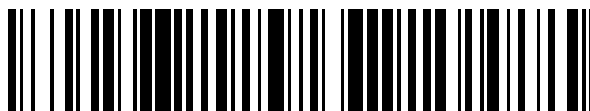


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 484 718**

51 Int. Cl.:

H03G 3/00 (2006.01)
H03G 5/00 (2006.01)
G10L 21/00 (2013.01)
G10L 25/90 (2013.01)
G10L 25/18 (2013.01)
G10H 1/08 (2006.01)
G10H 1/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.02.2011 E 11705571 (5)**

97 Fecha y número de publicación de la concesión europea: **02.07.2014 EP 2539885**

54 Título: **Aparato y procedimiento para modificar una señal de audio usando bloqueo de armónicos**

30 Prioridad:

03.09.2010 EP 10175282
26.02.2010 US 308513 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.08.2014

73 Titular/es:

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

72 Inventor/es:

DISCH, SASCHA

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 484 718 T3

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

DESCRIPCIÓN

Aparato y procedimiento para modificar una señal de audio usando bloqueo de armónicos

5 **[0001]** Las realizaciones según la invención se refieren al procesamiento de audio y particularmente a un aparato y un procedimiento para modificar una señal de audio.

[0002] Existe una demanda creciente de técnicas de procesamiento de señales digitales que se ocupen de la necesidad de manipulaciones de señales extremas con el fin de adecuar señales de audio pregrabadas, por ejemplo
10 tomadas de una base de datos, a un nuevo contexto musical. Con el fin de hacer eso, es necesario que se adapten las propiedades semánticas de alto nivel de la señal como la altura tonal, la clave musical y el modo de escala. Todas estas manipulaciones tienen en común que están dirigidas a alterar sustancialmente las propiedades musicales del material de audio original en tanto que preservando la calidad de sonido subjetiva tan buena como sea posible. En otras palabras, estas ediciones cambian totalmente el contenido musical del material de audio pero, no
15 obstante, se requiere que preserven la naturalidad de la muestra de audio procesada y mantengan así la verosimilitud. Esto requiere idealmente procedimientos de procesamiento de señales que sean ampliamente aplicables a diferentes clases de señales incluyendo contenido musical mezclado polifónico.

[0003] Hoy en día, se conocen muchos conceptos para modificar señales de audio. Algunos de estos
20 conceptos están basados en vocodificadores.

[0004] Por ejemplo, en el documento de S. Disch y B. Edler, "An amplitude- and frequency modulation vocoder for audio signal processing", Proc. of the Int. Conf. on Digital Audio Effects (DAFx), 2008, el documento de S. Disch y B. Edler, "Multiband perceptual modulation analysis, processing and Synthesis of audio signals", Proc. of the IEEE-
25 ICASSP, 2009, o el documento de S. Disch y B. Edler, "An iterative segmentation algorithm for audio signal spectra depending on estimated local centers of gravity", 12th International Conference on Digital Audio Effects (DAFx-09), 2009, se ha introducido el concepto de vocodificador de modulación (MODVOC) y se ha señalado su capacidad general para realizar una transposición selectiva significativa sobre el contenido musical polifónico. Esto hace que resulten posibles aplicaciones que están dirigidas a cambiar el modo de clave de muestras de música PCM pregrabadas (véase, por ejemplo, el documento de S. Disch y B. Edler, "Multiband perceptual modulation analysis, processing and Synthesis of audio signals", Proc. of the IEEE-ICASSP, 2009. También se dispone de un primer software disponible comercialmente que puede ocuparse de tal tarea de manipulación polifónica (el editor Melodyne de Celemony). El software implementa una tecnología que ha sido catalogado y comercializado por el término Direct Note Access (acceso directo a notas) (DNA). Recientemente se ha publicado una solicitud de patente (EP2099024,
30 P. Neubäcker, "Method for acoustic object-oriented analysis and note object-oriented processing of polyphonic sound recordings", septiembre de 2009), que cubre presumiblemente y por tanto desvela la funcionalidad esencial de DNA. Independientemente del procedimiento usado para modificar una señal de audio, se desea obtener una señal de audio con elevada calidad perceptiva.

40 **[0005]** El objeto de la presente invención es proporcionar un concepto mejorado para modificar una señal de audio, el cual permite aumentar la calidad perceptiva de la señal de audio modificada.

[0006] Este objeto se resuelve mediante un aparato según la reivindicación 1, un procedimiento según la reivindicación 12 o un programa informático según la reivindicación 13.
45

[0007] Una realización de la invención proporciona un aparato para modificar una señal de audio que comprende un procesador de banco de filtros, un determinador de tono fundamental, un determinador de sobretono, un procesador de señal y un combinador. El procesador de banco de filtros está configurado para generar una pluralidad de señales de paso de banda basándose en una señal de audio. Además, el determinador de tono
50 fundamental está configurado para seleccionar una señal de paso de banda de la pluralidad de señales de paso de banda para obtener una señal de paso de banda de tono fundamental. El determinador de sobretono está configurado para identificar una señal de paso de banda de la pluralidad de señales de paso de banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada para obtener una señal de paso de banda de sobretono asociada a la señal de paso de banda de tono fundamental seleccionada.
55 Además, el procesador de señal está configurado para modificar la señal de paso de banda de tono fundamental seleccionada basándose en un objetivo de modificación predefinido. Además, el procesador de señal está configurado para modificar una señal de paso de banda de sobretono identificada asociada a la señal de paso de banda de tono fundamental seleccionada dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada. Además, el combinador está configurado para combinar la pluralidad de señales de paso

de banda para obtener una señal de audio modificada.

[0008] Modificando los sobretonos de las frecuencias fundamentales y modificando los sobretonos del mismo modo que los tonos fundamentales correspondientes, puede evitarse una modificación diferente de los tonos fundamentales y sus sobretonos, de manera que el timbre de una señal de audio modificada puede preservarse con más exactitud en comparación con la señal de audio original. De este modo, la calidad perceptiva de la señal de audio modificada puede mejorarse significativamente. Por ejemplo, si se desea una transposición selectiva de altura tonal (por ejemplo, alterar el modo de clave de do mayor a do menor de una señal musical dada), la modificación de una señal de paso de banda de sobretono identificada está correlacionada con la modificación de la señal de paso de banda de tono fundamental. En comparación, los procedimientos conocidos modifican la zona de frecuencia de la señal de paso de banda que representa los sobretonos de manera diferente de la señal de paso de banda de tono fundamental. En otras palabras, una señal de paso de banda de sobretono identificada es bloqueada a la señal de paso de banda de tono fundamental usando el concepto descrito.

[0009] En algunas realizaciones de la invención, una señal de paso de banda de sobretono puede ser identificada comparando frecuencias de la señal de paso de banda de tono fundamental y señales de paso de banda de la pluralidad de señales de paso de banda, comparando un contenido de energía de la señal de paso de banda de tono fundamental y una señal de paso de banda de la pluralidad de señales de paso de banda y/o evaluando una correlación de una envolvente temporal de la señal de paso de banda de tono fundamental y la envolvente temporal de una señal de paso de banda de la pluralidad de señales de paso de banda. De este modo, pueden definirse uno o más criterios de sobretono para minimizar la identificación de sobretonos incorrectos.

[0010] Algunas realizaciones según la invención relacionadas con una determinación iterativa de señales de paso de banda de tono fundamental e identificación de señales de paso de banda de sobretono de la pluralidad de señales de paso de banda. Las señales de paso de banda de tono fundamental ya seleccionadas y las señales de paso de banda de sobretono ya identificadas pueden eliminarse del espacio de búsqueda o, en otras palabras, pueden no considerarse para la determinación de una señal de paso de banda de tono fundamental adicional o una señal de paso de banda de sobretono adicional. De este modo, cada señal de paso de banda de la pluralidad de señales de paso de banda puede seleccionarse como una señal de paso de banda de tono fundamental (y, por lo tanto, puede modificarse independientemente de las otras señales de paso de banda de tono fundamental) o una señal de paso de banda de sobretono (y, por lo tanto, puede modificarse dependiendo de la señal de paso de banda de tono fundamental seleccionada asociada).

[0011] Otra realización de la invención proporciona un aparato para modificar una señal de audio que comprende un determinador de forma de envolvente, un procesador de banco de filtros, un procesador de señal, un combinador y un conformador de envolvente. El determinador de forma de envolvente está configurado para determinar coeficientes de forma de envolvente basándose en una señal de audio en el dominio de la frecuencia que representa una señal de audio de entrada en el dominio del tiempo. Además, el procesador de banco de filtros está configurado para generar una pluralidad de señales de paso de banda en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia. El procesador de señal está configurado para modificar una señal de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en un objetivo de modificación predefinido. Además, el combinador está configurado para combinar al menos un subconjunto de la pluralidad de señales de paso de banda en el dominio de sub-banda para obtener una señal de audio en el dominio del tiempo. Además, el conformador de envolvente está configurado para conformar una envolvente de la señal de audio en el dominio del tiempo basándose en los coeficientes de forma de envolvente, para conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda que contienen la señal de paso de banda en el dominio de sub-banda modificada basándose en los coeficientes de forma de envolvente o para conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en los coeficientes de forma de envolvente antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal para obtener una señal de audio conformada.

[0012] Determinando los coeficientes de forma de envolvente de la señal de audio en el dominio de la frecuencia antes de que la señal de audio en el dominio de la frecuencia sea separada en una pluralidad de señales de paso de banda en el dominio de sub-banda, puede preservarse una información acerca de la coherencia espectral de la señal de audio y puede usarse para conformar la envolvente de la señal de audio en el dominio del tiempo después de la modificación de una o más señales de paso de banda en el dominio de sub-banda. De este modo, la coherencia espectral de la señal de audio modificada puede preservarse con más exactitud, aunque sólo se modifiquen algunas (o sólo una) señales de paso de banda en el dominio de sub-banda o se modifiquen de manera diferente las señales de paso de banda en el dominio de sub-banda, lo cual puede perturbar la coherencia espectral

de la señal de audio. De este modo, puede mejorarse significativamente la calidad perceptiva de la señal de audio modificada.

[0013] Algunas realizaciones según la invención se refieren a un procesador de señal configurado para modificar una segunda señal de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en un segundo objetivo de modificación predefinido. El objetivo de modificación predefinido y el segundo objetivo de modificación predefinido son diferentes. Aunque las señales de paso de banda se modifiquen de manera diferente, la coherencia espectral de la señal de audio modificada puede preservarse con más exactitud debido a la conformación de envolvente después de la modificación individual de las señales de paso de banda.

[0014] Las realizaciones según la invención se detallarán posteriormente haciendo referencia a los dibujos adjuntos, en los que:

- 15 La fig. 1 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la fig. 2 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la fig. 3 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- 20 la fig. 4 es un diagrama de bloques de una parte de un vocodificador de modulación que usa bloqueo de armónicos;
- la fig. 5 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- 25 la fig. 6a, 6b, 6c, 6d es un diagrama de bloques de un aparato para modificar una señal de audio;
- la fig. 7 es un diagrama de bloques de un procesador de banco de filtros;
- la fig. 8 es un diagrama de bloques de un conformador de envolvente;
- 30 la fig. 9 es una ilustración esquemática de un análisis de modulación con conformación de envolvente;
- la fig. 10 es una ilustración esquemática de una síntesis de modulación con conformación de envolvente;
- 35 la fig. 11 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- la fig. 12 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la fig. 13 es una ilustración esquemática de un análisis de modulación;
- 40 la fig. 14 es una ilustración esquemática de una implementación de un análisis de modulación;
- la fig. 16 es una ilustración esquemática de una transposición selectiva en un componente vocodificador de modulación;
- 45 la fig. 17 es una ilustración esquemática de un procedimiento para generar el conjunto de pruebas para evaluación de la calidad subjetiva del procesamiento del vocodificador de modulación para la tarea de transposición selectiva de altura tonal;
- 50 la fig. 18 es un diagrama que indica unas puntuaciones MUSHRA absolutas e intervalos de confianza del 95% de la prueba de escucha que se ocupa de la transposición selectiva de altura tonal;
- la fig. 19 es un diagrama que indica unas puntuaciones MUSHRA de diferencia con respecto a la condición del vocodificador de modulación e intervalos de confianza del 95% de la prueba de escucha que se ocupa de la transposición selectiva de altura tonal; y
- 55 la fig. 20 es un diagrama que indica unas puntuaciones MUSHRA de diferencia con respecto a la condición DNA e intervalos de confianza del 95% de la prueba de escucha que se ocupa de la transposición selectiva de altura tonal.

[0015] En lo que viene a continuación, se usan parcialmente los mismos números de referencia para objetos y unidades funcionales que tienen propiedades funcionales iguales o similares y la descripción de los mismos con respecto a una figura también se aplicará a otras figuras con el fin de reducir la redundancia en la descripción de las realizaciones.

5

[0016] Una modificación de banda de frecuencia selectiva, también denominada transposición de altura tonal selectiva, puede realizarse, por ejemplo, mediante un vocodificador o vocodificador de modulación.

[0017] Una descomposición de modulación multibanda (véase, por ejemplo, el documento de S. Ditch y B. Edler, "Multiband perceptual modulation analysis, processing and Synthesis of audio signals", Proc. of the IEEE-ICASSP, 2009) disecciona la señal de audio en un conjunto adaptativo de señales de paso de banda (analíticas), cada una de las cuales es dividida además en una portadora sinusoidal y su modulación de amplitud (AM) y modulación de frecuencia (FM). El conjunto de filtros de paso de banda puede computarse de manera que, por una parte, el espectro de la banda completa es cubierto sin interrupciones y, por otra parte, los filtros están alineados con los centros de gravedad totales (COG), por ejemplo. Además, la percepción auditiva humana puede tomarse en consideración eligiendo la anchura de banda de los filtros que coincida con una escala perceptiva, por ejemplo, la escala ERB (véase, por ejemplo, el documento de B. C. J. Moore y B. R. Glasberg, "A revision of Zwicker's loudness model" Acta Acustica, vol. 82, págs. 335-345, 1996).

[0018] Por ejemplo, el COG local corresponde a la frecuencia media que es percibida por un oyente debido a las contribuciones espectrales en esa zona de frecuencia. Por otra parte, las bandas centradas en las posiciones del COG local pueden corresponder a zonas de enganche de fase basado en influencia de vocodificadores de fase clásicos (véase, por ejemplo, el documento de J. Laroche y M. Dolson, "Improved phase vocoder timescale modification of audio", IEEE Transactions on Speech and Audio Processing, vol. 7, nº 3, págs. 323-332, 1999, o el documento de Ch. Duxbury, M Davies y M. Sandler, "Improved timescaling of musical audio using phase locking at transients", en la 112th AES Convention, 2002). Tanto la representación de la envolvente de señal de paso de banda como la zona tradicional de enganche de fase de influencia preservan la envolvente temporal de una señal de paso de banda: ya sea intrínsecamente o, en este caso, asegurando la coherencia de fase espectral local durante la síntesis. Con respecto a una portadora sinusoidal de una frecuencia que corresponde al COG local estimado, se captura tanto la AM como la FM en la envolvente de amplitud y la fase heterodinada de las señales de paso de banda analíticas, respectivamente. Un procedimiento de síntesis dedicado da la señal de salida a partir de las frecuencias portadoras, AM y FM.

[0019] En la figura 13 se representa un diagrama de bloques de una posible implementación 1300 de la descomposición de señales en señales portadoras y sus componentes de modulación asociadas. En la figura, se muestra el flujo de señales esquemático para la extracción de una de las componentes multibanda (señales de paso de banda). Todas las demás componentes se obtienen de manera similar. En primer lugar, una señal de entrada de banda ancha x se suministra a un filtro de paso de banda que ha sido diseñado adaptativamente a la señal produciendo una señal de salida. Después, la señal analítica se deduce mediante la transformada de Hilbert según la ecuación (1).

$$\hat{x}(t) = \tilde{x}(t) + jH(\tilde{x}(t)) \quad (1)$$

[0020] La AM (señal de modulación de amplitud) viene dada por la envolvente de amplitud de $\hat{x}$.

45

$$AM(t) = |\hat{x}(t)| \quad (2)$$

mientras que la FM (señal de modulación de frecuencia) se obtiene por la derivada de la fase de la señal analítica heterodinada por una portadora sinusoidal estacionaria con frecuencia angular ω_c . Se determina que la frecuencia portadora es una estimación del COG local. De ahí que la FM pueda interpretarse como la variación de IF (frecuencia instantánea) en la frecuencia portadora f_c .

50

$$\begin{aligned} \dot{x}(t) &= \hat{x}(t) \cdot \exp(j\omega_c t) \\ FM(t) &= \frac{1}{2\pi} \cdot \frac{d}{dt} \angle(\dot{x}(t)) \end{aligned} \quad (3)$$

[0021] La estimación del COG local y el diseño adaptativo de señal del banco de filtros del terminal de entrada se describe, por ejemplo, en una publicación dedicada (véase el documento de S. Disch y B. Edler, "An iterative segmentation algorithm for audio signal spectra depending on estimated local centers of gravity", 12h International Conference on Digital Audio Effects (DAFx-09), 2009).

[0022] Prácticamente, en un sistema de tiempo discreto, la extracción de componentes puede llevarse a cabo conjuntamente para todas las componentes tal como se ilustra en la figura 14. El esquema de procesamiento puede soportar computación en tiempo real. El procesamiento de un cierto bloque de tiempo sólo depende de parámetros de los bloques previos. Por consiguiente, no se requiere anticipación con el fin de mantener el retardo de procesamiento global lo más bajo posible. El procesamiento se computa basándose en bloque por bloque usando, por ejemplo, el 75% de solapamiento de bloques de análisis y la aplicación de una transformada discreta de Fourier (DFT) sobre cada bloque de señal de tipo ventana. La ventana puede ser una ventana de tipo "flat top" según la ecuación (4).

[0023] Esto asegura que las $N/2$ muestras centradas que se pasan para la síntesis de modulación subsiguiente que utiliza el 50% de solapamiento no se ven afectadas por los márgenes de la ventana de análisis. Puede usarse un grado más alto de solapamiento para una exactitud mejorada a costa de una mayor complejidad de cálculo.

$$window(i)_{analysis} = \begin{cases} \sin^2\left(\frac{2i\pi}{N}\right) \dots 0 < i < \frac{N}{4} \\ 1 \dots \frac{N}{4} \leq i < \frac{3N}{4} \\ \sin^2\left(\frac{2i\pi}{N}\right) \dots \frac{3N}{4} \leq i < N \end{cases} \quad (4)$$

[0024] Dada la representación espectral, a continuación se calcula un conjunto de funciones de ponderación de paso de banda espectral adaptativo de señal que está alineado con las posiciones del COG local. Después de la aplicación de la ponderación de paso de banda al espectro, la señal se transfiere al dominio del tiempo y la señal analítica puede deducirse por la transformada de Hilbert. Estas dos etapas de procesamiento pueden combinarse eficazmente mediante el cálculo de una IDFT unilateral en cada señal de paso de banda. Dada la señal de paso de banda de tiempo discreto, la estimación de la IF mediante la ecuación (3) se implementa por diferenciación de fase tal como se define en la ecuación (5), donde * indica el conjugado complejo. Esta expresión se usa convenientemente ya que evita las ambigüedades de fase y de ahí la necesidad de desenrollado de fase.

$$FM(n) = \angle(\hat{x}(n)\hat{x}(n-1)^*) \quad (5)$$

[0025] La señal se sintetiza sobre una base aditiva de todas las componentes. Los bloques sucesivos son combinados mediante solapamiento y suma (OLA) que se controla mediante el mecanismo de unión. La unión de componentes asegura una transición suave entre los bordes de bloques adyacentes aunque las componentes sean alteradas sustancialmente por un procesamiento de dominio de modulación. La unión sólo tiene en cuenta el bloque previo, permitiendo así potencialmente el procesamiento en tiempo real. La unión realiza esencialmente una concordancia por parejas de las componentes del bloque real con sus predecesores en el bloque previo. Para las componentes que no tienen coincidencia a través de los bloques de tiempo, se aplica una aparición gradual o desaparición gradual, respectivamente.

[0026] En la figura 15 se muestra la cadena de procesamiento para una componente. Detalladamente, primero se suma la señal FM a la frecuencia portadora estacionaria y la señal resultante se pasa a una fase OLA, cuya salida se integra temporalmente con posterioridad. A un oscilador sinusoidal se le suministra la señal de fase resultante. La señal AM es procesada por una segunda fase OLA. A continuación, la salida del oscilador es modulada en su amplitud por la señal AM para obtener la contribución aditiva de la componente a la señal de salida. En una etapa final, las contribuciones de todas las componentes se suman para obtener la señal de salida y.

[0027] En otras palabras, las figs. 13 y 14 ilustran un analizador de modulación 1300. El analizador de modulación

1300 comprende preferentemente un filtro de paso de banda 1320a, que proporciona una señal de paso de banda. Esta es introducida en un convertidor de señal analítica 1320b. La salida del bloque 1320b es útil para calcular la información de AM y la información de FM. Para calcular la información de AM, la magnitud de la señal analítica es calculada por el bloque 1320c. La salida del bloque de señal analítica 1320b es introducida en un multiplicador 1320d, que recibe, en su otra entrada, una señal de oscilador procedente de un oscilador 1320e, que es controlado por la frecuencia portadora real f_c 1310 del paso de banda 1320a. Después, la fase de la salida del multiplicador es determinada en el bloque 1320f. La fase instantánea es diferenciada en el bloque 1320g con el fin de obtener finalmente la información de FM. Además, la fig. 14 muestra un preprocesador 1410 que genera un espectro DFT de la señal de audio.

10

[0028] La descomposición de modulación multibanda disecciona la señal de audio en un conjunto adaptativo de señales de paso de banda (analíticas), cada una de las cuales está dividida además en una portadora sinusoidal y su modulación de amplitud (AM) y modulación de frecuencia (FM). El conjunto de filtros de paso de banda se computa de manera que, por una parte, el espectro de la banda completa es cubierto sin interrupciones y, por otra parte, los filtros están alineados con cada uno de los centros de gravedad totales (COG). Además, la percepción auditiva humana se toma en consideración eligiendo la anchura de banda de los filtros que coincida con una escala perceptiva, por ejemplo, la escala ERB (véase el documento de B. C. J. Moore y B. R. Glasberg, "A revision of Zwicker's loudness model" Acta Acustica, vol. 82, págs. 335-345, 1996).

15

[0029] El COG local corresponde a la frecuencia media que es percibida por un oyente debido a las contribuciones espectrales en esa zona de frecuencia. Por otra parte, las bandas centradas en las posiciones del COG local corresponden a zonas de enganche de fase basado en influencia de vocodificadores de fase clásicos (véase el documento de J. Laroche y M. Dolson, "Improved phase vocoder timescale modification of audio", IEEE Transactions on Speech and Audio Processing, vol. 7, nº 3, págs. 323-332, 1999, el documento de Ch. Duxbury, M Davies y M. Sandler, "Improved timescaling of musical audio using phase locking at transients", en la 112th AES Convention, 2002, el documento de A. Röbel, "A new approach to transient processing in the phase vocoder", Proc. Of the Int. Conf. on Digital Audio Effects (DAFx), págs. 344-349, 2003, el documento de A. Röbel. "Transient detection and preservation in the phase vocoder", Int. Computer Music Conference (ICMC'03), págs. 247-250, 2003. Tanto la representación de la envolvente de señal de paso de banda como la zona tradicional de enganche de fase de influencia preservan la envolvente temporal de una señal de paso de banda: ya sea intrínsecamente o, en este caso, asegurando la coherencia de fase espectral local durante la síntesis. Con respecto a una portadora sinusoidal de una frecuencia que corresponde al COG local estimado, se captura tanto la AM como la FM en la envolvente de amplitud y la fase heterodinada de las señales de paso de banda analíticas, respectivamente. Un procedimiento de síntesis dedicado da la señal de salida a partir de las frecuencias portadoras, AM y FM

35

[0030] En la figura 12 se representa un diagrama de bloques de la descomposición de señales en señales portadoras y sus componentes de modulación asociadas. En la imagen, se muestra el flujo de señales esquemático para la extracción de una componente. Todas las demás componentes se obtienen de manera similar. En la práctica, la extracción se lleva a cabo conjuntamente para todas las componentes a nivel de bloque por bloque usando, por ejemplo, un tamaño de bloque de $N = 2^{14}$ a frecuencia de muestreo de 48 kHz y el 75% de solapamiento de análisis – que corresponde aproximadamente a un intervalo de tiempo de 340 ms y un paso de 85 ms – mediante la aplicación de una transformada discreta de Fourier (DFT) en cada bloque de señal de tipo ventana. La ventana puede ser una ventana tipo "flat top" según la ecuación (a). Esto puede asegurar que las $N/2$ muestras centradas que se pasan para la síntesis de modulación subsiguiente no se ven afectadas por las pendientes de la ventana de análisis. Puede usarse un grado más alto de solapamiento para una exactitud mejorada a costa de una mayor complejidad de cálculo.

45

$$window(i)_{analysis} = \begin{cases} \sin^2\left(\frac{2i\pi}{N}\right) \dots 0 < i < \frac{N}{4} \\ 1 \dots \frac{N}{4} \leq i < \frac{3N}{4} \\ \sin^2\left(\frac{2i\pi}{N}\right) \dots \frac{3N}{4} \leq i < N \end{cases} \quad (a)$$

[0031] Dada la representación espectral, a continuación puede calcularse (mediante el determinante de frecuencia

portadora 1330 en cuanto a una estimación de frecuencia portadora o una estimación de frecuencia de COG de portadoras múltiples) un conjunto de funciones de ponderación espectral adaptativa de señal (que tiene característica de paso de banda) que está alineado con las posiciones del COG local. Después de la aplicación de la ponderación de paso de banda al espectro, la señal se transforma al dominio del tiempo y la señal analítica se deduce por la transformada de Hilbert. Estas dos etapas de procesamiento pueden combinarse eficazmente mediante el cálculo de una IDFT unilateral en cada señal de paso de banda. Posteriormente, cada señal analítica es heterodinada por su frecuencia portadora estimada. Por último, la señal es descompuesta además en su envolvente de amplitud y su pista de frecuencia instantánea (IF), obtenida computando la derivada de la fase, produciendo la señal AM y FM deseada (véase también el documento de S. Disch y B. Edler, "An amplitude- and frequency modulation vocoder for audio signal processing", Proc. of the Int. Conf. on Digital Audio Effects (DAFx), 2008).

[0032] Apropiadamente, la fig. 15 muestra un diagrama de bloques de un sintetizador de modificación 1500 una representación parametrizada de una señal de audio. Por ejemplo, una implementación ventajosa está basada en una operación de solapamiento y suma (OLA) en el dominio de la modulación, es decir, en el dominio antes de generar la señal de paso de banda en el dominio del tiempo. La señal de entrada que puede ser un flujo de bits, pero que también puede ser una conexión directa a un analizador o modificador también, es separada en la componente de AM 1502, la componente de FM 1504 y la componente de frecuencia portadora 1506. El sintetizador de AM comprende con preferencia un sumador de solapamiento 1510 y, además, un controlador de unión de componentes 1520 que, con preferencia no sólo comprende el bloque 1510 sino también el bloque 1530, que es un sumador de solapamiento dentro del sintetizador de FM. El sintetizador de FM además comprende un sumador de solapamiento de frecuencia 1530, un integrador de frecuencia instantánea 1532, un combinador de fases 1534 que, de nuevo, puede implementarse como un sumador regular y un desfaseador 1536 que es controlable mediante el controlador de unión de componentes 1520 con el fin de regenerar una fase constante de bloque a bloque de manera que la fase de una señal procedente de un bloque precedente sea continua con la fase de un bloque real. Por lo tanto, se puede decir que la adición de fase en los elementos 1534, 1536 corresponde a una regeneración de una constante que se perdió durante la diferenciación en el bloque 1520g en la fig. 13 en el lado del analizador. Desde una perspectiva de pérdida de información en el dominio perceptivo, ha de observarse que esta es la única pérdida de información, es decir, la pérdida de una porción constante por el dispositivo de diferenciación 1320g en la fig. 13. Esta pérdida puede compensarse sumando una fase constante determinada por el dispositivo de unión de componentes 1520.

[0033] Se aplica solapamiento y suma en el dominio paramétrico en lugar de sobre la señal inmediatamente sintetizada con el fin de evitar los efectos de pulsación entre bloques de tiempo adyacentes. El OLA se controla mediante un mecanismo de unión de componentes, que, dirigido por la proximidad espectral (medida en una escala ERB), realiza una concordancia por parejas de componentes del bloque real con sus predecesores en el bloque previo. Además, la unión alinea las fases de componentes absolutas del bloque real con las del bloque previo.

[0034] Detalladamente, en primer lugar la señal FM se suma a la frecuencia portadora y el resultado se pasa a la fase OLA, cuya salida se integra posteriormente. La señal de fase resultante se suministra a un oscilador sinusoidal 1540. La señal AM es procesada por una segunda fase OLA. Por último, la salida del oscilador es modulada 1550 en su amplitud por la señal AM resultante para obtener la contribución aditiva de la componente a la señal de salida 1560.

[0035] Debería ponerse énfasis en que una segmentación espectral apropiada de la señal dentro del análisis de modulación es de primordial importancia para un resultado convincente de cualquier procesamiento adicional de parámetros de modulación. Por lo tanto, en este documento, se describe un ejemplo para un algoritmo de segmentación adecuado.

[0036] Apropiadamente, la fig. 16 muestra un ejemplo 1600 para una aplicación para cambios de modo de clave polifónica. La figura muestra una transposición selectiva sobre componentes del vocodificador de modulación. Las frecuencias portadoras son cuantificadas a notas MIDI que se hacen corresponder con notas MIDI correspondientes apropiadas. La preservación de la modulación FM relativa mediante la multiplicación de las componentes asignadas por la relación de la frecuencia portadora original y la modificada.

[0037] La transposición de una señal de audio en tanto que manteniendo la velocidad de reproducción original es un desafío. Usando el sistema propuesto, esto se logra sencillamente mediante la multiplicación de todas las componentes portadoras por un factor constante. Como la estructura temporal de la señal de entrada únicamente es capturada por las señales AM no se ve afectada por el estiramiento del alargamiento espectral de la portadora.

[0038] Puede obtenerse un efecto aún más exigente mediante procesamiento selectivo. El modo de clave de una

pieza musical puede cambiarse de, por ejemplo, menor a mayor o viceversa. Por lo tanto, sólo un subconjunto de portadoras que corresponde a ciertos intervalos de frecuencia predefinidos se hace corresponder con nuevos valores adecuados. Para lograr esto, las frecuencias portadoras son cuantificadas 1670 a alturas tonales MIDI que se hacen corresponder posteriormente 1672 con nuevas alturas tonales MIDI apropiadas (usando un conocimiento a priori del modo y la clave de la pieza musical que ha de procesarse).

[0039] Después, las notas MIDI que se han hecho corresponder se vuelven a convertir 1574 con el fin de obtener las frecuencias portadoras modificadas que se usan para la síntesis. No se requiere una detección de comienzo/desplazamiento de notas MIDI dedicada ya que las características temporales están representadas predominantemente por la AM sin modificar y de este modo se preservan. Pueden definirse tablas de correspondencia arbitraria que permiten la conversión a y desde otros sabores menores (por ejemplo, armónico menor).

[0040] Una aplicación en el terreno de los efectos de audio es la transposición global de una señal de audio. El procesamiento requerido para este efecto de audio es una simple multiplicación de las portadoras con un factor de transposición constante. Multiplicando también la FM con el mismo factor se asegura que, para cada componente, se preserve la profundidad de modulación FM relativa. Como la estructura temporal de la señal de entrada únicamente es capturada por las señales AM no se ve afectada por el procesamiento. La transposición global cambia la clave original de una señal de música hacia una clave objetivo (por ejemplo, de do mayor a sol mayor) en tanto que preservando el tempo original.

[0041] Sin embargo, debido a la naturaleza adaptativa de la señal del análisis de modulación propuesto, el vocodificador de modulación tiene el potencial de ir más allá de esta tarea. Ahora, incluso la transposición de componentes seleccionadas de la música polifónica se vuelve factible, permitiendo aplicaciones que, por ejemplo, alteran el modo de clave (por ejemplo, de do mayor a do menor) de una señal de música dada (véase, por ejemplo, el documento de S. Disch y B. Edler, "Multiband perceptual modulation analysis, processing and Synthesis of audio signals", Proc. of the IEEE-ICASSP, 2009). Esto es posible debido al hecho de que cada portadora componente se corresponde estrechamente con la altura tonal percibida en su zona espectral. Si sólo las portadoras que se relacionan con ciertas alturas tonales originales se hacen corresponder con nuevos valores objetivo, se manipula el carácter musical global que se determina por el modo de clave.

[0042] El procesamiento necesario en los componentes del MODVOC se representa en la figura 16 tal como se mencionó anteriormente. Dentro del dominio de la descomposición del MODVOC, las frecuencias portadoras son cuantificadas a notas MIDI que posteriormente se hace corresponder con notas MIDI correspondientes apropiadas. Para una reasignación significativa de las alturas tonales midi y los nombres de notas, puede requerirse un conocimiento a priori del modo y la clave de la pieza musical original. No se actúa en absoluto sobre la AM de todas las componentes ya que estas no contienen información de altura tonal.

[0043] Específicamente, las frecuencias portadoras componentes f , que representan la altura tonal de las componentes, son convertidas en valores de altura tonal MIDI m según la ecuación 6, donde f_{std} indica la altura tonal estándar que corresponde a la altura tonal MIDI 69, la nota A0.

$$m(f) = 69 + 12 \cdot \log_2 \frac{|f|}{f_{std}}$$

$$n(f) = \text{round}(m(f)) \quad (6)$$

$$o(f) = m(f) - n(f)$$

$$n \rightarrow n'$$

$$n \rightarrow n'$$

$$f' = f_{std} \cdot 2^{(n' + o(f) - 69)/12} \quad (7)$$

[0044] Posteriormente, las alturas tonales MIDI son cuantificadas a notas MIDI $n(f)$ y, además, se determina el desplazamiento de altura tonal $o(f)$ de cada nota. Mediante la utilización de una tabla de correspondencia de notas MIDI que depende sólo de la clave, el modo original y el modo objetivo, estas notas MIDI son transformadas en

valores objetivo apropiados n' . En la tabla de más adelante, se ofrece una correspondencia ejemplar para la clave de do de mayor a menor natural. La tabla muestra una tabla de correspondencia de notas MIDI para una transformación de modo de escala de do mayor a do menor natural. La correspondencia se aplica para las notas de todas las octavas.

5

Nota original	Nota objetivo
do	do
re	re
mi	mi bemol
fa	fa
sol	sol
la	la bemol
si	si bemol

[0045] Por último, las notas MIDI que se han hecho corresponder incluyendo sus desplazamientos de altura tonal se vuelven a convertir en frecuencia f con el fin de obtener las frecuencias portadoras modificadas que se usan para la síntesis (ecuación 7). Además, con el fin de preservar la profundidad de modulación FM relativa, la FM de una componente que se ha hecho corresponder es multiplicada por el factor de transposición de altura tonal individual que se obtiene como la relación de la frecuencia portadora original y la modificada. Puede no requerirse una detección de comienzo/desplazamiento de notas MIDI dedicada ya que las características temporales están representadas predominantemente por la AM sin modificar y de este modo se preservan.

10

[0046] El vocodificador de modulación descrito es una posibilidad para modificar diferentes intervalos de frecuencia (señales de paso de banda) de señales de audio de manera diferente, lo cual se mencionó como transposición selectiva de altura tonal. El concepto inventivo permite la mejora de la calidad perceptiva de tales señales de audio modificadas. Aunque algunas realizaciones del concepto inventivo se describen en relación con un vocodificador o un vocodificador de modulación, también puede usarse generalmente para mejorar la calidad perceptiva de las señales de audio modificadas independientemente de la utilización de un vocodificador.

15

[0047] La fig. 1 muestra un diagrama de bloques de un aparato 100 para modificar una señal de audio 102 según una realización de la invención. El aparato 100 comprende un procesador de banco de filtros 110, un determinador de tono fundamental 120, un determinador de sobretono 130, un procesador de señal 140 y un combinador 150. El procesador de banco de filtros 110 está conectado al determinador de tono fundamental 120, el determinador de sobretono 130 y el procesador de señal 140 así como el determinador de tono fundamental 120 está conectado al determinador de sobretono 130 y el procesador de señal 140. Además, el determinador de sobretono 130 está conectado al procesador de señal 140 y el procesador de señal 140 está conectado al combinador 150. El procesador de banco de filtros 110 genera una pluralidad de señales de paso de banda 112 basándose en una señal de audio 102. Además, el determinador de tono fundamental selecciona una señal de paso de banda 112 de la pluralidad de señales de paso de banda para obtener una señal de paso de banda de tono fundamental 122. El determinador de sobretono identifica una señal de paso de banda 112 de la pluralidad de señales de paso de banda que cumple un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada 122 para obtener una señal de paso de banda de sobretono 132 asociada a la señal de paso de banda de tono fundamental seleccionada 122. Además, el procesador de señal 140 modifica la señal de paso de banda de tono fundamental seleccionada 122 basándose en un objetivo de modificación predefinido. Además, el procesador de señal 140 modifica una señal de paso de banda de sobretono identificada 132 asociada a la señal de paso de banda de tono fundamental seleccionada 122 dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada 122. Además, el combinador 150 combina la pluralidad de señales de paso de banda que contienen la señal de paso de banda de tono fundamental seleccionada y la señal de paso de banda de sobretono identificada modificada para obtener una señal de audio modificada 152.

25

30

35

40

[0048] Modificando la señal de paso de banda de tono fundamental 122 y la señal de paso de banda de sobretono identificada 132 asociada a la señal de paso de banda de tono fundamental 122 del mismo modo, puede preservarse un comportamiento común de estos armónicos, aunque otras señales de paso de banda de la pluralidad de señales de paso de banda pueden modificarse de diferentes modos. De este modo, el timbre de la señal de audio original 102 puede mantenerse con más exactitud, de manera que la calidad perceptiva de la señal de audio modificada puede mejorarse significativamente. Por ejemplo, la mayoría de los instrumentos excitan sonidos armónicos constituidos por una parte de frecuencia fundamental y sus armónicos. Si la parte de frecuencia fundamental debe modificarse, entonces una modificación correlacionada de los armónicos según el concepto descrito puede producir una calidad perceptiva significativamente mejor de la señal de audio modificada. Además, la

45

50

señal de audio puede modificarse en tiempo real, ya que puede no ser necesaria una información a priori acerca de toda la señal de audio entera (por ejemplo, el título de toda la música polifónica).

[0049] La señal de audio 102 puede ser, por ejemplo, una señal de audio de entrada en el dominio del tiempo o una señal de audio en el dominio de la frecuencia que representa una señal de audio de entrada en el dominio del tiempo.

[0050] El determinador de tono fundamental 120 puede proporcionar la señal de paso de banda de tono fundamental seleccionada 122 al procesador de señal 140 para su modificación o puede proporcionar una señal de activación 122 (por ejemplo, un índice $i \in [0.../-1]$ de la señal de paso de banda de tono fundamental seleccionada, en la que i es el número de señales de paso de banda de la pluralidad de señales de paso de banda) para activar el procesador de señal 140 para modificar la señal de paso de banda seleccionada de la pluralidad de señales de paso de banda según el objetivo de modificación predefinido. Por consiguiente, también el determinador de sobretono 130 puede proporcionar al procesador de señal 140 la señal de paso de banda de sobretono identificada 132 para la modificación o puede proporcionar una señal de activación 132 (por ejemplo, un índice que indica la señal de paso de banda de la pluralidad de señales de paso de banda que es identificada como la señal de paso de banda de sobretono) para activar el procesador de señal 140 para modificar la señal de paso de banda identificada de la pluralidad de señales de paso de banda.

[0051] El criterio de sobretono puede comprender una o más reglas para identificar un sobretono del tono fundamental. Puede haber uno o más criterios de sobretono que han de cumplirse para identificar una señal de paso de banda de la pluralidad de señales de paso de banda como un sobretono de la señal de paso de banda de tono fundamental seleccionada 122.

[0052] El objetivo de modificación predefinido puede ser diferente para las señales de paso de banda que comprenden diferente intervalo de frecuencia y puede depender de la modificación deseada de la señal de audio 102. Por ejemplo, la clave original de una señal de audio debería cambiarse hacia una clave objetivo. Mediante la tabla anterior se dio una correspondencia ejemplar para la clave de do de mayor a menor natural. Por ejemplo, si un intervalo de frecuencia de una señal de paso de banda de la pluralidad de señales de paso de banda corresponde a una nota do original, la nota objetivo también sería do, de manera que esta señal de paso de banda no se modifica (excepto que se identifique como señal de paso de banda de sobretono de una señal de paso de banda de tono fundamental asociada, que se modifica). En este caso, el objetivo de modificación es mantener esta señal de paso de banda sin modificar. Por otra parte, puede modificarse una señal de paso de banda de la pluralidad de señales de paso de banda que comprende un intervalo de frecuencia que se correlaciona con una nota la original, de manera que la señal de paso de banda modificada puede contener un intervalo de frecuencia que se correlaciona con una nota la bemol objetivo (excepto el caso, la señal de paso de banda se identifica como una señal de paso de banda de sobretono de una señal de paso de banda de tono fundamental que ha de modificarse según otro objetivo de modificación). Además, las señales de paso de banda de sobretono identificadas (señales de paso de banda que comprenden un intervalo de frecuencia correlacionado con un sobretono de la nota la original) pueden modificarse de manera que la señal de paso de banda de sobretono modificada comprende un intervalo de frecuencia correlacionado con un sobretono de la nota la bemol objetivo.

[0053] Todas las señales de paso de banda 112 de la pluralidad de señales de paso de banda pueden comprender una frecuencia portadora. La frecuencia portadora puede ser una frecuencia característica del intervalo de frecuencia representado por o contenido por una señal de paso de banda, como, por ejemplo, una frecuencia media del intervalo de frecuencia, una frecuencia de corte superior del intervalo de frecuencia, una frecuencia de corte inferior del intervalo de frecuencia o un centro de gravedad del intervalo de frecuencia de la señal de paso de banda. La frecuencia portadora de una señal de paso de banda puede ser diferente de la frecuencia portadora de cada una de las otras señales de paso de banda. Estas frecuencias portadoras pueden ser usadas por el determinador de sobretono 130 para identificar señales de paso de banda de sobretono. Por ejemplo, el determinador de sobretono 130 puede comparar la frecuencia portadora de una señal de paso de banda 112 de la pluralidad de señales de paso de banda con la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada 122. Como un sobretono puede ser aproximadamente un múltiplo de la frecuencia fundamental, puede cumplirse un criterio de sobretono, si la frecuencia portadora de una señal de paso de banda 112 es un múltiplo de la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada 122 (con una tolerancia de frecuencia portadora predefinida, por ejemplo, 100 Hz, 50 Hz, 20 Hz o menos). En otras palabras, un criterio de sobretono puede ser, por ejemplo, que la frecuencia portadora de una señal de paso de banda 112 sea un múltiplo de la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada 122 con una tolerancia de frecuencia portadora predefinida.

[0054] Además o alternativamente, el determinador de sobretono 130 puede comparar un contenido de energía de la señal de paso de banda 112 de la pluralidad de señales de paso de banda con un contenido de energía de la señal de paso de banda de tono fundamental seleccionada 122. En este ejemplo, puede cumplirse un criterio de sobretono, si una relación del contenido de energía de la señal de paso de banda 112 y el contenido de energía de la señal de paso de banda de tono fundamental seleccionada 122 puede estar dentro de un intervalo de tolerancia de energía predefinido. Este criterio de sobretono tiene en cuenta que normalmente los amónicos presentan energía más baja que los tonos fundamentales. El intervalo de tolerancia de energía predefinido puede ser, por ejemplo, de 0,3 a 0,9, 0,5 a 0,8, 0,6 a 0,7 u otro intervalo. Este criterio de sobretono basado en el contenido de energía puede combinarse con el criterio de sobretono basado en la frecuencia portadora mencionado anteriormente.

[0055] Además o alternativamente, el determinador de sobretono 130 puede calcular el valor de correlación que indica una correlación de una envolvente temporal de la señal de paso de banda 112 de la pluralidad de señales de paso de banda con una envolvente temporal de la señal de paso de banda de tono fundamental seleccionada 122. En este caso, puede cumplirse un criterio de sobretono, si el valor de correlación es superior a un umbral de correlación predefinido. Este criterio de sobretono considera el hecho de que un tono fundamental y su armónico comparten una envolvente temporal bastante similar. El umbral de correlación predefinido puede ser, por ejemplo, 0,2, 0,3, 0,4 o más. La criterio de sobretono basado en correlación descrito puede combinarse con el criterio de sobretono basado en frecuencia portadora y/o el criterio de sobretono basado en contenido de energía mencionados anteriormente.

[0056] El determinador de tono fundamental 120 puede seleccionar una señal de paso de banda adicional 112 de la pluralidad de señales de paso de banda sin considerar todas las señales de paso de banda de tono fundamental ya seleccionadas 122 y todas las señales de paso de banda de sobretono ya identificadas 132. En otras palabras, el determinador de tono fundamental 120 puede seleccionar señales de paso de banda de tono fundamental iterativas 122 de un conjunto de señales de paso de banda, el cual contiene señales de paso de banda no señales de paso de banda de tono fundamental ya seleccionadas y señales de paso de banda de sobretono ya identificadas 132. Esto puede hacerse hasta que todas las señales de paso de banda de la pluralidad de señales de paso de banda puedan seleccionarse como una señal de paso de banda de tono fundamental o identificarse como un sobretono de una señal de paso de banda de tono fundamental. Por consiguiente, el determinador de sobretono 130 puede identificar una señal de paso de banda 112 de la pluralidad de señales de paso de banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada adicional sin considerar todas las señales de paso de banda de sobretono ya identificadas y sin considerar todas las señales de paso de banda de tono fundamental ya seleccionadas 122.

[0057] Además, el procesador de señal 140 puede modificar la señal de paso de banda de tono fundamental seleccionada adicional 122 basándose en el objetivo de modificación predefinido adicional e independientemente de todas las demás señales de paso de banda de tono fundamental seleccionadas. En otras palabras, para cada señal de paso de banda de tono fundamental o para algunas de las señales de paso de banda de tono fundamental seleccionadas, pueden definirse diferentes objetivos de modificación. Por ejemplo, los objetivos de modificación pueden definirse mediante una tabla mencionada anteriormente que indica una transición de una nota clave a otra. Como las señales de paso de banda de tono fundamental pueden modificarse independientemente unas de otras, por ejemplo, también pueden modificarse selectivamente sólo los tonos fundamentales y los armónicos de un instrumento específico para cambiar el modo de clave o la sonoridad de este instrumento.

[0058] La señal de paso de banda 112 puede seleccionarse mediante el determinador de tono fundamental 120 basándose en un criterio de energía. Por ejemplo, puede seleccionarse la señal de paso de banda con el contenido de energía más elevado o uno de los más elevados (por ejemplo, superior al 70% o más de las otras señales de paso de banda). En este ejemplo, una señal de paso de banda de tono fundamental ya seleccionada puede excluirse de una nueva selección estableciendo un parámetro de contenido de energía que indica el contenido de energía de la señal de paso de banda de tono fundamental seleccionada igual a cero. Para la selección de la señal de paso de banda 112, el contenido de energía de cada señal de paso de banda (indicado, por ejemplo, por un parámetro de contenido de energía determinado por el determinador de tono fundamental) puede ponderarse (por ejemplo, mediante ponderación A) para enfatizar la selección de señales de paso de banda importantes perceptivamente.

[0059] El procesador de señal 140 puede modificar las señales de paso de banda de tono fundamental seleccionadas 132 y las señales de paso de banda de sobretono asociadas 132 de diversas maneras. Por ejemplo, el procesador de señal 140 puede modificar la señal de paso de banda de tono fundamental seleccionada 122

multiplicando una frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada 122 con un factor de transposición (por ejemplo, dependiendo del cambio de modo de clave) o sumando una frecuencia de transposición a la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada 122. Además, el modificador de señal 140 puede modificar la señal de paso de banda de sobretono identificada 132 multiplicando una frecuencia portadora de la señal de paso de banda identificada 132 con el factor de transposición (por ejemplo, con una tolerancia del 20%, el 10% , el 5%, el 1% o inferior) o sumando un múltiplo de la frecuencia de transposición (por ejemplo, con una tolerancia del 20%, el 10%, el 5%, el 1% o inferior) a la frecuencia portadora de la señal de paso de banda de sobretono identificada 132. En otras palabras, por ejemplo, un cambio de modo de clave puede realizarse multiplicando el tono fundamental y los armónicos asociados por el mismo factor de transposición o sumando una frecuencia de transposición al tono fundamental y un múltiplo de la frecuencia de transposición al sobretono. De este modo, la señal de paso de banda de sobretono identificada 132 es modificada dependiendo del mismo modo que la señal de paso de banda de tono fundamental seleccionada 122.

[0060] La fig. 2 muestra un diagrama de bloques de un aparato 200 para modificar una señal de audio 102 según una realización de la invención. El aparato 200 es similar al aparato mostrado en la fig. 1, pero comprende además un determinador de frecuencia portadora 260 y el procesador de banco de filtros 110 comprende un banco de filtros 212 y un convertidor de señal 214. El banco de filtros 212 está conectado al convertidor de señal 214 y el convertidor de señal 214 está conectado al procesador de señal 140. El determinador de frecuencia portadora opcional 260 está conectado al banco de filtros 212 del procesador de banco de filtros 110 y el procesador de señal 140.

[0061] El banco de filtros 212 puede generar señales de paso de banda basándose en la señal de audio 102 y el convertidor de señal 214 puede convertir las señales de paso de banda generadas a un dominio de sub-banda para obtener la pluralidad de señales de paso de banda proporcionadas al determinador de tono fundamental 120, el determinador de sobretono 130 y el procesador de señal 140. El convertidor de señal 214 puede realizarse, por ejemplo, como una unidad de transformada discreta de Fourier inversa unilateral, de manera que cada señal de paso de banda 112 de la pluralidad de señales de paso de banda puede representar una señal analítica. En este dominio de sub-banda, el determinador de tono fundamental 120 puede seleccionar una de estas señales de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda para obtener la señal de paso de banda de tono fundamental 122. Además, el determinador de sobretono puede identificar una de estas señales de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda.

[0062] Además, el determinador de frecuencia portadora 260 puede determinar una pluralidad de frecuencias portadoras basándose en la señal de audio 102 y el banco de filtros 212 del procesador de banco de filtros 11 puede generar las señales de paso de banda, de manera que cada señal de paso de banda comprende un intervalo de frecuencia que contiene una frecuencia portadora diferente 262 de la pluralidad de frecuencias portadoras para obtener una señal de paso de banda asociada a cada frecuencia portadora 262 de la pluralidad de frecuencias portadoras. En otras palabras, la anchura de banda y las frecuencia medias de las señales de paso de banda generadas por el banco de filtros 212 pueden controlarse mediante el determinador de frecuencia portadora 260. Esto puede hacerse de varios modos, por ejemplo, calculando el centro de gravedad (COG) de la señal de audio 102, como se describió anteriormente.

[0063] Como ya se mencionó anteriormente, las señales de paso de banda 112 pueden modificarse de varios modos. Por ejemplo, el procesador de señal 140 puede generar una señal de modulación de amplitud (AM) y una señal de modulación de frecuencia (FM) para cada señal de paso de banda 112 de la pluralidad de señales de paso de banda. Como cada señal de paso de banda representa una señal analítica en el dominio de sub-banda, el procesador de señal 140 puede generar la señal de modulación de amplitud y la señal de modulación de frecuencia tal como se mencionó anteriormente en relación con el vocodificador de modulación, por ejemplo. Además, el procesador de señal 140 puede modificar la señal de modulación de amplitud o la señal de modulación de frecuencia de la señal de paso de banda de tono fundamental seleccionada 122 basándose en el objetivo de modificación predefinido y puede modificar la señal de modulación de amplitud o la señal de modulación de frecuencia de la señal de paso de banda de sobretono 132 asociada a la señal de paso de banda de tono fundamental seleccionada 122 dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada 122.

[0064] El procesador de banco de filtros 110, el determinador de tono fundamental 120, el determinador de sobretono 130, el procesador de señal 140, el combinador 150 y/o el determinador de frecuencia portadora 260 pueden ser, por ejemplo, unidades de hardware individuales o parte de un procesador de señal digital, un ordenador o microcontrolador así como un programa informático o producto de software configurado para ejecutarse en un procesador de señal digital, ordenador o microcontrolador.

[0065] Algunas realizaciones según la invención se refieren a un procedimiento 300 para modificar una señal de audio según una realización de la invención. El procedimiento 300 puede comprender generar 310 una pluralidad de señales de paso de banda basándose en una señal de audio y seleccionar 320 una señal de paso de banda de la pluralidad de señales de paso de banda para obtener una señal de paso de banda de tono fundamental. Además, el procedimiento 300 puede comprender identificar 330 una señal de paso de banda de la pluralidad de señales de paso de banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada para obtener una señal de paso de banda de sobretono asociada a la señal de paso de banda de tono fundamental seleccionada. Además, la señal de paso de banda de tono fundamental seleccionada es modificada 340 basándose en un objetivo de modificación predefinido y la señal de paso de banda de sobretono identificada asociada a la señal de paso de banda de tono fundamental seleccionada es modificada 350 dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada. Además, el procedimiento 300 puede comprender combinar 360 la pluralidad de señales de paso de banda que contiene la señal de paso de banda de tono fundamental seleccionada modificada y la señal de paso de banda de sobretono identificada modificada para obtener una señal de audio modificada.

[0066] Opcionalmente, el procedimiento 300 puede comprender etapas adicionales que representan las características opcionales del concepto inventivo mencionado anteriormente y mencionado en lo que viene a continuación.

[0067] En lo que viene a continuación, el concepto descrito se ilustra con más detalle mediante un ejemplo para una implementación que usa un vocodificador de modulación, aunque el concepto propuesto también puede usarse de modo más general también para otras implementaciones.

[0068] La mayoría de los instrumentos excitan sonidos armónicos constituidos por una parte de frecuencia fundamental y sus armónicos que son aproximadamente múltiplos enteros de la frecuencia fundamental. Como los intervalos musicales obedecen una escala logarítmica, cada sobretono armónico se asemeja a un intervalo musical diferente con respecto al fundamental (y sus octavas). La tabla de más abajo enumera la correspondencia de números de armónicos e intervalos musicales para los primeros siete armónicos.

[0069] La tabla muestra números de armónicos e intervalos musicales relacionados con respecto al fundamental y sus octavas.

Número de armónicos			Nombre de intervalo
1	2	4	unísono justo (P1)
			segunda menor (m2)
		9	segunda mayor (M2)
			tercera menor (m3)
	5		tercera mayor (M3)
			cuarta justa (P4)
			tritono
	3	6	quinta justa (P5)
			sexta menor (m6)
			sexta mayor (M6)
		7	séptima menor (m7)
			séptima mayor (M7)

[0070] Por lo tanto, en la tarea de la transposición selectiva de contenido musical polifónico, existe una ambigüedad inherente con respecto a la función musical de una componente MODVOC. Si la componente tiene su origen en un tono fundamental tiene que ser transpuesto según la correspondencia de escala deseada, si es dominado por un armónico que ha de atribuirse a un tono fundamental tiene que ser transpuesto junto con este fundamental con el fin de preservar de la mejor manera el timbre original del tono. De esto surge la necesidad de una asignación de cada componente MODVOC (señal de paso de banda) con el fin de seleccionar el factor de transposición más apropiado.

[0071] Para lograr esto, el esquema de procesamiento simple presentado anteriormente fue ampliado por una funcionalidad de bloqueo de armónicos. El bloqueo de armónicos examina todas las componentes del MODVOC antes de la transposición si una componente (señal de paso de banda) ha de atribuirse a un tono fundamental o ha

de considerarse como una entidad independiente. Esto puede realizarse mediante un algoritmo iterativo. El diagrama de flujo de este algoritmo se representa en la figura 5. El algoritmo evalúa 510 relaciones de frecuencia, relaciones de energía y correlaciones cruzadas de envolvente de una componente de prueba t (señal de paso de banda de tono fundamental) con respecto a todas las demás componentes (señales de paso de banda) indexadas por $i \in [0 \dots I - 1]$ con I indicando el número total de componentes (número de señales de paso de banda de la pluralidad de señales de paso de banda). La sucesión de componentes de prueba (señal de paso de banda de tono fundamental) durante la iteración se determina por su energía con ponderación A 520 de manera que el orden de evaluación es en secuencia de energía decreciente. La ponderación A (ANSI, "Ansi standard s1.4-1983", 1983), (ANSI, "Ansi standard s1.42-2001", 2001) se aplica para modelar la importancia perceptiva de cada componente en cuanto a su sonoridad (véase, por ejemplo, el documento de H. Fletcher y W.A. Munson, "Loudness, its definition, measurement and calculation", J. Acoust Soc Amer., vol. 5, págs. 82-108, 1933).

[0072] Una concordancia de frecuencia portadora armónica, una falta de concordancia de frecuencia portadoras armónica, una energía de componentes y/o una correlación de envoltentes de amplitud normalizada con desfase cero pueden examinarse mediante formación de umbrales.

[0073] La concordancia y falta de concordancia de frecuencia pueden definirse según la ecuación 8 con f_t siendo la frecuencia portadora de componente de prueba (frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada) y siendo f_i la componente con índice i (una señal de paso de banda de la pluralidad de señales de paso de banda). Para la concordancia de frecuencia, todos los múltiplos mayores que 1 son armónicos potenciales. Un valor umbral adecuado (umbral de frecuencia portadora) para la falta de concordancia de frecuencia admisible para un armónico potencial es, por ejemplo, 22 Hz.

$$\begin{aligned} match_i &= round\left(\frac{f_i}{f_t}\right) \\ mismatch_i &= |f_i - (match_i \cdot f_t)| \end{aligned} \quad (8)$$

[0074] Puede requerirse que la relación de energía de componentes con ponderación A (ecuación 9) de los armónicos frente al fundamental sea menor que un umbral predefinido que refleja el hecho de que para la inmensa mayoría de los instrumentos los armónicos presentan energía más baja que el tono fundamental. Un valor umbral adecuado (intervalo de tolerancia de energía), por ejemplo, es la relación de 0,6.

$$nrgRatio_i = \frac{nrg_i}{nrg_t} \quad (9)$$

[0075] La correlación cruzada de desfase cero normalizada de la envolvente de la componente de prueba env_t y la envolvente env_i de la componente con índice i está definida por la ecuación 10. Esta medida aprovecha el hecho de que un tono fundamental y sus armónicos comparten una envolvente temporal bastante similar dentro de la longitud de bloque M . Se determinó que un valor umbral adecuado (umbral de correlación) es 0,4 mediante experimentos informales.

$$xcorr_i = \frac{\sum_{m=0}^{M-1} env_i(m) \cdot env_t(m)}{\sqrt{\sum_{m=0}^{M-1} env_i^2(m) \sum_{m=0}^{M-1} env_t^2(m)}} \quad (10)$$

[0076] Después de ser examinados, todas las componentes i que satisfacen 570 todas las condiciones umbrales son etiquetadas 580 como armónicos que han de ser bloqueados con respecto a la componente de prueba y posteriormente son eliminados de la búsqueda. A continuación, la componente de prueba también es excluida de iteraciones adicionales estableciendo 542 su energía en cero. El algoritmo se repite hasta que todas las componentes han sido asignadas, lo cual se indica por ser cero la máxima energía de componentes.

[0077] La figura 4 muestra el esquema de procesamiento mejorado de transposición selectiva mediante el MODVOC que incorpora bloqueo de armónicos. En contraposición a la figura 16, sólo las componentes no bloqueadas entran en la fase de transposición mientras que las componentes bloqueadas son modificadas en una segunda fase por el mismo factor de transposición que se ha aplicado a sus tonos fundamentales atribuidos.

5

[0078] En otras palabras, la fig. 5 muestra un diagrama de flujo del bloqueo de armónicos descrito (procedimiento 500 para modificar una señal de audio). Las componentes que se ajustan a las condiciones de ser armónicos de un tono fundamental de prueba (señal de paso de banda de tono fundamental seleccionada) son etiquetadas iterativamente y eliminadas del espacio de búsqueda. Para esto, cada señal de paso de banda de la pluralidad de señales de paso de banda comprende una frecuencia portadora, un contenido de energía, y una envolvente temporal o la frecuencia portadora, el contenido de energía y/o la envolvente temporal (parámetros de envolvente temporal) se determinan 510 para cada señal de paso de banda de la pluralidad de señales de paso de banda. Además, el contenido de energía (parámetro de contenido de energía) de cada señal de paso de banda es sometido a ponderación A 520. Después, se selecciona 530 una señal de paso de banda (fundamental de prueba f_1) que comprende una energía máxima (parámetro de contenido de energía). Como todas las señales de paso de banda de tono fundamental ya seleccionadas se establecen en cero y todas las señales de paso de banda de sobretono identificadas son excluidas del espacio de búsqueda, la señal de paso de banda de tono fundamental seleccionada puede comprender un parámetro de contenido de energía igual a cero, de manera que los algoritmos iterativos se detienen 540 en este punto. De lo contrario, se comparan 560 la concordancia (o falta de concordancia) de frecuencia, el contenido de energía y/o la correlación cruzada de envolvente temporal de la señal de paso de banda de tono fundamental seleccionada y las señales de paso de banda restantes de la pluralidad de señales de paso de banda. Si se satisface alguna o todas las condiciones (criterios de sobretono) 570, la señal de paso de banda respectiva es identificada 580 como señal de paso de banda de sobretono y pueden generarse datos de bloqueo de armónicos (por ejemplo, almacenando un índice de la señal de paso de banda identificada en una lista de sobretonos) así como la señal de paso de banda de sobretono identificada es eliminada del espacio de búsqueda. Los datos de bloqueo de armónicos pueden guardarse 590 con referencia a la señal de paso de banda de tono fundamental seleccionada asociada. Después de identificar todas las señales de paso de banda de sobretono de la señal de paso de banda de tono fundamental seleccionada, la energía (el parámetro de contenido de energía) de la señal de paso de banda de tono fundamental seleccionada se establece 592 en cero y se selecciona 530 la siguiente señal de paso de banda de tono fundamental que comprende la energía más elevada.

[0079] El procesador de señal puede usar los datos de bloqueo de armónicos para modificar las señales de paso de banda. En la fig. 4 se muestra una posible implementación. En esta implementación, por ejemplo, el procesador de señal comprende un dispositivo de correspondencia MIDI 1600 y un modificador de sobretono 400. El dispositivo de correspondencia MIDI 1600 puede modificar la frecuencia portadora de cada señal de paso de banda de tono fundamental seleccionada según el objetivo de modificación individual (que también puede incluir que no se modifique una señal de paso de banda de tono fundamental). El dispositivo de correspondencia MIDI 1600 puede implementarse, por ejemplo, tal como se muestra y describe en la fig. 16. El modificador de sobretono 400 puede comprender un controlador de modificación de sobretono 410, un multiplicador de sobretono 420 y un proveedor de modificación de sobretono 430. El controlador de modificación de sobretono 410 puede estar conectado al multiplicador de sobretono 420 y el proveedor de modificación de sobretono 430 y el multiplicador de sobretono 420 puede estar conectado al proveedor de modificación de sobretono 430. El multiplicador de sobretono 420 puede multiplicar la frecuencia portadora f de una señal de paso de banda de sobretono identificada con el mismo factor de transposición (con la tolerancia mencionada anteriormente) con el que se multiplica la señal de paso de banda de tono fundamental asociada, y puede proporcionar la frecuencia portadora modificada f' al proveedor de modificación de sobretono 430. El controlador de modificación de sobretono 410 puede activar el proveedor de modificación de sobretono 430 para proporcionar la frecuencia portadora modificada de la señal de paso de banda de sobretono identificada, si el modificador de sobretono 400 identifica la frecuencia portadora como una frecuencia portadora de una señal de paso de banda de sobretono identificada (por ejemplo, basándose en los datos de bloqueo de armónicos). De lo contrario, el proveedor de modificación de sobretono 430 puede proporcionar la salida del dispositivo de correspondencia MIDI 1600. Además, la fig. 4 muestra una implementación del concepto propuesto en un vocodificador, de manera que además de la frecuencia portadora de la señal de paso de banda también se modifica la señal de modulación de frecuencia correspondiente (FM) mediante una multiplicación con una relación de la frecuencia portadora antes de la modificación, y la frecuencia portadora modificada. Alternativamente a una modificación de frecuencia, o además de una modificación de frecuencia, la sonoridad de la señal de audio puede ser modificada selectivamente según la señal de paso de banda. Para esto, puede modificarse la señal de modulación de amplitud (AM) de una señal de paso de banda.

[0080] En otras palabras, la fig. 4 muestra una transposición selectiva mejorada sobre componentes del

vocodificador de modulación (señales de paso de banda) usando bloqueo de armónicos (modificando las señales de paso de banda de sobretono identificadas dependiendo de la señal de paso de banda de tono fundamental asociada) Sólo las frecuencias portadoras no bloqueadas (que pueden entonces ser señales de paso de banda de tono fundamental) son cuantificadas a notas MIDI que se hacen corresponder con notas MIDI correspondientes apropiadas (según el objetivo de modificación individual). Las componentes bloqueadas (señales de paso de banda de sobretono identificadas) pueden ser transpuestas mediante la multiplicación por la relación de la frecuencia portadora original y modificada del tono fundamental atribuido (paso de banda de tono fundamental asociado).

[0081] La fig. 6a muestra un diagrama de bloques de un aparato 600 para modificar una señal de audio según una realización de la invención. El aparato 600 comprende un determinador de forma de envolvente 610, un procesador de banco de filtros 620, un procesador de señal 630, un combinador 640 y un conformador de envolvente 650. El determinador de forma de envolvente 610 está conectado al conformador de envolvente 650, el procesador de banco de filtros 620 está conectado al procesador de señal 630, el procesador de señal 630 está conectado al combinador 640 y el combinador 640 está conectado al conformador de envolvente 650. El determinador de forma de envolvente 610 determina coeficientes de forma de envolvente 612 basándose en una señal de audio en el dominio de la frecuencia 602 que representa una señal de audio de entrada en el dominio del tiempo. Además, el procesador de banco de filtros 620 genera una pluralidad de señales de paso de banda 622 en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia 602. El procesador de señal 630 modifica una señal de paso de banda en el dominio de sub-banda 622 de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en un objetivo de modificación predefinido. Además, el combinador 640 combina al menos un subconjunto de la pluralidad de señales de paso de banda en el dominio de sub-banda (por ejemplo que contiene la señal de paso de banda en el dominio de sub-banda) para obtener una señal de audio en el dominio del tiempo 642. El conformador de envolvente 650 conforma una envolvente de la señal de audio en el dominio del tiempo 642 basándose en los coeficientes de forma de envolvente 612 para obtener una señal de audio conformada 652.

[0082] Alternativamente, el conformador de envolvente 650 puede estar situado entre el procesador de señal 630 y el combinador 640 (el procesador de señal 630 está conectado al conformador de envolvente 650 y el conformador de envolvente 650 está conectado al combinador 640) y puede conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda que contiene la señal de paso de banda en el dominio de sub-banda modificada basándose en los coeficientes de forma de envolvente 612.

[0083] Extrayendo los coeficientes de forma de envolvente 612 antes de que la señal de audio sea procesada selectivamente según la señal de paso de banda y usando los coeficientes de forma de envolvente 612 para conformar la envolvente de la señal de audio después de modificar una o más señales de paso de banda, la coherencia espectral de las señales de paso de banda modificadas de diferente manera puede preservarse con más exactitud. Además, especialmente para señales transitorias, también puede conformarse mediante el conformador de envolvente 650 un ruido de cuantificación extendido a lo largo del tiempo. De este modo, la calidad perceptiva de la señal de audio modificada puede mejorarse significativamente. Además, la señal de audio puede modificarse en tiempo real, ya que puede no ser necesaria una información a priori acerca de toda la señal de audio (por ejemplo, por ejemplo, el título de toda la música polifónica).

[0084] Además, alternativamente, el conformador de envolvente 650 puede estar situado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 está conectado al conformador de envolvente 650 y el conformador de envolvente 650 está conectado al procesador de señal 630) y puede conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en los coeficientes de forma de envolvente 612 antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal 630 para obtener una señal de audio conformada 652.

[0085] Extrayendo los coeficientes de forma de envolvente 612 antes de que la señal de audio sea procesada selectivamente según la señal de paso de banda y usando los coeficientes de forma de envolvente 612 para conformar la envolvente de la pluralidad de señales de paso de banda 622 después de que la pluralidad de señales de paso de banda 622 sea generada por el procesador de banco de filtros 620 en el dominio de sub-banda, puede implementarse un banco de filtros adaptativo, lo cual puede aumentar la coherencia local especialmente para señales transitorias (véase, por ejemplo, el documento de J. Herre y J. D. Johnston, "A continuously signal-adaptive filterbank for high-quality perceptual audio coding", IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk, 1997). En este caso, no es conformada la señal modificada (o la señal de paso de banda modificada), pero la calidad de las señales de paso de banda generadas en cuanto a reproducción de transitorios puede aumentarse antes de la modificación.

[0086] La señal de audio en el dominio de la frecuencia 602 puede proporcionarse, por ejemplo, desde un preprocesador que genera la señal de audio en el dominio de la frecuencia 602 basándose en una señal de audio de entrada en el dominio del tiempo (por ejemplo, mediante una transformada discreta de Fourier) o puede proporcionarse desde una unidad de almacenamiento. Los coeficientes de forma de envolvente 612 determinados por el determinador de forma de envolvente 610 pueden ser, por ejemplo, coeficientes de predicción lineal u otros coeficientes que parametrizan el espectro de la señal de audio en el dominio de la frecuencia 602.

[0087] El procesador de señal 630 puede modificar una, algunas o todas las señales de paso de banda en el dominio de sub-banda 622 de la pluralidad de señales de paso de banda en el dominio de sub-banda. El objetivo de modificación predefinido puede ser diferente, por ejemplo, para todas o para alguna de las señales de paso de banda en el dominio de sub-banda. Por ejemplo, para cambiar un modo de clave de la señal de audio, los objetivos de modificación predefinidos de las señales de paso de banda en el dominio de sub-banda pueden definirse como ya se mencionó en relación con la tabla anterior.

[0088] La señal de audio en el dominio de la frecuencia 602 puede comprender líneas espectrales obtenidas, por ejemplo, mediante la transformada de Fourier. La diferencia entre líneas espectrales de la señal de audio en el dominio de la frecuencia (que también pueden suponerse como señales de paso de banda) y una señal de paso de banda generada por el procesador de banco de filtros 620 puede ser que una línea espectral de la señal de audio en el dominio de la frecuencia 602 representa una anchura de banda más estrecha que una anchura de banda representada por una señal de paso de banda en el dominio de sub-banda 622 generada por el procesador de banco de filtros 620. Por ejemplo, la señal de audio en el dominio de la frecuencia 602 indica un espectro de frecuencia obtenido mediante una transformada discreta de Fourier, el cual es dividido en la pluralidad de señales de paso de banda por el procesador de banco de filtros 620, en el que un número de señales de paso de banda (por ejemplo, 10, 16, 20 o más) de la pluralidad de señales de paso de banda es significativamente inferior a un número de valores espectrales o líneas espectrales del espectro de frecuencia (por ejemplo, 512 o más valores espectrales).

[0089] El determinador de forma de envolvente 610 puede determinar los coeficientes de forma de envolvente basándose en una predicción sobre la frecuencia de la señal de audio en el dominio de la frecuencia 602, lo cual puede realizarse, por ejemplo, como ya se mencionó mediante una determinación de coeficientes de predicción lineal.

[0090] El procesador de banco de filtros 620 puede proporcionar la pluralidad de señales de paso de banda, representando cada señal de paso de banda 622 un intervalo de frecuencia específico de la señal de audio en el dominio de la frecuencia 602. Alternativamente, el procesador de banco de filtros 620 puede comprender un filtro de predicción 710, un sustractor de señal 720 y un banco de filtros 730 para obtener la pluralidad de señales de paso de banda 622 basándose en una señal de audio residual 722 tal como se muestra en la fig. 7. Para esto, el filtro de predicción 710 puede generar una señal de audio de predicción 712 basándose en una señal de audio en el dominio de la frecuencia 602 y los coeficientes de forma de envolvente 612 (por ejemplo, el filtro de predicción lineal). Además, el sustractor de señal 720 puede restar la señal de audio de predicción 712 de la señal de audio en el dominio de la frecuencia 602 para obtener una señal de audio residual 722. Esta señal de audio residual 722 puede ser usada por el banco de filtros 730 para generar señales de paso de banda para obtener la pluralidad de señales de paso de banda.

[0091] Además, el procesador de banco de filtros 620 puede comprender un convertidor de señal opcional. Este convertidor de señal (por ejemplo, un transformador para transformada discreta de Fourier inversa unilateral) puede convertir las señales de paso de banda generadas por el banco de filtros 730 al dominio de sub-banda para obtener la pluralidad de señales de paso de banda 622. Alternativamente, el convertidor de señal también puede ser parte del procesador de señal 630.

[0092] En algunas realizaciones según la invención, una parte de baja frecuencia de la señal de audio de entrada puede excluirse de una posible modificación para evitar una generación de artefactos en la parte de baja frecuencia de la señal de audio modificada. Para esto, un aparato 680 para modificar una señal de audio puede comprender un filtro de paso alto/paso bajo, como se muestra, por ejemplo, en la fig. 6b. El filtro de paso alto/paso bajo 660 filtra con paso alto la señal de audio de entrada en el dominio del tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo, de manera que el determinador de forma de envolvente 610 determina los coeficientes de forma de envolvente 612 basándose en la señal de audio en el dominio de la frecuencia de paso alto 602 y el procesador de banco de filtros 620 genera la pluralidad de señales de paso de banda 622 en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia de paso alto 602. Además, el filtro de paso alto/paso bajo 660 filtra con paso bajo la señal de audio de entrada en el dominio del

- tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo para obtener una señal de audio de paso bajo 662. Además, el aparato 680 comprende un proveedor de señal de banda completa 670 configurado para combinar la señal de audio conformada 652 y la señal de audio de paso bajo 662 para obtener una señal de audio de banda completa. En otras palabras, el filtro de paso alto/paso bajo 660 puede separar la señal de audio de entrada en el dominio del tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo en una señal de audio de paso alto y una señal de audio de paso bajo. La señal de audio de paso alto o una representación en el dominio de la frecuencia de la señal de audio de paso alto pueden proporcionarse al determinador de forma de envolvente 610 y el procesador de banco de filtros 620. Esto depende de si el filtro de paso alto/paso bajo se implementa en el dominio del tiempo seguido por un procesador de señal que genera la señal de audio en el dominio de la frecuencia basándose en la señal de audio de paso alto o el filtro de paso alto/paso bajo se implementa en el dominio de la frecuencia que ya recibe una señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo.
- 15 **[0093]** El filtro de paso alto/paso bajo 660 puede filtrar la señal de audio de entrada en el dominio del tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo, de manera que la señal de audio de paso bajo contiene frecuencias hasta una frecuencia umbral predefinida (por ejemplo, 100 Hz o más). Por consiguiente, la señal de audio de paso alto puede comprender frecuencias hasta la frecuencia umbral predefinida. En otras palabras, las frecuencias mayores que la frecuencia umbral predefinida pueden atenuarse mediante el filtro de paso alto/paso bajo 660 para proporcionar la señal de audio de paso bajo 662 y las frecuencias inferiores a la frecuencia umbral predefinida pueden atenuarse mediante el filtro de paso alto/paso bajo 660 para proporcionar la señal de paso alto.
- 20 **[0094]** Alternativamente, el conformador de envolvente 650 está situado entre el procesador de señal 630 y el combinador 640 tal como se muestra en la fig. 6c. En este caso, el filtro de paso alto/paso bajo 660 proporciona la señal de audio de paso bajo al combinador 640. El combinador 640 combina la pluralidad de señales de paso de banda en el dominio de sub-banda que contiene la señal de paso de banda en el dominio de sub-banda modificada y la señal de audio de paso bajo 662 para obtener una señal de audio en el dominio del tiempo 642. En este caso, el conformador de envolvente 650 puede determinar un conjunto de coeficientes de forma de envolvente de paso de banda basándose en los coeficientes de forma de envolvente 612 (por ejemplo, mediante el convertidor de coeficientes 810) para cada señal de paso de banda en el dominio de sub-banda que corresponde a la señal de paso de banda en el dominio de sub-banda respectiva (por ejemplo, que corresponde a la zona de frecuencia contenida por la señal de paso de banda en el dominio de sub-banda respectiva). Después, por ejemplo, cada muestra de tiempo de una señal de paso de banda en el dominio de sub-banda puede multiplicarse con un coeficiente de forma de envolvente de paso de banda del conjunto correspondiente de coeficientes de forma de envolvente. Por ejemplo, en la implementación de vocodificador mostrada en la fig. 15, el conformador de envolvente 650 puede estar situado entre el multiplicador 1550 y el combinador 1560.
- 30 **[0095]** Más alternativamente, el conformador de envolvente 650 puede estar situado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 está conectado al conformador de envolvente 650 y el conformador de envolvente 650 está conectado al procesador de señal 630) y puede conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en los coeficientes de forma de envolvente 612 antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal 630 para obtener una señal de audio conformada 652.
- 35 **[0096]** En algunas realizaciones según la invención, una parte de baja frecuencia de la señal de audio de entrada puede excluirse de la conformación de envolvente para evitar una generación de artefactos en la parte de baja frecuencia de la señal de audio modificada. Para esto, un aparato 680 para modificar una señal de audio puede comprender un filtro de paso alto/paso bajo, como se muestra, por ejemplo, en la fig. 6d. El filtro de paso alto/paso bajo 660 filtra con paso alto la señal de audio de entrada en el dominio del tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo. Además, el filtro de paso alto/paso bajo 660 filtra con paso bajo la señal de audio de entrada en el dominio del tiempo o la señal de audio en el dominio de la frecuencia que representa la señal de audio de entrada en el dominio del tiempo para obtener una señal de audio de paso bajo 662. El determinador de forma de envolvente 610 determina los coeficientes de forma de envolvente 612 basándose en la señal de audio en el dominio de la frecuencia de paso alto 602 sin considerar la señal de audio de paso bajo 622. El procesador de banco de filtros 620 genera la pluralidad de señales de paso de banda 622 en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia de paso alto 602 y la señal de audio de paso bajo 622. Si se usa un filtro de predicción, como se muestra, por ejemplo, en la fig. 7, sólo la señal de audio en el dominio de la frecuencia de paso alto 602 se proporciona al filtro de predicción y el

5 sustractor de señal para generar una señal de audio residual de paso alto. La señal de audio de paso bajo 622 puede proporcionarse directamente al banco de filtros para generar señales de paso de banda en el dominio de sub-banda. El procesador de señal 630 puede modificar una señal de paso de banda en el dominio de sub-banda que corresponde a la señal de audio en el dominio de la frecuencia de paso alto 602 o la señal de audio de paso bajo 622. Alternativamente, el procesador de señal 630 puede modificar una señal de paso de banda en el dominio de sub-banda que corresponde a la señal de audio en el dominio de la frecuencia de paso alto 602 y una señal de paso de banda en el dominio de sub-banda que corresponde a la señal de audio de paso bajo 622. El combinador 640 puede combinar sólo la señal de paso de banda en el dominio de sub-banda que corresponde a la señal de audio en el dominio de la frecuencia de paso alto 602, de manera que sólo las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio en el dominio de la frecuencia de paso alto 602 (y no las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio de paso bajo 622) pueden ser conformadas por el conformador de envolvente 650.

15 **[0097]** Además, el aparato 680 comprende un proveedor de señal de banda completa 670 configurado para combinar la señal de audio conformada 652 y las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio de paso bajo 662 para obtener una señal de audio de banda completa. Para esto, el procesador de señal 630 puede proporcionar las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio de paso bajo 662 al proveedor de señal de banda completa 670.

20 **[0098]** Alternativamente, el conformador de envolvente 650 está situado entre el procesador de señal 630 y el combinador 640. En este caso, el procesador de señal 630 puede proporcionar las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio de paso bajo 662 al combinador 640. El combinador 640 combina la pluralidad de señales de paso de banda en el dominio de sub-banda (las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio de paso bajo 662 y las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio en el dominio de la frecuencia de paso alto 602) que contiene la señal de paso de banda en el dominio de sub-banda modificada para obtener una señal de audio en el dominio del tiempo 642. En este caso, el conformador de envolvente 650 puede determinar un conjunto de coeficientes de forma de envolvente de paso de banda basándose en los coeficientes de forma de envolvente 612 (por ejemplo, mediante el convertidor de coeficientes 810) para cada señal de paso de banda en el dominio de sub-banda que corresponde a la señal de paso de banda en el dominio de sub-banda respectiva (por ejemplo, que corresponde a la zona de frecuencia contenida por la señal de paso de banda en el dominio de sub-banda respectiva) de las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio en el dominio de la frecuencia de paso alto 602. Después, por ejemplo, cada muestra de tiempo de una señal de paso de banda en el dominio de sub-banda puede multiplicarse con un coeficiente de forma de envolvente de paso de banda del conjunto correspondiente de coeficientes de forma de envolvente. Por ejemplo, en la implementación de vocodificador mostrada en la fig. 15, el conformador de envolvente 650 puede estar situado entre el multiplicador 1550 y el combinador 1560.

40 **[0099]** Además, alternativamente, el conformador de envolvente 650 puede estar situado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 está conectado al conformador de envolvente 650 y el conformador de envolvente 650 está conectado al procesador de señal 630) y puede conformar una envolvente de las señales de paso de banda en el dominio de sub-banda que corresponden a la señal de audio en el dominio de la frecuencia de paso alto 602 basándose en los coeficientes de forma de envolvente 612 antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal 630 para obtener una señal de audio conformada 652.

50 **[0100]** De este modo, una porción de baja frecuencia de la señal de audio de entrada puede ser eximida de la conformación de envolvente. Sin embargo, la porción de baja frecuencia es encaminada al procesamiento restante (por ejemplo, la modificación de una señal de paso de banda en el dominio de sub-banda). Además, un filtro de predicción (por ejemplo, como se muestra en la fig.7) puede ser aplicable sólo por encima de la frecuencia umbral predefinida. Alternativamente, si la separación de paso alto/paso bajo ya está realizada en el lado de análisis, la envolvente de señal de paso alto puede modificarse en el dominio del tiempo mediante un recíproco de los coeficientes de forma de envolvente.

55 **[0101]** Por ejemplo, en aplicaciones para transposición selectiva, la colocación mostrada puede proporcionar resultados equivalentes a una colocación después del procesamiento, ya que la AM no puede modificarse.

[0102] Según un aspecto, el conformador de envolvente 650 puede determinar una relación de energía de un contenido de energía E_{FDAS} de la señal de audio en el dominio de la frecuencia 602 y un contenido de energía E_{RAS}

de la señal de audio residual 722. Basándose en esta relación de energía, el conformador de envolvente 650 puede interrumpir la conformación de la envolvente de la señal de audio en el dominio del tiempo 642, si la relación de energía es inferior a un umbral de energía predefinido PET (0,1, 0,2, 0,5, 0,8, 1, 2 o aún más o menos).

5

$$PET = \frac{E_{FDAS}}{E_{RAS}}$$

[0103] En otras palabras, el empleo de la conformación de envolvente puede activarse o desactivarse adaptativamente a la señal dependiendo de la bondad de la predicción. La bondad de la predicción puede medirse por la ganancia de predicción la cual puede definirse como la relación de energía de la señal (señal de audio en el dominio de la frecuencia) y el error de predicción (señal de audio residual). Si la conformación de la envolvente de la señal de audio en el dominio del tiempo 642 se interrumpe, la señal de audio conformada 652 puede ser igual a la señal de audio en el dominio del tiempo 642 proporcionada por el combinador 640.

[0104] El conformador de envolvente 650 puede implementarse de diversos modos. En la fig. 8 se muestra un ejemplo. El conformador de envolvente 650 puede comprender un convertidor de coeficientes 810 y un multiplicador 820. El convertidor de coeficientes 810 puede convertir los coeficientes de forma de envolvente 612 al dominio del tiempo, de manera que los coeficientes de forma de envolvente convertidos 812 pueden multiplicarse con la señal de audio en el dominio del tiempo 642 para conformar la envolvente temporal de la señal de audio en el dominio del tiempo y para obtener la señal de audio conformada 652. Esto puede hacerse mediante el multiplicador 820. Por ejemplo, un bloque de tiempo de la señal de audio en el dominio del tiempo 642 puede contener 512 (o más) muestras de tiempo y el convertidor de coeficientes 810 puede proporcionar 512 (o más) coeficientes de forma de envolvente convertidos 812 para multiplicar cada muestra de tiempo con un coeficiente de forma de envolvente convertido 812.

[0105] Como ya se mencionó, el aparato 600 puede modificar diferentes señales de paso de banda en el dominio de sub-bandas de manera diferente. Más en general, esto significa que el procesador de señal 630 puede modificar una segunda señal (o señal adicional) de paso de banda en el dominio de sub-banda 622 de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en un segundo objetivo (u objetivo adicional) de modificación predefinido. El ya mencionado o primer objetivo de modificación predefinido y el objetivo adicional (o segundo objetivo) de modificación predefinido pueden ser diferentes.

[0106] En algunas realizaciones, el concepto descrito puede usarse en relación con vocodificadores o vocodificadores de modulación. En este caso, el procesador de señal 630 puede generar una señal de modulación de amplitud (AM) y una señal de modulación de frecuencia (FM) para cada señal de paso de banda en el dominio de sub-banda 622 de la pluralidad de señales de paso de banda en el dominio de sub-banda. Además, el procesador de señal 630 puede modificar la señal de modulación de amplitud o la señal de modulación de frecuencia de la señal de paso de banda en el dominio de sub-banda para que sea modificada basándose en el objetivo de modificación predefinido.

[0107] Además, el aparato 600 puede comprender opcionalmente un determinador de frecuencia portadora como ya se describió para el aparato 200 y se mostró en la fig. 2. El determinador de frecuencia portadora puede determinar una pluralidad de frecuencias portadoras basándose en la señal de audio en el dominio de la frecuencia 602. Estas frecuencias portadoras determinadas pueden ser usadas por el procesador de banco de filtros 620 o en la implementación mostrada en la fig. 7 por el banco de filtros 730 del procesador de banco de filtros 620 para generar señales de paso de banda en el dominio de sub-banda, de manera que cada señal de paso de banda en el dominio de sub-banda comprende un intervalo de frecuencia que contiene una frecuencia portadora diferente de la pluralidad de frecuencias portadoras para obtener una señal de paso de banda en el dominio de sub-banda asociada a cada frecuencia portadora de la pluralidad de frecuencias portadoras. Esto puede hacerse, por ejemplo, determinando los centros de gravedad de la señal de audio en el dominio de la frecuencia como se mencionó anteriormente.

[0108] El determinador de forma de envolvente 610, el procesador de banco de filtros 620, el procesador de señal 630, el combinador 640 y/o el conformador de envolvente 650 pueden ser, por ejemplo, unidades de hardware individuales o parte de un procesador de señal digital, un ordenador o microcontrolador así como un programa informático o producto de software configurado para ejecutarse en un procesador de señal digital, ordenador o microcontrolador.

[0109] Algunas realizaciones según la invención se refieren a una implementación del concepto descrito en un vocodificador de modulación. Para este ejemplo, el concepto se describe con más detalle en lo que viene a continuación. Las características mencionadas también pueden usarse en otras implementaciones o aplicaciones.

5 [0110] Se ha afirmado anteriormente que el procesamiento con MODVOC preserva la coherencia espectral en el área de paso de banda que rodea las ubicaciones de las portadoras. Sin embargo, la coherencia espectral global de banda ancha no se preserva. Para señales cuasi-estacionarias esto puede tener sólo un impacto menor sobre la calidad perceptiva de la señal sintetizada. Si la señal contiene transitorios prominentes como, por ejemplo, ritmos de batería o castañuelas, la preservación de la coherencia global puede mejorar en gran medida la calidad de
10 reproducción de estas señales.

[0111] La preservación de la coherencia global puede mejorarse mediante predicción lineal en el dominio espectral. Algunos procedimientos se utilizan en codecs de audio, por ejemplo mediante la herramienta de conformación de ruido temporal (TNS) (véase, por ejemplo, el documento de J Herre y J. D. Johnston, "Enhancing
15 the performance of perceptual audio coders by using temporal noise shaping (tns)", 101st AES convention, Los Angeles, nº preimpresión 4384, 1996) en la codificación de audio avanzada (AAC) de MPEG 2/4. En el documento de J Herre y J. D. Johnston, "A continuously signal-adaptive filterbank for high-quality perceptual audio coding", IEEE-ASSP Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk, 1997", se muestra la combinación de una transformada de tiempo-frecuencia de alta resolución y predicción espectral que corresponden
20 esencialmente a una transformada adaptativa de señal.

[0112] La figura 9 explica a grandes rasgos la integración del concepto descrito en el esquema de procesamiento con MODVOC. En el análisis, posteriormente a la DFT inicial de la señal de entrada x , los coeficientes de predicción lineal (LPC) de un predictor avanzado a lo largo de la frecuencia que tiene la respuesta de impulso $h(w)$ se
25 deducen, por ejemplo, mediante el procedimiento de autocorrelación minimizando el error de predicción en un sentido de mínimos cuadrados. Posteriormente, el filtro se aplica a los valores espectrales y la señal residual se procesa nuevamente mediante el algoritmo MODVOC. Los coeficientes de filtro, que representan la envolvente global, se llevan a la fase de síntesis. En la síntesis, la envolvente global, deducida por evaluación del filtro de
predicción sobre el círculo de radio unidad $|H(e^{it})|$, se recupera mediante una aplicación multiplicativa de la
30 misma a la señal suma produciendo la señal de salida y tal como se ilustra en la figura 10.

[0113] En otras palabras, las figs. 9 y 10 muestran una implementación del concepto descrito en un vocodificador de modulación. La fig. 9 muestra la parte de analizador de modulación que comprende un preprocesador 910, que realiza, por ejemplo, una transformada discreta de Fourier de una señal de audio en el dominio del tiempo para
35 obtener una señal de audio en el dominio de la frecuencia 602 y proporciona la señal de audio en el dominio de la frecuencia 602 al determinador de forma de envolvente 610, el filtro de predicción 710 (por ejemplo, un filtro LPC $h(w)$), el sustractor de señal 710 y el determinador de frecuencia portadora 920. El sustractor de señal 720 puede proporcionar la señal de audio residual 722 al banco de filtros 730. El determinador de frecuencia portadora 920 puede estimar múltiples frecuencias portadoras de centro de gravedad y proporcionar estas frecuencias portadoras
40 al banco de filtros 730 para controlar las ponderaciones espectrales de paso de banda. El banco de filtros 730 puede proporcionar las señales de paso de banda a un convertidor de señal 930 para realizar una transformada discreta de Fourier inversa unilateral para cada señal de paso de banda para proporcionar la pluralidad de señales de paso de banda en el dominio de sub-banda al procesador de señal. Las componentes del vocodificador de modulación ya están descritas anteriormente con mayor detalle. Además, la fig. 10 muestra la parte de síntesis del vocodificador de
45 modulación. Comprende el combinador 640 y el conformador de envolvente que comprende un convertidor de coeficientes 810 y un multiplicador 820. Los detalles adicionales de los componentes del vocodificador de modulación y el conformador de envolvente ya están explicados anteriormente.

[0114] La fig. 11 muestra un diagrama de flujo del procedimiento 1100 para modificar una señal de audio según una realización de la invención. El procedimiento 1100 comprende determinar 1110 coeficientes de forma de
50 envolvente basándose en una señal de audio en el dominio de la frecuencia que representa una señal de audio de entrada en el dominio del tiempo y generar 1120 una pluralidad de señales de paso de banda en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia. Además, el procedimiento 1100 comprende modificar 1130 una señal de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de
55 banda en el dominio de sub-banda basándose en un objetivo de modificación predefinido. Además, al menos un subconjunto de la pluralidad de señales de paso de banda en el dominio de sub-banda se combina 1140 para obtener una señal de audio en el dominio del tiempo. Además, el procedimiento 110 comprende conformar 1150 una envolvente de la señal de audio en el dominio del tiempo basándose en los coeficientes de forma de envolvente,

conformar 1150 una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda que contiene la señal de paso de banda en el dominio de sub-banda modificada basándose en los coeficientes de forma de envolvente o conformar 1150 una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en los coeficientes de forma de envolvente antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal para obtener una señal de audio conformada.

[0115] Opcionalmente, el procedimiento 1100 puede comprender etapas adicionales que representan las características del concepto descrito mencionado anteriormente.

10 **[0116]** Algunas realizaciones según la invención se refieren a un aparato para modificar una señal de audio que combina las características del aparato mostrado en la fig. 1 o 2 con las características del aparato mostrado en la fig. 6. Aproximadamente, la fig. 12 muestra un diagrama de bloques de un aparato 1200 según una realización de la invención.

15 **[0117]** Partiendo del aparato mostrado en la fig. 1, el aparato 1200 comprende además un determinador de forma de envolvente 610 y un conformador de envolvente 650. A este respecto, la señal de audio puede ser una señal de audio en el dominio de la frecuencia que representa una señal de audio de entrada en el dominio del tiempo, que puede ser usada por el determinador de forma de envolvente para determinar coeficientes de forma de envolvente basándose en la señal de audio en el dominio de la frecuencia. Además, la pluralidad de señales de paso de banda generadas por el banco de filtros pueden generarse en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia. Después de combinar la pluralidad de señales de paso de banda en el dominio de sub-banda que contiene la señal de paso de banda de tono fundamental seleccionada modificada y la señal de paso de banda de sobretono identificada modificada, la señal de audio en el dominio del tiempo obtenida 152, 642 puede proporcionarse al conformador de envolvente 650. El conformador de envolvente 650 puede conformar una
20 envolvente de la señal de audio en el dominio del tiempo basándose en los coeficientes de forma de envolvente 612 para obtener la señal de audio conformada 652.

[0118] Si no, partiendo del aparato mostrado en la fig. 6, el aparato 1200 comprende además un determinador de tono fundamental 120 y un determinador de sobretono 130 tal como se describió en relación con el aparato
30 mostrado en la fig. 1. El determinador de tono fundamental 120 puede seleccionar una señal de paso de banda en el dominio de sub-banda de la pluralidad de señales de paso de banda en el dominio de sub-banda para obtener las señales de paso de banda de tono fundamental 122. Además, el determinador de sobretono 130 puede identificar una señal de paso de banda en el dominio de sub-banda 112 de la pluralidad de señales de paso de banda en el dominio de sub-banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono
35 fundamental seleccionada 122 para obtener una señal de paso de banda de sobretono 132 asociada a la señal de paso de banda de tono fundamental seleccionada 122. El procesador de señal 140, 630 puede modificar la señal de paso de banda de tono fundamental seleccionada basándose en un objetivo de modificación predefinido y modificar una señal de paso de banda de sobretono identificada 132 asociada a la señal de paso de banda de tono fundamental seleccionada 122 dependiendo de la modificación de la señal de paso de banda de tono fundamental
40 seleccionada 122 tal como se mencionó anteriormente.

[0119] De este modo, los sobretonos de los tonos fundamentales y los sobretonos pueden tratarse de manera equivalente durante la modificación de la señal de audio y la coherencia espectral de la pluralidad de señales de paso de banda puede preservarse con mucha exactitud conformando la señal de audio en el dominio del tiempo
45 modificada basándose en los coeficientes de forma de envolvente deducidos antes de la modificación de las señales de paso de banda. De este modo, la calidad perceptiva de la señal de audio modificada puede mejorarse significativamente.

[0120] El aparato 1200 puede realizar características adicionales de los diferentes ejemplos de implementación
50 mencionados anteriormente.

[0121] En lo que viene a continuación la mejora de la calidad perceptiva de las señales de audio modificadas se muestra por los resultados de pruebas de escucha. Para esta prueba de escucha se usó una implementación basada en vocodificador de modulación (MODVOC), pero los resultados también son válidos en general para el
55 concepto propuesto.

[0122] Con el fin de evaluar la calidad de audio subjetiva del vocodificador de modulación (MODVOC) para la aplicación de transposición selectiva de altura tonal y, por otra parte, el mérito de las mejoras propuestas sobre el principio de MODVOC básico, se ha ensamblado y procesado en consecuencia un conjunto de archivos de audio

ejemplares. Además, la tecnología MODVOC se compara con un software de audio disponible comercialmente para manipulación de audio polifónico. El editor Melodyne de Celemony que está en venta desde finales de 2009.

[0123] Puesto que el procesamiento bajo prueba altera drásticamente el contenido de audio de una señal, una comparación directa de la señal original y la procesada – normalmente una parte inherente en las pruebas de escucha estándar – no es aparentemente oportuna en este caso. Con el fin de medir no obstante la calidad de audio subjetiva de un modo significativo, se ha aplicado un procedimiento de prueba de escucha especial: el conjunto de pruebas de escucha tiene su origen en datos MIDI simbólicos que se interpretan en formas de onda usando un expansor MIDI de alta calidad. Este planteamiento permite una comparación directa de archivos de audio alterados de manera similar dentro de la prueba y permite una investigación sobre el efecto del procesamiento selectivo de altura tonal en aislamiento. El procedimiento de generación del conjunto de pruebas se resume en la figura 17. Las señales de prueba originales se preparan en representación de datos MIDI simbólicos (arriba a la izquierda). Una segunda versión de estas señales se genera mediante un procesamiento MIDI simbólico que se asemeja al procesamiento objetivo bajo prueba sobre el audio original interpretado en forma de onda (arriba a la derecha). Posteriormente, estos pares de señales son interpretados por un expansor MIDI de alta calidad en archivos de forma de onda (WAV) (abajo a la izquierda y a la derecha). En la prueba de escucha, se comparan la forma de onda interpretada a partir del archivo MIDI procesado y varias versiones procesadas por el vocodificador de modulación (MODVOC) del archivo MIDI original interpretado (abajo a la izquierda). Además, la salida del MODVOC se compara con la salida del editor Melodyne.

[0124] Aparte de las condiciones procesadas con MODVOC, la prueba incluye una condición obtenida usando el editor Melodyne que es actualmente la única aplicación comercial que se ocupa de este tipo de manipulación de audio y, por lo tanto, puede aceptarse como el estándar de la industria. El editor Melodyne realiza inicialmente un análisis automático del archivo de audio entero. Después de la fase de inicialización, Melodyne sugiere una descomposición del archivo de audio. Por interacción del usuario, esta descomposición puede refinarse más. En aras de una comparación justa con los resultados de procesamiento del MODVOC, la evaluación está basada en el resultado de este análisis inicial automático ya que, aparte del conocimiento a priori de la clave y la altura tonal estándar, la descomposición del MODVOC también es totalmente automática.

[0125] La configuración de la prueba de escucha estaba basada en un ensayo multiestímulo con referencia y anclaje ocultos (MUSHRA) según la recomendación ITU BS.1534 (ITU-R, "Method for the subjective assessment of intermediate sound quality (mushra)", 2001). MUSHRA es una prueba de escucha a ciegas. Sólo se somete a la prueba a una persona a la vez. Para cada pieza, la prueba presenta todas las condiciones de prueba junto con la referencia oculta y un anclaje filtrado con paso bajo oculto para el oyente de una manera alineada con respecto al tiempo. La referencia oculta y el anclaje oculto se incluyen con el fin de comprobar la fiabilidad de los oyentes. El cambio entre las condiciones mientras se escucha está permitido y así lo está el configurar un bucle sobre particiones de la pieza seleccionadas arbitrariamente tal como se sugiere en la recomendación BS.1116-1 (ITU-R, "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems", 1994-1997) y también es aplicable a pruebas MUSHRA. No existe límite del número de repeticiones que los sujetos de prueba podrían escuchar antes de valorar la pieza y pasar a la siguiente pieza de prueba, permitiendo así una comparación muy estrecha y un examen exhaustivo de las diferentes condiciones. La calidad perceptiva de las piezas se valora sobre una escala comprendida entre "excelente" (100 puntos) pasando por "bueno" y "aceptable" hasta "pobre" (0 puntos). La secuencia de piezas de prueba se ordena aleatoriamente y, por otra parte, el orden de las condiciones de cada pieza también es aleatorio.

[0126] Las ocho piezas de prueba proceden del proyecto MUTOPIA (<http://www.mutopiaproject.org/>), que proporciona partituras gratuitas para uso público. Los pasajes adecuados que tienen una duración aproximada de 20 segundos como máximo han sido extraídos de diversas piezas de música clásica, que contienen tanto instrumentos individuales (por ejemplo, G, E) como partes de orquesta completa densa (por ejemplo, F). También se incluyen en el conjunto de pruebas melodías de solos instrumentales dominantes acompañadas por otros instrumentos (por ejemplo, C). Además de las partes tonales cuasi-estacionarias a corto plazo, en varias piezas también están contenidos elementos de percusión (comienzos de punteo de guitarra en C y piano en G) que plantean un reto especial sobre la respuesta transitoria del sistema bajo prueba. La siguiente tabla enumera todas las piezas del conjunto.

nombre	descripción	instrumentos	modo de clave
A	Concierto para violín, J. S. Bach, BWV1041	Orquesta	La menor
B	Eine kleine Nachtmusik, W. A. Mozart, KV525 Mv1	Cuarteto de cuerda	Sol mayor
C	Berceuse, G. Fauré, Op56	Flauta y guitarra	Mi mayor
D	Nocturno, F. Strauss, Op7	Trompa y piano	Re bemol mayor
E	Vals, F. Carulli, Op241 N°1	Guitarra	Do mayor
F	Ein Musikalischer Spass, W. A. Mozart, KY522 Mv1	Trompas, violín, viola, violonchelo	Fa mayor
G	Ode an die Freude, L. V. Beethoven	Piano	Sol mayor
H	Trío de piano, L. V. Beethoven, Op11 Mv3	Clarinete, violonchelo y piano	Si bemol mayor

[0127] El procesamiento MIDI para obtener las señales transpuestas originales se ha efectuado en Sonar8 fabricado por Cakewalk. La interpretación de formas de onda de alta calidad se ha realizado usando Band-stand de Native Instruments en la versión de biblioteca de sonidos 1.0.1 R3. El procesamiento con MODVOC fue evaluado en tres combinaciones diferentes con las dos etapas de procesamiento de mejora siendo bloqueo de armónicos y conformación de envolvente. Para comparación con el editor Melodyne, se utilizó la versión 1.0.11. Todas las condiciones se enumeran en la siguiente tabla.

condición	nombre	descripción
1	* preference	original transpuesta MIDI
2	* 3k5Hz_reference	original filtrada con paso bajo de 3,5 kHz (anclaje)
3	* MODVOC	MODVOC
4	* MODVOC_hann	MODVOC con bloqueo de armónicos
5	MODVOC_harm_cs	MODVOC con bloqueo de armónicos y conformación de envolvente
6	* dna	Modo totalmente automático con editor Melodyne (DNA)

10

[0128] Las pruebas de escucha subjetiva se llevaron a cabo en un laboratorio de escucha aislado acústicamente que está diseñado para permitir pruebas de escucha de alta calidad en un entorno similar a un salón "ideal". Los oyentes estaban equipados con auriculares electrostáticos STAX que eran manejados desde una interfaz de sonido Edirol USB conectada a un Apple MAC mini. El software de prueba de escucha era wavswitch de Fraunhofer IIS, operado en modo MUSHRA, que proporciona una GUI (interfaz gráfica de usuario) sencilla para apoyar al oyente en la realización de la prueba. Los oyentes puede cambiar entre la referencia (1) y las diferentes condiciones (2-7) durante la retransmisión. Cada oyente puede decidir individualmente cuánto tiempo escuchar cada pieza y la condición. Durante el cambio real, la retransmisión de sonido se silencia. En la GUI, las barras verticales visualizan la calificación atribuida a cada condición. Se eligieron oyentes experimentados familiarizados con la codificación de audio pero que también tienen una experiencia musical con el fin de obtener, por una parte, un juicio educado sobre artefactos típicos del procesamiento de señal como pre-ecos y post-ecos o dispersión de transitorios y, por otra parte, parámetros musicales tales como la altura tonal espectral, la melodía y el timbre. Además, se pidió a los oyentes que proporcionaran sus observaciones e impresiones informales.

25 [0129] Quince sujetos en total contribuyeron al resultado de la prueba, mientras que un oyente tuvo que ser examinado a posteriori debido a no identificar obviamente con éxito el original oculto (calificándolo con 64 puntos).

[0130] La figura 18 resume los resultados de la prueba de escucha. La calidad perceptiva para las piezas procesadas mediante transposición selectiva de altura tonal va de aceptable a buena. El anclaje inferior fue valorado entre pobre y malo de manera que la distancia desde las piezas procesadas y el anclaje asciende a aproximadamente 40 puntos MUSHRA.

[0131] Las puntuaciones absolutas proporcionan información que cuantifica la calidad perceptiva de cada pieza (en cada una de las condiciones de prueba) y de ese modo valoran implícitamente la diferencia de calidad entre las piezas del conjunto de pruebas, pero son inadecuadas para comparar las diferentes condiciones dentro de la prueba de escucha ya que las calificaciones de estas condiciones no son independientes. Para una comparación directa de las condiciones que tienen su origen en los diferentes esquemas de procesamiento de transposición selectiva, en lo que viene a continuación se consideran diferencias de puntuación.

[0132] La figura 19 representa el resultado basándose en diferencias de puntuación de las variantes de MODVOC mejorado (condiciones 4 y 5) con respecto a los resultados del MODVOC corriente (condición 3). Aquí, todas las variantes de MODVOC mejorado puntúan considerablemente mejor que el procesamiento con MODVOC corriente (todas las puntuaciones están claramente situadas por encima de cero). Existe relevancia en el sentido de confianza del 95% para todas las piezas y condiciones excepto la aplicación de bloqueo de armónicos únicamente en la pieza A y C.

[0133] La figura 20 muestra las puntuaciones de prueba como diferencias de puntuación con respecto a la condición 6 (editor Melodyne). Para la pieza C, el MODVOC en la condición 5 puntúa significativamente mejor que el editor Melodyne mientras que la condición 4, si bien es cierto que es ligeramente positiva, y la condición 3 son inconcluyentes en un sentido de confianza del 95% (los intervalos de confianza se superponen con 0). Para las piezas B (condición 2), F, G (condición 5) tampoco puede extraerse ninguna conclusión significativa, pero también puede observarse una tendencia a un mejor rendimiento del MODVOC para la pieza C en la condición 4 y la pieza F en las condiciones 4 y 5. En todos los demás casos el MODVOC puntúa significativamente peor que el editor Melodyne.

[0134] La puntuación refleja un juicio de calidad global que comprende aspectos como los artefactos sonoros innaturales como la degradación de transitorios por pre-ecos y post-ecos, exactitud de la altura tonal, corrección de la melodía y preservación del timbre. Con el fin de interpretar los resultados con más detalle, se pidió a los oyentes que anotaran sus observaciones informales junto con la anotación de la puntuación real. A partir de estas observaciones puede concluirse que la preservación del timbre y la ausencia de artefactos sonoros innaturales estaban representadas en la puntuación global en un mayor grado que, por ejemplo, la bondad de la preservación de la melodía. Por otra parte, si una cierta melodía es desconocida por el oyente parece ser que las personas de prueba no eran capaces de memorizar la melodía de referencia con poca antelación durante la prueba y, por lo tanto, estaban indecisos acerca de la melodía verdadera. Esto puede ser una explicación de la calificación global más alta de las piezas procesadas por el editor Melodyne, que tienen una fidelidad más alta con respecto a la preservación del timbre, especialmente de sonidos que tienen su origen en instrumentos individuales. Sin embargo, esto se consigue a costa de que se produzcan accidentalmente errores graves en la melodía que pueden ocurrir presumiblemente debido a una clasificación errónea. El MODVOC es más robusto en ese sentido ya que no confía predominantemente en técnicas de clasificación basadas en características.

[0135] Algunas realizaciones según la invención se refieren a un vocodificador de modulación mejorado para transposición selectiva de altura tonal. Se ha introducido el concepto del vocodificador de modulación (MODVOC) y se ha señalado su capacidad general para realizar una transposición selectiva sobre contenido musical polifónico. Esto hace que resulten posibles aplicaciones que están dirigidas a cambiar el modo de clave de muestras de música PCM pregrabadas. Se proponen dos técnicas de mejora para la transposición selectiva de altura tonal mediante el MODVOC. El rendimiento de la aplicación de transposición selectiva y el mérito de estas técnicas se evalúa por los resultados obtenidos de una metodología de pruebas de escucha diseñada especialmente que es capaz de controlar cambios extremos en cuanto a la altura tonal con respecto a los estímulos auditivos originales. Los resultados de esta evaluación de calidad perceptiva subjetiva se presentan para piezas que han sido convertidas entre el modo de clave menor y mayor por el MODVAC y, además, por el primer software disponible comercialmente que también es capaz de ocuparse de esta tarea.

[0136] Merece la pena observar que aunque el editor Melodyne realiza internamente un análisis automático de todo el archivo de audio permitiendo cualquier manipulación, el MODVAC funciona sobre la base de bloque por bloque, permitiendo así potencialmente un funcionamiento en tiempo real.

[0137] Se han propuesto técnicas de mejora para el vocodificador de modulación (MODVOC) para transposición selectiva de altura tonal. A partir de los resultados de las pruebas de escucha para señales de prueba interpretadas a partir de MIDI puede concluirse que la calidad perceptiva del MODVOC corriente se mejora de hecho mediante bloqueo de armónicos y conformación de envolvente. Sobre todas las piezas, puede esperarse un aumento de hasta 10 puntos MUSHRA. Una parte principal de la mejora es producto del bloqueo de armónicos.

[0138] Por otra parte, la comparación del MODVOC con un software disponible comercialmente (el editor Melodyne) reveló que el nivel de calidad general que puede alcanzarse en la transposición selectiva de altura tonal, en este momento, puede situarse entre "aceptable" y "bueno". El MODVOC es más robusto respecto a la interpretación errónea de la melodía ya que esencialmente no confía principalmente en decisiones de clasificación.

[0139] A diferencia del análisis de pasadas múltiples realizado por el editor Melodyne sobre todo el archivo de

audio antes de la manipulación, el MODVOC está basado únicamente en un procesamiento de bloques en una sola pasada que permite potencialmente escenarios de funcionamiento en flujo continuo o en tiempo real.

[0140] Aunque algunos aspectos del concepto descrito se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa del procedimiento o una característica de una etapa del procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa del procedimiento también representan una descripción de un bloque o pieza o característica correspondiente de un aparato correspondiente.

[0141] La señal de audio codificada inventiva puede almacenarse en un medio de almacenamiento digital o puede transmitirse por un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión cableada tal como Internet.

[0142] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo un disquete, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tienen señales de control legibles electrónicamente almacenadas en los mismos, que cooperan (o son capaces de cooperar) con un sistema informático programable de manera que se realice el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

[0143] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de manera que se realice uno de los procedimientos descritos en este documento.

[0144] Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un soporte legible por una máquina.

[0145] Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en este documento, almacenado en un soporte legible por una máquina.

[0146] En otras palabras, una realización del procedimiento inventivo es, por ejemplo, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en este documento, cuando el programa informático se ejecuta en un ordenador.

[0147] Una realización adicional del procedimiento inventivo es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los procedimientos descritos en este documento.

[0148] Una realización adicional del procedimiento inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representa el programa informático para realizar uno de los procedimientos descritos en este documento. El flujo de datos o la secuencia de señales puede configurarse, por ejemplo, para transferirse a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

[0149] Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en este documento.

[0150] Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los procedimientos descritos en este documento.

[0151] En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo una matriz de puertas programable in situ) para realizar algunas o todas las funcionalidades de los procedimientos descritos en este documento. En algunas realizaciones, una matriz de puertas programable in situ puede cooperar con un microprocesador con el fin de realizar uno de los procedimientos descritos en este documento. Generalmente, los procedimientos se realizan con preferencia mediante algún aparato de hardware.

[0152] Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que para otros expertos en la materia resultarán evidentes modificaciones y variaciones de las disposiciones y los detalles descritos en este documento. Por lo tanto, su propósito sólo ha de estar limitado por el alcance de las reivindicaciones de patente inminente y no por los detalles específicos presentados a título de
5 descripción y explicación de las realizaciones de este documento.

REIVINDICACIONES

1. Aparato (100, 200) para modificar una señal de audio (102), que comprende:
 - 5 un procesador de banco de filtros (110) configurado para generar una pluralidad de señales de paso de banda (112) basándose en una señal de audio (102);
un determinador de tono fundamental (120) configurado para seleccionar una señal de paso de banda (112) de la pluralidad de señales de paso de banda para obtener una señal de paso de banda de tono fundamental (122);
 - 10 un determinador de sobretono (130) configurado para identificar una señal de paso de banda (112) de la pluralidad de señales de paso de banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono fundamental (122) para obtener una señal de paso de banda de sobretono (132) asociada a la señal de paso de banda de tono fundamental seleccionada (122);
 - 15 un procesador de señal (140) configurado para modificar la señal de paso de banda de tono fundamental (122) basándose en un objetivo de modificación predefinido y configurado para modificar una señal de paso de banda de sobretono identificada (132) asociada a la señal de paso de banda de tono fundamental seleccionada (122) dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada (122);
 - 20 en el que el procesador de señal (140) está configurado para generar una señal modulada en amplitud (AM) y una señal modulada en frecuencia (FM) para cada señal de paso de banda (112) de la pluralidad de señales de paso de banda, en el que el procesador de señal (140) está configurado para modificar la señal modulada en frecuencia (FM) de la señal de paso de banda de tono fundamental seleccionada (122) basándose en el objetivo de modificación
 - 25 predefinido, y en el que el procesador de señal (140) está configurado para modificar la señal modulada en frecuencia (FM) de la señal de paso de banda de sobretono identificada (132) asociada a la señal de paso de banda de tono fundamental seleccionada (122) dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada (122); y
 - 30 un combinador (150) configurado para combinar la señal de paso de banda de tono fundamental modificada (122), la señal de paso de banda de sobretono modificada (132) y las señales de paso de banda no seleccionadas de la pluralidad de señales de paso de banda para obtener una señal de audio modificada (152).
2. Aparato según la reivindicación 1, en el que cada señal de paso de banda (112) de la pluralidad de
35 señales de paso de banda comprende una frecuencia portadora, en el que el determinador de sobretono (130) está configurado para comparar la frecuencia portadora de una señal de paso de banda (112) de la pluralidad de señales de paso de banda con la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada (122), en el que se cumple un criterio de sobretono si la frecuencia portadora del paso de banda (112) es un múltiplo de la frecuencia portadora de la señal de paso de banda de tono fundamental seleccionada (122) con una tolerancia
40 de frecuencia portadora predefinida.
3. Aparato según la reivindicación 1 o 2, en el que el determinador de sobretono (130) está configurado para comparar un contenido de energía de una señal de paso de banda de la pluralidad de señales de paso de banda con un contenido de energía de la señal de paso de banda de tono fundamental seleccionada (122), en el que se
45 cumple un criterio de sobretono si una relación del contenido de energía de la señal de paso de banda (112) y el contenido de energía de la señal de paso de banda de tono fundamental seleccionada (122) está dentro de un intervalo de tolerancia de energía predefinido.
4. Aparato según una de las reivindicaciones 1 a 3, en el que el determinador de sobretono (130) está
50 configurado para calcular un valor de correlación que indica una correlación de una envolvente temporal de una señal de paso de banda (112) de la pluralidad de señales de paso de banda con una envolvente temporal de la señal de paso de banda de tono fundamental seleccionada (122), en el que se cumple un criterio de sobretono si el valor de correlación es superior a un umbral de correlación predefinido.
- 55 5. Aparato según una de las reivindicaciones 1 a 4, en el que el determinador de tono fundamental (120) está configurado para seleccionar una señal de paso de banda adicional (112) de la pluralidad de señales de paso de banda sin considerar todas las señales de paso de banda de tono fundamental ya seleccionadas (122) y todas las señales de paso de banda de sobretono ya identificadas (132) para obtener una señal de paso de banda de tono fundamental adicional (122).

6. Aparato según la reivindicación 5, en el que el determinador de sobretono (130) está configurado para identificar una señal de paso de banda (112) de la pluralidad de señales de paso de banda que cumple un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada adicional (122) sin considerar todas las señales de paso de banda de sobretono ya identificadas (132) para obtener una señal de paso de banda de sobretono (132) asociada a la señal de paso de banda de tono fundamental seleccionada adicional (122).
7. Aparato según la reivindicación 5 o 6, en el que el procesador de señal (140) está configurado para modificar la señal de paso de banda de tono fundamental seleccionada adicional (122) basándose en un objetivo de modificación predefinido adicional.
8. Aparato según una de las reivindicaciones 1 a 7, en el que el determinador de tono fundamental (120) está configurado para seleccionar la señal de paso de banda (112) basándose en un criterio de energía.
9. Aparato según una de las reivindicaciones 1 a 8, en el que el determinador de tono fundamental (120) está configurado para determinar un contenido de energía con ponderación A de cada señal de paso de banda (112) de la pluralidad de señales de paso de banda y configurado para seleccionar una señal de paso de banda (112) que comprende el contenido de energía con ponderación A más elevado para obtener la señal de paso de banda de tono fundamental (122).
10. Aparato según una de las reivindicaciones 1 a 9, que comprende un determinador de frecuencia portadora (260), en el que el procesador de banco de filtros (110) comprende un banco de filtros (212) y un convertidor de señal (214), en el que el banco de filtros (212) está configurado para generar señales de paso de banda basándose en la señal de audio (102), en el que el convertidor de señal (214) está configurado para convertir las señales de paso de banda generadas a un dominio de sub-banda para obtener la pluralidad de señales de paso de banda, en el que el determinador de frecuencia portadora (260) está configurado para determinar una pluralidad de frecuencias portadoras basándose en la señal de audio (102), en el que el banco de filtros (212) del procesador de banco de filtros (110) está configurado para generar las señales de paso de banda, de manera que cada señal de paso de banda comprende un intervalo de frecuencia que contiene una frecuencia portadora diferente de la pluralidad de frecuencias portadoras para obtener una señal de paso de banda asociada a cada frecuencia portadora de la pluralidad de frecuencias portadoras.
11. Aparato según una de las reivindicaciones 1 a 10, que comprende un determinador de forma de envolvente y un conformador de envolvente, en el que el determinador de forma de envolvente está configurado para determinar coeficientes de forma de envolvente basándose en la señal de audio (102), en el que la señal de audio (102) es una señal de audio en el dominio de la frecuencia que representa una señal de audio de entrada en el dominio del tiempo, en el que el procesador de banco de filtros (110) está configurado para generar la pluralidad de señales de paso de banda en un dominio de sub-banda basándose en la señal de audio en el dominio de la frecuencia, en el que el combinador está configurado para combinar al menos un subconjunto de la pluralidad de señales de paso de banda para obtener la señal de audio modificada que representa una señal de audio en el dominio del tiempo, en el que el conformador de envolvente está configurado para conformar una envolvente de la señal de audio en el dominio del tiempo basándose en los coeficientes de forma de envolvente, para conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda que contiene la señal de paso de banda en el dominio de sub-banda modificada basándose en los coeficientes de forma de envolvente o para conformar una envolvente de la pluralidad de señales de paso de banda en el dominio de sub-banda basándose en los coeficientes de forma de envolvente antes de que una señal de paso de banda en el dominio de sub-banda sea modificada por el procesador de señal para obtener una señal de audio conformada.
12. Procedimiento (300) para modificar una señal de audio, que comprende:
generar (310) una pluralidad de señales de paso de banda basándose en una señal de audio;
seleccionar (320) una señal de paso de banda de la pluralidad de señales de paso de banda para obtener una señal de paso de banda de tono fundamental;
identificar (330) una señal de paso de banda de la pluralidad de señales de paso de banda que cumpla un criterio de sobretono respecto a la señal de paso de banda de tono fundamental seleccionada para obtener una señal de paso de banda de sobretono asociada a la señal de paso de banda de tono fundamental seleccionada;

modificar (340) la señal de paso de banda de tono fundamental seleccionada basándose en un objetivo de modificación predefinido generando una señal modulada en amplitud (AM) y una señal modulada en frecuencia (FM) para cada señal de paso de banda (112) de la pluralidad de señales de paso de banda y modificando la señal modulada en frecuencia (FM) de la señal de paso de banda de tono fundamental seleccionada (122) basándose en el objetivo de modificación predefinido,

modificar (350) una señal de paso de banda de sobretono identificada asociada a la señal de paso de banda de tono fundamental seleccionada dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada modificando la señal modulada en frecuencia (FM) de la señal de paso de banda de sobretono identificada (132) asociada a la señal de paso de banda de tono fundamental seleccionada (122) dependiendo de la modificación de la señal de paso de banda de tono fundamental seleccionada (122); y

combinar (360) la señal de paso de banda de tono fundamental modificada (122), la señal de paso de banda de sobretono modificada (132) y las señales de paso de banda no seleccionadas de la pluralidad de señales de paso de banda para obtener una señal de audio modificada.

13. Programa informático con un código de programa para realizar el procedimiento según la reivindicación 12, en el que el programa informático se ejecuta en un procesador de señal digital, un ordenador o un microcontrolador.

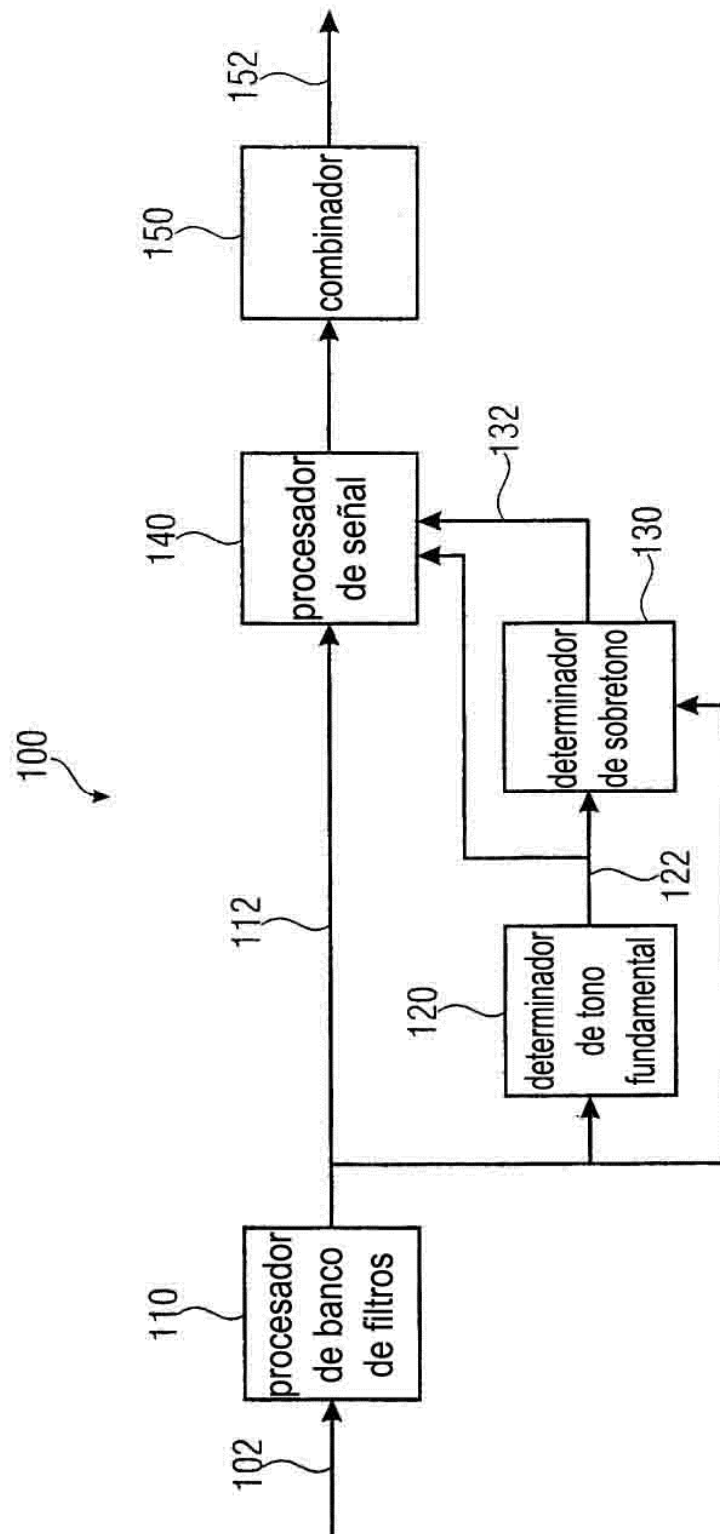


FIGURA 1

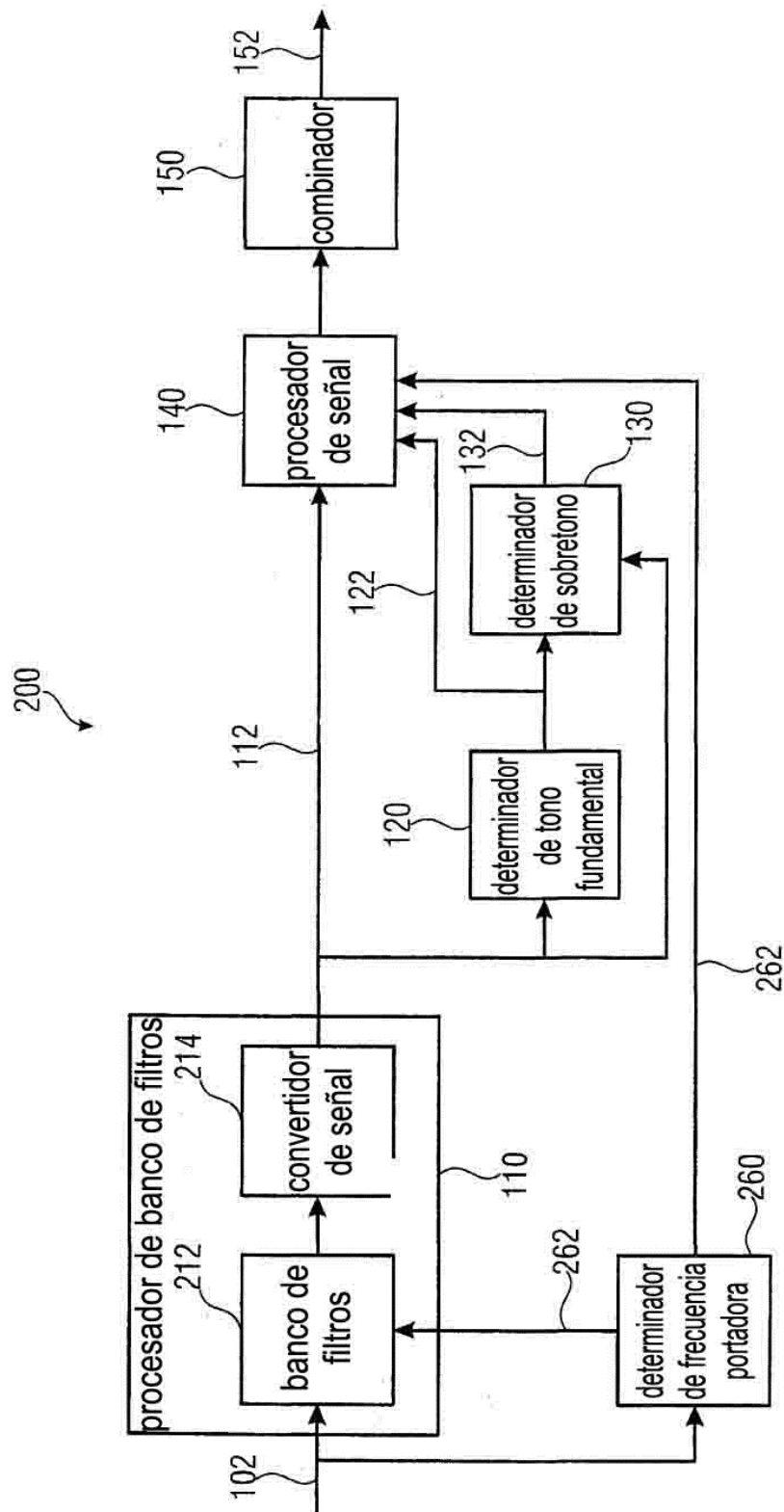


FIGURA 2

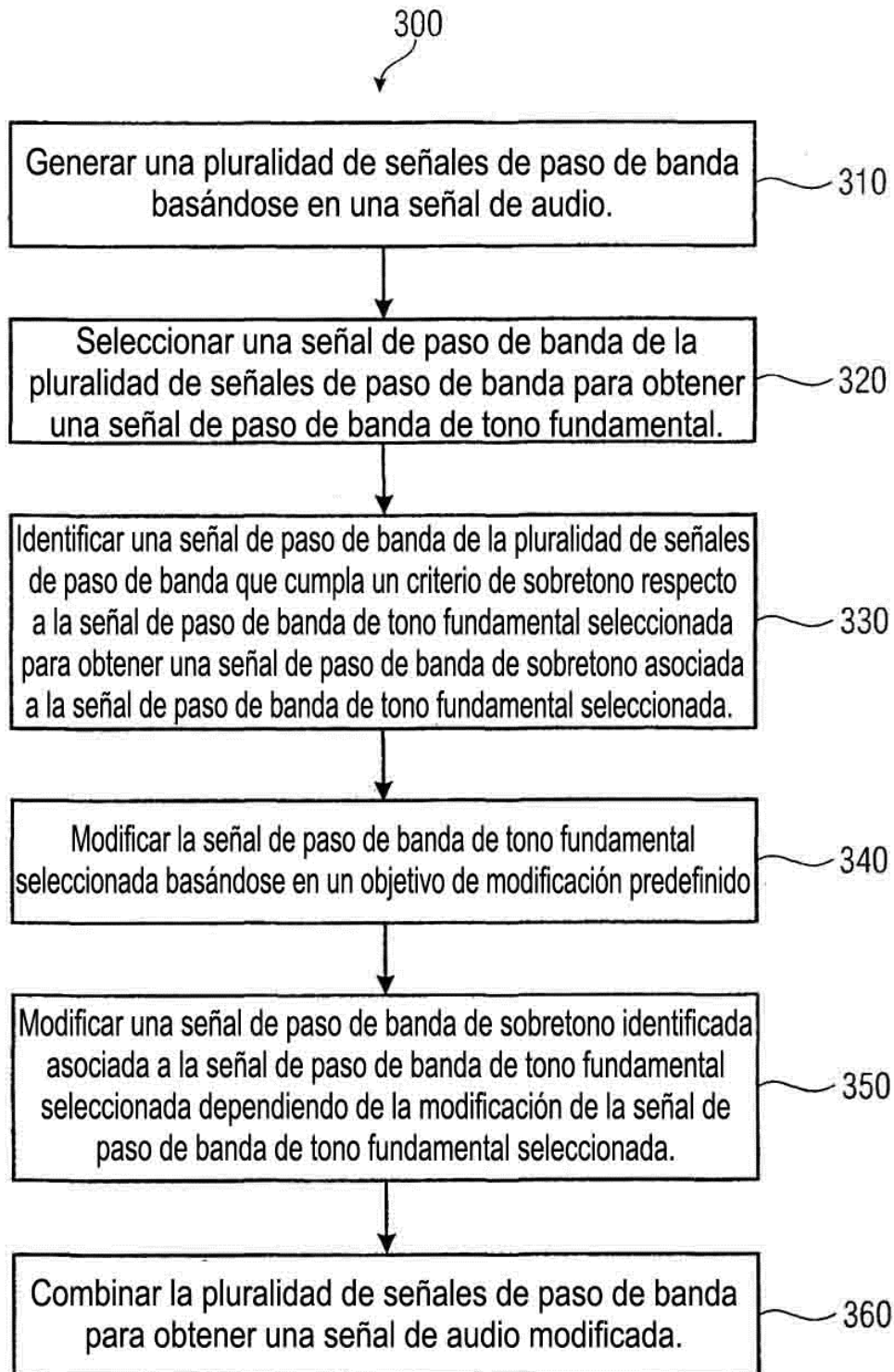


FIGURA 3

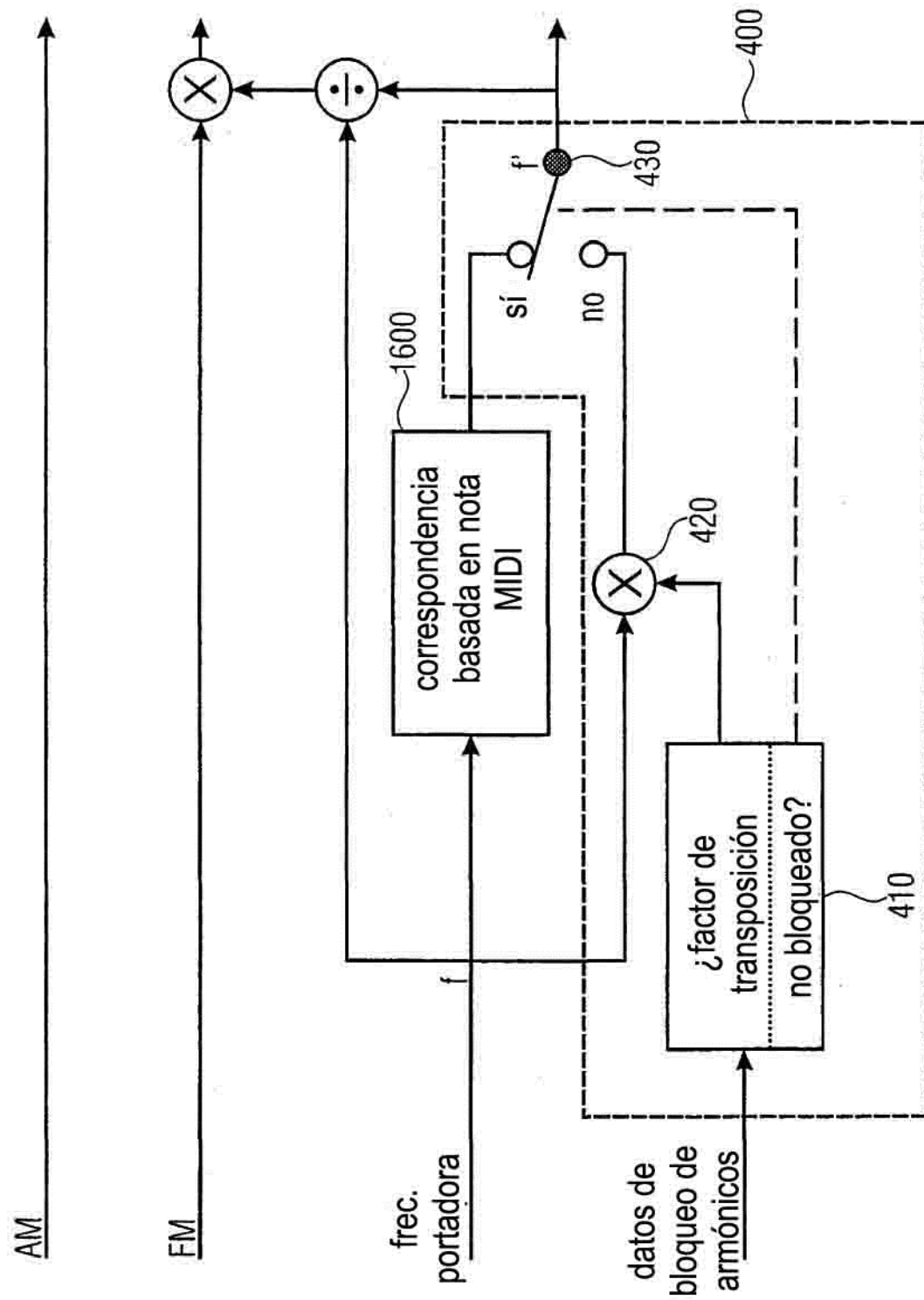


FIGURA 4

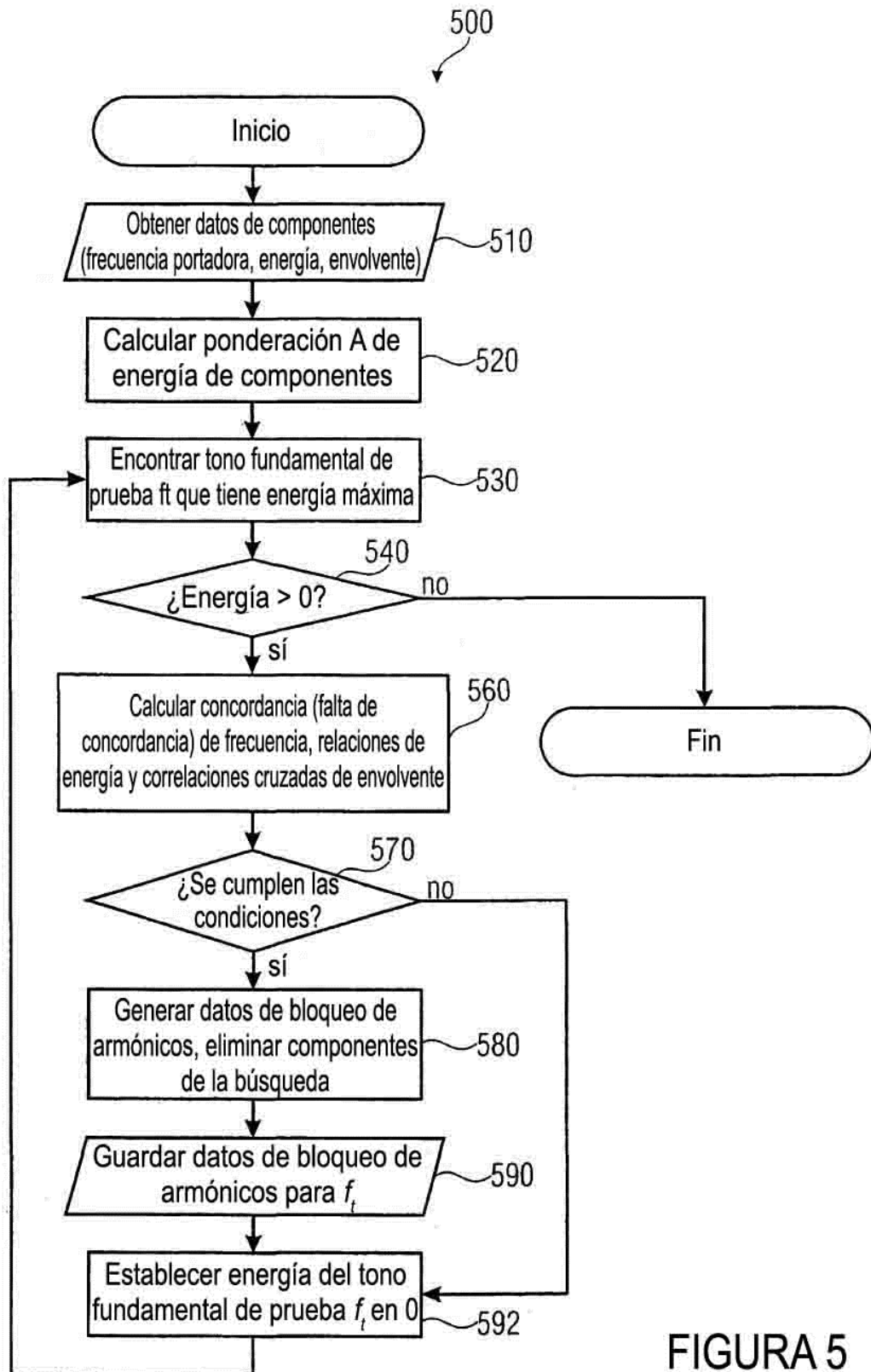


FIGURA 5

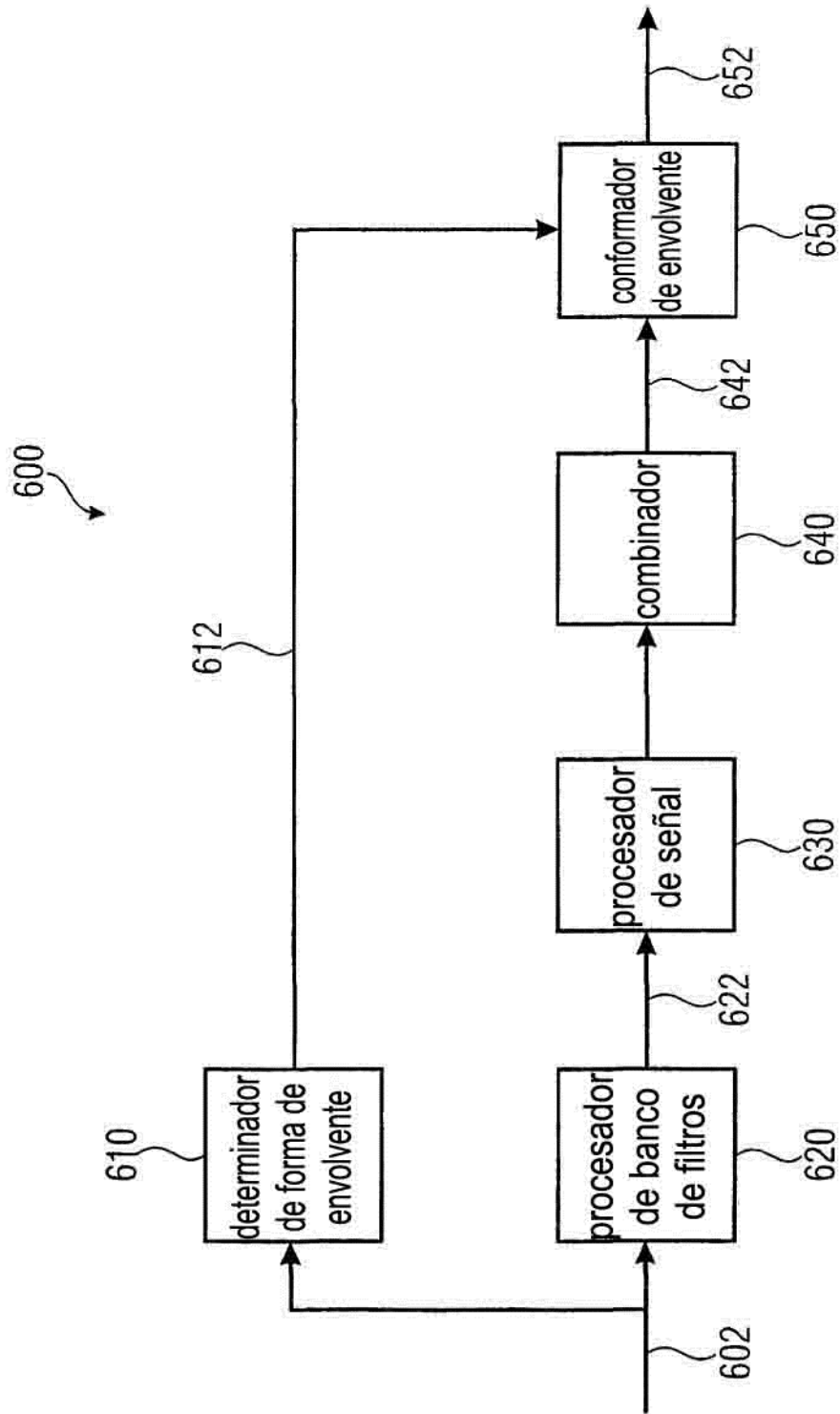


FIGURA 6A

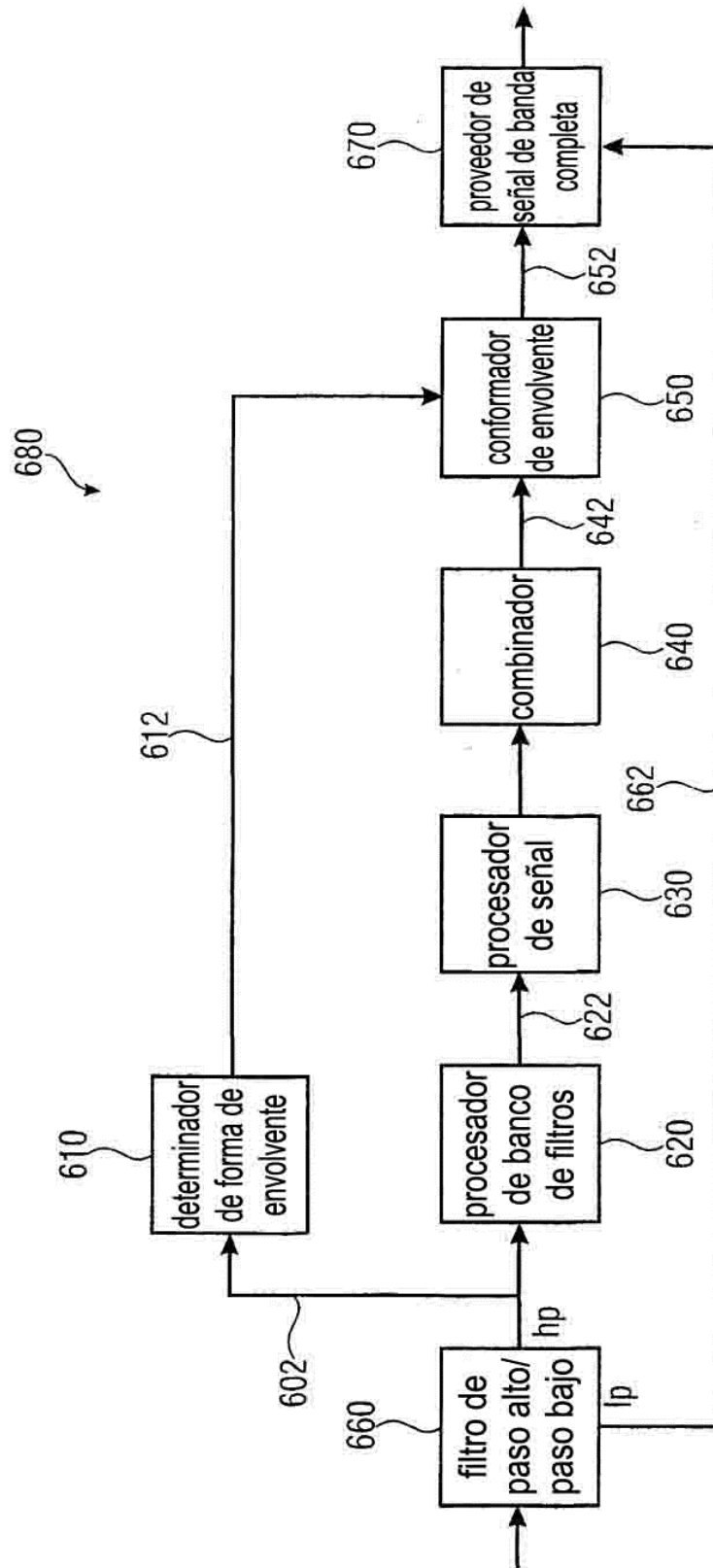


FIGURA 6B

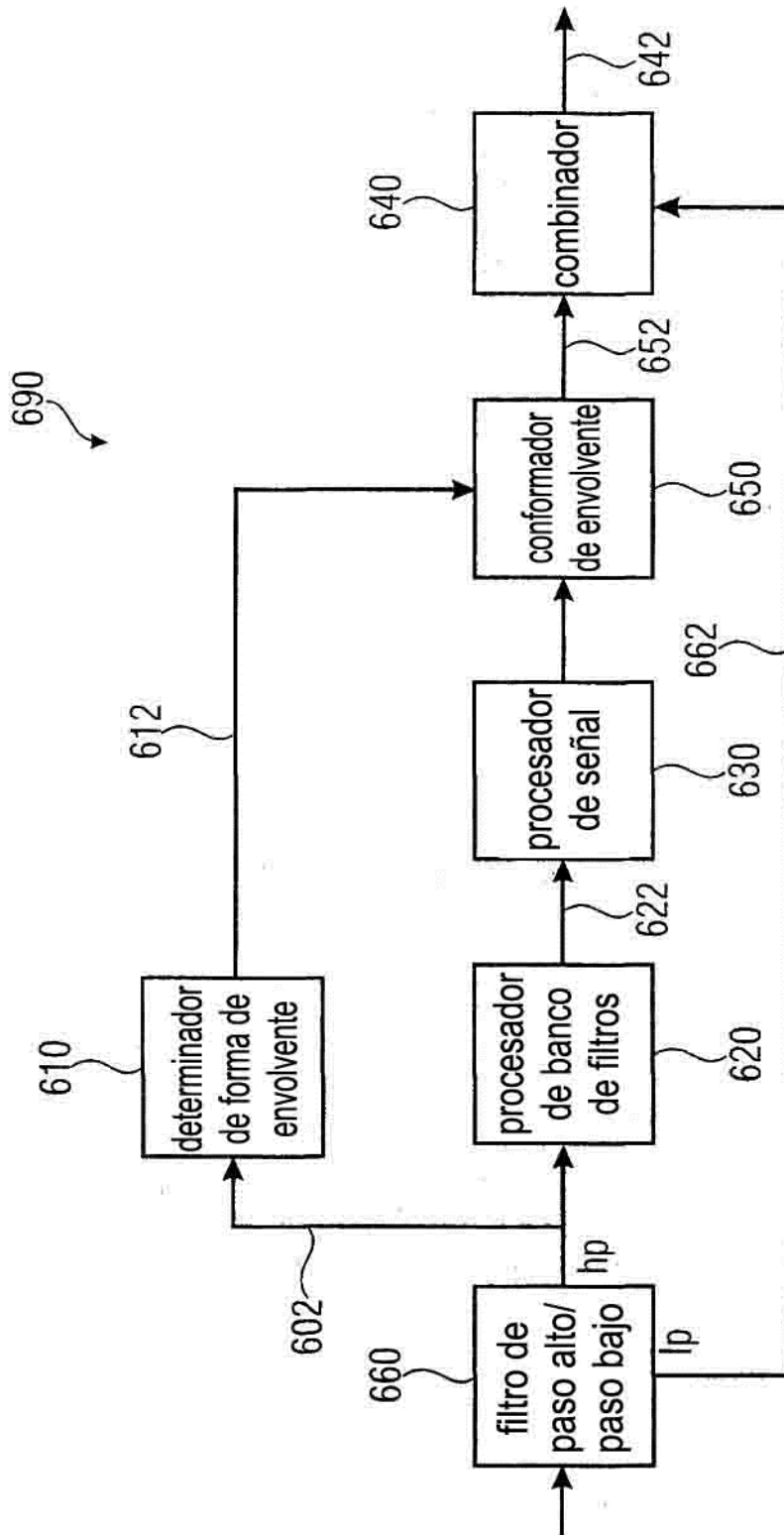


FIGURA 6C

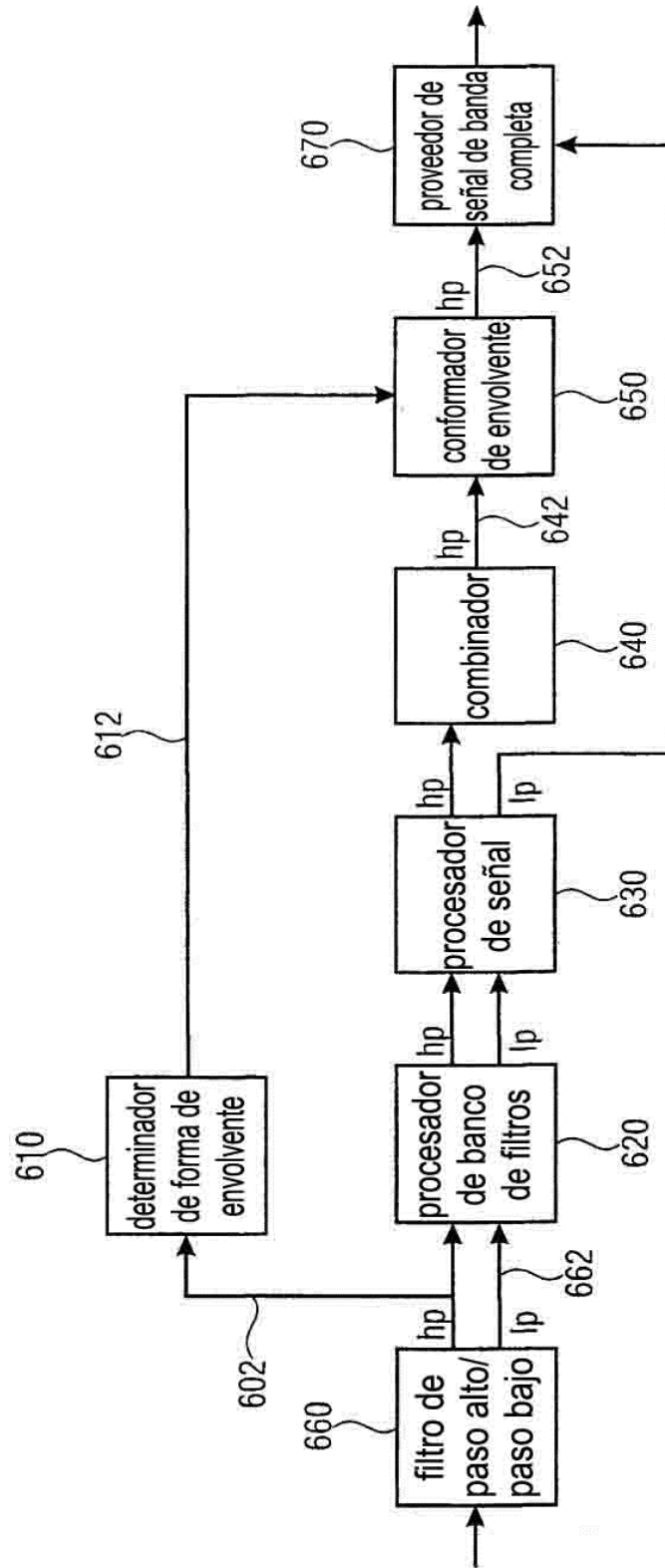


FIGURA 6D

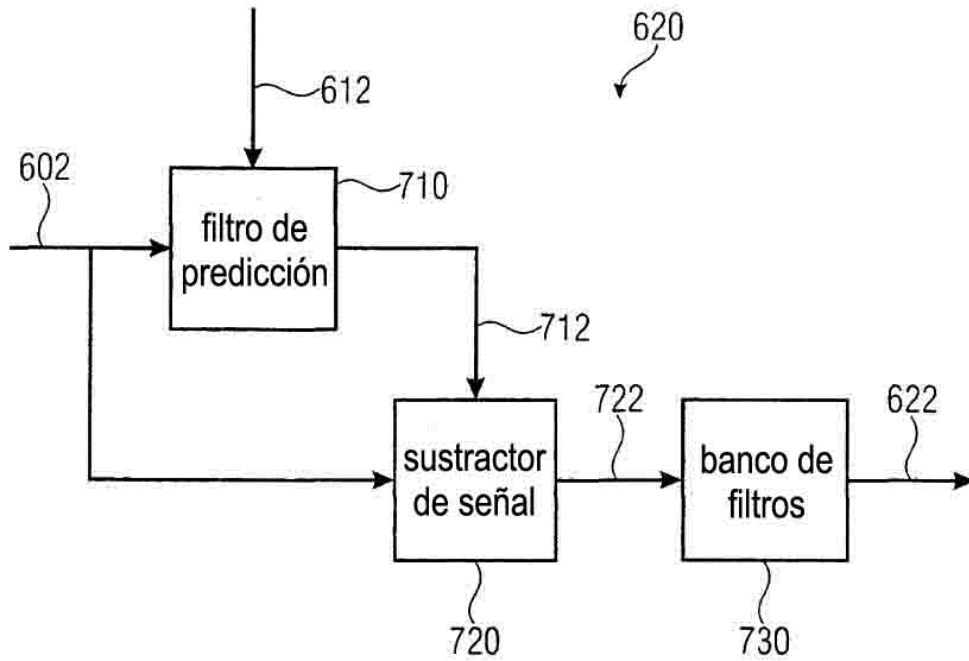


FIGURA 7

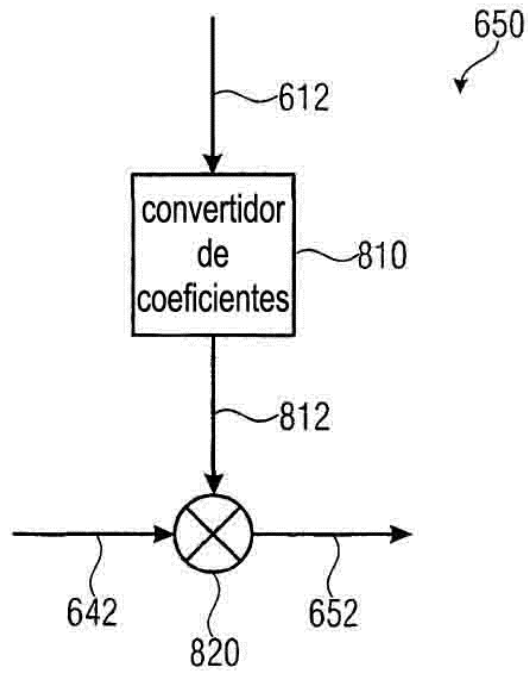


FIGURA 8

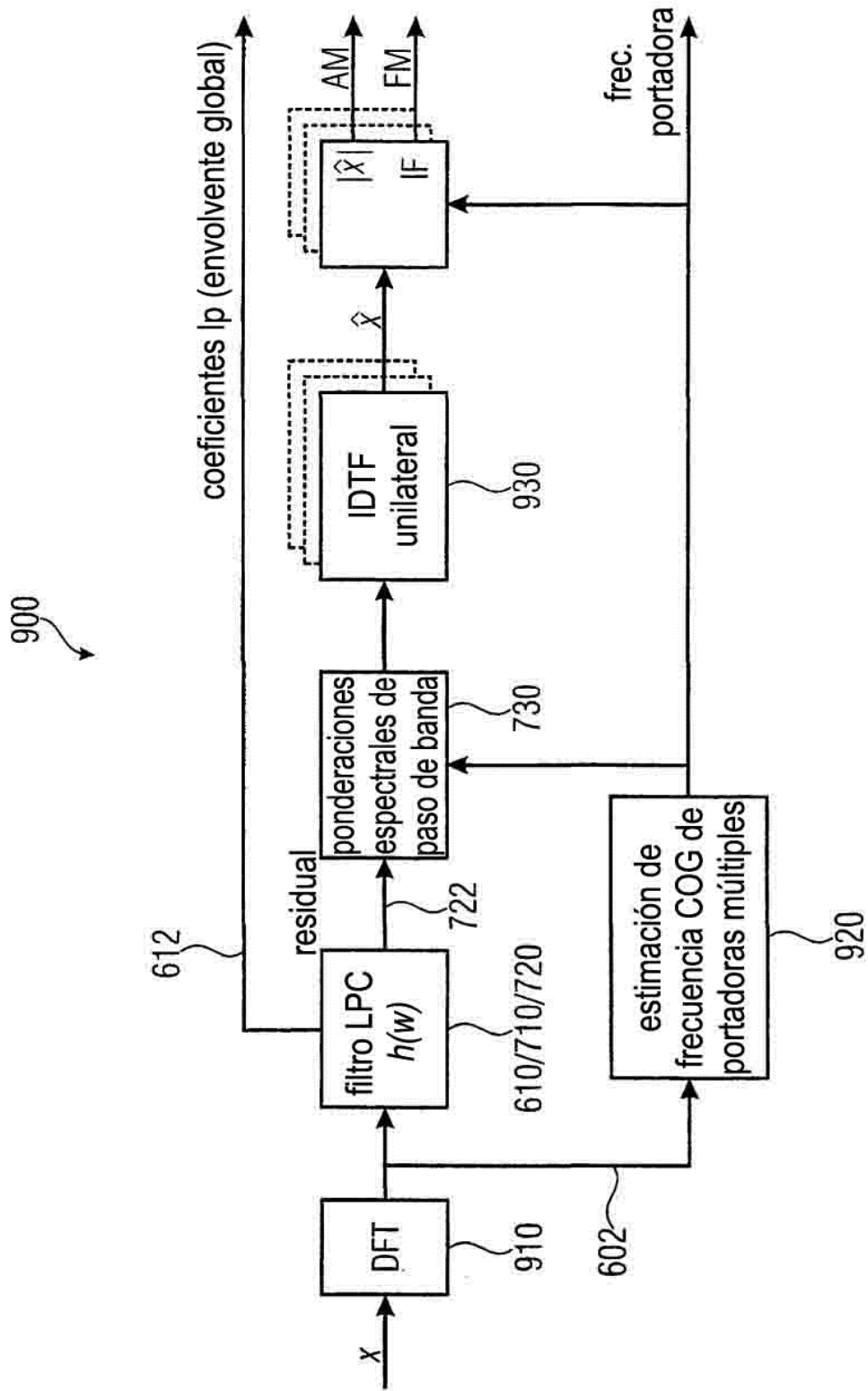


FIGURA 9

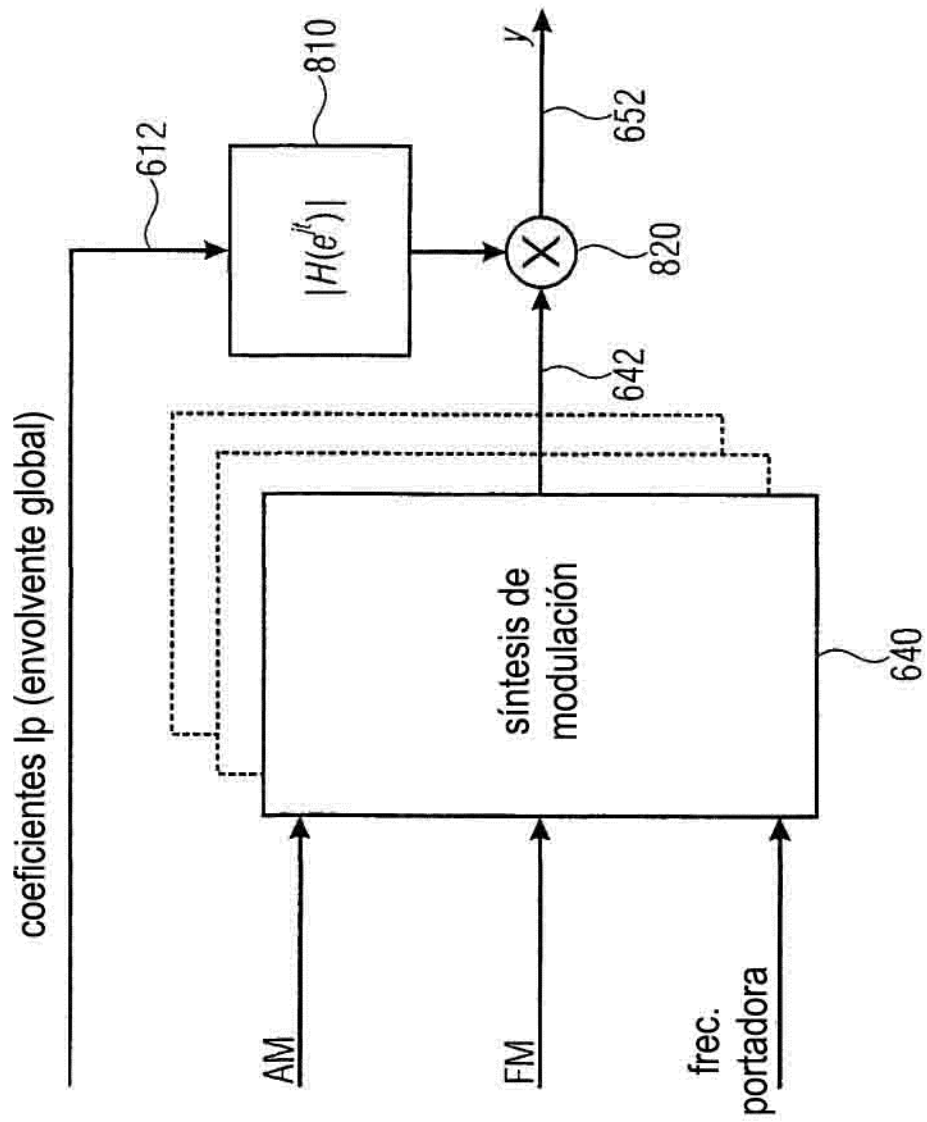


FIGURA 10

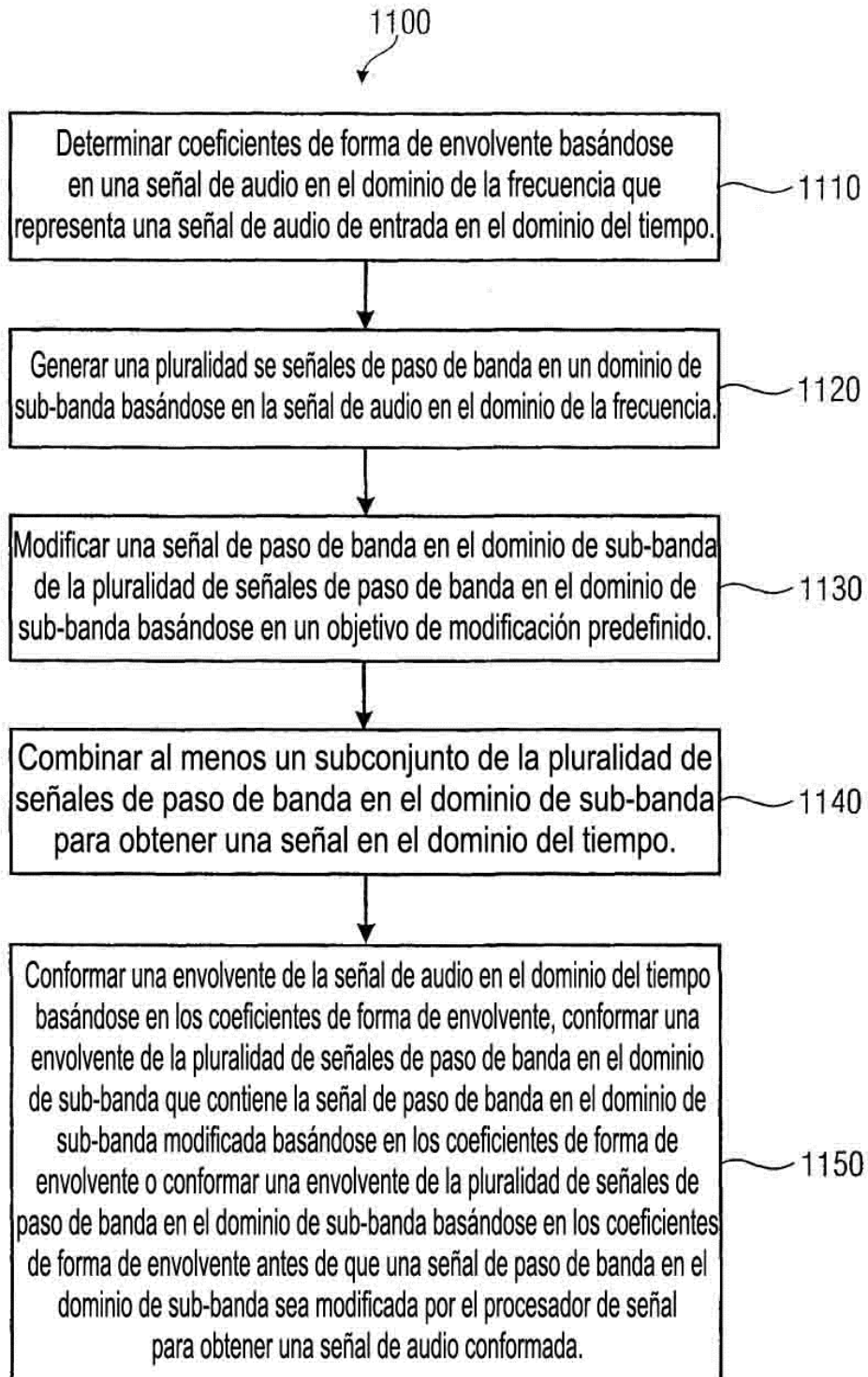


FIGURA 11

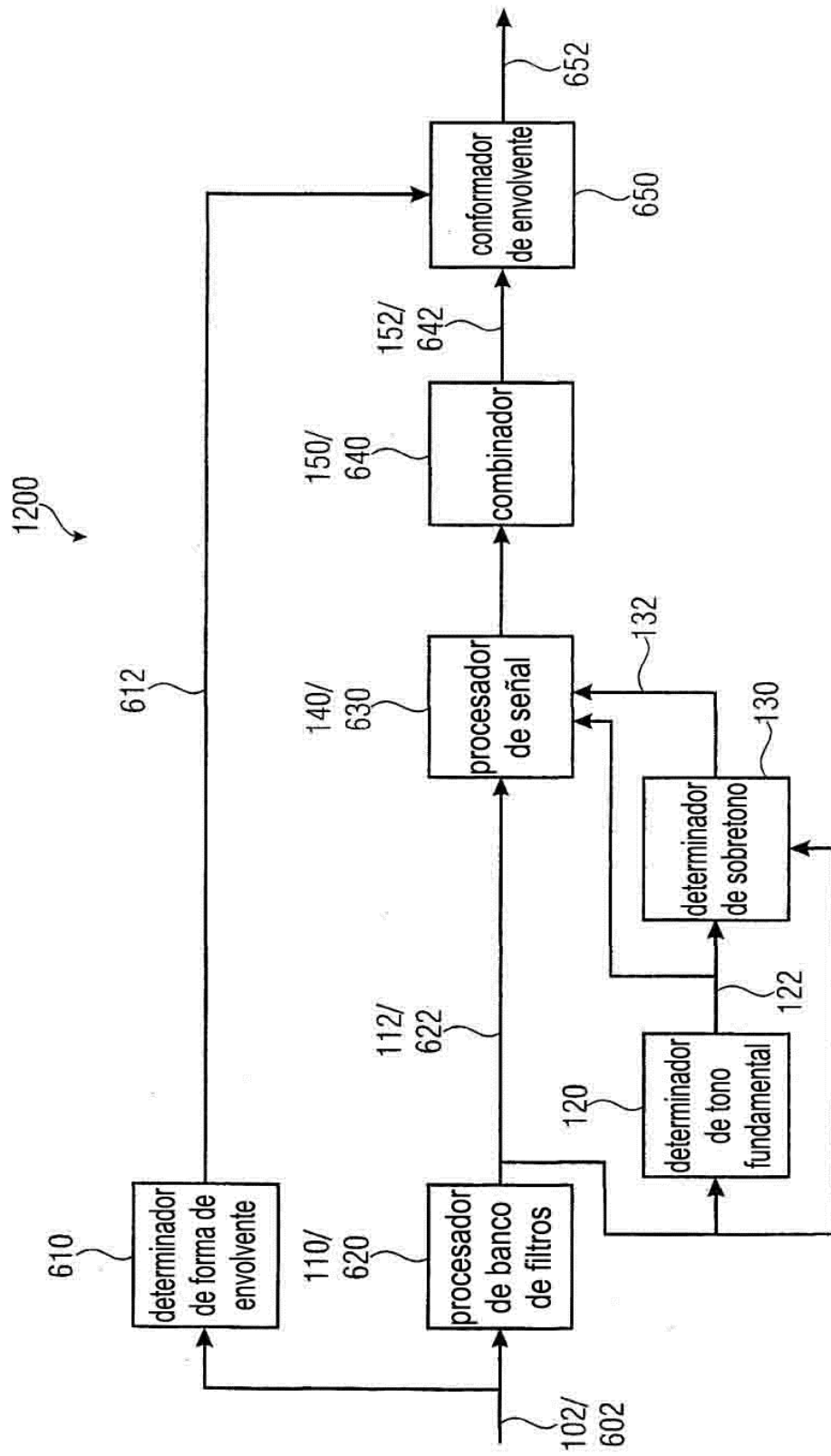


FIGURA 12

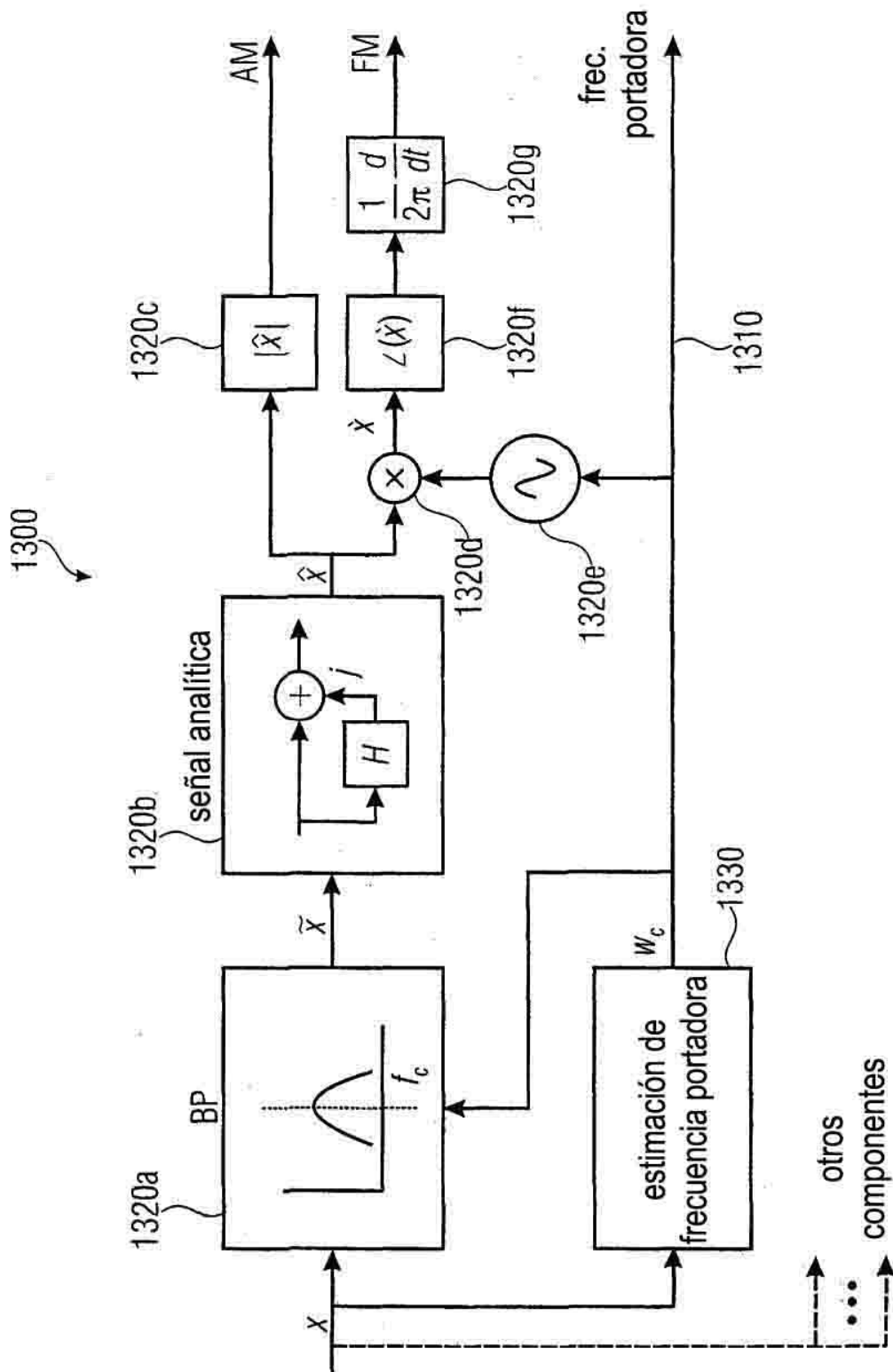


FIGURA 13

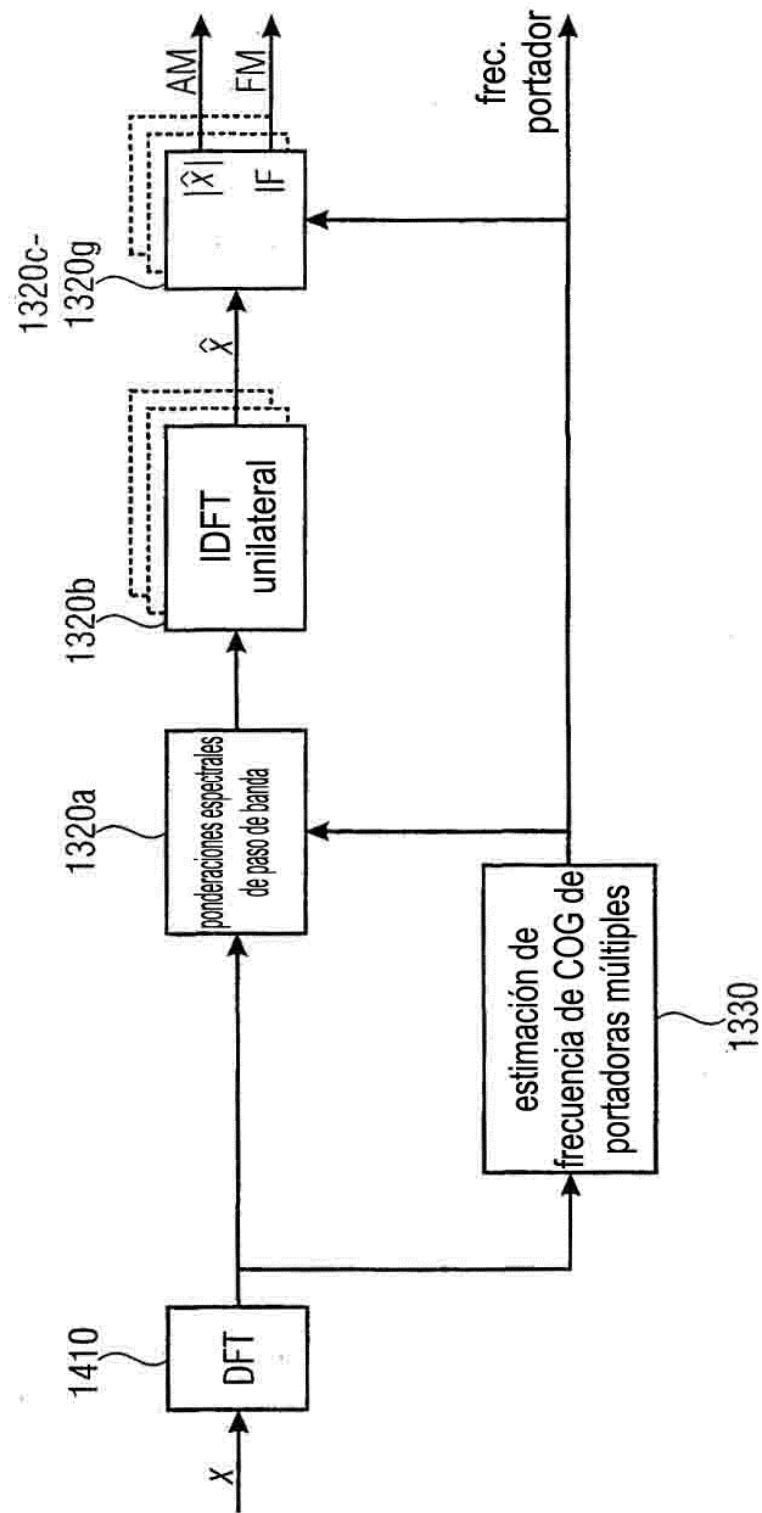


FIGURA 14

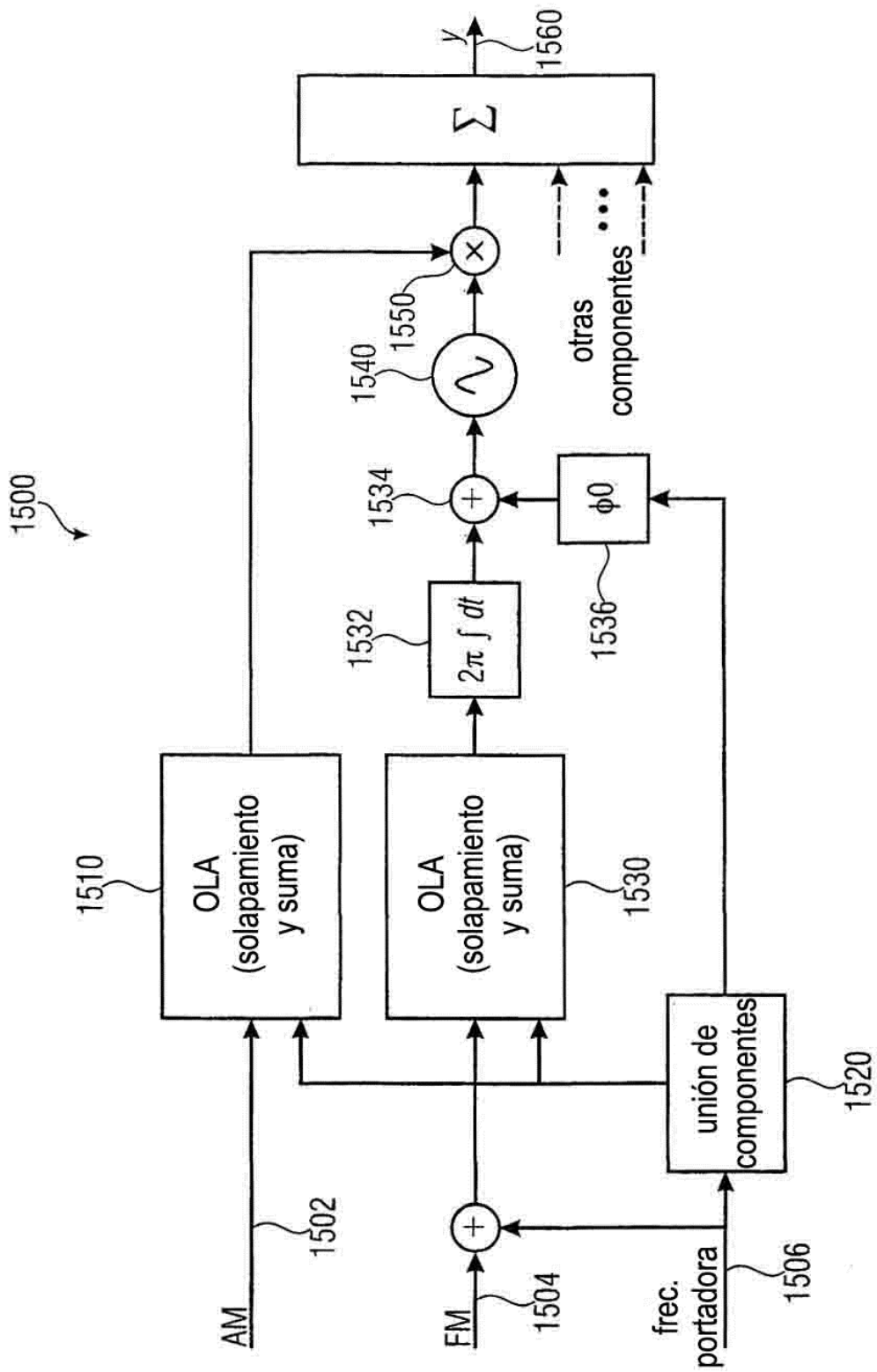


FIGURA 15

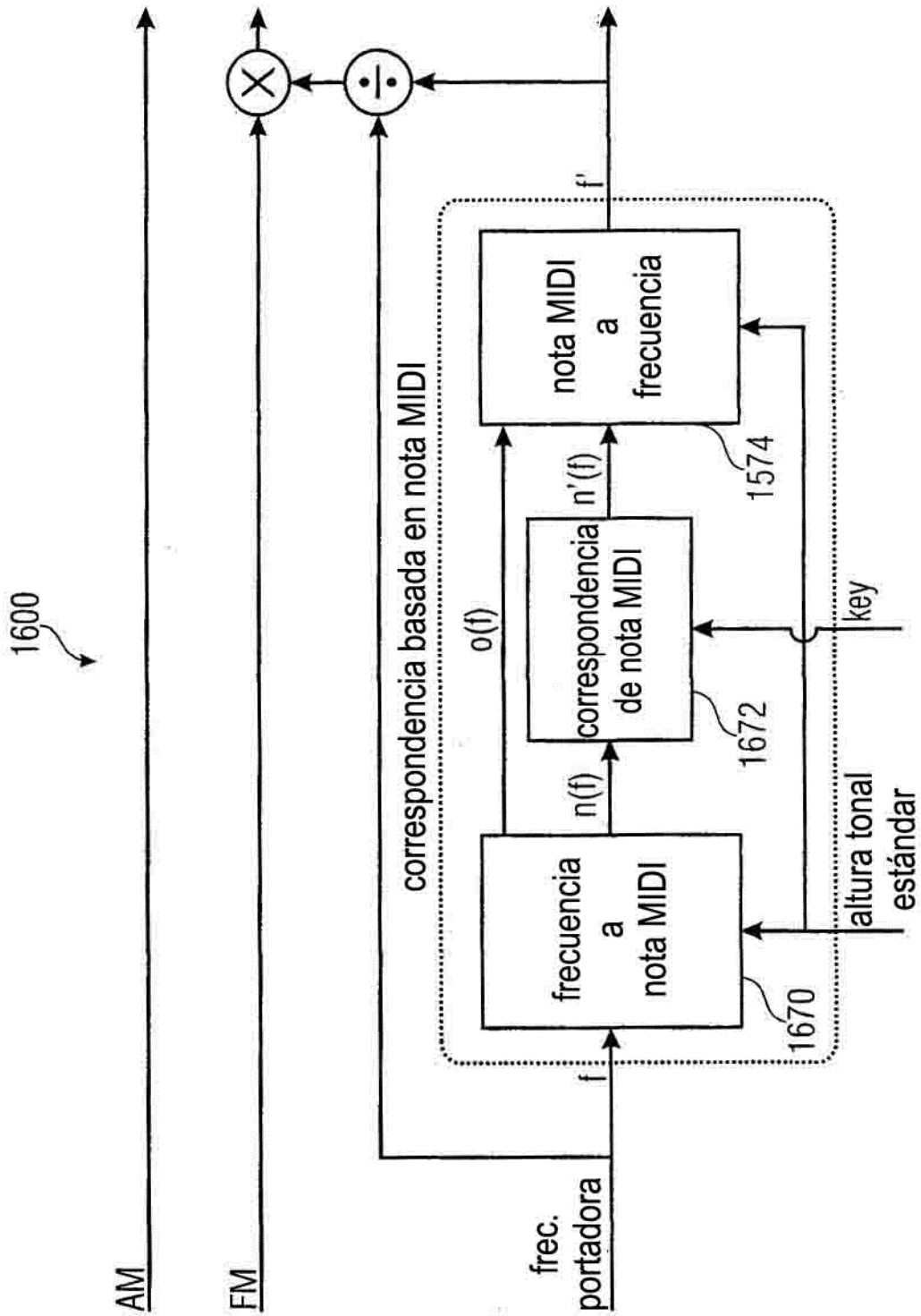


FIGURA 16

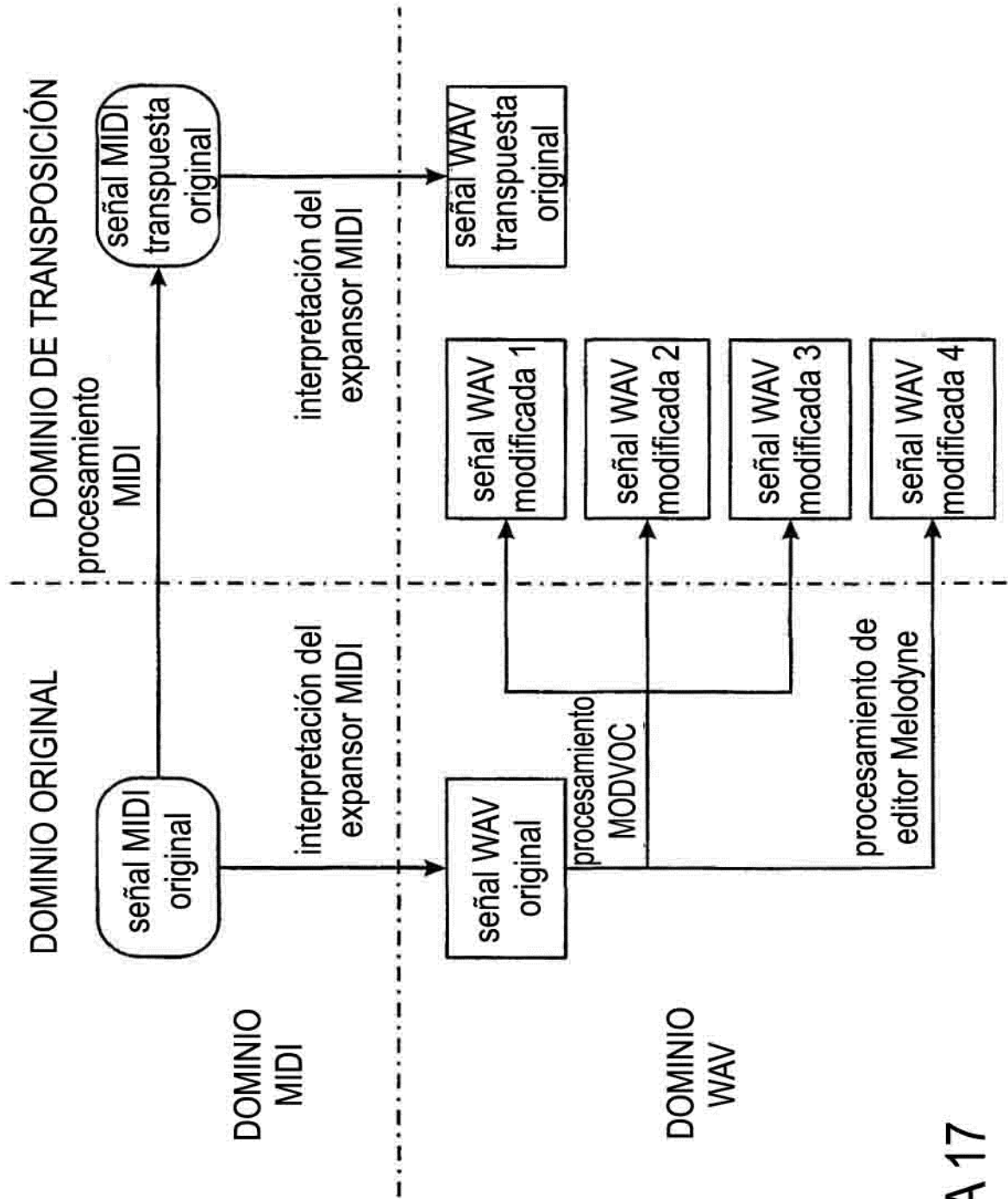


FIGURA 17

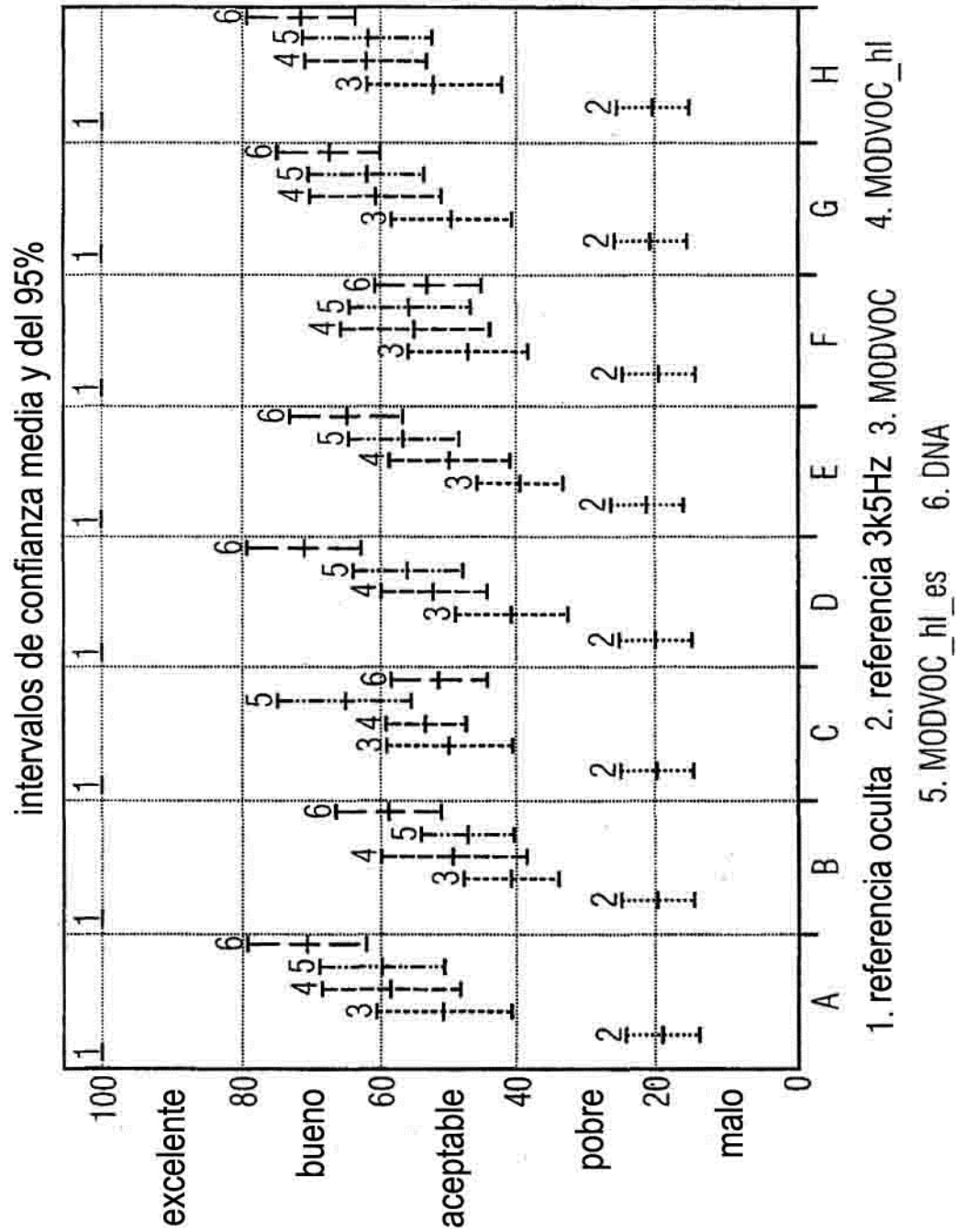


FIGURA 18

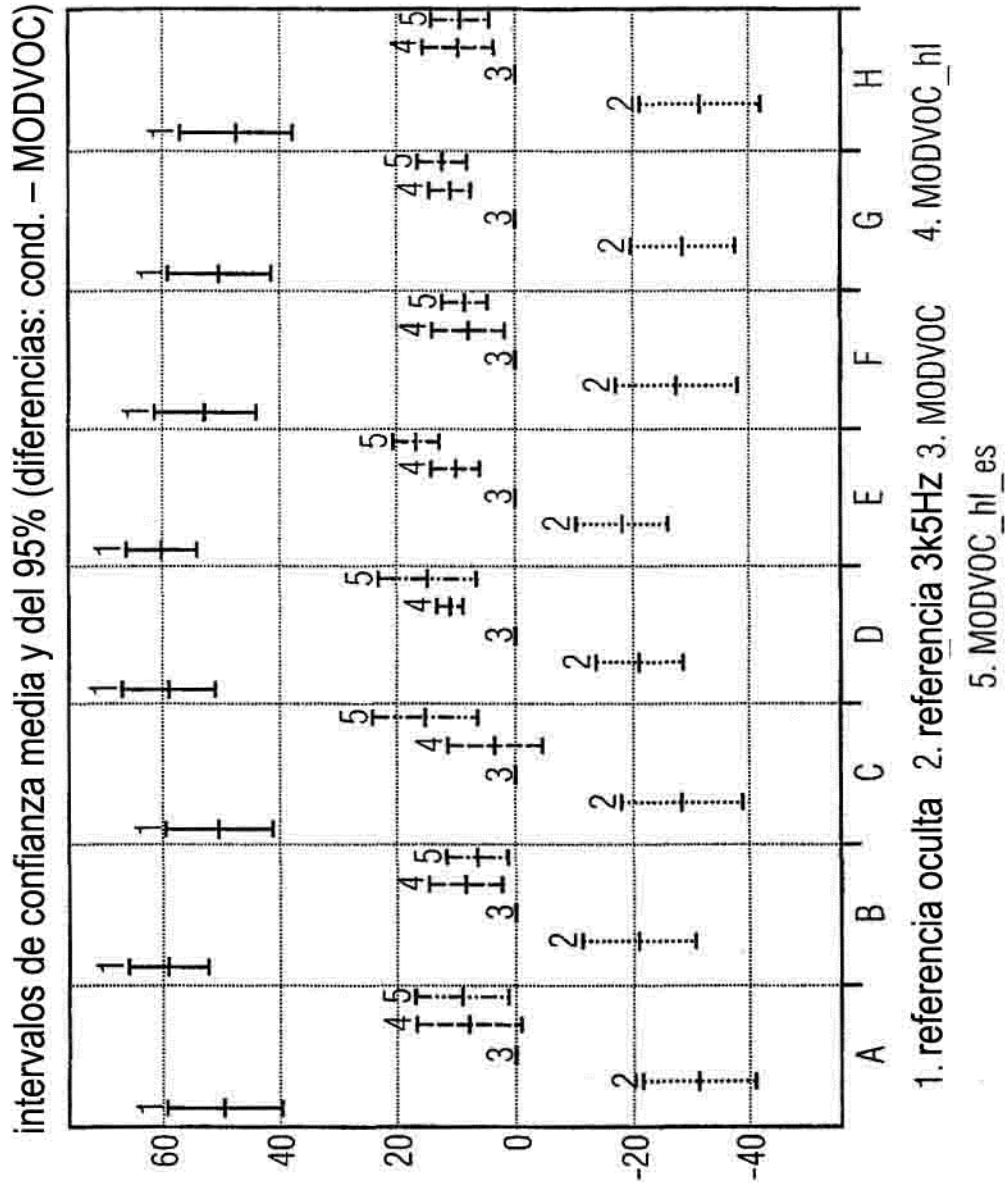


FIGURA 19

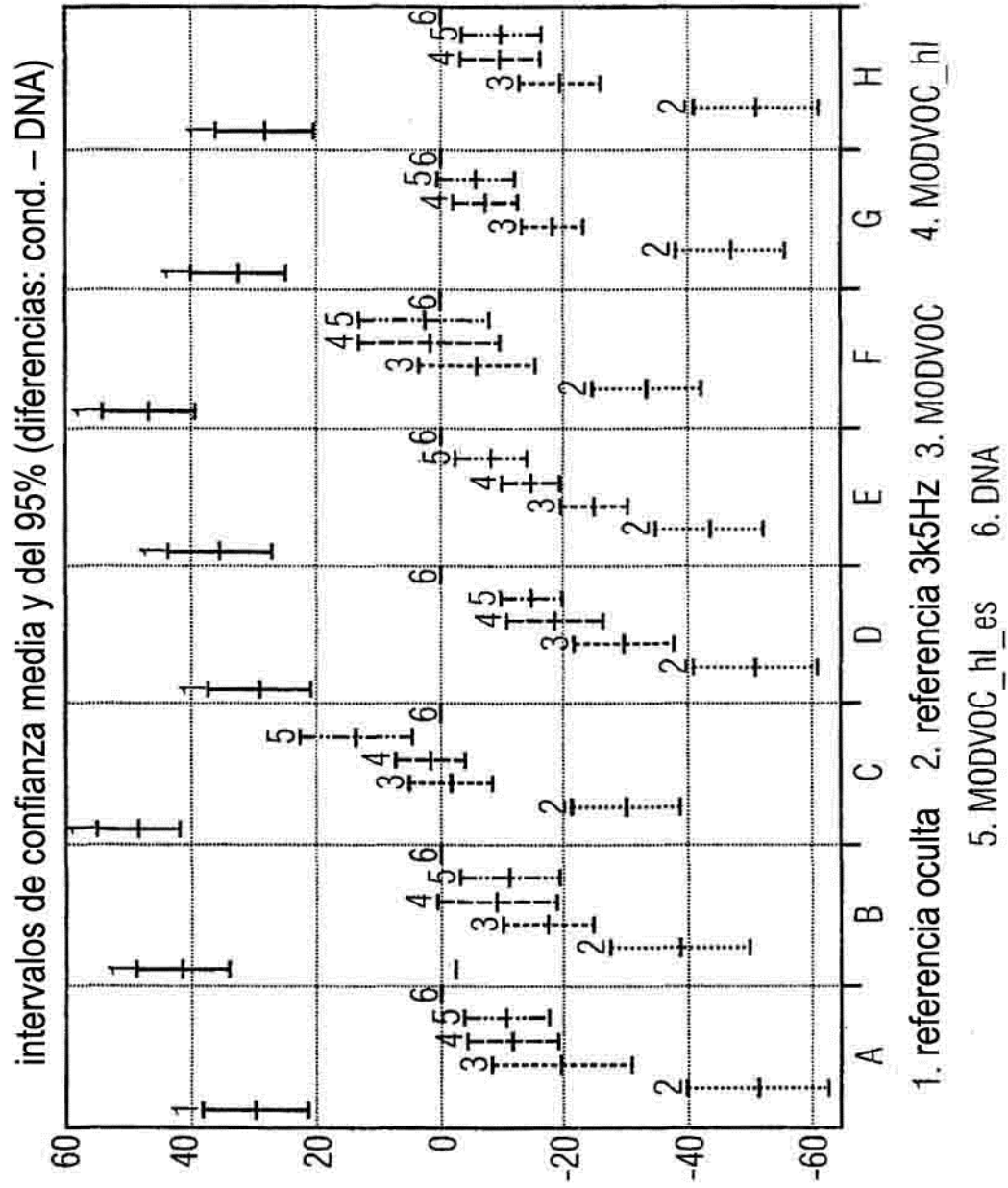


FIGURA 20