MDCT/MDST complejo. La fase aquí se gira aquí mediante un desplazamiento en frecuencia normalizado φ , en función del curso temporal de la señal. Esta rotación temporal del ángulo de fase permite así un desplazamiento muy fino del espectro. Lo siguiente es válido:

$$55 \quad \underline{\mathbf{X}}_{mod}(b) = \underline{\mathbf{X}}(b) \cdot e^{-j \cdot b \cdot \varphi \cdot \frac{\pi}{180^\circ}} \quad (4.12)$$

Así, $\underline{\mathbf{X}}(b)$ es la b -ésima columna de la matriz compleja $\underline{\mathbf{X}} = \underline{\mathbf{X}}_{MDCT} + j \cdot \underline{\mathbf{X}}_{MDST}$ y φ es el desplazamiento en

frecuencia normalizado en grados. Teóricamente, cualquier ángulo puede ser usado para φ , pero por razones prácticas el rango de valores está muy limitado y se sitúa dentro del intervalo $[-90, 90] \subset \mathbb{Z}$. Al usar este intervalo, es posible calcular modulaciones que cubren el ancho de banda de una banda MDCT. Al ajustar el desplazamiento en frecuencia normalizado al intervalo indicado, el espectro puede ser desplazado por la mitad del ancho de banda MDCT respectivamente hacia frecuencias altas o bajas.

Para la función exponencial compleja $e^{-j \cdot b \cdot \varphi \cdot \frac{\pi}{180^\circ}}$, la cual depende de b y φ , puede calcularse primero un valor de resultado. φ es un ángulo en grados el cual es dependiente de la diferencia de frecuencia por la que se van a desplazar los primeros valores de subbanda de las subbandas. Un segundo valor de subbanda que ha de determinarse puede establecerse entonces al multiplicar uno de los primeros valores de subbanda en $\underline{X}(b)$ por el valor de resultado.

A continuación se explica la adaptación de subbandas de número entero.

El rango limitado de valores del ángulo de fase φ permite, al usar la modulación introducida, únicamente desplazar el espectro como máximo el ancho de banda de una banda MDCT. Para desplazamientos del espectro mayores que el ancho de banda de una banda MDCT, este desplazamiento es dividido en dos partes, en un múltiplo entero del ancho de banda de la banda MDCT y una fracción de dicho ancho de banda. Primero, el espectro es modulado por la frecuencia necesaria más pequeña que el ancho de banda de una banda MDCT según la ecuación 4.12, y posteriormente el espectro es desplazado por valores espectrales enteros.

A continuación se considerará el desplazamiento que corresponde a exactamente un múltiplo del ancho de banda de una banda MDCT. En este caso, existe un ángulo de fase φ' que es un múltiplo entero de 180° . Así, el desplazamiento del espectro por los valores espectrales de MDCT enteros puede considerarse que son un caso especial del método presentado anteriormente para subbandas no de número entero. La evaluación de la función de modulación compleja en la ecuación 4.12 proporciona los siguientes resultados. Si el producto del índice de bloque b y φ' es un múltiplo entero de 180° , el resultado de la función de modulación siempre será 1, de lo contrario será -1. Sabiendo esto, no es necesario evaluar la función en la ecuación 4.12 para la adaptación de subbandas de número entero sino que basta con una simple diferenciación de casos. Es válido que:

$$\underline{x}_{\text{mod}}(b, 0 : N - \tau - 1) = \begin{cases} -\underline{x}(b, \tau : N - 1) & \text{si } (b \cdot \frac{\varphi'}{180^\circ}) \in [1, 3, 5, \dots] \\ \underline{x}(b, \tau : N - 1) & \text{si no} \end{cases} \quad (4.13)$$

con el índice de modulación entero τ :

$$\tau = \frac{\varphi'}{180^\circ} \quad (4.14)$$

Así, $\underline{X}(b, \tau : N - 1)$ de nuevo es la b -ésima columna de la matriz compleja $\mathbf{X} = \mathbf{X}_{\text{MDCT}} + j \cdot \mathbf{X}_{\text{MDST}}$, con la diferencia de que solo los elementos de vector que comienzan en τ hasta el último elemento N son usados aquí. Este recorte de elementos de vector corresponde a la filtración de paso alto mencionada anteriormente del espectro MDCT/MDST complejo.

Para aplicar la modulación, la frecuencia de modulación es transferida al índice de modulación τ y al ángulo de fase φ en función de f_{lag} . Para ello, en primer lugar, la frecuencia f_{lag} es normalizada a la mitad de la frecuencia de muestreo f_s . A continuación se establece el desplazamiento equivalente en bandas MDCT φ_{lag} y se calculan el índice de modulación τ y el ángulo de fase φ de la siguiente manera:

$$\omega_{\text{lag}} = f_{\text{lag}} \cdot 2 / f_s \quad (4.15)$$

$$\varphi_{\text{lag}} = N \cdot \omega_{\text{lag}} \quad (4.16)$$

$$\tau = \lceil \varphi_{lag} \rceil \quad (4.17)$$

$$\varphi = \lceil (\varphi_{lag} - \tau) \cdot 180^\circ \rceil \quad (4.18)$$

5 Al combinar los dos métodos, es posible implementar relaciones de parche variables. La relación de parche es así la relación de la frecuencia de señal máxima posible f_{max} y la frecuencia de corte de la banda base f_g . Una relación de parche de 2 : 1, por ejemplo, expresa que una única copia de la banda base es establecida y modulada (véase la Fig. 2). Las relaciones de parche mayores que 2 : 1 ocurren a tasas de transmisión bajas o fluctuantes. Tales relaciones se implementan, de manera similar a CM-BWE (véase más arriba), al copiar y modular la banda base
10 varias veces. Deberá también tenerse en cuenta aquí que la frecuencia de retraso necesaria aquí se incrementa por f_{lag} con cada copia de la banda base, como se ilustra en la Fig. 8 para una relación de 2,5 : 1.

La Fig. 8 ilustra un esquema de la reconstrucción de HF por HSBE ampliada. La Fig. 8 muestra a la izquierda el copiado y el desplazamiento de la banda base. La Fig. 8 muestra a la derecha el espectro después de haberse adaptado la envolvente espectral .
15

A continuación se describirán conceptos para las suprimir componentes espurias que se producen. Los conceptos descritos aquí pueden aplicarse por ejemplo a la unidad de filtro 330 de la Fig. 3.

20 La modulación del espectro en el dominio MDCT no puede realizarse fácilmente. Una reconstrucción perfecta con la MDCT inversa ya no es posible debido a la modulación del espectro. La causa es la formación de componentes de solapamiento de dominio de tiempo. La energía de estas componentes espurias es redistribuida mediante la modulación del espectro. La característica TDAC de la MDCT se infringe de este modo y ya no puede cancelar estas componentes con la transformada inversa. Cuando se considera la señal modulada después de la MDCT inversa,
25 pueden encontrarse componentes espurias en la respuesta de frecuencia de valor absoluto de la DFT por esta razón. Cuando se desplaza el espectro por $\varphi = 0^\circ$ y $\tau > 0$, estas componentes de solapamiento exhiben únicamente una muy baja amplitud y se sitúan en la primera o la última banda MDCT. En este caso, no es necesario reducir las componentes. Con factores de desplazamiento $\varphi \neq 0^\circ$, la amplitud de las componentes espurias que se forman es considerablemente mayor. En este caso, son audibles claramente. De manera consecuente, estas componentes han de tratarse.
30

La Fig. 9 muestra componentes de solapamiento para $\varphi = 45^\circ$: y_{ref} es un tono de seno original; y_{mod} es un tono de seno modulado que incluye componentes de solapamiento, el banco de filtros de DCT-IV se muestra expandido para una mejor ilustración.

35 En particular, la Fig. 9 ilustra la respuesta en frecuencia en valor absoluto de un tono de seno (y_{mod}) desplazado con $\varphi = 45^\circ$. La frecuencia del tono de seno original (y_{ref}) corresponde al centro de banda de duodécima banda MDCT. El espectro completo está modulado en un cuarto del ancho de banda de una banda MDCT hacia frecuencias altas por el ángulo de fase seleccionado. Como puede verse, las ocho componentes solapamiento dominantes están ubicadas en cada segunda banda por debajo y por encima de la duodécima banda MDCT. Esta característica de las componentes de solapamiento se da para cualquier señal. La razón de esto es que cada señal puede descomponerse en una suma ponderada de oscilaciones de seno y de coseno (véase más arriba). Para cada una de estas suboscilaciones aparece este patrón especial de componentes de solapamiento con la modulación según la ecuación 4.12. Sabiendo esto, se puede desarrollar un método que permita liberar cualquier señal de las componentes espurias indeseadas. Así, basta con analizar y cancelar las componentes de solapamiento que se forman mediante la modulación de una señal de seno.
40
45

A continuación se proporcionarán los conceptos para la filtración antisolapamiento.

Se forman fracciones adicionales en el dominio de frecuencia mediante la superposición temporal de los bloques para TDA. Éstas están presentes como fracciones espurias en el espectro de la señal de ancho de banda ampliado ya que no son canceladas con la transformada inversa mediante desplazamiento en el dominio de frecuencia. En una MDCT, estas componentes espurias reconocibles como picos en el espectro FFT (véase la Fig. 9) se representan mediante la atenuación de banda de rechazo baja del banco de filtros de la DCT-IV de únicamente alrededor de 15 dB por una suma de fracciones en varias de las bandas MDCT superpuestas. La energía de las componentes espurias en el espectro de la DFT de alta resolución puede así considerarse como el cálculo de la suma de la energía de varias bandas MDCT.
50
55

Debido esta asociación, se proporciona un filtro para reducir las componentes espurias en el dominio MDCT. El filtro está basado en una suma secuencial de los valores de frecuencia ponderados por los coeficientes de filtrado. La extensión del filtro en un valor de frecuencia centrado representa el rango de frecuencia en el cual se cancelan las componentes espurias. Para cada componente de solapamiento dominante se requiere un coeficiente de filtrado que la minimiza. El filtro depende del desplazamiento en frecuencia φ . Es válido que:
60

$$\underline{X}_{\text{AntiAlias}}(b) = \underline{X}(b) * \underline{h}(\varphi) \quad (4.19)$$

5 Así, $\underline{h}(\varphi)$ es el filtro antisolapamiento real para un cierto ángulo de fase φ y $\underline{X}(b)$ es el espectro MDCT/MDST complejo. El espectro después del filtrado ($\underline{X}_{\text{AntiAlias}}(b)$) aquí es más largo que el espectro original $\underline{X}(b)$. Esto significa que el espectro tiene que recortarse a fin de que corresponda de nuevo a la longitud de transformada N. Se elimina la parte del espectro en la que el filtro se asienta y decae. Así, la longitud del filtro se corta por la mitad tanto al inicio como al final del producto de convolución en el dominio MDCT/MDST complejo.

10 En la Fig. 10 puede verse la respuesta de impulso de filtro del filtro antisolapamiento (*anti-aliasing-filter*, AAF) para $\varphi = 90^\circ$. En el ejemplo de un único tono de seno es posible cancelar con el filtro mostrado un total de siete componentes de solapamiento dominantes. A este respecto, tres componentes están por debajo de la frecuencia del tono de seno. Estas componentes, las cuales se denominan conforme a su posición respecto a la frecuencia del tono de seno componentes hasta el tercer orden, son tratadas por los coeficientes de filtrado (tapas de filtro, *filtertaps*) 0, 2 y 4. Las tapas de filtro 8, 10, 12 y 14 cancelan cuatro componentes espurias en frecuencias por encima del tono de seno, es decir, componentes hasta el cuarto orden. En total el filtro comprende 15 coeficientes, siendo cada segundo valor igual a cero. Esto corresponde a la observación anterior de que las componentes de solapamiento se presentan únicamente en cada segunda banda.

20 Los coeficientes de filtrado de la Fig. 10 presentan así un cierto orden. Cada coeficiente de filtrado en este orden que sigue a un coeficiente de filtrado distinto de cero tiene el valor cero.

25 Con una estructura de filtro de este tipo es posible en general suprimir cualquier número de componentes de solapamiento. Sin embargo, es suficiente con cancelar las componentes hasta el cuarto orden. Esto permite conseguir una distancia señal a ruido de al menos 70 dB, lo que puede considerarse suficiente. Además, las componentes de solapamiento de alto orden se vuelven evidentes únicamente con ángulos de fase φ muy grandes. La limitación a la cancelación hasta el cuarto orden así es un buen compromiso entre una SNR alcanzable y el esfuerzo computacional para la cancelación de solapamiento.

30 A continuación se explica la optimización de los filtros antisolapamiento.

35 Un componente importante de la cancelación antisolapamiento descrita aquí son los filtros antisolapamiento usados. La reducción de la amplitud de las componentes espurias individuales que puede alcanzarse depende de manera decisiva de la elección de los coeficientes de filtrado apropiados. Así, es necesario optimizar estos filtros de modo que se garantice la supresión más alta posible. Un método fiable para ello es la optimización numérica de los coeficientes de filtrado por medio de aproximación sucesiva.

40 La aproximación sucesiva es un método de iteración de matemáticas numéricas y se refiere al proceso de aproximar gradualmente un problema de cálculo a la solución exacta. Así, un método de cálculo es aplicado repetidamente y el resultado de un paso es usado como valor de partida para el paso siguiente respectivo. La secuencia de resultados deberá ser convergente. Cuando el error aceptable con respecto a la solución exacta es mínimo, el resultado se ha determinado con un grado suficientemente preciso (Jordan-Engeln y Reutter, 1978).

45 Al comienzo del proceso de optimización se modula una señal de análisis con la ecuación 4.12 por un determinado ángulo de fase φ . La señal de análisis es un tono de seno, por las razones citadas anteriormente. La frecuencia del tono se sitúa idealmente en un cuarto de la frecuencia de muestreo subyacente. La ventaja de esto es que las componentes de solapamiento que se forman hasta el cuarto orden presentan la distancia más grande posible con respecto a los bordes del espectro y no interfieren con otras componentes espurias. Para propósitos de optimización, una longitud de transformada MDCT de 32 muestras es ideal. De ello se deriva que la frecuencia del tono de seno corresponde al centro de banda de la 16-ésima banda MDCT. La limitación a esta longitud de transformada ofrece varias ventajas. Por una parte, esto permite reducir la complejidad de cálculo de la MDCT. Por otra parte, las componentes de solapamiento hasta el cuarto orden se generan sin interfaces con una distancia máxima entre sí. Esto es ventajoso en particular para el reconocimiento de pico de señal necesario. El reconocimiento de pico de señal detecta automáticamente las componentes de solapamiento que han de suprimirse en la respuesta en frecuencia en valor absoluto de la DFT de alta resolución.

55 Después de modular la señal de análisis, las componentes de solapamiento se optimizan una tras otra en un orden alterno. Esto es necesario ya que las componentes espurias se influyen mutuamente. Así, el orden es desde la componente más débil de cuarto orden a la más dominante de primer orden. Esto asegura que las componentes de solapamiento de primer orden reciban la atenuación más grande posible. Para la componente directa, es decir el valor espectral para el cual han de cancelarse las componentes de solapamiento, el filtro se fija a uno. Este valor no se cambia durante la optimización.

La optimización numérica propiamente dicha se realiza de acuerdo con el principio de aproximación sucesiva expuesto. Así, cuando se asigna un valor de partida al coeficiente de filtrado que va a optimizarse, todos los demás coeficientes, excepto la componente directa, permanecen en cero. A continuación, el espectro MDCT/MDST complejo se convoluciona con este filtro y la respuesta en frecuencia en valor absoluto es examinada para ver si se ha producido la reducción de la componente espuria respectiva. Si este es el caso, el coeficiente de filtrado se incrementará conforme al tamaño de paso fijado. Este método de examen e incremento se repite hasta que ya no sea posible una supresión más fuerte de esta componente de solapamiento. A continuación se procede de la misma manera con los siguientes coeficientes de filtrado, manteniendo los coeficientes de filtrado ya optimizados.

5
10 Debido a la influencia mutua de las componentes de solapamiento, es práctico realizar varias iteraciones de este modo de proceder. El tamaño de paso mediante el cual los coeficientes de filtrado se incrementan se reduce con cada iteración. Esto significa que la calidad del filtro optimizado se incrementa con cada pasada. También se ha demostrado que tres iteraciones son suficientes para un conjunto de filtros óptimo que consiste en un filtro respectivo por cada ángulo de fase. Esto permite reducir las componentes de solapamiento a <-90 dB.

15 La Fig. 11 ilustra la influencia del filtrado antisolapamiento en una señal de seno para $\varphi = 90^\circ$. X_{Alias} es una señal de seno modulada por $\varphi = 90^\circ$; $X_{AntiAlias}$ es una señal filtrada con componentes espurias suprimidas.

20 En particular, la Fig. 11 muestra la influencia del filtrado antisolapamiento en una señal de seno modulada por $\varphi = 90^\circ$, en la respuesta en frecuencia en valor absoluto. X_{Alias} es el espectro de la señal modulada y $X_{AntiAlias}$ el de la señal modulada, convolucionado con el filtro optimizado para el ángulo de fase correspondiente. Los picos en el espectro identificados por el reconocimiento de picos son las componentes de solapamiento detectadas por el reconocimiento de picos de señal, incluyendo la componente directa (cuarto pico reconocido desde la izquierda). En este ejemplo, la optimización numérica de los filtros reduce las componentes espurias en promedio a -103 dB.

25 Es suficiente con establecer una vez un conjunto de filtros para cada ángulo de fase en el rango de valores. Para el filtrado de la señal, el filtro requerido puede cargarse entonces desde una base de datos.

30 Por ejemplo, los coeficientes de filtrado del filtro en función del ángulo de fase pueden leerse desde una base de datos o una memoria de un dispositivo para producir una señal de audio desplazada en frecuencia.

La Fig. 12 muestra la estructura de mariposa. Los pesos se determinan mediante aproximaciones sucesivas. La Fig. 12 muestra así la reducción de solapamiento para la subbanda X_4 (línea negra). El mismo método se ha realizado de manera correspondiente para todas las subbandas modificadas. Para reducir la componente de solapamiento causada por la modulación de X_4 , X_4 se tiene que multiplicarse por los pesos w_0 a w_4 y sumarse a las señales de subbanda X_0 , X_2 , X_4 , X_6 y X_8 . Deberá tenerse en cuenta que el peso w_2 siempre será igual a 1.

35 A la inversa, esto significa que a fin de producir un valor de subbanda filtrado de una de las subbandas, tiene que calcularse una suma del valor de subbanda no filtrado de esta subbanda y sumandos adicionales (el coeficiente peso/filtro w_2 que se aplicaría al valor de subbanda no filtrado de esta subbanda es $w_2=1$). Los sumandos adicionales son valores de subbanda ponderados, precisamente un valor de subbanda respectivo de las otras subbandas que se han multiplicado/ponderado por los otros coeficientes de peso/filtrado.

A continuación se describe la reconstrucción de la envolvente espectral.

45 La reconstrucción de la envolvente espectral se hace usando filtrado LPC. Así las fracciones tonales de la señal son eliminadas en el codificador mediante un filtro de predicción lineal y transmitidas de manera separada como coeficientes LPC. Los coeficientes de filtrado requeridos para esto pueden calcularse usando la recursión Levinson-Durbin (Larsen y Aarts, 2004). El resultado es que la banda base en el decodificador obtiene una característica espectral blanca. Después de la ampliación de ancho de banda por medio de HSBE, se realiza filtración inversa usando los coeficientes LPC y así la envolvente espectral original se le imprime de nuevo a la señal.

La Fig. 13 muestra una adaptación de curva envolvente HSBE-LPC. Así, X representa una señal BWE antes de la adaptación de curva envolvente. X_{LPC} es una señal BWE después de la adaptación de curva envolvente.

55 En particular, la Fig. 13 muestra las respuestas en frecuencia en valor absoluto de la DFT de una señal de ancho de banda ampliado usando HSBE. Antes de la reconstrucción de la envolvente espectral, la señal X presenta la característica de señal blanca mencionada. Después de la adaptación de la envolvente mediante filtración LPC inversa, la envolvente corresponde a la característica espectral original. Adicionalmente, la Fig. 13 muestra la función de transferencia del filtro LPC usado. Unos cuantos coeficientes de filtrado ya son suficientes para una descripción suficientemente precisa de la envolvente espectral; en este ejemplo se usaron 14 tapas de filtro LPC. La reconstrucción de curva envolvente no es un componente fijo de HSBE y puede reemplazarse por otro método.

En el marco de una evaluación, los conceptos proporcionados serán sometidos ahora a evaluación. La evaluación aquí significa tanto una comparación entre el método novedoso de la ampliación de banda espectral armónica por

medio de CM-BWE como el examen de la eficacia de HSBE con respecto a posibilidades y limitaciones, así como la complejidad de cálculo algorítmica.

Se presentará en primer lugar una comparación del método de replicación de banda espectral.

5 La ampliación de ancho de banda por medio de modulación de banda lateral única continua es un método de dominio de tiempo. Esto significa que una señal de tiempo siempre será necesaria para su aplicación. Dado que, después de la ampliación de ancho de banda tiene lugar una adaptación de curva envolvente y de tonalidad, para lo cual se requiere respectivamente una señal en el dominio espectral, cuando se aplica el método CM-BWE la señal de ancho de banda ampliado tiene que transferirse de vuelta al dominio de frecuencia. Esta transformada al dominio tiempo y de vuelta al dominio de frecuencia se omite con la ampliación de banda espectral armónica, ya que ésta opera en el dominio MDCT/MDST.

15 Adicionalmente, la señal de tiempo se deberá transferir a una señal de análisis antes de aplicar la modulación de banda lateral única continua. Es problemático calcular la señal de análisis requerida ya que esto se realiza usando un transformador de Hilbert. La función de transferencia ideal de la transformada de Hilbert es la función signo. Esta función puede representarse en el dominio de tiempo únicamente mediante un filtro de longitud infinita. Cuando se usa un filtro implementarle de una respuesta de pulso finita, el transformador de Hilbert ideal solo puede aproximarse. Además, la señal no es de un carácter perfectamente analítico después de la transformada de Hilbert aproximada. La calidad de la señal pseudoanalítica calculada así es dependiente de la longitud del filtro usado.

25 Aplicar la extensión de banda espectral armónica también requiere una transformada adicional. Dado que la HSBE opera en el dominio de frecuencia, un espectro complejo es necesario para el cálculo de la modulación de fase. Sin embargo, antes de aplicar la ampliación de ancho de banda, únicamente los coeficientes MDCET están presentes en la estructura de decodificador. Así, los coeficientes MDCT deben transferirse al dominio MDST a fin de obtener un espectro complejo y así obtener la información de fase necesaria. Esto es considerado de manera ejemplar aquí como realizado a través de una multiplicación de matrices que se ha formado al simplificar una transformada MDCT inversa y una transformada MDST posterior. Como se ha mostrado, la complejidad de cálculo requerida para esto puede minimizarse enormemente, y aun puede realizar un cálculo exacto de los coeficientes MDST.

30 Cuando se considera una organización esquemática de la respuesta en frecuencia en valor absoluto después de aplicar CM-BWE y HSBE, puede verse que parte del espectro en CM-BWE se ha rellenado con ruido blanco. La estructura armónica en esta región se perdió ya que los subtonos armónicos no pueden replicarse aquí. Este problema no se presenta cuando se aplica HSBE. La estructura armónica se continúa sin huecos.

35 La complejidad de la ampliación de banda espectral armónica se considerara a continuación.

40 La eficacia de la ampliación de banda espectral armónica novedosa depende de complejidad de cálculo y del espacio de memoria necesario. El examen estos factores está basado en implementar el algoritmo en el lenguaje de programación C. Con la realización algorítmica, se pone mucho énfasis en minimizar el número de pasos de cálculo. Sin embargo transformar los coeficientes MDCT a valores espectrales MDST y la filtración antisolapamiento están entre los pasos con mayor esfuerzo de cálculo. La modulación para formar una replicación correcta armónicamente del espectro es relativamente fácil ya que el desplazamiento por el índice de modulación τ corresponde únicamente a un proceso de copiado y la rotación de fase por el ángulo ϕ puede reducirse a una multiplicación compleja por cada valor espectral. La adaptación de la envolvente espectral no se toma en consideración aquí. Ya que una evaluación no es una parte fundamental del método HSBE, no se realiza algorítmicamente.

50 Todas las instrucciones relevantes, como adiciones (ADD), multiplicaciones (MULT), y la instrucción de multiplicación y acumulación (*multiply and accumulate*, MAC) realizada en la filtración antisolapamiento son consideradas en la evaluación. Tabla 5.1 es un compendio de los resultados para la modulación y filtración del espectro. Se refieren a la función local_HSBEpatching() en la cual los algoritmos correspondientes están implementados.

Tabla 5.1

N	2	4	8	16	32	64	128	256	512	1024	2048
ADD	4	8	16	32	64	128	256	512	1024	2048	4096
ABS	1	1	1	1	1	1	1	1	1	1	1
MULT	6	12	24	48	96	192	384	768	1536	3042	6144
MAC	16	32	64	128	256	512	1024	2048	4096	8192	16384

55 La tabla 5.1 ilustra la complejidad de la modulación HSBE y filtración antisolapamiento en una tabla. El listado contiene el número de operaciones relevantes en función de la longitud de transformada N . Un total de 10241

operaciones son requeridas para modular el espectro para $N = 2048$, entre las cuales están $2N$ adiciones y $3N$ multiplicaciones. La cancelación de solapamiento requerida es mucho más compleja. Se realizan para ello 16384 operaciones MAC. Esto corresponde al número de elementos no cero del filtro antisolapamiento multiplicado por la longitud de transformada, es decir en este caso $8N$ (véanse las explicaciones con relación a filtración antisolapamiento anteriores). Con este resultado, se obtiene una correlación lineal con la complejidad $O(N)$ para la complejidad cálculo de la modulación y AAF.

La Fig. 14 es una ilustración grafica de esta correlación. La Fig. 14 así ilustra la complejidad de la modulación y filtración antisolapamiento de HSBE.

Un elemento central de la ampliación de banda espectral armónica es transformar la señal por medio de MDCT y MDST. De manera similar a la transformada de Fourier rápida, se usa un algoritmo rápido para estas transformadas también. La estructura de la DCT-IV rápida subyacente se describirá aún más adelante. Los resultados de la evaluación se refieren a la función fmdcst1d() en la que la MDCT/MDST rápida se implementa de acuerdo con el enfoque descrito anteriormente. El análisis de MDCT y MDST está resumido en la Tabla 5.2.

Tabla 5.2 – Complejidad de MDCT/MDST rápida

N	2	8	32	128	512	2048
ADD	5	106	1046	6630	35974	181094
MULT	7	40	204	956	4348	19452

Cuando se evalúan estos resultados, uno puede ver que un total de aproximadamente $9 \cdot N \log N$ operaciones son requeridas para calcular una MDCT / MDST. Se realizan N multiplicaciones más en la transformada inversa. La razón de ello es que en este caso se efectúa el ajuste a escala requerido para la transformación. La evaluación correspondiente está resumida en la Tabla 5.3.

Tabla 5.3 – Complejidad de MDCT/MDST inversa rápida

N	2	8	32	128	512	2048
ADD	5	106	1046	6630	35974	181094
MULT	9	48	236	1084	4860	21500
DIV	1	1	1	1	1	1

El resultado es una complejidad de $O(N \log N)$ para las transformadas hacia adelante e inversas. El algoritmo implementado así cumple el requisito de una transformada rápida. Los resultados de esta evaluación se muestran en la Fig. 15.

La Fig. 15 muestra la complejidad de la MDCT/MDST rápida. Debe tenerse en cuenta que dos transformadas inversas deben calcularse para transferir la señal del dominio MDCT/MDST complejo al dominio de tiempo. El número de operaciones requerida se duplica.

La transferencia exacta de coeficientes MDCT a valores espectrales MDST es el proceso con mayor esfuerzo de cálculo en HSBE. No solo aparece aquí un retardo constante de un bloque, sino también el tiempo de cálculo más largo. Incluso cuando el esfuerzo requerido se reduce mediante el método introducido, esto resulta en el mayor número de operaciones, como se resume en la tabla 5.4.

Tabla 5.4 – Complejidad de la transformada MDCT/MDST exacta

N	2	8	32	128	512	2048
ADD	6	120	2016	32640	523776	8386560
MULT	12	144	2112	33024	525312	8392704

Para la multiplicación de las matrices ocupadas completamente H_0 y H_2 para el vector espectral, un total de

N^2 multiplicaciones y $N(N-1)$ adiciones son requeridas en cada caso. La matriz ocupada escasamente H_1 permite una implementación eficaz de modo que en este caso deben realizarse $2N$ multiplicaciones y N adiciones. Para la transferencia exacta de los valores MDCT a coeficientes MDST, se calculan en total $4N^2 + N$ adiciones y multiplicaciones. Para la complejidad del algoritmo implementado, existe una dependencia al cuadrado de la longitud de transformada. Esto puede expresarse mediante $O(N^2)$. El método simplificado requiere menos operaciones, aproximadamente, de manera correspondiente, en la relación $2\sigma + 1/N$. La correlación correspondiente está ilustrada en la Fig. 16. Así, la Fig. 16 muestra la complejidad de la transformada MDCT/MDST.

Aparte de la complejidad algorítmica, la utilización de memoria desempeña un papel importante cuando se considera la eficacia. Dependiendo del dispositivo terminal en el cual se implementa el codificador de audio, únicamente una cantidad muy limitada de memoria puede estar disponible. Así, es necesario mantener la utilización de la memoria principal tan baja como sea posible. La tabla 5.5 enumera los recursos utilizados. La descripción de módulo se refiere al archivo fuente C del mismo nombre. Únicamente las memorias intermedias más importantes requeridas para el almacenamiento y procesamiento de vectores de señal y matrices son consideradas aquí.

Tabla 5.5 - utilización de memoria por HSBE:

Módulo	Elementos	Tamaño en bytes	Tamaño en KiB
hsbelib	20480	81920	80
Hmatrix	4194304 (205)	16777216 (820)	16384 (0,80)
hsbeConv	15	60	0,05
fastDCSTIV	3072	12288	12
fastDCSTIII	4096	16384	16
AAF database	2715	10860	10,61
Total	4224682 (30583)	16898728 (122332)	16502,66 (119,46)

La información en (): contiene el cálculo simplificado de la transformada MDCT/MDST.

La implementación de la ampliación de banda espectral armónica está basada en la aritmética de coma flotante de precisión simple, lo que significa que el número de coma flotante es representado por 32 bits. El número indicado en la tabla 5.5 se refiere al número de números de coma flotante requeridos en este módulo. Como puede deducirse de la tabla, la utilización de memoria para el algoritmo HSBE propiamente dicho es relativamente pequeña con aproximadamente 109 KiB para modulación, la cancelación de solapamiento y MDCT/MDST. La base de datos para los filtros antisolapamiento también se almacena como una tabla de consulta y requiere casi 11 KiB para en total 2715 coeficientes de filtrado. De influencia decisiva para la necesidad de espacio de memoria es la matriz de transformación H_0 . Aproximadamente 16 MiB de memoria funcional son utilizados para esta matriz. Mediante las simplificaciones de la transformada MDCT/MDST presentadas anteriormente se reducen enormemente los requisitos de memoria necesarios de la matriz H_0 . Para el ejemplo tomado con $\sigma = [0, 1 \cdot N]$ se usan por tanto todavía alrededor de 0,8 KiB de memoria funcional.

Al usar la tecnología presentada, la ampliación de ancho de banda por medio de ampliación de banda espectral armónica proporcionada aquí permite una ampliación armónicamente correcta del espectro. A continuación se exponen las posibilidades y limitaciones de los conceptos proporcionados.

Generalmente pueden procesarse señales de diferentes longitudes. Sin embargo, para calcular la MDCT/MDST rápida implementada, es absolutamente necesario que la longitud de transformada N sea una potencia de número entero positivo de número dos. La máxima longitud de bloque posible está, similarmente a AAC, limitada a 2^{11} , es decir 2048. Con HSBE, también es posible variar la longitud de bloque durante el tiempo de ejecución. Esto es necesario particularmente para tratamiento de transitorios en codificadores de audio modernos. La SNR se determina de manera decisiva mediante la longitud de bloque. Longitudes de transformada grandes conducen a un mejor resultado que longitudes de bloque muy cortas. Esto es causado por las componentes de solapamiento que ocurren debido a la modulación. Aunque las componentes espurias hasta el cuarto orden son suprimidas por la filtración antisolapamiento, sin embargo algunas componentes indeseadas permanecen en la señal. Con longitudes de bloque muy cortas, estas componentes de solapamiento se extienden por un gran ancho de banda y son audibles claramente. Con longitudes de transformada grandes, el ancho de banda se vuelve correspondientemente más pequeño y las componentes de solapamiento son enmascaradas. Una longitud de bloque de 256 ha resultado ser un límite adecuado en diferentes pruebas. A partir de esta longitud, los artefactos son difícilmente percibidos, como se muestra en la Fig. 17.

La Fig. 17 muestra una restauración con HSBE en función de la longitud de transformada. La Fig. 17 muestra a la izquierda un espectro después de la MDCT inversa con $N = 32$ y $\varphi = 90$. La Fig. 17 muestra a la derecha un espectro después de la MDCT inversa con $N = 256$ y $\varphi = 90$.

Pueden verse las respuestas en frecuencia en valor absoluto de una señal de seno modulada con el ángulo de fase φ . Como puede reconocerse claramente, con pequeños tamaños de bloque (izquierda en la Fig. 17), claramente existen componentes espurias después de la cancelación de solapamiento. Estas se sitúan en un rango de aproximadamente -70 dB y son decisivas para la SNR. Con tamaños de bloque más grandes que 256 (derecha en la Fig. 17), estas fracciones son enmascaradas por la señal útil.

En definitiva, la ampliación de banda espectral armónica proporciona una reconstrucción altamente precisa de la estructura armónica de la señal. Para una longitud de transformada de 2156 y una frecuencia de muestreo de los datos de audio conforme al estándar de CD de 44100 Hz, la resolución del método desarrollado se sitúa en el rango de aproximadamente 0,5 Hz. Esto significa que el espectro puede modularse con una precisión de la mitad de un hercio. Para frecuencias de muestreo más pequeñas o longitudes de bloque más grandes, la resolución se incrementa y la modulación puede realizarse en incluso en rangos más precisos. El resultado de aplicar HSBE a una señal de mutiseno está ilustrado en la Fig. 18.

La Fig. 18 muestra una comparación de HSBE y SBR. REF freq. resp. se refiere a un espectro de la señal de mutiseno original. SBR freq. resp. se refiere a una señal de ancho de banda ampliado por medio de SBR; HSBE freq. resp. Se refiere a una señal de ancho de banda ampliado por medio de HSBE.

Como puede reconocerse en las respuestas de frecuencia de valor absoluto mostradas, el espectro es reconstruido de manera exacta con el método HSBE desarrollado. Después de procesar la señal de banda limitada usando HSBE (HSBE freq. resp.), el espectro se sitúa precisamente sobre el espectro original (REF freq. resp.). El espectro correspondiente que no se ha adaptado armónicamente (SBR freq. resp.) es mostrado con fines comparativos. Esta señal es calculada con el algoritmo HSBE, pero la frecuencia de retraso subyacente es cero. Esto provoca una ampliación de ancho de banda la cual básicamente corresponde al método SBR. Puede verse claramente aquí el desplazamiento de la estructura armónica en la frecuencia de corte $f_g = 6400$ Hz y el duplicado de esta frecuencia.

A continuación se proporcionan derivaciones de las mismas.

En primer lugar se realiza una derivación de la transferencia de DFT a DCT-II. En particular, se deriva la DCT-II a partir de la DFT en la ecuación 2.10 usando la ecuación 2.12 (véase también Rao y Yip, 2001). Es válido que:

$$\tilde{X}[k] = \sum_{n=0}^{2N-1} \tilde{x}[n] e^{-j\frac{2\pi}{2N}kn} \quad (\text{A.1a})$$

$$= \sum_{n=0}^{N-1} \tilde{x}[n] \left[e^{-j\frac{\pi}{N}kn} + e^{-j\frac{\pi}{N}k(2N-1-n)} \right] \quad (\text{A.1b})$$

$$= \sum_{n=0}^{N-1} \tilde{x}[n] e^{j\frac{\pi}{2N}k} \left[e^{-j\frac{\pi}{N}k(n+0,5)} + e^{-j\frac{\pi}{N}k(2N-0,5-n)} \right] \quad (\text{A.1c})$$

$$= 2e^{j\frac{\pi}{2N}k} \sum_{n=0}^{N-1} \tilde{x}[n] \cos\left(\frac{\pi(n+0,5)k}{N}\right), \quad \forall k \in [0, N-1]$$

con:

$$e^{-j\frac{\pi}{N}k(n+0,5)} + e^{-j\frac{\pi}{N}k(2N-0,5-n)} = \dots$$

$$\begin{aligned}
 &= \cos\left(\frac{\pi}{N}k(n+0,5)\right) - j \operatorname{sen}\left(\frac{\pi}{N}k(n+0,5)\right) + \dots \\
 &+ \cos\left(\frac{\pi}{N}k(2N-0,5-n)\right) - j \operatorname{sen}\left(\frac{\pi}{N}k(2N-0,5-n)\right) + \dots \\
 &= \cos\left(k\left(\frac{\pi}{N}n+0,5\frac{\pi}{N}\right)\right) + \cos\left(-k\left(\frac{\pi}{N}n+0,5\frac{\pi}{N}\right) + 2\pi k\right) + \dots \\
 &- j \operatorname{sen}\left(k\left(\frac{\pi}{N}n+0,5\frac{\pi}{N}\right)\right) - j \operatorname{sen}\left(-k\left(\frac{\pi}{N}n+0,5\frac{\pi}{N}\right) + 2\pi k\right) \quad (\text{A.2b}) \\
 &= 2 \cos\left(\frac{\pi}{N}k(n+0,5)\right) \quad (\text{A.2c})
 \end{aligned}$$

5 y las propiedades de la función de seno y coseno:

$$\cos(-x) = \cos(x) \quad (\text{A.3a})$$

$$\cos(2\pi + x) = \cos(x) \quad (\text{A.3b})$$

$$\operatorname{sen}(-x) = -\operatorname{sen}(x) \quad (\text{A.3c})$$

$$\operatorname{sen}(2\pi + x) = \operatorname{sen}(x) \quad (\text{A.3d})$$

10 Esta correlación también puede usarse para calcular la DCT de manera eficaz usando la FFT (véase también Ahmed *et al.*, 1974).

A continuación se comentan las matrices de convolución.

15 Las matrices de convolución \underline{E}_a y \underline{E}_s requeridas para calcular la MCDT (véase la ecuación 2.35) están compuestas por una matriz de retardo \underline{D} y la matriz ventana \underline{F} . La matriz ventana contiene los coeficientes de la función ventana $w[n]$, dispuestos en una estructura en forma de diamante. Es válido que:

$$\tilde{\mathbf{F}}_s = \tilde{\mathbf{F}}^T \cdot \tilde{\mathbf{D}}^{-1} \quad (\text{A.5a})$$

$$= \begin{bmatrix} 0 & w[0] & w[N]z^{-1} & 0 \\ \vdots & & & \vdots \\ w[\frac{N}{2}-1] & 0 & w[N+\frac{N}{2}-1]z^{-1} & \\ w[\frac{N}{2}] & & -w[N+\frac{N}{2}]z^{-1} & \\ \vdots & & & \vdots \\ 0 & w[N-1] & -w[2N-1]z^{-1} & 0 \end{bmatrix} \quad (\text{A.5b})$$

$$\tilde{\mathbf{F}}_a = \tilde{\mathbf{D}} \cdot \tilde{\mathbf{F}} \quad (\text{A.4a})$$

$$= \begin{bmatrix} 0 & w[\frac{N}{2}-1]z^{-1} & w[\frac{N}{2}]z^{-1} & 0 \\ \vdots & & & \vdots \\ w[0]z^{-1} & 0 & w[N-1]z^{-1} & \\ w[N] & & -w[2N-1] & \\ \vdots & & & \vdots \\ 0 & w[N+\frac{N}{2}-1] & -w[N+\frac{N}{2}] & 0 \end{bmatrix} \quad (\text{A.4b})$$

$$(\text{A.4c})$$

5 con:

$$\tilde{\mathbf{F}} = \begin{bmatrix} 0 & w[\frac{N}{2}-1] & w[\frac{N}{2}] & 0 \\ \vdots & & & \vdots \\ w[0] & 0 & w[N-1] & \\ w[N] & & -w[2N-1] & \\ \vdots & & & \vdots \\ 0 & w[N+\frac{N}{2}-1] & -w[N+\frac{N}{2}] & 0 \end{bmatrix} \quad (\text{A.6})$$

10 y:

$$\tilde{\mathbf{D}} = \begin{bmatrix} z^{-1} & & & & 0 \\ & \ddots & & & \\ & & z^{-1} & & \\ & & & 1 & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} \quad (\text{A.7a})$$

$$\tilde{\mathbf{D}}^{-1} = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & z^{-1} & \\ & & & & \ddots & \\ 0 & & & & & z^{-1} \end{bmatrix} \quad (\text{A.7b})$$

A fin de preservar causalidad, la matriz de retardo inversa $\tilde{\mathbf{D}}^{-1}$ se ha multiplicado por el retardo z^{-1} . Esto es el origen del retardo del banco de filtro MDCT (Schuller and Smith, 1996).

5 **Matriz de transformación H**

La matriz de transformación $\tilde{\mathbf{H}}$ es requerida para transferir el espectro MDCT al respectivo espectro MDST. Lo siguiente es válido:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{T}}_{DST} \cdot \tilde{\mathbf{D}} \cdot \tilde{\mathbf{F}}_{sen} \cdot \tilde{\mathbf{F}}^T \cdot \tilde{\mathbf{D}}^{-1} \cdot \tilde{\mathbf{T}}^{-1} \quad (\text{A.8})$$

10

La descomposición detallada se muestra con el ejemplo de $N = 4$:

$$\tilde{\mathbf{T}}' = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix} \cdot \begin{bmatrix} z^{-1} & 0 & 0 & 0 \\ 0 & z^{-1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{1,1}z^{-1} & a_{1,2}z^{-1} & a_{1,3} & a_{1,4} \\ a_{2,1}z^{-1} & a_{2,2}z^{-1} & a_{2,3} & a_{2,4} \\ a_{3,1}z^{-1} & a_{3,2}z^{-1} & a_{3,3} & a_{3,4} \\ a_{4,1}z^{-1} & a_{4,2}z^{-1} & a_{4,3} & a_{4,4} \end{bmatrix} = \tilde{\mathbf{T}}_{DST} \cdot \tilde{\mathbf{D}} \quad (\text{A.9a})$$

15

$$\begin{aligned} \underline{\mathbf{F}}' &= \begin{bmatrix} 0 & -w[1] & w[2] & 0 \\ -w[0] & 0 & 0 & w[3] \\ w[4] & 0 & 0 & w[7] \\ 0 & w[5] & w[6] & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & w[0] & w[4] & 0 \\ w[1] & 0 & 0 & w[5] \\ w[2] & 0 & 0 & -w[6] \\ 0 & w[3] & -w[7] & 0 \end{bmatrix} \\ &= \begin{bmatrix} w'_{1,1} & 0 & 0 & w'_{1,4} \\ 0 & w'_{2,2} & w'_{2,3} & 0 \\ 0 & w'_{3,2} & w'_{3,3} & 0 \\ w'_{4,1} & 0 & 0 & w'_{4,4} \end{bmatrix} = \underline{\mathbf{F}}_{sen} \cdot \underline{\mathbf{F}}^T \end{aligned} \quad (\text{A.9b})$$

$$\begin{aligned} \underline{\mathbf{T}}'^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & z^{-1} & 0 \\ 0 & 0 & 0 & z^{-1} \end{bmatrix} \cdot \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ b_{3,1} & b_{3,2} & b_{3,3} & b_{3,4} \\ b_{4,1} & b_{4,2} & b_{4,3} & b_{4,4} \end{bmatrix} \\ &= \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ b_{3,1}z^{-1} & b_{3,2}z^{-1} & b_{3,3}z^{-1} & b_{3,4}z^{-1} \\ b_{4,1}z^{-1} & b_{4,2}z^{-1} & b_{4,3}z^{-1} & b_{4,4}z^{-1} \end{bmatrix} = \underline{\mathbf{D}}^{-1} \cdot \underline{\mathbf{T}}^{-1} \end{aligned} \quad (\text{A.9c})$$

Las matrices con los elementos de retardo z^{-1} pueden representarse como adiciones de dos matrices cada una. Resultará lo siguiente:

$$\begin{aligned} \underline{\mathbf{T}}'(z) &= \underline{\mathbf{T}}'(z^0) + \underline{\mathbf{T}}'(z^{-1}) \\ &= \begin{bmatrix} 0 & 0 & a_{1,3} & a_{1,4} \\ 0 & 0 & a_{2,3} & a_{2,4} \\ 0 & 0 & a_{3,3} & a_{3,4} \\ 0 & 0 & a_{4,3} & a_{4,4} \end{bmatrix} + \begin{bmatrix} a_{1,1}z^{-1} & a_{1,2}z^{-1} & 0 & 0 \\ a_{2,1}z^{-1} & a_{2,2}z^{-1} & 0 & 0 \\ a_{3,1}z^{-1} & a_{3,2}z^{-1} & 0 & 0 \\ a_{4,1}z^{-1} & a_{4,2}z^{-1} & 0 & 0 \end{bmatrix} \end{aligned} \quad (\text{A.10a})$$

$$\begin{aligned} \underline{\mathbf{T}}'^{-1}(z) &= \underline{\mathbf{T}}'^{-1}(z^0) + \underline{\mathbf{T}}'^{-1}(z^{-1}) \\ &= \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ b_{3,1}z^{-1} & b_{3,2}z^{-1} & b_{3,3}z^{-1} & b_{3,4}z^{-1} \\ b_{4,1}z^{-1} & b_{4,2}z^{-1} & b_{4,3}z^{-1} & b_{4,4}z^{-1} \end{bmatrix} \end{aligned} \quad (\text{A.10b})$$

5

Con esta descomposición y la matriz $\underline{\mathbf{F}}'$, la ecuación A.8 puede expresarse de la siguiente manera:

$$\underline{\underline{\mathbf{H}}}(z) = \underline{\underline{\mathbf{T}}}'(z) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z) \quad (\text{A.11a})$$

$$= (\underline{\underline{\mathbf{T}}}'(z^0) + \underline{\underline{\mathbf{T}}}'(z^{-1})) \cdot \underline{\underline{\mathbf{F}}}' \cdot (\underline{\underline{\mathbf{T}}}^{-1}(z^0) + \underline{\underline{\mathbf{T}}}^{-1}(z^{-1})) \quad (\text{A.11b})$$

$$= (\underline{\underline{\mathbf{T}}}'(z^0) \cdot \underline{\underline{\mathbf{F}}}' + \underline{\underline{\mathbf{T}}}'(z^{-1}) \cdot \underline{\underline{\mathbf{F}}}') \cdot (\underline{\underline{\mathbf{T}}}^{-1}(z^0) + \underline{\underline{\mathbf{T}}}^{-1}(z^{-1})) \quad (\text{A.11c})$$

$$= \underline{\underline{\mathbf{T}}}'(z^0) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^0) + \underline{\underline{\mathbf{T}}}'(z^{-1}) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^0) + \dots \\ + \underline{\underline{\mathbf{T}}}'(z^0) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^{-1}) + \underline{\underline{\mathbf{T}}}'(z^{-1}) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^{-1}) \quad (\text{A.11d})$$

Los términos individuales de la ecuación son entonces sustituidos mediante su dependencia de z:

$$\underline{\underline{\mathbf{H}}}(z^0) = \underline{\underline{\mathbf{T}}}'(z^0) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^0) \quad (\text{A.12a})$$

$$\underline{\underline{\mathbf{H}}}(z^{-1}) = \underline{\underline{\mathbf{T}}}'(z^{-1}) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^0) + \underline{\underline{\mathbf{T}}}'(z^0) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^{-1}) \quad (\text{A.12b})$$

$$\underline{\underline{\mathbf{H}}}(z^{-2}) = \underline{\underline{\mathbf{T}}}'(z^{-1}) \cdot \underline{\underline{\mathbf{F}}}' \cdot \underline{\underline{\mathbf{T}}}^{-1}(z^{-1}) \quad (\text{A.12c})$$

5 e insertados en la ecuación A.11a:

$$\underline{\underline{\mathbf{H}}}(z) = \underline{\underline{\mathbf{H}}}(z^0) + \underline{\underline{\mathbf{H}}}(z^{-1}) + \underline{\underline{\mathbf{H}}}(z^{-2}) \quad (\text{A.13})$$

10 El resultado es la notación simplificada de la ecuación 4.7:

$$\underline{\underline{\mathbf{H}}} = \underline{\underline{\mathbf{H}}}_0 z^0 + \underline{\underline{\mathbf{H}}}_1 z^{-1} + \underline{\underline{\mathbf{H}}}_2 z^{-2} \quad (\text{A.14})$$

15 El análisis numérico de las tres submatrices $\underline{\underline{\mathbf{H}}}_0$, $\underline{\underline{\mathbf{H}}}_1$ y $\underline{\underline{\mathbf{H}}}_2$ proporciona algunas observaciones importantes las cuales contribuyen a reducir la complejidad de cálculo de la matriz $\underline{\underline{\mathbf{H}}}$ de manera considerable. Deberá notarse primero que $\underline{\underline{\mathbf{H}}}_1$ es una matriz ocupada escasamente. Siempre exhibirá la misma forma para diferentes longitudes de transformada N e incluye elementos 0,5 y -0,5 exclusivamente. Lo siguiente es válido:

$$\underline{\underline{\mathbf{H}}}_1 = \begin{bmatrix} 0,5 & -0,5 & 0 & & \dots & & 0 \\ 0,5 & 0 & -0,5 & 0 & & & \\ 0 & 0,5 & 0 & -0,5 & 0 & & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & 0 & 0,5 & 0 & -0,5 & 0 \\ & & & 0 & 0,5 & 0 & -0,5 \\ 0 & & \dots & & 0 & 0,5 & 0,5 \end{bmatrix} \quad (\text{A.15})$$

20 Existe una correlación directa entre las matrices $\underline{\underline{\mathbf{H}}}_0$ y $\underline{\underline{\mathbf{H}}}_2$, de modo que

$$\underline{\underline{\mathbf{H}}}_2 = \underline{\underline{\mathbf{\Pi}}}_N \cdot \underline{\underline{\mathbf{H}}}_0^T \cdot \underline{\underline{\mathbf{\Pi}}}_N \quad (\text{A.16})$$

25 con:

$$\underline{\mathbf{P}}_N = \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \dots & 0 & 0 \end{bmatrix} \quad (\text{A.17})$$

Así, $\underline{\mathbf{P}}_{\leq N}$ es una matriz $N \times N$ la cual, al multiplicarse desde la izquierda, revierte el orden de las líneas y, al multiplicarse desde la derecha, invierte el orden de las columnas. Así, la ecuación A.16 puede interpretarse como un

5 reflejo de los valores de la matriz $\underline{\mathbf{H}}_0$ en sus diagonales secundarias. Con estas propiedades, la complejidad requerida para calcular $\underline{\mathbf{H}}$ puede reducirse de las $4N^3$ operaciones necesarias originalmente (véase la ecuación A.11d) a un cuarto de las mismas.

10 La DCT-IV rápida se explica a continuación. La implementación de la DCT-IV depende del algoritmo de la DCT-IV rápida. La ventaja de esta realización es el cálculo eficaz de la transformada y el retardo algorítmico corto asociado a ello. El núcleo de la DCT-IV son dos transformadas DCT-III conectadas en paralelo de acuerdo con la ecuación 2.14c. De manera similar a la FFT, se compone de una denominada estructura de mariposa y una estructura de

15 tubería (Rao y Yip, 2001). La complejidad de este algoritmo es de $O(N \log N)$ y es comparable a la complejidad de cálculo requerida de la FFT. La estructura precisa de la DCT-III se ilustra en la Fig. 19. En particular, la Fig. 19 ilustra una estructura de DCT-III/DST-III rápida universal (Rao y Yip, 2001).

20 La secuencia de entrada x_n es transferida por tanto a los valores espectrales DCT-II $\underline{\mathbf{F}}_a$. Dado que la DCT-III representa la transformada inversa a DCT-II, las dos transformadas pueden calcularse mediante esta estructura. Esta estructura universal para calcular la DCT-III/DST-III y las transformadas inversas respectivas DCT-II y DST-II constituyen la base para la DCT-IV de acuerdo con la ecuación 2.14d. La Fig. 20 muestra la estructura de la DCT-IV implementada. Así, la Fig. 20 muestra una estructura de DCT-IV rápida (Rao y Yip, 2001).

25 Las estructuras mostradas pueden realizarse por ejemplo en el lenguaje de programación C. Las funciones correspondientes son `dct_processor()` para la DCT-III/DST-III (DST-II/DST-II) universal, y `fdctstiv1d()` para la DCT-IV/DST-IV rápida. Ninguna variante rápida de esta transformada es implementada en Matlab ya que el rendimiento del algoritmo no es importante aquí.

La siguiente sección sirve para describir las funciones Matlab así como funciones del lenguaje de programación C.

30 **Funciones Matlab:**

decoder_hsbe()

Descripción:

35 `decoder_hsbe()` es un decodificador para ampliación de banda espectral armónica (HSBE). Una señal de paso bajo en el dominio MDCT se amplía así espectralmente de modo que la estructura armónica de la señal se mantiene. Para reconstruir la curva envolvente se realiza filtración LPC inversa. La señal es transferida al dominio de tiempo por medio de una MDCT inversa.

40 Prototipo:

`[y varargout] = decoder_hsbe(X_core, T_idct, G, sideinfo, ...
aaf_dat, b_noisefill, b_useMDST);`

45 Salida:

Y
 varargout(1) señal de tiempo reconstruida
 varargout(2) espectro MDCT ampliado antes de la adaptación de curva envolvente
 50 varargout(3) espectro MDST ampliado antes de la adaptación de curva envolvente
 varargout(3) espectro complejo antes de la adaptación de curva envolvente

Entrada:

5 X_core banda base de la señal en el dominio MDCT
 T_idct matriz de transformación DCT-IV inversa
 G matriz de convolución inversa para la iMDCT en forma de diamante
 Sideinfo información lateral
 aaf_dat base de datos con coeficientes de filtrado antisolapamiento

10 Opcional:

b_noisefill booleano, indica si ha de añadirse ruido. Estándar: falso
 b_useMDST booleano, indica si ha de usarse la MDST precalculada. Si es verdadero, X_core debe ser complejo. Estándar: falso

15 **AAFCreator.m**

Descripción:

20 Herramienta de creación de filtro FIR - filtro antisolapamiento (AAF) "AAFCreator": La herramienta de creación de filtro FIR - AAF calcula los filtros antisolapamiento requeridos para cancelación de solapamiento con desplazamiento de subbanda del algoritmo HSBE. Los filtros son filtros FIR en una base de datos. La base de datos se almacena como un archivo .mat y contiene una estructura unidimensional. Cada entrada en la base de datos incluye una indicación del ángulo para el cual se usará el filtro FIR y los coeficientes de filtrado como vector.

25 Prototipo:

Script de Matlab, sin nombre funcional.

Salida:

30 .bmp mapas de bits con la respuesta de impulso AAF y el resultado de cancelación de solapamiento
 .mat base de datos con los filtros AAF en formato de archivo MAT
 .wav respuesta de impulso AAF como un archivo de audio RIFF

35 Entrada:

Fs frecuencia de muestreo de la señal de prueba
 NMDCT longitud de transformada N
 sig_typ tipo de señal de prueba
 40 f frecuencia de la señal de prueba
 t duración de la señal de prueba
 phi ángulo de fase ϕ , varios ángulos posibles
 num_opt número de operaciones de optimización
 num_alias orden más alto de las componentes de solapamiento que han de suprimirse
 45 b_save_wav booleano, indica si los archivos .wav han de almacenarse
 b_save_bmp booleano, indica si los archivos .bmp han de almacenarse
 b_break booleano, indica si se debe hacer un descanso con resultados intermedios
 path_bmp directorio para los archivos .bmp
 path_wav directorio para los archivos .wav
 50

AAFOptimizer()

Descripción:

55 AAFOptimizer() genera un filtro antisolapamiento (AAF) como filtro FIR. El algoritmo optimiza de manera numérica las componentes de solapamiento detectadas por un reconocimiento de picos por medio de aproximación sucesiva.

Prototipo:

60 [fir x_antialias] = AAFOptimizer(x, X_cmpx, T_imdct, ...T_imdst, varargin);

Salida:

Fir respuesta de impulso de filtro del filtro antisolapamiento en el dominio de tiempo

x_antialias señal de entrada x libre de solapamiento por fir

Entrada:

5 X señal de tiempo dotada de componentes de solapamiento
 X_cmpx representación compleja de x en el dominio MDCT/MDST
 T_imdct matriz MDCT inversa
 T_imdst matriz MDST inversa

10 Opcional:

 varargin(1) orden más alto de las componentes de solapamiento que han de suprimirse. Estándar: 4
 varargin(2) número de operaciones de optimización; estándar: 3
 varargin(3) ángulo de fase ϕ con el cual se ha modulado x; estándar: 0
 15 varargin(4) booleano, indica si han de establecerse trazados; estándar: falso
 varargin(5) booleano, indica si los trazados establecidos han de almacenarse; estándar: falso
 varargin(6) directorio para los archivos .bmp; estándar: cero

Funciones C

20

HSBEprocessing()

Descripción:

25 HSBEprocessing() es la interfaz principal para la biblioteca estática hsbe.lib establecida para la ampliación de ancho de banda por medio de HSBE. En la versión actual 1.0, un fragmento de señal en el dominio de tiempo es transformado por MDCT/MDST y expandido espectralmente de modo que la estructura armónica se mantiene. La señal de ancho de banda ampliado compleja es transferida de nuevo al dominio de tiempo y emitida. Una reconstrucción de curva envolvente, como se hace en la función Matlab decoder_hsbe(), no tiene lugar.

30

Prototipo:

 HSBE_RESULT HSBEprocessing(HSBE_HANDLE HSBEhandle, ...
 float *pSamplesIn, float *pSamplesOut, ...
 35 unsigned short nSamplesIn, float frequencyOffset);

Salida:

 HSBE_RESULT código de error
 40 pSamplesOut fasor para el vector de la señal de tiempo de ancho de banda ampliado

Entrada:

 HSBEhandle estructura de puntero inteligente (*handle*) de HSBE
 45 pSamplesIn fasor al vector de la señal de entrada
 nSamplesIn longitud de transformada N
 frequencyOffset frecuencia de retraso f_{lag} como una frecuencia normalizada

HSBEinit()

50

Descripción:

 HSBEinit() es la función para inicializar HSBE. Aquí se reserva la memoria funcional requerida, se calcula la función ventana y se calcula la matriz de transformación MDCT/MDST \mathbf{H} . Además, todos los parámetros requeridos para
 55 HSBEprocessing() se introducen en el puntero inteligente.

Prototipo:

 HSBE_RESULT HSBEinit(HSBE_HANDLE *HSBEhandle, ...
 60 unsigned short nSamplesIn, float f_core, float f_hsbe, ...
 unsigned char windowtype, char flags);

Salida:

HSBE_RESULT código de error

5 HSBEhandle fasor a la estructura de puntero inteligente de HSBE

Entrada:

10 HSBEhandle fasor a la estructura de puntero inteligente de HSBE
 nSamplesIn longitud de transformada N
 f_core frecuencia de corte f_g de la banda base como frecuencia normalizada
 f_hsbe frecuencia máxima que ha reconstruirse como frecuencia normalizada
 windowtype tipo de la función ventana que ha de usarse
 flags banderas de control

15 **HSBEfree()**

Descripción:
 HSBE free() libera los recursos demandados por HSBEinit().

20 Prototipo:

HSBE_RESULT HSBEfree(HSBE_HANDLE*HSBEhandle);

Salida:

25 HSBE_RESULT código de error

HSBEhandle fasor a la estructura de puntero inteligente de HSBE

30 Entrada:

HSBEhandle fasor a la estructura de puntero inteligente de HSBE

35 Aunque se han descrito algunos aspectos en conexión con un dispositivo, deberá entenderse que estos aspectos representan también una descripción del método correspondiente, de modo que un bloque o elemento constructivo de un dispositivo también deberá entenderse como un paso del método correspondiente o una característica de un paso del método. Por analogía, los aspectos que se han descrito en conexión con un paso de método o como un paso de método también representan una descripción de un bloque o detalle o característica correspondiente de un dispositivo correspondiente. Algunos o todos los pasos de método pueden realizarse mediante aparatos de hardware (o usando un aparato de hardware), tales como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunos ejemplos de realización, algunos o varios de los pasos de método más importantes pueden realizarse mediante un aparato de este tipo.

45 Según los requerimientos de implementación específicos, los ejemplos de realización de la invención pueden implementarse ya sea en hardware o software. La implementación puede realizarse usando un medio de almacenamiento digital, tal como, por ejemplo, un disco flexible, un DVD, un disco Blu-ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, una unidad de disco duro u otra memoria magnética u óptica en la cual están almacenadas señales de control legibles electrónicamente que pueden interactuar o interactúan con un sistema informático programable de modo que se ejecute el método respectivo. Así, el medio de almacenamiento digital puede ser legible por ordenador.

55 Algunos ejemplos de realización de acuerdo con la invención incluyen un soporte de datos que presenta señales de control legibles electrónicamente las cuales pueden de interactuar con un sistema informático programable de modo que se realiza uno de los métodos descritos aquí.

Generalmente, los ejemplos de realización de la presente invención pueden implementarse como un producto de programa informático que comprende un código de programa, siendo el código de programa efectivo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador.

60 El código de programa puede almacenarse por ejemplo en un soporte legible por máquina.

Otros ejemplos de realización comprenden el programa informático para realizar uno de los métodos descritos aquí, estando el programa informático almacenado en un soporte legible por máquina. En otras palabras, un ejemplo de realización del método inventivo es un programa informático que presenta un código de programa para realizar uno

de los métodos descritos aquí cuando el programa informático se ejecuta en un ordenador.

Así, otro ejemplo de realización del método inventivo es un soporte de datos (o un medio de almacenamiento digital o un medio legible por ordenador) en el cual se ha grabado el programa informático para realizar uno de los métodos descritos aquí.

Así, otro ejemplo de realización del método inventivo es una corriente de datos o una secuencia de señales que representa o representan el programa informático para realizar uno de los métodos descritos aquí. La corriente de datos o la secuencia de señales puede o pueden estar configuradas por ejemplo para transferirse por medio de un enlace de comunicaciones de datos, por ejemplo por medio de internet.

Otro ejemplo de realización comprende un equipo de procesamiento, como, por ejemplo, un ordenador o un dispositivo de lógica programable configurado o adaptado para realizar uno de los métodos descritos aquí.

Otro ejemplo de realización comprende un ordenador en el cual se ha instalado el programa informático para realizar uno de los métodos descritos aquí.

Otro ejemplo de realización de acuerdo con la invención comprende un dispositivo o un sistema configurado para transmitir un programa informático para realizar cuando menos uno de los métodos descritos aquí a un receptor. La transmisión puede tener lugar por ejemplo de manera electrónica u ópticamente. El receptor puede ser por ejemplo un ordenador, un dispositivo móvil, un dispositivo de almacenamiento o similar. El dispositivo o sistema puede comprender por ejemplo un servidor de archivos para transmitir el programa informático al receptor.

En algunos ejemplos de realización, puede usarse un dispositivo de lógica programable (por ejemplo una disposición de puertas programables en campo (FPGA, por sus siglas en inglés)) para realizar algunas o todas las funcionalidades de los métodos descritos aquí. En algunos ejemplos de realización, una disposición de puertas programables en campo puede interactuar con un microprocesador para realizar uno de los métodos descritos aquí. En general, en algunos ejemplos de realización, los métodos son realizados por parte de cualquier dispositivo de hardware. Este puede ser hardware de uso universal, como un procesador de ordenador (CPU), o hardware específico para el método, como, por ejemplo, un ASIC.

Los ejemplos de realización descritos anteriormente representan únicamente una ilustración de los principios de la presente invención. Deberá entenderse que modificaciones y variantes de las disposiciones y particularidades descritas aquí resultarán evidentes para otros expertos en la técnica. Por consiguiente, solo se pretende que la invención esté limitada por el alcance de las siguientes reivindicaciones, pero no por los detalles específicos que se han presentado aquí por medio de la descripción y explicación de los ejemplos de realización.

Significado de las abreviaturas:

40	AAC	Advanced Audio Coding
	AAF	Anti-Aliasing-Filter
	ac-3	Dolby Digital Adaptive Transform Coder 3
	BWE	ampliación de ancho de banda (Bandwidth Extension)
	CD	Compact Disc
45	CM-BWE	ampliación de ancho de banda por medio de modulación continua (Continuously Modulated Bandwidth Extension)
	DCT	transformada de coseno discreta
	DFT	transformada de Fourier discreta
	DST	transformada de seno discreta
50	DVD	Digital Versatile Disc
	FFT	transformada de Fourier rápida (Fast Fourier Transform)
	FIR	respuesta de impulso finito (Finite Impulse Response)
	HBE	ampliación de ancho de banda armónica (Harmonic Bandwidth Extension)
55	HE-AAC	Advanced Audio Coding de alta eficacia por MPEG-4 (MPEG-4 High Efficiency Advanced Audio Coding)
	HF	alta frecuencia
	HSBE	ampliación de banda espectral armónica (Harmonic Spectral Bandwidth Extension)
	JPEG	Joint Photographic Experts Group
	KiB	kibibyte 2^{10} bytes = 1024 bytes
60	LOT	transformada ortogonal superpuesta (Lapped Orthogonal Transform)
	LPC	codificación predictiva lineal (Linear Predictive Coding)
	LTI	invariante de tiempo lineal (linear time-invariant)
	MDCT	transformada de coseno discreta modificada
	MDST	transformada de seno discreta modificada

	MiB	mebibyte = 2^{20} bytes = 1048576 bytes
	mp3	MPEG-1 Audio Layer III
	PC	Personal Computer
	PCM	Pulse Code Modulation
5	PR	reconstrucción perfecta
	QMF	filtro de espejo en cuadratura (Quadrature Mirror Filter)
	SBR	replicación de banda espectral (Spectral Band Replication)
	SNR	relación señal a ruido (Signal-To-Noise Ratio)
	TDA	distorsión de dominio de tiempo (Time Domain Aliasing)
10	TDAC	compensación de solapamiento de dominio de tiempo (Time Domain Aliasing Cancelation)
	USAC	Unified Speech and Audio Coding
	VoIP	Voice over Internet Protocol

Significado de los símbolos usados:

15	e	número de Euler
	j	unidad imaginaria
	Im	parte imaginaria de una función
	ld	logaritmo de base dos
20	log	logaritmo de base diez
	Re	parte real de una función
	\mathcal{O}	símbolo Landau para estimación de la complejidad
	T	transpuesto
25	$\lceil \cdot \rceil$	operador de redondeo
	$\lfloor \cdot \rfloor$	operador de redondeo para redondear hacia abajo
	\mathbb{R}	cantidad de números reales
	\mathbb{N}	cantidad de números enteros no negativos
	\mathbb{N}^+	cantidad de números enteros positivos
30	\mathbb{Z}	cantidad de números enteros
	ϵ	factor de normalización, \tilde{H}_0
	σ	factor de estirado, índice de ventana, \tilde{H}_1
	τ	índice de modulación, \tilde{H}_2
	φ	ángulo de fase, \tilde{H}'_0
35	ω	frecuencia normalizada, F_α
	b	índice de bloque, F_s
	B	ancho de banda, D^{-1}
	f	frecuencia, $f \in \mathbb{R}$
	k	índice de frecuencia discreto, F'
40	m	índice de fase discreto, \mathcal{O}
	M	numero de fases, \mathbb{R}
	n	índice de tiempo discreto, \mathbb{N}
	N	longitud de transformada, \mathbb{N}^+
	t	tiempo, \mathbb{Z}
45	W	fasor de rotación complejo
	$h[n]$	respuesta de impulso de un filtro FIR
	$H[z]$	función de transferencia de $h[n]$ en el dominio z
	$\rho_m[n]$	fase emésima de una señal discreta
	$w[n]$	función ventana en el dominio de tiempo discreto
50	$x[n]$	señal real en el dominio de tiempo discreto
	$x(t)$	señal real en el dominio de tiempo continuo

	$X(f)$	señal en el dominio de frecuencia continua
	$X[k]$	señal en el dominio de frecuencia discreta
	\underline{h}	respuesta de impulso $h[n]$ en notación vectorial
	\underline{H}	$H[z]$ en notación vectorial
5	\underline{x}	señal $x[n]$ en notación vectorial
	\underline{D}	matriz de retardo
	\underline{F}	matriz ventana en forma de diamante
	$\underline{F}_{a,s}$	matriz de convolución
	\underline{H}	matriz de transformación polifásica
10	$\underline{P}_{a,s}$	matriz polifásica
	\underline{T}	matriz de transformación
	\underline{x}	matriz de bloques del vector de señal \underline{x}
	\underline{X}	señal transformada

Bibliografía:

- 5 [Ahmed *et al.* 1974] Ahmed, N. ; Natarajan, T. ; Rao, K.R.: Discrete Cosine Transform. En: *Computers, IEEE Transactions on C-23* (1974), enero, n.º 1
- [Bosi y Goldberg 2003] Bosi, M. ; Goldberg, R. E.: *Introduction a Digital Audio Coding and Standards*. 2nd edition. Boston; Dordrecht; Londres : Kluwer Academic Publishers, 2003
- 10 [Britanak *et al.*2007] Britanak, V. ; Yip, P. C. ; Rao, K. R.: *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms y Integer Approximations*. Ámsterdam; entre otros : Elsevier, 2007
- [Burrus y Parks 1985] Burrus, C. S. ; Parks, T.: *DFT/FFT y Convolution Algorithms: Theory and Implementation*. Nueva York : John Wiley & Sons Ltd., 1985
- 15 [Cheng 2004] Cheng, C.: Method for Estimating Magnitude and Phase in the MDCT Domain. En: *Audio Engineering Society Convention 116*, mayo 2004
- [Dolson 1986] Dolson, M: The Phase Vocoder: A Tutorial. En: *Computer Music Journal* 10 (1986), n.º 4
- 20 [Ekstrand 2002] Ekstrand, P.: Bandwidth Extension of Audio Signals by Spectral Band Replication. En: *Proceedings of 1st IEEE Benelux Workshop on MPCA, Leuven, Bélgica*. Tomo 1, noviembre 2002
- [Fastl y Zwicker 2007] Fastl, H. ; Zwicker, E.: *Psychoacoustics: Facts and Models*. 3ª edición. Berlín; Heidelberg; Nueva York : Springer, 2007
- 25 [Jordan-Engeln y Reutter 1978] Jordan-Engeln, G. ; Reutter, F.: *Numerische Mathematik für Ingenieure*. 2ª ed. revisada. Mannheim : Bibliographisches Institut, 1978
- [Kiencke y Jäkel 2005] Kiencke, U. ; Jäkel, H.: *Signale und Systeme*. 3ª ed. revisada. Múnich; Viena : Oldenburg Verlag, 2005
- 30 [Larsen y Aarts 2004] Larsen, E. ; Aarts, R. M.: *Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design*. Chichester : John Wiley & Sons Ltd., 2004
- 35 [Lochmann 1990] Lochmann, D.: *Digitale Nachrichtentechnik: Digitale Modulation und Signalverarbeitung*. Tomo 1. Berlín : VBE Verlag Technik, 1990
- [Malvar 1992] Malvar, H. S.: *Signal Processing with Lapped Transforms*. Boston; Londres: Artech House, 1992
- 40 [Malvar y Staelin 1989] Malvar, H. S. ; Staelin, D. H.: The LOT: transform coding without blocking effects. En: *Acoustics, Speech and Signal Processing, IEEE Transactions on* 37 (1989), abril, n.º 4
- [Nagel y Disch 2009] Nagel, F. ; Disch, S.: A harmonic bandwidth extension method for audio codecs. En: *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference*, abril 2009
- 45 [Nagel *et al.* 2010] Nagel, F. ; Disch, S. ; Wilde, S.: A continuous modulated single sideband bandwidth extension. En: *Acoustics Speech y Signal Processing (ICASSP), 2010 IEEE International Conference*, marzo 2010
- 50 [Princen y Bradley 1986] Princen, J. ; Bradley, A.: Analysis/Synthesis filter bank design based on time domain aliasing cancellation. En: *Acoustics, Speech y Signal Processing, IEEE Transactions on* 34 (1986), oct, n.º 5
- [Princen *et al.* 1987] Princen, J. ; Johnson, A. ; Bradley, A.: Subband/Transform coding using filter bank designs based on time domain aliasing cancellation. En: *Acoustics, Speech, y Signal Processing, IEEE International Conference on ICASSP '87*. Tomo 12, abril 1987
- 55 [Rao y Yip 2001] Rao, K. R. ; Yip, P. C.: *The Transform and Data Compression Handbook*. Londres; Nueva York; Washington, D.C. : CRC Press, 2001
- 60 [Schuller y Smith 1996] Schuller, G. D. T. ; Smith, M. J. T.: New framework for modulated perfect reconstruction filter banks. En: *Signal Processing, IEEE Transactions on* 44 (1996), agosto, n.º 8
- [Weinzierl 2008] Weinzierl, S.: *Handbuch der Audiotechnik*. Berlín; Heidelberg: Springer, 2008

[Wilde 2009] Wilde, Stephan: *Entwicklung von Zeitbereichsverfahren zur Bandbreitenerweiterung von Audiosignalen*, Friedrich-Alexander-Universität Erlangen-Nürnberg, tesis doctoral, 2009

[Yost 1994] Yost, W. A.: *Fundamentals of Hearing: An Introduction*. 3^a ed. San Diego; Nueva York; Boston; Londres; Sidney; Tokio : Academic Press, 1994

5 [Zölzer 2005] Zölzer, U.: *Digitale Audiosignalverarbeitung*. 3^a ed. revisada y ampliada. Stuttgart; Leipzig; Wiesbaden : Teubner, 2005

REIVINDICACIONES

1. Un dispositivo (100; 150; 300) para producir una señal de audio desplazada en frecuencia basándose en una señal de entrada de audio, en el que la señal de entrada de audio, para una pluralidad de primeras subbandas, está representada por uno o varios primeros valores de subbanda, comprendiendo el dispositivo (100; 150; 300):

una interfaz (110; 310) para recibir la señal de entrada de audio, y
 una unidad de desplazamiento en frecuencia (120; 320) para producir la señal de audio desplazada en frecuencia, presentando la señal de audio desplazada en frecuencia, para una pluralidad de segundas subbandas, en cada caso uno o varios segundos valores de subbanda, presentando cada uno de los primeros y los segundos valores de subbanda información de un ángulo de fase respectivo, y
 estando la unidad de desplazamiento en frecuencia (120; 320) configurada para producir uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de modo que el segundo ángulo de fase de este segundo valor de subbanda difiere del primer ángulo de fase de este primer valor de subbanda por una diferencia de ángulo de fase, dependiendo la diferencia de ángulo de fase de una información de frecuencia que indica por qué diferencia de frecuencia se va a desplazar la señal de entrada de audio a fin de obtener la señal de audio desplazada en frecuencia, **caracterizado por que** la diferencia de ángulo de fase depende de un ancho de banda de frecuencia de una de las primeras subbandas.

2. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 1, en el que la unidad de desplazamiento en frecuencia (120; 320) está configurada para determinar una de las primeras subbandas basándose en la información de frecuencia para cada una de las segundas subbandas, estando la unidad de desplazamiento en frecuencia (120; 320) configurada de manera adicional para producir el uno o varios de los segundos valores de subbanda de la segunda subbanda respectiva basándose en uno o varios de los primeros valores de subbanda de la primera subbanda respectiva determinada, estando la unidad de desplazamiento en frecuencia (120; 320) configurada para producir en cada caso uno de los segundos valores de subbanda de la respectiva segunda subbanda basándose en cada caso en una de las primeras subbandas de la respectiva primera subbanda determinada de modo que la información del ángulo de fase del respectivo segundo valor de subbanda depende de la información del ángulo de fase del respectivo primer valor de subbanda y de una relación de la diferencia de frecuencia por la que se va a desplazar la señal de entrada de audio con respecto al ancho de banda de frecuencia de esta subbanda.

3. Un dispositivo (100; 150; 300) de conformidad con las reivindicaciones 1 o 2, en el que cada uno de los primeros valores de subbanda y cada uno de los segundos valores de subbanda pueden ser representados como números complejos, estando la unidad de desplazamiento en frecuencia (120, 320) configurada para determinar un valor de resultado de una función exponencial compleja, y estando la unidad de desplazamiento en frecuencia (120; 320) configurada para producir al menos uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda al multiplicar un primer número complejo que representa uno de los primeros valores de subbanda por el valor de resultado a fin de obtener un segundo número complejo que representa uno de los segundos valores de subbanda.

4. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 3, en el que la unidad de desplazamiento en frecuencia (120; 320) está configurada para determinar el valor de resultado de la función exponencial compleja, presentando la función exponencial compleja la fórmula:

$$e^{-j \cdot b \cdot \varphi \cdot \frac{\pi}{180^\circ}}$$

donde j es una unidad imaginaria, b es un numero entero, y φ es un ángulo en grados, el cual depende de la diferencia de frecuencia por la que se van a desplazar los primeros valores de subbanda de las subbandas.

5. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones precedentes, en el que la unidad de desplazamiento en frecuencia (120; 320) está configurada para producir uno de los segundos valores de subbanda basándose en uno los primeros valores de subbanda, y siendo los primeros valores de subbanda primeros coeficientes MDCT de la señal de entrada de audio los cuales son coeficientes de una transformada de coseno discreta modificada de la señal de entrada de audio, o siendo los primeros valores de subbanda primeros coeficientes QMF de la señal de entrada de audio los cuales son coeficientes de una transformada QMF de la señal de entrada de audio.

6. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones precedentes, en el que el dispositivo (100; 150; 300) comprende adicionalmente una unidad de transformación MDCT/MDST (315), estando la unidad de transformación MDCT/MDST (315) configurada para obtener uno o varios primeros coeficientes MDCT de la señal de entrada de audio los cuales son coeficientes de una transformada de coseno discreta modificada de la señal de entrada de audio,

estando la unidad de transformación MDCT/MDST (315) configurada para determinar, basándose en uno o varios de los primeros coeficientes MDCT de la señal de entrada de audio, uno o varios primeros coeficientes MDST de la señal de entrada de audio los cuales son coeficientes de una transformada de seno discreta modificada, y estando la unidad de desplazamiento en frecuencia (120; 320) configurada para producir cada uno de los segundos valores de subbanda basándose en cada caso en uno de los primeros valores de subbanda, basándose cada uno de los primeros valores de subbanda en uno de los primeros coeficientes MDCT y uno de los primeros coeficientes MDST, el cual se ha determinado basándose en este primer coeficiente MDCT.

5
10
15

7. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 6, en el que la unidad de desplazamiento en frecuencia (120; 320) está configurada para producir cada uno de los segundos valores de subbanda basándose en cada caso en uno de los primeros valores de subbanda, siendo cada uno de los primeros coeficientes MDCT y cada uno de los primeros coeficientes MDST de valor real, y representando cada uno de los primeros valores de subbanda una suma de uno de los primeros coeficientes MDCT y uno de los primeros coeficientes MDST multiplicado por la unidad imaginaria j que se ha determinado basándose en este primer coeficiente MDCT.

20
25

8. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones 6 o 7, en el que la unidad de transformación MDCT/MDST (315) está configurada para determinar el uno o varios primeros coeficientes MDST de la señal de entrada de audio basándose en uno o varios de los primeros coeficientes MDCT de la señal de entrada de audio, pudiendo representarse los primeros coeficientes MDCT de la señal de entrada de audio como coeficientes de matriz de una matriz de coeficientes MDCT \underline{X} la cual comprende una o varias columnas, estando la señal de entrada de audio subdividida en bloques, representando cada una de las columnas de la matriz de coeficientes MDCT \underline{X} uno de los bloques de la señal de entrada de audio,

30

formando cada una de las columnas de la matriz de coeficientes MDCT \underline{X} un espectro MDCT del respectivo bloque de la señal de entrada de audio, y determinando la unidad de transformación MDCT/MDST (315) los coeficientes de matriz de una matriz de transformación \underline{H} de modo que, mediante una multiplicación de matrices de la matriz de transformación \underline{H} por la matriz de coeficientes MDCT \underline{X} , se obtiene una matriz de coeficientes MDST \underline{X}_{MDST} cuyos coeficientes de matriz representan los primeros coeficientes MDST.

35
40

9. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones 6 o 7, en el que la unidad de transformación MDCT/MDST (315) está configurada para determinar el uno o varios primeros coeficientes MDST de la señal de entrada de audio basándose en uno o varios de los primeros coeficientes MDCT de la señal de entrada de audio, pudiendo representarse los primeros coeficientes MDCT de la señal de entrada de audio como coeficientes de matriz de una matriz de coeficientes MDCT \underline{X} la cual comprende una o varias columnas, estando la señal de entrada de audio subdividida en bloques, representando cada una de las columnas de la matriz de coeficientes MDCT \underline{X} uno de los bloques de la señal de entrada de audio, y estando la unidad de transformación MDCT/MDST (315) configurada para aplicar una fórmula:

$$\underline{X}_{MDST}(b-1) = \underline{H}_0 \cdot \underline{X}(b) + \underline{H}_1 \cdot \underline{X}(b-1) + \underline{H}_2 \cdot \underline{X}(b-2)$$

45
50

a fin de determinar los primeros coeficientes MDST, donde $\underline{X}(b)$ es la columna b-ésima de la matriz \underline{X} , donde $\underline{X}(b-1)$ es la columna b-1-ésima de la matriz \underline{X} , donde $\underline{X}(b-2)$ es la columna b-2-ésima de la matriz \underline{X} , donde \underline{H}_0 es una primera submatriz, donde \underline{H}_1 es una segunda submatriz, donde \underline{H}_2 es una tercera submatriz, y donde $\underline{X}_{MDST}(b-1)$ representa un espectro MDST del bloque b-1 el cual comprende los primeros coeficientes MDST.

10. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 9, en el que la unidad de transformación MDCT/MDST (315) está configurada para aplicar la fórmula:

$$\underline{X}_{MDST}(b-1) = \underline{H}_0 \cdot \underline{X}(b) + \underline{H}_1 \cdot \underline{X}(b-1) + \underline{H}_2 \cdot \underline{X}(b-2)$$

5 a fin de determinar los primeros coeficientes MDST,

teniendo dos o varios coeficientes de matriz de la primera submatriz \underline{H}_0 el valor 0, teniendo dos o varios coeficientes de matriz de la segunda submatriz \underline{H}_1 el valor 0, y teniendo dos o más coeficientes de matriz de la tercera submatriz \underline{H}_2 el valor 0.

10 11. Un dispositivo (100; 150; 300) de conformidad con las reivindicaciones 9 o 10, en el que la unidad de transformación MDCT/MDST (315) está configurada para aplicar la fórmula:

$$\underline{X}_{MDST}(b-1) = \underline{H}_0 \cdot \underline{X}(b) + \underline{H}_1 \cdot \underline{X}(b-1) + \underline{H}_2 \cdot \underline{X}(b-2)$$

15 a fin de determinar los primeros coeficientes MDST,

presentando la segunda submatriz \underline{H}_1 la siguiente fórmula:

$$\underline{H}_1 = \begin{bmatrix} 0,5 & -0,5 & 0 & & \dots & & 0 \\ 0,5 & 0 & -0,5 & 0 & & & \\ 0 & 0,5 & 0 & -0,5 & 0 & & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & 0 & 0,5 & 0 & -0,5 & 0 \\ & & & 0 & 0,5 & 0 & -0,5 \\ 0 & & \dots & & 0 & 0,5 & 0,5 \end{bmatrix}$$

donde “...” significa que los mismos valores de coeficientes de matriz previos de la segunda submatriz \underline{H}_1 continúan

20 en coeficientes de matriz subsecuentes de la segunda submatriz \underline{H}_1 .

12. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones precedentes, en el que el dispositivo (100; 150; 300) comprende adicionalmente una unidad de filtro (330), estando la unidad de filtro (330) configurada para filtrar los segundos valores de subbanda aplicando un filtro que presenta una pluralidad de coeficientes de filtrado de modo que se reducen las componentes espurias de los segundos valores de subbanda que se formaron al producir los segundos valores de subbanda basándose en los primeros valores de subbanda.

13. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 12, en el que la unidad de desplazamiento en frecuencia (120; 320) está configurada para producir uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de tal manera que el segundo ángulo de fase de este segundo valor de subbanda difiere del primer ángulo de fase este primer valor de subbanda por la diferencia de ángulo de fase, estando la unidad de filtro (330) configurada para filtrar los segundos valores de subbanda aplicando el filtro, y dependiendo uno o varios de los coeficientes de filtrado de la diferencia de ángulo de fase.

14. Un dispositivo (100; 150; 300) de conformidad con la reivindicación 13, en el que el dispositivo (100; 150; 300) comprende adicionalmente una memoria o una base de datos, dependiendo uno o varios de los coeficientes de filtrado de la diferencia de ángulo de fase, estando la unidad de filtro (330) configurada para leer los coeficientes de filtrado de la memoria o la base de datos en función en la diferencia de ángulo de fase, y estando la unidad de filtro (330) configurada para filtrar los segundos valores de subbanda aplicando el filtro el cual presenta los coeficientes de filtrado leídos de la memoria o la base de datos.

15. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones 12 a 14, en el que la unidad de filtro (330) está configurada para filtrar cada uno del uno o varios segundos valores de

- subbanda de cada una de las segundas subbandas a fin de obtener uno o varios valores de subbanda filtrados para cada una de las segundas subbandas, estando la unidad de filtro (330) configurada para determinar cada uno de los valores de subbanda filtrados al calcular una suma de un primer sumando o uno o varios sumandos adicionales,
- 5 siendo el primer sumando uno de los segundos valores de subbanda de una de las segundas subbandas, siendo el uno o varios sumandos adicionales uno o varios valores de subbanda ponderados, y estando la unidad de filtro (330) configurada para determinar el uno o varios valores de sumandos ponderados en cada caso mediante multiplicación de uno de los coeficientes de filtrado por en cada caso un valor de subbanda adicional de otra de las subbandas adicionales.
- 10
16. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones 12 a 15, en el que la unidad de filtro (330) está configurada para filtrar los segundos valores de subbanda aplicando el filtro el cual comprende la pluralidad de coeficientes de filtrado de modo que se reducen las componentes espurias de los segundos valores de subbanda,
- 15 presentando los coeficientes de filtrado del filtro una secuencia ordenada, y, en esta secuencia, a cada coeficiente de filtrado distinto de cero le sigue un coeficiente de filtrado igual a cero.
17. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones precedentes, en el que la interfaz (110; 310) está configurada para recibir la información de frecuencia que indica por qué diferencia de frecuencia se van a desplazar los primeros valores de subbanda de las subbandas de la señal de entrada de audio.
- 20
18. Un dispositivo (100; 150; 300) de conformidad con una de las reivindicaciones precedentes, en el que el dispositivo (100; 150; 300) comprende adicionalmente una unidad de transformación de síntesis (340), y estando la unidad de transformación de síntesis configurada para obtener la señal de audio desplazada en frecuencia al transformar los segundos valores de subbanda a en una la pluralidad de muestras de dominio de tiempo.
- 25
19. Un dispositivo (150) de conformidad con una de las reivindicaciones precedentes, en el que el dispositivo (150) está configurado para producir una señal de audio de frecuencia ampliada, estando el dispositivo (150) configurado para producir la señal de audio de frecuencia ampliada al producir el dispositivo (150) los segundos valores de subbanda de la señal de audio desplazada en frecuencia, comprendiendo la señal de audio de frecuencia ampliada los primeros valores de subbanda de la señal de entrada de audio y los segundos valores de subbanda de la señal de audio desplazada en frecuencia.
- 30
20. Un dispositivo (150) de conformidad con la reivindicación 19, en el que el dispositivo (150) comprende adicionalmente una unidad de transformación de síntesis (340), y estando la unidad de transformación de síntesis configurada para obtener la señal de audio de frecuencia ampliada al transformar los primeros valores de subbanda y los segundos valores de subbanda en una pluralidad de muestras de dominio de tiempo.
- 35
21. Un dispositivo (150) de conformidad con una de las reivindicaciones precedentes 19 o 20, en el que el dispositivo (150) comprende adicionalmente una unidad de adaptación de curva envolvente (350) configurada para filtrar la señal de audio de frecuencia ampliada de modo que se le imprime una envolvente espectral a la señal de frecuencia ampliada.
- 40
22. Un dispositivo (150) de conformidad con la reivindicación 21, en el que la unidad de adaptación de curva envolvente (350) está configurada para filtrar la señal de audio de frecuencia ampliada por medio de filtración inversa con coeficientes LPC a fin de imprimirle la envolvente espectral a la señal de frecuencia ampliada.
- 45
23. Un método para producir una señal de audio desplazada en frecuencia basándose en una señal de entrada de audio, estando representada la señal de entrada de audio, para una pluralidad de primeras subbandas, por uno o varios primeros valores de subbanda, comprendiendo el método:
- 50
- recibir la señal de entrada de audio, y
- 55 producir la señal de audio desplazada en frecuencia, presentando la señal de audio desplazada en frecuencia, para una pluralidad de segundas subbandas, en cada caso uno o varios segundos valores de subbanda, presentando cada uno de los primeros y los segundos valores de subbanda información de un ángulo de fase respectivo, y
- 60 produciéndose uno de los segundos valores de subbanda basándose en uno de los primeros valores de subbanda de tal manera que el segundo ángulo de fase de este segundo valor de subbanda difiere del primer ángulo de fase de este primer valor de subbanda por una diferencia de ángulo de fase, dependiendo la diferencia de ángulo de fase de una información de frecuencia que indica por qué diferencia de frecuencia se va a desplazar la señal de entrada de audio a fin de obtener la señal de audio desplazada en frecuencia, **caracterizado por que** la diferencia de ángulo de fase depende de un ancho de banda de frecuencia

de una de las primeras subbandas.

24. Programa informático configurado para realizar el método de conformidad con la reivindicación 23 cuando el programa informático es ejecutado en un ordenador o procesador de señales.

5

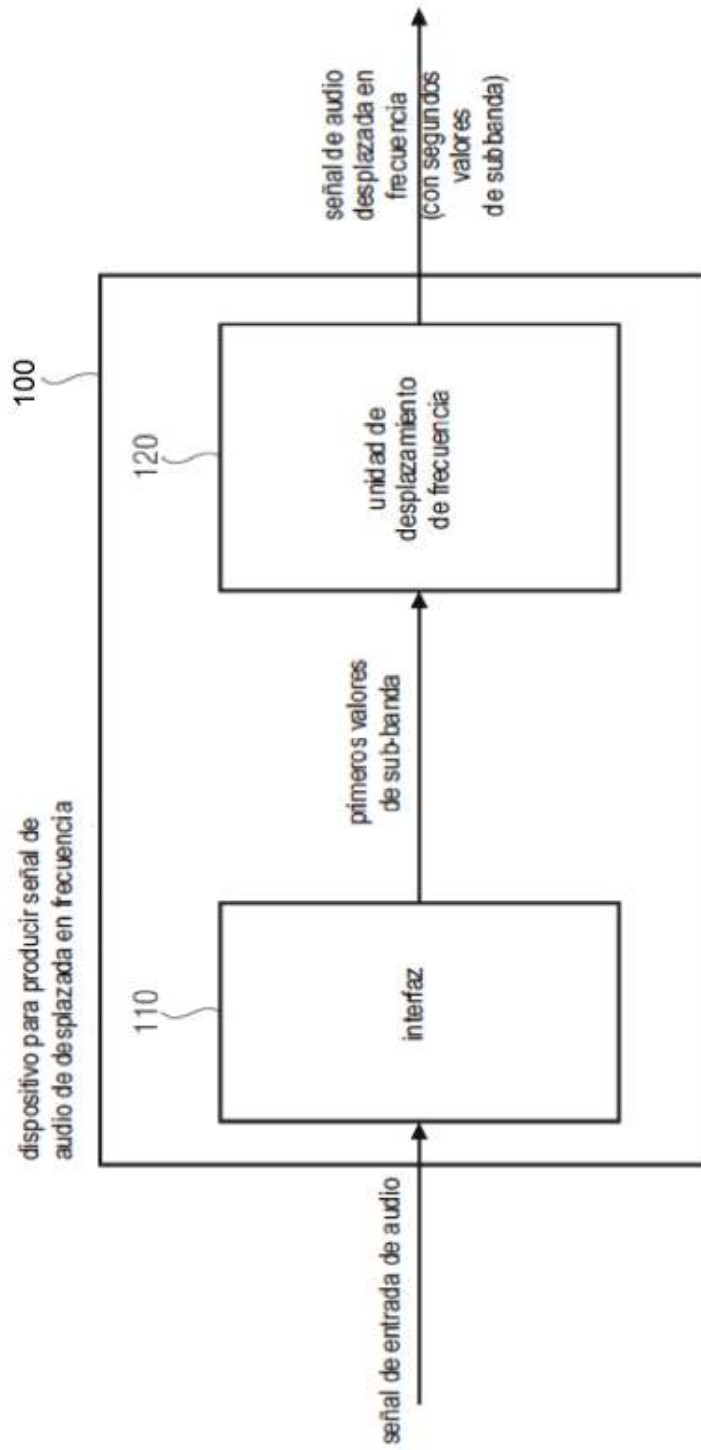


FIGURA 1A

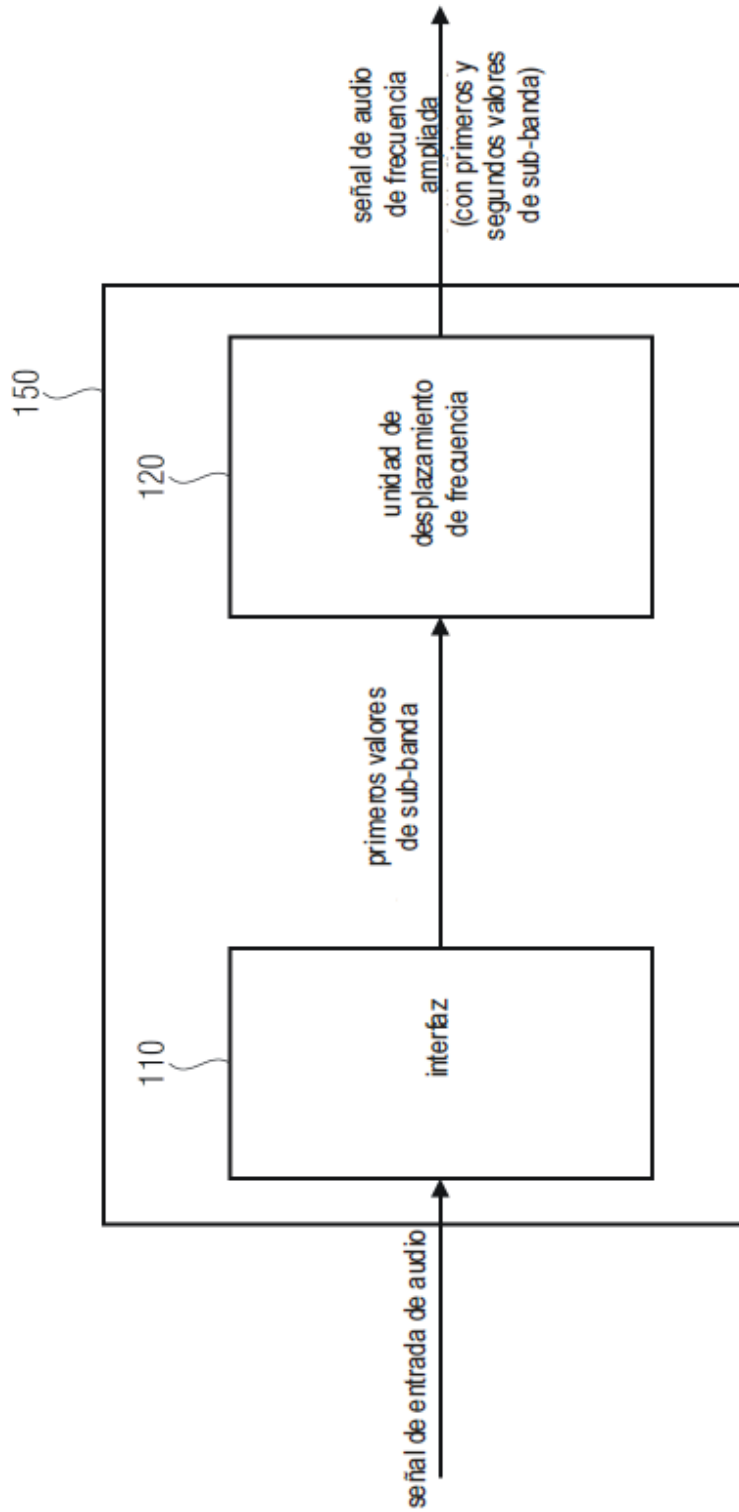


FIGURA 1B

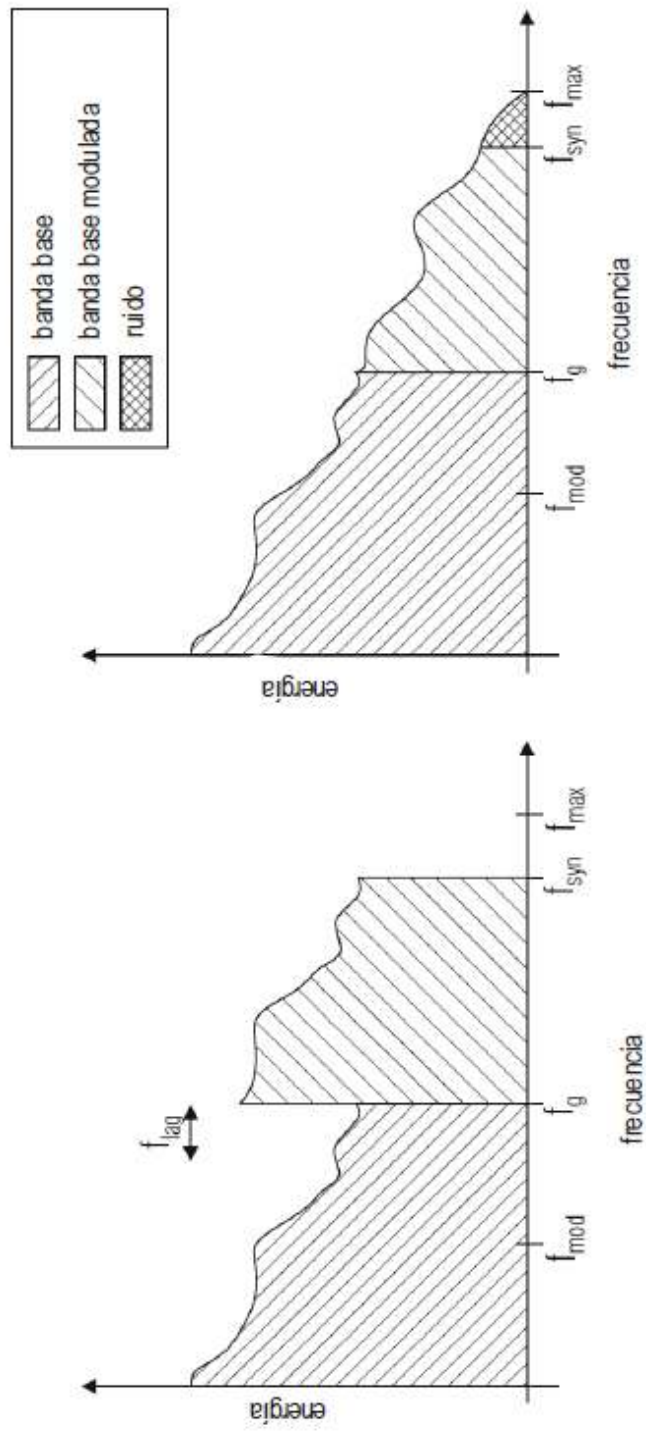


FIGURA 2

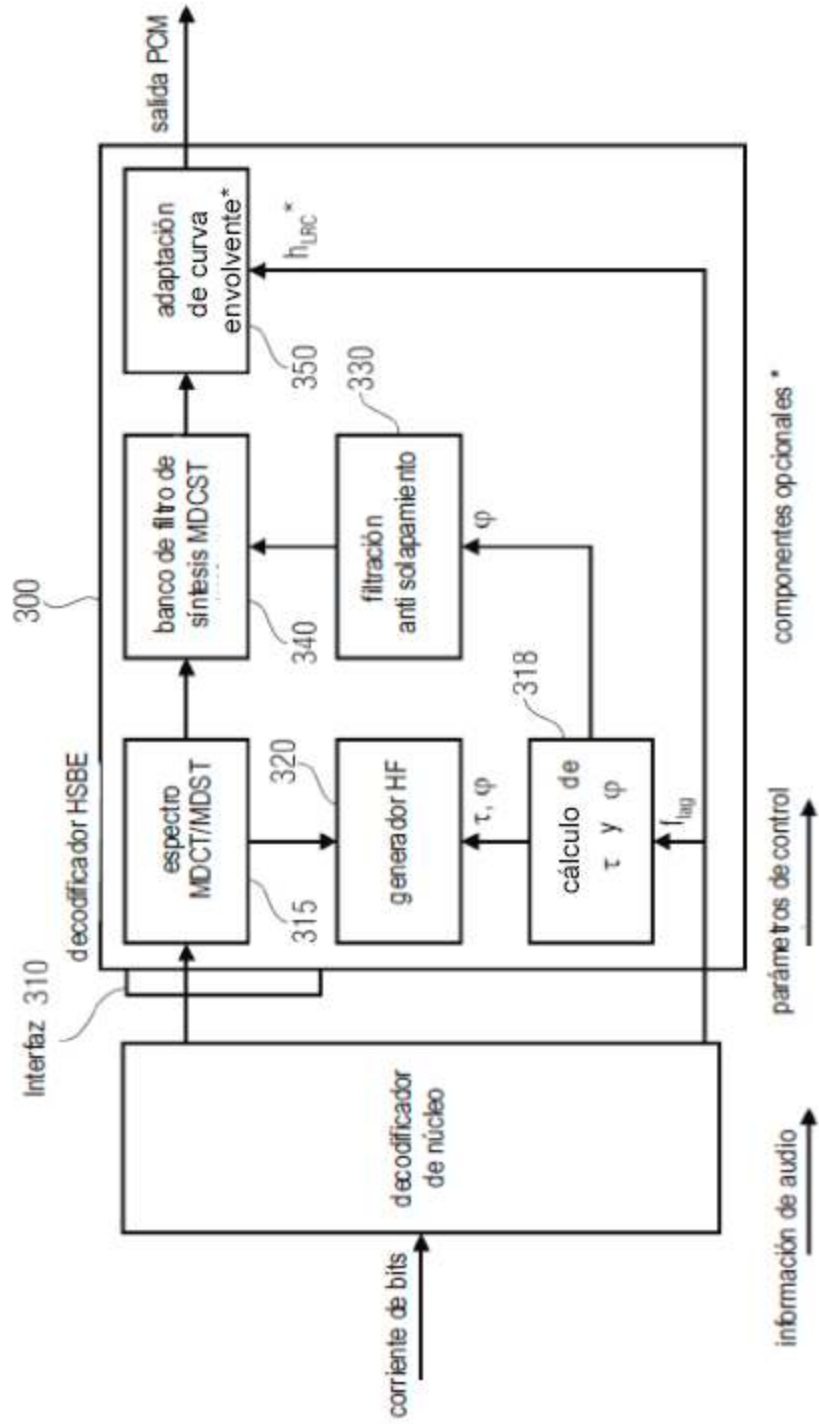
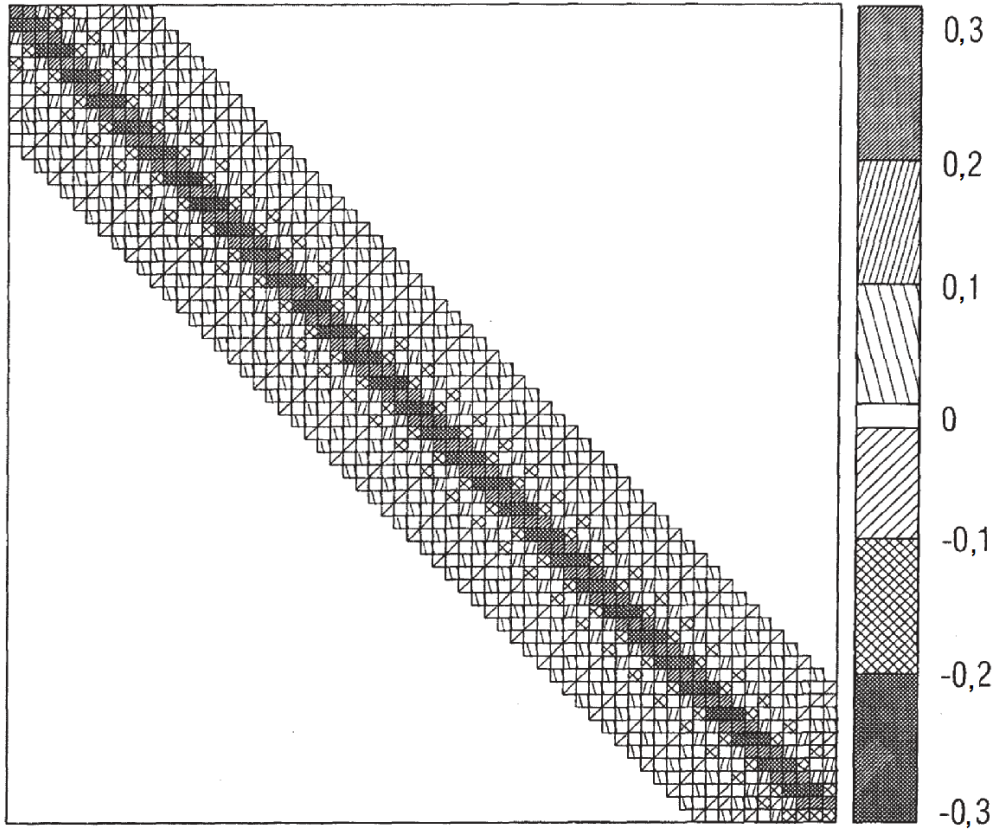


FIGURA 3

H_0



Matriz de transformación calculada de manera exacta H_0

FIGURA 4A

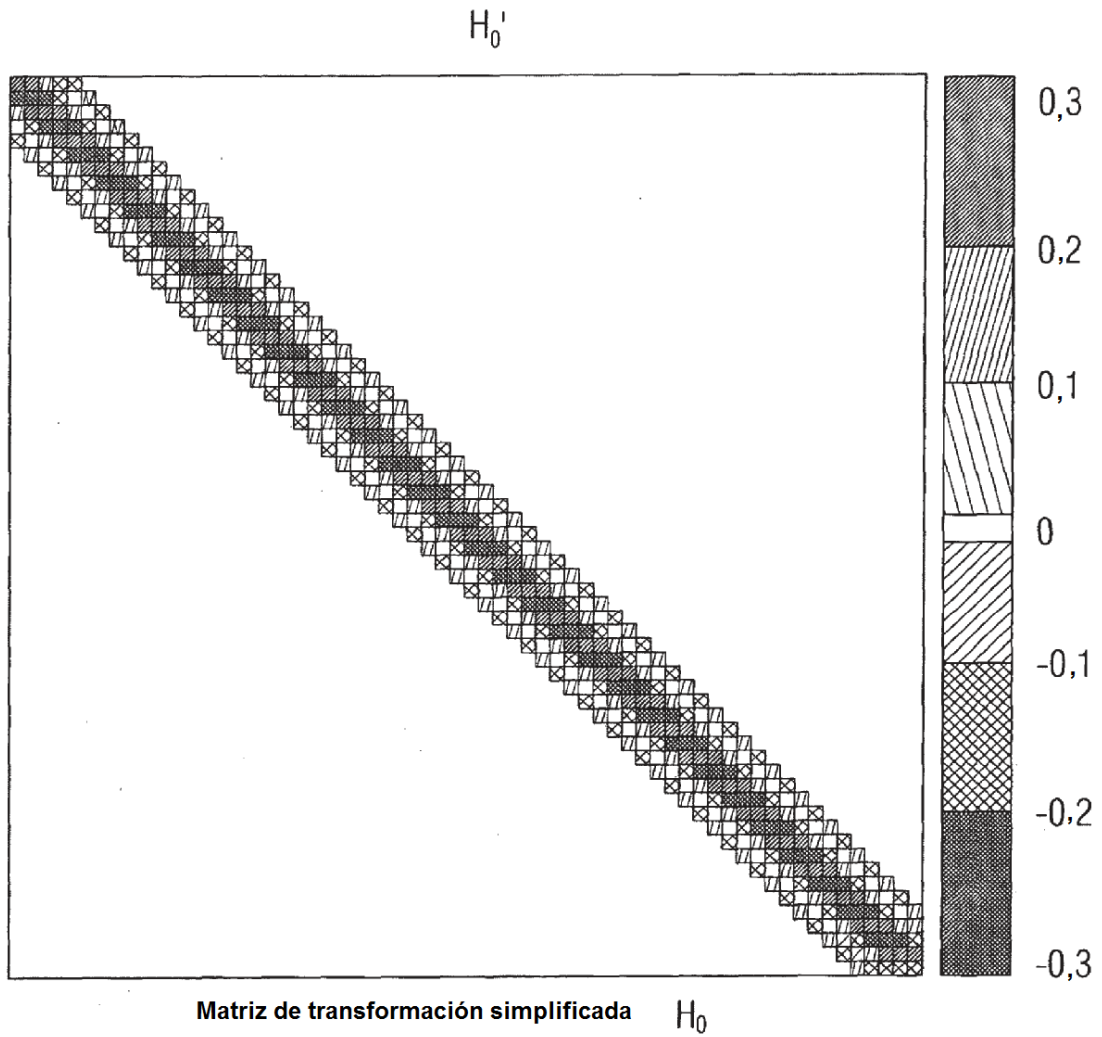


FIGURA 4B

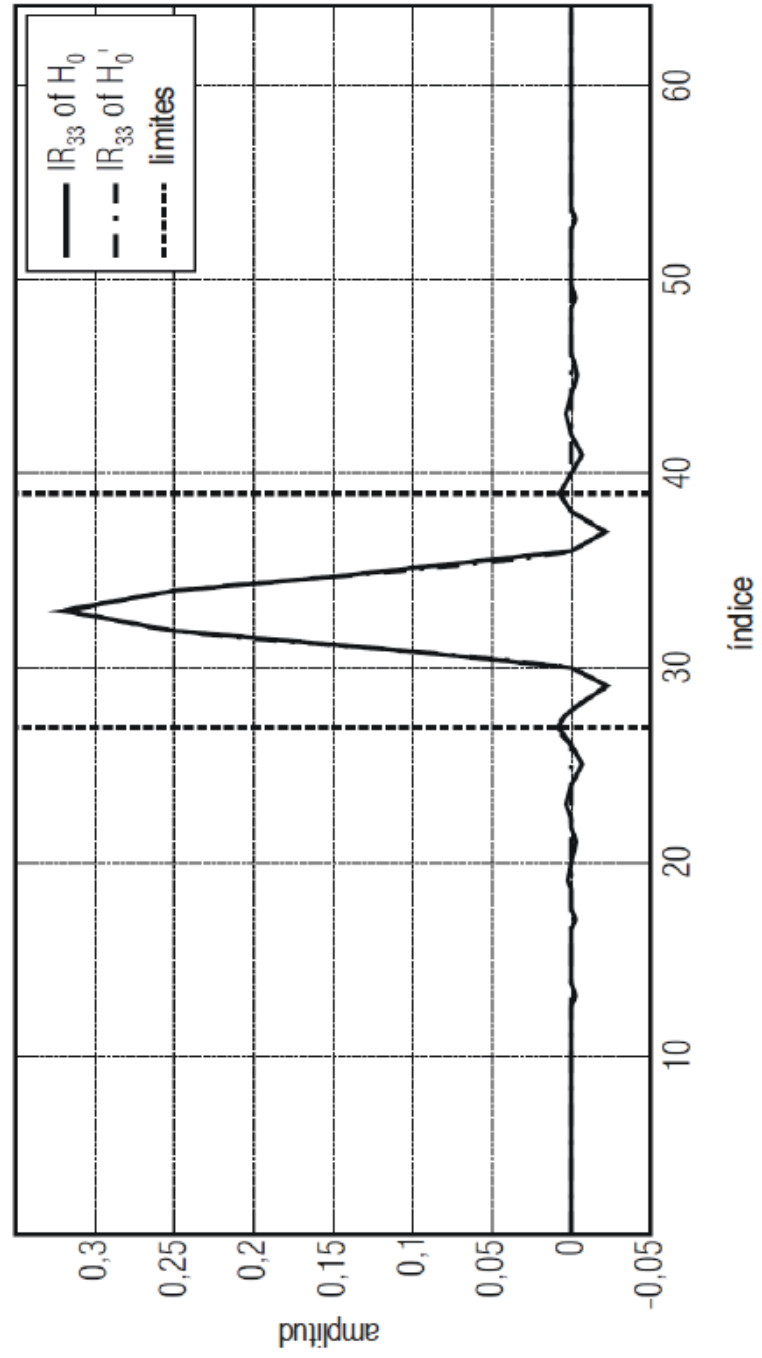


FIGURA 5

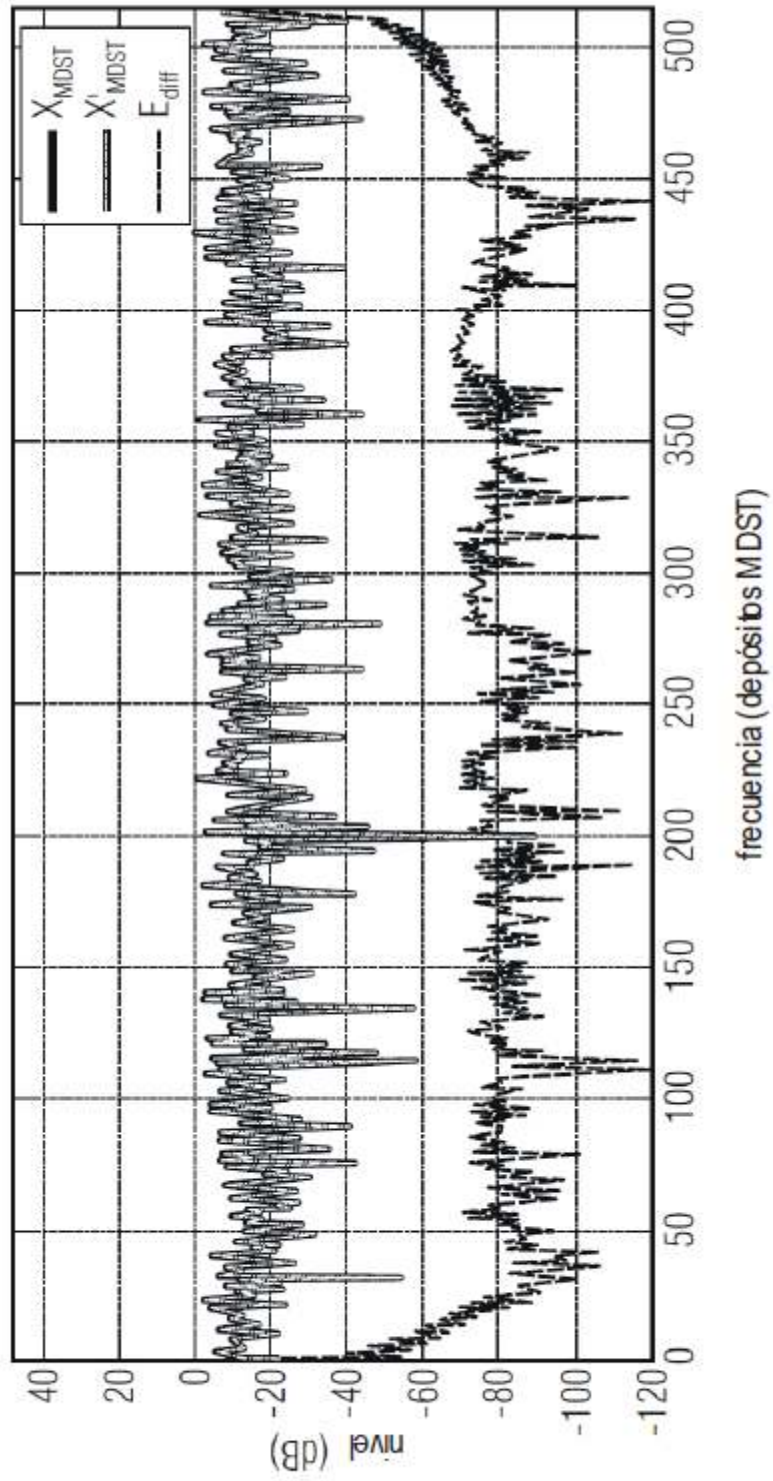


FIGURA 6

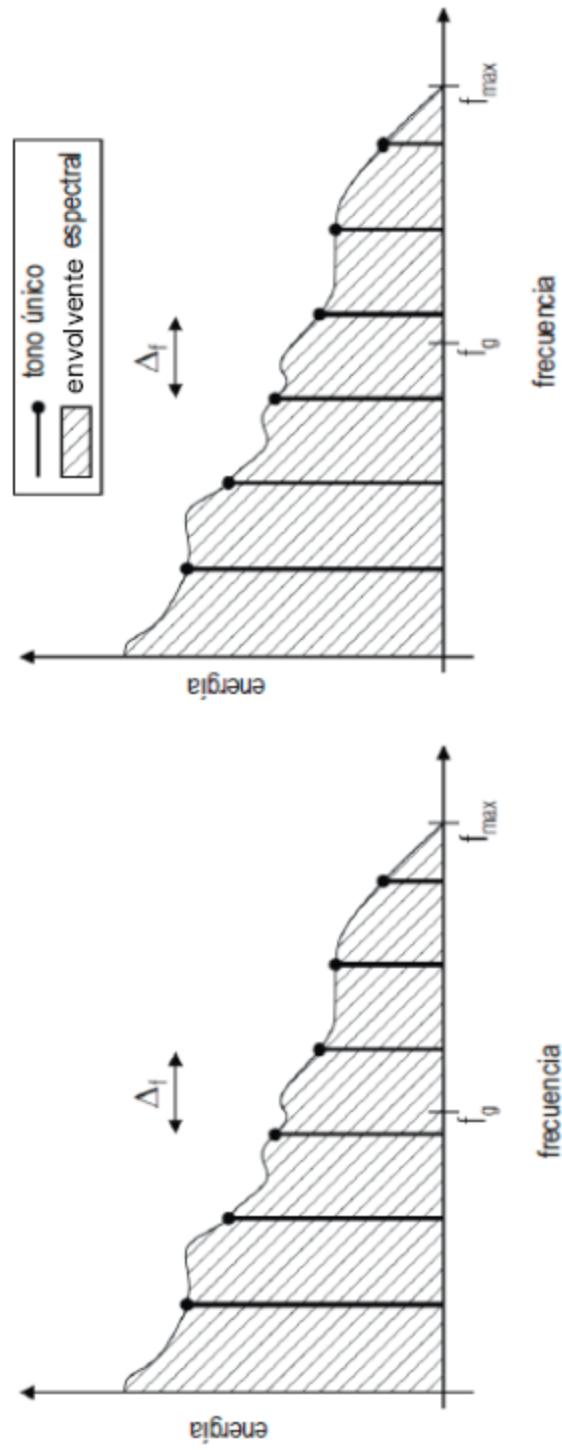


FIGURA 7

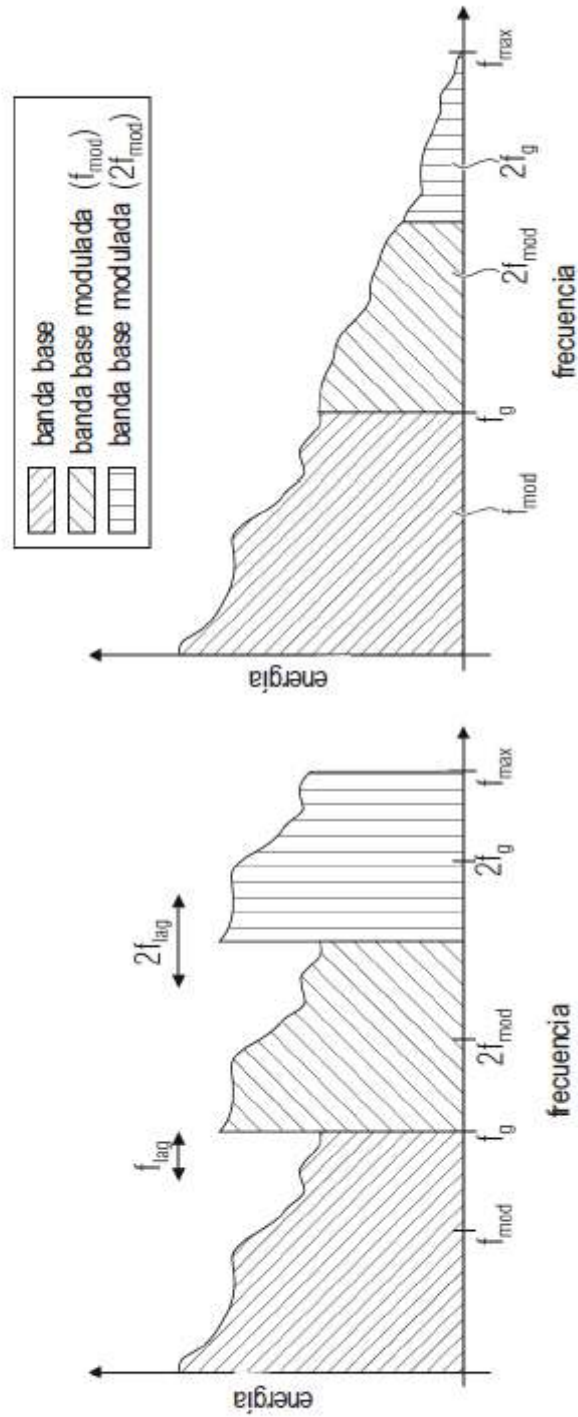


FIGURA 8

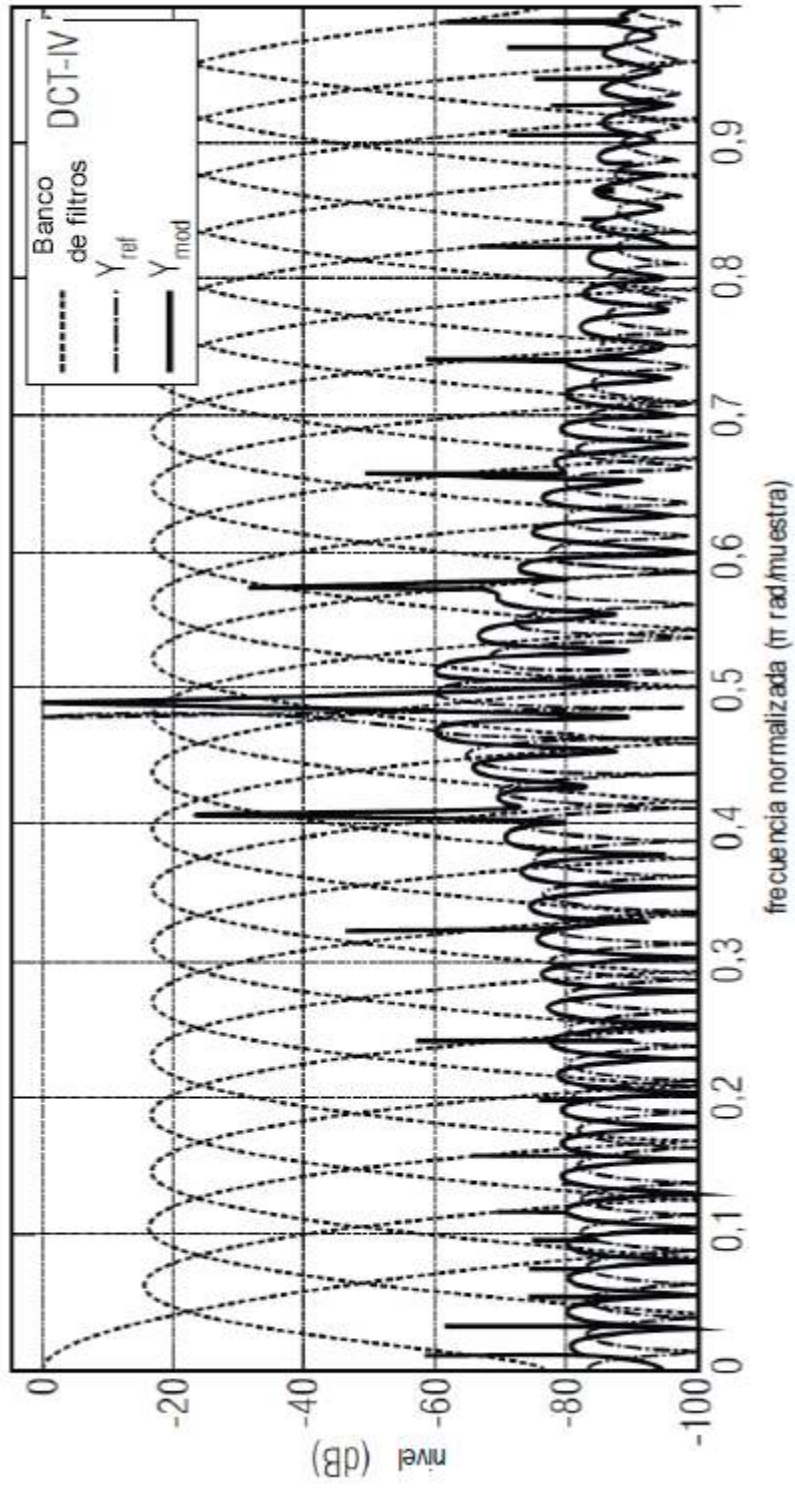


FIGURA 9

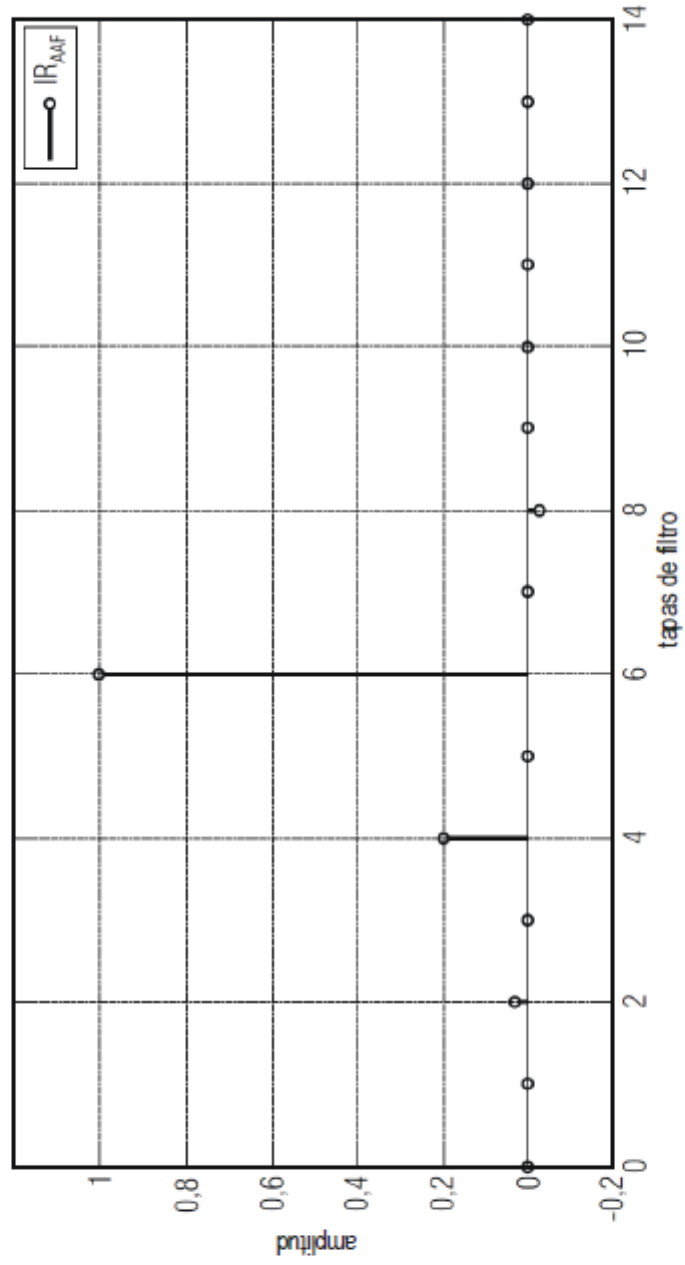


FIGURA 10

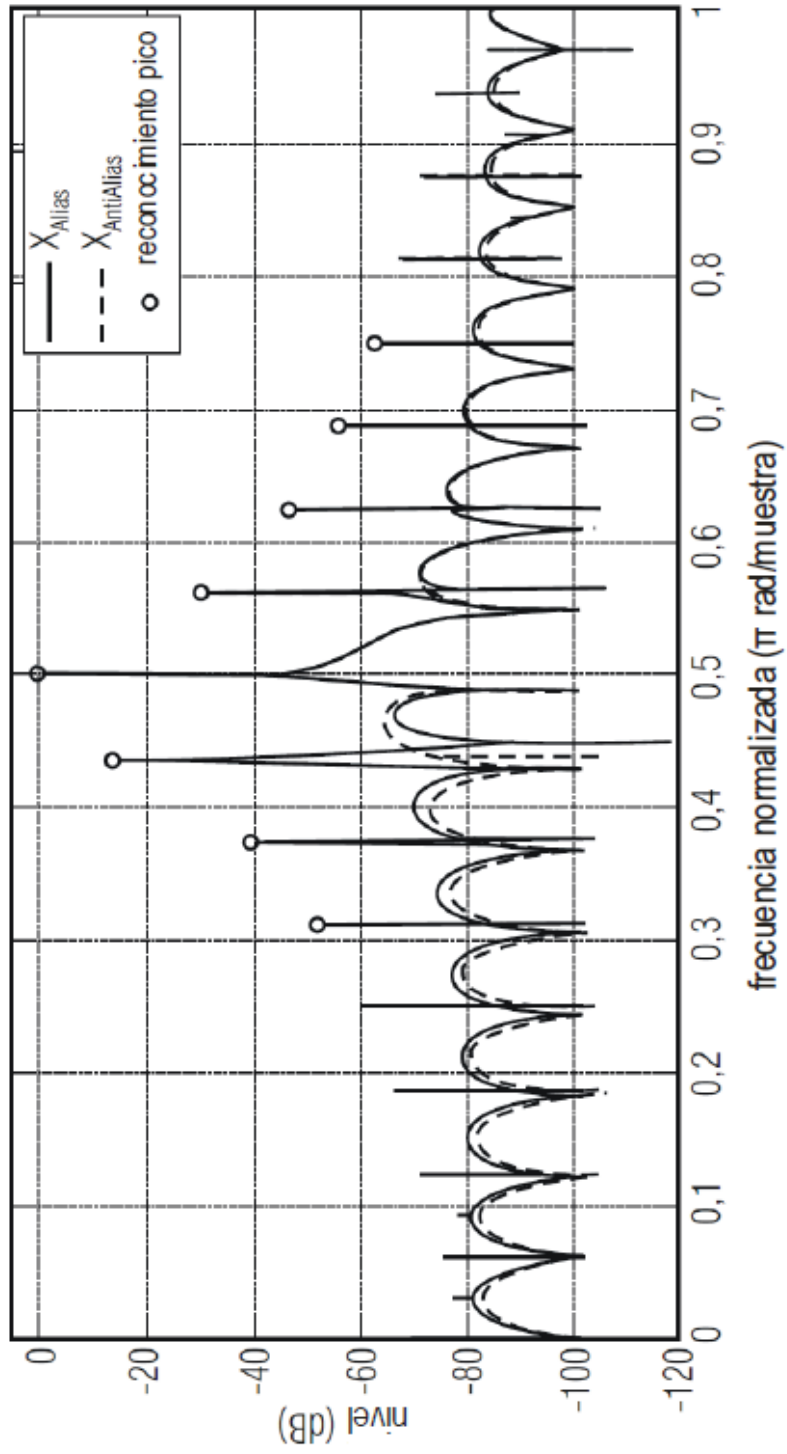


FIGURA 11

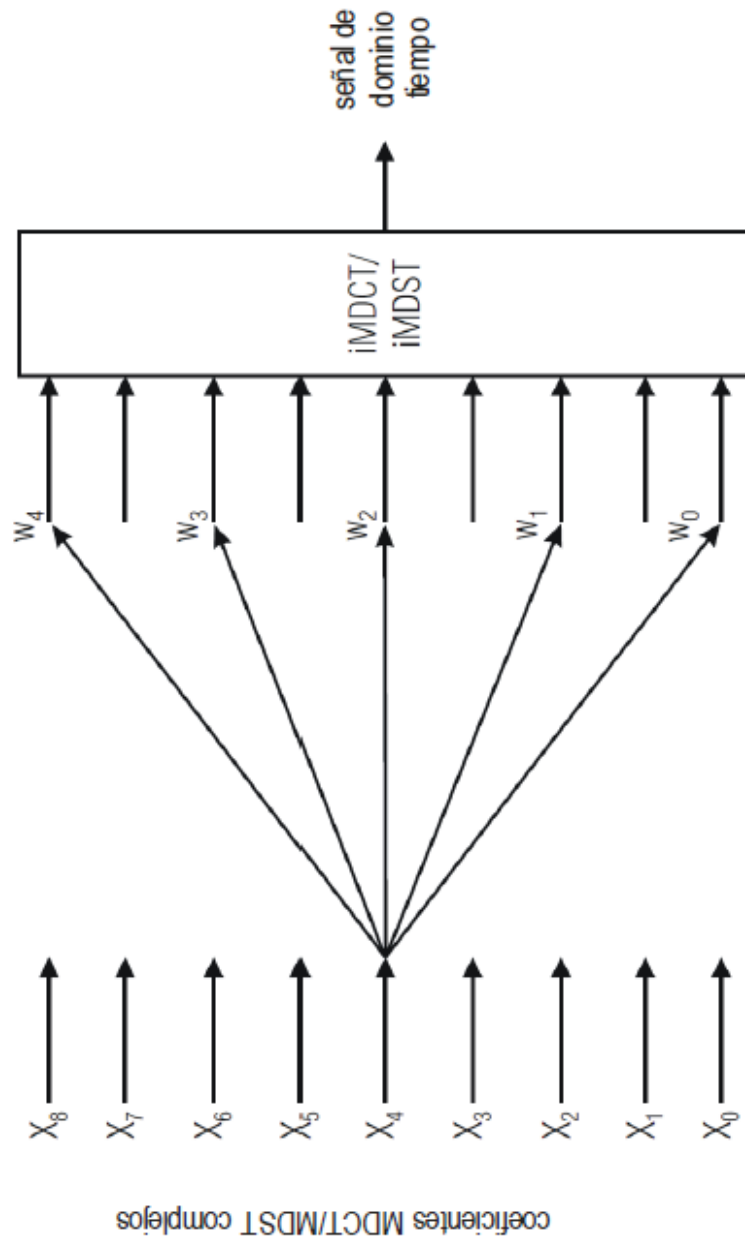


FIGURA 12

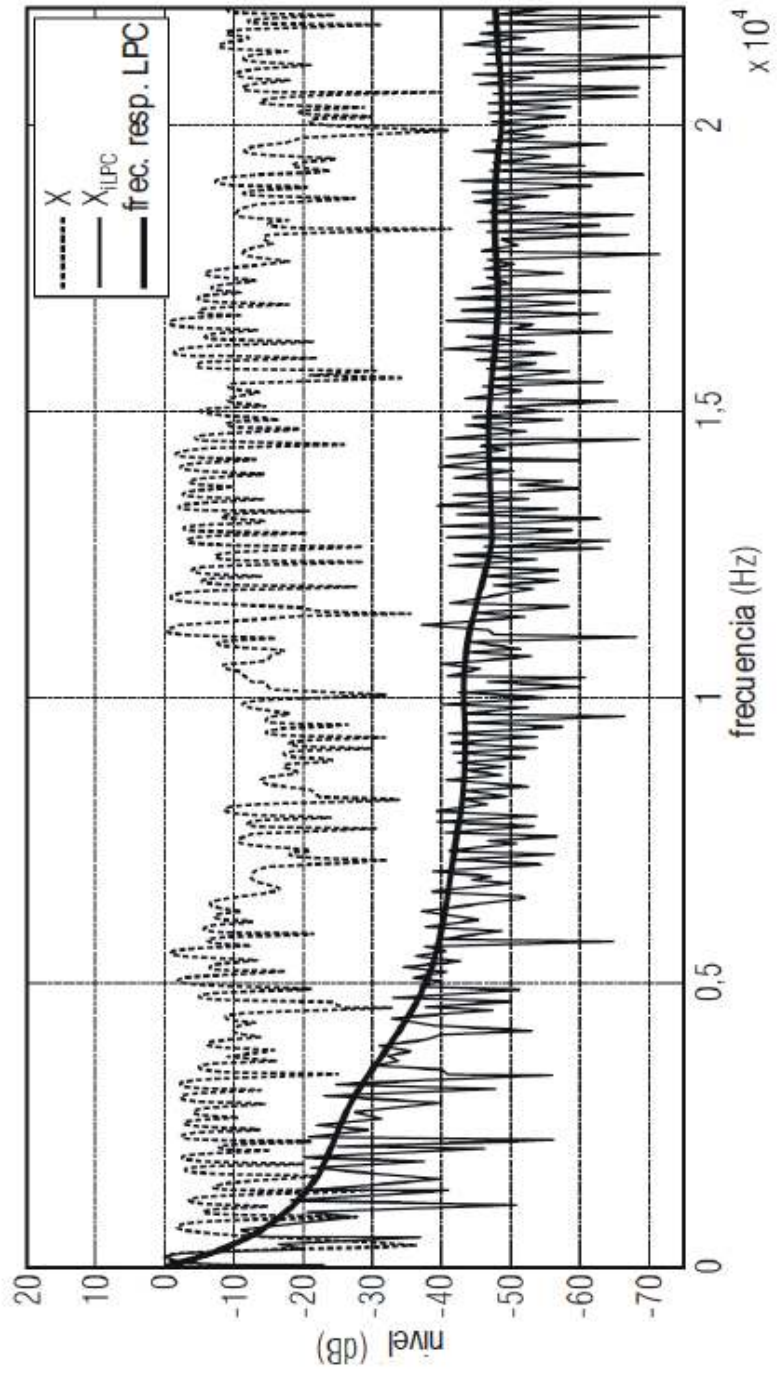


FIGURA 13

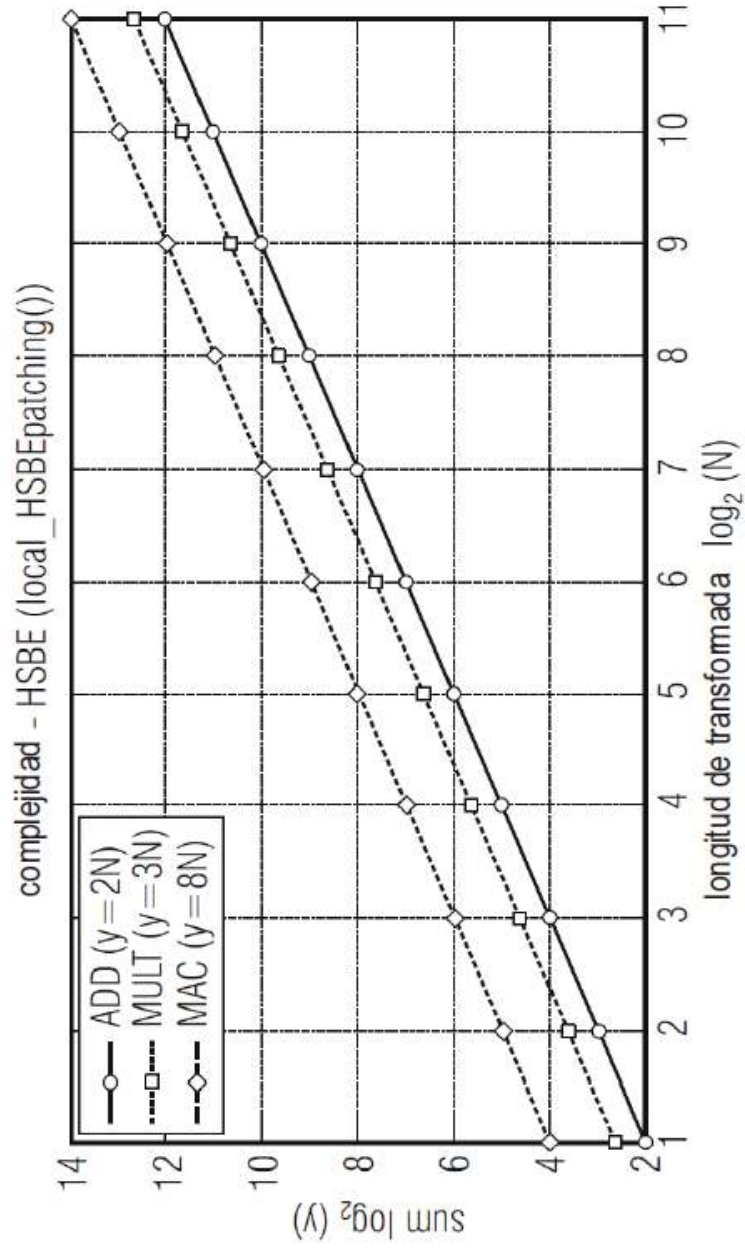


FIGURA 14

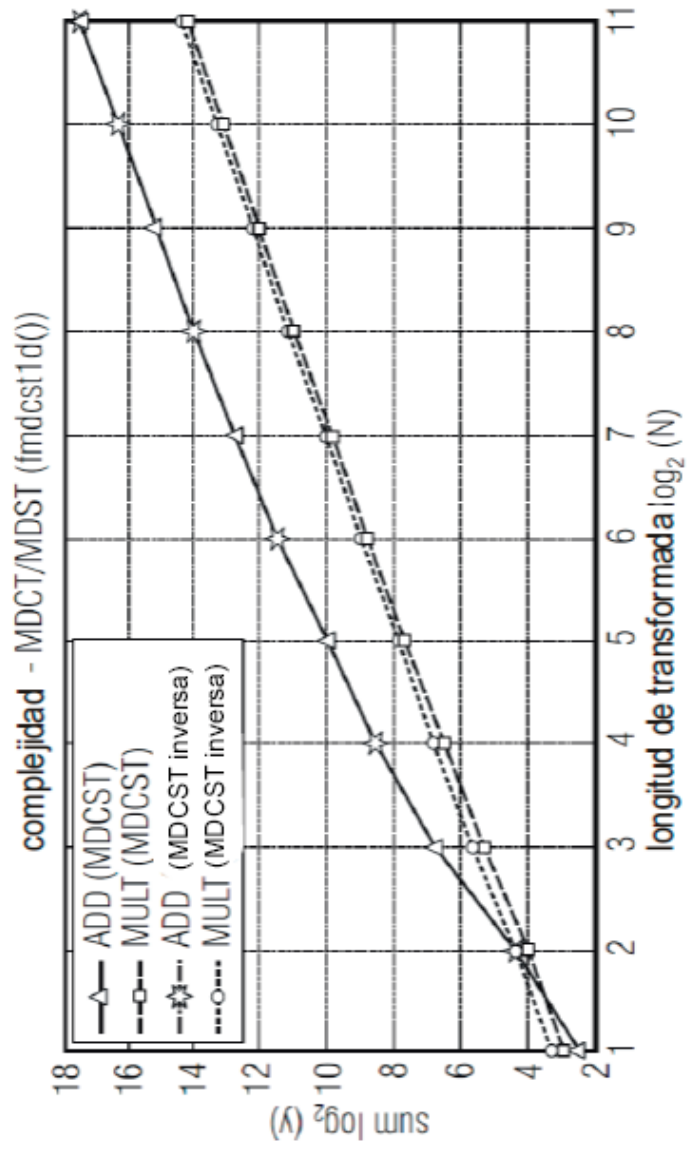


FIGURA 15

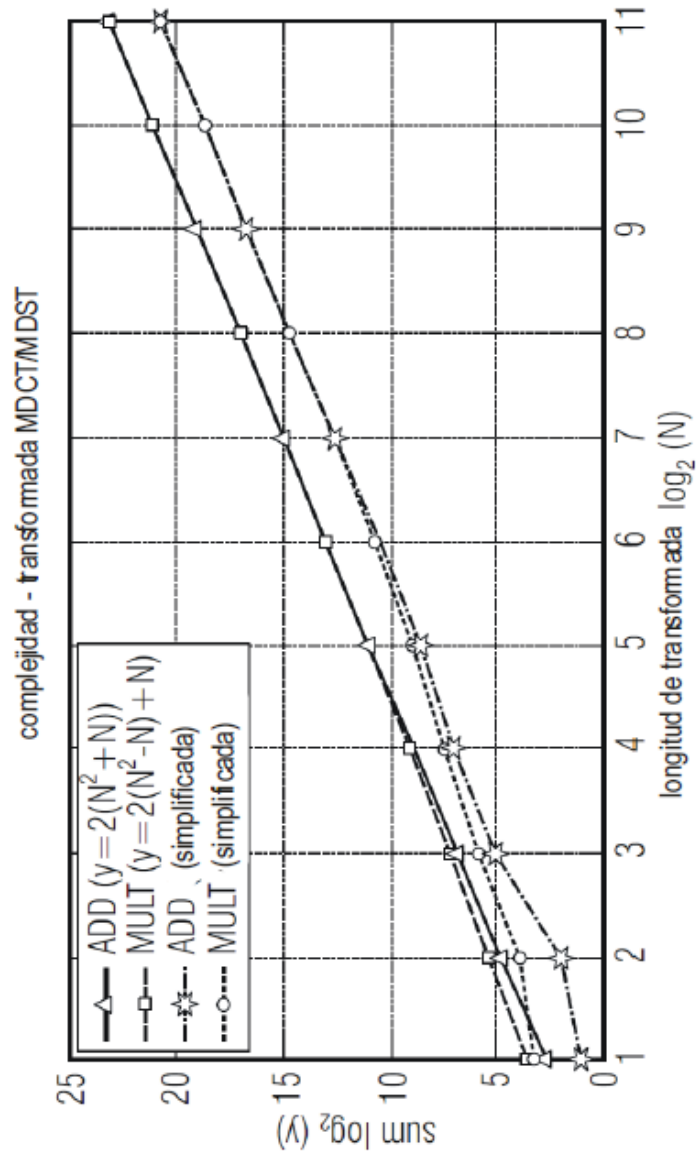


FIGURA 16

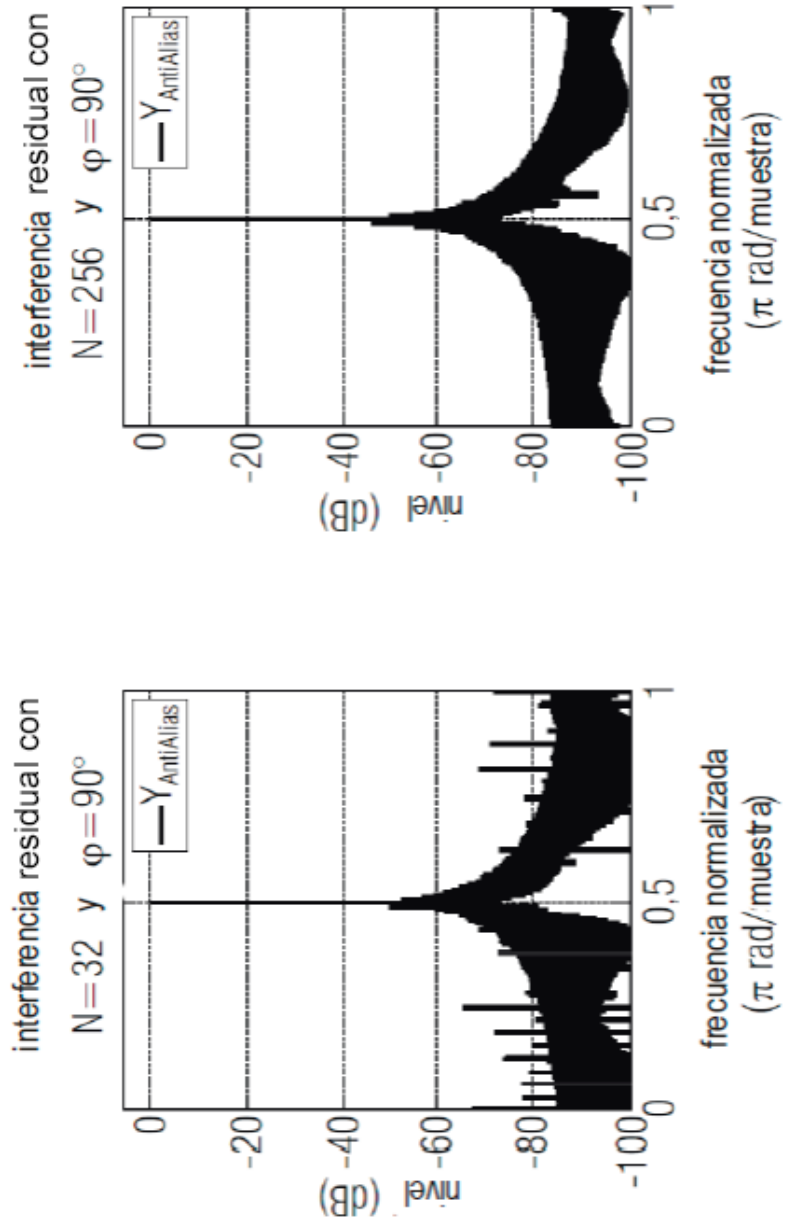


FIGURA 17

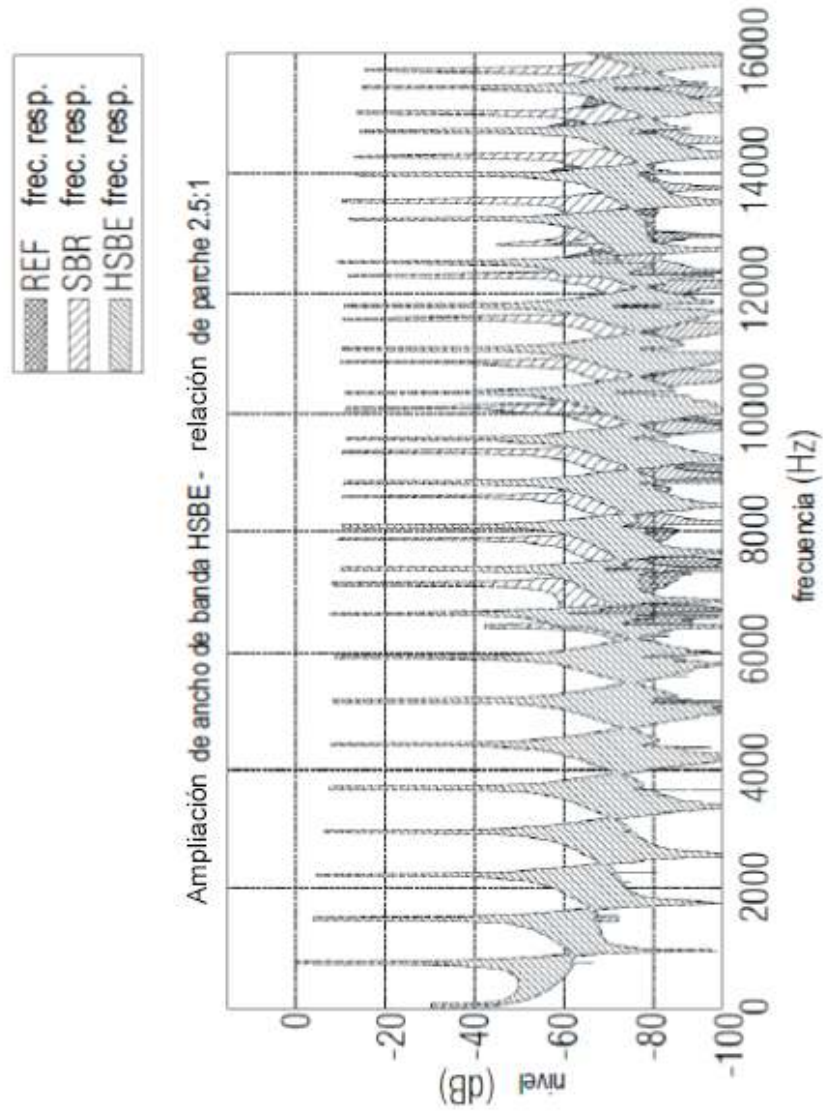


FIGURA 18

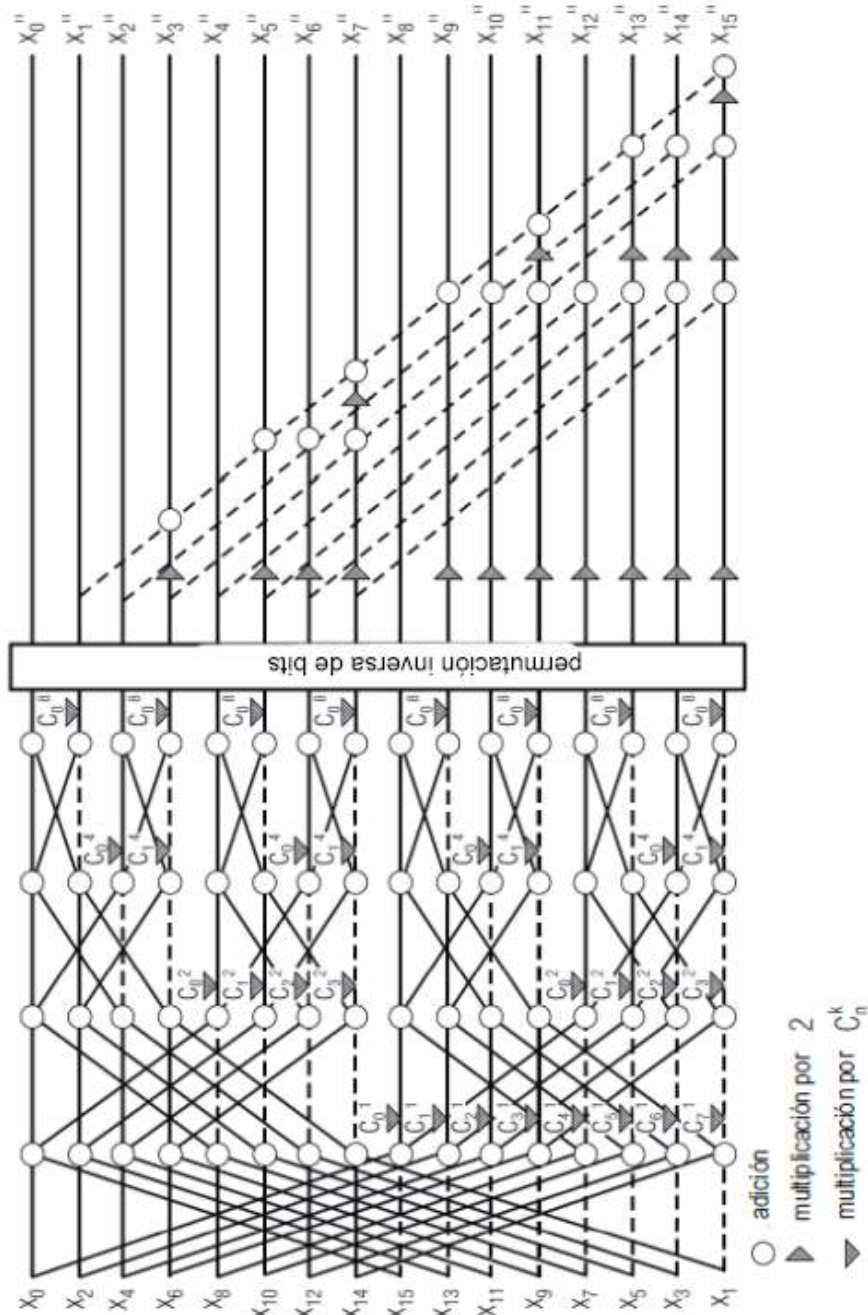


FIGURA 19

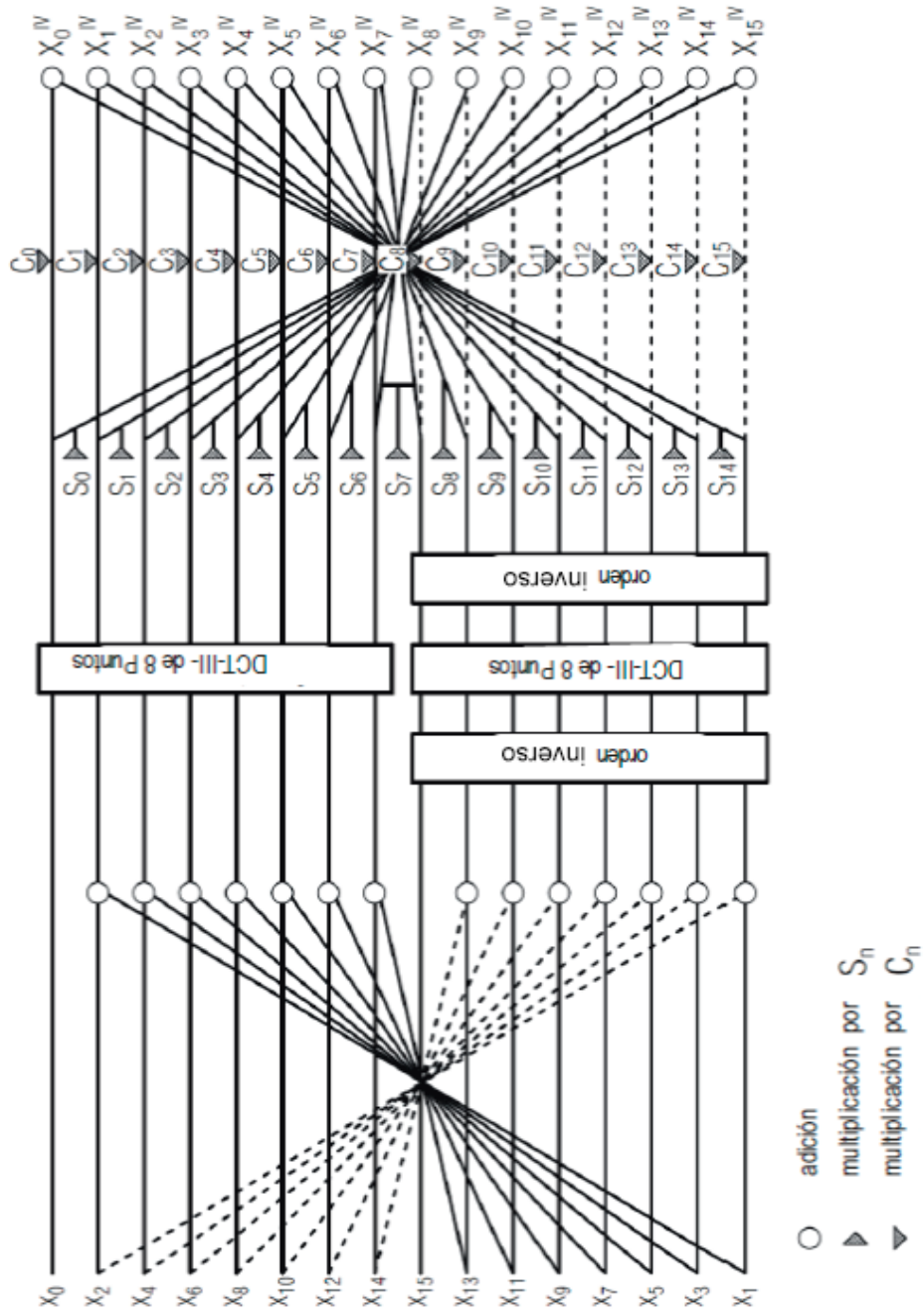


FIGURA 20

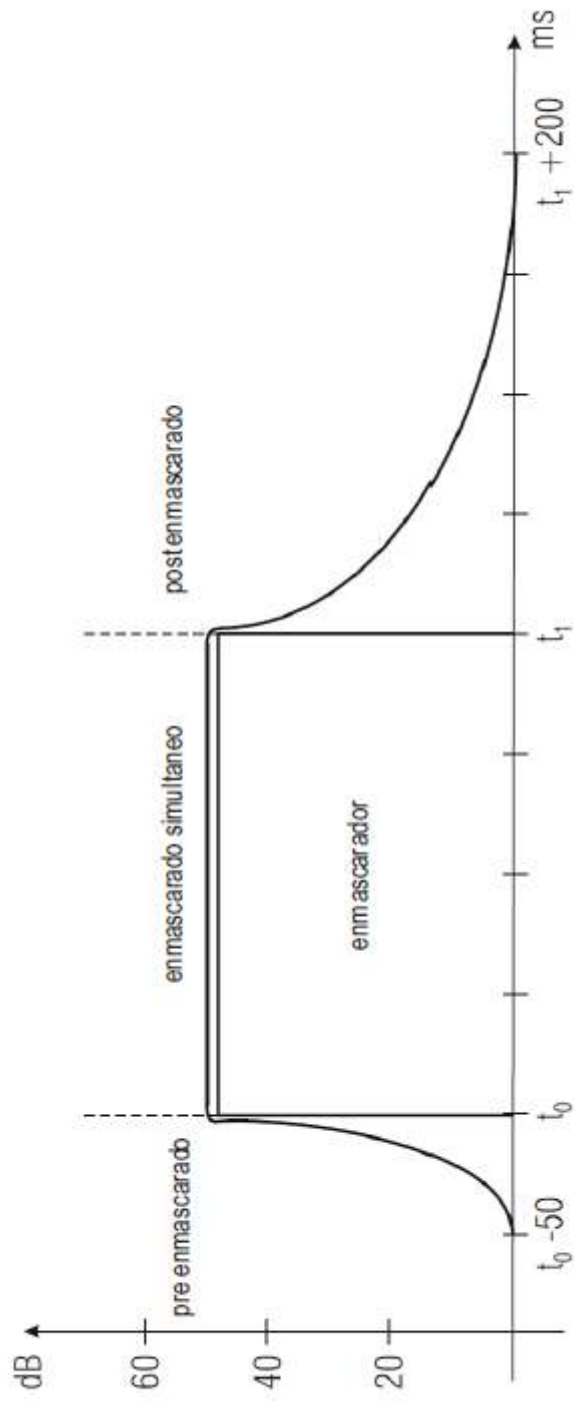


FIGURA 21

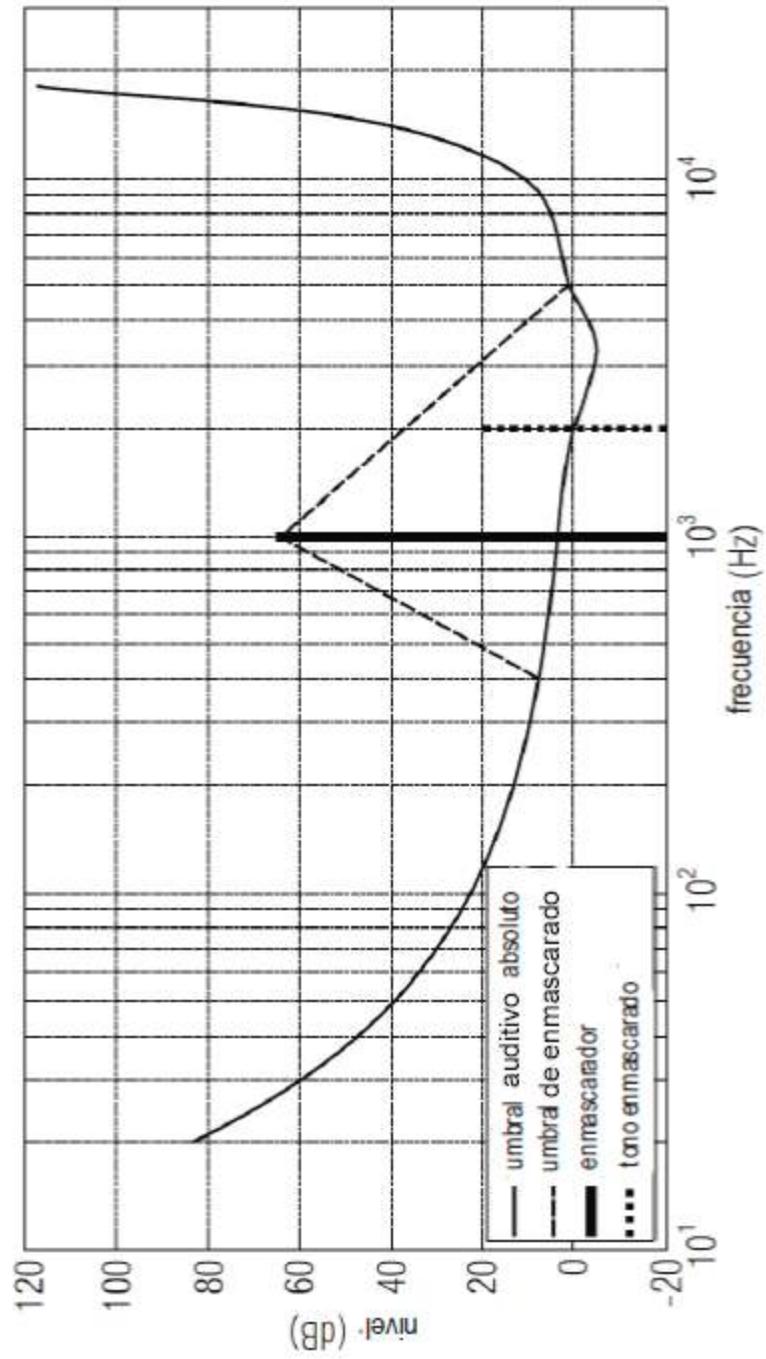


FIGURA 22

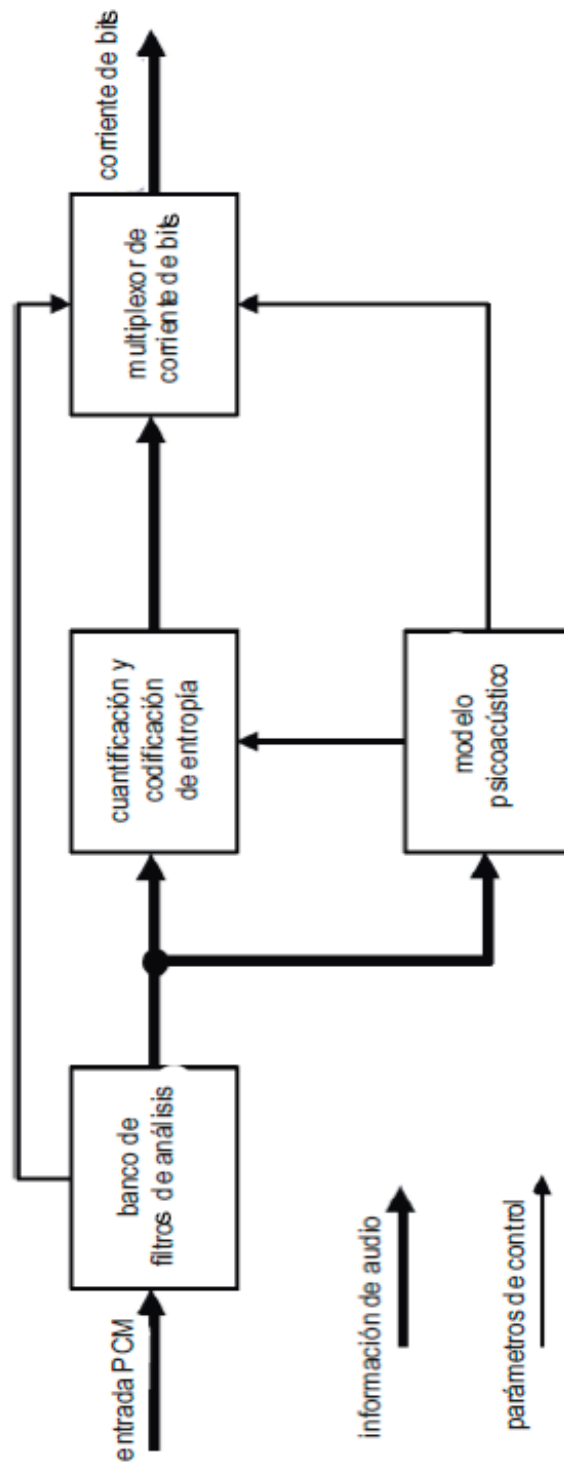


FIGURA 23

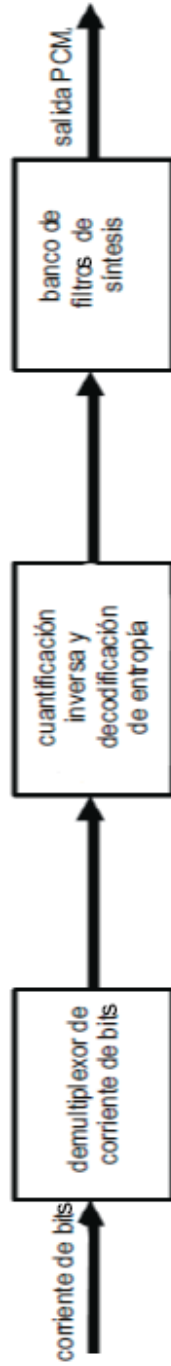


FIGURA 24

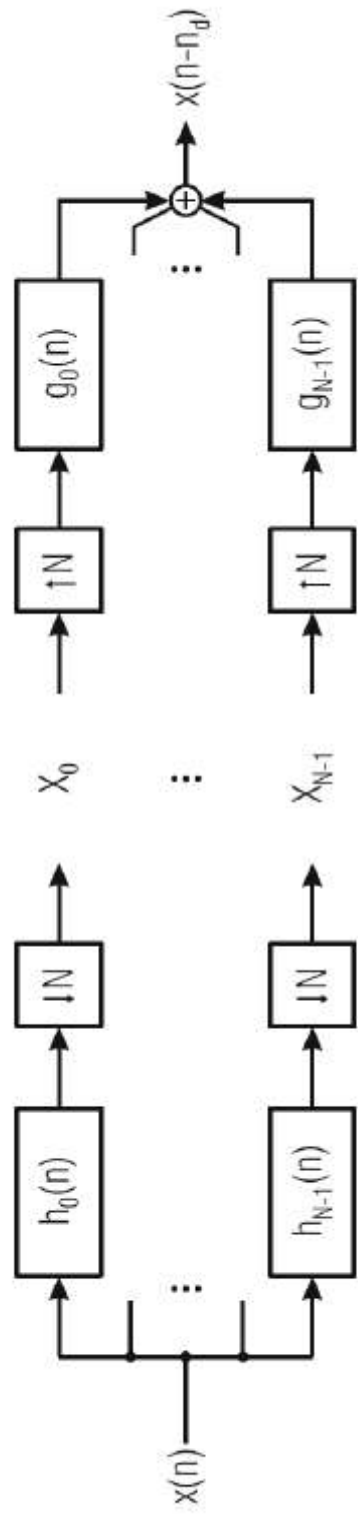


FIGURA 25

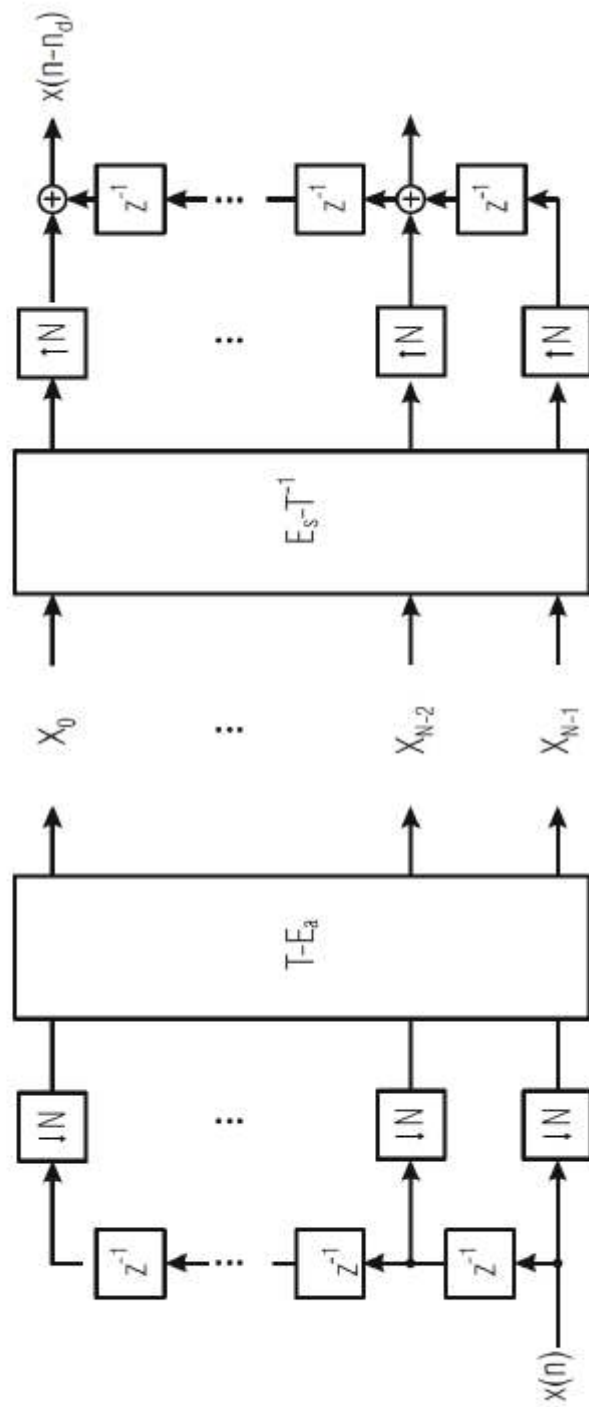


FIGURA 26

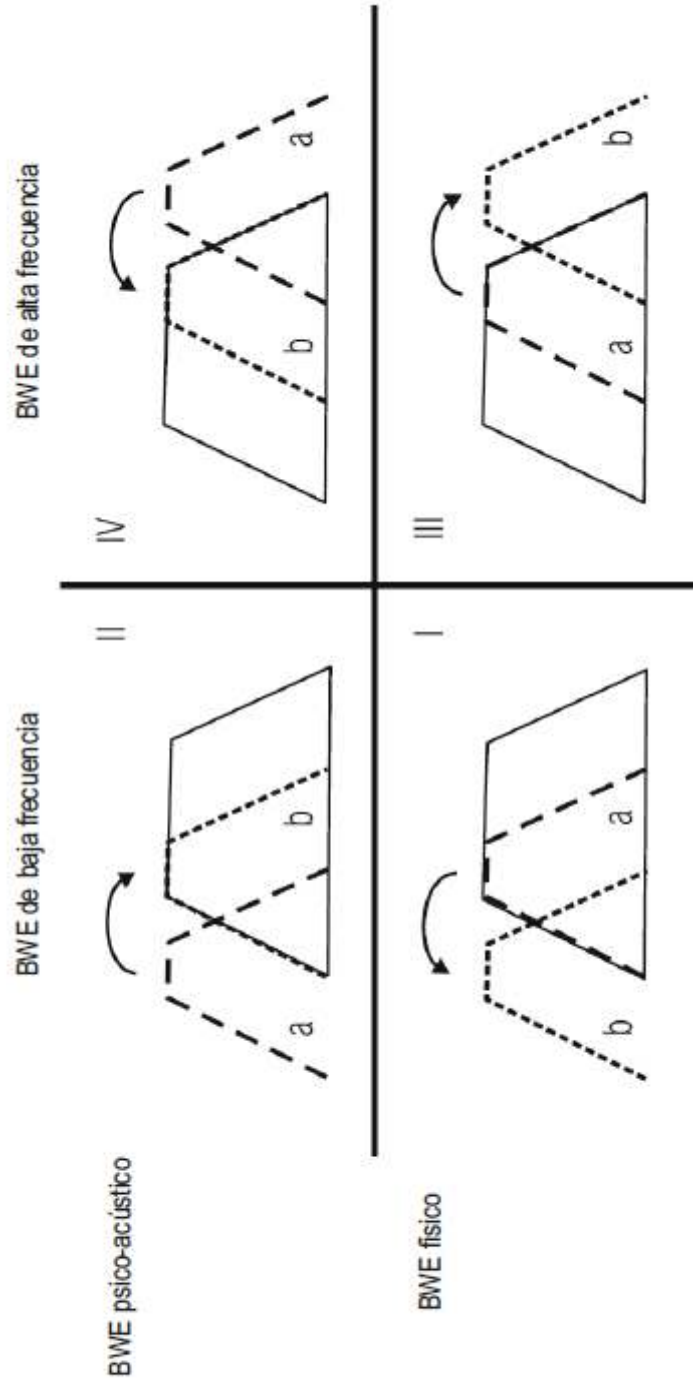


FIGURA 27

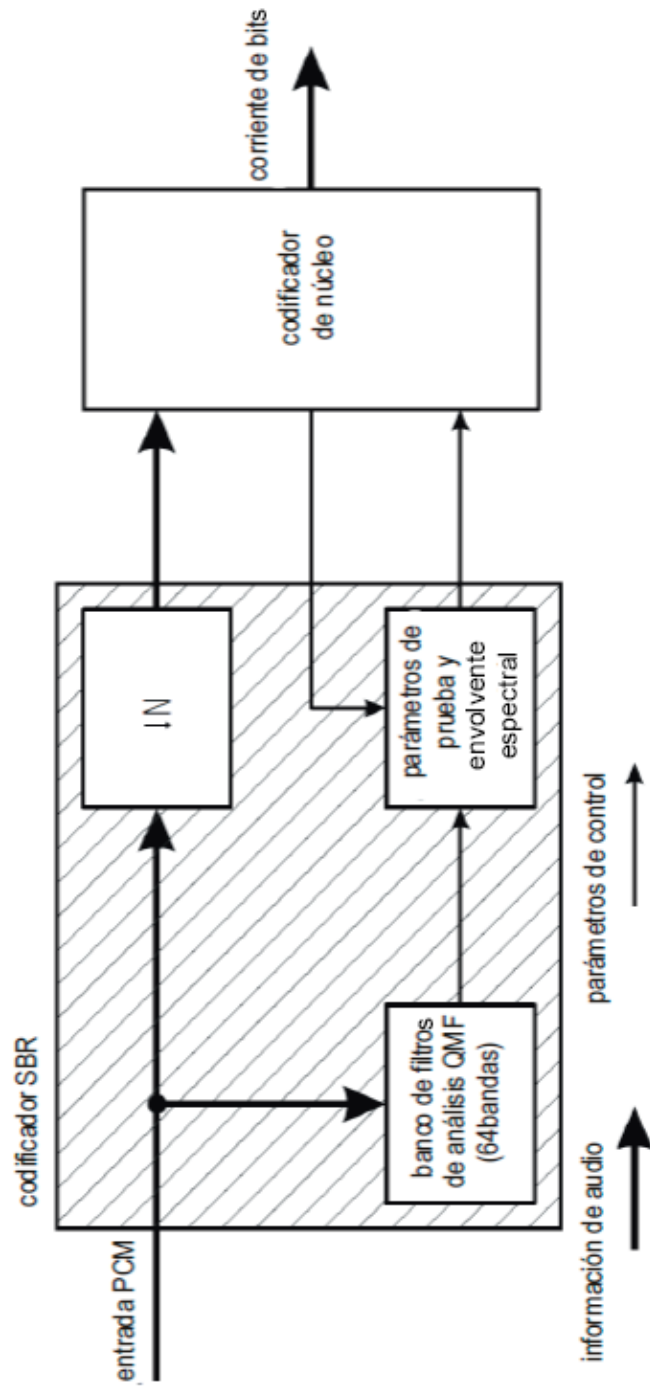


FIGURA 28

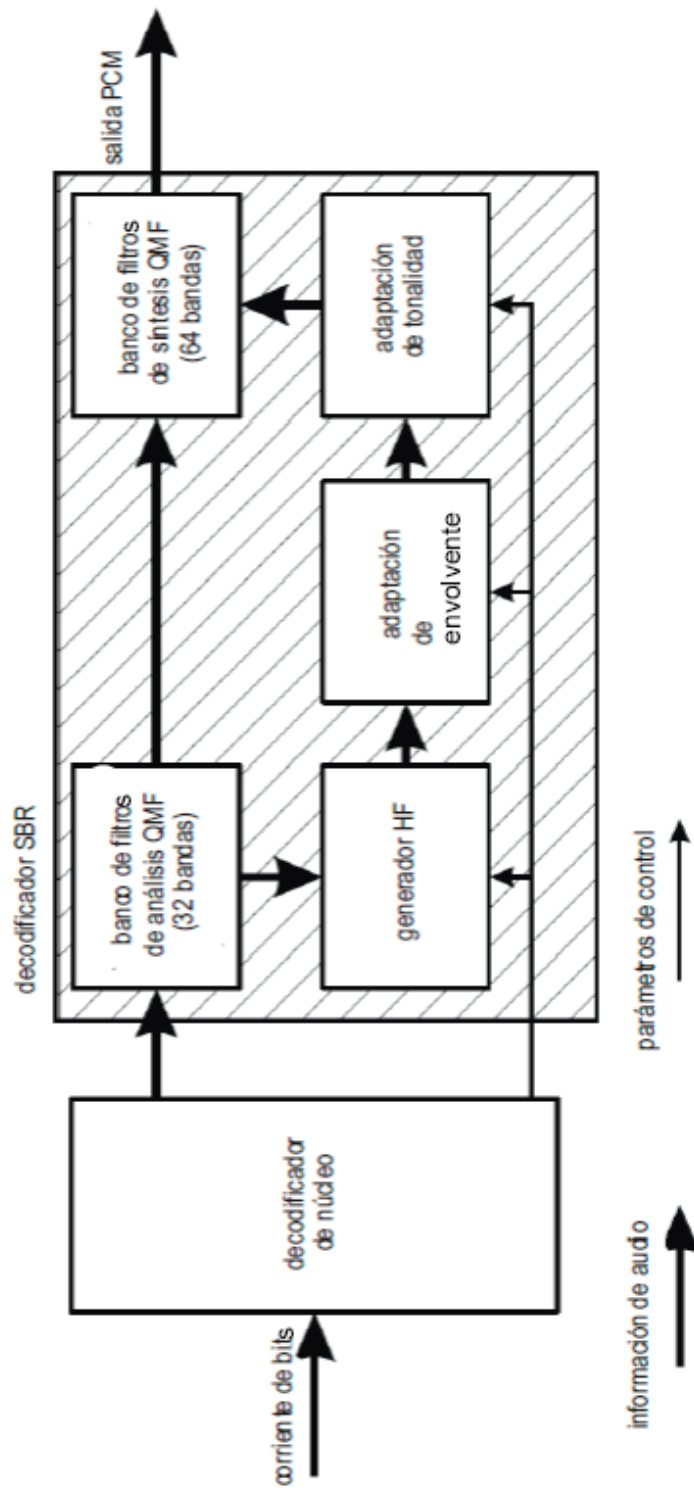


FIGURA 29

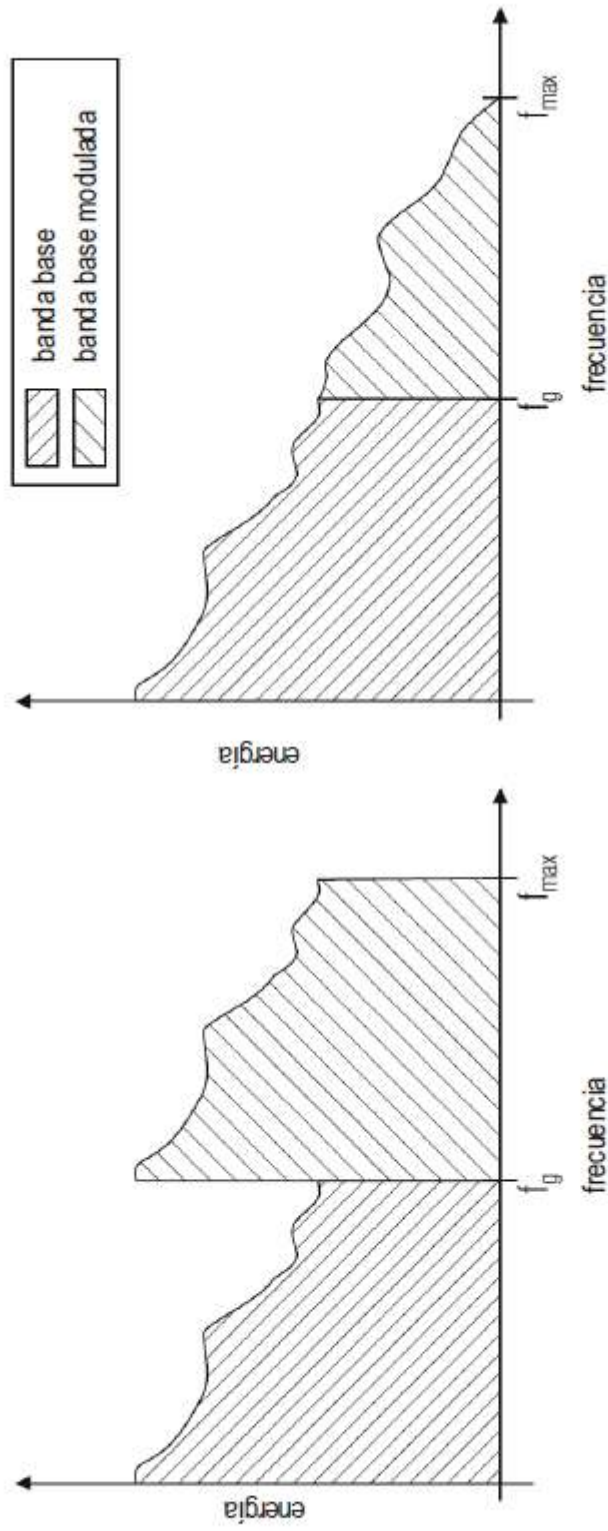


FIGURA 30

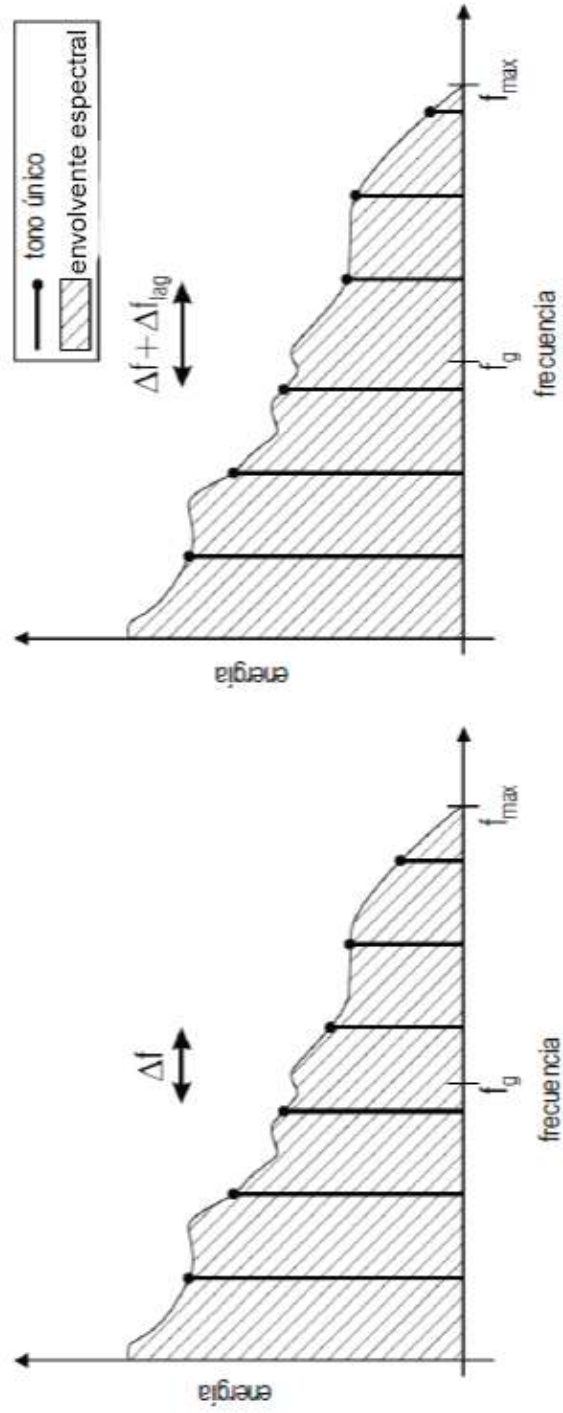


FIGURA 31

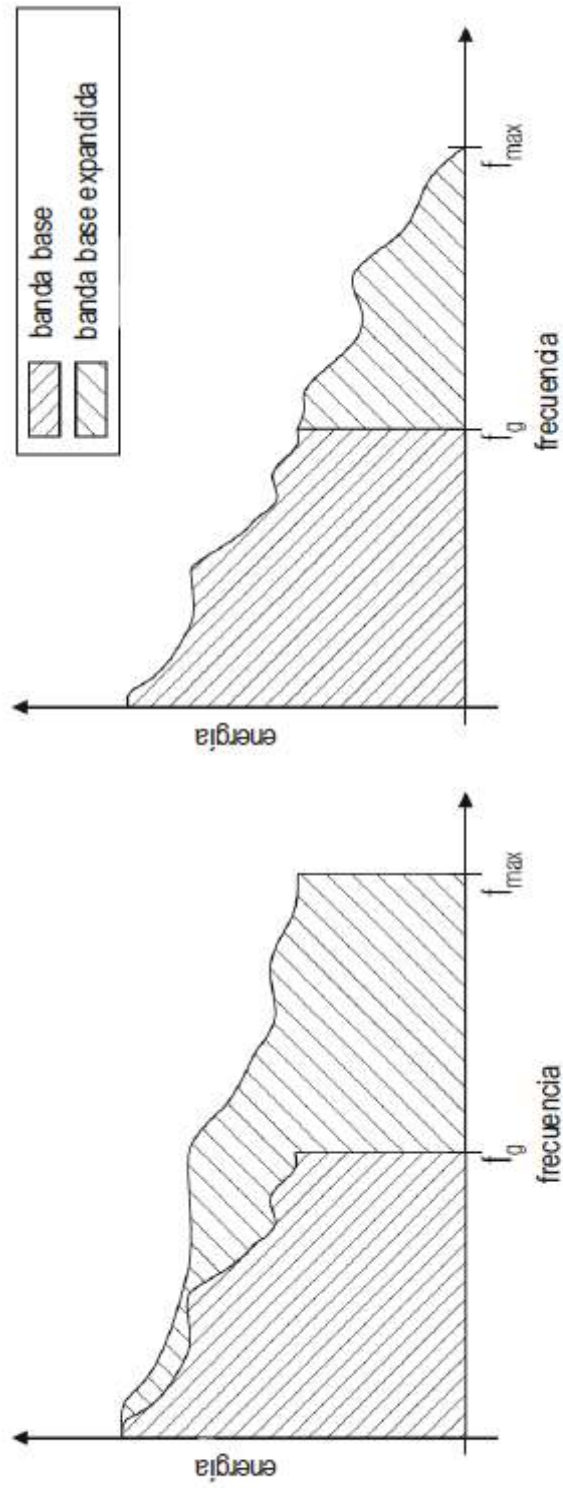


FIGURA 32

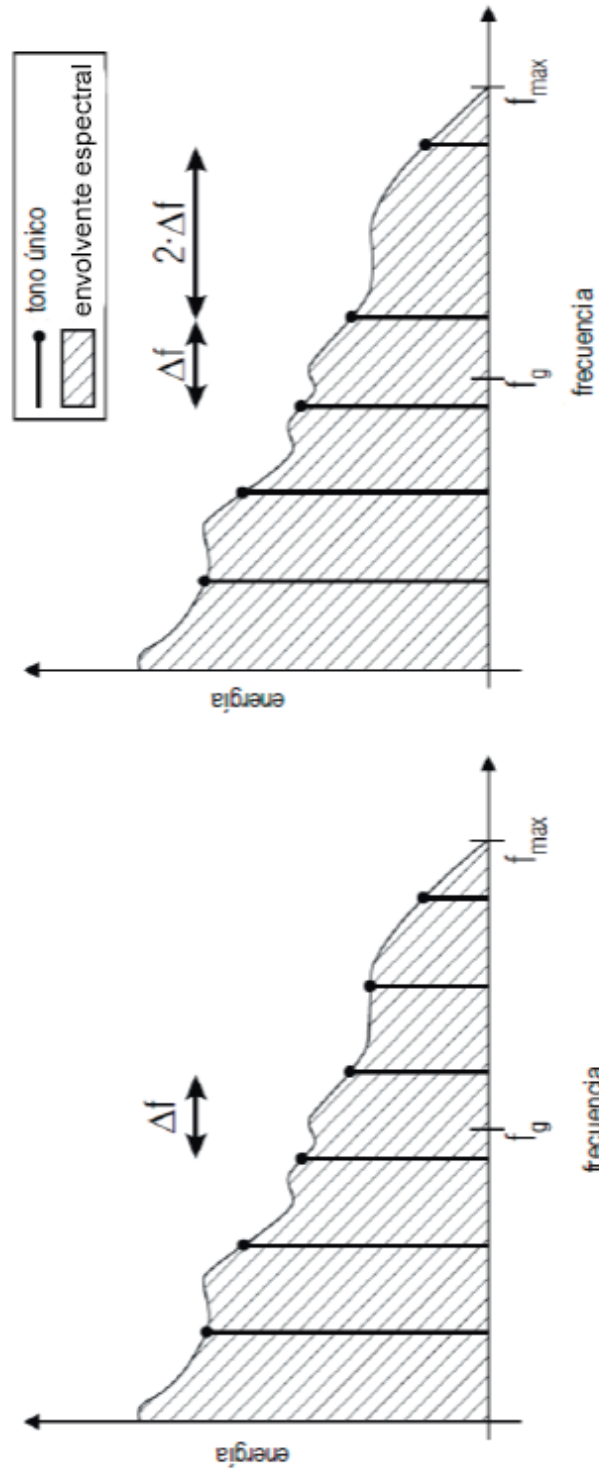


FIGURA 33

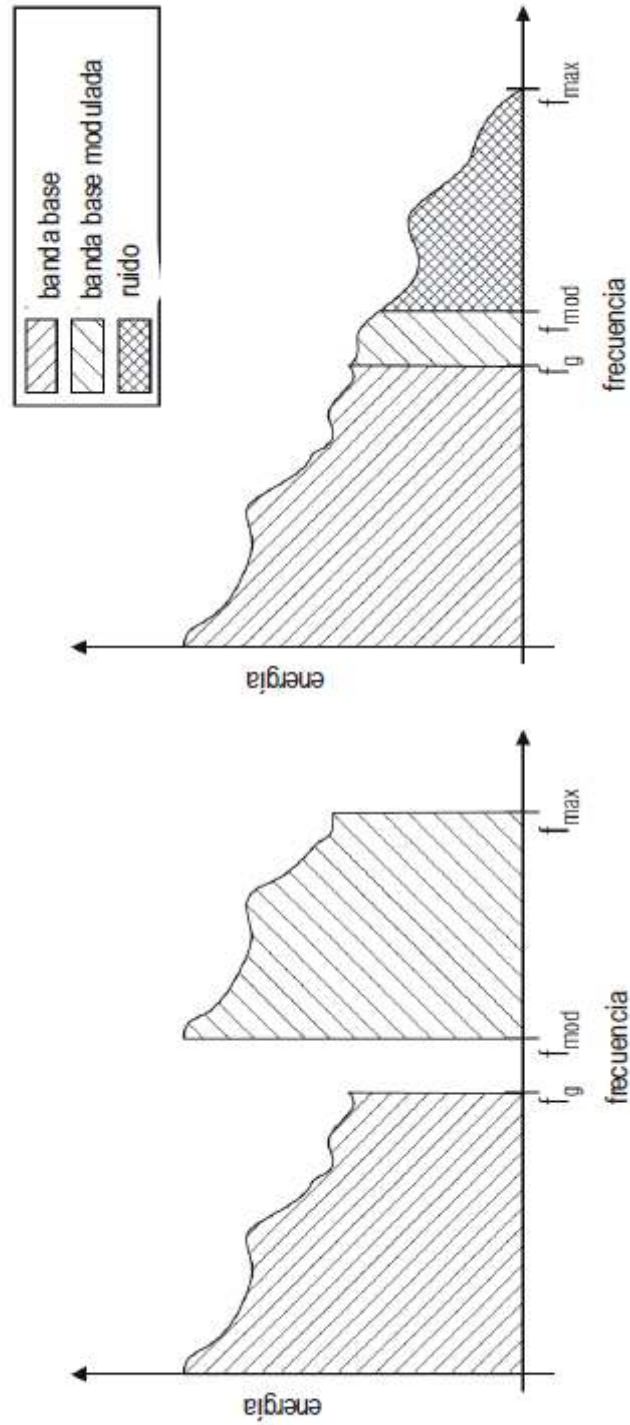


FIGURA 34

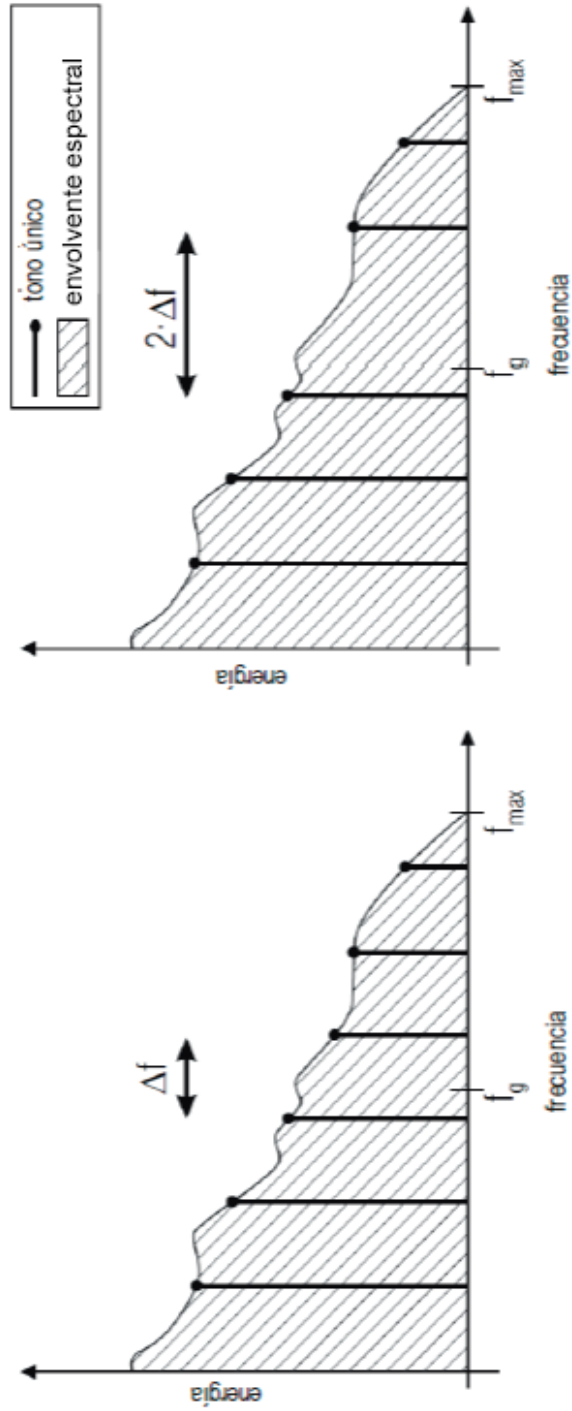


FIGURA 35