

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 523 800**

51 Int. Cl.:

**G10L 21/00** (2013.01)  
**H03G 3/00** (2006.01)  
**H03G 5/00** (2006.01)  
**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 11705876 (8)**

97 Fecha y número de publicación de la concesión europea: **13.08.2014 EP 2539886**

54 Título: **Aparato y procedimiento para modificar una señal de audio usando modelado de envolvente**

30 Prioridad:

**03.09.2010 EP 10175302**  
**26.02.2010 US 308513 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**01.12.2014**

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 2 523 800 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 modelado de envolvente

5 **[0001]** Las realizaciones de acuerdo con el invento se refieren a procesamiento de audio y particularmente a un aparato y un procedimiento para modificar una señal de audio.

10 **[0002]** Hay una creciente demanda de técnicas de procesamiento de señal digital que apunta a la necesidad de manipulaciones de señal extremas para adecuar señales de audio pre-grabadas, por ejemplo, tomadas de una base de datos, a un nuevo contexto musical. Para hacerlo, es necesario adaptar propiedades semánticas de la señal de alto nivel tal como tono, clave musical y modo de escala. Todas estas manipulaciones tienen en común que apuntan a alterar substancialmente las propiedades musicales del material de audio original mientras que conservan lo mejor posible la calidad de sonido subjetiva. En otras palabras, estos editores cambian fuertemente el contenido musical de material de audio pero, no obstante, se requiere que conserven la naturaleza de la muestra de audio procesada y que por ende mantengan la credibilidad. Idealmente esto requiere procedimientos de procesamiento de señal que sean ampliamente aplicables a diferentes clases de señales incluyendo contenido de música polifónica mezclada.

15 **[0003]** Hoy en día se conocen muchos conceptos para modificar señales de audio. Algunos de estos conceptos se basan en vocoders.

20 **[0004]** Por ejemplo, en el documento de "S. Disch y B. Edler, "Un vocoder de modulación de amplitud y de frecuencia para procesamiento de señal de audio" ("*An amplitude- and frequency modulation vocoder for audio signal processing*") Acta. de la Conf. Intern. sobre Efectos de Audio Digital (DAFx), 2008", en el documento de S. Disch y B. Edler, "Análisis, procesamiento y Síntesis de modulación perceptual multibanda, de señales de audio" ("*Multiband perceptual modulation analysis, processing and Synthesis of audio signals*") Acta de IEEE-ICASSP, 2009, o en el documento de S. Disch y B. Edler, "Un algoritmo de segmentación iterativa para espectro de señal de audio dependiendo de centros de gravedad locales estimados" ("*An iterative segmentation algorithm for audio signal spectra depending on estimated local centers of gravity*") 12<sup>o</sup> Conferencia Internacional sobre Efectos de Audio Digital (DAFx-09), 2009.", se presentó el concepto de vocoder de modulación (MODVOC) y se destacó su capacidad general de realizar una significativa transposición selectiva sobre contenido de música polifónica. Esto posibilita aplicaciones que apuntan a cambiar el modo clave de muestras de música PCM pre-grabadas (ver por ejemplo, Disch y B. Edler, "*Multiband perceptual modulation analysis, processing and Synthesis of audio signals*," ("Análisis, procesamiento y Síntesis de modulación perceptual multibanda, de señales de audio") Proc. de IEEE-ICASSP, 2009."). También está disponible comercialmente un primer software que puede manejar una tarea así de manipulación polifónica (editor Melodyne de Celemony). El software implementa una tecnología que se denomina y se comercializa mediante el término acceso directo a nota (*direct note access*) (DNA). Hace poco se ha publicado una solicitud de patente (EP2099024, P. Neubäcker, "*Method for acoustic object-oriented analysis and note object-oriented processing of polyphonic sound recordings*," ("Procedimiento para análisis acústico orientado a objeto, de grabaciones de sonido polifónico") septiembre de 2009.) que presumiblemente cubre y por ende revela la funcionalidad esencial del DNA. Independientemente del procedimiento usado para modificar una señal de audio, se desea obtener una señal de audio con alta calidad perceptual.

35 **[0005]** Es un objetivo del presente invento proveer un concepto mejorado para modificar una señal de audio, el cual permite aumentar la calidad perceptual de la señal de audio modificada.

40 **[0006]** Este objetivo se resuelve mediante un aparato de acuerdo con la reivindicación 1, un procedimiento de acuerdo con la reivindicación 13 o un programa de computadora de acuerdo con la reivindicación 14.

45 **[0007]** Una realización del invento provee un aparato para modificar una señal de audio que comprende un procesador de banco de filtros, un elemento de determinación de fundamental, un elemento de determinación de sobretono, un procesador de señal y un combinador. El procesador de banco de filtros está configurado para generar una pluralidad de señales pasabanda en base a una señal de audio. Además, el elemento de determinación de fundamental está configurado para seleccionar una señal pasabanda de la pluralidad de señales pasabanda para obtener una señal pasabanda fundamental. El elemento de determinación de sobretono está configurado para identificar una señal pasabanda de la pluralidad de señales pasabanda que cumplen un criterio de sobretono con relación a la señal pasabanda fundamental seleccionada para obtener una señal pasabanda de sobretono asociada a la señal pasabanda fundamental seleccionada. Además, el procesador de señal está configurado para modificar la señal pasabanda fundamental seleccionada en base a un blanco de modificación predefinido. Adicionalmente, el procesador de señal está configurado para modificar una señal pasabanda de sobretono identificada asociada a la señal pasabanda fundamental seleccionada dependiendo de la modificación de la señal pasabanda fundamental seleccionada. Además, el combinador está configurado para combinar la pluralidad de señales pasabanda para obtener una señal de audio modificada.

5 **[0008]** Identificando sobretonos de frecuencias fundamentales y modificando los sobretonos de la misma manera que las correspondientes fundamentales, se puede evitar una modificación diferente de las fundamentales y sus sobretonos, de modo que se puede conservar más exactamente el timbre de una señal de audio modificada en comparación con la señal de audio original. De esta manera se puede mejorar significativamente la calidad perceptual de la señal de audio modificada. Por ejemplo, si se desea una transposición de tono (por ejemplo, alterar el modo clave de C mayor a C menor de una dada señal de música), la modificación de una señal pasabanda de sobretono identificada está correlacionada con la modificación de la señal pasabanda fundamental. En comparación, los procedimientos conocidos modifican la región de frecuencia de la señal pasabanda que representa sobretonos de manera diferente de la señal pasabanda fundamental. En otras palabras, una señal pasabanda de sobretono identificada es ligada a la señal pasabanda fundamental usando el concepto descrito.

15 **[0009]** En algunas realizaciones del invento, una señal pasabanda de sobretono puede ser identificada comparando frecuencias de la señal pasabanda fundamental y señales pasabanda de la pluralidad de señales pasabanda, comparando un contenido de energía de la señal pasabanda fundamental y una señal pasabanda de la pluralidad de señales pasabanda y/o evaluando una correlación de una envolvente temporal de la señal pasabanda fundamental y una envolvente temporal de una señal pasabanda de la pluralidad de señales pasabanda. De esta manera se puede definir uno o más criterios de sobretono para minimizar la identificación de sobretonos erróneos.

20 **[0010]** Algunas realizaciones de acuerdo con el invento se refieren a una determinación iterativa de señales pasabanda fundamentales e identificación de señales pasabanda de sobretono entre la pluralidad de señales pasabanda. Las señales pasabanda fundamentales ya seleccionadas y las señales pasabanda de sobretono ya identificadas pueden ser retiradas del espacio de búsqueda o, en otras palabras, pueden no ser consideradas para la determinación de una señal pasabanda fundamental adicional o una señal pasabanda de sobretono adicional. De esta manera, cada señal pasabanda de la pluralidad de señales pasabanda puede ser seleccionada como una señal pasabanda fundamental (y, por lo tanto, puede ser modificada independientemente de las otras señales pasabanda fundamentales) o una señal pasabanda de sobretono (y, por lo tanto, puede ser modificada dependiendo de la señal pasabanda fundamental asociada seleccionada).

30 **[0011]** Otra realización del invento provee un aparato para modificar una señal de audio que comprende un elemento de determinación de modelado de envolvente, un procesador de banco de filtros, un combinador y un modelador de envolvente. El elemento de determinación de modelado de envolvente está configurado para determinar coeficientes de modelado de envolvente en base a una señal de audio del dominio de la frecuencia que representa una señal de audio de entrada del dominio del tiempo. Además, el procesador de banco de filtros está configurado para generar una pluralidad de señales pasabanda en un dominio subbanda en base a la señal de audio del dominio de la frecuencia. El procesador de señal está configurado para modificar una señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un blanco de modificación predefinido. Además, el combinador está configurado para combinar por lo menos un subconjunto de la pluralidad de señales pasabanda del dominio subbanda para obtener una señal de audio del dominio del tiempo. Adicionalmente, el modelador de envolvente está configurado para modelar una envolvente de la señal de audio del dominio del tiempo en base a los coeficientes de forma de envolvente, para modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda conteniendo la señal pasabanda del dominio subbanda modificada en base a los coeficientes de forma de envolvente o para moldear una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base a coeficientes de forma de envolvente antes que una señal pasabanda de dominio subbanda sea modificada por el procesador de señal para obtener una señal de audio moldeada.

45 **[0012]** Determinando coeficientes de forma de envolvente de la señal de audio del dominio de la frecuencia antes de que sea separada la señal de audio del dominio de la frecuencia en una pluralidad de señales pasabanda del dominio subbanda, se puede conservar una información acerca de la coherencia espectral de la señal de audio y se puede usar para modelar la envolvente de la señal de audio del dominio del tiempo después de la modificación de una o más señales pasabanda del dominio subbanda. De esta manera se puede conservar más exactamente la coherencia espectral de la señal de audio modificada, a pesar de que sólo algunas (o sólo una) señales pasabanda del dominio subbanda son modificadas o señales pasabanda del dominio subbanda son modificadas de manera diferente, lo cual puede perturbar la coherencia espectral de la señal de audio. De esta manera se puede mejorar significativamente la calidad perceptual de la señal de audio modificada.

60 **[0013]** Algunas realizaciones de acuerdo con el invento se refieren a un procesador de señal configurado para modificar una segunda señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un segundo blanco de modificación predefinido. El blanco de modificación predefinido y el segundo blanco de modificación predefinido son diferentes. A pesar de que las señales pasabanda son modificadas de manera diferente, la coherencia espectral de la señal de audio modificada puede ser conservada más exactamente debido al modelado de envolvente después de la modificación individual de las señales pasabanda.

**[0014]** Se detallarán subsiguientemente realizaciones de acuerdo con el invento con referencia a los dibujos anexos, en los cuales:

- 5 la Figura 1 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la Figura 2 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la Figura 3 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- 10 la Figura 4 es un diagrama de bloques de una parte de un vocoder de modulación que usa bloqueo armónico (*harmonic locking*);
- la Figura 5 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- 15 la Figura 6a, 6b, 6c, 6d es un diagrama de bloques de un aparato para modificar una señal de audio;
- la Figura 7 es un diagrama de bloques de un procesador de banco de filtros;
- la Figura 8 es un diagrama de bloques de un modelador de envolvente;
- 20 la Figura 9 es una ilustración esquemática de un análisis de modulación con modelado de envolvente;
- la Figura 10 es una ilustración esquemática de un síntesis de modulación con modelado de envolvente;
- 25 la Figura 11 es un diagrama de flujo de un procedimiento para modificar una señal de audio;
- la Figura 12 es un diagrama de bloques de un aparato para modificar una señal de audio;
- la Figura 13 es una ilustración esquemática de un análisis de modulación;
- 30 la Figura 14 es una ilustración esquemática de una implementación de un análisis de modulación;
- la Figura 15 es una ilustración esquemática de una síntesis de modulación;
- 35 la Figura 16 es una ilustración esquemática de una transposición selectiva sobre un componente de vocoder de modulación;
- la Figura 17 es una ilustración esquemática de un procedimiento para generar el conjunto de prueba para evaluación de la calidad subjetiva del procesamiento de vocoder de modulación para la tarea de transposición selectiva de tono;
- 40 la Figura 18 es un diagrama que indica puntajes de MUSHRA absolutos e intervalos de confianza de 95% de prueba de audición que apunta a transposición selectiva de tono;
- la Figura 19 es un diagrama que indica una diferencia de puntajes de MUSHRA con respecto a una condición de vocoder de modulación e intervalos de confianza de 95% de prueba de audición que apunta a transposición selectiva de tono; y
- 45 la Figura 20 es un diagrama que indica una diferencia de puntajes de MUSHRA con respecto a condición DNA e intervalos de confianza de 95% de prueba de audición que apunta a transposición selectiva de tono.
- 50

**[0015]** En lo que sigue se usan parcialmente los mismos números de referencia para objetos y unidades funcionales que tienen las mismas o similares propiedades funcionales y la descripción de ellos con relación a una figura también se aplicará a otras figuras para reducir redundancia en la descripción de las realizaciones.

55 **[0016]** Se puede realizar una modificación de banda de frecuencia selectiva, también llamada transposición selectiva de tono, por ejemplo, mediante un vocoder o un vocoder de modulación.

60 **[0017]** Una descomposición de modulación multibanda (ver por ejemplo "S. Disch y B. Edler, "*Multiband perceptual modulation analysis, processing and Synthesis of audio signals*," ("Análisis, procesamiento y Síntesis de modulación perceptual multibanda, de señales de audio") Proc.de IEEE-ICASSP, 2009.") divide la señal de audio en un conjunto señal-adaptable (*signal adaptive set*) de señales pasabanda (analíticas), cada una de las cuales además es dividida en una portadora sinusoidal y su modulación de amplitud (AM) y modulación de frecuencia (FM). El conjunto de filtros pasabanda puede ser computado tal que, por un lado, se cubra todo el espectro de manera continua y por el otro, los filtros estén alineados con los centros de gravedad totales (COGs), por ejemplo. Adicionalmente, la

percepción del auditorio humano puede ser tenida en cuenta eligiendo el ancho de banda de los filtros para que coincidan con una escala perceptual, por ejemplo, la escala ERB (ver, por ejemplo, "B. C. J. Moore y B. R. Glasberg, "A revision of zwicker's loudness model" ("Una revisión del modelo de volumen de Zwicker" Acta Acústica, vol. 82, pp. 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 región de frecuencia. Asimismo, las bandas centradas en posiciones de COG locales pueden corresponder a bloqueos de fase en base a regiones de influencia de vocoders de fase clásicos (ver por ejemplo, de J. Laroche y M. Dolson, "Modificación del audio de escala de tiempo de vocoder de fase mejorado" ("*Improved phase vocoder timescale modification of audio*") Transacciones en Procesamiento de Voz y Audio, IEEE (*Transactions on Speech and Audio Processing*) vol. 7, no. 3, pp. 323–332, 1999." o de C. Duxbury, M. Davies, y M. Sandler, "Escaleo de tiempo mejorado de audio musical usando bloqueo de fase en componentes transitorios" ("*Improved timescaling of musical audio using phase locking at transients*") en la 12<sup>o</sup> Convención de AES, 2002.). Tanto la representación de envolvente de señal pasabanda como el bloqueo de fase de la región de influencia tradicional preservan la envolvente temporal de una señal pasabanda: ya sea intrínsecamente o, en el último caso, asegurando coherencia de fase espectral local durante la síntesis. Con respecto a la portadora sinusoidal de una frecuencia correspondiente al COG local estimado, ambas, AM y FM son capturadas en la envolvente de amplitud y la fase heterodinada de las señales pasabanda analíticas, respectivamente. Un procedimiento de síntesis dedicado convierte la señal de salida desde las frecuencias de la portadora, AM y FM.

**[0019]** Un diagrama de bloques de una posible implementación 1300 de la descomposición de señal en señales portadoras y sus componentes de modulación asociados, está representado en la Figura 13. En la figura se muestra el flujo esquemático de señal para la extracción de una de las componentes multibanda (señales pasabanda). Todas las otras componentes se obtienen de una manera similar. Primero se entrega una señal de entrada de banda ancha  $x$  al filtro pasabanda que ha sido diseñado para proveer, de manera señal-adaptable, una señal de salida  $A$  continuación se establece la señal analítica mediante la transformación de Hilbert de acuerdo con la Ecuación (1).

$$\hat{x}(t) = \tilde{x}(t) + j\mathcal{H}(\tilde{x}(t)) \quad (1)$$

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

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

**[0021]** mientras que la FM (señal de modulación de frecuencia) se obtiene mediante la derivada de fase de la señal analítica heterodinada por una portadora sinusoidal estacionaria con frecuencia angular  $\omega_c$ . Se determina la frecuencia de la portadora para que sea una estimación del COG local. Por ende, la FM puede ser interpretada como la variación de IF (frecuencia instantánea) a la frecuencia de la portadora  $f_c$ .

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

**[0022]** La estimación del COG local y el diseño adaptativo de señal del banco de filtros del extremo frontal se describe, por ejemplo, en un publicación específica (ver "S. Disch y B. Edler, "An iterative segmentation algorithm for audio signal spectra depending on estimated local centers of gravity," ("Un algoritmo de segmentación iterativa para espectro de señal de audio dependiendo de centros de gravedad locales estimados") 12<sup>o</sup> Conferencia Internacional sobre Efectos de Audio Digital (DAFx-09). 2009").

**[0023]** En la práctica, en un sistema de tiempo discreto, la extracción de componentes puede ser llevada a cabo conjuntamente para todas las componentes 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 es dependiente de parámetros de bloques previos. Por ende, no se requiere mirar adelante para mantener lo más bajo posible la demora de procesamiento. El procesamiento es computado base por base usando, por ejemplo 75% de superposición de bloque de análisis y aplicación de una transformación de Fourier discreta (DFT) sobre cada bloque de señal ventaneada. La ventana puede ser una ventana de parte superior plana de acuerdo con la Ecuación (4).

**[0024]** Esto asegura que las  $N/2$  muestras centradas que se hacen pasar para la subsiguiente síntesis de modulación utilizando superposición de 50% no sean afectadas por los faldones de la ventana de análisis. Se puede usar un grado más alto de superposición para mejorar exactitud a costo de mayor complejidad computacional.

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

5 [0025] Dada la representación espectral, a continuación se calcula un conjunto de funciones de ponderación pasabanda espectral señal–adaptable, que está alineado con posiciones de COG locales. Después de la aplicación de la ponderación pasabanda al espectro, la señal es transferida al dominio del tiempo y la señal analítica puede ser establecida mediante la transformación de Hilbert. Estos dos pasos de procesamiento puede ser eficientemente combinados mediante el cálculo de un IDFT de una solo lado en cada señal pasabanda. Dada una señal pasabanda de tiempo discreta, la estimación de la IF mediante la ecuación (3) es implementada mediante diferenciación de fase según se define en la Ecuación (5) donde \* denota el complejo conjugado. Esta expresión se usa convenientemente ya que evita ambigüedades de fase y por ende la necesidad de desrapinado de fase.

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

15 [0026] La señal es sintetizada sobre una base aditiva de todos los componentes. Los sucesivos bloques son combinados mediante superposición–adición (*overlap-add*) (OLA) el cual es controlado por el mecanismo de vinculación. La vinculación de componentes asegura una transición suave entre los bordes de bloques adyacentes incluso si los componentes son substancialmente alterados por un procesamiento de dominio de modulación. La vinculación sólo toma en cuenta el bloque previo permitiendo potencialmente procesamiento de dominio de modulación. La vinculación esencialmente realiza un apareamiento a modo de pares de los componentes del bloque actual con sus predecesores en el bloque previo. Adicionalmente, la vinculación alinea las fases de componente absolutas del bloque actual con las del bloque previo. Para componentes que no tienen coincidencia a través de bloques de tiempo, se aplica un fundido de entrada o fundido de salida, respectivamente.

25 [0027] En la Figura 15 se muestra la cadena de procesamiento para un componente. En detalle, primero se agrega la señal FM a la frecuencia de la portadora estacionaria y se hace pasar la señal resultante a una etapa de OLA, la salida de la cual es subsiguientemente integrada temporalmente. Se alimenta un oscilador sinusoidal mediante la señal de fase resultante. La señal AM es procesada por una segunda etapa de 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 un paso final, se suman las contribuciones de todos los componentes para obtener la señal de salida y.

35 [0028] En otras palabras, las Figuras 13 y 14 ilustran un analizador de modulación 1300. El analizador de modulación 1300 preferiblemente comprende un filtro pasabanda 1320a, el cual provee una señal pasabanda. Ésta es ingresada en un convertidor de señal analítico 1320b. La salida del bloque 1320b es útil para calcular información de AM e información de FM. Para calcular información de AM, la magnitud de la señal analítica es calculada por el bloque 1320c. La salida del bloque de señal analítica 1320b es ingresada en un multiplicador 1320d, el cual recibe, en su otra entrada, una señal de oscilador desde un oscilador 1320e, el cual es controlado por la frecuencia de la portadora actual  $f_c$  1310 del pasabanda 1320a. Luego, en el bloque 1320f es determinada la fase la salida del multiplicador. La fase instantánea es diferenciada en el bloque 1320g para obtener finalmente la información de FM. Además, la Figura 14 muestra un procesador 1410 que genera un espectro de DFT de la señal de audio.

45 [0029] La descomposición de modulación multibanda divide la señal de audio en un conjunto señal–adaptable de señales pasabanda (analíticas), cada una de las cuales es a su vez dividida en una portadora sinusoidal y su modulación de amplitud (AM) y modulación de frecuencia (FM). El conjunto de filtros pasabanda es computado tal que, por un lado, se cubra todo el espectro de manera continua y por el otro, los filtros estén alineados con cada COG local. Adicionalmente, la percepción del auditorio humano es tenida en cuenta eligiendo el ancho de banda de los filtros para que coincidan con una escala perceptual, por ejemplo, la escala ERB (ver, B. C. J. Moore y B. R. Glasberg, "A revision of Zwicker's loudness model" ("Una revisión del modelo de volumen de Zwicker" Acta Acústica, vol. 82, pp. 335–345, 1996").

55 [0030] El COG local corresponde a la frecuencia media que es percibida por un oyente debido a las contribuciones espectrales en esa región de frecuencia. Asimismo, las bandas centradas en posiciones de COG locales corresponden a bloqueos de fase en base a regiones de influencia de vocoders de fase clásicos (ver, de J. Laroche y M. Dolson, "Modificación del audio de escala de tiempo de vocoder de fase mejorado" ("*Improved phase vocoder timescale modification of audio*") Transacciones en Procesamiento de Voz y Audio, IEEE (*Transactions on Speech and Audio Processing*) vol. 7, no. 3, pp. 323–332, 1999" o de C. Duxbury, M. Davies, y M. Sandler, "Escaleo de tiempo mejorado de audio musical usando bloqueo de fase en componentes transitorios" ("*Improved timescaling of*

musical audio using phase locking at transients") en la 12<sup>o</sup> Convención de AES, 2002, de A. Röbel, "Un nuevo enfoque al procesamiento de componente transitorio en el vocoder de fase" ("A new approach to transient processing in the phase vocoder") Acta de la Conf. Int. sobre Efectos de Audio Digital (DAFx), pp. 344–349, 2003", de A. Röbel, "Detección y conservación de componente transitorio en el vocoder de fase" ("Transient detection and preservation in the phase vocoder", Conferencia Int. de Música de Computadora (ICMC'03), pp. 247–250, 2003"). Tanto la representación de envolvente de señal pasabanda como el bloqueo de fase de la región de influencia tradicional preservan la envolvente temporal de una señal pasabanda: ya sea intrínsecamente o, en el último caso, asegurando coherencia de fase espectral local durante la síntesis. Con respecto a la portadora sinusoidal de una frecuencia correspondiente al COG local estimado, ambas, AM y FM son capturadas en la envolvente de amplitud y la fase heterodinada de las señales pasabanda analíticas, respectivamente. Un procedimiento de síntesis dedicado convierte la señal de salida desde las frecuencias de la portadora, AM y FM.

[0031] Un diagrama de bloques de la descomposición de señal en señales portadoras y sus componentes de modulación asociados, está representado en la Figura 12. En la figura se muestra el flujo esquemático de señal para la extracción de una de componente. Todas las otras componentes se obtienen de una manera similar. En la práctica, la extracción es llevada a cabo conjuntamente para todas las componentes de modo bloque a bloque usando por ejemplo, un tamaño de bloque de  $N = 2^{14}$  frecuencia de muestreo a 48 kHz y superposición de análisis 75% – que aproximadamente corresponde a un intervalo de tiempo de 340 ms y un paso (*stride*) de 85 ms – mediante aplicación de una transformación discreta de Fourier (DFT) a cada bloque de señal ventaneado. La ventana puede ser una ventana "de parte superior plana" de acuerdo con la Ecuación (a). Esto puede asegurar que las  $N/2$  muestras centradas que se hacen pasar para la subsiguiente síntesis de modulación no sean afectadas por las laderas de la ventana de análisis. Se puede usar un grado más alto de superposición para mejor exactitud a costo de mayor complejidad computacional.

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

(a)

[0032] Dada la representación espectral, a continuación se puede calcular un conjunto de funciones de ponderación espectral señal–adaptables (teniendo característica pasabanda) que está alineado con posiciones de COG local (mediante el medio de determinación de frecuencias de la portadora 1330 en términos de una estimación de frecuencias de la portadora o una estimación de frecuencia de COG de la portadora múltiple. Después de la aplicación de la ponderación pasabanda al espectro, la señal es transformada al dominio del tiempo y la señal analítica es establecida mediante la transformación de Hilbert. Estos dos pasos de procesamiento puede ser eficientemente combinados mediante el cálculo de un IDFT de una solo lado en cada señal pasabanda. Subsiguientemente, cada señal analítica es heterodinada por su frecuencia de la portadora estimada. Finalmente, la señal es descompuesta adicionalmente en su envolvente de amplitud y su rastro de frecuencia instantánea (IF), obtenida computando la derivada de fase, produciendo la señal deseada AM y FM (ver también, de S. Disch y B. Edler, "Un vocoder de modulación de amplitud y de frecuencia, para procesamiento de señal de audio" ("An amplitude– and frequency modulation vocoder for audio signal processing" Acta de la Conf. Int. sobre Efectos de Audio Digital (DAFx), 2008).

[0033] Oportunamente, la Figura 15 muestra un diagrama de bloques de un sintetizador de modificación 1500 para una representación parametrizada de una señal de audio. Por ejemplo, una implementación ventajosa se basa en una operación de superposición–adición (OLA) en el dominio de modulación, esto es, en el dominio antes de generar la señal pasabanda del dominio del tiempo. La señal de entrada la cual puede ser una serie de bits en el tiempo, pero la cual puede también ser una conexión directa a un analizador o modificador, es separada en la componente AM 1502, la componente FM 1504 y la componente de frecuencia portador 1506. El sintetizador AM preferiblemente comprende un superposición–adicionador 1510 y, adicionalmente, un controlador de vinculación de componente 1520 el cual, preferiblemente no sólo comprende el bloque 1510 sino también el bloque 1530, el cual es una superposición–adicionador adentro del sintetizador FM. El sintetizador FM adicionalmente comprende un superposición–adicionador de frecuencia 1530, un integrador de frecuencia instantánea 1532, un combinador de fase 1534 el cual, nuevamente, puede ser implementado como un adición regular y un corredor de fase 1536 el cual es controlable mediante el controlador de vinculación de componente 1520 para regenerar una fase constante de bloque a bloque de modo que la fase de una señal de un bloque precedente sea continua la fase de un bloque actual. Por lo tanto, uno puede decir que la adición de fase en los elementos 1534, 1536 corresponde a una regeneración de una constante que fue perdida durante la diferenciación en el bloque 1520 en la Figura 13 del lado del analizador. Desde una perspectiva de pérdida de información en el dominio perceptual, se ha de notar que ésta es la única pérdida de información, esto es, la pérdida de una porción constante mediante el dispositivo de

diferenciación 1320g de la Figura 13. Esta pérdida puede ser compensada adicionando una fase constante determinada por el dispositivo de vinculación de componente 1520.

5 **[0034]** Se aplica superposición–adición (OLA) en el dominio de parámetro en lugar de sobre la señal fácilmente sintetizada para evitar efectos de latido entre bloques de tiempo adyacentes. La OLA es controlada por un mecanismo de vinculación de componente, que, gobernado por vecindad espectral (medida en una escala ERB), realiza un apareamiento a modo de pares de componentes del bloque actual con sus predecesores de bloques previos. Adicionalmente, la vinculación alinea las fases de componente absolutas del bloque actual con las del bloque previo.

10 **[0035]** En detalle, primero se agrega la señal FM a la frecuencia de la portadora y se hace pasar el resultado a la etapa de OLA, la salida de la cual es subsiguientemente integrada. Se alimenta un oscilador sinusoidal 1540 mediante la señal de fase resultante. La señal AM es procesada por una segunda etapa de OLA. Finalmente, 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.

15 **[0036]** Se debe enfatizar que una apropiada segmentación espectral de la señal adentro del análisis de modulación es de primordial importancia para un resultado convincente de cualquier otro procesamiento de parámetro de modulación. Por lo tanto, aquí se describe un ejemplo para un ejemplo de segmentación.

20 **[0037]** Oportunamente, la Figura 16 muestra un ejemplo 1600 de una aplicación para cambios de modo clave polifónica. La figura muestra una transposición selectiva sobre componentes de vocoder de modulación. Las frecuencias de la portadora son cuantificadas a notas MIDI las cuales son mapeadas a correspondientes notas MIDI apropiadas. Preservación de modulación FM relativa mediante multiplicación de las componentes mapeadas por la proporción de frecuencias de la portadora original y modificada.

25 **[0038]** La transposición de una señal de audio mientras se mantiene la velocidad de reproducción original es una tarea desafiante. Usando el sistema propuesto, se logra directamente mediante multiplicación de todas las componentes portadoras con un factor constante. Como la estructura temporal de la señal de entrada únicamente es capturada por las señales AM ésta no es afectada por el estiramiento del espaciamiento espectral de la portadora.

30 **[0039]** Un efecto incluso más exigente se puede obtener mediante procesamiento selectivo. El modo clave de una pieza musical puede ser cambiado por ejemplo, de menor a mayor o vice versa. Por lo tanto, sólo se mapea a nuevos valores adecuados un subconjunto de portadoras correspondientes a ciertos intervalos de frecuencia predefinidos. Para lograr esto, las frecuencias de la portadora son cuantificadas 1670 a tonos MIDI los cuales son subsiguientemente mapeados 1672 a nuevos tonos MIDI apropiados (usando conocimiento a–priori de modo y clave de la pieza musical a ser procesada).

35 **[0040]** Luego, las notas MIDI mapeadas son convertidas de vuelta 1574 para obtener las frecuencias de la portadora modificadas que son usadas para síntesis. No se requiere una detección de inicio/desplazamiento de nota MIDI especializada ya que las características temporales están predominantemente representadas por la AM no modificada y por ende se conservan. Se pueden definir tablas de mapeo arbitrarias para permitir conversión a y desde otros sabores menores (por ejemplo, armónico menor).

40 **[0041]** Una aplicación en el campo de 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. También multiplicando la FM con el mismo factor se asegura que, para cada componente, se conserva la profundidad de modulación FM relativa. Como la estructura temporal de la señal de entrada únicamente es capturada por las señales AM ésta no es afectada por el procesamiento. La transposición global cambia la clave original de una señal de música hacia una clave blando (por ejemplo, de C mayor a G mayor) mientras se conserva el tempo original.

45 **[0042]** Sin embargo, debido a la naturaleza señal adaptable del análisis de modulación propuesto, el vocoder de modulación tiene el potencial de ir más allá de esta tarea. Ahora se hace factible incluso la transposición de componentes seleccionados de música polifónica, permitiendo aplicaciones que, por ejemplo, alteran el modo clave (por ejemplo de C mayor a C menor) de una señal de música dada (ver por ejemplo, de S. Disch y B. Edler, "Análisis, procesamiento y Síntesis de de modulación perceptual multibanda, de señales de audio" (*"Multiband perceptual modulation analysis, processing and Synthesis of audio signals"*) Acta de IEEE–ICASSP, 2009). Esto es posible debido al hecho de que cada portadora componente corresponde estrechamente al tono percibido en su región espectral. Si sólo portadoras que se relacionan con ciertos tonos originales son mapeadas hacia nuevos valores blanco, se manipula el carácter musical global que está determinado por el modo clave.

50 **[0043]** El procesamiento necesario sobre los componentes de MODVOC está representado en la Figura 16 como se mencionó antes. Dentro del dominio de descomposición MODVOC, las frecuencias de la portadora son cuantificadas

a notas MIDI las cuales subsiguientemente son mapeadas a correspondientes notas MIDI apropiadas. Para una reasignación con sentido de los tonos midi y nombres de nota, se puede requerir un conocimiento a-priori del modo y la clave de la pieza de música original. No se actúa en absoluto sobre la AM de todas las componentes porque no contienen información de tono.

5 [0044] Específicamente, las frecuencias de la portadora de componente  $f$ , las cuales representan el tono de componente, son convertidas a valores de tono MIDI  $m$  de acuerdo con la Ecuación 6, donde  $f_{std}$  denota el tono estándar que corresponde al tono MIDI 69, la nota A0.

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

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

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

$$n \rightarrow n'$$

10 (6)

$$n \rightarrow n'$$

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

15 (7)

[0045] Subsiguientemente, son cuantificados tonos MIDI a notar MIDI  $n(f)$  y, adicionalmente, se determina el corrimiento de tono  $o(f)$  de cada nota. Mediante la utilización de una tabla de mapeo de nota MIDI que es dependiente de la clave, el modo original y el modo blanco, estas notas MIDI son transformadas en apropiados valores blanco  $n'$ . En la Tabla de abajo se da un mapeo ejemplar para clave de C de mayor a menor natural. La tabla muestra una tabla de mapeo de notas MIDI para una transformación de modo de escala desde C mayor a C menor natural. El mapeo se aplica para notas de todas las octavas.

20

Nota original	Nota blanco
C	C
D	D
E	E <sub>b</sub>
F	F
G	G
A	A <sub>b</sub>
B	B <sub>b</sub>

25 [0046] Finalmente, las notas MIDI mapeadas que incluyen sus corrimientos de tono son convertidas de vuelta a frecuencia  $f'$  para obtener las frecuencias de la portadora modificadas que son usadas para síntesis (Ecuación 7). Adicionalmente, para conservar la relativa profundidad de modulación FM, la FM d una componente mapeada es multiplicada por el factor de transposición de tono individual el cual es obtenido como el cociente de la frecuencia de la portadora original y modificada. Puedo no requerirse una detección de inicio/desplazamiento de nota MIDI especializada ya que las características temporales están predominantemente representadas por la AM no modificada y por ende se conservan.

30

[0047] El vocoder de modulación descrito es una posibilidad de modificar diferentes rangos de frecuencia (señales pasabanda) de señal de audio diferentemente, lo que se mencionó como transposición selectiva de tono. El concepto inventivo permite el mejoramiento de la calidad perceptual de tales señales de audio modificadas. A pesar de que algunas realizaciones del concepto inventivo se describen en relación con un vocoder o un vocoder de modulación, se puede usar también en general para mejorar la calidad perceptual de señal de audio modificadas independientemente del uso de un vocoder.

35 40

[0048] La Figura 1 muestra un diagrama de bloques de un aparato 100 para modificar una señal de audio 102 de acuerdo con una realización del invento. El aparato 100 comprende un procesador de banco de filtros 110, un elemento de determinación de fundamental 120, un elemento de determinación de sobretono 130, un procesador de señal 140 y un combinador 150. El procesador de banco de filtros 110 se conecta al elemento de determinación de fundamental 120, el elemento de determinación de sobretono 130 y al procesador de señal 140 así como también el

elemento de determinación de fundamental 120 se conecta al elemento de determinación de sobretono 130 y al procesador de señal 140. Además, el elemento de determinación de sobretono 130 se conecta al procesador de señal 140 y el procesador de señal 140 se conecta al combinador 150. El procesador de banco de filtros 110 genera una pluralidad de señales pasabanda 112 en base a una señal de audio 102. Además, el elemento de determinación de fundamental selecciona una señal pasabanda 112 de la pluralidad de señales pasabanda para obtener una señal pasabanda fundamental 122. El elemento de determinación de sobretono identifica una señal pasabanda 112 de la pluralidad de señales pasabanda que cumplen un criterio de sobretono con relación a la señal pasabanda fundamental 122 seleccionada para obtener una señal pasabanda de sobretono 132 asociada a la señal pasabanda fundamental seleccionada 122. Además, el procesador de señal 140 modifica la señal pasabanda fundamental seleccionada 122 en base a un blanco de modificación predefinido. Adicionalmente, el procesador de señal 140 modifica una señal pasabanda de sobretono 132 identificada asociada a la señal pasabanda fundamental 122 seleccionada dependiendo de la modificación de la señal pasabanda fundamental seleccionada 122. El combinador 150 combina la pluralidad de señales pasabanda que contienen la señal pasabanda fundamental seleccionada modificada y la señal pasabanda de sobretono identificada modificada para obtener una señal de audio modificada 152.

**[0049]** Modificando la señal pasabanda fundamental 122 y la señal pasabanda de sobretono identificada 132 asociada con la señal pasabanda fundamental 122 de la misma manera, se puede conservar un comportamiento común de estos armónicos, a pesar de que otras señales pasabanda de la pluralidad de señales pasabanda pueden ser modificadas de diferentes maneras. De esta manera se puede mantener el timbre de la señal de audio original 102 con mayor exactitud, de modo que se puede mejorar significativamente a calidad perceptual de la señal de audio modificada. Por ejemplo, la mayoría de los instrumentos excitan sonidos armónicos que consisten en una parte de frecuencia fundamental y sus armónicos. Si la parte de frecuencia fundamental debe ser modificada, entonces una modificación correlacionada de los armónicos de acuerdo con el concepto descrito puede proveer una calidad perceptiva significativamente mejor de la señal de audio modificada. Además, la señal de audio puede ser modificada en tiempo real, ya que puede no ser necesaria una información a-priori acerca de la señal de audio completa (por ejemplo, el título de música polifónico completo).

**[0050]** 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 representando una señal de audio de entrada en el dominio del tiempo.

**[0051]** El elemento de determinación de fundamental 120 puede proveer la señal pasabanda fundamental seleccionada 122 al procesador de señal 140 para modificación o puede proveer una señal de disparo 122 (por ejemplo, un índice  $i \in [0 \dots I - 1]$  de la señal pasabanda fundamental seleccionada, en donde I es el número de señales pasabanda de la pluralidad de señales pasabanda) para disparar el procesador de señal 140 para modificar la señal pasabanda seleccionada de la pluralidad de señales pasabanda de acuerdo con el blanco de modificación predefinido. Consecuentemente, también el elemento de determinación de sobretono 130 puede proveer la señal pasabanda de sobretono identificada 132 para modificación al procesador de señal 140 o puede proveer una señal de disparo 132 (por ejemplo, un índice que indica la señal pasabanda de la pluralidad de señales pasabanda que está siendo identificada como señal pasabanda de sobretono) para disparar el procesador de señal 140 para modificar la señal pasabanda identificada de la pluralidad de señales pasabanda.

**[0052]** El criterio de sobretono puede comprender una o más reglas para identificar un sobretono de la fundamental. Puede haber uno o más criterios de sobretono a ser cumplidos para identificar una señal pasabanda de la pluralidad de señales pasabanda como un sobretono de la señal pasabanda fundamental seleccionada 122.

**[0053]** El blanco de modificación predefinido puede ser diferente para señales pasabanda que comprenden diferente rango 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 debe ser cambiada hacia una clave blanco. Se dio un mapeo ejemplar para la clave de C desde mayor a menor natural mediante la tabla de arriba. Por ejemplo, si un rango de frecuencia de una señal pasabanda de la pluralidad de señales pasabanda corresponde a una nota original C, la nota blanco sería C también, de modo que esta señal pasabanda no es modificada (excepto por ser identificada como señal pasabanda de sobretono de una señal pasabanda fundamental asociada, la cual es modificada). En este caso, el blanco de modificación se mantiene esta señal pasabanda no modificada. Por el otro lado, una señal pasabanda de la pluralidad de señales pasabanda que comprende un rango de frecuencia que se correlaciona con una nota original A puede ser modificada, de modo que la señal pasabanda modificada puede contener un rango de frecuencia que se correlaciona con una nota blanco Ab (excepto el caso que la señal pasabanda es identificada como una señal pasabanda de sobretono de una señal pasabanda fundamental a ser modificada de acuerdo con otro blanco de modificación). Además, las señales pasabanda de sobretono (señales pasabanda que comprende un rango de frecuencia correlacionado con un sobretono de una nota original A) pueden ser modificadas de modo que la señal pasabanda de sobretono modificada comprende un rango de frecuencia correlacionado con un sobretono de la nota blanco Ab.

[0054] Todas las señales pasabanda 112 de la pluralidad de señales pasabanda pueden comprender una frecuencias de la portadora. La frecuencias de la portadora puede ser una frecuencia características del rango de frecuencia representado o contenido por una señal pasabanda, como, por ejemplo, una frecuencia media del rango de frecuencia, una frecuencia de corte superior del rango de frecuencia, una frecuencia de corte inferior del rango de frecuencia o un centro de gravedad del rango de frecuencia de la señal pasabanda. La frecuencias de la portadora de la señal pasabanda puede ser diferente de la frecuencias de la portadora de cada una de las otras señales pasabanda. Estas frecuencias de la portadora pueden ser usadas por el elemento de determinación de sobretono 130 para identificar las señales pasabanda de sobretono. Por ejemplo, el elemento de determinación de sobretono 130 puede comparar la frecuencias de la portadora de una señal pasabanda 112 de la pluralidad de señales pasabanda con la frecuencias de la portadora de la señal pasabanda fundamental seleccionada 122. Como un sobretono puede ser aproximadamente un múltiplo de la frecuencia fundamental, un criterio de sobretono puede ser cumplido si la frecuencias de la portadora de la señal pasabanda 112 es un múltiplo de la frecuencias de la portadora de la señal pasabanda fundamental seleccionada 122 (con una tolerancia de frecuencias de la 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 de la portadora de una señal pasabanda 112 es un múltiplo de la frecuencia de la portadora de la señal pasabanda fundamental seleccionada 122 con una tolerancia de frecuencias de la portadora predefinida.

[0055] Adicionalmente o alternativamente, el elemento de determinación de sobretono 130 puede comparar un contenido de energía de la señal pasabanda 112 de la pluralidad de señales pasabanda con un contenido de energía de la señal pasabanda fundamental seleccionada 122. En este ejemplo un criterio de sobretono puede ser cumplido si un cociente del contenido de energía de la ratio 112 y el contenido de energía de la señal pasabanda fundamental seleccionada 122 puede estar dentro de un rango de tolerancia de energía predefinido. Este criterio de sobretono toma en cuenta que usualmente los armónicos exhiben menor energía que las fundamentales. El rango de tolerancia de energía predefinido puede ser, por ejemplo, desde 0,3 a 0,9, desde 0,5 a 0,8, desde 0,6 a 0,7 u otro rango. Este criterio de sobretono basado en contenido de energía puede ser combinado con el criterio de sobretono basado en frecuencias de la portadora mencionado arriba.

[0056] Adicionalmente o alternativamente, el elemento de determinación de sobretono 130 puede calcular un valor de correlación indicando una correlación de una envolvente temporal de la ratio 112 de la pluralidad de señales pasabanda con una envolvente temporal de la señal pasabanda fundamental seleccionada 122. En este caso se puede cumplir un criterio de sobretono si el valor de correlación es mayor que un umbral de correlación predefinido. Este criterio de sobretono considera el hecho de que una fundamental y su armónica comparten una envolvente temporal bastante similar. El umbral de correlación predefinido puede ser, por ejemplo 0,2; 0,3; 0,4 ó más. La correlación descrita basada en criterio de sobretono puede ser combinada con el criterio de sobretono basado en frecuencias de la portadora y/o el criterio de sobretono basado en contenido de energía mencionado arriba.

[0057] El elemento de determinación de fundamental 120 puede seleccionar una señal pasabanda 112 adicional de la pluralidad de señales pasabanda sin considerar todas las señales pasabanda fundamentales ya seleccionadas 122 y todas las señales pasabanda de sobretono ya identificadas 132. En otras palabras, el elemento de determinación de fundamental 120 puede seleccionar señales pasabanda fundamentales 122 iterativamente de un conjunto señales pasabanda, el cual contiene señales pasabanda que no son señales pasabanda fundamentales ya seleccionadas ni señales pasabanda de sobretono ya identificadas. Esto puede hacerse hasta que todas las señales pasabanda de la pluralidad de señales pasabanda pueden ser seleccionadas como una señal pasabanda fundamental o identificadas como un sobretono de una señal pasabanda fundamental. Consecuentemente, el elemento de determinación de sobretono 130 puede identificar una señal pasabanda 112 de la pluralidad de señales pasabanda que cumplen un criterio de sobretono con relación a la señal pasabanda fundamental adicional seleccionada sin considerar todas las señales pasabanda de sobretono ya identificadas y sin considerar todas las señales pasabanda fundamentales ya seleccionadas 122. Además, el procesador de señal 140 puede modificar la señal pasabanda fundamental seleccionada 122 adicional en base al blanco de modificación predefinido adicional e independiente de toda otra señal pasabanda fundamental seleccionada. En otras palabras, para cada señal pasabanda fundamental seleccionada o para algunas señales pasabanda fundamentales seleccionadas, se pueden definir diferentes blancos de modificación. Por ejemplo, los blancos de modificación pueden ser definidos mediante una tabla mencionada arriba indicando una transición desde un nota clave a otra. Como las señales pasabanda fundamentales pueden ser modificadas independientemente una de otra, por ejemplo, también selectivamente sólo fundamentales y armónicos de un instrumento específico pueden ser modificados para cambiar el modo clave o el volumen de este instrumento.

[0058] La señal pasabanda 112 puede ser seleccionada mediante el elemento de determinación de fundamental 120 en base a un criterio de energía. Por ejemplo, se puede seleccionar la señal pasabanda con el más alto o uno de los más altos contenidos de energía (por ejemplo, más alto que 70% o más de las otras señales pasabanda). En este ejemplo, una señal pasabanda fundamental ya seleccionada puede ser excluida de una selección ulterior fijando el parámetro de contenido de energía que indica el contenido de energía de la señal pasabanda fundamental seleccionada, igual a cero. Para la selección de la señal pasabanda 112, el contenido de energía de cada señal

pasabanda (indicado por ejemplo mediante un parámetro de contenido de energía determinado por el elemento de determinación de fundamental) puede ser ponderado (por ejemplo, mediante un ponderador) para enfatizar la selección de señales pasabanda perceptualmente importantes.

5 **[0059]** El procesador de señal 140 puede modificar las señales pasabanda fundamentales seleccionadas 132 y las  
 10 señales pasabanda de sobretono asociadas 132 de diversas maneras. Por ejemplo, el procesador de señal 140  
 puede modificar la señal pasabanda fundamental seleccionada 122 multiplicando una frecuencias de la portadora de  
 la señal pasabanda fundamental seleccionada 122 con un factor de transposición (por ejemplo, dependiendo del  
 cambio de modo clave) o sumando una frecuencia de transposición a la frecuencias de la portadora de la señal  
 pasabanda fundamental seleccionada 122. Además, el modificador de señal 140 puede modificar la señal  
 pasabanda de sobretono identificada 132 multiplicando una frecuencias de la portadora de la señal pasabanda  
 identificada 132 con el factor de transposición (por ejemplo, con una tolerancia de 20%, 10%, 5%, 1% o más baja) o  
 sumando un múltiplo de la frecuencia de transposición (por ejemplo con una tolerancia de 20%, 10%, 5%, 1% o más  
 15 baja) a la frecuencias de la portadora de la señal pasabanda de sobretono identificada 132. En otras palabras, por  
 ejemplo, un cambio de modo clave puede ser realizado multiplicando la fundamental y los armónicos asociados por  
 el mismo factor de transposición o sumando una frecuencia de transposición a la fundamental y un múltiplo de la  
 frecuencia de transposición al sobretono. De esta manera, la señal pasabanda de sobretono identificada 132 es  
 modificada dependiendo (de alguna manera) según la señal pasabanda fundamental seleccionada 122.

20 **[0060]** La Figura 2 muestra un diagrama de bloques de un aparato 200 para modificar una señal de audio 102 de  
 acuerdo con una realización del invento. El aparato 200 es similar al aparato mostrado en la Figura 1, pero  
 comprende adicionalmente un elemento de determinación de frecuencias de la portadora 260 y el portador de bando  
 de filtros 110 comprende un banco de filtros 212 y un conversor de señal 214. El banco de filtros 212 se conecta al  
 conversor de señal 214 y el conversor de señal 214 se conecta al procesador de señal 140. El elemento de  
 25 determinación de frecuencias de la portadora opcional 260 se conecta al bando de filtros 212 del portador de banco  
 de filtros 110 y al procesador de señal 140.

**[0061]** El banco de filtros 212 puede generar señales pasabanda en base a la señal de audio 102 y el conversor de  
 30 señal 214 puede convertir las señales pasabanda generadas a un dominio de subbanda para obtener la pluralidad  
 de señales pasabanda provista al elemento de determinación de fundamental 120, al elemento de determinación de  
 sobretono 130 y al procesador de señal 140. El conversor de señal 214 puede ser realizado, por ejemplo, como una  
 unidad de transformación de Fourier discreta inversa de un solo lado, de modo que cada señal pasabanda 112 de la  
 pluralidad de señales pasabanda puede representar una señal analítica. En este dominio de subbanda, el elemento  
 de determinación de fundamental 120 puede seleccionar una de estas señales pasabanda del dominio subbanda de  
 35 la pluralidad de señales pasabanda para obtener la señal pasabanda fundamental 122. Además, el elemento de  
 determinación de sobretono puede identificar una de estas señales pasabanda del dominio subbanda de la  
 pluralidad de señales pasabanda.

**[0062]** Adicionalmente, el elemento de determinación de frecuencias de la portadora 260 puede determinar una  
 40 pluralidad de frecuencias de la portadora en base a la señal de audio 102 y el banco de filtros 212 del procesador de  
 banco de filtros 110 puede generar las señales pasabanda, de modo que cada señal pasabanda comprende un  
 rango de frecuencia que contiene una frecuencias de la portadora diferente 262 de la pluralidad de frecuencias de  
 portadora para obtener una señal pasabanda asociada a cada frecuencias de la portadora 262 de la pluralidad de  
 45 frecuencias de la portadora. En otras palabras, el ancho de banda y las frecuencias medias de las señales  
 pasabanda generadas por el banco de filtros 212 pueden ser controlados por el elemento de determinación de  
 frecuencias de la portadora 260. Esto se puede hacer de varias maneras, por ejemplo, calculando centros de  
 gravedad (COG) de la señal de audio 102, como se describió arriba.

**[0063]** Como ya se mencionó arriba, las señales pasabanda 112 pueden ser modificadas de diversas maneras. Por  
 50 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 pasabanda 112 de la pluralidad de señales pasabanda. Como cada  
 señal pasabanda representa una señal analítica en el dominio de subbanda, el procesador de señal 140 puede  
 generar la señal de modulación de amplitud y la señal de modulación de frecuencia como se mencionó antes en  
 relación con el vocoder de modulación, por ejemplo. Además, el procesador de señal 140 puede modificar la señal  
 55 de modulación de amplitud o la señal de modulación de frecuencia de la señal pasabanda fundamental seleccionada  
 122 en base al blanco de modificación predefinido y puede modificar la señal de modulación de amplitud de la señal  
 de modulación de frecuencia de la señal pasabanda de sobretono identificada 132 asociada a la señal pasabanda  
 fundamental seleccionada 122 dependiendo de la modificación de la señal pasabanda fundamental seleccionada  
 122.

60 **[0064]** El procesador de bando de filtros 110, el elemento de determinación de fundamental 120, el elemento de  
 determinación de sobretono 130, el procesador de señal 140, el combinador 150 y/o el elemento de determinación  
 de frecuencias de la portadora 260 pueden ser, por ejemplo, unidades de hardware individuales o parte de un  
 procesador de señal digital, una computadora o un microcontrolador así como también un programa de computadora

o producto de software configurado para correr en un procesador de señal digital, una computadora o un microcontrolador.

5 **[0065]** Algunas realizaciones de acuerdo con el invento se refieren a un procedimiento 300 para modificar una señal de audio de acuerdo con una realización del invento. El procedimiento 300 puede comprender generar 310 una pluralidad de señales pasabanda en base a una señal de audio y seleccionar 320 una señal pasabanda de la pluralidad de señales pasabanda para obtener una señal pasabanda fundamental. Además, el procedimiento 300 puede comprender identificar 330 una señal pasabanda de la pluralidad de señales pasabanda que cumplen un criterio de sobretono con relación a la señal pasabanda fundamental seleccionada para obtener una señal pasabanda de sobretono asociada a la señal pasabanda fundamental seleccionada. Además, la señal pasabanda fundamental seleccionada es modificada 340 en base a un blanco de modificación predefinido y la señal pasabanda de sobretono identificada asociada a la señal pasabanda fundamental seleccionada es modificada 350 dependiendo de la modificación de la señal pasabanda fundamental seleccionada. Además, el procedimiento 300 puede comprender combinar 360 la pluralidad de señales pasabanda que contienen la señal pasabanda fundamental seleccionada y la señal pasabanda de sobretono identificada modificada para obtener una señal de audio modificada.

20 **[0066]** Opcionalmente, el procedimiento 300 puede comprender pasos adicionales que representan rasgos opcionales del concepto inventivo mencionado arriba y mencionado en lo que sigue.

**[0067]** En lo que sigue se ilustra el concepto descrito en más detalle mediante un ejemplo para una implementación que usa un vocoder de modulación, a pesar de que el concepto propuesto puede ser usado también más en general también para otras implementaciones.

25 **[0068]** La mayoría de los instrumentos excitan sonidos armónicos que consisten en una parte de frecuencia fundamental y sus armónicos que son aproximadamente múltiplos enteros de la frecuencia fundamental. Como los intervalos musicales siguen una escala logarítmica, cada sobretono armónico se parece a intervalo musical diferente con respecto a la fundamental (y sus octavas). La tabla de abajo lista la correspondencia de número armónicos e intervalos musicales para los primeros siete armónicos. La tabla muestra número de armónicos e intervalos musicales relacionados con respecto a la fundamental y sus octavas.

Número de armónico			Nombre de intervalo
1	2	4	unisono perfecto (P1)
			segundo menor (m2)
		9	segundo mayor (M2)
			tercer menor (m3)
	5		tercer mayor (M3)
			cuarto perfecto (P4)
			tritono
	3	6	quinto perfecto (P5)
			sexto menor (m6)
			sexto mayor (M6)
		7	septimo menor (m7)
			septimo mayor (M7)

35 **[0069]** Así, para la tarea de transposición selectiva de contenido de música polifónica, existe una ambigüedad inherente con respecto a la función musical de un componente MODVOC. Si el componente se origina de una fundamental, tiene que ser transpuesto de acuerdo con el mapeo de escala deseado, si es dominado por un armónico a ser atribuido a una fundamental tiene que ser transpuesto junto con esta fundamental para la mejor conservación del timbre original del tono. De esto emerge la necesidad de una asignación de cada componente MODVOC (señal pasabanda) para seleccionar el factor de transposición más apropiada.

40 **[0070]** Para lograrlo, el simple esquema de procesamiento presentado antes, fue extendido mediante una funcionalidad de bloque armónico. El bloqueo armónico examina todas las componentes de MODVOC antes de la transposición si un componente (señal pasabanda) ha de ser atribuido a una fundamental o ha de ser considerado como una entidad independiente. Esto puede ser realizado mediante un algoritmo iterativo. El diagrama de flujo de este algoritmo está representado en la Figura 5. El algoritmo evalúa 510 cocientes de frecuencias, cocientes de energía y correlaciones cruzadas de envolvente de un componente de prueba t (señal pasabanda fundamental) con respecto a toda otra componente (señales pasabanda) con índice  $i \in [0 \dots l - 1] \setminus t$  con l que denota el número total de componentes (número de señales pasabanda de la pluralidad de señales pasabanda). La sucesión de componentes de prueba (señal pasabanda fundamental) durante la iteración está determinada por su energía A-ponderada 520 tal que el orden de evaluación está en secuencia de energía decreciente. La A-ponderación (ANSI, "Ansi standard sl.4-1983," 1983.), (ANSI, "Ansi standard s1.42-2001," 2001.) se aplica para modelar el prominencia

perceptual de cada componente en términos de su volumen (ver por ejemplo, de H. Fletcher y W.A. Munson, "Volumen, su definición, medición y cálculo" ("Loudness, its definition, measurement and calculation") J. Acúst Soc Amer., vol. 5, pp. 82–108, 1933).

5 **[0071]** Mediante segmentación con umbral (*thresholding*) se puede examinar una coincidencia de frecuencias de la portadora armónica, falta de coincidencia de frecuencias de la portadora armónica, una energía de componente y/o una correlación de envolvente de amplitud normalizada a retardo cero.

10 **[0072]** La coincidencia y la falta de coincidencia de frecuencia puede ser definida de acuerdo con la Ecuación 8 con  $f_t$  que es la frecuencias de la portadora de componente de prueba (frecuencias de la portadora de la señal pasabanda fundamental seleccionada) y  $f_i$  que la componente con índice  $i$  (una señal pasabanda de la pluralidad de señales pasabanda). Para la coincidencia de frecuencia, todos los múltiplos mayores que 1 son armónicos potenciales. Un valor de umbral adecuado (umbral de frecuencias de la portadora) para la falta de coincidencia de frecuencia permitido para una armónico potencial es, por ejemplo, 22 Hz.

$$match_i = \text{round} \left( \frac{f_i}{f_t} \right)$$

$$mismatch_i = |f_i - (match_i \cdot f_t)|$$

20 (8)

**[0073]** Se puede requerir que el cociente de energía de componente a–ponderado (Ecuación 9) de armónicos contra fundamental sea menor que un umbral predefinido reflejando el hecho de que para la gran mayoría de instrumentos, los armónicos exhiben menor energía que la fundamental. Un valor de umbral adecuado (rango de tolerancia de energía), por ejemplo, es el cociente de 0,6.

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

30 **[0074]** La correlación cruzada de retardo cero normalizada de la envolvente de la componente de prueba  $env_t$  y le envolvente  $env_i$  de la componente con índice  $i$ , está definid por la Ecuación 10. Esta medida explota el hecho de que una fundamental y sus armónicos comparten una envolvente temporal bastante similar dentro de la longitud de bloque  $M$ . Se determinó que un valor de umbral adecuado (umbral de correlación) es 0,4 mediante experimentos informales.

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

40 **[0075]** Después de ser examinadas, todas las componentes  $i$  que cumple 570 todas las condiciones de umbral son rotuladas 580 como armónicos para ser bloqueadas con respecto a la componente de prueba y subsiguientemente son retiradas de la búsqueda. A continuación. la componente de prueba también es excluida de posteriores iteraciones poniendo 542 su energía en cero. El algoritmo es repetido hasta que todas las componentes han sido asignadas, lo cual es indicado por la máxima energía de componente que es cero.

45 **[0076]** La Figura 4 muestra el esquema de procesamiento mejorado de transposición selectiva mediante el MODVOC que incorpora bloqueo armónico. En contraste con la Figura 16, sólo las componentes no bloqueadas entran a la etapa de transposición mientras que las componentes bloqueadas son modificadas en una segunda etapa por el mismo factor de transposición que se aplicó a sus fundamentales atribuidas.

50 **[0077]** En otras palabras, la Figura 5 muestra un diagrama de flujo del bloque armónico descrito (procedimiento 500 para modificar una señal de audio). Las componentes que coinciden con las condiciones de ser armónicos de una fundamental de prueba (señal pasabanda fundamental seleccionada) son rotuladas iterativamente y retiradas del espacio de búsqueda. Para ello, cada señal pasabanda de la pluralidad de señales pasabanda comprende una frecuencias de la portadora, un contenido de energía, y una envolvente temporal o la frecuencias de la portadora, el

contenido de energía y/o la envolvente temporal (parámetros de envolvente temporal) son determinados 510 para cada señal pasabanda de la pluralidad de señales pasabanda. Además, el contenido de energía (parámetro de contenido de energía) de cada señal pasabanda es a-ponderado 520. Luego se selecciona 530 una señal pasabanda fundamental (fundamental de prueba ft) que comprende una energía máxima (parámetro de contenido de energía). Como todas las señales pasabanda fundamentales ya seleccionadas se ponen en cero y todas las señales pasabanda de sobretono identificadas son excluidas del espacio de búsqueda, la señal pasabanda fundamental seleccionada puede comprender un parámetro de contenido de energía igual a cero, de modo que los algoritmos iterativos paran 540 en este punto. En otro caso, se compara 560 la coincidencia (o falta de coincidencia) de frecuencia, el contenido de energía y/o la correlación cruzada de envolvente temporal de la señal pasabanda fundamental seleccionada y del resto de las señales pasabanda de la pluralidad de señales pasabanda. Si se cumple 570 una, algunas o todas las condiciones (criterios de sobretono), la respectiva señal pasabanda es definida 580 como señal pasabanda de sobretono y se pueden generar datos de bloque armónico (por ejemplo almacenando un índice de la señal pasabanda identificada en una lista de sobretonos) así como también la señal pasabanda de sobretono identificada es retirada del espacio de búsqueda. Los datos de bloqueo armónico pueden ser guardados 590 con referencia a la señal pasabanda fundamental seleccionada asociada. Después de identificar todas las señales pasabanda de sobretono de la señal pasabanda fundamental seleccionada, la energía (el parámetro de contenido de energía) de la señal pasabanda fundamental seleccionada es puesta 592 en cero y se selecciona 530 la siguiente señal pasabanda fundamental que comprende la energía más alta.

**[0078]** El procesador de señal puede usar los datos de bloqueo armónico para modificar las señales pasabanda. En la Figura 4 se muestra una posible implementación. En esta implementación, por ejemplo, el procesador de señal comprende un mapeador MIDI 1600 y un modificador de sobretono 400. El mapeador MIDI 1600 puede modificar la frecuencia de la portadora de cada señal pasabanda fundamental seleccionada de acuerdo con el blanco de modificación individual (el cual también puede incluir que una señal pasabanda fundamental no sea modificada). El mapeador MIDI 1600 puede ser implementado, por ejemplo, como se muestra y se describe en la Figura 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 pueden estar conectados al proveedor de modificación de sobretono 430. El multiplicador de sobretono 420 puede multiplicar la frecuencias de la portadora  $f$  de una señal pasabanda de sobretono identificada con el mismo factor de transposición (con tolerancia mencionada arriba) con el que se multiplica la señal pasabanda fundamental asociada, y puede proveer la frecuencias de la portadora modificada  $f'$  al proveedor de modificación de sobretono 430. El controlador de modificación de sobretono 410 puede disparar el proveedor de modificación de sobretono 430 para proveer la frecuencias de la portadora modificada de la señal pasabanda de sobretono identificada, si el modificador de sobretono 400 identifica la frecuencias de la portadora como una frecuencias de la portadora de una señal pasabanda de sobretono identificada (por ejemplo, en base a datos de bloqueo armónico). En otro caso, el proveedor de modificación de sobretono 430 puede proveer la salida del mapeador MIDI 1600. Además, la Figura 4 muestra una implementación del concepto propuesto en un vocoder, de modo que adicionalmente a la frecuencias de la portadora de la señal pasabanda también la correspondiente señal de modulación de frecuencia (FM) también es modificada mediante una multiplicación con un cociente de la frecuencias de la portadora antes de la modificación y la frecuencias de la portadora modificada. Alternativamente, a una modificación de frecuencia o, adicionalmente, a una modificación de frecuencia, el volumen de la señal de audio puede ser modificado selectivamente en función de la señal pasabanda. Para ello, la señal de modulación de amplitud (AM) de una señal pasabanda puede ser modificada.

**[0079]** En otras palabras, la Figura 4 muestra una transposición selectiva mejorada sobre componentes de vocoder de modulación (señales pasabanda) usando bloqueo armónicos (modificando señales pasabanda de sobretono identificadas dependiendo de la modificación de la señal pasabanda fundamental asociada). Sólo frecuencias de la portadora no bloqueadas (las cuales luego pueden ser señales pasabanda fundamentales) son cuantificadas a notas MIDI las cuales son mapeadas a correspondientes notas MIDI apropiadas (de acuerdo con el blanco de modificación individual). Las componentes bloqueadas (señales pasabanda de sobretono identificadas) pueden ser transpuestas mediante multiplicación por el cociente de la frecuencia original y modificada de la fundamental atribuida (pasabanda fundamental asociada).

**[0080]** La Figura 6a muestra un diagrama de bloques de un aparato 600 para modificar una señal de audio de acuerdo con una realización del invento. El aparato 600 comprende un elemento de determinación de modelado de envolvente 610, un procesador de banco de filtros 620, un procesador de señal 630, un combinador 640 y un modelador de envolvente 650. El elemento de determinación de modelado de envolvente 610 está conectado al modelador de envolvente 650, el procesador de banco de filtro 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 modelador de envolvente 650. El elemento de determinación de modelado de envolvente 610 determina coeficientes de modelado de envolvente 612 en base a una señal de audio del dominio de la frecuencia 602 que representa una señal de audio de entrada del dominio del tiempo. Además, el procesador de banco de filtros 620 genera una pluralidad de señales pasabanda 622 en un dominio subbanda en base a la señal de audio del dominio de la frecuencia 602. El

procesador de señal 630 modifica una señal pasabanda 622 del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un blanco de modificación predefinido. Además, el combinador 640 combina por lo menos un subconjunto de la pluralidad de señal pasabanda del dominio subbanda (por ejemplo, que contiene la señal pasabanda del dominio subbanda modificada) para obtener una señal de audio del dominio del tiempo 642. El modelador de envolvente 650 modela una envolvente de la señal de audio del dominio del tiempo 642 en base a los coeficientes de modelado de envolvente 612 para obtener una señal de audio modelada 652.

**[0081]** Alternativamente, el modelador de envolvente 650 puede estar ubicado entre el procesador de señal 630 y el combinador 640 (el procesador de señal 630 se conecta al modelador de envolvente 650 y el modelador de envolvente 650 se conecta al combinador 640) y puede modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda que contiene la señal pasabanda del dominio subbanda modificada en base a los coeficientes de modelado de envolvente 612.

**[0082]** Extrayendo los coeficientes de modelado de envolvente 612 antes de que la señal de audio sea procesada de manera selectiva como señal pasabanda y usando los coeficientes de modelado de envolvente 612 para modelar la envolvente de la señal de audio después de modificar una o más señales pasabanda, se puede conservar con más exactitud la coherencia espectral de señales pasabanda modificadas diferentemente. Además, especialmente para señales con componente transitorio, también una dispersión de ruido de cuantificación con el tiempo puede ser modelado mediante el modelador de envolvente 650. De esta manera se puede mejorar significativamente la calidad perceptual de la señal de audio modificada. Además, la señal de audio puede ser modificada en tiempo real, ya que puede no ser necesaria una información a-priori acerca de la señal de audio completa (por ejemplo, el título de música polifónico completo).

**[0083]** Además alternativamente, el modelador de envolvente 650 puede estar ubicado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 se conecta al modelador de envolvente 650 y el modelador de envolvente 650 se conecta al procesador de señal 630) y puede modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base a los coeficientes de modelado de envolvente 612 antes de que una señal pasabanda del dominio subbanda sea modificada mediante el procesador de señal 630 para obtener una señal de audio modelada 652.

**[0084]** Extrayendo los coeficientes de modelado de envolvente 612 antes de que la señal de audio sea procesada de manera selectiva como señal pasabanda y usando los coeficientes de modelado de envolvente 612 para modelar la envolvente de la pluralidad de señales pasabanda 622 después que la pluralidad de señales pasabanda 622 es generada por el procesador de banco de filtros 620 en el dominio subbanda, se puede implementar un banco de filtros adaptable, el cual puede incrementar la coherencia local especialmente para señales que contienen componente transitorio (ver, por ejemplo, de J. Herre y J. D. Johnston, "Un banco de filtros continuamente señal-adaptable para codificación de audio de alta calidad perceptual" ("*A continuously signal-adaptive filterbank for high-quality perceptual audio coding*") Taller IEEE ASSP sobre Aplicaciones de Procesamiento de Señal a Audio y Acústica, Mohonk, 1997). En este caso, no se modela la señal modificada (o la señal pasabanda modificada), sino que se puede incrementar la calidad de las señales pasabanda generadas en términos de reproducción componentes transitorios.

**[0085]** La señal de audio del dominio de la frecuencia 602 puede ser provista, por ejemplo, desde un preprocesador que genera la señal de audio del dominio de la frecuencia 602 en base a una señal de audio de entrada del dominio del tiempo (por ejemplo, mediante una transformación de Fourier discreta) o puede ser provista desde una unidad de almacenamiento. Los coeficientes de modelado de envolvente 612 determinados por el elemento de determinación de modelado de envolvente 610 pueden ser, por ejemplo, coeficientes de predicción lineal u otros coeficientes que parametrizan el espectro de la señal de audio del dominio de la frecuencia 602.

**[0086]** El procesador de señal 630 puede modificar una, algunas o todas las señales pasabanda del dominio subbanda 622 de la pluralidad de señales pasabanda del dominio subbanda. El blanco de modificación predefinido puede ser diferente, por ejemplo, para todas o para algunas señales pasabanda del dominio subbanda. Por ejemplo, para cambiar un modo clave de la señal de audio, los blancos de modificación predefinidos de las señales pasabanda del dominio subbanda pueden ser definidos como ya se mencionó con relación a la tabla de arriba.

**[0087]** La señal de audio del dominio de la frecuencia 602 puede comprender líneas espectrales obtenidas, por ejemplo, mediante transformación de Fourier. La diferencia entre líneas espectrales de la señal de audio del dominio de la frecuencia (las cuales también pueden ser supuestas como señales pasabanda) y una señal pasabanda generada por el procesador de banco de filtros 620 puede ser que una línea espectral de la señal de del dominio de la frecuencia 602 representa un ancho de banda más angosto que un ancho de banda representado por una señal pasabanda del dominio subbanda 622 generada por el procesador de banco de filtros 620. Por ejemplo, la señal de audio del dominio de la frecuencia 602 indica un espectro de frecuencia obtenido mediante una transformación de Fourier discreta, la cual es dividida en la pluralidad de señales pasabanda mediante el procesador de banco de filtros 620, en donde un número de señales pasabanda (por ejemplo, 10, 16, 20 ó más) de la pluralidad de señales

pasabanda, es significativamente menor que un número de valores espectrales o líneas espectrales del espectro de frecuencia (por ejemplo, 512 o más valores espectrales).

5 **[0088]** El elemento de determinación de modelado de envolvente 610 puede determinar los coeficientes de modelado de envolvente en base a una predicción sobre la frecuencia de la señal de audio del dominio de la frecuencia 602, lo cual puede ser realizado, por ejemplo, como ya se mencionó, mediante una determinación de los coeficientes de predicción lineal.

10 **[0089]** El procesador de banco de filtros 620 puede proveer la pluralidad de señales pasabanda, representando cada señal pasabanda 622 un rango de frecuencia específico de la señal de audio del 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 pasabanda 622 en base a una señal de audio residual 722 como se muestra en la Figura 7. Para ello, el filtro de predicción 710 puede generar una señal de audio de predicción 712 en base a una señal de audio del dominio de la frecuencia 602 y los coeficientes de modelado de envolvente 612 (por ejemplo, filtro de predicción lineal). Además, el sustractor de señal 720 puede sustraer la señal de audio de predicción 712 de la señal de audio del dominio de la frecuencia 602 para obtener una señal de audio residual 722. La señal de audio residual 722 puede ser usada por el banco de filtros 730 para generar señales pasabanda para obtener la pluralidad de señales pasabanda.

20 **[0090]** Además, el procesador de banco de filtros 620 puede comprender un conversor de señal opcional. Este conversor de señal (por ejemplo, transformador de Fourier discreta inversa de un solo lado) puede convertir las señales pasabanda generadas por el banco de filtro 730 al dominio subbanda para obtener la pluralidad de señales pasabanda 622. Alternativamente, el conversor de señal también puede ser parte del procesador de señal 630.

25 **[0091]** En algunas realizaciones de acuerdo con el invento, una parte de baja frecuencia de la señal de audio de entrada puede ser excluida de una posible modificación para evitar una generación de artefactos en la parte de baja frecuencia de la señal de audio modificada. Para ello, un aparato 680 para modificar una señal de audio puede comprender un filtro pasa alto/pasa bajo, como por ejemplo se muestra en la Figura 6b. El filtro pasa alto/pasa bajo 660 filtra pasa alto la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo, de modo que el elemento de determinación de modelado de envolvente 610 determina los coeficientes de modelado de envolvente 612 en base a la señal de audio del dominio de la frecuencia para alto 602 y el procesador de banco de filtros 620 genera la pluralidad de señales pasabanda 622 en un dominio subbanda en base a la señal de audio del dominio de la frecuencia pasa alto 602. Además, el filtro pasa alto/pasa bajo 660 filtra pasa bajo la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo para obtener una señal de audio de banda completa 670 configurado para combinar la señal de audio modelada 652 y la señal de audio de pasa bajo 662 para obtener una señal de audio de banda completa. En otras palabras, el filtro pasa alto/pasa bajo 660 puede separar la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo en una señal de audio para alto y una señal de audio para bajo. La señal de audio de pasa alto o una representación del dominio de la frecuencia de la señal de audio de pasa alto, puede ser provista al elemento de determinación de modelado de envolvente 610 y al procesador de banco de filtros 620. Esto depende de si el filtro pasa alto/pasa bajo está implementado en el dominio del tiempo seguido por un procesador de señal que genera la señal de audio del dominio de la frecuencia en base a la señal de audio para alto, o el filtro pasa alto/pasa bajo está implementado en el dominio de la frecuencia que recibe ya una señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo.

50 **[0092]** El filtro pasa alto/pasa bajo 660 puede filtrar la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo, de modo que la señal de audio de pasa bajo contiene frecuencias hasta una frecuencia umbral predefinida (por ejemplo, 100 Hz o más). Consecuentemente, la señal de audio de pasa alto puede comprender frecuencias bajas hasta la frecuencia umbral predefinida. En otras palabras, las frecuencias mayores que la frecuencia umbral predefinida pueden ser atenuadas mediante el filtro pasa alto/pasa bajo 660 para proveer la señal de audio de pasa bajo 662 y las frecuencias menores que la frecuencia umbral predefinida pueden ser atenuadas por el filtro pasa alto/pasa bajo 660 para proveer la señal pasa alto.

60 **[0093]** Alternativamente, el modelador de envolvente 650 se ubica entre el procesador de señal 630 y el combinador 640 como se muestra en la Figura 6c. En este caso, el filtro pasa alto/pasa bajo 660 provee la señal de audio de pasa bajo al combinador 640. El combinador 640 combina la pluralidad de señales pasabanda del dominio subbanda que contienen la señal pasabanda del dominio subbanda modificada y la señal de audio de pasa bajo 662 para obtener una señal de audio del dominio del tiempo 642. En este caso, el modelador de envolvente 650 puede determinar un conjunto de coeficientes de modelado de envolvente pasabanda en base a los coeficientes de modelado de envolvente 612 (por ejemplo, mediante el conversor de coeficientes 810) para cada señal pasabanda

del dominio subbanda que corresponde a la respectiva señal pasabanda del dominio subbanda (por ejemplo, que corresponde a la región de frecuencia contenida por la respectiva señal pasabanda del dominio subbanda). Entonces, por ejemplo, cada muestra de tiempo de la señal pasabanda del dominio subbanda puede ser multiplicada con un coeficiente de modelado de envolvente del correspondiente conjunto de coeficientes de modelado de envolvente. Por ejemplo, en la implementación de vocoder mostrada en la Figura 15, el modelador de envolvente 650 puede ser ubicado entre el multiplicador 1550 y el combinador 1560.

**[0094]** Además alternativamente, el modelador de envolvente 650 puede estar ubicado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 se conecta al modelador de envolvente 650 y el modelador de envolvente 650 se conecta al procesador de señal 630) y puede modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base a los coeficientes de modelado de envolvente 612 antes de que una señal pasabanda del dominio subbanda sea modificada mediante el procesador de señal 630 para obtener una señal de audio modelada 652.

**[0095]** En algunas realizaciones de acuerdo con el invento, una parte de baja frecuencia de la señal de audio de entrada puede ser excluida del modelado de envolvente para evitar una generación de artefactos en la parte de baja frecuencia de la señal de audio modificada. Para ello, un aparato 680 para modificar una señal de audio puede comprender un filtro pasa alto/pasa bajo, como por ejemplo se muestra en la Figura 6d. El filtro pasa alto/pasa bajo 660 filtra pasa alto la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo. Además, el filtro pasa alto/pasa bajo 660 filtra pasa bajo la señal de audio de entrada del dominio del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo para obtener una señal de audio de pasa bajo 662. El elemento de determinación de modelado de envolvente 610 determina los coeficientes de modelado de envolvente 612 en base a la señal de audio del dominio de la frecuencia para alto 602 sin considerar la señal de audio de pasa bajo 622. El procesador de banco de filtros 620 genera la pluralidad de señales pasabanda 622 en un dominio subbanda en base a la señal de audio del dominio de la frecuencia pasa alto 602 y la señal de audio de pasa bajo 622. Si se usa un filtro de predicción, como por ejemplo se muestra en la Figura 7, se provee sólo la señal de audio del dominio de la frecuencia de pasa alto 602 al filtro de predicción y al sustractor de señal para generar una señal de audio residual de pasa alto. La señal de audio pasa bajo 622 puede ser provista directamente al banco de filtros para generar señales pasabanda del dominio subbanda. El procesador de señal 630 puede modificar una señal pasabanda del dominio subbanda que corresponde a la señal de audio del dominio de la frecuencia de pasa alto 602 o a la señal de audio de pasa bajo 622. Alternativamente, el procesador de señal 630 puede modificar una señal pasabanda del dominio subbanda que corresponde a la señal de audio del dominio de la frecuencia de pasa alto 602 y una señal pasabanda del dominio subbanda correspondiente a la señal de audio de pasa bajo 622. El combinador 640 puede combinar sólo la señal pasabanda del dominio subbanda que corresponde a la señal de audio del dominio de la frecuencia de pasa alto 602, de modo que sólo las señales pasabanda del dominio subbanda que corresponden a la señal de audio del dominio de la frecuencia de pasa alto 602 (y no las señales de pasabanda del dominio subbanda correspondientes a la señal de audio de pasa bajo 622) pueden ser modeladas por el modelador de envolvente 650.

**[0096]** Además, el aparato 680 comprende un proveedor de señal de banda completa 670 configurado para combinar la señal de audio modelada 652 y las señales pasabanda del dominio subbanda que corresponden a la señal de audio de pasa bajo 662 para obtener una señal de audio de banda completa. Para ello, el procesador de señal 630 puede proveer las señales pasabanda del dominio subbanda que corresponden a la señal de audio de pasa bajo 662 al proveedor de señal de banda completa 670.

**[0097]** Alternativamente, el modelador de envolvente 650 se ubica entre el procesador de señal 630 y el combinador 640. En este caso, el procesador de señal 630 puede proveer las señales pasabanda del dominio subbanda que corresponden a la señal de audio de pasa bajo 662 al combinador 640. El combinador 640 combina la pluralidad de señales pasabanda del dominio subbanda (las señales pasabanda del dominio subbanda que corresponden a la señal de audio de pasa bajo 662 y las señales pasabanda del dominio subbanda que corresponden a la señal de audio del dominio de la frecuencia de pasa alto 602) que contienen la señal pasabanda del dominio subbanda modificada para obtener una señal de audio del dominio del tiempo 642. En este caso, el modelador de envolvente 650 puede determinar un conjunto de coeficientes de modelado de envolvente pasabanda en base a los coeficientes de modelado de envolvente 612 (por ejemplo, mediante el conversor de coeficientes 810) para cada señal pasabanda del dominio subbanda que corresponde a la respectiva señal pasabanda del dominio subbanda (por ejemplo, que corresponde a la región de frecuencia contenida por la respectiva señal pasabanda del dominio subbanda) de las señales pasabanda del dominio subbanda que corresponden a la señal de audio del dominio de la frecuencia de para alto 602. Entonces, por ejemplo, cada muestra de tiempo de la señal pasabanda del dominio subbanda puede ser multiplicada con un coeficiente de modelado de envolvente del correspondiente conjunto de coeficientes de modelado de envolvente. Por ejemplo, en la implementación de vocoder mostrada en la Figura 15, el modelador de envolvente 650 puede ser ubicado entre el multiplicador 1550 y el combinador 1560.

**[0098]** Además alternativamente, el modelador de envolvente 650 puede estar ubicado entre el procesador de señal 630 y el procesador de banco de filtros 620 (el procesador de banco de filtros 620 se conecta al modelador de envolvente 650 y el modelador de envolvente 650 se conecta al procesador de señal 630) y puede modelar una envolvente de las señales pasabanda del dominio subbanda que corresponden a la señal de audio del dominio de la frecuencia de pasa alto en base a los coeficientes de modelado de envolvente 612 antes de que una señal pasabanda del dominio subbanda sea modificada mediante el procesador de señal 630 para obtener una señal de audio modelada 652.

**[0099]** De esta manera, una porción de baja frecuencia de la señal de audio de entrada, puede ser eximida del modelado de envolvente. Sin embargo, la porción de baja frecuencia es enviada al procesamiento restante (por ejemplo, modificación de una señal pasabanda del dominio subbanda). Además, un filtro de predicción (por ejemplo el mostrado en la Figura 7) sólo puede ser aplicado arriba de la frecuencia umbral predefinida. Alternativamente, si ya se realizó la separación pasa alto/para bajo en el lado del análisis, la envolvente de señal de pasa alto puede ser modificada en el dominio del tiempo mediante una recíproca de los coeficientes de modelado de envolvente.

**[0100]** Por ejemplo, en aplicaciones para transposición selectiva, la ubicación mostrada puede proveer resultados equivalentes a una ubicación después del procesamiento, ya que la AM no puede ser modificada.

**[0101]** De acuerdo con un aspecto, el modelador de envolvente 650 puede determinar un cociente de energía de un contenido de energía EFDAS de la señal de audio del dominio de la frecuencia 602 y el contenido de energía ERAS de la señal de audio residual 722. En base a este cociente de energía, el modelador de envolvente 650 puede interrumpir el modelado de la envolvente de la señal de audio del dominio del tiempo 642, si el cociente de energía es menor que un umbral de energía predefinido PET (0,1; 0,2; 0,5; 0,8; 1; 2 o incluso más o menos)

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

**[0102]** En otras palabras, el contacto del moldeado de envolvente puede ser encendido o apagada señal-adaptablemente dependiendo de la bondad de la predicción. La bondad de la predicción puede medirse mediante la ganancia de predicción la cual puede ser definida para ser el cociente de energía de la señal (señal de audio del dominio de la frecuencia) y el error de predicción (señal de audio residual). Si el modelado de la envolvente de la señal de audio del dominio del tiempo 642 es interrumpido, la señal de audio modelada 652 puede ser igual a la señal de audio del dominio del tiempo 642 provista por el combinador 640.

**[0103]** El modelador de envolvente 650 puede ser implementado de varias maneras. Un ejemplo se muestra en la Figura 8. El modelador de envolvente 650 puede comprender un conversor de coeficientes 810 y un multiplicador 820. El conversor de coeficientes 810 puede convertir los coeficientes de modelado de envolvente 612 al dominio del tiempo, de modo que los coeficientes de modelado de envolvente convertidos 812 pueden ser multiplicados con la señal de audio del dominio de tiempo 642 para modelar la envolvente temporal de la señal de audio del dominio del tiempo y para obtener la señal de audio modelada 652. Esto se puede hacer mediante el multiplicador 820. Por ejemplo, un bloque de tiempo de la señal de audio del dominio del tiempo 642 puede contener 512 (o más) muestras de tiempo y el conversor de coeficientes 810 puede proveer 512 (o más) coeficientes de modelado de envolvente convertidos 812 para multiplicar cada muestra de tiempo con un coeficiente de modelado de envolvente convertido. 812.

**[0104]** Como se ya mencionó, el aparato 600 puede modificar diferentes señales pasabanda del dominio subbanda, diferentemente. Más generalmente, esto significa que el procesador de señal 630 puede modificar una segunda o ulterior señal pasabanda del dominio subbanda 622 de la pluralidad de señales pasabanda del dominio subbanda en base a un segundo o ulterior blanco de modificación predefinido. El ya mencionado o primer blanco de modificación y el ulterior o segundo blanco de modificación predefinido pueden ser diferentes.

**[0105]** En algunas realizaciones, el concepto descrito puede ser usado en relación con a vocoders o vocoders 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 pasabanda del dominio subbanda 622 de la pluralidad de señales pasabanda del dominio subbanda. Además, el procesador de señal 630 puede modificar la señal de modulación de amplitud o la señal de modulación de frecuencia de la señal pasabanda del dominio subbanda a ser modificada en base al blanco de modificación predefinido.

**[0106]** Además, el aparato 600 puede comprender opcionalmente, un elemento de determinación de frecuencias de la portadora como ya se describió para el aparato 200 y mostrado en la Figura 2. El elemento de determinación de frecuencias de la portadora puede determinar una pluralidad de frecuencias de la portadora en base a la señal de audio del dominio de la frecuencia 602. Estas frecuencias de la portadora determinadas pueden ser usadas por el procesador de banco de filtros 620 o en la implementación mostrada en la Figura 7 por el banco de filtros 730 del

procesador de banco de filtros 620 para generar señales pasabanda del dominio subbanda, de modo que cada señal pasabanda del dominio subbanda comprende un rango de frecuencia que contiene una frecuencias de la portadora diferente de la pluralidad de frecuencias de la portadora para obtener una señal pasabanda del dominio subbanda asociada a cada frecuencias de la portadora de la pluralidad de frecuencias de la portadora. Esto se puede hacer, por ejemplo, determinando los centros de gravedad de la señal de audio del dominio de la frecuencia como se mencionó arriba.

**[0107]** El elemento de determinación de fundamental 610, el procesador de banco de filtros 620, el procesador de señal 630, el combinador 640 y/o el modelador de envolvente 650 pueden ser, por ejemplo, unidades de hardware individuales o parte de un procesador de señal digital, una computadora o un microcontrolador así como también un programa de computadora o producto de software configurado para correr en un procesador de señal digital, una computadora o un microcontrolador.

**[0108]** Algunas realizaciones de acuerdo con el invento se refieren a una implementación del concepto descrito en un vocoder de modulación. Para este ejemplo, el concepto está descrito en más detalle a continuación. Los rasgos mencionados también pueden usarse en otras implementaciones o aplicaciones.

**[0109]** Se ha establecido antes que el procesamiento MODVOC conserva la coherencia espectral en el área pasabanda que rodea las ubicaciones de las portadoras. Sin embargo, no se conserva la coherencia espectral global de amplio espectro. Para señales cuasi-estacionarias esto puede tener sólo un impacto menor sobre la calidad perceptual de la señal sintetizada. Si la señal contiene componentes transitorios prominentes, como por ejemplo golpes de tambor o castañuelas, la conservación de la coherencia global puede mejorar mucho la calidad de reproducción de estas señales.

**[0110]** La conservación de la coherencia global puede ser mejorada mediante predicción lineal en el dominio espectral. Algunos enfoques son utilizados para codecos de audio, por ejemplo, mediante la herramienta de modelado de ruido en el tiempo (TNS) (ver, por ejemplo, de J. Herre y J. D. Johnston, "Mejorando el rendimiento de coders de audio perceptual usando modelado de ruido en el tiempo (tns)" ("*Enhancing the performance of perceptual audio coders by using temporal noise shaping (tns)*") 101ra. convención de AES, Los Angeles, N°. Preimpr. 4384, 1996) en codificación de audio avanzada MPEG 2/4 (AAC). En el documentos de J. Herre y J. D. Johnston, "Un banco de filtros continuamente señal-adaptable para codificación de audio de perceptual de alta calidad" ("*A continuously signal-adaptive filterbank for high-quality perceptual audio coding*" Taller de IEEE ASSP sobre Aplicaciones de Procesamiento de Señal para Audio y Acústica, Mohonk, 1997, la combinación de transformación de tiempo-frecuencia de alta resolución y la predicción espectral, muestra corresponder esencialmente a una transformación señal-adaptable.

**[0111]** La Figura 9 da un panorama de la integración del concepto descrito en el esquema de procesamiento MODVOC. En el análisis, subsiguientemente a la DFT inicial de la señal de entrada  $x$ , se establece los coeficientes de predicción lineal (LPC) de un predictor hacia adelante a lo largo de la frecuencia que tiene la respuesta de impulso  $h(\omega)$ , por ejemplo, mediante el procedimiento de autocorrelación que minimiza el error de predicción en un sentido de cuadrados mínimos. Subsiguientemente, se aplica el filtro a los valores espectrales y la señal residual el procesada ulteriormente mediante el algoritmo MODVOC. Los coeficientes de filtro, que representan la envolvente global, son llevados a la etapa de síntesis. En la síntesis, la envolvente global, establecida mediante evaluación del

filtro de predicción sobre el círculo unitario  $\left|H(e^{j\omega})\right|$ , es restablecida mediante aplicación multiplicativa de la misma señal suma que provee la señal de salida y como se ilustra en la Figura 10.

**[0112]** En otras palabras, las Figuras 9 y 10 muestran una implementación del concepto descrito en un vocoder de modulación. La Figura 9 muestra la parte del analizador de modulación que comprende un preprocesador 910, el cual realiza, por ejemplo, una transformación de Fourier discreta de una señal de audio del dominio del tiempo para obtener una señal de audio del dominio de la frecuencia 602 y provee la señal de audio del dominio de la frecuencia 602 al elemento de determinación de modelado de envolvente 610, al filtro de predicción 710 (por ejemplo, filtro LPC  $h(\omega)$ ), al sustractor de señal 710 y al elemento de determinación de frecuencias de la portadora 920. El sustractor de señal 720 puede proveer la señal de audio residual 722 al banco de filtros 730. El elemento de determinación de frecuencias de la portadora 920 puede estimar múltiples frecuencias de centro de gravedad de portadora y proveer estas frecuencias de la portadora al banco de filtros 730 para controlar los pesos espectrales de pasabanda. El banco de filtros 730 puede proveer las señales pasabanda a un conversor de señal 930 que realiza una transformación de Fourier discreta inversa de una solo lado para cada señal pasabanda para proveer la pluralidad de señales pasabanda del dominio subbanda al procesador de señal. Los componentes del vocoder de modulación ya se describieron arriba en mayor detalle. Además la Figura 10 muestra la parte de síntesis del vocoder de modulación. Éste comprende el combinador 640 y el modelador de envolvente que comprende un conversor de coeficiente 810 y un multiplicador 820. Detalles adicionales para los componentes del vocoder de modulación y el modelador de envolvente ya se explicaron arriba.

5 [0113] La Figura 11 muestra un diagrama de flujo de un procedimiento 1100 para modificar una señal de audio de acuerdo con una realización del invento. El procedimiento 1100 comprende determinar 1110 coeficientes de modelador de envolvente en base a una señal de audio del dominio de la frecuencia que representa una señal de audio de entrada del dominio del tiempo y generar 1120 una pluralidad de señales pasabanda en un dominio subbanda en base a la señal de audio del dominio de la frecuencia. Además, el procedimiento 1100 comprende modificar 1130 una señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un blanco de modificación predefinido. Adicionalmente, por lo menos un subconjunto de la pluralidad de señales pasabanda del dominio subbanda es combinado 1140 para obtener una señal de audio del dominio del tiempo. Además, el procedimiento 110 comprende modelar 1150 una envolvente de la señal de audio del dominio del dominio del tiempo en base a los coeficientes de forma de envolvente, modelar 1150 una envolvente de la pluralidad de señales pasabanda del dominio subbanda conteniendo la señal pasabanda del dominio subbanda modificada en base a los coeficientes de forma de envolvente o moldear 1150 una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base los coeficientes de forma de envolvente antes que la señal pasabanda de dominio subbanda sea modificada por el procesador de señal para obtener una señal de audio moldeada.

[0114] Opcionalmente, el procedimiento 1100 puede comprender otros pasos que representan los rasgos del concepto mencionado arriba.

20 [0115] Algunas realizaciones de acuerdo con el invento se refieren a un aparato para modificar una señal de audio que combina los rasgos del aparato mostrado en la Figura 1 o 2 con los rasgos del aparato mostrado en la Figura 6. Oportunamente, la Figura 12 muestra un diagrama de bloques de un aparato 1200 de acuerdo con una realización del invento.

25 [0116] A partir del aparato mostrado en la Figura 1, el aparato 1200 comprende adicionalmente un elemento de determinación de modelado de envolvente 610 y un modelador de envolvente 650. En esta conexión la señal de audio puede ser una señal de audio del dominio de la frecuencia que representa una señal de audio de entrada del dominio del tiempo, la cual puede ser usada por el elemento de determinación de modelado de envolvente para determinar coeficientes de modelado de envolvente en base a la señal de audio del dominio de la frecuencia. Además, la pluralidad de señales pasabanda generada por el banco de filtro puede ser generada en un dominio subbanda en base a la señal de audio del dominio de la frecuencia. Después de combinar la pluralidad de señales pasabanda del dominio subbanda que contienen la señal pasabanda fundamental seleccionada modificada y la señal pasabanda de sobretono identificada modificada, la señal de audio del dominio del tiempo obtenida 152, 642 puede ser provista al modelador de envolvente 650. El modelador de envolvente 650 puede modelar una envolvente de la señal de audio del dominio del tiempo en base a los coeficientes de modelado de envolvente 612 para obtener la señal de audio modelada 652.

40 [0117] De otro modo, a partir del aparato mostrado en la Figura 6, el aparato 1200 comprende adicionalmente un elemento de determinación de fundamental 120 y un elemento de determinación de sobretono 130 como se describe con relación al aparato mostrado en la Figura 1. El elemento de determinación de fundamental 120 puede seleccionar una señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda para obtener las señales pasabanda fundamental 122. Además, el elemento de determinación de sobretono 130 puede identificar una señal pasabanda del dominio subbanda 112 de la pluralidad de señales pasabanda del dominio subbanda que cumplen un criterio de sobretono con relación a la señal pasabanda fundamental 122 seleccionada para obtener una señal pasabanda de sobretono 132 asociada a la señal pasabanda fundamental seleccionada 122. El procesador de señal 140, 630 puede modificar la señal pasabanda de fundamental seleccionada en base a un blanco de modificación predefinido y modificar una señal pasabanda de sobretono identificada 132 asociada a la señal pasabanda fundamental seleccionada 122 dependiendo de la modificación de la señal pasabanda fundamental seleccionada 122 como se mencionó arriba.

50 [0118] De esta manera, sobretonos de fundamentales y sobretonos pueden ser tratados en forma equivalente durante la modificación de la señal de audio y la coherencia espectral de la pluralidad de señales pasabanda puede ser conservada muy exactamente modelando la señal de audio del dominio del tiempo modificada en base a los coeficientes de modelado de envolvente establecidos antes de la modificación de las señales pasabanda. De esta manera se puede mejorar significativamente la calidad perceptual de la señal de audio modificada.

[0119] El aparato 1200 puede realizar otras tareas de diferentes ejemplos de implementación mencionados arriba.

60 [0120] En lo que sigue se muestra la mejora de la calidad perceptual de las señales de audio modificadas mediante los resultados de pruebas de audición. Para esta prueba de audición se usó una implementación basada en vocoder de modulación (MODVOC), pero los resultados también son válidos en general para el concepto propuesto.

[0121] Para evaluar la calidad de audio subjetiva del vocoder de modulación (MODVOC) para la aplicación de transposición selectiva de tono y, asimismo, el mérito de las mejoras propuestas al principio de MODVOC básico, un

conjunto de archivos de audio ha sido ensamblado y procesado correspondientemente. Adicionalmente, la tecnología de MODVOC es comparada con un software de audio disponible para manipulación de audio polifónico. El editor Melodyne de Celemony el cual está en venta desde fines de 2009.

5 **[0122]** Como el procesamiento bajo prueba altera drásticamente el contenido de audio de una señal, una comparación directa de señal original y procesada — usualmente una parte inherente en pruebas de audición estándar — es claramente no conveniente en este caso. Para de todos modos medir la calidad de audio subjetiva de una manera que tenga sentido, se ha aplicado un procedimiento de prueba de audición especial: el conjunto de prueba de audición se origina a partir de datos MIDI simbólicos que se convierten en formas de onda usando un expansor MIDI de alta calidad. Este enfoque permite una comparación directa de archivos de audio similarmente alterados dentro de la prueba y permite una investigación dentro del efecto del procesamiento selectivo de tono en aislamiento. El procedimiento de generar el conjunto de prueba está sintetizado en la Figura 17. Las señales de prueba originales son preparadas en representación de datos MIDI simbólicos (arriba a la izquierda). Una segunda versión de estas señales es generada mediante un procesamiento MIDI simbólico el cual se parece al procesamiento blanco bajo prueba sobre el audio original convertido en forma de onda (arriba a la derecha). Subsiguientemente, estos pares de señales son convertidos mediante el expansor MIDI de alta calidad en archivos de forma de onda (WAV) (abajo izquierda y derecha) En esta prueba de audición, son comparadas la forma de onda convertida proveniente del archivo MIDI procesado y diversas versiones procesadas de vocoder de modulación (MODVOC) del archivo MIDI original (abajo a la derecha). Adicionalmente, la salida del MODVOC es comparada con la salida del editor Melodyne.

15 **[0123]** Aparte de las condiciones procesadas por MODVOC, la prueba incluye una condición obtenida usando el editor Melodyne el cual es actualmente la única aplicación comercial que apunta a este tipo de manipulación de audio y por ende puede verse como el estándar de la industria. El editor Melodyne inicialmente realiza un análisis automático de todo el archivo de audio. Después de la fase de inicialización, Melodyne sugiere una descomposición del archivo de audio. Mediante interacción con el usuario, esta descomposición puede ser más refinada. En aras de una comparación justa con los resultados de procesamiento de MODVOC, la evaluación se basa en el resultado de este análisis inicial automático ya que, aparte del conocimiento a-priori de la clave y el tono estándar, la descomposición de MODVOC también es totalmente automática.

25 **[0124]** La preparación de la prueba de audición se basó en una prueba de Estímulo Múltiple estándar con Referencia Oculta y Ancla (*standard Multiple Stimuli with Hidden Reference and Anchor, (MUSHRA)*) de acuerdo con la recomendación de ITU BS.1534 (ITU-R, "Procedimiento para la evaluación subjetiva de calidad de sonido intermedio" ("*Method for the subjective assessment of intermediate sound quality (mushra)*", 2001). MUSHRA es una prueba de audición a ciegas. Una sola persona por vez es sometida a la prueba. Para cada tema, la prueba presenta todas las condiciones junto con la referencia oculta y un ancla filtrado por pasa bajo oculto al oyente de una manera alineada en el tiempo. La referencia oculta y el ancla inferior se incluyen para verificar la confiabilidad de los oyentes. Está permitido conmutar entre condiciones mientras se escucha y también fijar un bucle sobre particiones seleccionadas arbitrariamente del tema, como se sugiere en BS.1116-1 (ITU-R, "Procedimientos para evaluación subjetiva de pequeños deterioros en sistemas de audio incluyendo Sistemas de sonido multicanal" ("*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 hay límite del número de repeticiones que los sujetos de prueba pueden escuchar antes de calificar el tema y proceder al siguiente tema de prueba, permitiendo así una comparación muy estrecha y a lo largo de toda la examinación de las diferentes condiciones. La calidad perceptual de los temas es calificada en una escala en el rango desde "excelente" (100 puntos) vía "bueno" y "aceptable" hasta "pobre" (0 puntos). La secuencia de temas de prueba es aleatoriamente ordenada y asimismo el orden de las condiciones de cada tema también es randomizado.

30 **[0125]** Los ocho temas de prueba se han tomado del proyecto MUTOPIA (<http://www.mutopiaproject.org/>), el cual provee partituras gratis para uso público. Se han extraído extractos adecuados que tiene una duración aproximada de 20 segundos como máximo de diversas piezas de música clásica, conteniendo tanto instrumentos solos (por ejemplo, G, E) como partes de orquesta completa (por ejemplo F). También se incluyen en el conjunto de prueba, melodías de solo instrumental dominante acompañado por otros instrumentos (por ejemplo C). Además de partes tonales cuasi-estacionarias de corto plazo, también están contenidos elementos percusivos en varios temas (inicios de guitarra punteada en C y piano en G) que plantean un desafío especial en la respuesta de componente transitorio del Sistema bajo prueba. La siguiente tabla lista todos los temas del conjunto.

Nombre	Descripción	Instrumentos	Modo clave
A	Violon concerto J.S. Bach, BWV 1041	Orquesta	Amen
B	Eine kleine Nachtmusik W.A. Mozart, KV 525 Mv1	Cuarteto de cuerda	Gmay
C	Berceuse G. Fauré, Op 56	Flauta y guitarra	Emay
D	Nocturno F. Strauss Op 7	Trompeta y piano	Dbmay

E	Waltz F. Carulli, Op 241 No. 1	Guitarra	Cmay
F	Ein musikalischer Spass W.A. Mozart, KV 522 Mv1	Trompetas, violín, viola y chelo	Fmay
G	Ode an die Freude L. v. Beethoven	Piano	Gmay
H	Piano Trio L. v. Beethoven, Op 11 Mv 3	Clarinete, chelo y pinao	Bbmay

5 [0126] El procesamiento MIDI para obtener las señales transpuestas originales ha sido fabricado en Sonar8 por Cakewalk. Las conversiones de formas de onda de alta calidad se han realizado usando Band-stand de Native Instruments en la versión de biblioteca de sonido 1.0.1 R3. Se evaluó el procesamiento MODVOC en tres diferentes combinaciones con los dos pasos de procesamiento de mejoramiento siendo bloqueo armónico y modelado de envolvente. Se utilizó el editor Melodyne, versión 1.0.11 para comparación. Todas las condiciones están listadas en la tabla de abajo.

Condición	Nombre	Descripción
1	*_reference	Original transpuesto por MIDI
2	*_3k5Hz_reference	Original filtrado por pasa bajo de 3,5 kHz (ancla)
3	*_MODVOC	MODVOC
4	*_MODVOC_harm	MODVOC con un bloqueo armónico
5	*_MODVOC_harm_es	MODVOC con un bloqueo armónico y un modelado de envolvente
6	*_dna	Editor Melodyne (DNA) [0001] Modo completamente automático

10 [0127] Las pruebas de audición subjetiva se realizaron en un laboratorio de audición acústicamente aislado que está diseñado para permitir pruebas de audición de alta calidad en un ambiente similar a un salón "ideal". Los oyentes fueron equipados con auriculares electroacústicos STAX que se accionaron desde una interfaz de sonido Edirol USB conectada a un mini MAC Apple. El Software de prueba de audición fue WavSwitch de Fraunhofer IIS, operado en modo MUSHRA, proveyendo una GUI simple para ayudar al oyente a realizar la prueba. Los oyentes pueden conmutar entre la referencia (1) y las diferentes condiciones (2–7) durante la emisión. Cada oyente puede decidir individualmente cuánto tiempo escuchar cada tema y la condición. Durante la conmutación actual, la reproducción de sonido es silenciada. En el GUI, las Barras verticales visualizan la calificación atribuida a cada condición. Se seleccionaron oyentes experimentados que están familiarizados con codificación de audio pero también tienen conocimiento musical para obtener, por un lado, un juicio educado sobre típicos artefactos de procesamiento de señal tipo pre – y post – ecos o dispersión de componentes transitorios y por el otro lado, parámetros musicales tales como tono espectral melodía y timbre. Además, se les pidió a los oyentes que proveyeran observaciones e impresiones informales.

25 [0128] En total quince personas contribuyeron al resultado de la prueba, mientras que un oyente tuvo que ser retirado después de la audición debido a que obviamente no identificaba exitosamente el original oculto (otorgándole 64 puntos).

30 [0129] La Figura 18 sintetiza los resultados de la prueba de audición. La calidad perceptual para los temas procesados por transposición selectiva de tono está en el rango de aceptable a bueno. El ancla inferior fue calificado entre pobre y malo de modo que la distancia desde los temas procesados y el ancla llega a aproximadamente 40 puntos MUSHRA.

35 [0130] Los puntajes absolutos proveen información que cuantifica la calidad perceptual de cada tema (en cada una de las condiciones de prueba) y de ese modo implícitamente calificar la diferencia de calidad entre los temas en conjunto de prueba, pero no son adecuados para comparar las diferentes condiciones dentro de la prueba de audición ya que las calificaciones de estas condiciones no son independientes. Para una comparación directa de las condiciones que se originan de diferentes esquemas de procesamiento de transposición selectiva, a continuación se consideran diferencias de puntaje.

40 [0131] La Figura 19 representa el resultado en base a diferencias de puntaje de las variantes MODVOC mejoradas (condiciones 4 y 5) con respecto a los resultados de MODVOC básico (condición 3). Aquí, todas las variantes de MODVOC mejoradas tienen puntajes considerablemente mejores que el procesamiento MODVOC básico (todos los puntajes están bien por arriba de cero). Hay significado en el sentido de confianza de 95% para todos los temas y condiciones excepto para la aplicación de bloqueo armónico sólo en el tema A y C.

45

- 5 [0132] La Figura 20 presenta los puntajes de prueba como diferencias de puntajes con respecto a la condición 6 (editor Melodyne). Para el tema C, el MODVOC en la condición 5 obtiene significativamente mejor puntaje que el editor Melodyne mientras que la condición 4, no obstante ser levemente positivo, y la condición 3 no son concluyentes en un sentido de intervalo de confianza de 95% (intervalos de confianza se superponen con 0). Para los temas B (condición 2), F, G (condición 5) tampoco se puede extraer conclusión significativa, pero se puede ver una tendencia a mejor desempeño de MODVOC también para el tema C en la condición 4 y el tema F en las condiciones 4 y 5. En todos los otros casos, MODVOC alcanza puntajes significativamente peores que el editor Melodyne.
- 10 [0133] El puntaje refleja un juicio de calidad total que comprende aspectos como artefactos que suenan no naturales como degradación de componentes transitorios por pre- o post-ecos, exactitud de tono, corrección de melodía y conservación de timbre. Para interpretar los resultados en más detalle, se les pidió a los oyentes que anoten sus observaciones informales junto con la nota del puntaje actual. A partir de estas observaciones se puede concluir que la conservación del timbre y la ausencia de artefactos de sonido no natural estuvo representada en la fuente total hasta un mayor grado que, por ejemplo, la bondad de preservación de melodía. Asimismo, si una cierta melodía es desconocida para el oyente, parece que las personas de la prueba no pudieron memorizar la melodía de referencia con rapidez durante la prueba y por ende estaban inseguros acerca de la melodía real. Ésta puede ser una explicación de la calificación global más alta de los temas procesados por editor Melodyne, que tienen un fidelidad más alta con respecto a conservación de timbre, especialmente de sonidos que se originan de instrumentos solos. Sin embargo esto es a costo de que accidentalmente ocurren severos errores de melodía que pueden pasar presumiblemente debido a mala clasificación. El MODVOC es más robusto en ese sentido ya que no se apoya predominantemente en técnicas de clasificación basadas en rasgos.
- 15 [0134] Algunas realizaciones de acuerdo con el invento se refieren a un vocoder de modulación mejorado para transposición selectiva de tono. Se ha presentado el concepto de vocoder de modulación (MODVOC) y se ha destacado su capacidad general para realizar transposición selectiva sobre contenido de música polifónica. Esto hace posible aplicaciones que apuntan a cambiar el modo clave de muestras de música PCM pre-grabadas. Se proponen dos técnicas de mejoramiento para transposición selectiva de tono mediante el MODVOC. El desempeño de la aplicación de transposición selectiva y el mérito de estas técnicas evaluadas comparativamente mediante resultados obtenidos de una metodología de prueba de audición especialmente diseñada que es capa de gobernar cambios en términos de tono con respecto a los estímulos de audio originales. Los resultados de esta evaluación de calidad perceptual subjetiva son presentados para temas que han sido convertidos entre modo clave menor y mayor mediante el MODVOC y, adicionalmente, mediante el primer software comercialmente disponible el cual también es capa de manejar esta tarea.
- 20 [0135] Vale la pena notar que mientras el editor Melodyne inicialmente realiza un análisis automático de todo el archivo de audio antes de encargarse de cualquier manipulación, el MODVOC opera sobre una base de bloque por bloque permitiendo potencialmente por eso operación en tiempo real.
- 25 [0136] Se han propuesto técnicas de mejoramiento para el vocoder de modulación (MODVOC) para transposición selectiva del tono. A partir de los resultados de la prueba obtenidos para señales de prueba convertidas desde MIDI, se puede concluir que la calidad perceptual del MODVOC básico es mejorada por el bloque armónico y el modelado de envolvente. Sobre todos los temas, se puede esperar un aumento de hasta 10 puntos MUSHRA. Una gran parte del mejoramiento se origina en el bloque armónico.
- 30 [0137] Asimismo, la comparación del MODVOC con un Software comercialmente disponible (editor Melodyne) reveló que el nivel de calidad general que pueden alcanzarse en transposición selectiva de tono, en este momento, puede ser ubicado entre «aceptable» y «bueno». El MODVOC es más robusto para mala interpretación de melodía ya que esencialmente no se apoya principalmente en decisiones de clasificación.
- 35 [0138] En contraste con el análisis multi-paso realizado por el editor Melodyne sobre todo el archivo de audio antes de la manipulación, el MODVOC sólo se basa en un procesamiento a modo de bloque de paso simple que potencialmente permite transmisión o escenarios de operación en tiempo real.
- 40 [0139] A pesar de que algunos aspectos del concepto descrito han sido descritos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a un paso de procedimiento o a un rasgo de un paso de procedimiento. Análogamente, aspectos descritos en el contexto de un paso de procedimiento también representan una descripción de una correspondiente bloque o componente o rasgo de un correspondiente aparato.
- 45 [0140] La señal de audio codificada inventiva puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida a través de un medio de transmisión tal como un medio de transmisión inalámbrico o una medio de transmisión físico tal como Internet.
- 50
- 55
- 60

- 5 [0141] Dependiendo de ciertos requerimientos de implementación, realizaciones del invento pueden ser implementadas en hardware o en software. La implementación puede ser realizada usando una medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, una memoria de sólo lectura, una PROM, una EEPROM o una memoria FLASH, teniendo señales de control legibles electrónicamente almacenadas en las mismas, las cuales cooperan (o son capaces de cooperar) con un sistema de computadora programable tal que se ejecute el respectivo procedimiento. Por lo tanto, el medio de almacenamiento digital puede ser legible por computadora.
- 10 [0142] Algunas realizaciones de acuerdo con el invento comprenden un portador de datos que tiene señales de control legibles electrónicamente, las cuales son capaces de cooperar con una sistema de computadora programable, tal que uno de los procedimientos descrito en la presente sea ejecutado.
- 15 [0143] Generalmente, realizaciones del presente invento pueden ser implementadas como un programa de computador con un código de programa, siendo código de programa operativo para ejecutar uno de los procedimientos cuando el producto de programa de computadora corre en una computadora. El código de programa puede ser almacenado, por ejemplo, sobre un portador legible por una máquina.
- 20 [0144] Otras realizaciones comprenden el programa de computadora para ejecutar uno de los procedimientos descritos en la presente, almacenado en un portador legible por una máquina.
- 25 [0145] En otras palabras, una realización del procedimiento inventivo es, por lo tanto, un programa de computadora que un código de programa para ejecutar uno de los procedimientos descritos en la presente, cuando el programa de computadora corre en una computadora.
- 30 [0146] Una realización adicional de los procedimientos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa de computadora para ejecutar uno de los procedimientos descritos en la presente.
- 35 [0147] Una realización adicional del procedimiento inventivo es, por lo tanto, una transmisión de datos o una secuencia de señales que representan el programa de computador para ejecutar uno de los procedimientos descritos en la presente. La transmisión de datos o la secuencia de señales pueden ser configuradas, por ejemplo, para ser transferidos vía una conexión de comunicación de datos, por ejemplo, vía Internet.
- 40 [0148] Una realización adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los procedimientos descritos en la presente.
- 45 [0149] Una realización adicional comprende una computadora que tiene instalado en ella el programa de computadora para ejecutar uno de los procedimientos descritos en la presente.
- 50 [0150] En algunas realizaciones se puede usar un dispositivo de lógica programable (por ejemplo un arreglo de compuesta programable de campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en la presente. En algunas realizaciones, el arreglo de compuerta programable de campo puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en la presente. Generalmente, los procedimientos preferiblemente son realizados mediante algún aparato de hardware.
- [0151] Las realizaciones que se describieron más arriba son puramente ilustrativas para los principios del presente invento. Se entiende que las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en la presente serán evidentes para los expertos en la materia. Por lo tanto, es la intención que el invento esté limitado sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados por la descripción y la explicación de las realizaciones en la presente.

REIVINDICACIONES

1. Un aparato (600) para modificar una señal de audio, que comprende:

- 5 un elemento de determinación de modelado de envolvente (610) configurado para determinar coeficientes de modelado de envolvente (612) en base a una señal de audio del dominio de la frecuencia (602) que representa una señal de audio de entrada del dominio del tiempo;
- 10 un procesador de banco de filtros (620) configurado para generar una pluralidad de señales pasabanda (622) en un dominio subbanda en base a la señal de audio del dominio de la frecuencia (602);
- un procesador de señal (630) configurado para modificar una señal pasabanda del dominio subbanda (622) de la pluralidad de señales pasabanda del dominio subbanda en base a un blanco de modificación predefinido;
- 15 un combinador (640) configurado para combinar por lo menos un subconjunto de la pluralidad de señales pasabanda del dominio subbanda para obtener una señal de audio del dominio del tiempo (642); y
- 20 un modelador de envolvente (650) configurado para modelar una envolvente de la señal de audio del dominio del tiempo (642) en base a los coeficientes de modelado de envolvente (612), para modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda que contienen la señal pasabanda del dominio subbanda modificada en base a los coeficientes de modelado de envolvente (612) o para modelar una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base a los coeficientes de modelado de envolvente (612) antes que una señal pasabanda de dominio subbanda sea modificada por el
- 25 procesador de señal (630) para obtener una señal de audio modelada (652);

**caracterizado por el hecho de que** el elemento de determinación de modelado de envolvente (610) se configura para determinar los coeficientes de modelado de envolvente (612) en base a una predicción sobre la frecuencia de la señal de audio de dominio de la frecuencia (602).

- 30 2. Un aparato de acuerdo con la reivindicación 1, en donde el modelador de envolvente (650) comprende un conversor de coeficientes (810) y un multiplicador (820), en donde el conversor de coeficientes (810) está configurado para convertir los coeficientes de modelado de envolvente (612) al dominio del tiempo, en donde el multiplicador (820) está configurado para multiplicar la señal de audio del dominio del tiempo (642) con los
- 35 coeficientes de modelado de envolvente (612) convertidos.

3. Un aparato de acuerdo con una de las reivindicaciones 1 a 2, en donde el procesador de banco de filtros (620) comprende un filtro de predicción (710), un sustractor de señal (720) y un banco de filtros (730), en donde el filtro de predicción (710) está configurado para generar una señal de audio de predicción (712) en base a la señal de audio del dominio de la frecuencia (602) y los coeficientes de modelado de envolvente (612), en donde el sustractor de señal (720) está configurado para sustraer la señal de audio de predicción (712) de la señal de audio del dominio de la frecuencia (602) para obtener una señal de audio residual (722), en donde el banco de filtros (730) está configurado para generar señales pasabanda para obtener la pluralidad de señales pasabanda en base a la señal de audio residual (722).

4. Un aparato de acuerdo con la reivindicación 3, en donde el modelador de envolvente (650) está configurado para determinar un cociente de energía de un contenido de energía de la señal de audio del dominio de la frecuencia (602) y un contenido de energía de la señal de audio residual (722), en donde el modelador de envolvente (650) está configurado para interrumpir el modelado de la envolvente de la señal de audio del dominio del tiempo (642), si el cociente de energía es menor que un umbral de energía predefinido.

5. Un aparato de acuerdo con la reivindicación 3 o 4, en donde el procesador de banco de filtros (620) comprende un conversor de señal (930) configurado para convertir las señales pasabanda generadas por el banco de filtros (730) al dominio subbanda para obtener la pluralidad de señales pasabanda.

6. Un aparato de acuerdo con una de las reivindicaciones 1 a 5, en donde el procesador de señal (630) está configurado para modificar una segunda señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un segundo blanco de modificación predefinido, en donde el blanco de modificación predefinido y el segundo blanco de modificación predefinido son diferentes.

7. Un aparato de acuerdo con una de las reivindicaciones 1 a 6, en donde el procesador de señal (630) está configurado para generar una señal de modulación de amplitud (AM) y una señal de modulación de frecuencia (FM) para cada señal pasabanda del dominio subbanda (622) de la pluralidad de señales pasabanda del dominio subbanda, en donde el procesador de señal (630) está configurado para modificar la

señal de modulación de amplitud (AM) o la señal de modulación de frecuencia (FM) de la señal pasabanda del dominio subbanda (622) a ser modificada en base al blanco de modificación predefinido.

5 8. Un aparato de acuerdo con una de las reivindicaciones 1 a 7, que comprende un elemento de  
determinación de frecuencias de la portadora (920) configurado para determinar una pluralidad de  
frecuencias de la portadora en base a la señal de audio del dominio de la frecuencia (602), en donde el  
procesador de banco de filtros (620) está configurado para generar señales pasabanda, de modo que cada  
10 señal pasabanda comprende un rango de frecuencia que contiene una frecuencia de la portadora diferente de  
la pluralidad de frecuencias de la portadora para obtener una señal pasabanda asociada a cada frecuencia de  
la portadora de la pluralidad de frecuencias de la portadora.

15 9. Un aparato de acuerdo con una de las reivindicaciones 1 a 8, que comprende un filtro pasa alto/pasa bajo  
(660) configurado para filtrar pasa alto la señal de audio de entrada del dominio del tiempo o la señal de audio  
del dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo, en donde el  
filtro pasa alto/pasa bajo (660) está configurado para filtrar pasa bajo la señal de audio de entrada del dominio  
del tiempo o la señal de audio del dominio de la frecuencia que representa la señal de audio de entrada del  
dominio del tiempo para obtener una señal de audio pasa bajo (662), en donde el elemento de determinación  
de modelado de envolvente (610) está configurado para determinar los coeficientes de modelado de  
20 envolvente (612) en base a la señal de audio del dominio de la frecuencia (602) de pasa alto, en donde el  
procesador de banco de filtros (620) está configurado para generar la pluralidad de señales pasabanda (622)  
en un dominio subbanda en base a la señal de audio del dominio de la frecuencia (602) de pasa alto y la  
señal de audio de pasa bajo (622), en donde el procesador de señal (630) está configurado para modificar  
una señal pasabanda del dominio de la subbanda asociada a la señal de audio de dominio de frecuencia de  
25 pasa alto (602) o la señal de audio de pasa bajo (622), en donde el combinador (640) está configurado para  
combinar la señal pasabanda del dominio subbanda correspondiente a la señal de audio del dominio de la  
frecuencia (602) de pasa alto para obtener la señal de audio del dominio del tiempo (642) o para combinar la  
señal pasabanda del dominio subbanda correspondiente a la señal de audio del dominio de la frecuencia  
(602) de pasa alto y la señal pasabanda del dominio subbanda (622) correspondiente a la señal de audio de  
30 pasa bajo (662) para obtener la señal de audio del dominio del tiempo (642).

35 10. Un aparato de acuerdo con una de las reivindicaciones 1 a 9, que comprende un filtro pasa alto/pasa bajo  
(660) configurado para filtrar pasa alto la señal de audio de entrada del dominio del tiempo o la señal de audio  
de dominio de la frecuencia que representa la señal de audio de entrada del dominio del tiempo de modo que  
el elemento de determinación de modelado de envolvente (610) determina los coeficientes de modelado de  
envolvente (612) en base a la señal de audio del dominio de la frecuencia (602) de pasa alto y el procesador  
de banco de filtros (620) genera la pluralidad de señales pasabanda (622) en un dominio subbanda en base a  
la señal de audio del dominio de la frecuencia (602) de pasa alto, en donde el filtro pasa alto/pasa bajo (660)  
está configurado para filtrar por pasa bajo la señal de audio de entrada del dominio del tiempo o la señal de  
40 audio del dominio de la frecuencia que representan la señal de audio de entrada del dominio del tiempo para  
obtener una señal de audio de pasa bajo, en donde el combinador (640) está configurado para combinar la  
pluralidad de señales pasabanda del dominio subbanda que contienen la señal pasabanda de dominio de  
subbanda modificada y la señal de audio de pasa bajo para obtener una señal de audio del dominio del  
tiempo (642) o un proveedor de señal de banda completa (670) está configurado para combinar la señal de  
45 audio modelada (652) y la señal de audio de pasa bajo para obtener una señal de audio de banda completa.

50 11. Un aparato de acuerdo con una de las reivindicaciones 1 a 10, que comprende un elemento de  
determinación de fundamental y un elemento de determinación de sobretono, en donde el elemento de  
determinación de fundamental está configurado para seleccionar una señal pasabanda del dominio subbanda  
(622) de la pluralidad de señales pasabanda del dominio subbanda para obtener una señal pasabanda de la  
fundamental, en donde el elemento de determinación de sobretono está configurado para identificar una  
55 señal pasabanda del dominio subbanda (622) de la pluralidad de señales pasabanda del dominio subbanda  
que cumplen un criterio de sobretono con relación a la señal pasabanda de la fundamental seleccionada para  
obtener una señal pasabanda de sobretono asociada a la señal pasabanda de la fundamental seleccionada,  
en donde el procesador de señal (630) está configurado para modificar la señal pasabanda de la fundamental  
seleccionada en base a un blanco de modificación predefinido y configurado para modificar la señal  
pasabanda de sobretono identificada asociada con la señal pasabanda de la fundamental seleccionada  
dependiendo de la modificación de la señal pasabanda de la fundamental seleccionada.

60 12. Un aparato de acuerdo con la reivindicación 11, en donde cada señal pasabanda del dominio subbanda  
(622) comprende una frecuencia de la portadora, en donde el elemento de determinación de sobretono está  
configurado para comparar la frecuencia de la portadora de una señal pasabanda del dominio subbanda  
(622) de la pluralidad de señales pasabanda del dominio subbanda con la frecuencia de la portadora de la  
señal pasabanda de la fundamental seleccionada, en donde se cumple un criterio de sobretono, si la  
65 frecuencia de la portadora de la señal pasabanda del dominio subbanda (622) es un múltiplo de la frecuencia  
de la portadora de la señal pasabanda de la fundamental seleccionada con una tolerancia de frecuencia de la  
portadora predefinida.

13. Un procedimiento (1100) para modificar una señal de audio, que comprende:

determinar (1110) coeficientes de modelado de envolvente en base a una señal de audio del dominio de la frecuencia (602) que representa una señal de audio de entrada del dominio del tiempo;

5

generar (1120) una pluralidad de señales pasabanda en un dominio subbanda en base a la señal de audio del dominio de la frecuencia;

10

modificar (1130) una señal pasabanda del dominio subbanda de la pluralidad de señales pasabanda del dominio subbanda en base a un blanco de modificación predefinido;

combinar (1140) por lo menos un subconjunto de la pluralidad de señales pasabanda del dominio subbanda para obtener una señal de audio del dominio del tiempo; y

15

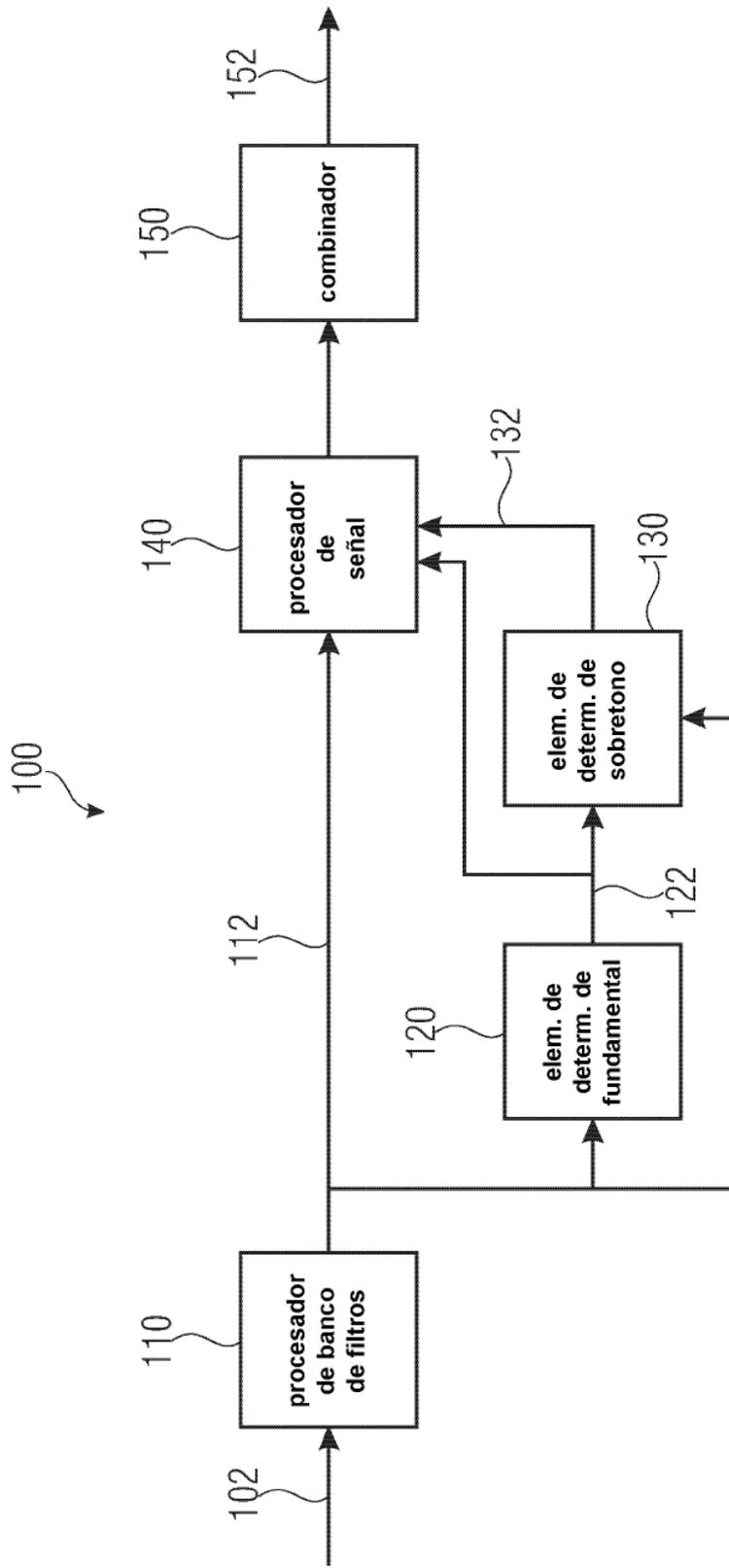
modelar (1150) una envolvente de la señal de audio del dominio del tiempo en base a los coeficientes de modelado de envolvente, modelar (1150) una envolvente de la pluralidad de señales pasabanda del dominio subbanda que contiene la señal pasabanda del dominio subbanda modificada en base a los coeficientes de modelado de envolvente o moldear (1150) una envolvente de la pluralidad de señales pasabanda del dominio subbanda en base los coeficientes de modelado de envolvente antes que la señal pasabanda de dominio subbanda sea modificada por el procesador de señal para obtener una señal de audio moldeada;

20

**caracterizado por el hecho de que** los coeficientes de modelado de envolvente (612) se determinan en base a una predicción sobre la frecuencia de la señal de audio del dominio de la frecuencia (602).

25

14. Un programa de computadora con un código de programa para ejecutar el procedimiento de acuerdo con la reivindicación 13, cuando el programa de computadora corre sobre un procesador de señal digital, una computadora o un microcontrolador.



**FIG. 1**

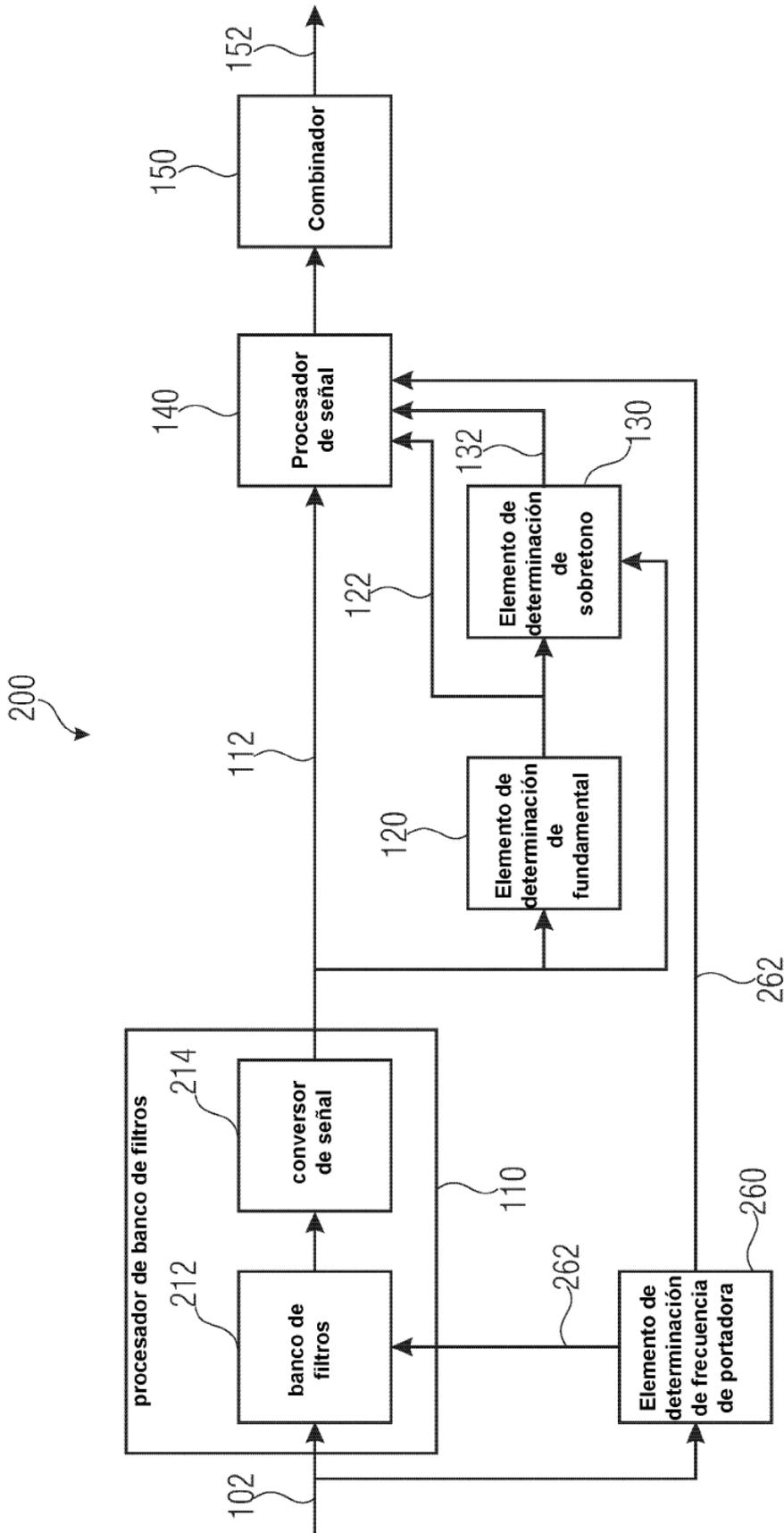
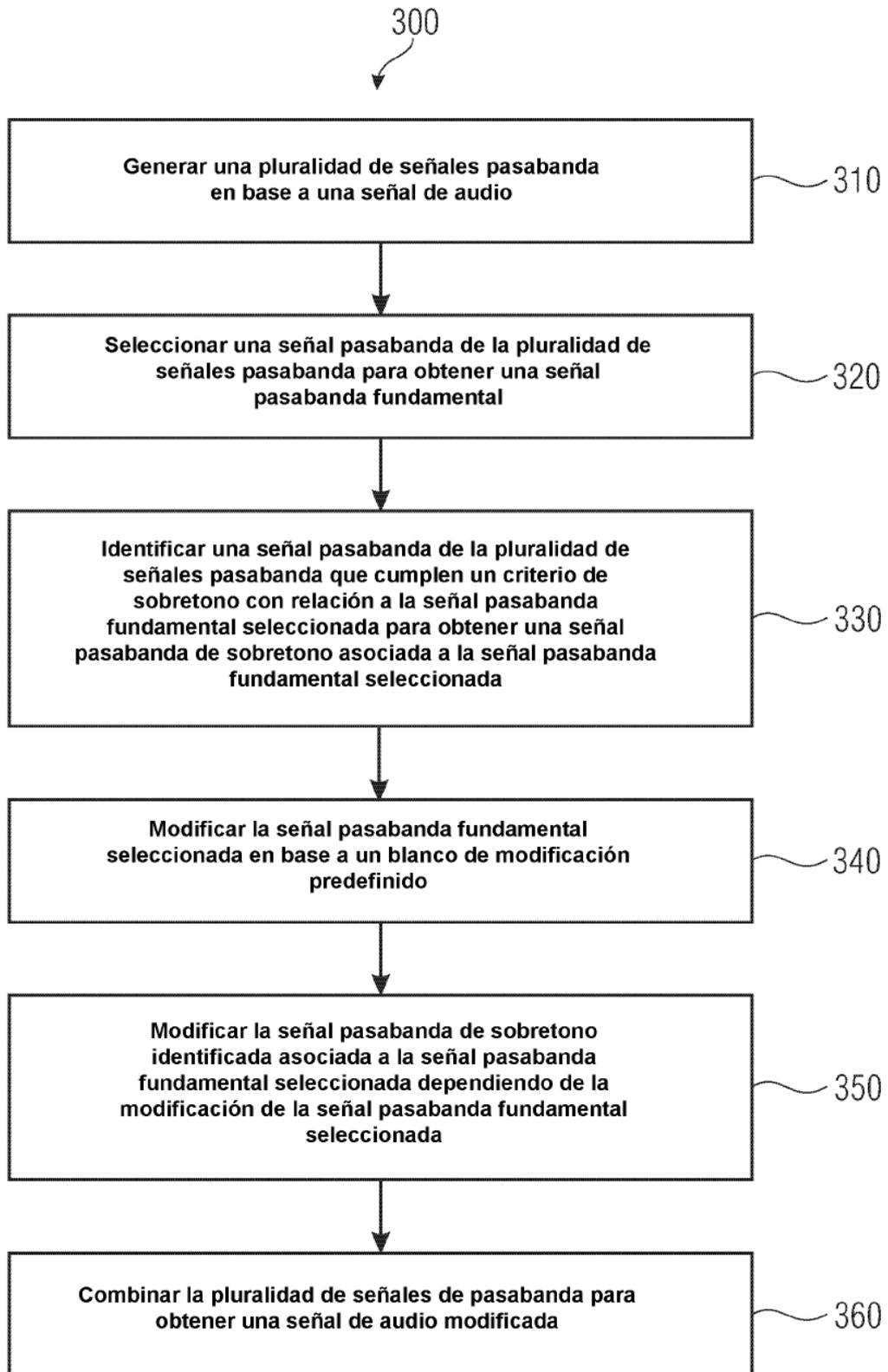


FIG. 2



**FIG. 3**

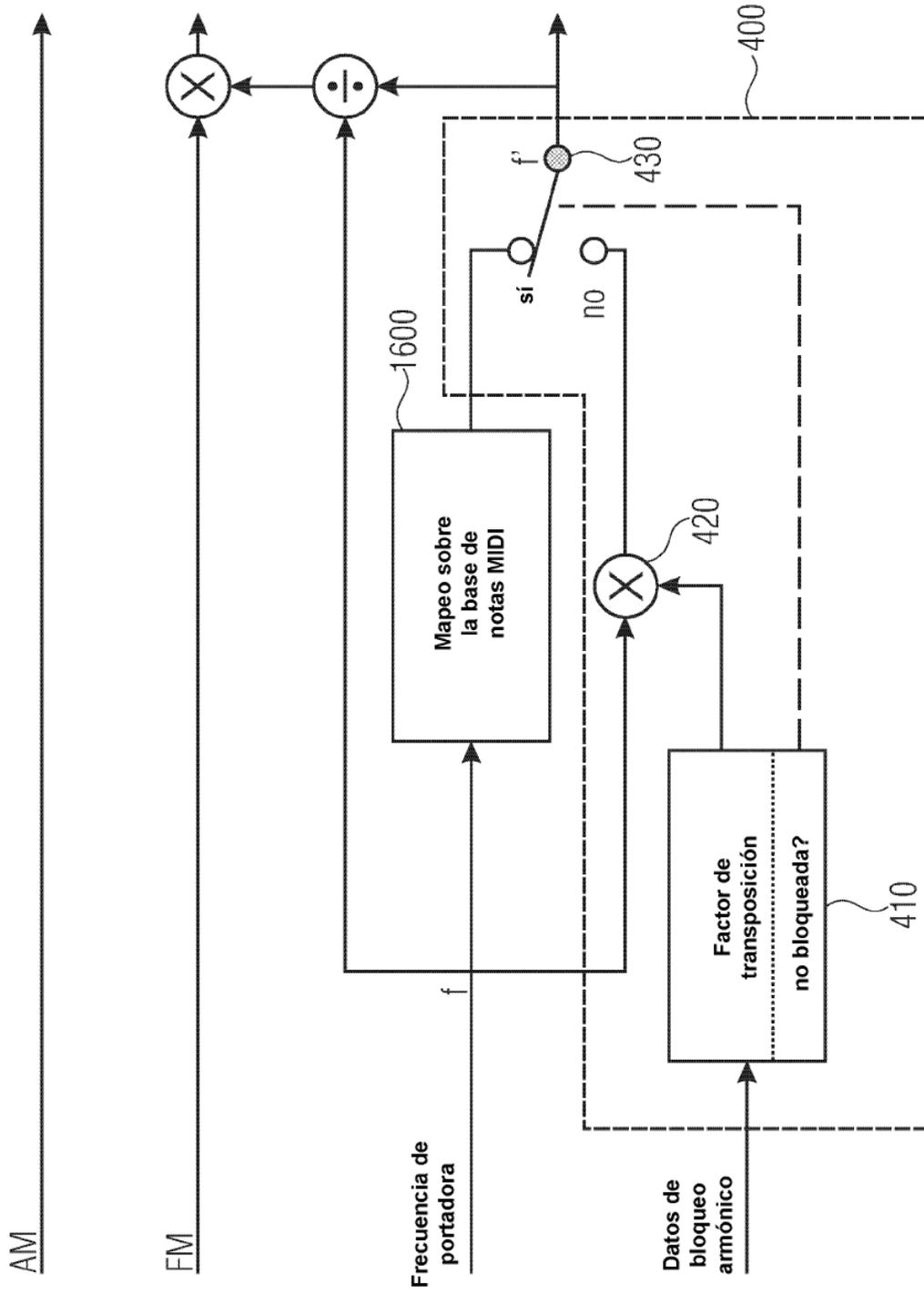


FIG. 4

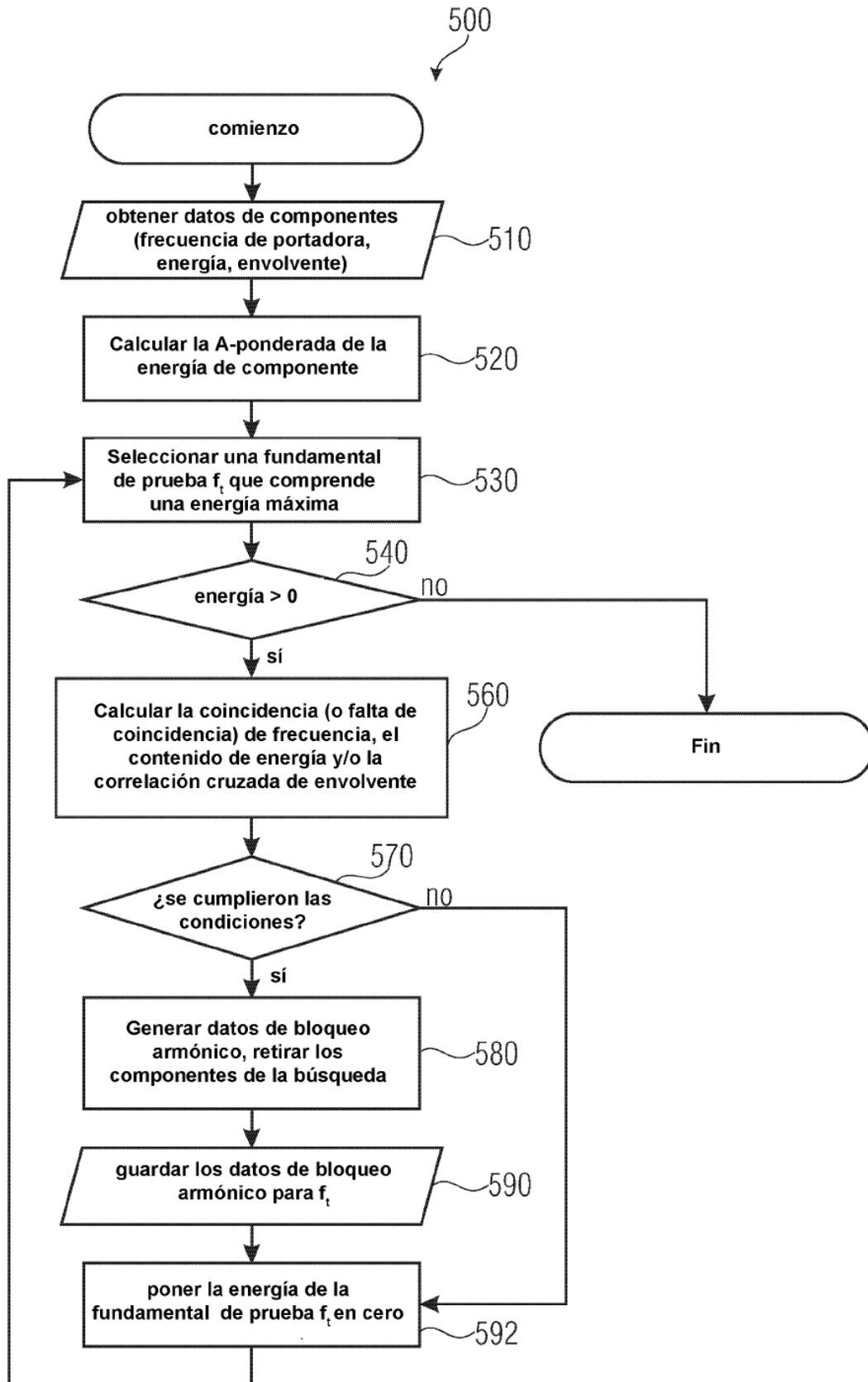


FIG. 5

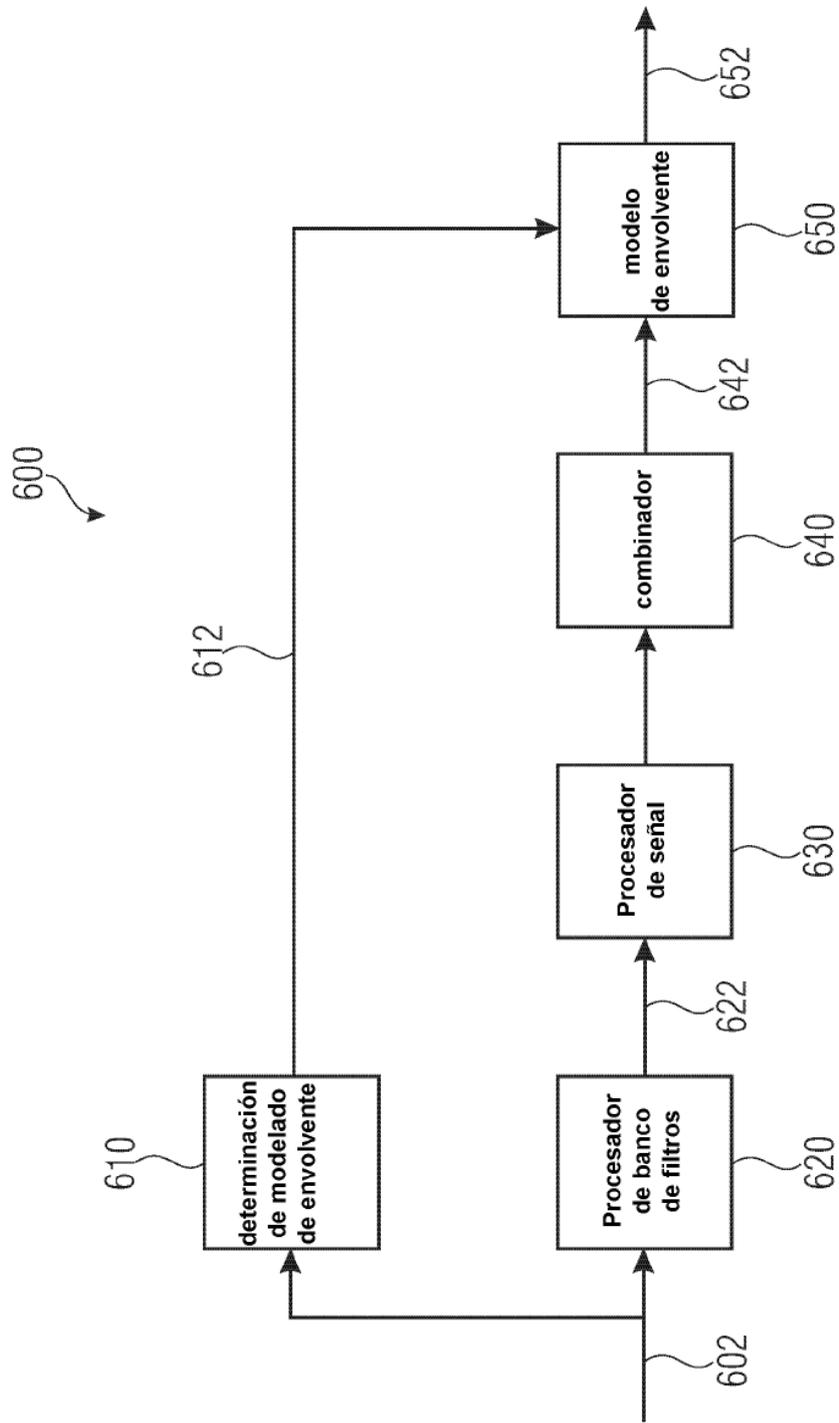


FIG. 6 A

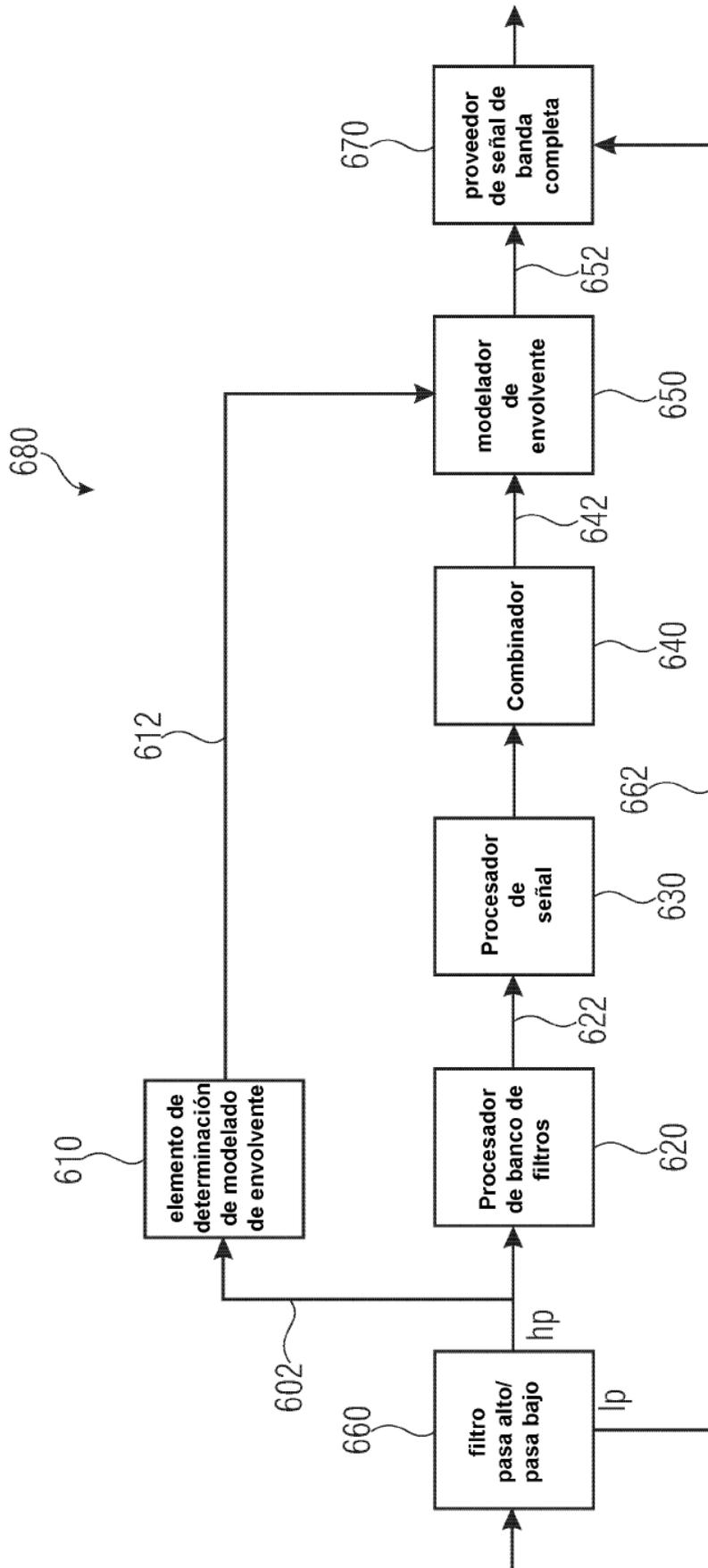


FIG. 6 B

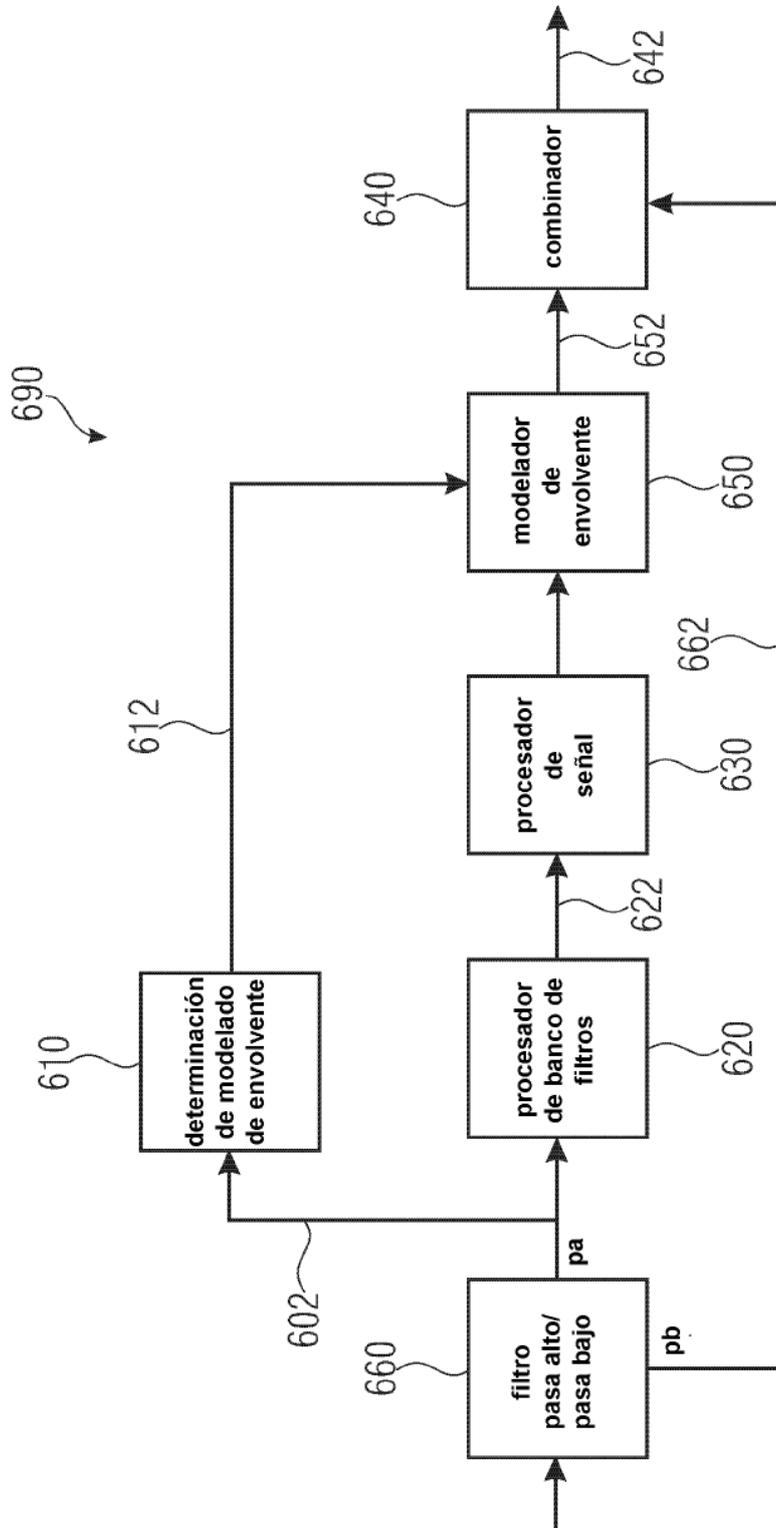


FIG. 6 C

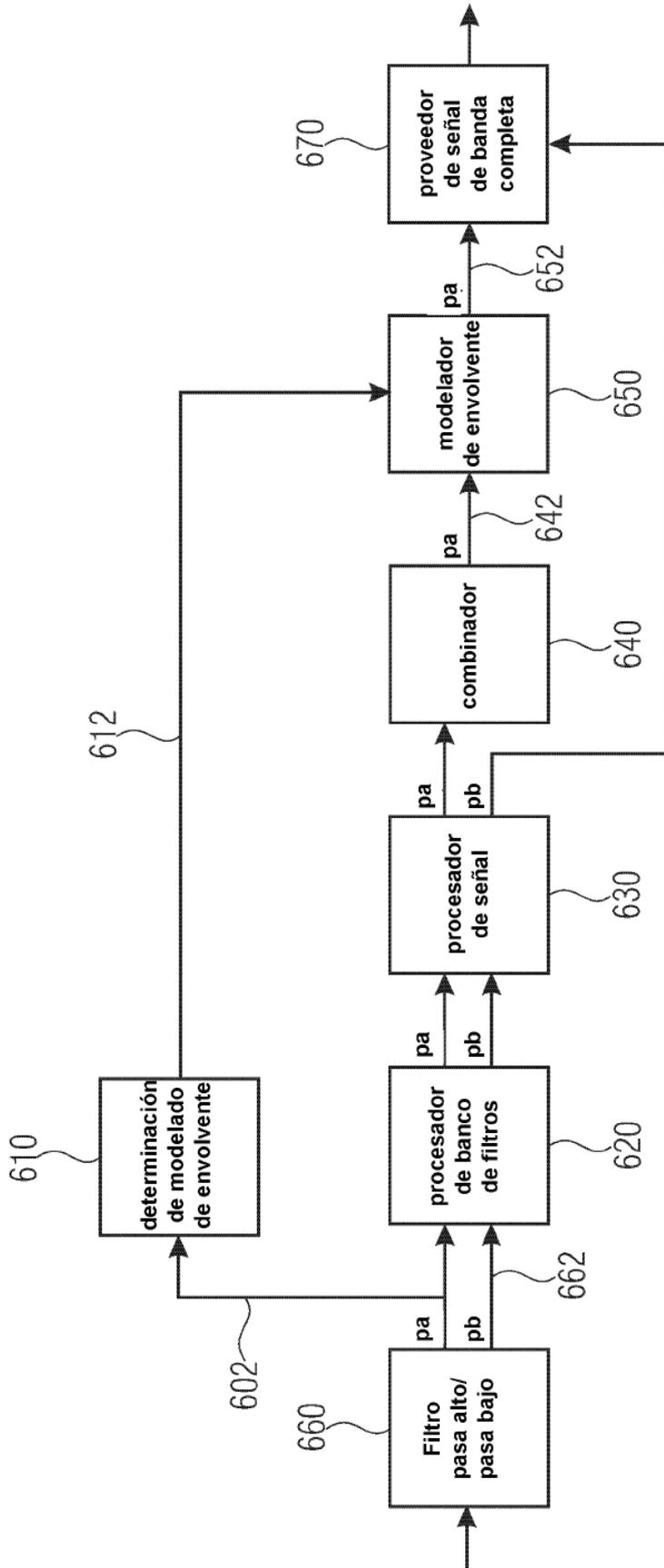
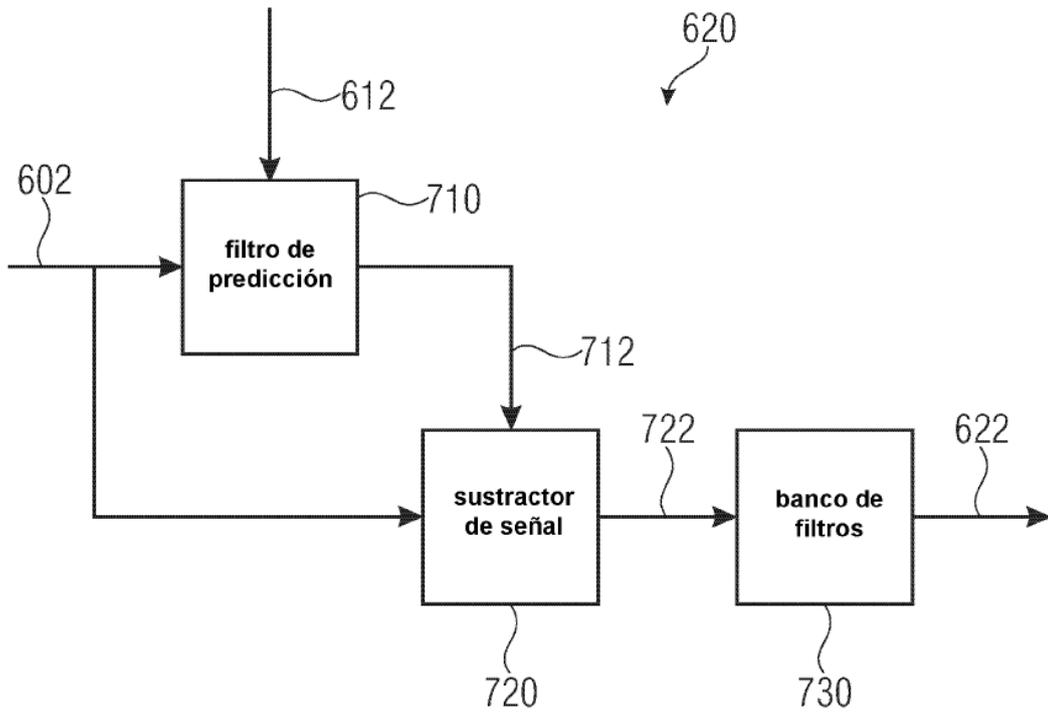
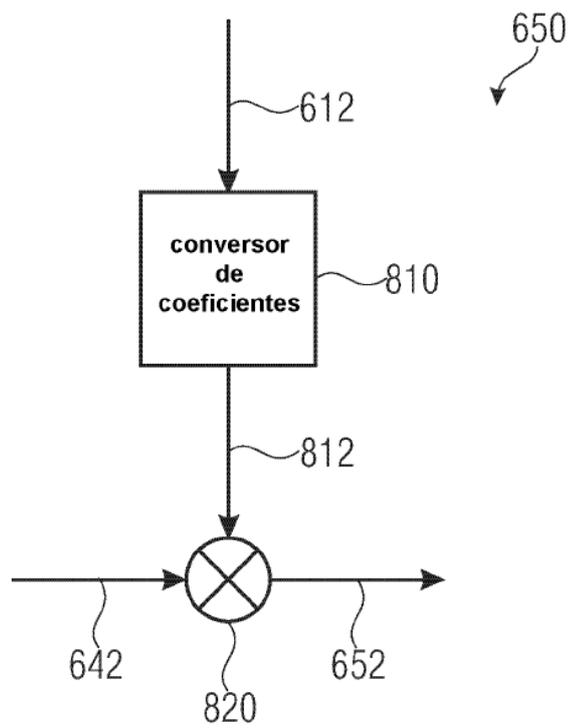


FIG. 6 D



**FIG. 7**



**FIG. 8**

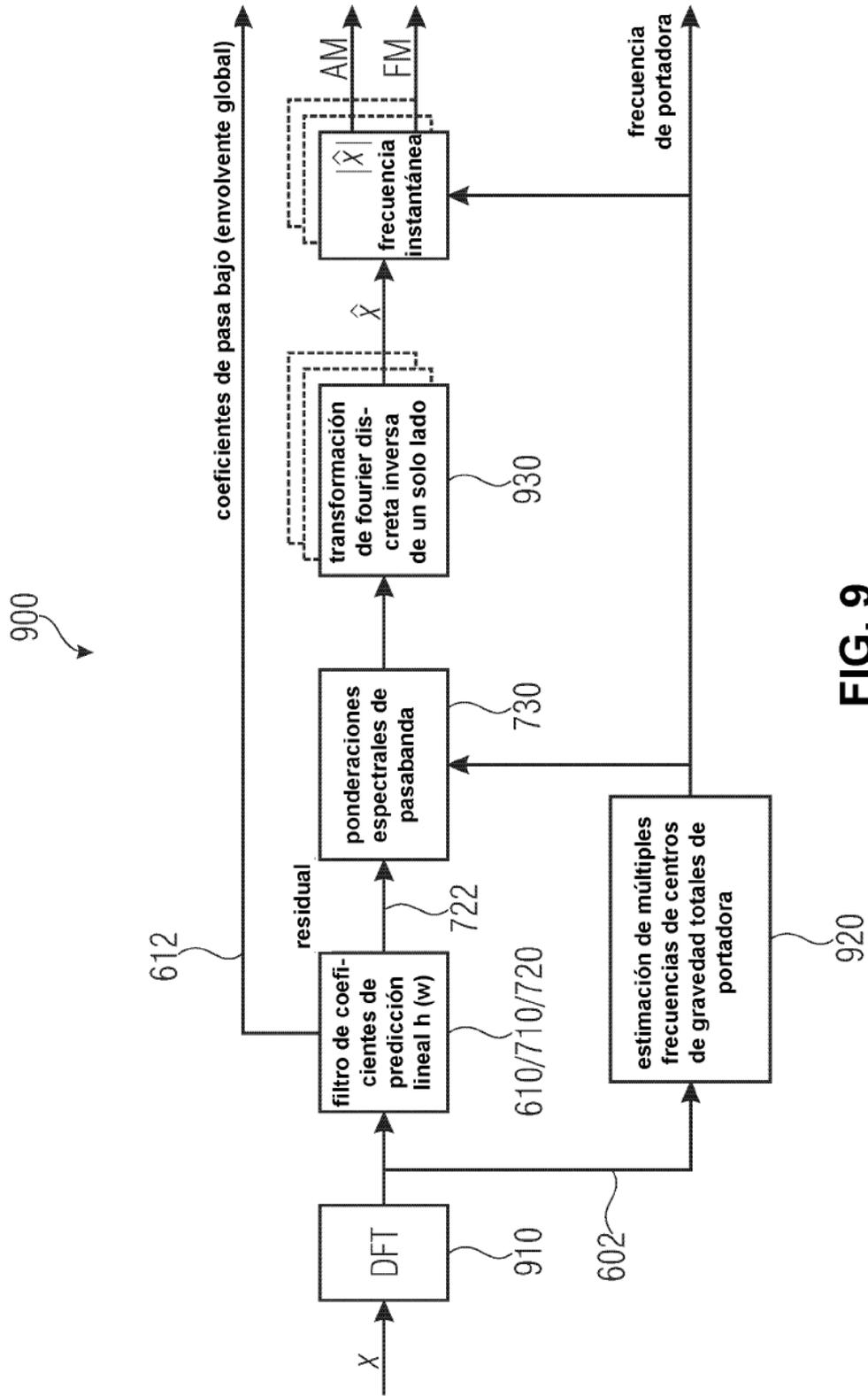


FIG. 9

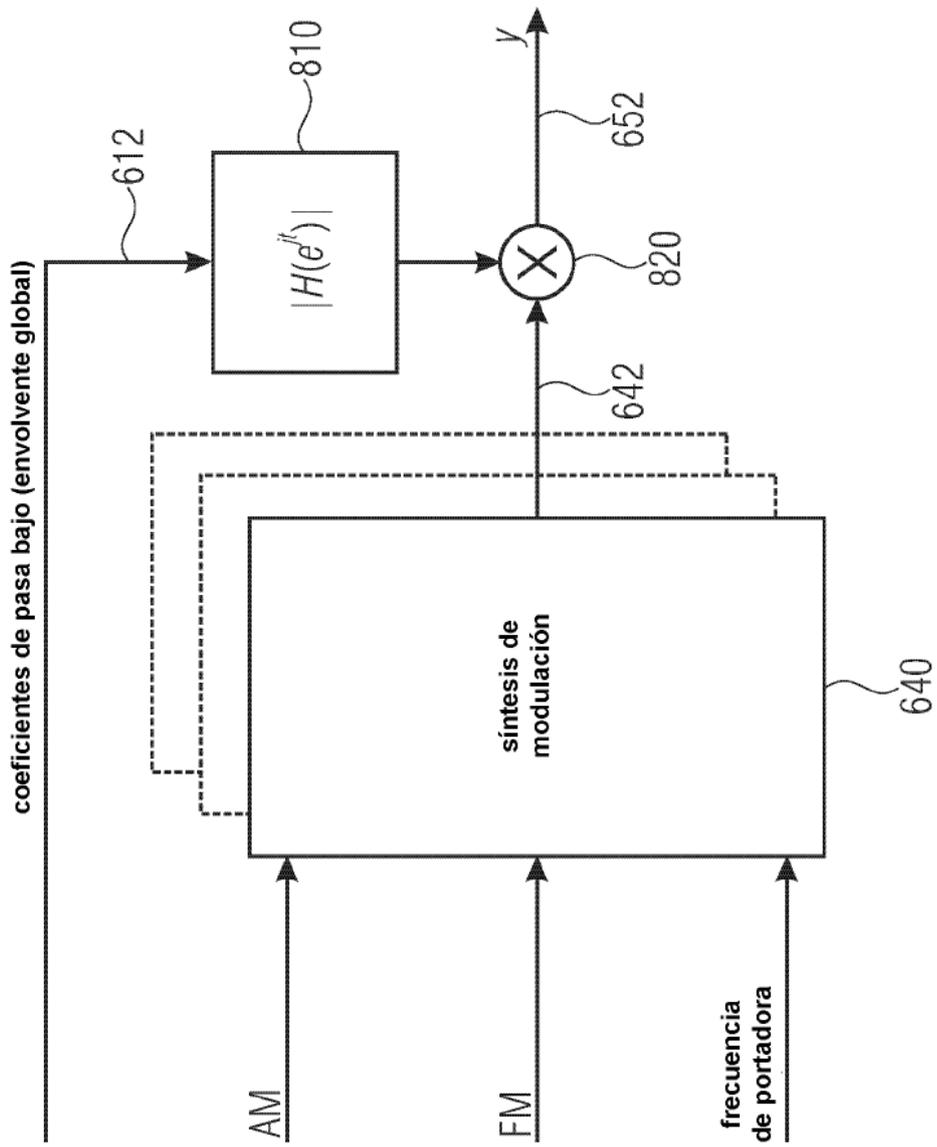
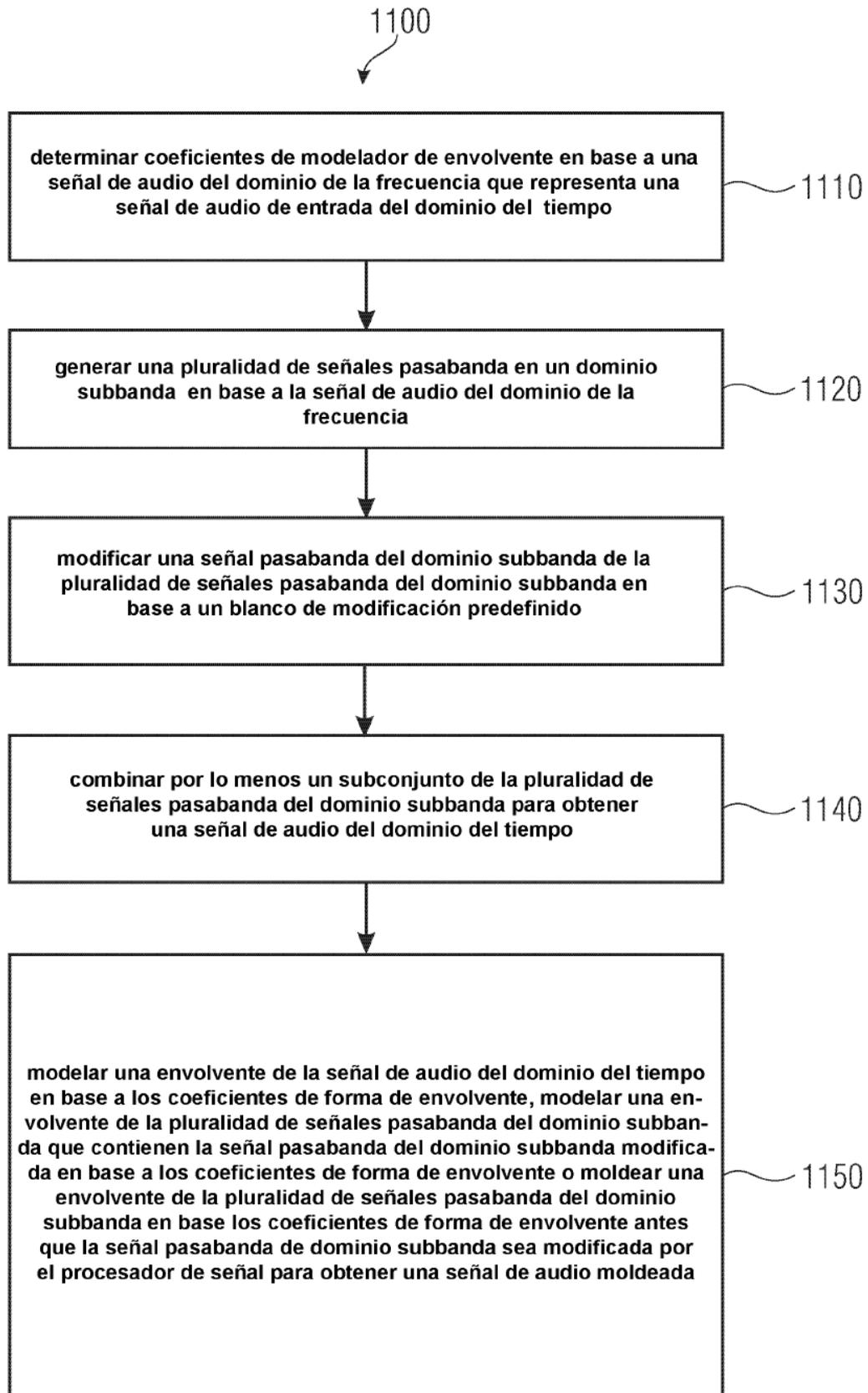


FIG. 10



**FIG. 11**

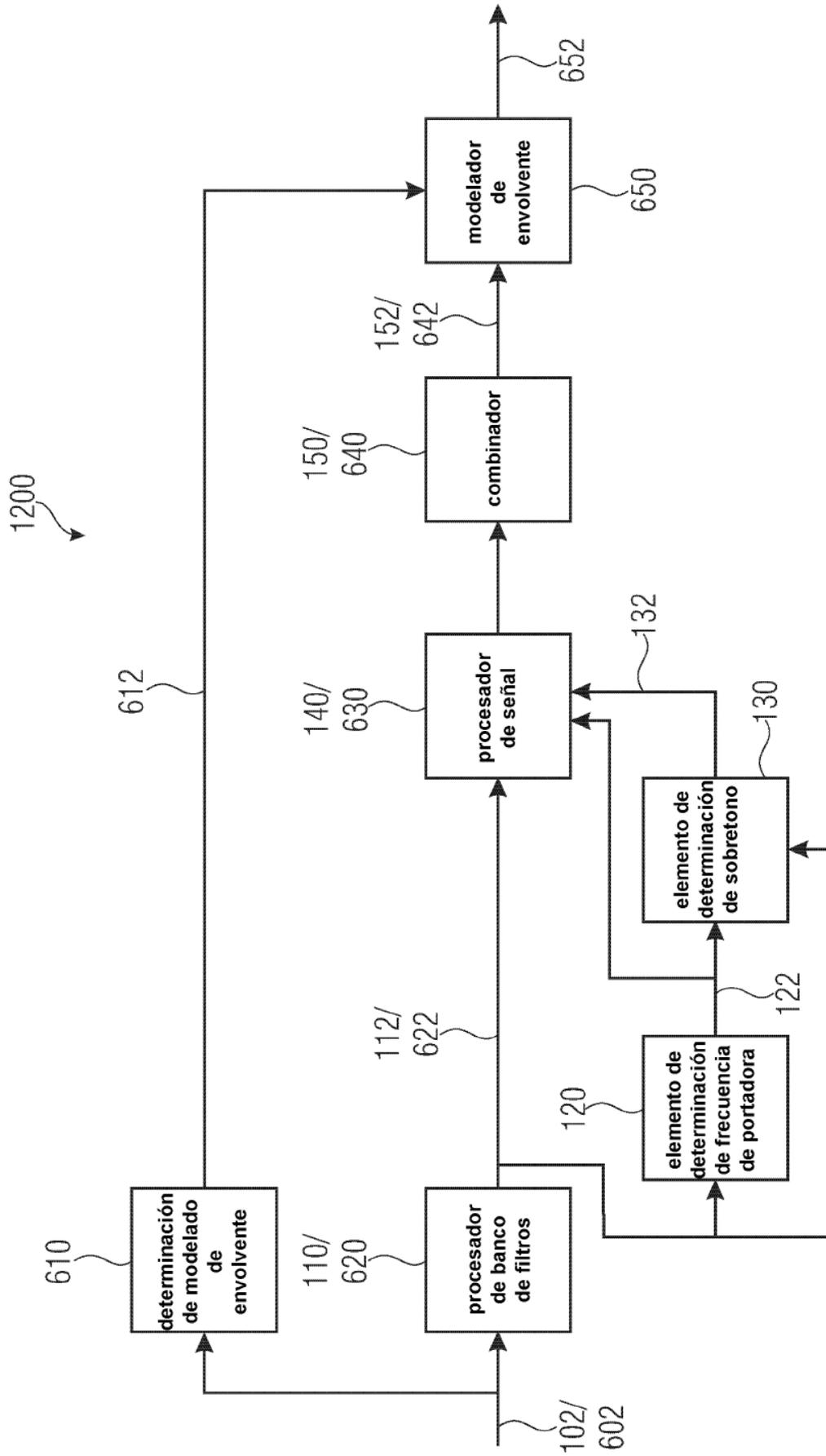


FIG. 12

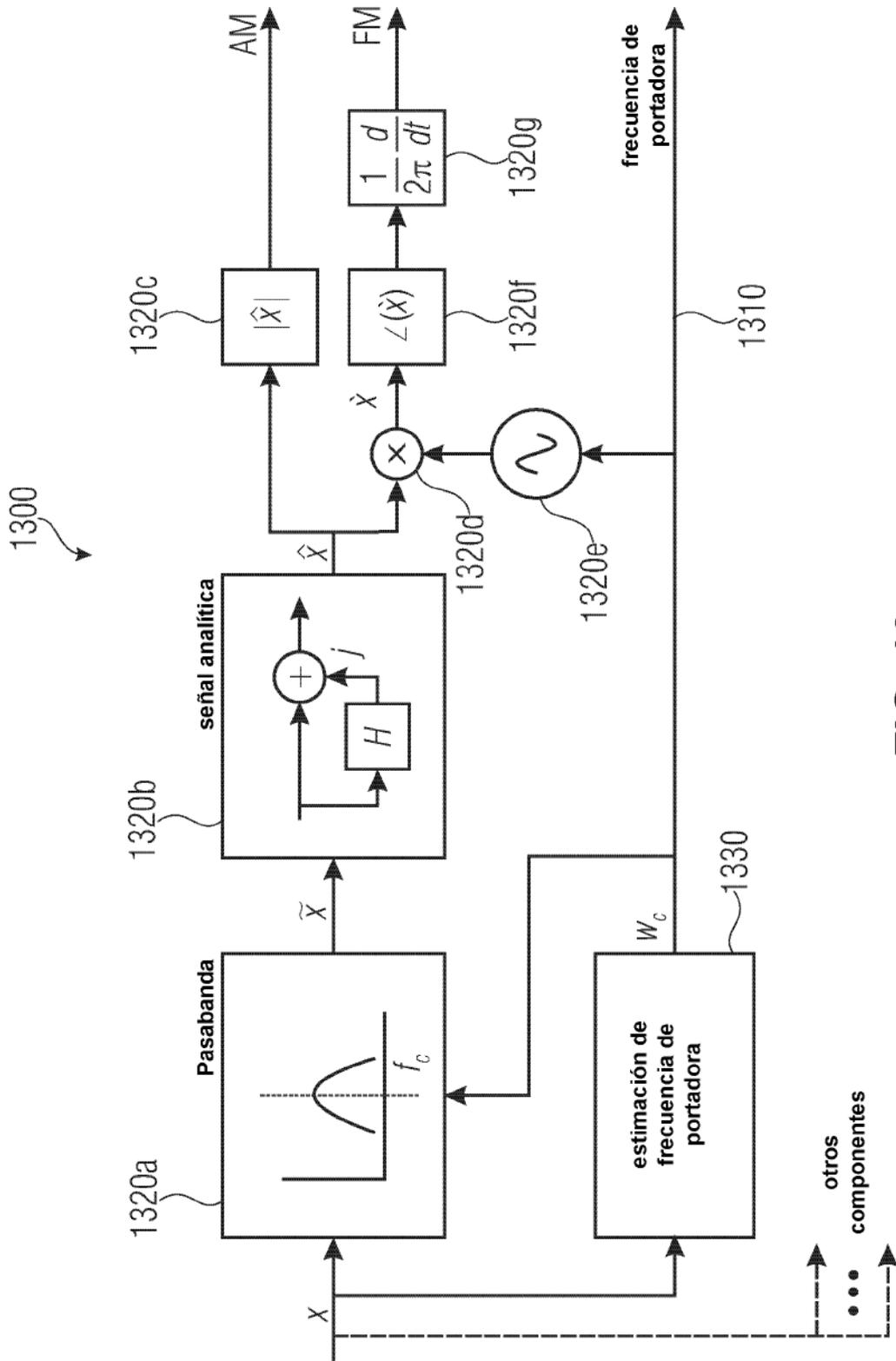


FIG. 13

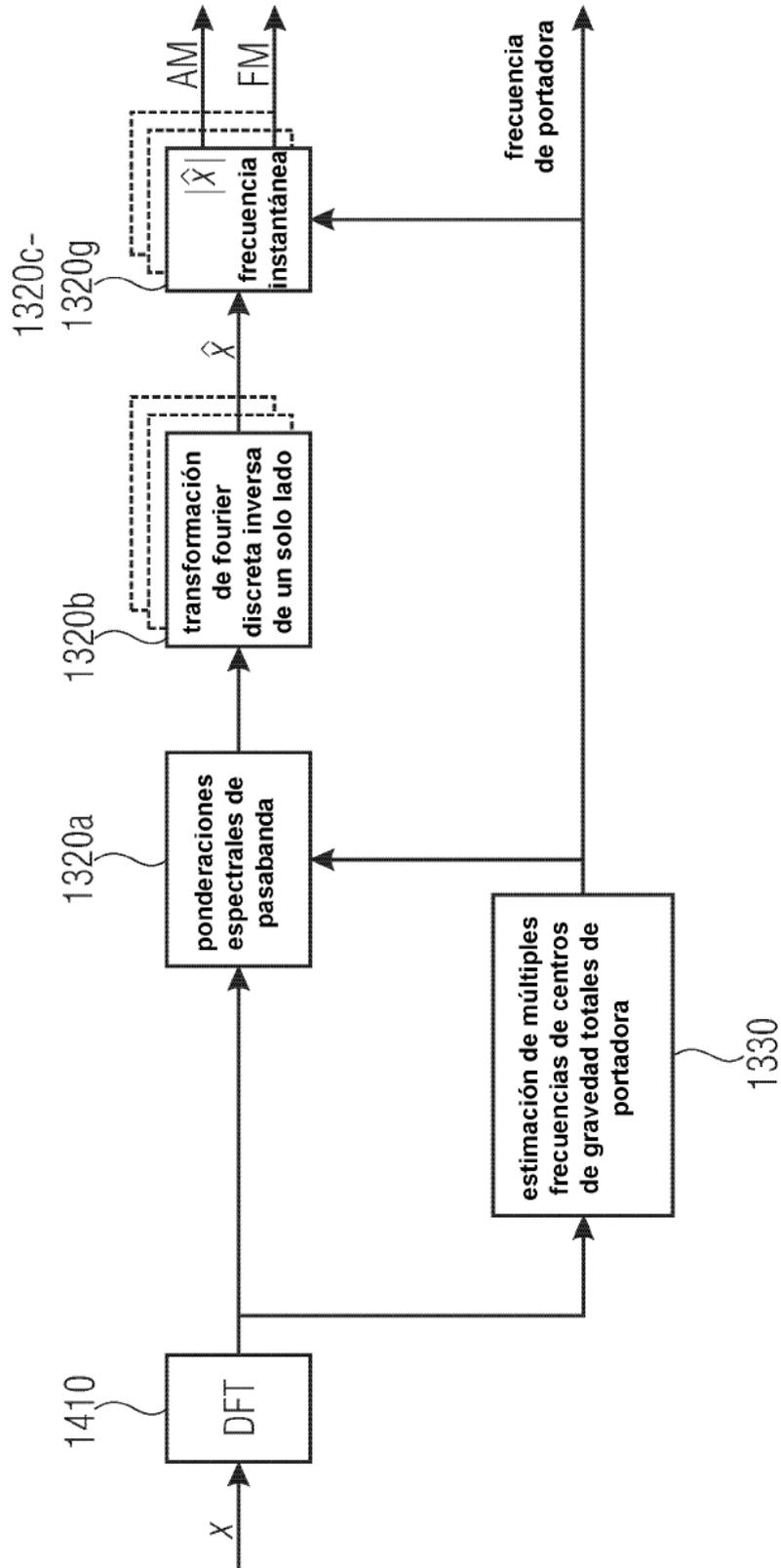
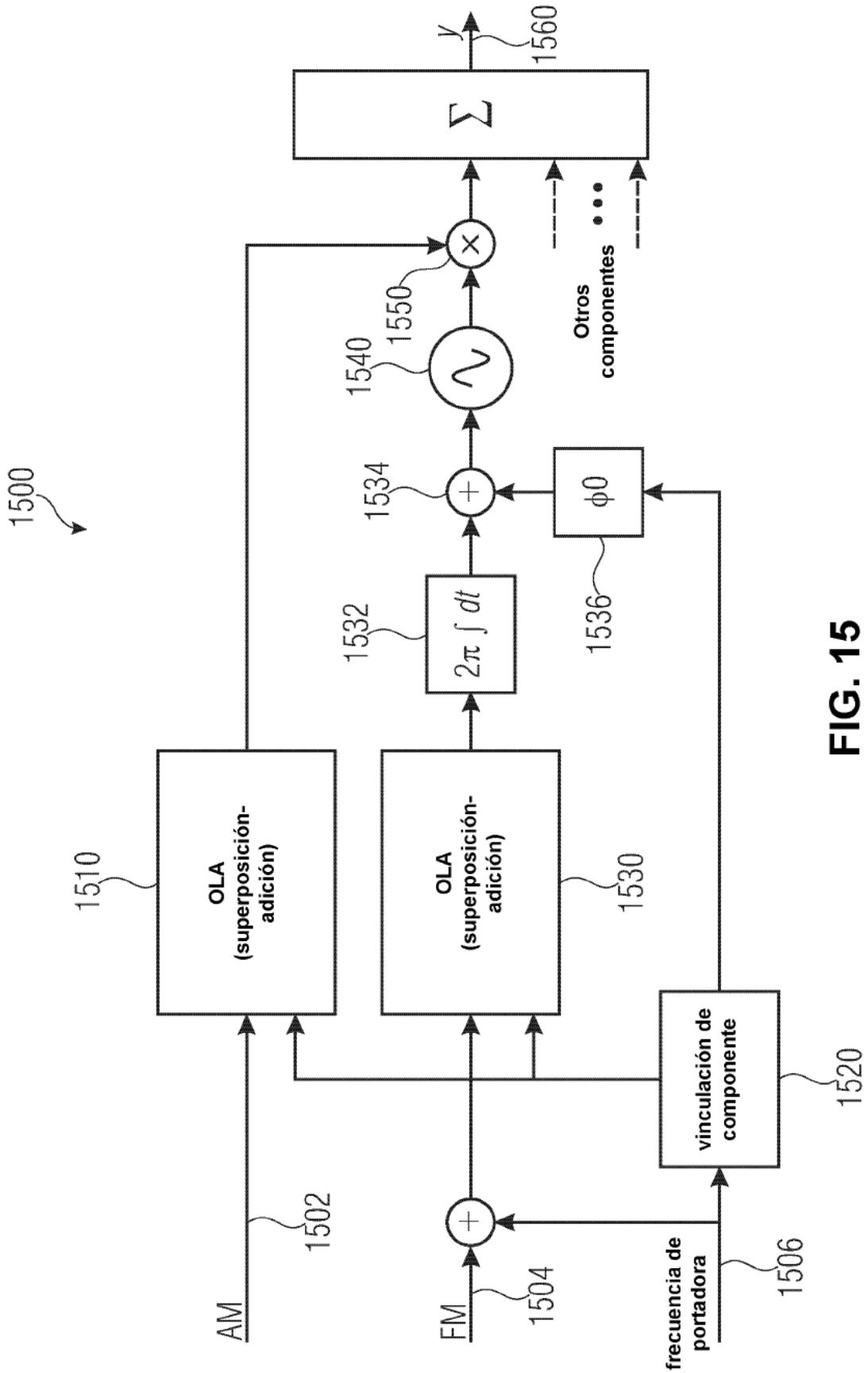


FIG. 14



**FIG. 15**

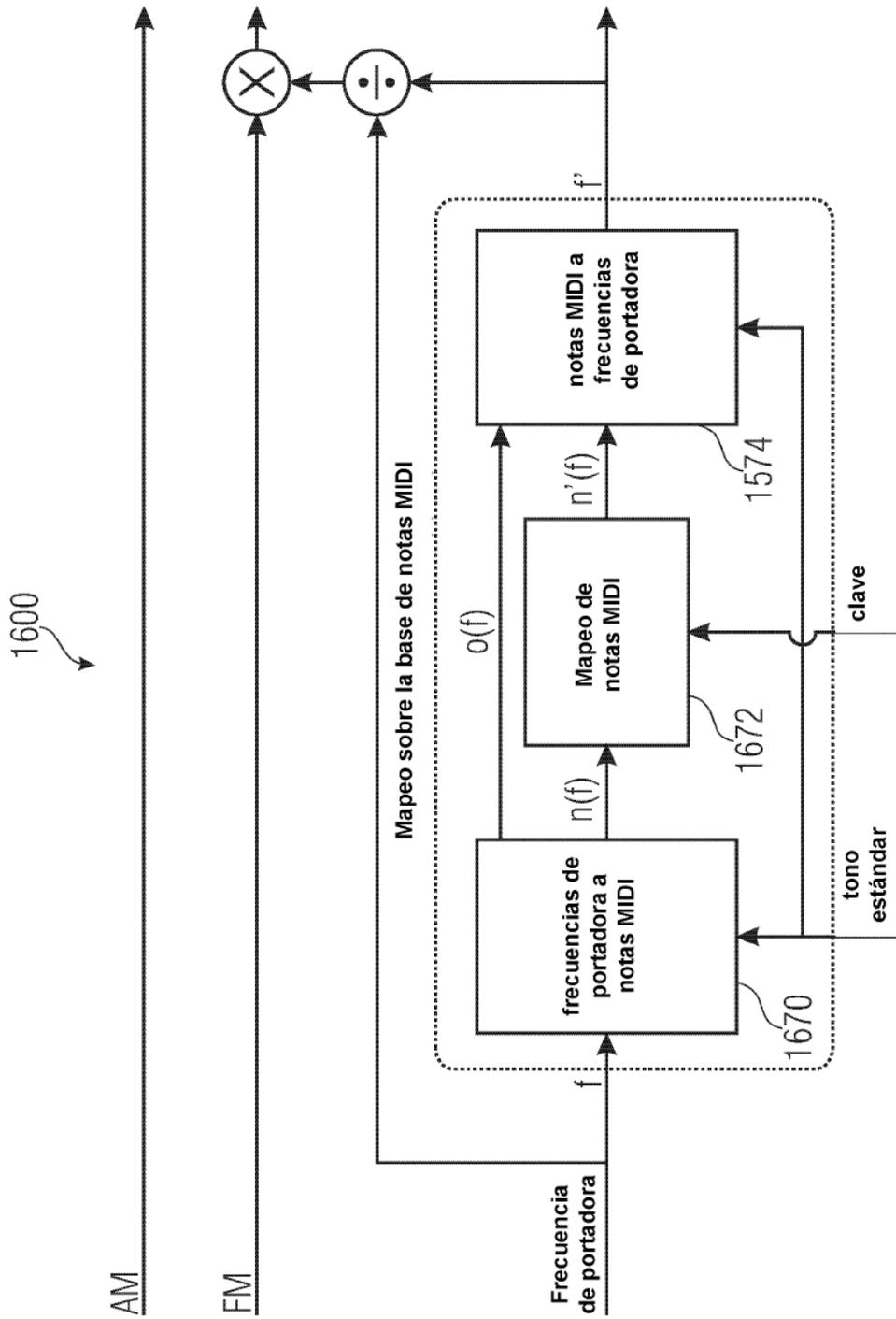


FIG. 16

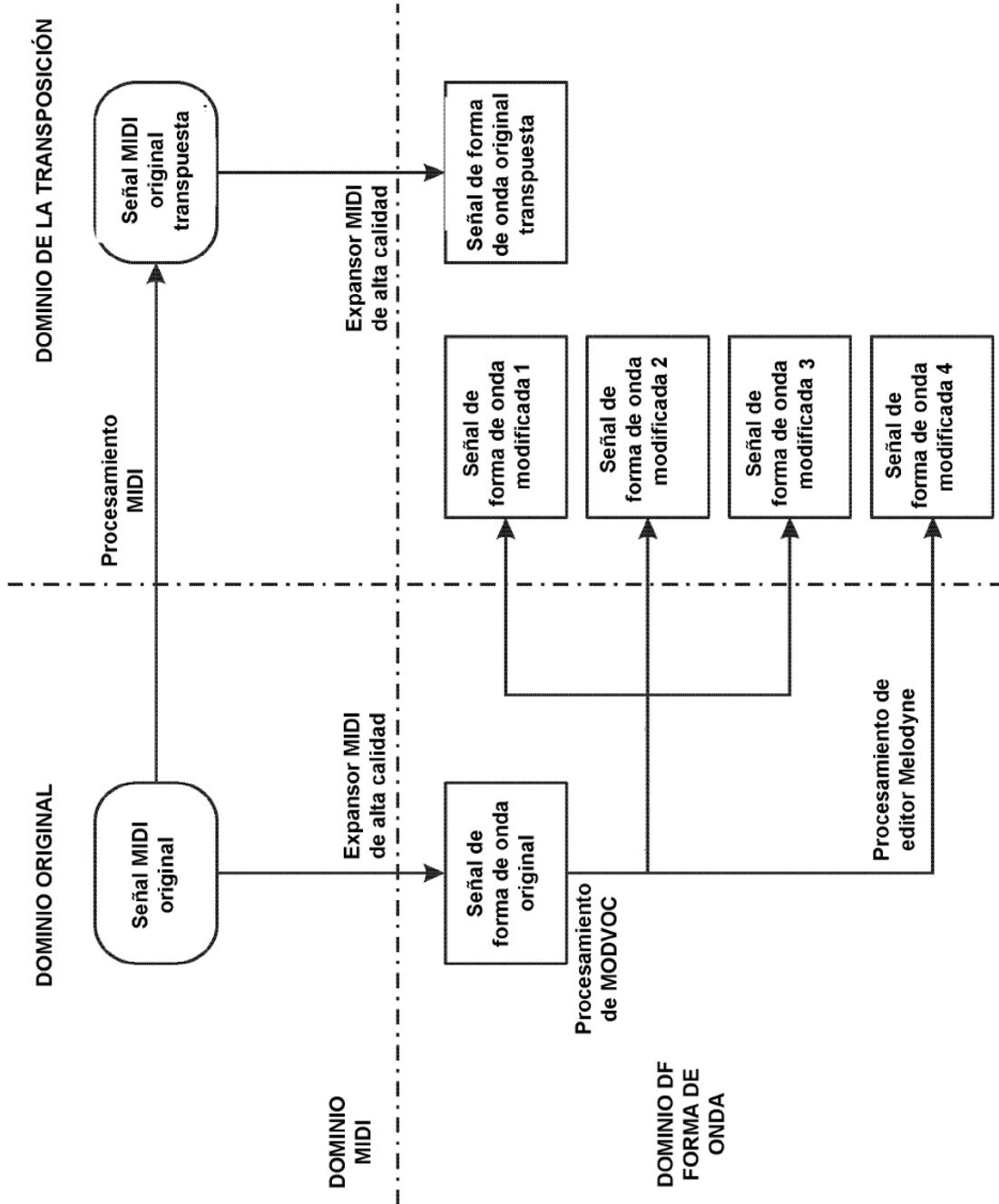
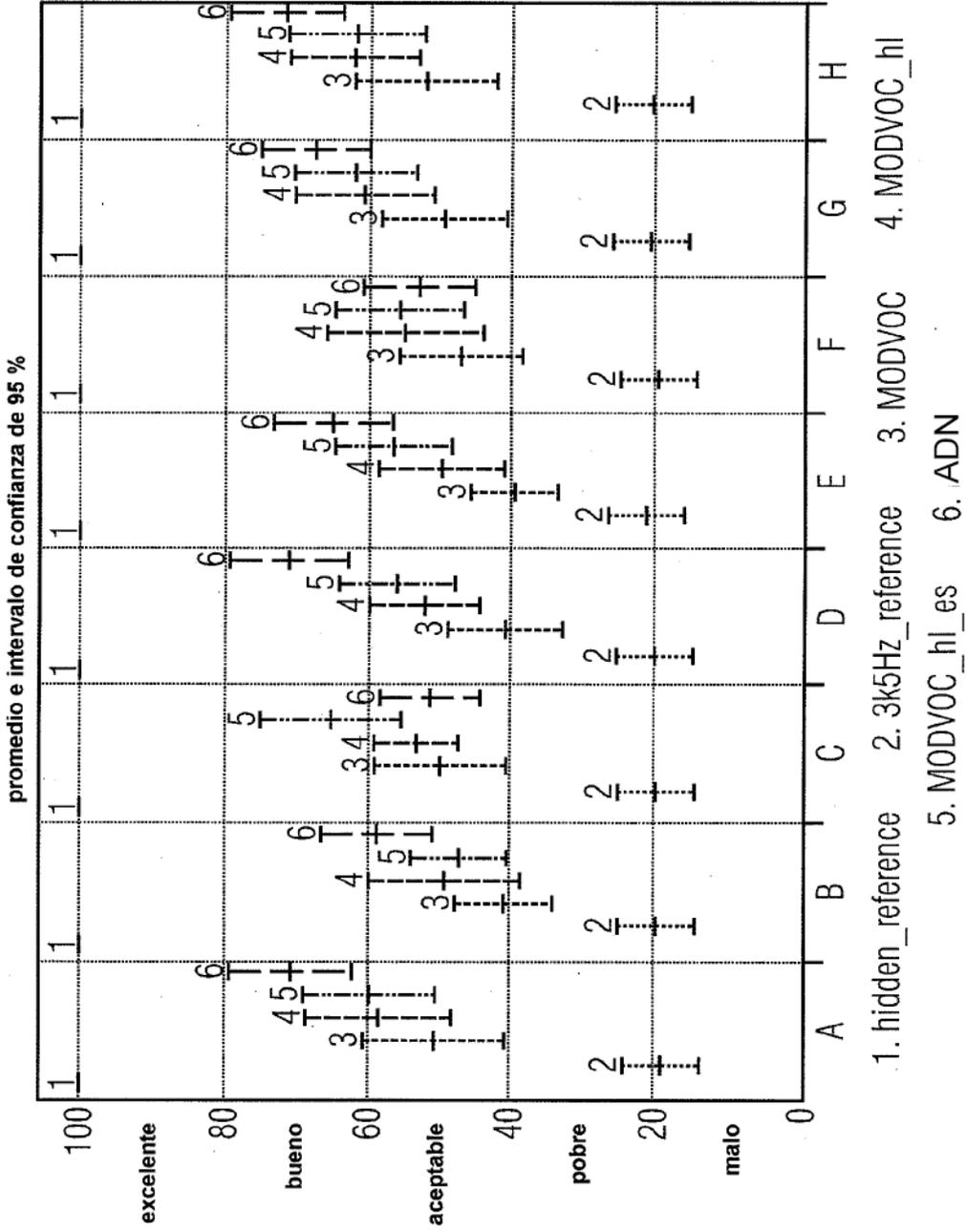
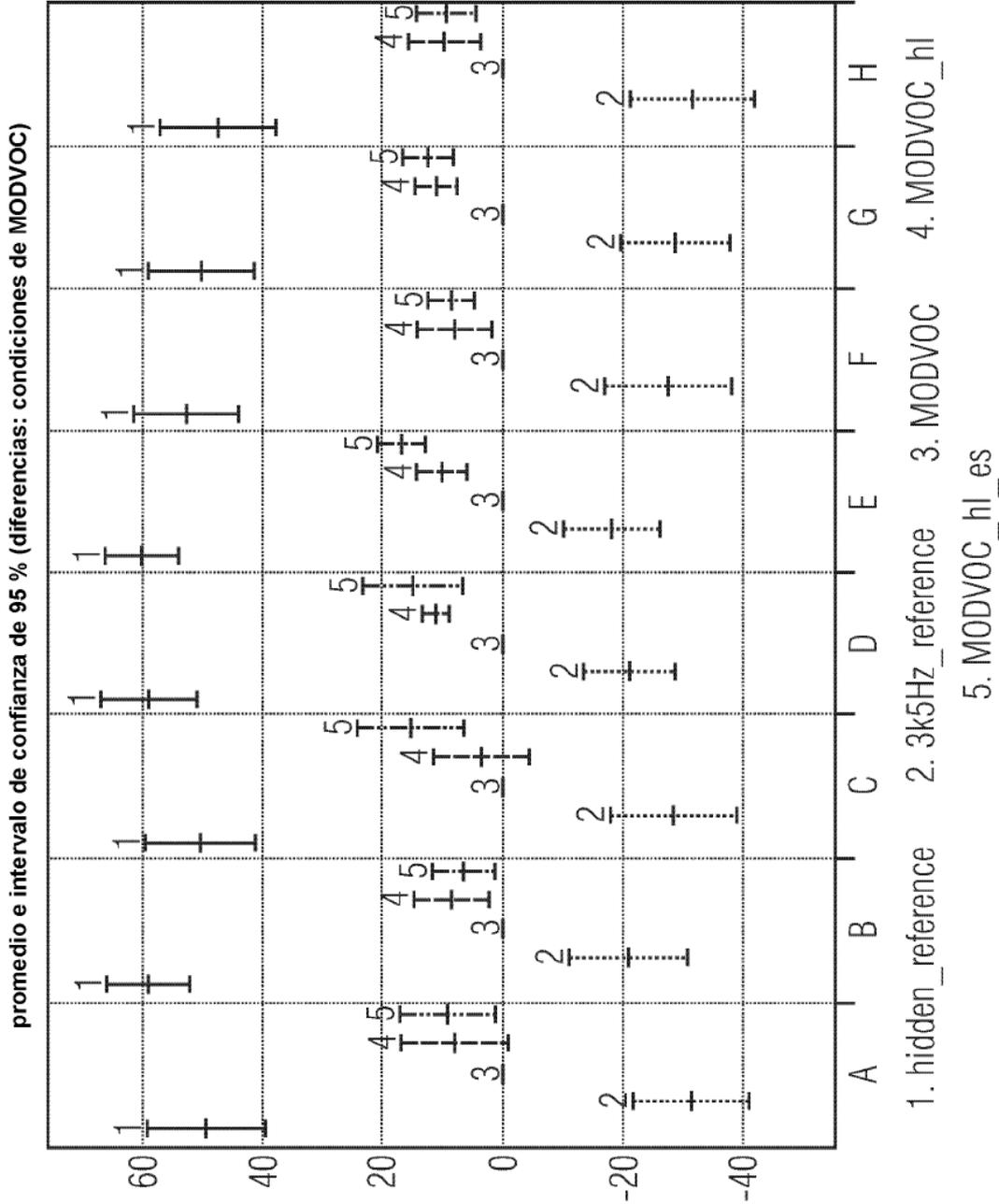


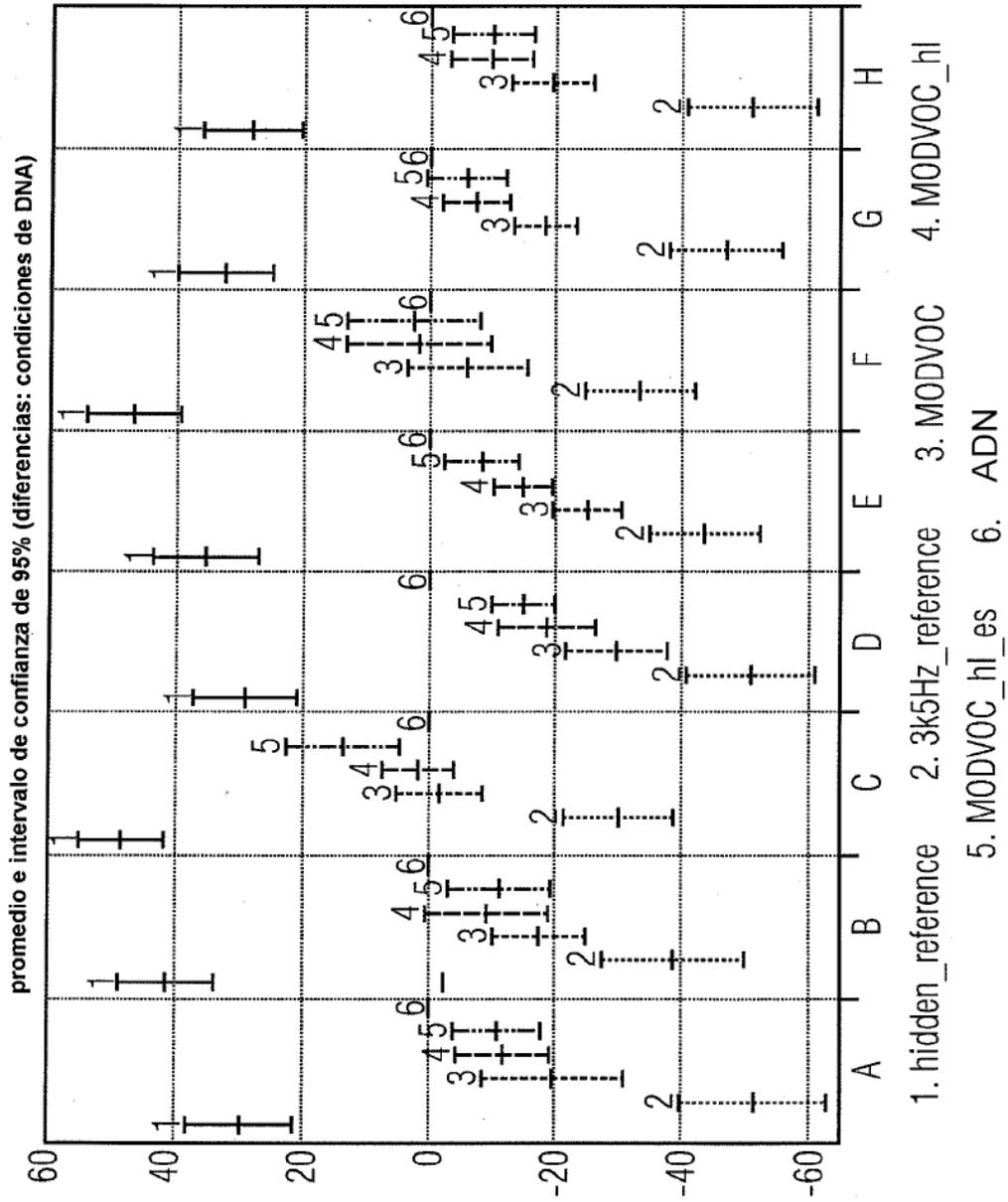
FIG. 17



**FIG. 18**



**FIG. 19**



**FIG. 20**