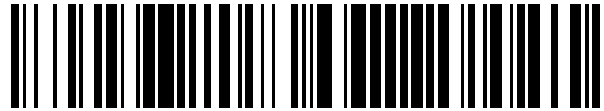


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 566 927**

51 Int. Cl.:

G10L 19/025 (2013.01)

G10L 21/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.01.2010 E 10700048 (1)**

97 Fecha y número de publicación de la concesión europea: **30.12.2015 EP 2392004**

54 Título: **Aparato, proceso y programa informático para la manipulación de una señal de audio que comprende un evento transitorio**

30 Prioridad:

30.01.2009 US 148759 P

05.08.2009 US 231563 P

30.09.2009 EP 09012410

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

18.04.2016

73 Titular/es:

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

72 Inventor/es:

**NAGEL, FREDERIK;
WALTHER, ANDREAS;
FUCHS, GUILLAUME;
LECOMTE, JÉRÉMIE;
POPP, HARALD y
WIK, TILO**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 566 927 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, proceso y programa informático para la manipulación de una señal de audio que comprende un evento transitorio

5 **Antecedentes de la Invención**

Los modos de realización según la invención se refieren a un aparato, un proceso y un programa informático para la manipulación de una señal de audio que comprende un evento transitorio.

10 A continuación se describen escenarios de aplicación típicos en los cuales pueden aplicarse modos de realización según la invención.

15 En los sistemas de procesamiento de señal de audio actuales, las señales audio son a menudo procesadas mediante técnicas digitales. Las porciones de señal específicas, tales como transitorios, por ejemplo, imponen requisitos especiales para el procesamiento de señales digitales.

20 Los eventos transitorios (o „transitorios“) son eventos en una señal durante los cuales la energía de la señal en toda la banda o en una determinada gama de frecuencias cambia rápidamente, es decir, su energía aumenta rápidamente o disminuye rápidamente. Las características de transitorios específicos (eventos transitorios) pueden hallarse en la distribución de la energía de señal en el espectro. De modo típico, la energía de la señal de audio durante un evento transitorio está distribuida sobre toda la gama de frecuencias, mientras que en porciones de señal no transitorias la energía se concentra normalmente en una porción de baja frecuencia de la señal de audio o en una o varias bandas específicas. Eso significa que una porción de señal no transitoria, también llamada porción de señal estacionaria o “tonal”, tiene un espectro que no es plano. Asimismo, el espectro de la porción de señal transitoria es típicamente caótico y “no previsible” (por ejemplo cuando se conoce un espectro de una porción de señal que precede a la porción de señal transitoria). En otras palabras, la energía de la señal está incluida en un número comparativamente pequeño de líneas espectrales o bandas espectrales que son altamente acentuadas respecto a un ruido de fondo de una señal de audio. Sin embargo, en una porción transitoria, la energía de la señal de audio será distribuida sobre muchas bandas de frecuencias distintas y será, de modo específico, distribuida en una porción de alta frecuencia, de modo que un espectro para la porción transitoria de la señal de audio será comparativamente plano y será, de modo típico, más plano que un espectro de una porción tonal de la señal de audio. No obstante, conviene notar que existen otros tipos de señales que tienen un espectro plano, tales como, por ejemplo, las señales de tipo ruido, señales que no representan un transitorio. Sin embargo, mientras que los bins espectrales de señales de tipo ruido tienen valores de fase no correlacionados o poco correlacionados, hay a menudo una correlación de fase muy importante de los bins espectrales en presencia de un transitorio.

35 De modo típico, un evento transitorio es un fuerte cambio en una representación en el dominio temporal de la señal de audio, lo que significa que la señal comportará muchas componentes de frecuencia superior cuando se realiza una descomposición de Fourier. Una característica importante de estos muchos armónicos superiores es el hecho de que las fases de estos armónicos superiores presentan una relación mutua muy específica, de modo que la superposición de todos los armónicos resulte en un cambio rápido de la energía de señal (considerada en el dominio temporal). En otras palabras, existe una fuerte correlación en todo el espectro a proximidad de un evento transitorio. La situación de fase específica entre todos los armónicos puede también llamarse “coherencia vertical”. Esta “coherencia vertical” está relacionada con una representación de espectrograma tiempo/frecuencia de la señal en la cual una dirección horizontal corresponde a una evolución de la señal en el tiempo y donde una dimensión vertical describe la dependencia sobre la frecuencia de las componentes espectrales en un espectro de corta duración respecto sobre la frecuencia.

40 Por ejemplo, si se producen cambios en grandes dominios temporales, por ejemplo por cuantificación, dichos cambios influirán sobre el bloque entero. Dado que los transitorios se caracterizan por un incremento de energía a corto plazo, esta energía será probablemente maculada, cuando cambia el bloque, en toda la región representada por el bloque.

50 El problema también se hace particularmente evidente cuando la velocidad de reproducción de una señal cambia mientras que se mantiene el tono o cuando la señal se transpone mientras que se mantiene la duración original de la reproducción. Ambos pueden realizarse mediante un vocodificador de fase o un método tal como (P)SOLA (refiérase a las referencias [A1] a [A4] al respecto). Este último se obtiene reproduciendo la señal alargada, acelerada por el factor del alargamiento en el tiempo. Para la representación de señal discreta en el tiempo, eso corresponde a un submuestreo de la señal por el factor de alargamiento mientras que se mantiene la frecuencia de muestreo. Los métodos de alargamiento en el tiempo, tales como el vocodificador de fase, convienen actualmente sólo para señales estacionarias o casi estacionarias, dado que los transitorios son “maculados” en el tiempo por dispersión. El vocodificador de fase daña las propiedades llamadas de coherencia vertical (relacionadas con una representación de espectrograma tiempo/frecuencia) de la señal.

60 El alargamiento en el tiempo de señales audio desempeña un papel importante tanto en el entretenimiento como en las

artes. Los algoritmos comunes se basan en técnicas de solapamiento y adición (OLA), tales como Phase Vocodificador (PV), Synchronous Overlap Add (SOLA), Pitch Synchronous Overlap Add (PSOLA), y Waveform Similarity Overlap Add (WSOLA). Mientras que estos algoritmos son capaces de cambiar la velocidad de reproducción de señales audio al tiempo que se preserva su tono original, los transitorios no se preservan bien. El alargamiento en el tiempo de una señal de audio sin modificar su tono mediante OLA requiere el procesamiento por separado de los transitorios y de las porciones de señal sostenidas para evitar una dispersión de transitorios [B1] y un repliegue en el dominio temporal que ocurre a menudo con WSOLA y SOLA. La tarea de alargar una combinación de una señal muy tonal tal como un diapasón y una señal de percusión tal como castañetas constituye un desafío.

A continuación nos referiremos a algunos enfoques convencionales, para presentar el historial de la presente invención.

Algunos métodos actuales alargan el tiempo alrededor de los transitorios más intensamente, de modo que se no deba efectuar ningún, o sólo un pequeño, alargamiento en el tiempo en toda la duración del transitorio (véanse, por ejemplo, las referencias [5] a [8]).

Los siguientes artículos y patentes describen procesos de manipulación del tiempo y/o del tono: [A1], [A2], [A3], [A4], [A5], [A6], [A7], [A8].

En [B2] se propone un proceso que preserve aproximadamente la envolvente de una señal en la versión alargada en el tiempo así como sus características espectrales. Este enfoque prevé que un evento de percusión dilatado en el tiempo desvanecerá más lentamente que el original.

Varios procesos ampliamente conocidos permiten un procesamiento distinto de las componentes de señal transitorias y estacionarias, por ejemplo, la modelización de una señal como suma de senos, transitorios, y ruido (S+T+N) [B4, B5]. Para preservar los transitorios tras la modificación de la escala de tiempo, las tres partes son alargadas por separado. Esta técnica es capaz de preservar perfectamente las componentes transitorias de señales audio. Sin embargo, el sonido resultante es a menudo percibido como poco natural.

Otros enfoques cambian la cantidad de alargamiento en el tiempo y la ajustan a uno durante el tiempo del transitorio o bloquean la fase en el evento transitorio [B3, B6, B7].

El documento [B8] muestra cómo los transitorios pueden preservarse durante el alargamiento del tiempo y de la frecuencia con el PV. En este enfoque, los transitorios fueron eliminados de la señal antes de alargarla. La eliminación de las partes transitorias resultó en lapsos libres dentro de la señal que fueron alargados por el proceso PV. Tras el alargamiento, los transitorios se reañadieron a la señal con una envoltura adaptada a los lapsos libres alargados.

En Hamdy, K.N. et al.: "Time-scale modification of audio signals with combined harmonic and wavelet representations", ICASSP-97, 21-24 Abril 1997, vol. 1, páginas 439-442, se describe una estrategia para escalado en el tiempo de señales de audio que comprenden transitorios.

Visto lo que precede, existe una necesidad de un concepto de manipulación de una señal de audio que comprende un evento transitorio que ofrezca una señal de salida de una calidad percibida mejorada.

Breve Descripción de la Invención

Un modo de realización según la invención crea un aparato para la manipulación de una señal de audio que comprende un evento transitorio de acuerdo con la reivindicación 1, 2 o 3.

El modo de realización descrito más arriba se basa en la descubierta de que el procesador de señal suministra una señal de salida de calidad mejorada si la porción de señal transitoria es sustituida por una porción de señal de sustitución de la cual una energía de señal es adaptada a las características de energía de señal de la señal de audio original, al tiempo que se reduce o elimina el evento transitorio. Este concepto evita grandes cambios paso a paso de la energía de la señal entrada en el procesador de señal que pudieran ser provocados al eliminar sencillamente la porción de señal transitoria de la señal de audio, y también evitar, o al menos reducir, el efecto perjudicial de un transitorio sobre el procesador de señal.

Así, al eliminar o reducir el evento transitorio en la señal de audio (para obtener la señal de audio de transitorios reducidos), y al limitar un cambio de la energía de la señal de audio de transitorios reducidos comparado a la señal de audio de entrada, el procesador de señal recibe una señal de entrada adecuada, de modo que su señal de salida se aproxima a una señal de salida deseada en la ausencia de un evento transitorio.

En un modo de realización preferido, el sustituidor de señal transitoria es configurado para suministrar la porción de señal de sustitución (o porción de señal de transitorios reducidos) de modo que la porción de señal de sustitución represente una señal en el tiempo que presenta una evolución en el tiempo alisada en comparación con la porción de señal transitoria, y de modo que una desviación entre una energía de la porción de señal de sustitución y una energía de una porción de señal no transitoria de la señal de audio que precede a la porción de señal transitoria o que sigue a la porción de señal transitoria sea inferior a un valor de umbral predeterminado. De este modo se puede obtener que la porción de señal de sustitución cumpla con dos condiciones, en particular una “condición de transitorio” y una “condición de energía”. La condición de transitorio indica que un evento transitorio, que es representado por un rediente o un pico en un dominio temporal, es limitado en cuanto a su intensidad (o altura de rediente, o altura de pico) en la porción de señal de sustitución. La condición de energía indica además que la señal de audio de transitorios reducidos (de la porción de señal de sustitución) debería presentar una suave evolución en el tiempo de la distribución de energía espectral. Las discontinuidades en la evolución en el tiempo de la distribución de energía espectral resultan típicamente en la generación de artefactos audibles. Por consiguiente, al limitar tales discontinuidades en el tiempo de la distribución de energía espectral se pueden evitar los artefactos audibles que pudieron resultar de una simple supresión (sin sustitución) de una porción de señal transitoria de la señal de audio de entrada.

En un modo de realización preferido, el sustituidor de señal transitoria es configurado para extrapolar los valores de amplitud de una o varias porciones de señal que preceden la porción de señal transitoria, para obtener valores de amplitud de la porción de señal de sustitución. El sustituidor de señal transitoria es también configurado para extrapolar los valores de fase de una o varias porciones de señal que preceden la porción de señal transitoria, para obtener valores de fase de la porción de señal de sustitución. Mediante este enfoque se puede obtener una suave evolución de amplitud de la señal de audio de transitorios reducidos. Además, las fases de las distintas componentes espectrales de la señal de audio de transitorios reducidos están bien controladas (mediante extrapolación), de modo que el evento transitorio, que se caracteriza por valores de fase específicos durante la porción de señal transitoria (distintos de los valores de fase de porciones de señal no transitorias), está suprimido.

En otras palabras, los valores de fase son forzados mediante extrapolación, los cuales son generados de modo distinto de los valores de fase que caracterizan el transitorio. La extrapolación ofrece también la ventaja que el conocimiento de las porciones de señal de audio que preceden la porción de señal transitoria es suficiente para efectuar la extrapolación. Sin embargo, es posible, naturalmente, aplicar además algunas informaciones laterales, por ejemplo parámetros de extrapolación, para efectuar la extrapolación.

En otro modo de realización preferido, el reinsertador de señal transitoria (150) es configurado para encadenar la versión procesada de la señal de audio de transitorios reducidos con la señal transitoria que representa, de forma original o procesada, un contenido transitorio de la porción de señal transitoria. En tal caso, la versión procesada de la señal de transitorios reducidos puede ser una versión alargada en el tiempo de la señal de audio de entrada. Por consiguiente, el transitorio puede ser reinsertado suavemente en una versión alargada de la señal de audio de entrada. En otras palabras, tras el alargamiento (en el tiempo) de la señal de audio de transitorios reducidos, los transitorios (de forma procesada o no procesada) son reañadidos a la señal con una envoltura adaptada a los lapsos libres alargados.

En otro modo de realización preferido, el sustituidor de señal transitoria es configurado para interpolar entre un valor de amplitud de una porción de señal que precede a la porción de señal transitoria y un valor de amplitud de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o varios valores de amplitud de la porción de señal de sustitución. El sustituidor de señal transitoria es, además, configurado para interpolar entre un valor de fase de una porción de señal que precede a la porción de señal transitoria y un valor de fase de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o varios valores de fase de la porción de señal de sustitución. Al efectuar una interpolación se puede obtener una suave evolución en el tiempo de los valores tanto de amplitud como de fase. La interpolación de la fase resulta también, típicamente, en una reducción o cancelación del evento transitorio, dado que los transitorios comprenden típicamente una distribución de fase muy específico a proximidad directa del transitorio, distribución de fase que es típicamente distinta de la distribución de fase a una determinada distancia del transitorio.

En un modo de realización preferido, el sustituidor de señal transitoria es configurado para aplicar un ruido ponderado (por ejemplo un espectro de una señal de tipo ruido, adaptado a las características de energía de señal de una o varias porciones de señal no transitorias de la señal de audio, o a una característica de energía de señal de la porción de señal transitoria), para obtener los valores de amplitud de la porción de señal de sustitución, y para aplicar un ruido ponderado para obtener los valores de fase de la porción de señal de sustitución. Eso es posible aplicando un ruido ponderado, para reducir más aún el transitorio al tiempo que se conserva suficientemente reducido el impacto sobre la energía.

En un modo de realización preferido, el sustituidor de señal transitoria es configurado para combinar componentes no transitorias de la porción de señal transitoria con los valores extrapolados o interpolados, para obtener la porción de señal de sustitución. Se ha constatado que se puede obtener una calidad mejorada de la señal de audio de transitorios reducidos

(y de su versión procesada, la cual se obtiene mediante el procesador de señal) si se mantienen las componentes no transitorias de la porción de señal transitoria. Por ejemplo, las componentes tonales de la porción de señal transitoria sólo pueden tener un impacto limitado sobre el transitorio (ya que un transitorio en el tiempo está provocado típicamente por una señal de banda ancha que presenta una distribución de fase específica sobre la frecuencia). Así, las componentes tonales
5 pueden tener un impacto limitado sobre el transitorio (ya que un transitorio en el tiempo está provocado típicamente por una señal de banda ancha que presenta una distribución de fase específica sobre la frecuencia). Así, las componentes tonales no transitorias de la porción de señal transitoria pueden llevar una información preciosa que puede realmente contribuir a una señal de salida deseable del procesador de señal. Así, mantener tales porciones de señal – al tiempo que se reduce el transitorio – puede contribuir a una mejora de la señal de audio procesada.

En un modo de realización de la invención, el sustituidor de señal transitoria es configurado para obtener porciones de señal de sustitución de longitud variable en función de una longitud de una porción de señal transitoria. Se ha constatado que la calidad de la señal de audio puede a veces ser mejorada adaptando la longitud de las porciones de señal de sustitución a una longitud variable de las porciones de señal transitorias. Por ejemplo, en algunas señales, las porciones de señal transitorias pueden ser de muy corta duración. En tal caso, una señal de audio procesada optimizada puede obtenerse sustituyendo sólo una porción relativamente corta de la señal de audio de entrada. Así, se puede mantener una mayor cantidad posible de informaciones (no transitorias) de la señal de audio de entrada original. Al mantener también cortas las porciones de señal de sustitución (según la longitud de la porción de señal transitoria) puede evitarse en muchas situaciones un solapamiento porciones de señal de sustitución sucesivas. Por lo tanto, en la mayoría de los casos resulta realizable que exista una porción de señal no transitoria original entre dos porciones de señal de sustitución sucesivas. Por lo tanto, la señal de audio procesada se genera con suficiente precisión, manteniendo una mayor cantidad posible de informaciones (no transitorias) de la señal de audio de entrada original.
10
15
20

En un modo de realización preferido, el procesador de señal es configurado para procesar la señal de audio de transitorios reducidos de modo que una porción de señal temporal dada de la versión procesada de la señal de audio de transitorios reducidos esté función de una pluralidad de porciones de señal temporal que no se solapan en el tiempo de la señal de audio de transitorios reducidos. En otras palabras, es preferible que el procesador de señal comprenda una memoria temporal durante la generación de las porciones de señal de la versión procesada de la señal de audio de transitorios reducidos. El procesamiento de señal mediante una memoria permite un procesamiento por bloque de la señal de audio de transitorios reducidos, o un filtrado temporal (por ejemplo filtrado FIR, o filtrado IIR) de la señal de audio de transitorios reducidos. También se ha constatado que el concepto según la invención de la sustitución de porciones de señal transitorias conviene muy bien para trabajar en cooperación con tal procesador de señal. Mientras que los transitorios tendrían normalmente un importante impacto negativo sobre el procesador de señal descrito que efectúa un procesamiento por bloque o que presenta una memoria temporal, las porciones de señal de sustitución según la invención reducen este efecto perjudicial del transitorio. Mientras que un transitorio tendría normalmente un impacto sobre múltiples porciones de señal suministradas por el procesador de señal – que se extienden más allá de los límites temporales de la porción de señal transitoria – el efecto perjudicial de un transitorio es reducido o incluso eliminado por el concepto de la invención. Al mantener una suave evolución en el tiempo de la energía de la señal de transitorios reducidos, cualquier degradación puede ser mantenida suficientemente suave. Por ejemplo, un bloque (del procesamiento por bloque del procesador de señal) que comprende una porción de señal de sustitución (por ejemplo además de una porción de señal no transitoria original) no se degrada seriamente, dado que la porción de señal de sustitución está adaptada en energía al resto del bloque. Así, el bloque en su totalidad se ve afectado sólo ligeramente por la eliminación o reducción del evento transitorio. Además, un filtrado temporal que sería afectado negativamente por un evento transitorio, y también por una eliminación completa (por ejemplo en forma de un forzado a cero) de la porción de señal transitoria, queda prácticamente inafectado por la eliminación (o reducción) de transitorios, debido al uso de una porción de señal de sustitución.
25
30
35
40

En un modo de realización preferido, el procesador de señal es configurado para efectuar un procesamiento a base de bloques en el tiempo de la señal de audio de transitorios reducidos, para obtener la versión procesada de la señal de audio de transitorios reducidos. El sustituidor de señal transitoria es también configurado para ajustar la duración de la porción de señal a sustituir por la porción de señal de sustitución con una resolución temporal más fina que la duración de un bloque temporal, o para sustituir una porción de señal transitoria que presenta una duración en el tiempo inferior a la duración del bloque temporal por una porción de señal de sustitución que presenta una duración en el tiempo inferior a la duración del bloque temporal. Así, la sustitución sugerida aquí permite un procesamiento con poca distorsión de señales audio, incluso si la longitud de las porciones transitorias eliminadas es distinta de la longitud de los bloques en el tiempo.
45
50

En un modo de realización preferido, el procesador de señal es configurado para procesar la señal de audio de transitorios reducidos en función de la frecuencia, de modo que le procesamiento introduzca en la señal de audio de transitorios reducidos desfases función de la frecuencia que degradan los transitorios. Sin embargo, incluso tal procesamiento de señal que degrada los transitorios no tiene ningún impacto perjudicial importante sobre la señal de audio procesada, dado que los transitorios son típicamente procesados separadamente del procesamiento de la señal de audio de transitorios reducidos. Por consiguiente, mientras que un algoritmo de procesamiento de señal que degrada los transitorios puede aplicarse en el procesador de señal, la calidad de los transitorios puede mantenerse mediante un procesamiento separado del transitorio y una reinserción de los transitorios en una fase ulterior del procesamiento.
55
60

5 En un modo de realización preferido, el sustituidor de señal transitoria comprende un detector de transitorios, donde el detector de transitorios es configurado para suministrar un umbral de detección variable en el tiempo para la detección del transitorio en la señal de audio, de modo que el umbral de detección siga una envolvente de la señal de audio con una constante de tiempo de alisado ajustable. El detector de transitorios es configurado para cambiar la constante de tiempo de alisado en respuesta a la detección de un transitorio y/o en función de una evolución en el tiempo de la señal de audio. Por medio de tal detector de transitorios, es posible detectar transitorios de intensidades distintas, incluso si los transitorios son poco distantes en el tiempo. Por ejemplo, el concepto de la invención permite la detección de un transitorio débil, incluso si el transitorio débil sigue de cerca un transitorio más fuerte que precede. Por consiguiente, la detección de transitorios para la sustitución de transitorios puede efectuarse de modo fiable y preciso.

15 En un modo de realización preferido, el aparato comprende un procesador de transitorios configurado para recibir una información de transitorio que representa el contenido transitorio de la porción de señal transitoria. En tal caso, el procesador de transitorios puede ser configurado para obtener, en base a la información de transitorio, una señal transitoria procesada en la cual las componentes tonales son reducidas. El reinsertador de señal transitoria puede ser configurado para combinar la versión procesada de la señal de audio de transitorios reducidos con la señal transitoria procesada suministrada por el procesador de transitorios. Así, el procesamiento separado de la señal de audio de transitorios reducidos y de la componente transitoria de la señal de entrada audio (representada por la información de transitorio) puede efectuarse de modo que una combinación sucesiva de las distintas porciones de señal resulte en una señal de salida total apropiada. Estas componentes de señal de la porción de señal transitoria que fueron procesadas por el procesador de señal "principal" (por ejemplo componentes de señal tonales) no deben ser incluidas en el procesamiento separado del transitorio. Por consiguiente, se puede realizar una repartición apropiada del procesamiento de las componentes audio de la porción de señal transitoria.

25 Otros modos de realización según la invención crean un proceso de acuerdo con la reivindicación 15 y un programa informático de acuerdo con la reivindicación 16, para la manipulación de una señal de audio que comprende un evento transitorio.

30 **Breve Descripción de las Figuras**

A continuación son descritos modos de realización según la invención, con referencia a las figuras adjuntas, en las cuales:

35 La figura 1 ilustra un diagrama esquemático de un aparato para la manipulación de una señal de audio que comprende un evento transitorio, según un modo de realización de la presente invención;

La figura 2 ilustra un diagrama esquemático de un sustituidor de señal transitoria, según un modo de realización de la presente invención;

40 Las figuras 3a-3d ilustran diagramas esquemáticos de un procesador de señal, según modos de realización de la presente invención;

La figura 4 ilustra un diagrama esquemático de un reinsertador de señal transitoria, según un modo de realización de la presente invención;

45 La figura 5a ilustra una vista general de la implementación de un vocodificador a utilizar en el procesador de señal de la figura 1;

La figura 5b ilustra una implementación de partes (análisis) de un procesador de señal de la figura 1;

50 La figura 5c ilustra otras partes (alargamiento) de un procesador de señal de la figura 1;

La figura 6 ilustra una implementación de transformación de un vocodificador de fase a utilizar en el procesador de señal de la figura 1;

55 La figura 7 ilustra una representación esquemática del funcionamiento de un algoritmo de vocodificador de fase, siendo el tamaño de tramo de síntesis distinto del tamaño de tramo de análisis, por ejemplo de un factor de 2;

La figura 8 ilustra una representación gráfica de una evolución en el tiempo de la amplitud de una señal de audio;

60 La figura 9 ilustra una representación gráfica de una sincronización del procesamiento de señal en el aparato de la figura 1;

La figura 10 ilustra una representación gráfica de señales que pueden aparecer en un aparato según la figura 1;

La figura 11 ilustra otra representación gráfica de señales que pueden aparecer en un aparato según la figura 1;

La figura 12 ilustra un organigrama de un proceso para la manipulación de una señal de audio, según un modo de realización de la presente invención;

La figura 13 ilustra una representación gráfica de una eliminación de transitorios e interpolación, según un modo de realización de la invención;

La figura 14 ilustra una representación gráfica de un alargamiento en el tiempo y re inserción de transitorios, según un modo de realización de la invención;

La figura 15 ilustra una representación gráfica de formas de onda de señal que se presentan en distintas fases de la manipulación de transitorios de la invención en una aplicación de alargamiento en el tiempo con el vocodificador de fase; y

La figura 16 ilustra una representación gráfica de señales que son presentes en las distintas fases de un alargamiento en el tiempo.

Descripción Detallada de las Modalidades Preferidas

A continuación se describen algunos modos de realización según la invención. Un primer modo de realización de un aparato para la manipulación de una señal de audio que comprende un evento transitorio será descrito con referencia a la figura 1 que ilustra una vista general del primer modo de realización, también con referencia a las figuras 2, 3a a 3c, 4, 5a, 5b, 5c, 6 y 7, las cuales ilustran detalles de los componentes del primer modo de realización y el funcionamiento del vocodificador de fase (figura 7). Una señal transitoria es ilustrada en la figura 8, y su procesamiento es ilustrado en las figuras 9 a 11. La figura 12 ilustra un organigrama de un proceso correspondiente.

A continuación se describe el funcionamiento de un segundo modo de realización de un aparato para la manipulación de una señal de audio que comprende un evento transitorio, con referencia a las figuras 13 a 17.

Modo de realización según la figura 1

La figura 1 ilustra un diagrama esquemático de un aparato para la manipulación de una señal de audio que comprende un evento transitorio, según un modo de realización de la invención. El aparato ilustrado a la figura 1 está designado en su totalidad con 100. El aparato 100 es configurado para recibir una señal de audio 110 que comprende un evento transitorio, y para suministrar, en base a la misma, una señal de audio procesada 120 con un transitorio "natural" o sintetizado sin procesar. El aparato 100 comprende un sustituidor de señal transitoria 130 configurada para sustituir una porción de señal transitoria, que comprende el evento transitorio de la señal de audio 110, por una porción de señal de sustitución adaptada a las características de energía de señal de una o varias porciones de señal no transitorias de la señal de audio, o a una característica de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos 132. Opcionalmente, las características de fase de la porción de señal de sustitución pueden ser adaptadas a las características de fase de una o varias porciones de señal no transitorias de la señal de audio. El aparato 100 comprende además un procesador de señal 140 configurado para procesar la señal de audio de transitorios reducidos 132, para obtener una versión procesada 142 de la señal de audio de transitorios reducidos. El aparato 100 comprende además un reinsertador de señal transitoria 150 configurado para combinar la versión procesada 142 de la señal de audio de transitorios reducidos con una señal transitoria 152, para obtener la señal de audio procesada 120 con un transitorio "natural" o sintetizado sin procesar. La señal transitoria 152 puede representar, de forma original o procesada, un contenido transitorio de la porción de señal transitoria que fue sustituida por la porción de señal de sustitución por el sustituidor de señal transitoria 130.

El sustituidor de señal transitoria 130 puede además, opcionalmente, suministrar una información sobre transitorios 134 que representa el contenido transitorio de la porción de señal transitoria (que es sustituida por la porción de señal de sustitución en la señal de audio de transitorios reducidos 132). Por consiguiente, la información sobre transitorios 134 puede servir para "salvar" el contenido transitorio de la señal de audio 110, el cual es reducido o incluso totalmente suprimido en la señal de audio de transitorios reducidos 132. La información sobre transitorios 134 puede mandarse directamente al reinsertador de señal transitoria 150, para servir de señal transitoria 152. Sin embargo, el aparato 100 puede comprender además un procesador de transitorios opcional 160 que es configurado para procesar la información sobre transitorios 134, para derivar de la misma la señal transitoria 152. Por ejemplo, el procesador de transitorios 160 puede ser configurado para efectuar una

transposición de frecuencia de transitorios, un desfase de frecuencia de transitorios, o una síntesis de transitorios.

El aparato 100 puede además comprender, opcionalmente, un acondicionador de señal 170 configurado par acondicionar la señal de audio procesada 120, para obtener una señal de audio acondicionada para la reproducción.

En cuanto a la funcionalidad del aparato 100, se puede decir de modo general que el aparato 100 permite un procesamiento separado de un contenido audio no transitorio de la señal de audio 110 (representado por la señal de audio de transitorios reducidos 132), y de un contenido audio transitorio de la señal de audio 110 (representado por la información sobre transitorios 134). Los eventos transitorios son reducidos, o incluso suprimidos, en la señal de audio de transitorios reducidos 132, de modo que el procesador de señal 140 pueda efectuar un procesamiento de señal que degradaría los eventos transitorios y/o que resultaría afectado de modo perjudicial por los eventos transitorios. Sin embargo, al sustituir las porciones de señal transitorias por porciones de señal de sustitución adaptadas en energía, el sustituidor de señal transitoria 130 sirve para evitar artefactos audibles que serían introducidos por el procesador de señal 140 si las porciones de señal transitorias fuesen sencillamente puestas a cero.

Una impresión de escucha apropiada también se obtiene mediante una reinserción de transitorios por el reinsertador de señal transitoria 150. Obviamente, una impresión de escucha sería, típicamente, seriamente degradada si los eventos transitorios fuesen sencillamente eliminados. Por este motivo, los transitorios son reinsertados en la señal de audio procesada 142. Los transitorios reinsertados pueden ser idénticos a los transitorios eliminados de la señal de audio 110 por el sustituidor de señal transitoria 130. Alternativamente, un procesamiento de dichos transitorios eliminados (o sustituidos) puede realizarse, por ejemplo, en forma de transposición de frecuencia o desfase de frecuencia. Sin embargo, en algunos modos de realización, los transitorios reinsertados pueden incluso generarse de modo sintético, por ejemplo en base a parámetros transitorios que describen un tiempo y una intensidad de los transitorios a reinsertar.

Detalles del sustituidor de señal transitoria

A continuación se describe la funcionalidad del sustituidor de señal transitoria 130, con referencia a la figura 2, donde la figura 2 ilustra un diagrama esquemático de un modo de realización del sustituidor de señal transitoria 130. El sustituidor de señal transitoria 130 recibe la señal de audio 110 y suministra, en base a la misma, la señal de audio de transitorios reducidos 132.

A tal fin, el sustituidor de señal transitoria 130 puede, por ejemplo, comprender un detector de transitorios 130a que es configurado para detectar un transitorio y suministrar una información sobre una sincronización del transitorio. Por ejemplo, el detector de transitorios 130a puede suministrar una información 130b que describe un momento de inicio y un momento de fin de una porción de señal transitoria. Distintos conceptos de detección de un transitorio son conocidos en el arte, de modo que una descripción detallada se omitirá aquí. Sin embargo, en algunos casos, el detector de transitorios 130a puede ser configurado para distinguir transitorios de longitud distinta, de modo que la longitud de una porción de señal transitoria reconocida puede variar en función de la forma de señal real.

Alternativamente, el sustituidor de señal transitoria puede comprender un extractor de informaciones laterales 130c, por ejemplo, si una información lateral que describe una duración de transitorios es asociada con la señal de audio 110. En tal caso, el detector de transitorios 130a puede obviamente omitirse. El extractor de informaciones laterales 130c puede, opcionalmente, además ser configurado para suministrar uno o varios parámetros de interpolación, parámetros de extrapolación y/o parámetros de sustitución en base a las informaciones laterales asociadas con la señal de audio 110. El sustituidor de transitorios 130 comprende además un sustituidor de porción transitoria 130d, por ejemplo un interpolador de porción transitoria o un extrapolador de porción transitoria. El sustituidor de porción transitoria 130e es configurado para recibir la señal de audio 110 y la información sobre la duración del transitorio 130b (suministrada por el detector de transitorios 130a o por el extractor de informaciones laterales 130c) y para sustituir una porción transitoria de la señal de audio 110 por una porción de señal de sustitución.

A continuación se describen detalles relacionados con la detección y sustitución (o eliminación) de transitorios. En particular, distintos métodos de eliminación de transitorios son descritos en detalle.

Los transitorios (por ejemplo el comienzo de un instrumento o de señales de percusión) pueden generalmente describirse como siendo un corto intervalo de tiempo durante el cual la señal se desarrolla rápidamente de un modo imprevisible. Por ejemplo, un transitorio puede ser detectado (mediante el detector de transitorios 130a) evaluando una representación en el dominio temporal de la señal de audio 110. Si la representación en el dominio temporal de la señal de audio 110 excede un umbral (que puede ser variable en el tiempo), la presencia de un evento transitorio puede ser indicada. Una región temporal que comprende el evento transitorio puede ser considerada como una porción de señal transitoria, y puede ser descrita por una información de duración de transitorio 130b.

Dado que tales porciones de señal (es decir, transitorios o intervalos de tiempo durante los cuales la señal se desarrolla rápidamente de un modo imprevisible), de modo ideal, no deben ser alargadas en el tiempo, resulta ventajoso eliminar de la señal “un período de tiempo transitorio” antes del alargamiento en el tiempo (que puede ser realizado por el procesador de señal 140). La supresión puede ocurrir durante todo el período de tiempo que es considerado “no estacionario”. Para los instrumentos de percusión, este período de tiempo consiste lo más a menudo en todo el evento sonoro (por ejemplo un solo toque de hi-hat). Para el comienzo de un instrumento, una envolvente ADSR (Attack Decay Sustain Release) puede servir para ilustrar el período de tiempo transitorio.

La figura 8 ilustra una representación gráfica 800 de una evolución en el tiempo de una amplitud de señal. Una abscisa 810 describe un tiempo, y una ordenada 812 describe una amplitud. Una curva 814 describe una evolución en el tiempo de la amplitud. Tal y como se puede ver en la figura 8, la evolución en el tiempo de la amplitud comprende un intervalo de ataque, un intervalo de decaimiento, un intervalo de sostenido y un intervalo de liberación. El intervalo de ataque y el intervalo de decaimiento pueden, por ejemplo, ser considerados como una “región transitoria” o porción de señal transitoria.

Sin embargo, se ha constatado que para un procesamiento de señal ulterior (por ejemplo en el procesador de señal 140), el lapso libre en la señal de audio causado por la supresión de transitorios debería llenarse de modo que, cuando se escucha la señal procesada (=señal de síntesis) (por ejemplo procesada mediante el procesador de señal 140), el auditorio tenga la sensación de una señal libre, transitoria, continua sin pausas disruptivas y modulaciones de amplitud.

Para el caso de aplicación específico descrito aquí, es preferible suprimir todas las porciones transitorias de la señal original (por ejemplo, señal 110) en la señal de síntesis (por ejemplo, en la señal 132 suministrada por el procesador de señal 140 o, por consiguiente, en la señal 142 suministrada por el procesador de señal 140), mientras que las porciones tonales y las componentes de ruido no transitorias siguen existiendo.

En este respecto, hay varios enfoques que ya existen, pero que nunca tienen por objeto una señal de transitorios ajustados de alta calidad (o de transitorios eliminados). En cuanto a este punto, refiérase a la publicación [Edler], por ejemplo.

En cuanto a la eficacia de los métodos de detección de transitorios y la descomposición en varias componentes, tales como por ejemplo “transitorios + ruido”, las siguientes conclusiones pueden derivarse de las respectivas publicaciones especializadas [Bello] y [Daudet], que ofrecen una buena visión global de los métodos comunes: ninguno de los métodos es claramente superior a los demás; la selección deberá hacerse según la respectiva aplicación y la potencia de cálculo disponible.

De allí resulta que la selección de los métodos de detección y descomposición específicos puede influir de modo importante sobre el resultado del método de la invención. Para el hombre del arte, resulta fácilmente posible aplicar cualquiera de los métodos conocidos para crear la mejor condición posible para el respectivo escenario de aplicación.

Conceptos de sustitución de porción transitoria

Algunos escenarios de aplicación se refieren a la generación de porciones de señal que no deben ser evaluadas como “correctas” o “erróneas” mediante comprobación con una señal de referencia, sino sólo en base a su buen sonido general. Eso significa que los modos de realización según la invención no son limitados a la separación de las porciones, y a la omisión de las componentes transitorias, pero pueden generar ellos mismos señales de síntesis que tienen propiedades específicas.

La generación de señales de síntesis (por ejemplo, generación de una señal de transitorios reducidos 132 por el sustituidor de señal transitoria 130d) puede, por lo tanto, ser una combinación de descomposición de señal y de generación de señal (en el sentido de una interpolación y/o extrapolación de la señal supuesta) durante el período de tiempo transitorio. Las componentes no transitorias de la señal original pueden ser mezcladas con las componentes interpoladas/extrapoladas, o pueden sustituir estas últimas.

En algunos modos de realización según la presente invención, la extrapolación puede ser igual a una generación de señal de síntesis mediante valores pasadas. Por consiguiente, la extrapolación puede ser capaz de tiempo real. Por otro lado, en algunos modos de realización, la interpolación puede ser igual a una generación de señal de síntesis mediante valores precedentes y sucesivos. Así, en algunos casos, la interpolación puede requerir una predicción.

Para resumir lo que precede, distintos conceptos pueden aplicarse en el sustituidor de porciones transitorias 130d, para obtener la señal de audio de transitorios reducidos 132.

Por ejemplo, el sustituidor de porciones transitorias 130d puede ser configurado para reducir las componentes transitorias

de la señal de audio 110, para obtener la señal de audio de transitorios reducidos. En tal caso, el sustituidor de porciones transitorias 130d puede ser configurado para asegurar que una energía suficiente permanezca en la porción de señal de sustitución, sustituyendo a la porción de señal transitoria. Por ejemplo, las componentes de frecuencia que comprenden una característica de fase transitoria pueden ser eliminadas de la señal de audio 110, mientras que otras componentes de frecuencia que no comprenden la característica de fase transitoria (por ejemplo, componentes de frecuencia tonales) pueden ser transferidas de la porción de señal transitoria a la porción de señal de sustitución. Por consiguiente, se puede asegurar que la porción de señal de sustitución comprenda una energía de señal suficiente, que no se desvía demasiado de la energía de señal de las porciones de señal precedentes y sucesivas.

Alternativamente, el sustituidor de porciones transitorias 130d puede ser configurado para obtener la porción de señal de sustitución destruyendo la relación de fase de formación de transitorio en la porción de señal transitoria. Por ejemplo, el sustituidor de porciones transitorias puede ser configurado para aleatorizar o ajustar (de modo determinista) la fase de las distintas componentes de frecuencia de la porción de señal transitoria. Por consiguiente, la porción de señal de sustitución obtenida de este modo puede comprender (al menos aproximadamente) la misma energía que la porción de señal transitoria (dado que la modificación de fase de las componentes de frecuencia no cambia la energía). Sin embargo, la evolución en el tiempo con formación de transitorios de la señal en el tiempo descrita por la porción de señal de sustitución puede perderse debido a la evolución en el tiempo de los transitorios en base a una relación de fase específica de distintas componentes de frecuencia que es destruida.

Alternativamente, sin embargo, el sustituidor de porciones transitorias 130d puede interpolar, por ejemplo, una evolución en el tiempo de la energía en distintas bandas de frecuencias en base a una porción de señal no transitoria que precede a la porción de señal transitoria. Por consiguiente, el contenido de la porción de señal de sustitución puede sencillamente basarse en una extrapolación del contenido de una porción de señal no transitoria que precede a la porción de señal transitoria. Por consiguiente, el contenido de la porción de señal transitoria puede ser ignorado por completo.

Alternativamente, sin embargo, el contenido de la porción de señal de sustitución puede obtenerse, mediante el sustituidor de porciones transitorias 130d, interpolando entre un contenido de una porción de señal no transitoria que precede a la porción de señal transitoria y una porción de señal no transitoria que sigue a la porción de señal transitoria. Aquí también, el contenido de la porción de señal transitoria puede ser ignorado por completo. La interpolación puede efectuarse, por ejemplo, en un dominio temporal-frecuencial.

Alternativamente, sin embargo, una combinación de los métodos descritos más arriba puede utilizarse para obtener el contenido de la porción de señal de sustitución. Por ejemplo, un contenido no transitorio de la porción de señal transitoria (extraída por ejemplo eliminando el contenido transitorio o destruyendo la relación de fase de formación de transitorio) puede ser combinado con un contenido de señal de audio obtenido interpolando o extrapolando una o varias porciones de señal transitorias. Como otro ejemplo, una relación de fase de formación de transitorio en una porción de señal transitoria puede ser destruida y una energía de la porción de señal transitoria puede ser modulada para adaptarse a una energía de porciones de señal no transitorias adyacentes.

Visto lo que precede, se puede decir que la porción de señal de sustitución es sintetizada o bien sólo en base a porciones de señal no transitorias (por ejemplo que preceden y/o siguen la porción de señal transitoria) (sin utilizar el contenido de la porción de señal transitoria), sólo en base a la porción de señal transitoria, o en base a una combinación de una o varias porciones de señal no transitorias y de la porción de señal transitoria.

Otro concepto de generación de la señal de audio de transitorios reducidos - fundamentos

A continuación se describe otro concepto de generación de la señal de audio de transitorios reducidos 132, aspectos del cual pueden aplicarse en cualquiera de los modos de realización descritos aquí. En cuanto al proceso de detección y sustitución, refiérase al documento WO 2007/118533.

El documento WO 2007/118533 A1 describe un aparato y un proceso para la producción de una señal de zona circundante. Este documento describe un detector de transitorios que está previsto para detectar un período de tiempo transitorio. El detector de transitorios descrito en el documento WO 2007/118533 A1 puede, por ejemplo, ser utilizado para implementar (o sustituir) el detector de transitorios 130a descrito aquí. Dicha publicación describe además un generador de señal de síntesis que produce una señal de síntesis que cumple con una condición de transitorio y una condición de continuidad. El generador de síntesis descrito en el documento WO 2007/118533 A1 puede, por ejemplo, ser utilizado para implementar el sustituidor de porciones transitorias 130d, o puede incluso ocupar el lugar del sustituidor de porciones transitorias 130d. Así, el concepto descrito en el documento WO 2007/118533 A1 para la generación de una señal de síntesis puede ser utilizado para la generación de la señal de audio de transitorios reducidos 132 en algunos modos de realización de la presente invención.

Otro concepto de generación de la señal de audio de transitorios reducidos - extensiones

Dado que, en la aplicación descrita aquí (procesamiento de una señal que comprende un transitorio, al tiempo que se mantiene una buena impresión de escucha), una alta calidad audio de la señal resultante es sustancialmente más crítica que en la aplicación del documento WO 2007/118533 (Ambient Señal Generation), el método descrito en el documento WO 2007/118533 está expandido con algunas etapas, para mejorar la calidad de la señal de audio.

Por ejemplo, además de la extrapolación de amplitud, un modo de realización según la presente invención puede comprender también la extrapolación o la interpolación de los valores de fase de modo que se obtenga una señal de síntesis de calidad mejorada que no presenta porciones transitorias.

La extrapolación o interpolación se efectúa, por ejemplo, mediante una predicción lineal o una codificación de predicción lineal (LPC), o linealmente y/o con splines o similares + ruido ponderado.

En algunos modos de realización, la generación descrita más arriba de la señal de audio de transitorios reducidos puede ser particularmente ventajosa cuando se utiliza en combinación con un vocodificador de fase, que puede ser parte del procesador de señal, o que puede constituir el procesador de señal. En algunos modos de realización se explota la propiedad del vocodificador de fase – que se considera usualmente ser un grand problema [8] – que consiste en que no existe ninguna relación previsible con las tramas precedentes durante los transitorios. En algunos modos de realización, este mismo hecho se explota para suprimir el transitorio en el sentido de que el transitorio es borrado forzando una relación con los bins precedentes. En otras palabras, la fase de distintos coeficientes que describen los distintos bins tiempo-frecuencia de la porción de señal de sustitución (por ejemplo en forma de números complejos) es, por ejemplo, ajustada mediante extrapolación a partir de bins tiempo-frecuencia precedentes (de una porción de señal no transitoria precedente), o interpolación entre bins tiempo-frecuencia correspondientes de una porción de señal no transitoria precedente y de una porción de señal no transitoria siguiente. En la publicación [Maher] se describe un método de interpolación comparable. El método presentado en [Maher] no es capaz de tiempo real, dado que se requieren también las porciones que siguen el lapso libre de señal. Además, [Maher] sólo describe el procesamiento de los “picos” en una señal de audio (al revés, algunos modos de realización según la invención procesan todas las líneas de frecuencia), y tampoco son tratadas de modo explícito las componentes de ruido. En otras palabras, en algunos modos de realización, el concepto descrito en [Maher] para puentear los lapsos libres en una señal de audio puede aplicarse con la presente aplicación para obtener la señal de audio de transitorios reducidos, en base a la señal de entrada audio original. En vez de puentear una porción “que falta” de una señal de audio, una porción identificada como porción de señal transitoria puede ser sustituida mediante el método descrito en [Maher]. Sin embargo, la interpolación/extrapolación puede realizarse independientemente para cada “bin” de frecuencia. Opcionalmente, la amplitud y la fase pueden ser interpoladas (por ejemplo por separado).

Detector de transitorios 130a

A continuación se describen algunos detalles presentes relacionados con el detector de transitorios 130a. Se notará sin embargo que se pueden utilizar muchas implementaciones distintas del detector de transitorios 130a, de modo que los siguientes detalles deberán ser considerados como ejemplos de una implementación ventajosa. En algunos modos de realización, umbrales adaptativos son preferidos para el reconocimiento de los períodos de tiempo transitorio. Normalmente, los umbrales adaptativos son versiones alisadas de una función de detección que pueden resultar en grandes fluctuaciones y, por lo tanto, en una falta de detección de pequeños picos en el entorno de grandes picos. Para los detalles, refiérase a la publicación [Bello]. Este problema puede solucionarse, por ejemplo, por una adaptación adecuada de las constantes de alisado en función de la condición detectada actual (región transitoria / región no transitoria) y del desarrollo de la función de detección (por ejemplo ataque, decaimiento).

A continuación son dadas algunas referencias de literatura relacionadas con los arriba mencionados aspectos: [Edler], [Bello], [Goodwin], [Walther], [Maher], [Daudet].

Extractor de porción transitoria 130e

Además de las funcionalidades descritas más arriba, el sustituidor de señal transitoria 130 puede comprender además un extractor de porción transitoria 130e, extractor de porción transitoria 130e que puede ser configurado para recibir la señal de audio 110 (o al menos la porción de señal transitoria de la misma), y para suministrar la información sobre transitorios 134. El extractor de porción transitoria 130e puede ser configurado para suministrar la información sobre transitorios 134 en cualquier forma posible, por ejemplo en forma de señal temporal de porción de señal transitoria, en forma de representación en el dominio frecuencial de porción de señal transitoria, o en forma de parámetros transitorios (por ejemplo una información sobre la duración del transitorio y/o una información sobre la intensidad del transitorio y/o una información sobre la pendiente del transitorio y/o cualquier otra información apropiada sobre transitorios).

En particular, el extractor de porción transitoria 130e puede ser configurado para suministrar la información sobre transitorios 134 sólo para las porciones de señal que han sido eliminados de la señal de audio 110 para obtener la señal de audio de transitorios reducidos 132, para mantener el caudal de datos razonablemente pequeño.

Alternativas de Implementación del Procesador de Señal 140 – Visión global

A continuación se describen varios conceptos básicos para la implementación del procesador de señal 140. La figura 3a ilustra una implementación preferida del procesador de señal 140 de la figura 1. Esta implementación comprende un analizador selectivo en frecuencia 310 y un dispositivo de procesamiento selectivo en frecuencia 312 conectado a valle que es implementado de modo que suministre una influencia negativa sobre la "coherencia vertical" de la señal de audio original. Un ejemplo de este procesamiento selectivo en frecuencia es el alargamiento de una señal en el tiempo o el acortamiento de una señal en el tiempo, donde este alargamiento o acortamiento es aplicado de modo selectivo en frecuencia, de modo que, por ejemplo, el procesamiento introduzca desfases en la señal de audio procesada, los cuales son distintos para bandas de frecuencias distintas. Los desfases pueden, por ejemplo, ser introducidos de modo que los transitorios sean degradados. El procesador de señal 140 ilustrado en la figura 3a pueden, opcionalmente, comprender además un combinador de frecuencias 314 que es configurado para combinar las distintas componentes de frecuencia de la señal de audio procesada suministrada por el procesamiento selectivo en frecuencia 312, para obtener una sola señal (por ejemplo una señal en el dominio temporal).

Tanto el analizador selectivo en frecuencia 310, que puede dividir la señal de audio de transitorios reducidos 132 en una pluralidad de componentes de frecuencia (por ejemplo coeficientes espectrales de valor complejo), como el combinador de frecuencias 314, que puede ser configurado para obtener la representación en el dominio temporal de la señal de audio procesada 142 en base a una pluralidad de coeficientes espectrales de valor complejo para distintas bandas de frecuencias, pueden ser configurados para realizar un procesamiento por bloque. Por ejemplo, el analizador selectivo en frecuencia 310 puede procesar un bloque (por ejemplo dividido en ventanas) de muestras de la señal de audio 132, para obtener un conjunto de coeficientes espectrales de valor complejo que representan el contenido audio del bloque de muestras de señal de audio. Asimismo, el combinador de frecuencias opcional 314 puede recibir un conjunto de coeficientes de valor complejo (por ejemplo uno para cada banda de frecuencias de entre una pluralidad de bandas de frecuencias) y para suministrar, en base al mismo, una representación en el dominio temporal sobre un intervalo de tiempo limitado que comprende una pluralidad de muestras en el dominio temporal.

Otro procesamiento de señal preferido es ilustrado en la figura 3b en el contexto de un procesamiento de vocodificador de fase. De modo general, un vocodificador de fase comprende un analizador de sub-banda/transformación 320, un procesador conectado a valle 322 para realizar un procesamiento selectivo en frecuencia de una pluralidad de señales de salida suministradas por el analizador 320, y, a continuación, un combinador de sub-banda/transformación 324 que combina las señales procesadas por el procesador 322, para finalmente obtener una señal procesada 142 en el dominio temporal a una salida 326. La señal procesada 142 en el dominio temporal es, aquí también, una señal de anchura de banda total para una señal de filtro de paso bajo, a la condición de que la anchura de banda de la señal procesada 142 sea superior a la anchura de banda representada por una sola rama entre los elementos 322 y 324, dado que el combinador de sub-banda/transformación 324 efectúa una combinación de señales selectivas en frecuencia.

Otros detalles relacionados con el vocodificador de fase serán descritos a continuación, en relación con las figuras 5a, 5b, 5c y 6.

La figura 3c ilustra otra implementación posible del procesador de señal 140. Tal y como se puede ver, la señal de audio de transitorios reducidos 132 puede incluso ser procesada en el dominio temporal en algunos modos de realización. De modo típico, el procesamiento en el dominio temporal 330 puede comprender una memoria, de modo que un transitorio en la señal 132 tendría un impacto de larga duración sobre la señal de audio procesada 142. En algunos casos, la señal de audio de transitorios reducidos 132 causaría una respuesta transitoria en la señal de audio procesada 142, que es considerablemente más larga (por ejemplo de un factor de 2, o incluso de un factor de 5, o incluso de un factor de 10) que la duración del transitorio (o la duración de la porción de señal transitoria). En tal caso, los transitorios en la señal de audio 132 degradarían considerablemente, de modo no deseado, la señal de audio procesada 142, por ejemplo produciendo ecos audibles. Por otro lado, una supresión completa de una porción de señal transitoria tendría también un impacto de larga duración sobre la señal de audio procesada 142, ya que una supresión completa de una porción de señal transitoria causa un transitorio en sí.

Implementación del Procesador de Señal mediante un Vocodificador - Implementación de Banco de Filtros

A continuación son ilustradas, con referencia a las figuras 5 y 6, implementaciones preferidas de un vocodificador que pueden ser utilizadas para una implementación del procesador de señal 140, o que puede ser parte del procesador de señal 140. La figura 5a ilustra una implementación de banco de filtros de un vocodificador de fase, donde una señal de entrada

audio (por ejemplo la señal de audio de transitorios reducidos 132) es alimentada a una entrada 500 y una señal de audio procesada (por ejemplo la señal de audio procesada 142) es obtenida a una salida 510. En particular, cada canal del banco de filtros esquemático ilustrado en la figura 5a comporta un filtro de paso de banda 501 y un oscilador a valle 502. Las señales de salida de todos los osciladores de cada canal son combinadas por un combinador, que es implementado, por ejemplo, como sumador e indicado en 503, para obtener la señal de salida a la salida 510. Cada filtro 501 es implementado de modo que suministre una señal de amplitud, por un lado, y una señal de frecuencia, por otro lado. La señal de amplitud y la señal de frecuencia son señales temporales que ilustran un desarrollo de la amplitud en un filtro 501 en el tiempo, mientras que la señal de frecuencia representa un desarrollo de la frecuencia de la señal filtrada por un filtro 501.

Un montaje esquemático del filtros 501 es ilustrado en la figura 5b. Cada filtro 501 de la figura 5a puede ser montado tal y como ilustrado en la figura 5b, donde, sin embargo, sólo las frecuencias f_i alimentadas hacia los dos mezcladores de entrada 551 y el sumador 552 son distintas de un canal a otro. Las señales de salida de mezclador son ambas filtradas paso bajo por filtros de paso bajo 553, donde las señales de paso bajo son distintas, en la medida en que fueron generadas por señales de oscilador local que son desfasadas de 90° . El filtro de paso bajo superior 553 suministra una señal en cuadratura 554, mientras que el filtro inferior 553 suministra una señal en fase 555. Estas dos señales, es decir I y Q, son alimentadas hacia un transformador de coordenadas 556 que genera una representación de fase de magnitud a partir de la representación rectangular. La señal de magnitud, o respectivamente la señal de amplitud, de la figura 5a en el tiempo es salida a una salida 557. La señal de fase es alimentada hacia un desacondicionador de fase 558. A la salida del elemento 558, ya no hay ningún valor de fase que se sitúe siempre entre 0 y 360° , sino un valor de fase que aumenta linealmente. Este valor de fase "desacondicionado" es alimentado hacia un convertidor de fase/frecuencia 559 que puede ser implementado, por ejemplo, como un mero formador de diferencia de fase que sustrae una fase de un punto anterior en el tiempo de una fase a un punto actual en el tiempo, para obtener un valor de frecuencia para el punto actual en el tiempo. Este valor de frecuencia es añadido al valor de frecuencia constante f_i del canal de filtro i , para obtener un valor de frecuencia variable en el tiempo a la salida 560. El valor de frecuencia a la salida 560 tiene una componente directa $= f_i$ y una componente alternativa $=$ la desviación de frecuencia de la cual una frecuencia actual de la señal en el canal de filtro se desvía de la frecuencia media f_i .

Así, tal y como ilustrado en las figuras 5a y 5b, el vocodificador de fase logra una separación de la información espectral y de la información de tiempo. La información espectral se encuentra en el canal especial o en la frecuencia f_i que suministra la porción directa de la frecuencia para cada canal, mientras que la información de tiempo está contenida en la desviación de frecuencia o respectivamente la magnitud en el tiempo.

La figura 5c ilustra una manipulación que puede realizarse en el vocodificador en el sitio del vocodificador trazado con trazos interrumpidos en la figura 5a.

Para la modulación en el tiempo, por ejemplo las señales de amplitud $A(t)$ en cada canal o la frecuencia de las señales $f(t)$ en cada señal pueden ser decimadas o respectivamente interpoladas. A los fines de la transposición, ya que es útil para la presente invención, se efectúa una interpolación, es decir, una extensión en el tiempo de las señales $A(t)$ y $f(t)$, para obtener señales extendidas $A'(t)$ y $f'(t)$, donde la interpolación es controlada por un factor de extensión. La frecuencia de cada oscilador individual 502 en la figura 5a no cambia por la interpolación de la variación de fase, es decir, el valor antes de la adición de la frecuencia constante por el sumador 552. El cambio en el tiempo de la señal de audio total es sin embargo decelerado, es decir, por el factor 2. El resultado es un tono extendido en el tiempo que presenta el tono original, es decir, la onda fundamental original con sus armónicos.

Para la transposición de frecuencia puede utilizarse el siguiente concepto. Al realizar el procesamiento de señal ilustrado en la figura 5c, donde tal procesamiento es realizado en cada canal de banda de filtro en la figura 5a, y al decimar la señal temporal resultante en un decimador, la señal de audio puede ser retractada a su duración original, mientras que todas las frecuencias son simultáneamente dobladas. Eso conduce a una transposición de tono por el factor 2, donde, sin embargo, se obtiene una señal de audio que tiene la misma longitud que la señal de audio original, es decir, el mismo número de muestras.

Implementación del Procesador de señal mediante un Vocodificador - Implementación de Transformación

Como alternativa a la implementación de banco de filtros ilustrada en la figura 5a, también se puede utilizar una implementación de transformación de un vocodificador de fase, tal y como ilustrado en la figura 6. Aquí, la señal de audio 132 es alimentada hacia un procesador de FFT o, más generalmente, un procesador de Transformación de Fourier de Corta Duración 600 como secuencia de muestras temporales. El procesador FFT 600 es implementado de modo esquemático en la figura 6 para efectuar una división en ventanas temporales de una señal de audio para, a continuación, mediante una FFT, calcular la magnitud y la fase del espectro, donde este cálculo se efectúa para espectros sucesivos que están relacionados con bloques de la señal de audio que se solapan fuertemente.

En un caso extremo, para cada nueva muestra de señal de audio puede calcularse un nuevo espectro, donde un nuevo espectro puede también calcularse, por ejemplo, sólo para cada veintésima muestra. Esta distancia a de muestras entre dos espectros es dada, preferiblemente, por un controlador 602. El controlador 602 está implementado además para alimentar un procesador IFFT 604 que está implementado para funcionar en una operación de solapamiento. En particular, el procesador IFFT 604 está implementado de modo que efectúe una Transformación de Fourier de corta duración inversa efectuando una IFFT por espectro en base a la magnitud y la fase de un espectro modificado, para efectuar a continuación una operación de solapamiento-adición, a partir de la cual se obtiene la señal temporal resultante. La operación de solapamiento-adición elimina los efectos de la ventana de análisis.

Un alargamiento de la señal temporal es obtenido por que la distancia b entre dos espectros, a medida de su procesamiento por el procesador IFFT 604, es superior a la distancia a entre los espectros durante la generación de los espectros FFT. La idea de base es la de alargar la señal de audio por el mero hecho de que las FFT inversas son más distanciadas que las FFT de análisis. Por consiguiente, los cambios en el tiempo en la señal de audio sintetizada se producen más lentamente que en la señal de audio original.

Sin embargo, sin modulación de fase en el bloque 606, esto conduciría a artefactos. Si se toma en consideración, por ejemplo, un solo bin de frecuencia para el cual son implementados valores de fase sucesivos de 45° , esto implica que la señal en este banco de filtros aumenta en la fase a una velocidad de $1/8$ de ciclo, es decir, de 45° por intervalo de tiempo, donde el intervalo de tiempo es aquí el intervalo de tiempo entre FFT sucesivas. Si las FFT inversas son más espaciadas entre sí, esto significa que el aumento de fase de 45° se produce en un intervalo de tiempo más largo. Esto significa que, debido al desfase, se produce una desadaptación en el proceso de solapamiento-adición que sigue, conduciendo a una cancelación de señal no deseada. Para eliminar este artefacto, la fase es modulada de exactamente el mismo factor que aquel del cual la señal de audio fue alargada en el tiempo. La fase de cada valor espectral de FFT es así aumentada del factor b/a , de modo que se elimina esta desadaptación.

Mientras que en el modo de realización ilustrado en la figura 5c el alargamiento por interpolación de las señales de control de amplitud/frecuencia se obtuvo para un oscilador de señal en la implementación de banco de filtros de la figura 5a, el alargamiento en la figura 6 se obtiene por que la distancia entre dos espectros IFFT es superior a la distancia entre dos espectros FFT, es decir, que b es superior a a , donde, sin embargo, para prevenir un artefacto, se efectúa una remodelación de fase según b/a .

En cuanto a una descripción detallada de vocodificadores de fase, refiérase a los siguientes documentos:

"The phase Vocoder: A tutorial", Mark Dolson, Computer Music Journal, volumen 10, nº 4, páginas 14 - 27, 1986, o "New phase Vocoder techniques for pitch-shifting, harmonizing and other exotic effects", L. Laroche and M. Dolson, Proceedings 1999 IEEE Workshop on applications of signal processing to audio and acoustics, New Paltz, New York, 17 - 20 de octubre de 1999, páginas 91 a 94; "New approach to transient processing interphase vocoder", A. Röbel, Proceeding of the 6th international conference on digital audio effects (DAFx-03), Londres, RU, 8-11 de septiembre de 2003, páginas DAFx-1 a DAFx-6; "Phase-locked Vocoder", Meller Puckette, Proceedings 1995, IEEE ASSP, Conference on applications of signal processing to audio and acoustics, o Solicitud de Patente US Número 6.549.884.

A continuación se describe brevemente un ejemplo de la funcionalidad del vocodificador de fase en base a una transformación, con referencia a la figura 7. La figura 7 ilustra una representación esquemática del funcionamiento de un algoritmo de vocodificador de fase, siendo el tamaño de tramo de síntesis distinto del tamaño de tramo de análisis, por ejemplo de un factor de 2.

El algoritmo de vocodificador de fase (PV) es utilizado para modificar la duración de una señal sin alterar su tono [B9]. Divide una señal en granos que designan cortes divididos en ventanas de la señal de una longitud típicamente del orden de algunas decenas de milisegundos. Los granos son re-dispuestos en un proceso de solapamiento y adición (OLA) con un tamaño de tramo de síntesis que difiere del tamaño de tramo de análisis. Para alargar la señal por un factor de dos, por ejemplo, el tamaño de tramo de síntesis es de dos veces el tamaño de tramo de análisis. La figura 7 ilustra el algoritmo.

Reinsertador de señal transitoria

A continuación se describe una implementación preferida del reinsertador de señal transitoria 150 ilustrado en la figura 1, con referencia a la figura 4.

El reinsertador de señal transitoria 150 comprende, como componente clave, un combinador de señal 150a. El combinador de señal 150a es configurado para recibir tanto la señal de audio procesada 142 como la señal transitoria 152, y para suministrar, en base a las mismas, la señal de audio procesada 120. El combinador de señal 150a puede, por ejemplo, ser configurado para efectuar una dura sustitución de conmutación de una porción de la señal de audio procesada 142 por una

porción de la señal transitoria 152. Sin embargo, en un modo de realización preferido, el combinador de señal 150a puede ser configurado para formar un fundido encadenado entre la señal de audio procesada 142 y la señal transitoria 152, de modo que exista una transición suave entre dichas señales 142, 152 dentro de la señal de audio procesada 120.

5 Sin embargo, el reinsertador de señal transitoria 150 puede ser configurado para determinar un coeficiente de inserción óptimo. Por ejemplo, el reinsertador de señal transitoria 150 puede comprender un calculador 150b para calcular una longitud de la porción de reinsertión de transitorio. El cálculo de esta longitud de la porción de reinsertión de transitorio puede, por ejemplo, ser importante si la longitud de la porción transitoria sustituida (tal y como determinada, por ejemplo, por el detector de transitorios 130a) es variable en función de las características de señal. En el caso de que la señal de audio
10 procesada 142 comprende una longitud distinta (o un número distinto de muestras por segundo, o un número distinto de muestras generales), comparado con la señal de entrada audio original 110, un factor de alargamiento o factor de compresión puede ser tomado en consideración por el calculador 150b para determinar la longitud de la porción de reinsertión de transitorio. Una descripción detallada de esta variación de longitud será dada a continuación con referencia a las figuras 10 y 11.

15 El reinsertador de señal transitoria 150 puede, además, comprender un calculador 150c para calcular una posición de reinsertión. En algunos casos, el cálculo de la posición de reinsertión puede tener en cuenta un alargamiento o una compresión de la señal de audio procesada 142. En algunos casos, es preferible que una relación entre un contenido de señal de audio no transitoria y un contenido de señal transitoria (por ejemplo, relación temporal) en la señal de audio
20 procesada 120 sea al menos aproximadamente idéntica a la relación temporal de dicho contenido audio no transitorio y dicho contenido audio transitorio en la señal de entrada audio original 110. Sin embargo, además de un precálculo de la posición de señal transitoria reinsertión apropiada se puede efectuar un ajuste fino de dicha posición de reinsertión. Por ejemplo, el calculador 150c para calcular las posiciones de reinsertión puede ser configurado para leer tanto la señal de audio procesada 142 como la señal transitoria 152, y para determinar un momento de reinsertión en base a una
25 comparación de la señal de audio procesada 142 y de la señal transitoria 152. Los detalles relacionados con el posible cálculo de la posición de reinsertión serán descritos a continuación, con referencia a los ejemplos ilustrados en las figuras 10 y 11.

Posible relación de sincronización

30 A continuación se describen los detalles relacionados con una posible relación de sincronización, con referencia a la figura 9. La figura 9 ilustra una representación gráfica de un procesamiento de los distintos bloques de la señal de entrada audio original 110. Una primera representación gráfica 910 describe una evolución en el tiempo de la señal de entrada audio original 110, donde una abscisa 912 designa el tiempo. La señal de entrada audio 110 comprende una porción de señal transitoria 920 cuya longitud puede ser variable. Como referencia de sincronización, intervalos de procesamiento, o bloques
35 de procesamiento 922a, 922b, 922c, del procesador de señal 140 son ilustrados en la representación gráfica 910. Tal y como se puede ver, la duración de la porción de señal transitoria 920 puede ser inferior a la duración en el tiempo de los intervalos de procesamiento 922a, 922b, 922c. En algunos casos, sin embargo, la duración en el tiempo de la porción de señal transitoria puede incluso ser superior a la duración en el tiempo de los intervalos de procesamiento, o extender sobre
40 más de un solo intervalo de procesamiento. En algunos casos, los intervalos de procesamiento 922a, 922b, 922c pueden también solaparse en el tiempo.

Una representación gráfica 930 representa la señal de audio de transitorios reducidos 132 que puede obtenerse por la sustitución de transitorios realizada por el sustituidor de señal transitoria 130. Tal y como se puede ver, la porción de señal transitoria 920 fue sustituida por una porción de señal de sustitución.

Una representación gráfica 950 describe la señal de audio procesada 142 que puede obtenerse, por ejemplo, mediante un procesamiento por bloque de la señal de audio de transitorios reducidos 132. El procesamiento puede, por ejemplo, efectuarse mediante un vocodificador de fase y una bajada en frecuencia. En este procesamiento, los bloques pueden,
50 opcionalmente, ser divididos en ventanas, solapándose, opcionalmente, también los bloques.

Otra representación gráfica 970 representa la señal de audio procesada 120 en la cual el transitorio (o una versión modificada del mismo) fue reinsertado por el reinsertador de señal transitoria 150.

55 Es importante notar que la porción de señal transitoria 920 tendría un impacto sobre todo el bloque 1" si la porción de señal transitoria 920 fuese tomada en consideración en el procesamiento por bloque, dado que la energía de transitorio se expandiría, típicamente, a todo el bloque en tal procesamiento por bloque. Así, si the porción de señal transitoria debiese ser tomada en consideración en el procesamiento por bloque, la energía total del bloque sería posiblemente falseada por la energía del transitorio. Además, el transitorio se expandiría típicamente (es decir, se extendería), si el transitorio fuese
60 afectado por el procesamiento por bloque. Al revés, el procesamiento por separado del transitorio permite limitar el impacto del transitorio a un intervalo de tiempo 1" de la señal de audio procesada 120 que está asociado con el transitorio. Una

expansión de la porción de señal transitoria hacia un bloque entero del procesamiento por bloque de la señal en el procesador de señal 140 puede ser evitado. Al contrario, la duración de la porción de señal transitoria en la señal de audio procesada 120 puede ser determinada por el procesamiento del transitorio realizado por el procesador de transitorios 160. Alternativamente, resulta posible insertar la porción de señal transitoria 920 en la señal de audio procesada 142 en su duración original, si tal fuese deseado. Así, una expansión no deseada de la energía del transitorio en el procesador de señal 140 puede ser evitada.

Alargamiento del tiempo de la señal de audio

Tal y como se puede ver de la descripción más arriba, el concepto de la invención para la manipulación de una señal de audio que comprende un evento transitorio puede aplicarse en distintas aplicaciones. Por ejemplo, dicho concepto puede aplicarse en cualquier procesamiento de señal de audio en el cual los transitorios serían degradados por el procesamiento de señal y en el cual es, no obstante, deseable mantener los transitorios. Por ejemplo, muchos tipos de procesamiento de señal de audio no lineal resultarían en resultados seriamente degradados en presencia de transitorios. Algunos tipos de filtrado temporal serían, además, considerablemente afectados por la presencia de transitorios. Además, cualquier procesamiento por bloque de una señal de audio sería, típicamente, degradado por la presencia de transitorios, dado que la energía de los transitorios sería maculada en un bloque de procesamiento entero, resultando así en artefactos audibles.

No obstante, un alargamiento en el tiempo de señales audio puede ser considerado como una aplicación particularmente importante del presente concepto de manipulación de una señal de audio que comprende un evento transitorio. Es la razón por la cual se describirán a continuación los detalles relacionados con esta aplicación.

A continuación se describen algunos inconvenientes de los conceptos convencionales para el alargamiento en el tiempo de señales, para permitir una comprensión de las ventajas del concepto de la invención. El alargamiento en el tiempo de señales audio por el vocodificador de fase comprende un "maculado" de porciones de señal transitorias por dispersión, dado que se degrada la coherencia vertical (en el sentido de una relación de fase específica entre componentes de distintas bandas de frecuencias) de la señal. Los procesos que funcionan con métodos de solapamiento-adición (OLA) pueden generar pre-ecos disruptivos y ecos retrasados de eventos sonoros transitorios. Estos problemas pueden en efecto producirse con un alargamiento más pronunciado en el tiempo en el entorno de transitorios. Sin embargo, si una transposición debe tener lugar, el factor de transposición ya no será constante en el entorno de los transitorios, es decir, el tono de constituyentes de señal (posiblemente tonales) superpuestos cambiará y será percibido como siendo disruptivo.

Si los transitorios son eliminados y el lapso libre resultante es alargado, un lapso libre muy grande deberá ser llenado a continuación. Si los transitorios se suceden de cerca, los grandes lapsos libres pueden posiblemente solaparse.

A continuación se describe un nuevo procedimiento de transformación de señales. El procedimiento presentado aquí resuelve los arriba mencionados problemas.

Según un aspecto de este procedimiento, un segmento dividido en ventanas que contiene el transitorio es interpolado o extrapolado a partir de la señal a manipular (por ejemplo, la señal de entrada audio original 110). Si la aplicación es crítica en cuanto al tiempo, es decir, si se debe evitar un retraso, se podrá elegir preferiblemente una extrapolación. Si se conoce el futuro como predicción y el retraso no juega un papel demasiado importante, se preferirá una interpolación.

En algunos modos de realización, el procedimiento puede consistir esencialmente en las siguientes etapas, y es ilustrado en las figuras 10 y 11.

1. Reconocimiento del transitorio;
2. Determinación de la longitud del transitorio;
3. Se guarda el transitorio;
4. Extrapolación y/o interpolación;
5. Aplicación del procedimiento propiamente dicho, por ejemplo vocodificador de fase;
6. Reinserción del transitorio guardado; y
7. Posiblemente (opcionalmente) nuevo muestreo (para modificación de la tasa de muestreo).

Una vez realizada esta secuencia, la duración del transitorio es acortada durante la bajada en frecuencia. Si esto no es

deseado, el transitorio puede ser modulado de modo que se sitúe en la banda de frecuencias deseada antes de que sea reinsertado tras la modulación por deslizamiento (etapas 6 y 7 intercambiadas).

5 A continuación se describen algunos detalles, con referencia a la figura 10. La figura 10 ilustra una representación gráfica de distintas señales que pueden aparecer en un modo de realización del aparato 100 según la figura 1. La representación de la figura 10 está designada en su totalidad con 1000. Una representación de señal 1010 describe una evolución en el tiempo de la señal de entrada audio original 110. Tal y como se puede ver, la señal de entrada audio 110 comprende una porción de señal transitoria 1012 cuya anchura variable (o duración) puede ser determinada por el detector de transitorios 130a de modo adaptado a la señal. La porción de señal transitoria 1012 puede ser eliminada por el sustituidor de señal transitoria 10
10 130, y puede ser sustituida por una porción de señal de sustitución. Por consiguiente, se puede obtener una señal de audio de transitorios reducidos 132, la cual es ilustrada en una representación señal 1020. Una porción de señal de sustitución es ilustrada en el número de referencia 1022, sustituyendo la porción de señal transitoria 1012. La señal de audio de transitorios reducidos 132 puede ser procesada en un modo por bloque, donde distintas ventanas de procesamiento (que determinan la granularidad del procesamiento por bloque, y son también designadas como "granos") son ilustradas en una
15 representación de señal 1030. Por ejemplo, para cada bloque (o "grano") se puede obtener un conjunto de coeficientes espectrales, de modo que se forme una representación en el dominio temporal-frecuencial de la señal de audio de transitorios reducidos 132. Un procesamiento por vocodificador de fase puede ser aplicado en la representación en el dominio temporal-frecuencial de la señal de audio de transitorios reducidos 132, de modo que se obtenga una señal de duración incrementada. A tal fin, se puede obtener coeficientes en el dominio temporal-frecuencial interpolados. Los
20 coeficientes en el dominio temporal-frecuencial pueden entonces ser utilizados para construir una señal en el dominio temporal cuya duración en el tiempo está extendida, comparado con la señal de entrada audio original, al tiempo que se mantiene el tono. En otras palabras, el número de períodos de señal es incrementado. La señal obtenida por la operación de vocodificador de fase es ilustrada en una representación de señal 1040. Tal y como se puede ver en la representación gráfica 1040, una "zona de transitorio recortada", en la cual se ha insertado una porción de señal de sustitución en sustitución de la porción de señal transitoria, es desplazada en el tiempo respecto a una posición en el tiempo de la porción de señal transitoria en la señal de entrada audio original 110 (considerada con referencia al inicio de la señal de entrada audio).

30 A continuación, la porción de señal transitoria, que fue previamente sustituida, es reinsertada, por ejemplo por el reinsertador de señal transitoria 150. Por ejemplo, la porción de señal transitoria descrita por la señal transitoria 152 puede ser encadenada en la versión procesada 142 de la señal de audio de transitorios reducidos. Un resultado de la reinsertión de transitorio es ilustrado en una representación gráfica 1050.

35 En un submuestreo sucesivo, una duración en el tiempo de la señal de audio procesada 120 puede ser reducida. El submuestreo puede, por ejemplo, ser realizado por el acondicionador de señal 170. El submuestreo puede, por ejemplo, comprender un cambio de la escala de tiempo. Alternativamente, un número de puntos de muestreo puede ser reducido. Por consiguiente, una duración en el tiempo de la señal submuestreada es reducida, comparada con una señal suministrada por el vocodificador de fase. Al mismo tiempo, un número de períodos puede ser mantenido por la bajada en frecuencia, comparado con la señal suministrada por el vocodificador de fase. Por consiguiente, el tono de la señal bajada en
40 frecuencia, que es ilustrada en una representación de señal 1050, puede ser incrementado, comparado con la señal suministrada por el vocodificador de fase (ilustrada en la representación de señal 1040).

45 La figura 11 ilustra otra representación de señal que representa las señales que aparecen en otro modo de realización del aparato 100 de la figura 1. El procesamiento es similar al procesamiento explicado con referencia a la figura 10, de modo que se describirán aquí sólo las diferencias en el orden del procesamiento, y de modo que representaciones de señal y características de señal idénticas serán designadas con números de referencia idénticos en las figuras 10 y 11.

50 En el procesamiento de señal representado en la representación de señal 1100, el submuestreo se efectúa antes de la reinsertión de señal transitoria. Así, una representación de señal 1150 ilustra la señal submuestreada sin porción de señal transitoria insertada. Sin embargo, la porción de señal transitoria es deslizada en frecuencia mediante una operación de deslizamiento en frecuencia del transitorio 1160, lo que puede ser realizado por el procesador de transitorio 160. La señal transitoria deslizada en frecuencia (deslizada en frecuencia respecto a la porción de señal transitoria sustituida por el sustituidor de señal transitoria 130) puede ser reinsertada en la señal de audio procesada submuestreada 142 por el reinsertador de señal transitoria 150. El resultado de la reinsertión de transitorio es ilustrado en una representación de señal
55 1170.

Ajuste de la porción de señal transitoria

60 A continuación se describe cómo la señal transitoria 152 puede ser combinada con la señal de audio procesada 142 mediante el insertador de señal transitoria 150. Por ejemplo, el insertador de señal transitoria 150 puede ser configurado para cortar una zona transitoria de la señal de audio procesada 142, zona transitoria en la cual debe insertarse la señal

transitoria 152. Se puede considerar aquí que las porciones límites de la señal transitoria 152 pueden solapar en el tiempo con las porciones límites del área transitoria cortada. En esta porción límite en solapamiento puede tener lugar un fundido encadenado entre la señal de audio procesada 142 y la señal transitoria 152. La señal transitoria 152 puede también ser deslizada en el tiempo respecto a la señal de audio procesada 142, de modo que la forma de onda de las porciones límites de la zona transitoria recubierta sea llevada en buena concordancia con la forma de onda de las porciones límites de la señal transitoria 152.

Se puede realizar un ajuste preciso calculando el máximo de la correlación cruzada de los bordes del receso resultante con los bordes de la porción transitoria (donde el receso puede ser causado por el corte de la zona transitoria de la señal de audio procesada 142). De este modo, la calidad audio subjetiva del transitorio ya no resulta dañada por efectos de dispersión y de eco.

Se puede realizar una determinación precisa de la posición del transitorio con fines de seleccionar un corte adecuado, por ejemplo mediante un cálculo de centro de gravedad flotante de la energía durante un período de tiempo apropiado.

Un ajuste óptimo del transitorio de conformidad con la correlación cruzada máxima puede requerir un ligero desplazamiento en el tiempo respecto a la posición original del mismo. Sin embargo, debido a la existencia de un preenmascarado en el tiempo y, en particular, a efectos de post-enmascarado, la posición del transitorio reinsertado no debe coincidir exactamente con la posición original. Debido al período de acción más largo del post-enmascarado, se debe favorecer un desplazamiento del transitorio en la dirección de tiempo positiva en este contexto. Al insertar la porción de señal original, un cambio de la tasa de muestreo conduce a un cambio del timbre, o del tono. Sin embargo, éste es generalmente enmascarado por el transitorio mediante mecanismos de enmascarado psíquico-acústico.

Procesamiento del Transitorio

Si el transitorio debe ser menos tonal antes de la reinsertión que después del corte, por ejemplo, porque debe sencillamente ser añadido a la señal procesada, la correspondiente porción transitoria dividida en ventanas deberá ser procesada de modo adecuado. En este contexto se puede realizar un filtrado inverso (LPC).

A continuación se describe brevemente un enfoque alternativo:

1. Determinar la Transformación de Fourier de corta duración (STFT) (por ejemplo de la porción de señal transitoria descrito por la información sobre transitorios 134), para obtener un espectro;
2. Determinar el Cepstro (por ejemplo del espectro de la porción de señal transitoria);
3. Filtrar paso alto el cepstro (los primeros coeficientes son ajustados a 0), para obtener un filtrado paso alto del espectro;
4. Dividir el espectro (por ejemplo de la porción de señal transitoria) por el espectro filtrado (por ejemplo de la porción de señal transitoria), para obtener un espectro alisado; y
5. Transformación inversa (por ejemplo del espectro alisado) al dominio temporal (por ejemplo para obtener la señal procesada transitoria 152).

La señal resultante presenta (al menos aproximadamente) la misma envolvente espectral que la señal de salida, pero ha perdido porciones tonales.

Proceso

Un modo de realización según la invención comprende un proceso para la manipulación de una señal de audio que comprende un evento transitorio. La figura 12 ilustra un organigrama de tal proceso 1200.

El proceso 1200 comprende una etapa 1210 que consiste en sustituir una porción de señal transitoria que comprende el evento transitorio de la señal de audio por una porción de señal de sustitución adaptada a las características de energía de señal de una o varias de las porciones de señal no transitorias de la señal de audio o a una característica de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos.

El proceso 1200 comprende además una etapa 1220 que consiste en procesar la señal de audio de transitorios reducidos, para obtener una versión procesada de la señal de audio de transitorios reducidos.

El proceso 1200 comprende además una etapa 1230 que consiste en combinar la versión procesada de la señal de audio de transitorios reducidos con una señal transitoria que representa, de forma original o procesada, un contenido transitorio de la porción de señal transitoria.

5 El proceso 1200 puede ser completado con cualquiera de las características o funcionalidades descritas aquí en relación también con el arriba mencionado aparato de la invención.

En otras palabras, si bien algunos aspectos hayan sido descritos en el contexto de un aparato, está claro que estos aspectos representan también una descripción del proceso correspondiente, donde un bloque o dispositivo corresponde a una etapa de proceso o una característica de una etapa de proceso. De modo similar, los aspectos descritos en el contexto de una etapa de proceso representan también una descripción de un bloque o ítem o una característica correspondiente de un aparato correspondiente.

15 Programa Informático

En función de determinados requisitos de implementación, los modos de realización de la invención pueden implementarse en hardware o en software. La implementación puede realizarse mediante un soporte de almacenamiento digital, por ejemplo un disquete, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que presenta señales de control legibles electrónicamente memorizadas en el mismo y que cooperan (o capaces de cooperar) con un sistema informático programable de modo que se realice el proceso respectivo. Por lo tanto, el soporte de almacenamiento digital puede ser legible por ordenador.

Algunos modos de realización según la invención comprenden un soporte de datos que presenta señales de control legibles electrónicamente capaces de cooperar con un sistema informático programable de modo que se realice uno de los procesos descritos aquí.

Generalmente, los modos de realización de la presente invención pueden implementarse como producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los procesos cuando se ejecuta el producto de programa informático en un ordenador. El código de programa puede por ejemplo ser memorizado en un soporte legible en máquina.

Otros modos de realización comprenden el programa informático para realizar uno de los procesos descritos aquí y memorizado en un soporte legible en máquina.

En otras palabras, un modo de realización del proceso de la invención es por lo tanto un programa informático que presenta un código de programa para realizar uno de los procesos descritos aquí cuando se ejecuta el programa informático en un ordenador.

Otro modo de realización de los procesos de la invención es por lo tanto un soporte de datos (o un soporte de almacenamiento digital, o un soporte legible en máquina) que comprende, memorizado en el mismo, el programa informático para realizar uno de los procesos descritos aquí.

Otro modo de realización del proceso de la invención es por lo tanto un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los procesos descritos aquí. El flujo de datos o la secuencia de señales pueden, por ejemplo, ser configurados para ser transferidos vía una conexión de comunicación de datos, por ejemplo vía Internet.

Otro modo de realización comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procesos descritos aquí.

Otro modo de realización comprende un ordenador que presenta, instalado en el mismo, el programa informático para realizar uno de los procesos descritos aquí.

En algunos modos de realización, un dispositivo lógico programable (por ejemplo, una matriz de puertas programable in-situ) puede utilizarse para realizar algunas o todas las funcionalidades de los procesos descritos aquí. En algunos modos de realización, una matriz de puertas programable in-situ puede cooperar con un microprocesador, para realizar uno de los procesos descritos aquí. Generalmente, los procesos se realizan preferiblemente mediante cualquier aparato de hardware.

60 Conclusión

Para resumir lo que precede, los modos de realización según la presente invención comprenden un nuevo proceso para tratar los eventos sonoros que no deben o no pueden ser procesados mediante la rutina de procesamiento actual (por ejemplo mediante el procesador de señal). En algunos modos de realización, el proceso de la invención consiste esencialmente en extrapolar o interpolar la porción de señal que contiene los eventos sonoros que deben procesarse por separado. Tras el procesamiento, las porciones transitorias tratadas por separada son de nuevo añadidas. Este procesamiento no se limita al alargamiento en el tiempo o en frecuencia, pero puede utilizarse generalmente en el procesamiento de señal cuando el procesamiento actual de la señal es perjudicial para la porción de señal transitoria (o es afectado negativamente por las porciones de señal transitorias).

A continuación se describen algunas ventajas del nuevo proceso que pueden ser obtenidas en algunos de los modos de realización. Con el nuevo proceso son presentados efectivamente los artefactos (tales como dispersión, pre-eco, y ecos retrasados) que pueden producirse durante el procesamiento del transitorio mediante métodos de alargamiento en el tiempo y transposición. Una potencial degradación de la calidad de porciones de señal (posiblemente tonales) es evitada.

Los modos de realización según la invención pueden aplicarse en distintos campos de aplicación. El método conviene, por ejemplo, para cualquier aplicación audio en la cual las velocidades de reproducción de señales audio, o sus tonos, deben ser cambiadas.

Para resumir lo que precede se han descrito un medio y un proceso para un procesamiento por separado de eventos sonoros en señales audio para evitar artefactos.

Modo de realización 2

A continuación se describe otro modo de realización de la invención, con referencia a las figuras 13-16.

En primer lugar se describirán detalles relacionados con una detección de transitorio. A continuación se explicará la manipulación de transitorios, con referencia a las figura 13 y 14. Los resultados de la manipulación de transitorios serán descritos con referencia a la figura 15. Mejoras adicionales de la manipulación de transitorios serán explicadas con referencia a la figura 16. Además, se dará una evaluación de rendimiento del modo de realización, y se presentarán algunas conclusiones.

Modo de realización 2 – Detección de Transitorios

Para la implementación del concepto de la invención, es importante detectar la presencia de transitorios, para permitir una sustitución de transitorios y para una manipulación por separado de los transitorios.

Además de la aplicación de alargamiento en el tiempo a mano, una amplia gama de procesos de procesamiento de señal requieren un conocimiento del contenido transitorio de una señal de audio. Ejemplos prominentes son las decisiones de longitud de bloque (B. Edler, "Coding of audio signals with overlapping block transform and adaptive window functions (en alemán)," *Frequenz*, volumen 43, nº 9, páginas 252-256, septiembre de 1989), o la codificación por separado de señales transitorias y estacionarias (Oliver Niemeyer and Bernd Edler, "Detection and extraction of transients for audio coding," en *AES 120th Convention*, Paris, Francia, 2006) en codecs de audio de transformación, la modificación de componentes transitorios (M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society*, volumen 54, páginas 827-840, 2006.) y la segmentación de señal de audio (P. Brossier, J.P. Bello, and M.D. Plumbley, "Real-time temporal segmentation of note objects in music signals," en *ICMC*, Miami, USA, 2004). Los enfoques para la detección de transitorios son igual de numerosos que sus aplicaciones. Lo más habitual es que la detección se realice calculando una función de detección (J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, volumen 13, nº 5, páginas 1035-1047, septiembre de 2005), es decir, una función donde los máximos locales coinciden con la ocurrencia de transitorios. Varios procesos propuestos derivan tal función de detección investigando la magnitud (ponderada) o envolvente de energía de señales de sub-banda, la señal de banda ancha, su función de derivada o de diferencia relativa (véanse, por ejemplo, las referencias (A. Klapuri, "Sound onset detection by applying psychoacoustic knowledge," en *ICASSP*, 1999) y (P. Masri and A. Bateman, "Improved modelling of attack transients in music analysis-resynthesis," en *ICMC*, 1996).

Otros procesos calculan la desviación entre la fase medida y una fase prevista (véase, por ejemplo, C. Duxbury, M. Davies, and M. Sandler, "Separation of transient information in musical audio using multiresolution analysis techniques," en *DAFX*, 2001), un examen combinado tanto de la fase como de las magnitudes de señales de sub-banda (véase, por ejemplo, C. Duxbury, M. Sandler, and M. Davies, "A hybrid approach to musical note onset detection," en *DAFX*, 2002), o el error cometido por un predictor lineal adaptativo (véase, por ejemplo, W-C. Lee and C-C. J. Kuo, "Musical onset detection based on adaptive linear prediction," en *ICME*, 2006). Mediante selección de picos, la presencia de un transitorio y su localización

en el tiempo son derivados o bien como decisión binaria, o bien se aplica la función de detección continua para controlar el comportamiento de la unidad de modificación (véase, por ejemplo, Ref. M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society*, volumen 54, páginas 827-840, 2006).

Con una decisión binaria, atribuciones erróneas debidas a malas clasificaciones en la fase de detección pueden causar degradaciones severas en algunas aplicaciones. Para el presente algoritmo, un falso negativo (es decir, falta de detección de un transitorio) sería peor que un falso positivo (es decir, detección de un transitorio no existente). Lo primero conduciría a una componente transitoria maculada, mientras que lo último sólo produce una interpolación superflua si la interpolación se realiza correctamente.

Los valores absolutos ponderados resumidos de bloques de transformación de Fourier de corta duración se utilizan para la detección de zonas transitorias. Esta función muestra las subidas marcadas durante transitorios de ataque y también es capaz de indicar el decaimiento de señales de percusión y la reverberación asociada. La selección de picos en la función de detección alisada se realizó mediante un umbral adaptativo en base a un cálculo percentil tal y como descrito, por ejemplo, en la referencia J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, volumen 13, nº 5, páginas 1035-1047, septiembre de 2005.

Para resumir lo que precede, distintos conceptos de detección de un transitorio son conocidos en el arte y pueden aplicarse en un aparato de la invención. Por ejemplo, el concepto descrito más arriba para la detección de un transitorio puede utilizarse en el detector de transitorios 130a del sustituidor de señal transitoria 130.

Modo de realización 2 – Manipulación de transitorio

A continuación se describe la manipulación de un transitorio, con referencia a las figuras 13 y 14. La figura 13 ilustra una representación gráfica de una eliminación e interpolación de transitorio. La figura 14 ilustra una representación gráfica de un alargamiento en el tiempo y reinserción de transitorio. Así, las representaciones esquemáticas en las figures 13 y 14 ilustran la secuencia de etapas de procesamiento del algoritmo presentado.

Una primera fila 1310 de la figura 13 ilustra la señal original (es decir, la señal de audio 110) que contiene un evento transitorio 1312. En respuesta a (o mediante) la detección de este transitorio 1312 se define (por ejemplo, por el detector de transitorios 130a) una zona transitoria (por ejemplo, que se extiende de una posición de inicio de zona transitoria 1314 a una posición de fin de zona transitoria 1316), la cual es a continuación sustraída de la señal. En otras palabras, en primer lugar, el transitorio es detectado y dividido en ventanas. En segundo lugar, es sustraído de la señal. Una señal en la cual el transitorio es sustraído es ilustrada en la Referencia [B20]. El transitorio es a su vez almacenado para utilización ulterior. Hasta esta etapa, el algoritmo es idéntico a aquel descrito en la Referencia [B8] no obstante el hecho de que la ventana cortada utilizada aquí sea rectangular (trazo gordo interrumpido). Para el almacenamiento del transitorio, un intervalo de guarda de unos milisegundos precede y sigue y la ventana es cónica (trazo continuo fino), para definir las zonas de fundido encadenado para una reinserción suave del transitorio almacenado en las señales libres de transitorio suprimido en el tiempo.

A continuación se aplica la característica más importante del algoritmo de la invención según el presente modo de realización – la interpolación para rellenar el lapso libre. En otras palabras, finalmente, el lapso libre resultante se rellena mediante interpolación. Se puede ver un resultado de la interpolación en una fila inferior de la figura 13, en el número de referencia 1330. Dado que la señal es, típicamente, casi estacionaria tras la interpolación, ésta puede ahora ser alargada sin introducir artefactos molestos. Un resultado de este alargamiento es ilustrado en una primera fila de la figura 14, en el número de referencia 1410. La región transitoria en la posición transpuesta es identificada y preparada para reinserción del transitorio dividido en ventanas anteriormente almacenado. Por lo tanto, la ventana cónica (que se aplicó para extracción y/o almacenamiento del transitorio, y que es ilustrada mediante un trazo continuo fino en la representación gráfica en el número de referencia 1310) es invertida y aplicada a la señal para permitir añadir de nuevo el transitorio. Un resultado de este proceso es ilustrado en el número de referencia 1420. Finalmente, el transitorio almacenado es añadido a la señal alargada, tal y como visible en la representación gráfica en el número de referencia 1430.

Para resumir lo que precede, la eliminación de transitorio y la interpolación del lapso libre, que es causado por la eliminación de transitorio, son ilustradas en la figura 13. En primer lugar, el transitorio es detectado y dividido en ventanas. En segundo lugar, es sustraído de la señal. Finalmente, el lapso libre resultante es relleno mediante la interpolación. La figura 14 ilustra el alargamiento en el tiempo y la reinserción de transitorio que sigue a la eliminación de transitorio e interpolación. En primer lugar, la señal casi estacionaria es alargada, por ejemplo, mediante el vocodificador descrito aquí. A continuación, la posición para el transitorio en la señal alargada en el tiempo es preparada por multiplicación por la ventana inversa de la que se utilizó para el almacenamiento del transitorio en la figura 14. Finalmente, el transitorio es reañadido a la señal. En otras

palabras, finalmente, el transitorio almacenado es añadido a la señal alargada.

Modo de realización 2 – Resultados de la manipulación de transitorios

A continuación se describen algunos resultados de la manipulación de transitorios de la invención, con referencia a la figura 15. La figura 15 ilustra una representación gráfica de etapas de la manipulación de transitorios de la invención en una aplicación de alargamiento en el tiempo con el vocodificador de fase. Una primera fila contiene la señal sin alargamiento, y una segunda fila contiene partes alargadas. Se notarán los distintos intervalos de tiempo utilizados en las representaciones gráficas de la primera fila y en la segunda fila.

La figura 15 muestra los resultados de las distintas etapas algorítmicas en base a castañetas mezcladas con un diapason.

Un trazado de la forma de onda de la señal de entrada original con una indicación de las zonas transitorias detectadas es ilustrado en la figura 15a. La figura 15b ilustra las zonas transitorias cortadas que son interpoladas (en una etapa siguiente) para producir la señal estacionaria libre de transitorio visualizada en la figura 15c. La figura 15d contiene las zonas transitorias inclusive los intervalos de guardia de fundido encadenado, mientras que la figura 15e ilustra la señal interpolada (y, típicamente, alargada en el tiempo) que es atenuada por la ventana de fundido encadenado inversa en las posiciones de transitorio suprimidas en el tiempo. Para ser completo, la figura 15f visualiza el resultado final del algoritmo de alargamiento en el tiempo.

Así, la figura 15a representa la señal de audio 110. La figura 15e representa la señal de audio de transitorios reducidos 132. La figura 15d representa la señal transitoria 152. La figura 15f representa la señal de audio procesada 120.

Modo de realización 2 – Mejoras de la manipulación de transitorios

Se ha constatado que distintos conceptos relacionados con la interpolación de las zonas transitorias cortadas pueden, en algunos casos, ser importantes. Por ejemplo, la interpolación sobre una zona transitoria puede resultar difícil si la señal antes del transitorio difiere considerablemente de la señal después del transitorio. En tal caso, la implicación de la señal durante el evento transitorio es, en algunos casos, apenas previsible. La figura 16 ilustra tal situación, por ejemplo simplificada mediante la posible evaluación de sólo uno o dos parciales. El algoritmo (por ejemplo, el algoritmo para la realización de la interpolación para rellenar el lapso libre) debe decidir una implicación del tono (de la señal interpolada para rellenar el lapso libre). Lo mismo vale para señales de banda ancha más complejas. Una posible solución para resolver el problema consiste en una predicción hacia adelante y hacia atrás con fundido encadenado entre ellas. Así, tal predicción hacia adelante y hacia atrás con fundido encadenado entre ellas puede aplicarse cuando se calcula la señal interpolada para rellenar el lapso libre.

Este problema es ilustrado en la figura 16 y una solución según un aspecto de la invención es presentada. La figura 16 ilustra que la interpolación del transitorio (es decir, la interpolación del lapso libre causado por una eliminación del transitorio) resulta difícil si la señal cambia de modo notable durante el transitorio. Existen infinitas maneras de contornar el tono durante la gama de interpolación (es decir, el lapso libre causado por la eliminación del transitorio). La figura 16a ilustra una representación gráfica de una señal que contiene un evento transitorio en forma de una representación en tiempo-frecuencia. Una gama de transitorios, es decir, un intervalo de tiempo que fue identificado como intervalo de tiempo transitorio, es designado con 1610. La figura 16b ilustra una representación gráfica de distintas posibilidades para obtener una porción temporal de la señal de entrada audio durante la cual se ha detectado y eliminado un transitorio. Tal y como se puede ver, si existe un primer tono que precede en el tiempo el intervalo de tiempo 1620 durante el cual el transitorio es eliminado de la señal de entrada audio, y un segundo tono en el tiempo después del intervalo de tiempo 1620, es necesario determinar una evolución del tono para rellenar el lapso libre dejado por la eliminación del intervalo de tiempo transitorio 1620. Tal y como se puede ver, es posible, por ejemplo, extrapolar hacia adelante (en la dirección del tiempo) el tono que precede el intervalo de tiempo 1620, para obtener el tono durante el intervalo de tiempo 1620 (véase el trazo interrumpido 1630). Alternativamente, es posible extrapolar hacia atrás (en la dirección del tiempo) un tono que es presente después del intervalo de tiempo 1620, al intervalo de tiempo 1620 (véase el trazo interrumpido 1632). Alternativamente, es posible interpolar, durante el intervalo de tiempo 1620, entre un tono que es presente antes del intervalo de tiempo 1620 y un tono que es presente después del intervalo de tiempo 1620 (véase el trazo interrumpido 1634). Naturalmente, distintos esquemas para obtener una evolución del tono durante el intervalo de tiempo 1620 (lapso libre causado por la eliminación de transitorio) son posibles.

Un impacto de la señal de audio procesada finalmente obtenida, tras la reinserción de señal transitoria, es ilustrado en la figura 16c. Tal y como se puede ver, la porción de señal transitoria reinsertada (que refleja un contenido transitorio original o procesado de la porción de señal transitoria) puede ser más corto en el tiempo que la señal de audio procesada (por ejemplo alargada en el tiempo) 142 que fue procesada sin el contenido transitorio. Así, la elección del concepto para rellenar

5 el lapso libre causado por la eliminación de transitorio en la señal de audio 132 puede, en realidad, tener un impacto audible sobre la señal de audio procesada 120, incluso tras inserción del transitorio, por ejemplo si la porción transitoria reinsertada (descrita por la señal transitoria 152) es más corta que el resultado procesado del relleno de lapso libre en la señal de audio procesada 142. Refiérase al intervalo de tiempo 140 que precede el transitorio reinsertado y un intervalo de tiempo 142 que sigue el transitorio reinsertado.

10 Para resumir lo que precede, se ha mostrado, con referencia a la figura 16, que la interpolación de la zona transitoria requiere alguna consideración si la señal cambia de forma notable durante el transitorio. Existen infinitas maneras de contornar el tono durante la gama de interpolación. La figura 16a ilustra una señal que contiene un evento transitorio. La figura 16b ilustra distintas posibilidades para interpolaciones de la gama de transitorios que son indicados mediante trazos interrumpidos. La figura 16c ilustra una señal alargada. A medida que las regiones interpoladas alargadas se extienden más allá de las partes transitorias, la señal interpolada es audible y puede conducir a artefactos de percepción.

15 Modo de realización 2 – Evaluación del rendimiento

Para tener alguna idea del rendimiento de percepción del proceso propuesto, se ha realizado una escucha informal. Las señales seleccionadas comportaban ítems con características de señal tanto transitoria como estacionaria, para evaluar el beneficio del nuevo esquema de señales transitorias, y, al mismo tiempo, asegurar que las señales estacionarias no son degradadas.

20 Este ensayo informal reveló un beneficio importante para la arriba mencionada combinación de diapasón y castañetas en comparación con el algoritmo de alargamiento en el tiempo del software del estado del arte. El resultado mostró una preferencia para los algoritmos de alargamiento en el tiempo en base a PV sobre WSOLA cuando el enfoque es dirigido a señales transitorias.

25 Las señales del mundo real alargadas con el nuevo proceso eran a veces también preferidas sobre otros procesos.

Conclusión

30 Para resumir lo que precede, se ha descrito un nuevo esquema de manipulación de transitorios que puede utilizarse ventajosamente para algoritmos de alargamiento en el tiempo. El cambio de la velocidad o del tono de señales audio sin afectar el otro respectivo es a menudo utilizado para la producción musical y la reproducción creativa, tal como remix. También se utiliza a otros fines tales como extensión de la anchura de banda y mejora de la velocidad. Mientras que las señales estacionarias pueden ser alargadas sin dañar la calidad, los transitorios a menudo no son bien mantenidos tras el alargamiento cuando se utilizan los algoritmos convencionales. La presente invención muestra un enfoque para la manipulación de transitorios en algoritmos de alargamiento en el tiempo. Las regiones transitorias son sustituidas por señales estacionarias. Los transitorios así eliminados son guardados y reinsertados en la señal de audio estacionaria dilatada en el tiempo tras el alargamiento en el tiempo.

40 La tarea de combinar una señal muy tonal tal como un diapasón y una señal de percusión tal como castañetas constituye un desafío.

45 Mientras que algunos procesos convencionales preservan aproximadamente la envolvente de una señal en la versión alargada en el tiempo así como sus características espectrales, y prevén que un evento de percusión dilatado en el tiempo decaiga más lentamente que el original, la presente invención sigue a la suposición opuesta que, para una modulación en el tiempo de señales musicales, el objetivo es él de preservar la envolvente de eventos transitorios. Por lo tanto, algunos modos de realización según la invención alargan sólo la componente sostenida para obtener un efecto que suene como el mismo instrumento tocado a un temperamento distinto (véase, por ejemplo, la Referencia [B3]). A tal fin, las componentes de señales transitorias y estacionarias son tratadas por separado según la invención.

50 Los modos de realización según la invención se basen en un concepto que ha sido descrito en la publicación [B8], en la cual se ha mostrado cómo los transitorios pueden preservarse en el alargamiento en el tiempo y en frecuencia con el vocodificador de fase. En este enfoque, los transitorios son cortados de la señal antes de su alargamiento. La eliminación de la parte transitoria resulta en lapsos libres dentro de la señal que son alargados por el proceso de vocodificador de fase. Tras el alargamiento, los transitorios son reañadidos a la señal con una envoltura que se adapta a los lapsos libres alargados. Sin embargo, se ha constatado que la solución comprende algunas ventajas para muchas señales. Sin embargo, también se ha constatado que al cortar los transitorios surgen nuevos artefactos, dado que los lapsos libres introducen nuevas partes no estacionarias en la señal, en particular en los límites de los lapsos libres introducidos. Tales no estacionalidades pueden verse, por ejemplo, en la figura 15b.

60 Los modos de realización del proceso de la invención descritos aquí presentan la ventaja, respecto a las técnicas descritas,

5 por ejemplo, en las publicaciones [B3], [B6], [B7], que permiten un alargamiento en el tiempo sin necesidad de cambiar el factor de alargamiento en el entorno de un transitorio. El proceso de la invención tiene elementos en común con los procesos descritos, por ejemplo, en las referencias [B8] y [B5]. El esquema de la invención divide la señal en una parte transitoria y una señal casi estacionaria sin transitorio. Contrariamente al proceso descrito en [B8], los lapsos libres que se producen debido al corte de los transitorios son sustituidos por señales estacionarias. Un proceso de interpolación es utilizado para estimar una continuación de las señales que rodean el período de lapso libre en todo el lapso libre. La parte casi estacionaria resultante conviene muy bien para los algoritmos de alargamiento en el tiempo. Debido a que esta señal ya no comporta entonces (es decir, tras la interpolación o extrapolación) ni transitorios ni lapsos libres, los artefactos tanto de transitorios alargados como de lapsos libres pueden evitarse. Tras ejecución del alargamiento, los transitorios sustituyen partes de la señal interpolada. La técnica se basa tanto en la detección correcta de transitorios como en una interpolación correcta desde el punto de vista de la percepción de la parte estacionaria. Sin embargo, aparte la interpolación, otras técnicas de rellenado pueden utilizarse tal y como descrito más arriba.

15 Para resumir mejor lo que precede, en algunos modos de realización descritos más arriba, el objetivo era el de alargar una combinación de una señal estrictamente tonal y una señal transitoria, tal como diapason plus castañetas, sin artefactos de percepción. Se ha mostrado que la presente invención ofrece un progreso importante hacia este objetivo. Uno de los aspectos importantes de la presente invención reside en la correcta identificación de un evento transitorio, en particular de su comienzo exacto, y, aún más difícil, su decaimiento y su reverberación asociada. Dado que el decaimiento y una reverberación de un evento transitorio son cubiertos por las partes estacionarias de la señal, estas porciones requieren una manipulación meticulosa para evitar fluctuaciones de percepción tras su readición a las partes alargadas de la señal.

25 Algunos oyentes tienden a preferir versiones en las cuales la reverberación es alargada junto con las partes de señal sostenidas. Esta preferencia es en contradicción con el objetivo actual de considerar un transitorio y los sonidos asociados como una entidad. Por lo tanto, en algunos casos es necesaria una mayor revelación de la preferencia de los oyentes.

30 Sin embargo, la idea y el enfoque de principio, según la presente invención, han demostrado su valor y aplicación para un caso especial. No obstante, se prevé que la gama de aplicaciones de la presente invención puede incluso extenderse. Debido a su estructura, el algoritmo de la invención puede fácilmente ser adaptado para su uso para una manipulación de la parte transitoria, por ejemplo cambiando su nivel comparado con las partes de señal estacionarias.

35 Otra aplicación posible del proceso de la invención consistiría en atenuar de modo arbitrario o ganar transitorios para la reproducción. Esto podría ser explotado para cambiar la sonoridad de eventos transitorios tales como tambores o incluso para eliminarlos totalmente, dado que una separación de la señal en parte transitorio y parte estacionaria es inherente al algoritmo.

40 Los modos de realización descritos más arriba son meramente ilustrativos de los principios de la presente invención. Está claro que modificaciones y variaciones de las disposiciones y de los detalles descritos aquí serán obvias para el hombre del arte. Por lo tanto, la intención es que la invención sea limitada sólo por el marco de las reivindicaciones independientes y no por los detalles específicos presentados a título de descripción y explicación de los modos de realización dados aquí.

Referencias

- [A1] J.L. Flanagan and R.M. Golden, "The Bell System Technical Journal, noviembre de 1966", páginas 1394 a 1509;
- 5 [A2] Patente de los Estados Unidos 6.549.884, Laroche, J. & Dolson, M.: "Phase-vocoder pitch-shifting";
- [A3] Jean Laroche and Mark Dolson, "New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing and Other Exotic Effects", by Proc.
- 10 [A4] Zölzer, U: "DAFX: Digital Audio Effects", Wiley & Sons, Edition: 1 (26 de febrero de 2002), páginas 201-298;
- [A5] Laroche L., Dolson M.: "Improved phase-vocoder timescale modification of audio", IEEE Trans. Speech and Audio Processing, volumen 7, nº 3, páginas 323-332;
- 15 [A6] Emmanuel Ravelli, Mark Sandler and Juan P. Bello: "Fast implementation for non-linear time-scaling of stereo audio", Proc. of the 8th Int. Conference on Digital Audio Effects (DAFx'05), Madrid, España, 20-22 de septiembre de 2005;
- [A7] Duxbury, C., M. Davies, and M. Sandler (2001, diciembre): "Separation of transient information in musical audio using multiresolution analysis techniques". En: Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), Limerick, Irlanda;
- 20 [A8] Röbel A.: "A NEW APPROACH TO TRANSIENT PROCESSING IN THE PHASE-VOCODER", Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), Londres, RU, 8-11 de septiembre de 2003.
- 25 [B1] T. Karrer, E. Lee, and J. Borchers, "Phavorit: A phase vocoder for real-time interactive time-stretching," in *Proceedings of the ICMC 2006 International Computer Music Conference*, New Orleans, USA, noviembre de 2006, páginas 708-715.
- [B2] T. F. Quatieri, R. B. Dunn, R. J. McAulay, and T. E. Hanna, "Time-scale modifications of complex acoustic signals in noise," Technical report, Massachusetts Institute of Technology, febrero de 1994.
- 30 [B3] C. Duxbury, M. Davies, and M. B. Sandler, "Improved time-scaling of musical audio using phase locking at transients," in *112th AES Convention*, Munich, 2002, Audio Engineering Society.
- [B4] S. Levine and Julius O. Smith III, "A sines+transients+noise audio representation for data compression and time/pitch scale modifications," 1998.
- 35 [B5] T. S. Verma and T. H. Y. Meng, "Time scale modification using a sines+transients+noise signal model," in *DAFX98*, Barcelona, España, 1998.
- 40 [B6] A. Röbel, "A new approach to transient processing in the phase vocoder," in *6th Conference on Digital Audio Effects (DAFx-03)*, Londres, 2003, páginas 344-349.
- [B7] A. Röbel, "Transient detection and preservation in the phase vocoder," in *Int. Computer Music Conference (ICMC 03)*, Singapur, 2003, páginas 247-250.
- 45 [B8] F. Nagel, S. Disch, and N. Rettelbach, "A phase vocoder driven bandwidth extension method with novel transient handling for audio codecs," in *126th AES Convention*, Munich, 2009.
- [B9] M. Dolson, "The phase vocoder: A tutorial," *Computer Music Journal*, volumen 10, nº 4, páginas 14-27, 1986.
- 50 [B10] B. Edler, "Coding of audio signals with overlapping block transform and adaptive window functions (en alemán)," *Frequenz*, volumen 43, nº 9, páginas 252-256, septiembre de 1989.
- [B11] Oliver Niemeyer and Bernd Edler, "Detection and extraction of transients for audio coding," in *AES 120th Convention*, Paris, Francia, 2006.
- 55 [B12] M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society*., volumen 54, páginas 827-840, 2006.
- 60 [B13] P. Brossier, J.P. Bello, and M.D. Plumbley, "Real-time temporal segmentation of note objects in music signals," in *ICMC*, Miami, USA, 2004.

- [B14] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, volumen 13, nº 5, páginas 1035-1047, septiembre 2005.
- 5 [B15] A. Klapuri, "Sound onset detection by applying psychoacoustic knowledge," in ICASSP, 1999.
- [B16] P. Masri and A. Bateman, "Improved modelling of attack transients in music analysis-resynthesis," in ICMC, 1996.
- [B17] C. Duxbury, M. Davies, and M. Sandler, "Separation of transient information in musical audio using multiresolution analysis techniques," in *DAFX*, 2001.
- 10 [B18] C. Duxbury, M. Sandler, and M. Davies, "A hybrid approach to musical note onset detection," " in *DAFX*, 2002.
- [B19] W-C. Lee and C-C. J. Kuo, "Musical onset detection based on adaptive linear prediction," in *ICME*, 2006.
- 15 [Edler] O. Niemeyer and B. Edler, "Detection and extraction of transients for audio coding", presentado en la AES 120th Convention, Paris, Francia, 2006;
- [Bello] J.P. Bello et al., "A Tutorial on Onset Detection in Music Signals", *IEEE Transactions on Speech and Audio Processing*, Volumen 13, Nº 5, septiembre de 2005;
- 20 [Goodwin] M. Goodwin, C. Avendano, "Enhancement of Audio Signals Using Transient Detection and Modification", presentado en la AES 117th Convention, USA, October 2004;
- [Walther] Walther et al., "Using Transient Suppression in Blind Multi-channel Upmix Algorithms", presentado en la AES 122th Convention, Austria, mayo de 2007;
- 25 [Maher] R.C. Maher, "A Method for Extrapolation of Missing Digital Audio Data", *JAES*, Volumen 42, nº 5, mayo de 1994;
- [Daudet] L. Daudet, "A review on techniques for the extraction of transients in musical signals", book series: *Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, Volume 3902/2006, Book: *Computer Music Modeling and Retrieval*, páginas 219-232.
- 30

REIVINDICACIONES

1. Aparato (100) para la manipulación de una señal de audio (110) que comprende un evento transitorio, comprendiendo el aparato (100):
- 5 un sustituidor de señal transitoria (130) configurado para sustituir una porción de señal transitoria, que comprende el evento transitorio, de la señal de audio por una porción de señal de sustitución adaptada a las características de energía de señal de una o varias porciones de señal no transitorias de la señal de audio, o a una característica de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos (132);
- 10 un procesador de señal (140) configurado para procesar la señal de audio de transitorios reducidos (132), para obtener una versión procesada (142) de la señal de audio de transitorios reducidos; y
- 15 un reinsertador de señal transitoria (150) configurado para combinar la versión procesada (142) de la señal de audio de transitorios reducidos (132) con una señal transitoria (152) que representa, de forma original o procesada, un contenido transitorio de la porción de señal transitoria;
- 20 donde el sustituidor de señal transitoria (130) está configurado para extrapolar valores de amplitud de una o más porciones de señal que preceden a la porción de señal transitoria, para obtener valores de amplitud de la porción de señal de reemplazo, y
- 25 donde el sustituidor de señal transitoria (130) está configurado para extrapolar valores de fase de una o más porciones de señal que preceden a la porción de señal transitoria para obtener valores de fase de la porción de señal de reemplazo.
2. Aparato (100) para manipular una señal de audio (110), que comprende un evento transitorio, el aparato (100) comprende:
- 30 un sustituidor de señal transitoria (130) configurado para sustituir una porción de señal transitoria, que comprende el evento transitorio, de la señal de audio con una porción de señal de reemplazo adaptada a características de energía de señal de una o más porciones de señal no transitoria de la señal de audio, o a una característica de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos (132);
- 35 un procesador de señal (140) configurado para procesar la señal de audio de transitorios reducidos (132), para obtener una versión procesada (142) de la señal de audio de transitorios reducidos; y
- 40 un reinsertador de señal transitoria (150) que está configurado para combinar la versión procesada (142) de la señal de audio de transitorios reducidos (132) con una señal transitoria (152) representando, en una forma original o procesada, un contenido transitorio de la porción de señal transitoria;
- 45 donde el sustituidor de señal transitoria (130) está configurado para interpolar entre un valor de amplitud de una porción de señal que precede a la porción de señal transitoria y un valor de amplitud de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o más valores de amplitud de la porción de señal de reemplazo, y
- 50 donde el sustituidor de señal transitoria (130) está configurado para interpolar entre un valor de fase de una porción de señal que precede a la porción de señal transitoria y un valor de fase de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o más valores de fase de la porción de señal de reemplazo.
3. Aparato (100) para manipular una señal de audio (110), que comprende un evento transitorio, el aparato (100) comprende:
- 55 un sustituidor de señal transitoria (130) configurado para sustituir una porción de señal transitoria, que comprende el evento transitorio, de la señal de audio con una porción de señal de reemplazo adaptada a características de energía de señal de una o más porciones de señal no transitoria de la señal de audio, o a una característica de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos (132);
- 60 un procesador de señal (140) configurado para procesar la señal de audio de transitorios reducidos (132), para obtener una versión procesada (142) de la señal de audio de transitorios reducidos; y
- un reinsertador de señal transitoria (150) configurado para combinar la versión procesada (142) de la señal de audio de transitorios reducidos (132) con una señal transitoria (152) representando, en una forma original o procesada, un contenido transitorio de la porción de señal transitoria;

5 donde el sustituidor de señal transitoria (130) está configurado para extrapolar, en un dominio de tiempo-frecuencia, coeficientes en el dominio temporal-frecuencial de valor complejo asociados con una porción de señal no transitoria de la señal de audio (110) precediendo a la porción de señal transitoria, para obtener coeficientes en el dominio de tiempo-frecuencia de la porción de señal de reemplazo, o

10 donde el sustituidor de señal transitoria (130) está configurado para interpolar, en un dominio de tiempo-frecuencia, entre coeficientes en el dominio temporal-frecuencial de valor complejo asociados con una porción de señal no transitoria de la señal de audio (110) precediendo a la porción de señal transitoria, y coeficientes en el dominio de tiempo-frecuencia de valores complejos asociados con una porción de señal no transitoria de la señal de audio que sigue a la porción de señal transitoria, para obtener coeficientes en el dominio de tiempo-frecuencia de la porción de señal de reemplazo;

15 en el cual el sustituidor de señal transitoria (130) comprende un detector de transitorios (130a, 130c) configurado para detectar una porción de señal transitoria de la señal de audio (110) en base a una monitorización de la señal de audio (110), o en base a información lateral que acompaña a la señal de audio, y para determinar una longitud de la porción de señal transitoria;

20 en el cual el sustituidor de señal transitoria (130) está configurado para tener en cuenta la longitud de la porción de señal transitoria determinada por el detector de transitorios (130a, 130c).

25 4. Aparato (100) según una cualquiera de las reivindicaciones 1 a 3, donde el sustituidor de señal transitoria (130) está configurado para suministrar la porción de señal de sustitución de modo que la porción de señal de sustitución represente una señal temporal que presenta una evolución en el tiempo alisada, comparada con la porción de señal transitoria, de modo que una desviación entre una energía de la porción de señal de sustitución y una energía de una porción de señal no transitoria de la señal de audio (110) que precede a la porción de señal transitoria o que sigue a la porción de señal transitoria sea inferior a un valor de umbral predeterminado.

30 5. Aparato (100) según una cualquiera de las reivindicaciones 1 a 4, donde el sustituidor de señal transitoria (130) está configurado para aplicar un ruido ponderado para obtener los valores de amplitud de la porción de señal de sustitución, o para aplicar un ruido ponderado para obtener los valores de fase de las porciones de señal de sustitución.

35 6. Aparato (100) según una cualquiera de las reivindicaciones 1 a 4, donde el sustituidor de señal transitoria (130) está configurado para combinar las componentes no transitorias de la porción de señal transitoria con los valores extrapolados o interpolados, para obtener la porción de señal de sustitución.

7. Aparato (100) según una de las reivindicaciones 1 a 6, donde el sustituidor de señal transitoria (130) está configurado para obtener porciones de señal de sustitución de longitud variable en función de una longitud de la presente porción de señal transitoria.

40 8. Aparato (100) según una cualquiera de las reivindicaciones 1 a 7, donde el procesador de señal (140) está configurado para procesar la señal de audio de transitorios reducidos (132) de modo que una porción de señal temporal dada de la versión procesada (142) de la señal de audio de transitorios reducidos dependa de una pluralidad de porciones de señal temporalmente desplazadas de la señal de audio de transitorios reducidos (132).

45 9. Aparato (100) según una cualquiera de las reivindicaciones 1 a 8, donde el procesador de señal (140) está configurado para efectuar un procesamiento en base a bloques en el tiempo de la señal de audio de transitorios reducidos (132), para obtener la versión procesada (142) de la señal de audio de transitorios reducidos; y

50 en el cual el sustituidor de señal transitoria (130) está configurado para ajustar la duración de la porción de señal transitoria a sustituir por la porción de señal de sustitución con una resolución en el tiempo que es más fina que la duración de un bloque temporal, o para sustituir una porción de señal transitoria que tiene una duración en el tiempo inferior a la duración del bloque temporal por una porción de señal de sustitución que tiene una duración en el tiempo inferior a la duración del bloque temporal.

55 10. Aparato (100) según una cualquiera de las reivindicaciones 1 a 9, donde el procesador de señal (140) está configurado para procesar la señal de audio de transitorios reducidos (132) en función de la frecuencia, de modo que el procesamiento introduzca en la señal de audio de transitorios reducidos (132) desfases en función de la frecuencia que degradan el transitorio.

60 11. Aparato (100) según una cualquiera de las reivindicaciones 1 a 10, donde el sustituidor de señal transitoria (130) comprende un detector de transitorios (130a), donde el detector de transitorios (130a) está configurado para suministrar un

umbral de detección variable en el tiempo para la detección del transitorio en la señal de audio (110) de modo que el umbral de detección siga una envolvente de la señal de audio con una constante de tiempo de alisado ajustable, y

5 en el cual el detector de transitorios está configurado para cambiar la constante de tiempo de alisado en respuesta a la detección de un transitorio y/o en función de una evolución en el tiempo de la señal de audio.

10 12. Aparato (100) según una cualquiera de las reivindicaciones 1 a 11, donde el aparato (100) comprende un procesador de transitorios (160) configurado para recibir una información sobre transitorios (134) y para obtener, en base a la información sobre transitorios (134), una señal procesada transitoria (152) en la cual las componentes tonales son reducidas, y

en el cual el reinsertador de señal transitoria (150) está configurado para combinar la versión procesada (142) de la señal de audio de transitorios reducidos (132) con la señal procesada transitoria (152) suministrada por el procesador de transitorios (160).

15 13. Aparato (100) según una cualquiera de las reivindicaciones 1 a 12,

20 en el cual el sustituidor de señal transitoria (130) comprende un detector de transitorios (130a, 130c) configurado para detectar una porción de señal transitoria de la señal de audio (110) en base a una monitorización de la señal de audio (110), o en base a información lateral que acompaña a la señal de audio, y para determinar una longitud de la porción de señal transitoria;

en el cual el sustituidor de señal transitoria (130) está configurado para tener en cuenta la longitud de la porción de señal transitoria determinada por el detector de transitorios (130a, 130c);

25 en el cual el sustituidor de señal transitoria (130) está configurado para extrapolar, en un dominio temporal-frecuencial, coeficientes en el dominio temporal-frecuencial de valor complejo asociados a una porción de señal no transitoria de la señal de audio (110) que precede a la porción de señal transitoria, para obtener coeficientes en el dominio temporal-frecuencial de la porción de señal de sustitución, o

30 en el cual el sustituidor de señal transitoria (130) está configurado para interpolar, en un dominio temporal-frecuencial, entre coeficientes en el dominio temporal-frecuencial de valor complejo asociados a una porción de señal no transitoria de la señal de audio (110) que precede a la porción de señal transitoria, y coeficientes en el dominio temporal-frecuencial de valor complejo asociados a una porción de señal no transitoria de la señal de audio que sigue a la porción de señal transitoria, para obtener coeficientes en el dominio temporal-frecuencial de la porción de señal de sustitución;

35 en el cual el procesador de señal (140) está configurado para efectuar un procesamiento de señal de audio que degrada el transitorio mediante alargamiento en el tiempo o compresión en el tiempo, de modo que la señal procesada (142) suministrada por el procesador de señal (140) comprenda una duración superior, o inferior, a una duración de la señal sin procesar (132) recibida por el procesador de señal de audio; y

40 en el cual el aparato (100) está configurado para adaptar una modulación en el tiempo o una de tasa de muestreo de la señal obtenida por el reinsertador de señal transitoria (150) de modo que al menos las componentes no transitorias de la señal obtenida por el reinsertador de señal transitoria (150) estén transpuestas en frecuencia en comparación con la señal de audio (110) que entra en el sustituidor de señal transitoria (130).

45 14. Aparato (100) según una cualquiera de las reivindicaciones 1 a 13, donde el reinsertador de señal transitoria (150) está configurado para encadenar la versión procesada (142) de la señal de audio de transitorios reducidos (132) con una señal transitoria (152) que representa, de forma original o procesada, un contenido transitorio de la porción de señal transitoria.

50 15. Proceso (1200) para la manipulación de una señal de audio que comprende un evento transitorio, comprendiendo el proceso:

55 sustituir (1210) una porción de señal transitoria, que comprende el evento transitorio, de la señal de audio por una porción de señal de sustitución adaptada a las características de energía de señal de una o varias porciones de señal no transitorias de la señal de audio, o a características de energía de señal de la porción de señal transitoria, para obtener una señal de audio de transitorios reducidos;

60 procesar (1220) la señal de audio de transitorios reducidos, para obtener una versión procesada de la señal de audio de transitorios reducidos; y

combinar (1230) la versión procesada de la señal de audio de transitorios reducidos con una señal transitoria que

representa, de forma original o procesada, un contenido transitorio de la porción de señal transitoria;

5 donde los valores de amplitud de una o más porciones de señal que preceden a la porción de señal transitoria son extrapolados para obtener valores de amplitud de la porción de señal de reemplazo, y donde los valores de fase de una o más porciones de señal que preceden a la porción de señal transitoria son extrapolados para obtener valores de fase de la porción de señal de reemplazo; o

10 donde una interpolación se lleva a cabo entre un valor de amplitud de una porción de señal que precede a la porción de señal transitoria y un valor de amplitud de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o más valores de amplitud de la porción de señal de reemplazo, y

15 donde una interpolación se lleva a cabo entre un valor de fase de una porción de señal que precede a la porción de señal transitoria y un valor de fase de una porción de señal que sigue a la porción de señal transitoria, para obtener uno o más valores de fase de la porción de señal de reemplazo; o

20 donde los coeficientes en el dominio temporal-frecuencial de valor complejo asociados con una porción de señal no transitoria de la señal de audio que precede a la porción de señal transitoria son extrapolados en un dominio temporal-frecuencial, para obtener coeficientes en el dominio temporal-frecuencial de la porción de señal de reemplazo; o

25 donde una interpolación se lleva a cabo, en un dominio temporal-frecuencial, entre coeficientes en el dominio temporal-frecuencial de valor complejo asociados con una porción de señal no transitoria de la señal de audio que precede a la porción de señal transitoria, y coeficientes en el dominio temporal-frecuencial de valor complejo asociados con una porción de señal no transitoria de la señal de audio que siguen a la porción de señal transitoria, para obtener coeficientes en el dominio temporal-frecuencial de la porción de señal de reemplazo;

30 donde el proceso comprende detectar una porción de señal transitoria de la señal de audio (110) en base a una monitorización de la señal de audio (110), o en base a una información lateral que acompaña a la señal de audio, y determinar una longitud de la porción de señal de audio;

35 donde el proceso comprende tener en cuenta la longitud determinada de la porción de señal transitoria.

16. Programa informático para realizar el proceso según la reivindicación 15 cuando se ejecuta el programa informático en un ordenador.

35

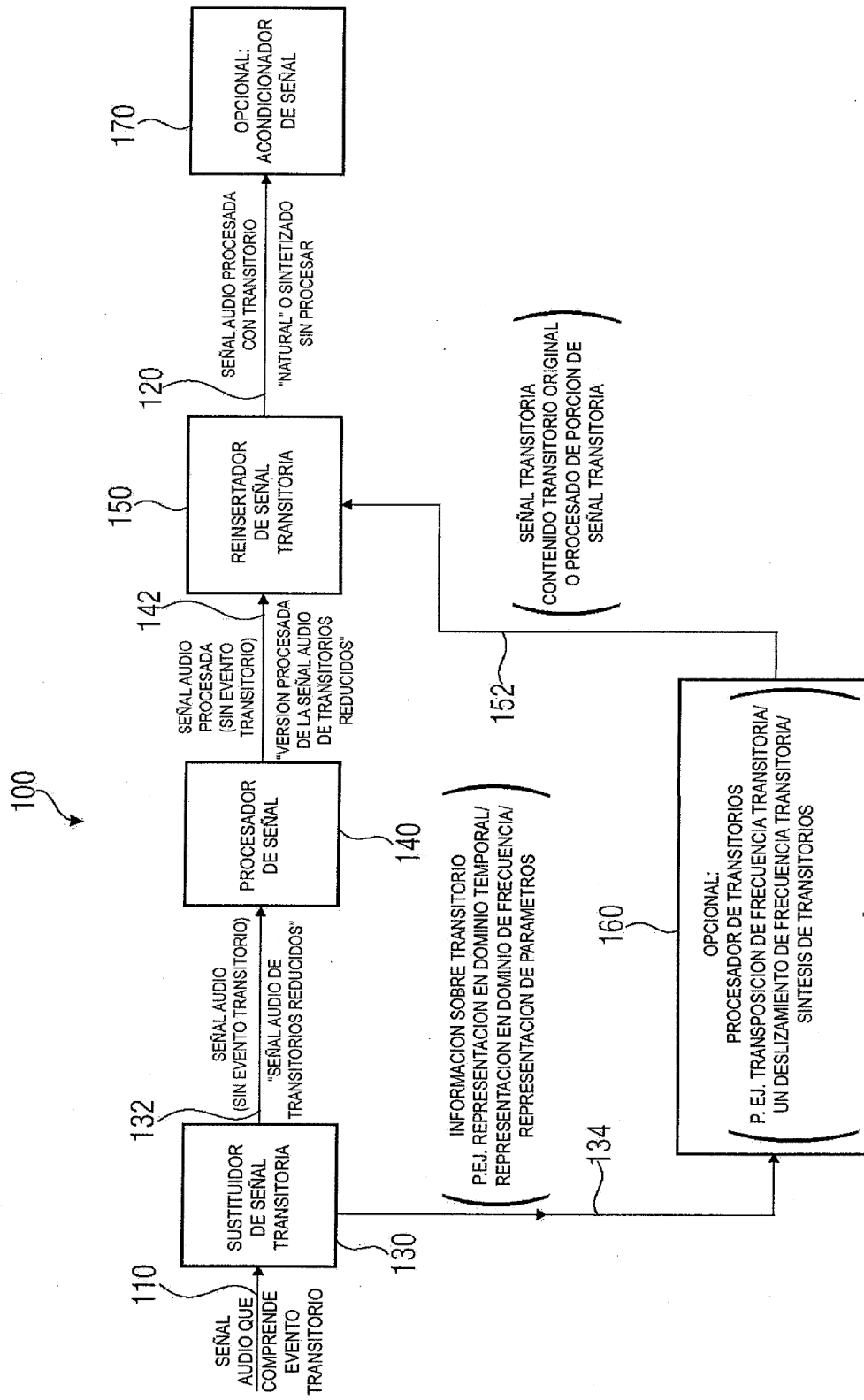


FIG 1

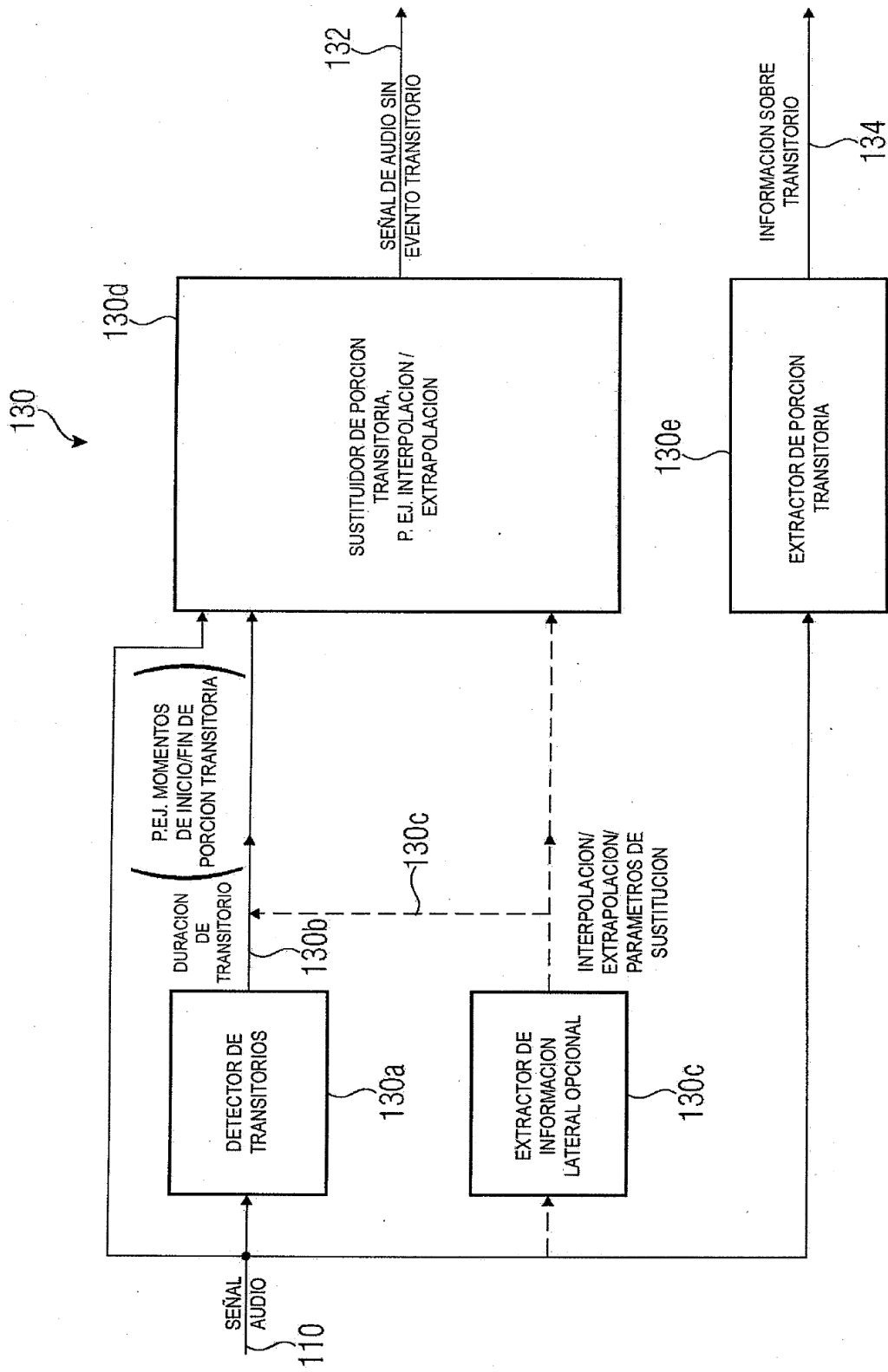


FIG 2

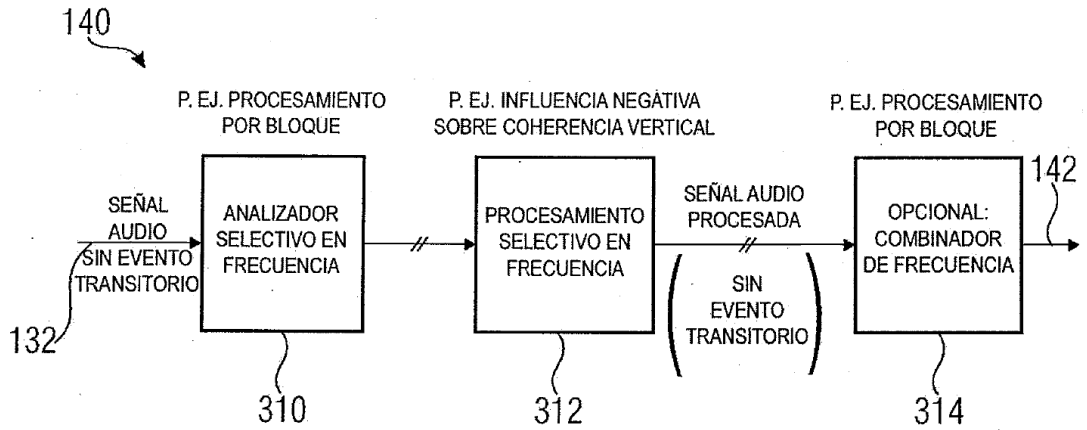


FIG 3A

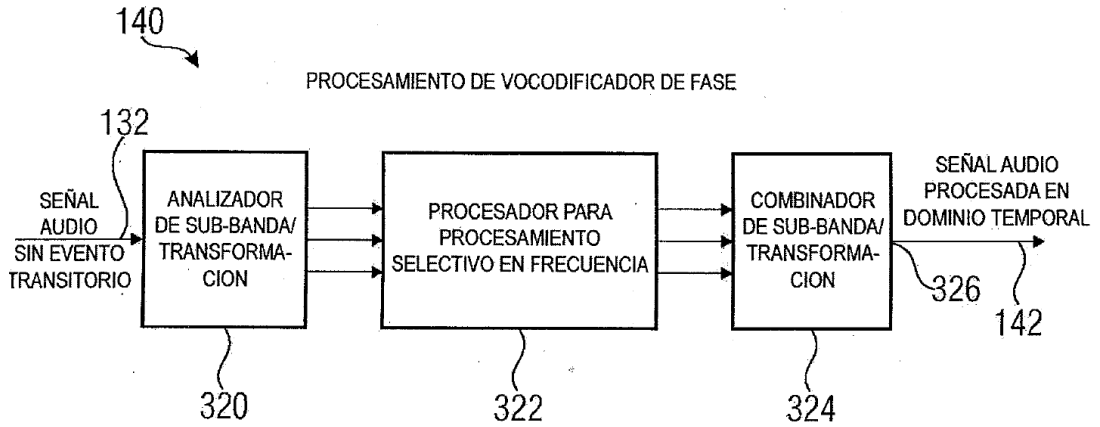


FIG 3B

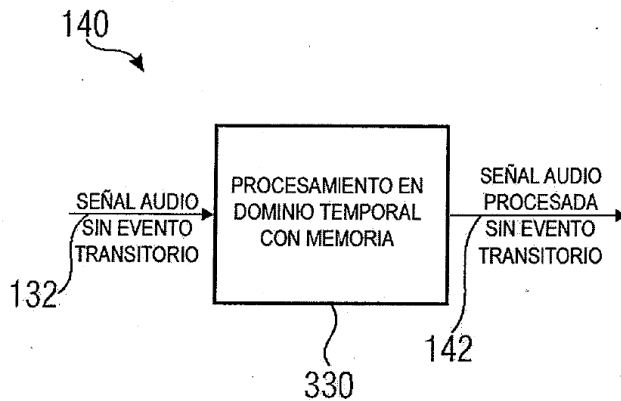


FIG 3C

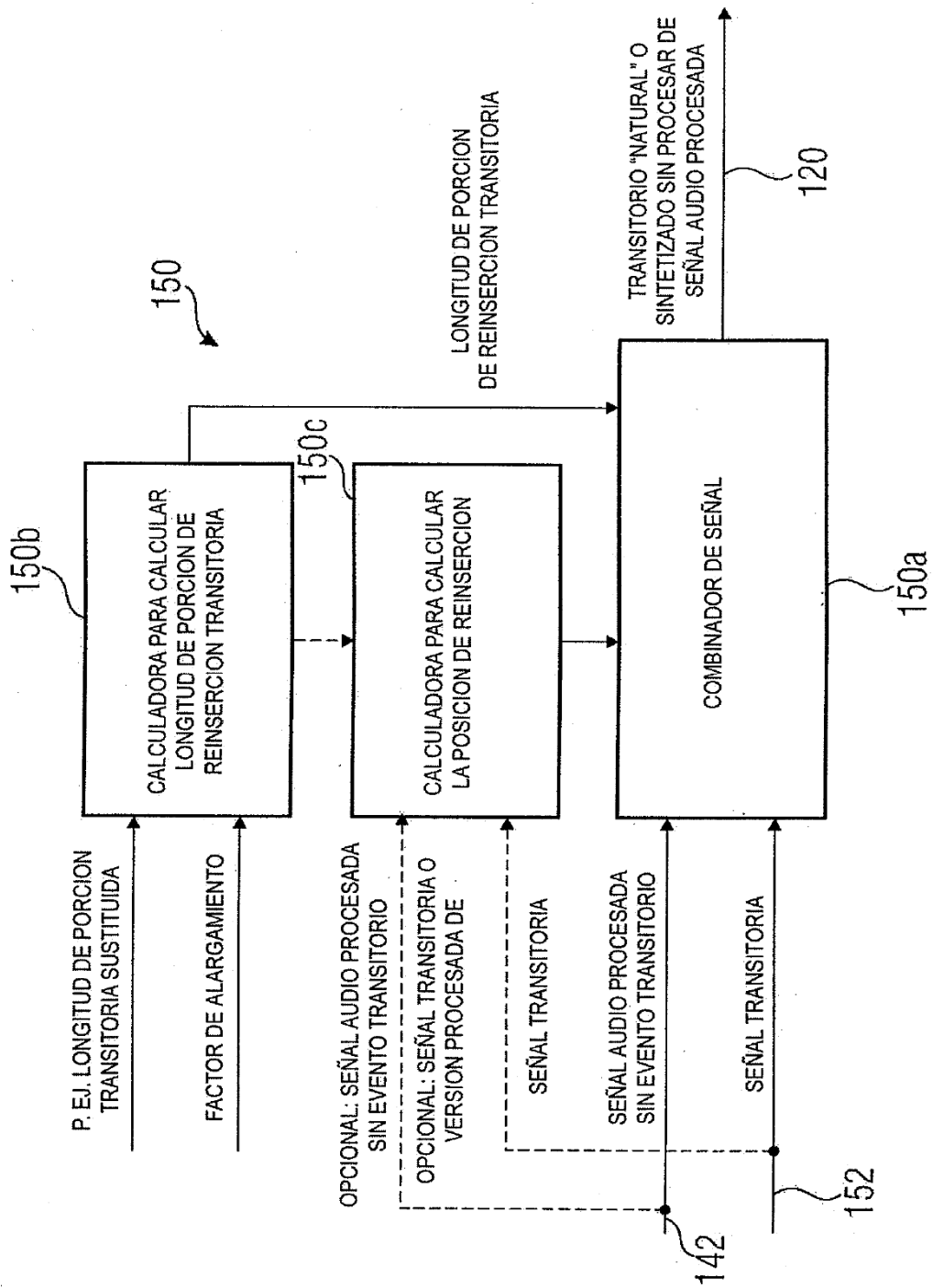


FIG 4

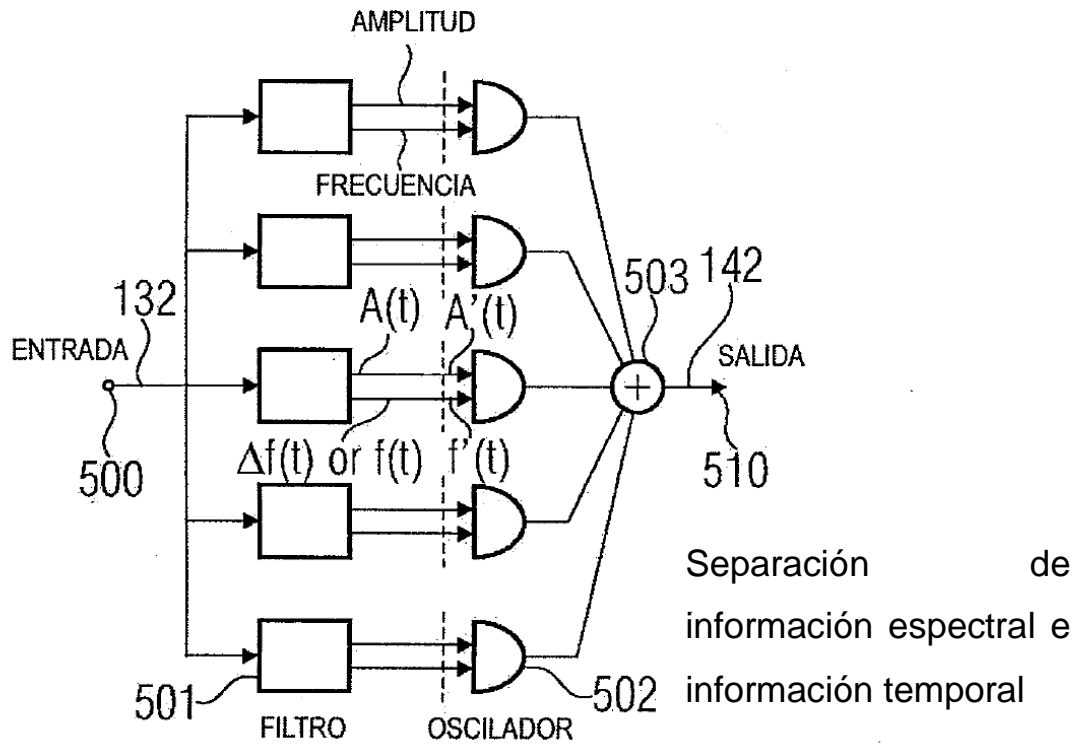


FIG 5A

(IMPLEMENTACION DE BANCO DE FILTROS)

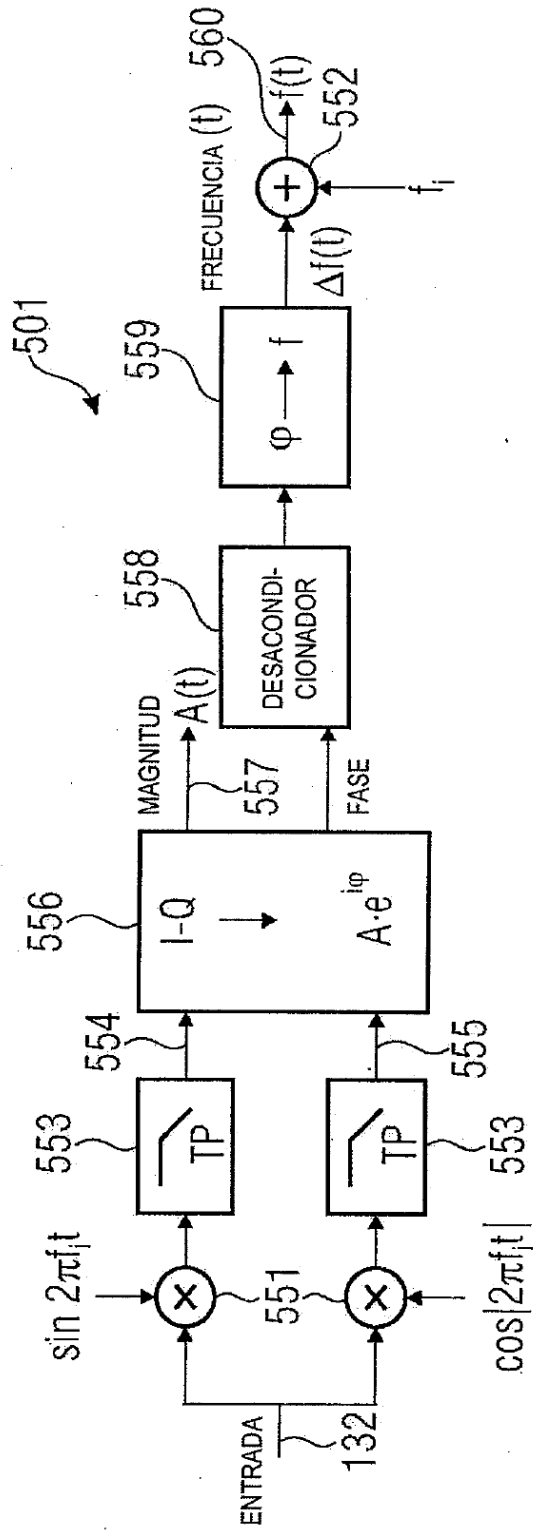


FIG 5B

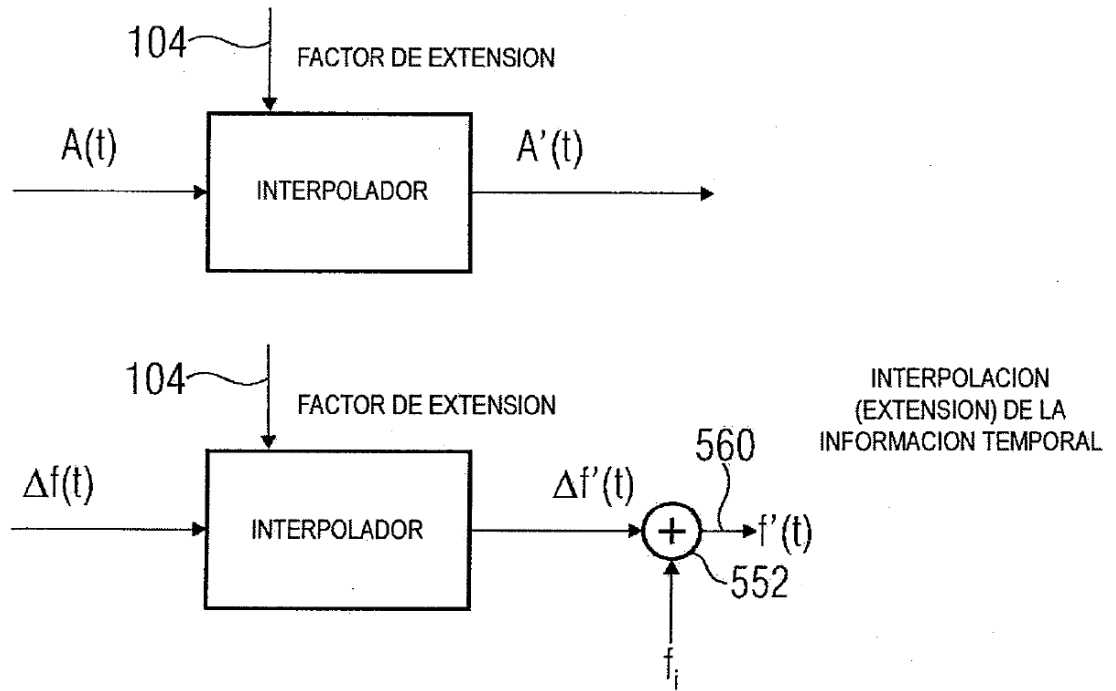


FIG 5C

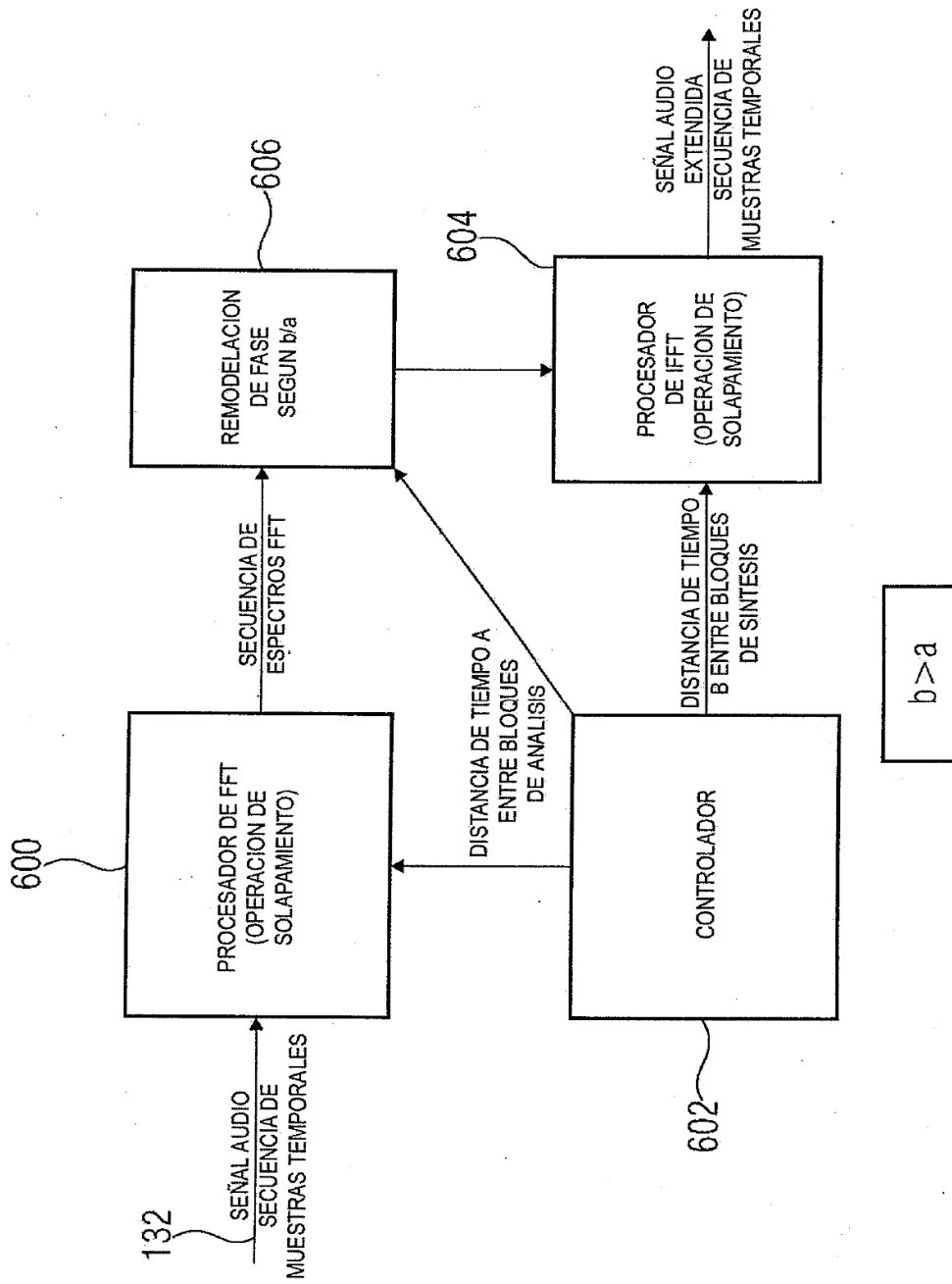


FIG 6

(IMPLEMENTACION DE TRANSFORMACION)

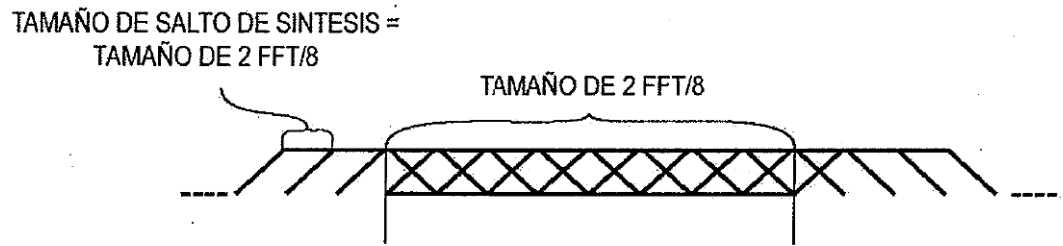
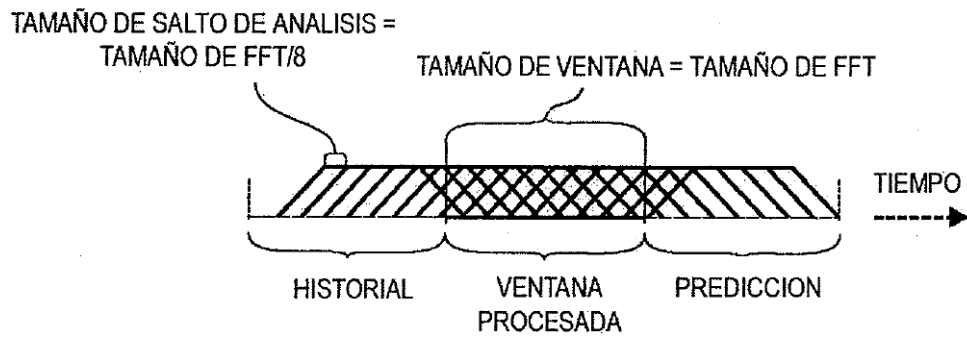


FIG 7

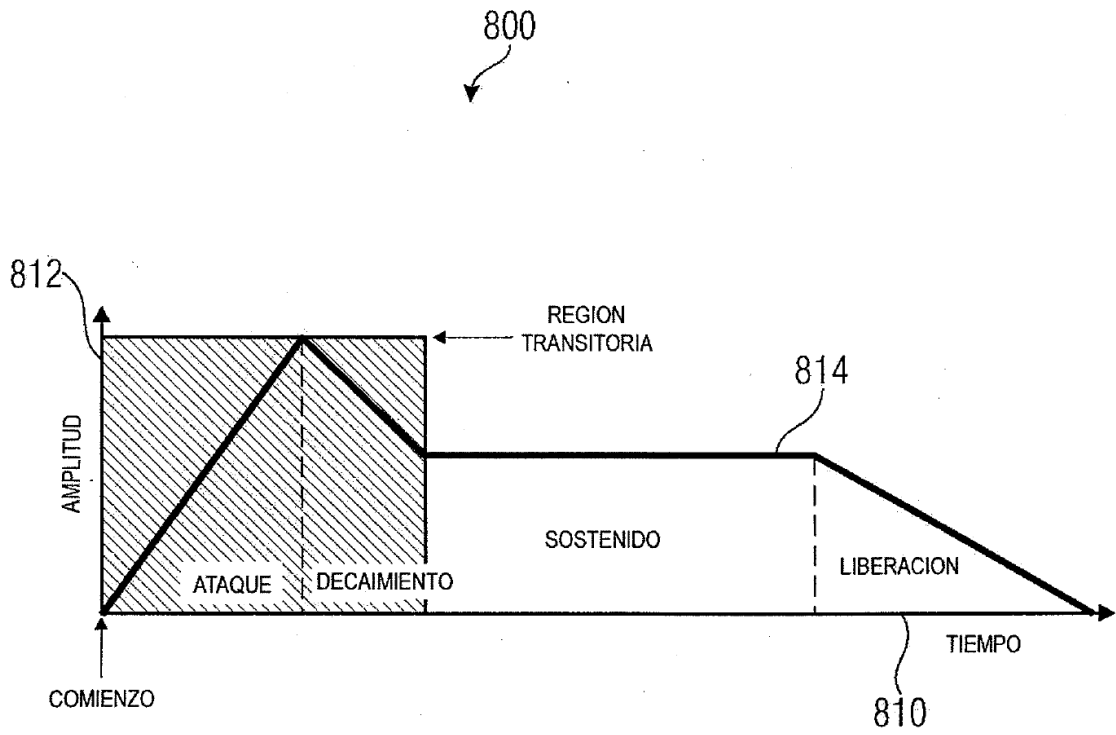
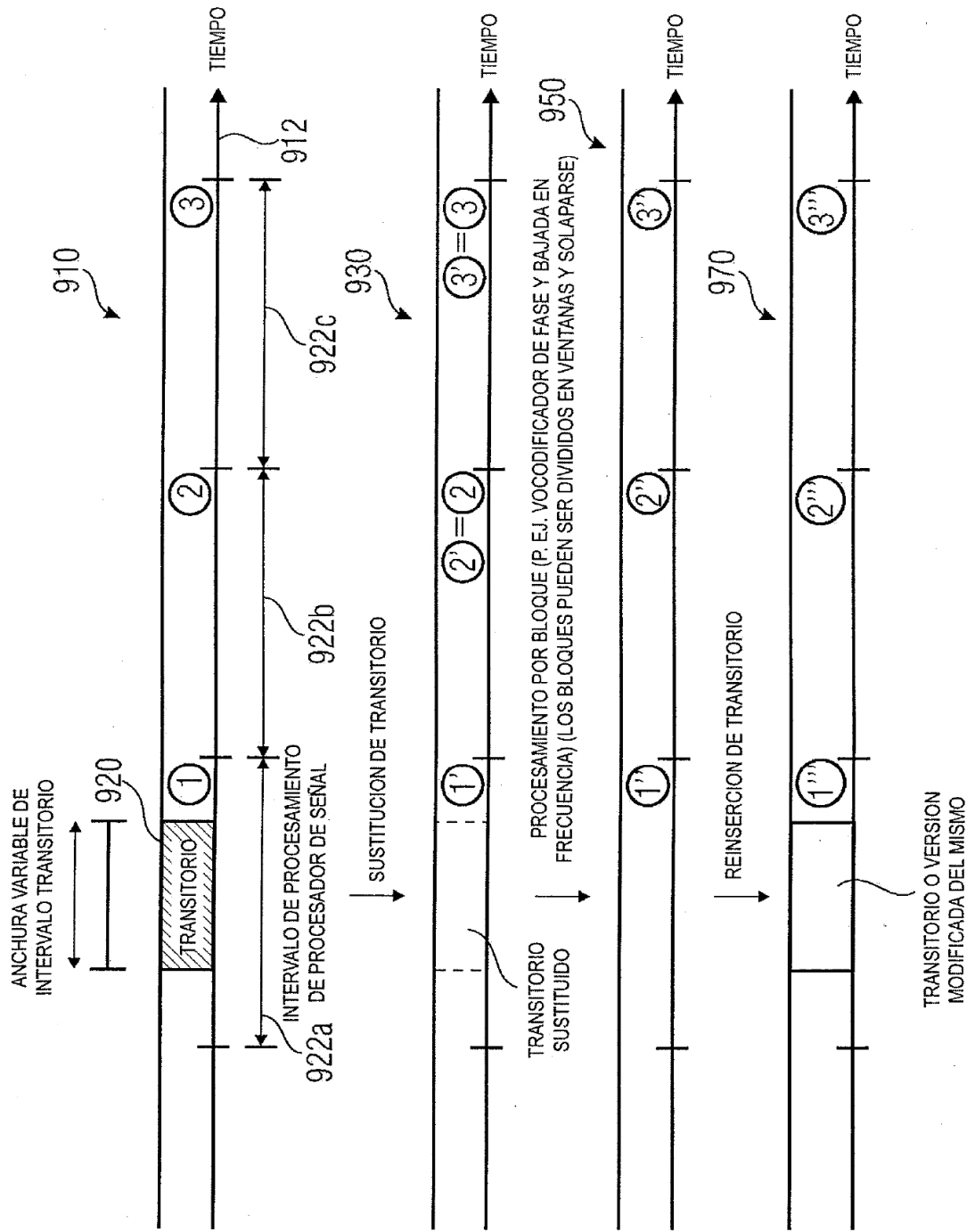


FIG 8

FIG 9



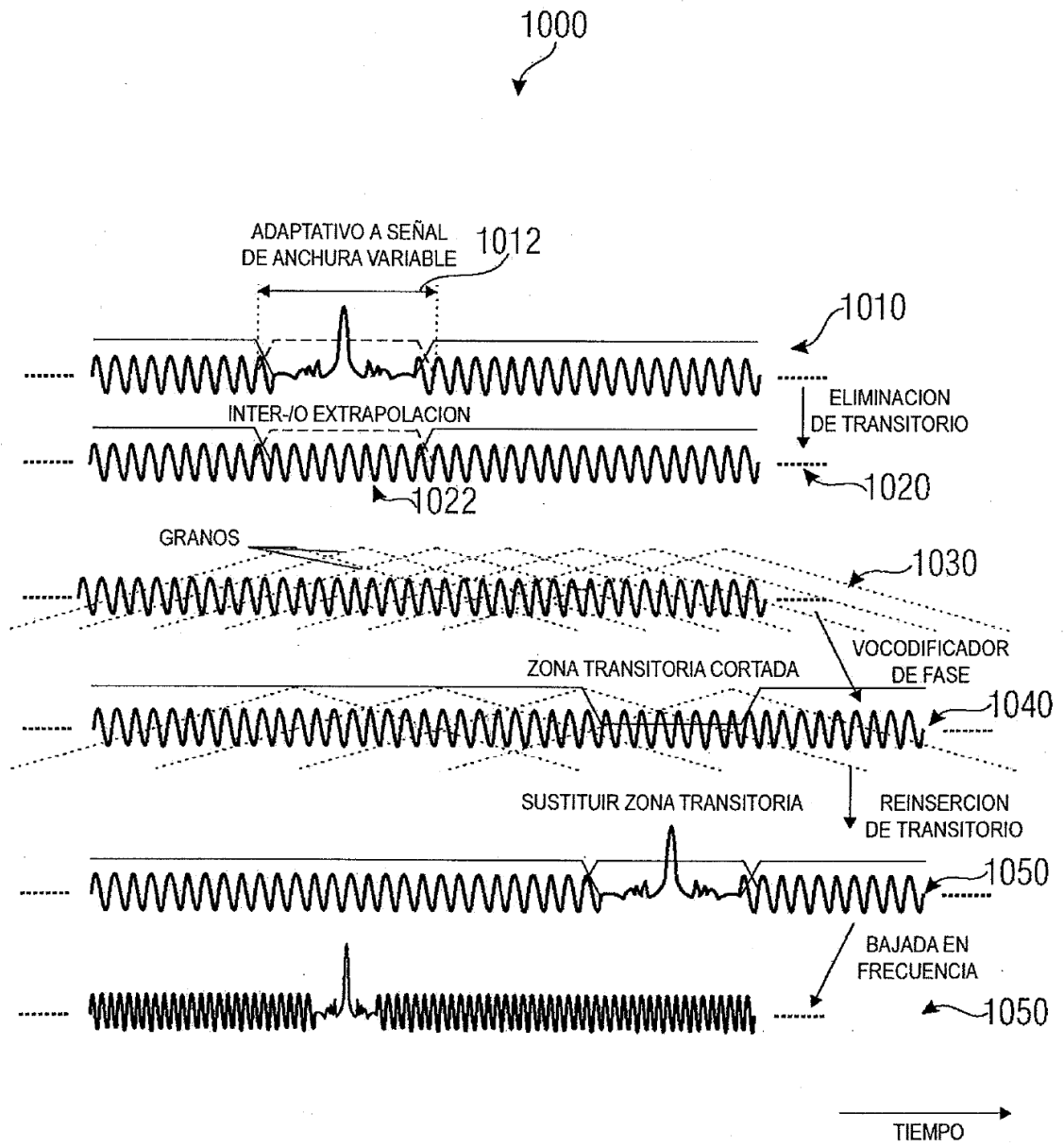


FIG 10

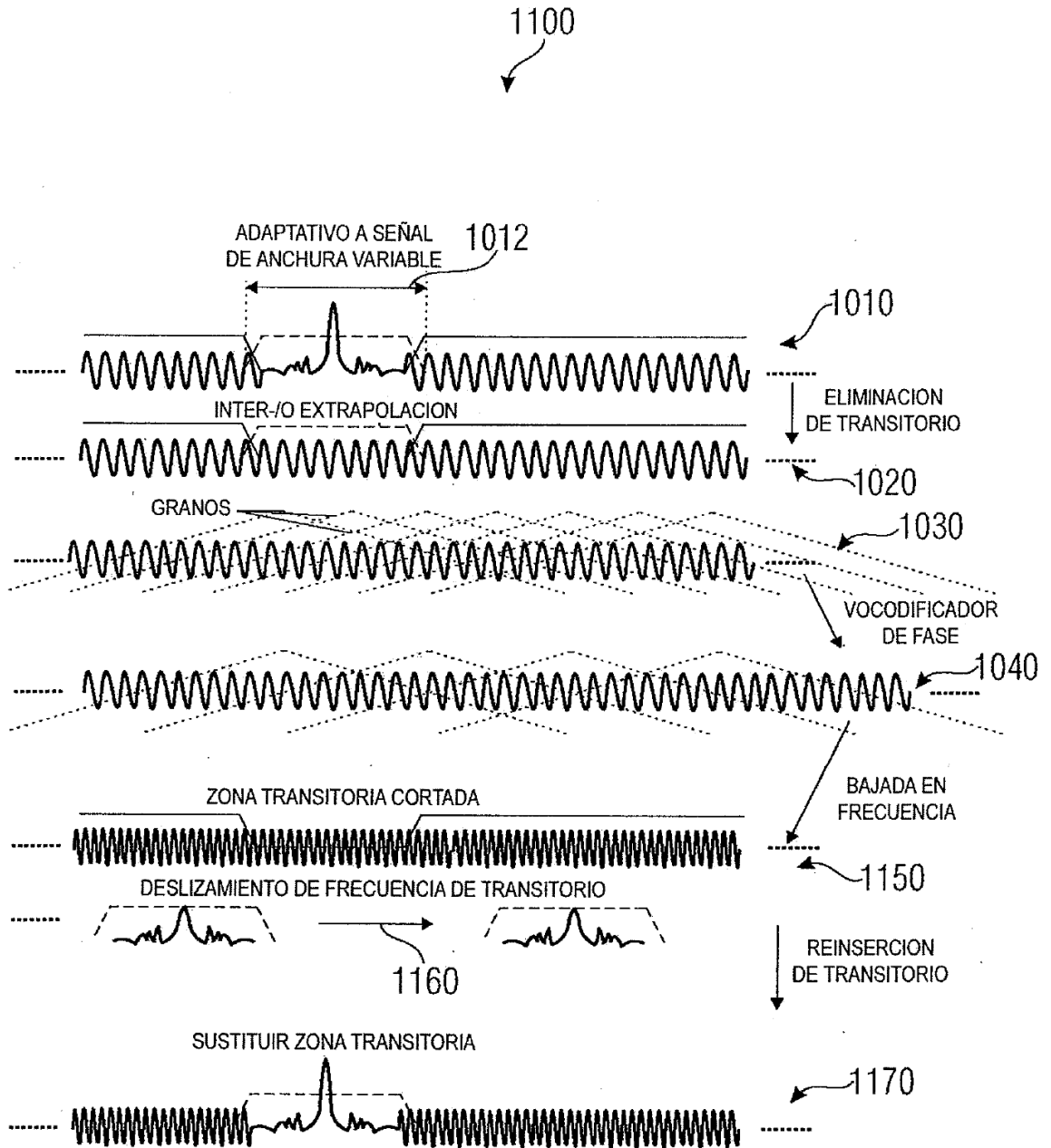


FIG 11

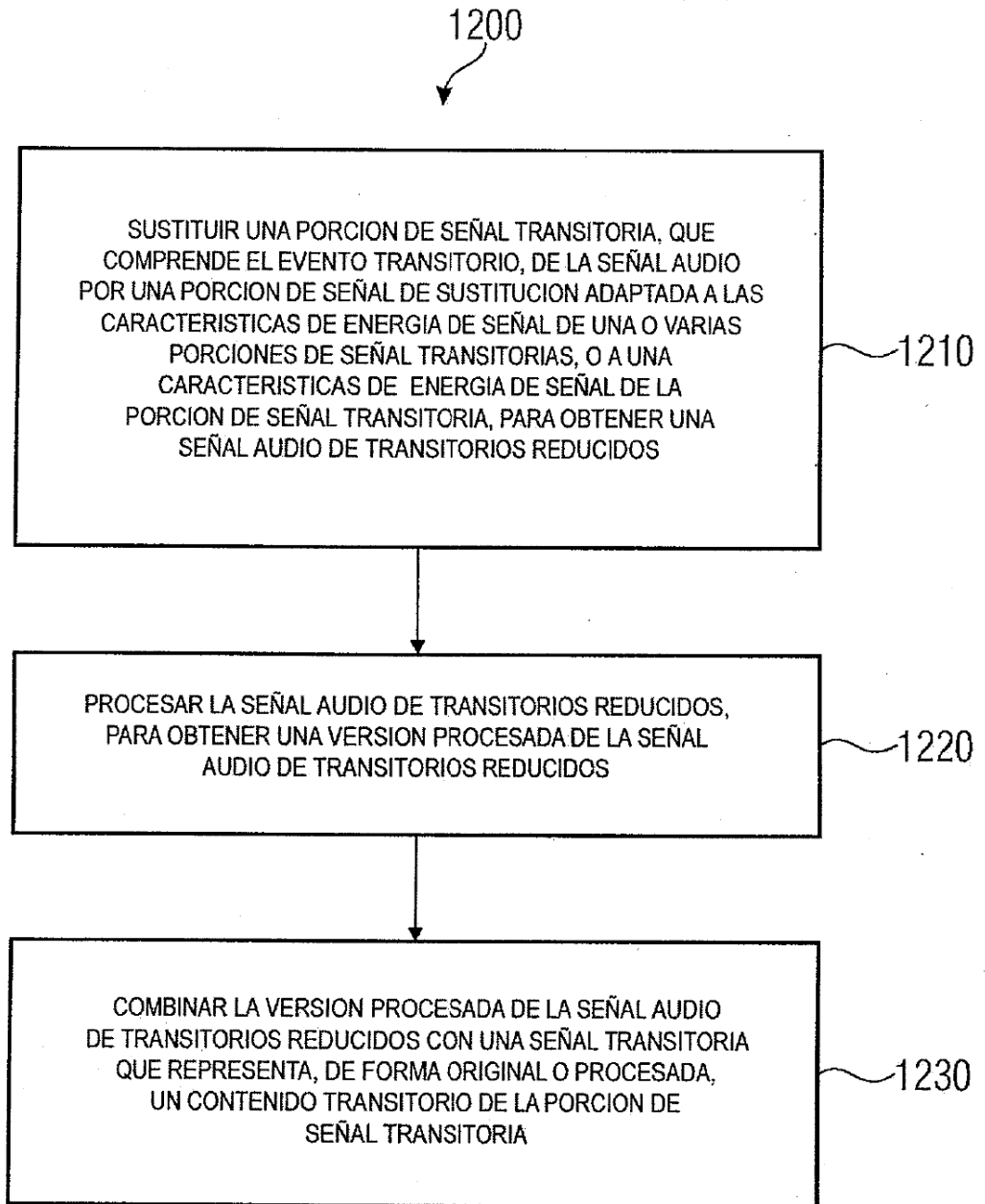


FIG 12

FIG 13

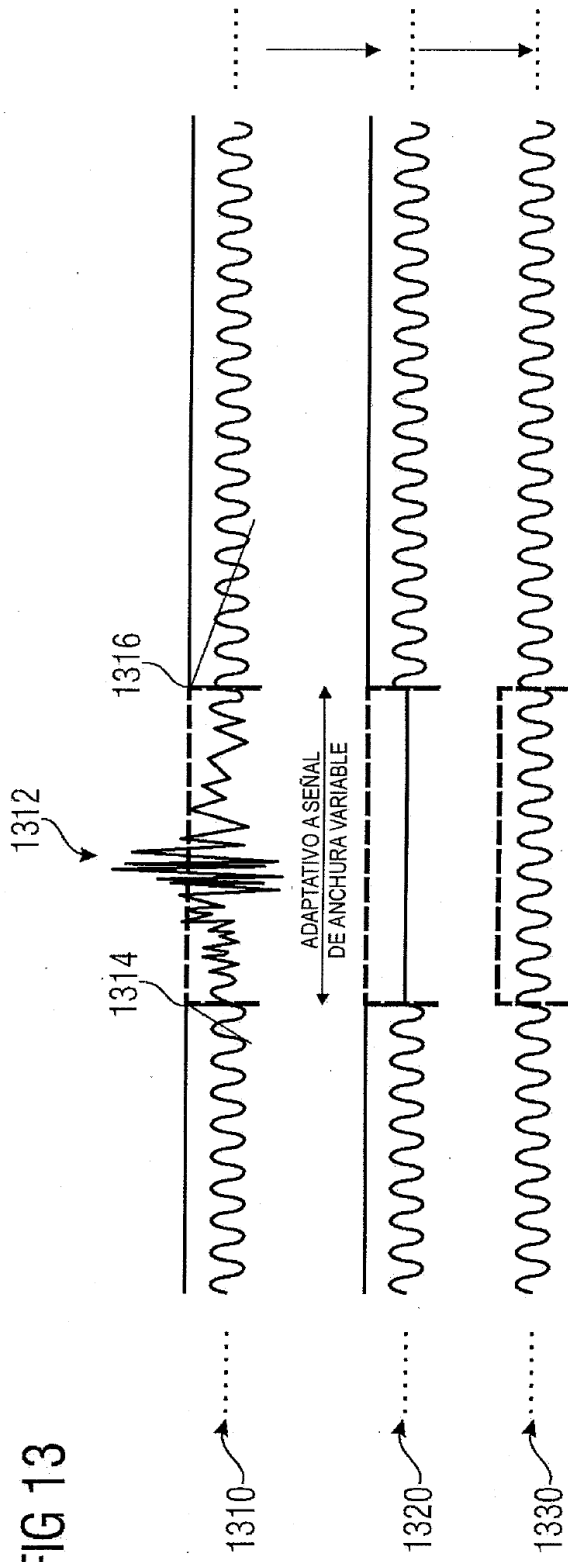


FIG 14

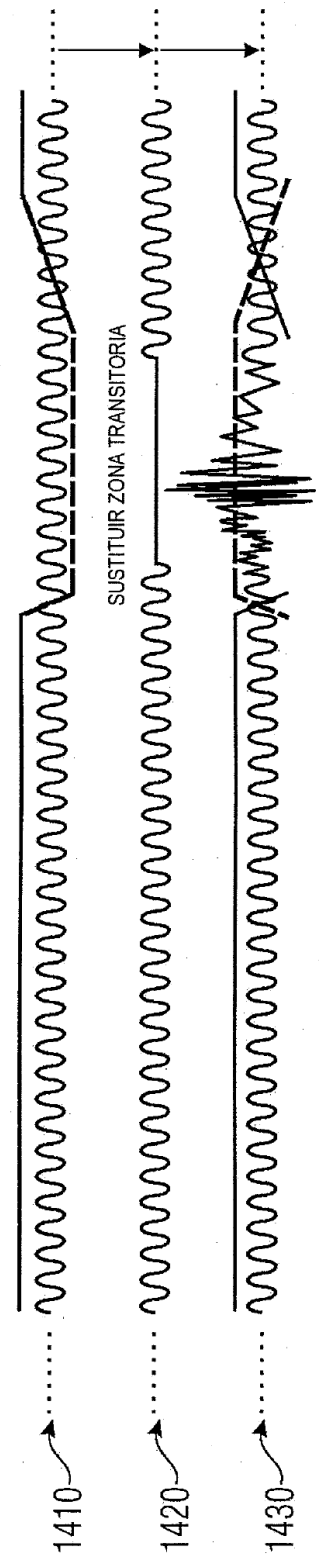
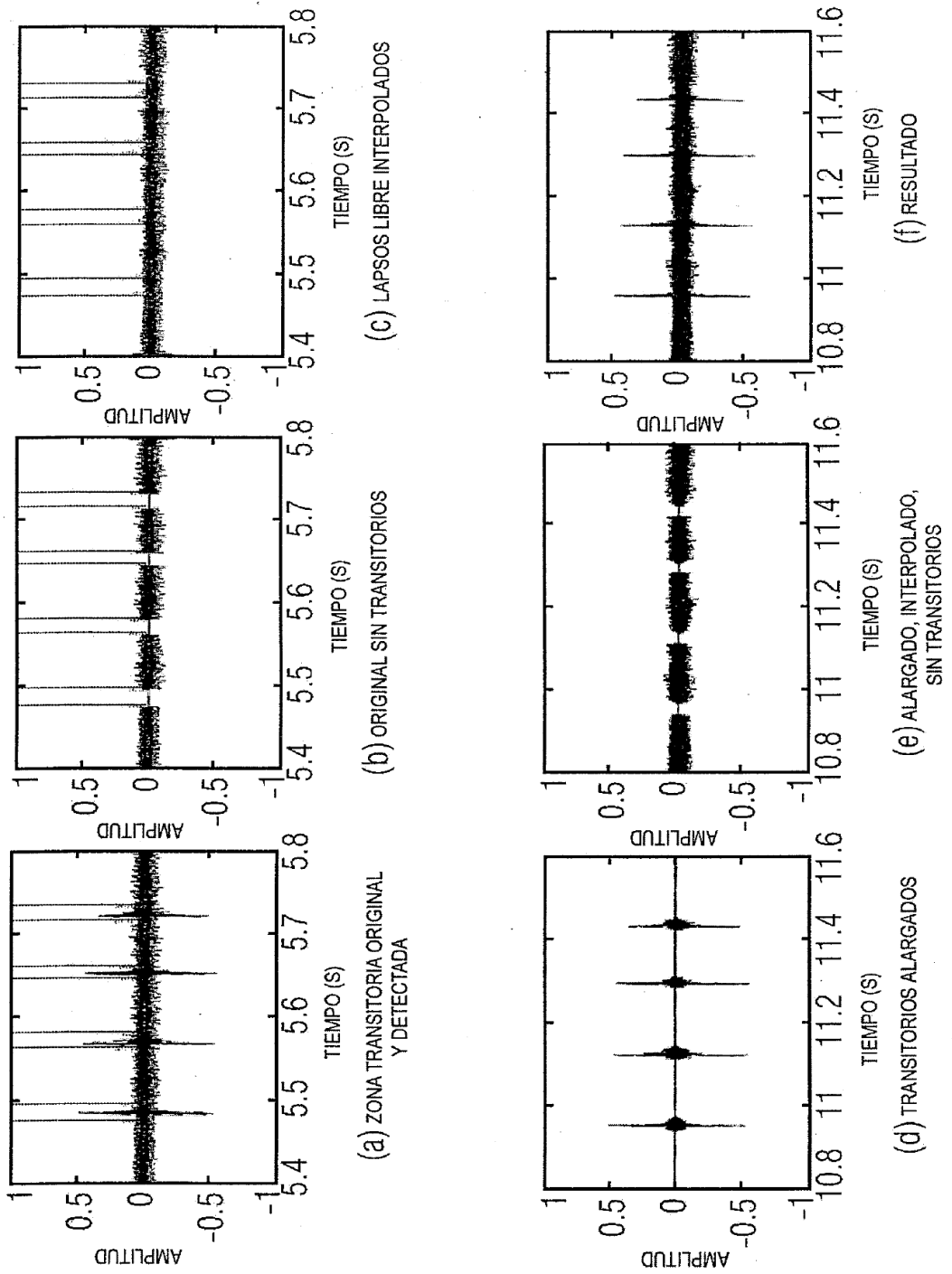


FIG 15



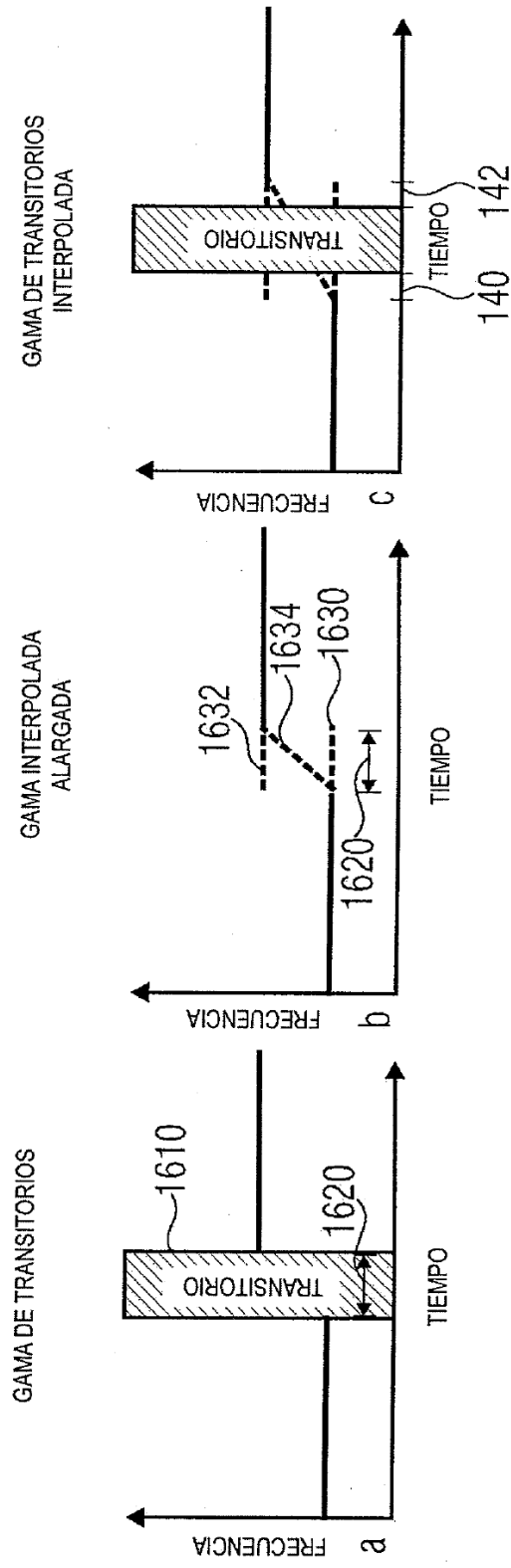


FIG 16