

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 132**

51 Int. Cl.:

G10L 21/0364 (2013.01)

G10L 19/26 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.06.2013 PCT/FR2013/051517**

87 Fecha y número de publicación internacional: **03.01.2014 WO14001730**

96 Fecha de presentación y número de la solicitud europea: **28.06.2013 E 13744654 (8)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 2867893**

54 Título: **Atenuación eficaz de preecos en una señal de audio digital**

30 Prioridad:

29.06.2012 FR 1256285

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

30.04.2019

73 Titular/es:

**ORANGE (100.0%)
78, rue Olivier de Serres
75015 Paris, FR**

72 Inventor/es:

**KOVESI, BALAZS y
RAGOT, STÉPHANE**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 711 132 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

Atenuación eficaz de preecos en una señal de audio digital

5 La invención se refiere a un procedimiento y a un dispositivo de procesamiento de atenuación de los preecos durante la decodificación de una señal de audio digital.

10 Para el transporte de las señales de audio digitales sobre las redes de transmisión, ya se trate, por ejemplo, de redes fijas o móviles, o para el almacenamiento de las señales, se recurre a unos procesos de compresión (o codificación fuente) que implementan unos sistemas de codificación del tipo codificación temporal o codificación de frecuencia por transformada.

15 El procedimiento y el dispositivo, objetos de la invención, tienen, de este modo, como campo de aplicación la compresión de las señales sonoras, en particular, las señales de audio digitales codificadas por transformada de frecuencia.

La figura 1 representa a título ilustrativo, un esquema de principio de la codificación y de la decodificación de una señal de audio digital por transformada que consta de un análisis-síntesis por adición/recubrimiento según la técnica anterior.

20 Algunas secuencias musicales, tales como las percusiones y algunos segmentos de palabra como las oclusivas (/k/, /t/, ...), están caracterizadas por unos ataques extremadamente bruscos que se traducen en unas transiciones muy rápidas y una variación muy fuerte de la dinámica de la señal en el espacio de algunas muestras. Un ejemplo de transición se da en la figura 1 a partir de la muestra 410.

25 Para el procesamiento de codificación/decodificación, la señal de entrada se recorta en bloques de muestras de longitud L, representados en la figura 1 por unos trazos verticales en punteados. La señal de entrada está anotada $x(n)$, donde n es el índice de la muestra. El recorte en bloques sucesivos conduce a definir los bloques $X_N(n) = [x(N.L) \dots x(N.L+L-1)] = [x_N(0) \dots x_N(L-1)]$, donde N es el índice de la trama, L es la longitud de la trama. En la figura 1 se tiene $L=160$ muestras. En el caso de la transformada modulada de coseno modificada MDCT (para "Modified Discrete Cosine Transform" en inglés), se analizan dos bloques $X_N(n)$ y $X_{N+1}(n)$ conjuntamente para dar un bloque de coeficientes transformados asociados a la trama de índice N.

35 La división en bloques, también llamados tramas, operada por la codificación por transformada es totalmente independiente de la señal sonora y las transiciones pueden aparecer, por lo tanto, en un punto cualquiera de la ventana de análisis. Ahora bien, después de decodificación por transformada, la señal reconstruida está manchada de "ruido" (o distorsión) generado por la operación de cuantificación (C)-cuantificación inversa (C^{-1}). Este ruido de codificación se reparte temporalmente de forma relativamente uniforme sobre todo el soporte temporal del bloque transformado, es decir, sobre toda la longitud de la ventana de longitud $2L$ de muestras (con recubrimiento de L muestras). La energía del ruido de codificación es, en general, proporcional a la energía del bloque y es función de la velocidad de codificación/decodificación.

40 Para un bloque que incluye un ataque (como el bloque 320-480 de la figura 1) la energía de la señal es elevada, por lo tanto, el ruido es, igualmente, de nivel elevado.

45 En codificación por transformada, el nivel del ruido de codificación es típicamente inferior al de la señal para los segmentos de fuerte energía que siguen inmediatamente a la transición, pero el nivel es superior al de la señal para los segmentos de energía más escasa, en concreto, sobre la parte que precede a la transición (muestras 160 - 410 de la figura 1). Para la parte anteriormente citada, la relación de señal a ruido es negativa y la degradación resultante puede aparecer muy molesta para la escucha. Se llama preeco al ruido de codificación anterior a la transición y posteco al ruido posterior a la transición.

Se puede observar en la figura 1 que el preeco afecta a la trama que precede a la transición, así como a la trama donde se produce la transición.

55 Los experimentos físico-acústicos han mostrado que el oído humano efectúa un preenmascaramiento temporal de los sonidos bastante limitado, del orden de algunos milisegundos. El ruido que precede al ataque, o preeco, es audible cuando la duración del preeco es superior a la duración del preenmascaramiento.

60 El oído humano efectúa, igualmente, un postenmascaramiento de una duración más larga, de 5 a 60 milisegundos, durante el paso de secuencias de fuerte energía a unas secuencias de escasa energía. La tasa o nivel de molestia aceptable para los postecos es, por lo tanto, más importante que para los preecos.

65 El fenómeno de los preecos, más crítico, es tanto más molesto en cuanto que la longitud de los bloques en número de muestras es importante. Ahora bien, en codificación por transformada, se conoce bien que, para las señales estacionarias, cuanto más aumenta la longitud de la transformada, más importante es la ganancia de codificación. A frecuencia de muestreo fijada y a velocidad fijada, si se aumenta el número de puntos de la ventana (por lo tanto, la

longitud de la transformada) se dispondrá de más bits por trama para codificar las rayas de frecuencia que se consideran útiles por el modelo psicoacústico, de donde procede la ventaja de utilizar unos bloques de gran longitud. La codificación MPEG AAC (Advanced Audio Coding), por ejemplo, utiliza una ventana de gran longitud que contiene un número fijo de muestras, 2048, esto es, sobre una duración de 64 ms a una frecuencia de muestreo de 32 kHz; el problema de los preecos se gestiona ahí permitiendo conmutar de estas ventanas largas a ventanas cortas por medio de ventanas intermedias (de transición), lo que necesita un cierto retardo en la codificación para detectar la presencia de una transición y adaptar las ventanas. La longitud de estas ventanas cortas es, por lo tanto, de 8 ms. A baja velocidad siempre se puede tener un preeco audible de algunos ms. La conmutación de las ventanas permite atenuar el preeco, pero no suprimirlo. Los codificadores por transformada utilizados para las aplicaciones conversacionales como UIT-T G.722.1, G.722.1C o G.719 utilizan a menudo una ventana de duración 40 ms a 16, 32 o 48 kHz (respectivamente) y una longitud de trama de 20 ms. Se puede anotar que el codificador UIT-T G.719 integra un mecanismo de conmutación de ventanas con detección de transitoria, sin embargo, el preeco no está completamente reducido a baja velocidad (típicamente a 32 kbit/s).

Con la finalidad de reducir el efecto molesto anteriormente citado del fenómeno de los preecos, se han propuesto diferentes soluciones al nivel del codificador y/o del decodificador.

La conmutación de ventanas se ha citado con anterioridad. Otra solución consiste en aplicar un filtrado adaptativo. En la zona que precede al ataque, la señal reconstruida se ve como la suma de la señal original y del ruido de cuantificación.

Una técnica de filtrado correspondiente se ha descrito en el artículo que lleva por título High Quality Audio Transform Coding at 64 kbits, IEEE Trans. on Communications Vol. 42, N.º 11, noviembre de 1994, publicado por Y. Mahieux y J. P. Petit.

La implementación de un filtrado de este tipo necesita el conocimiento de parámetros de los cuales algunos, como los coeficientes de predicción y la varianza de la señal corrompida por el preeco, se estiman en el decodificador a partir de las muestras con ruido. Por el contrario, unas informaciones tales como la energía de la señal de origen no pueden conocerse más que en el codificador y deben, por consiguiente, transmitirse. Esto necesita transmitir unas informaciones suplementarias, lo que a velocidad restringida disminuye el presupuesto relativo asignado a la codificación por transformada. Cuando el bloque recibido contiene una variación brusca de dinámica, se le aplica el procesamiento de filtrado.

El proceso de filtrado anteriormente citado no permite encontrar la señal de origen, pero procura una fuerte reducción de los preecos. No obstante, necesita transmitir los parámetros suplementarios al decodificador.

Se han propuesto diferentes técnicas de reducción de preeco sin transmisión específica de la información. Por ejemplo, se presenta una revisión de la reducción de preecos en el contexto de la codificación jerárquica en el artículo B. Kövesi, S. Ragot, M. Gartner, H. Taddei, "Pre-echo reduction in the ITU-T G.729.1 embedded coder," EUSIPCO, Lausana, Suiza, agosto de 2008.

Se describe un ejemplo típico de procedimiento de atenuación de preecos en la solicitud de patente francesa FR 08 56248. En este ejemplo, se determinan unos factores de atenuación por subbloque, en los subbloques de escasa energía que preceden a un subbloque en el que se ha detectado una transición o ataque.

El factor de atenuación por subbloques $g(k)$ se calcula, por ejemplo, en función de la relación $R(k)$ entre la energía del subbloque de energía más fuerte y la energía del k-ésimo subbloque en cuestión:

$$g(k)=f(R(k))$$

donde f es una función decreciente de valores entre 0 y 1 y k es el número del subbloque. Son posibles otras definiciones del factor $g(k)$, por ejemplo, en función de la energía $En(k)$ en el subbloque corriente y de la energía $En(k-1)$ en el subbloque precedente.

Si la variación de la energía con respecto a la energía máxima es escasa, entonces, no es necesaria ninguna atenuación. El factor $g(k)$ está fijado, entonces, a un valor de atenuación que inhibe la atenuación, es decir, 1. Si no, el factor de atenuación está comprendido entre 0 y 1.

En la mayor parte de los casos, sobre todo cuando el preeco es molesto, la trama que precede a la trama de preeco tiene una energía homogénea que corresponde a la energía de un segmento de escasa energía (típicamente, un ruido de fondo). Según el experimento, no es útil ni tampoco deseable que después del procesamiento de atenuación de preeco, la energía de la señal se vuelva inferior a la energía media por subbloque de la señal que precede a la zona de procesamiento (típicamente, la de la trama precedente \bar{En} o la de la segunda mitad de la trama precedente \bar{En}').

Para el subbloque k que hay que procesar, se puede calcular el valor límite del factor $\lim g(k)$, con el fin de obtener exactamente la misma energía que la energía media por subbloque del segmento que precede al subbloque que hay

que procesar. Este valor está, por supuesto, limitado a un máximo de 1, puesto que, en el presente documento, se presta interés a los valores de atenuación. De manera más precisa:

$$\text{lim}_g(k) = \min\left(\sqrt{\frac{\max(\overline{E_n}, \overline{E_{n'}})}{E_n(k)}}, 1\right)$$

donde la energía media del segmento precedente se aproxima por $\max(\overline{E_n}, \overline{E_{n'}})$.

El valor $\text{lim}_g(k)$ obtenido de este modo sirve de límite inferior en el cálculo final del factor de atenuación del subbloque:

$$g(k) = \max(g(k), \text{lim}_g(k))$$

Los factores de atenuación (o ganancias) $g(k)$ determinados por subbloques se normalizan, a continuación, por una función de normalización aplicada muestra por muestra para evitar unas variaciones bruscas del factor de atenuación en las fronteras de los bloques.

Por ejemplo, en primer lugar, se puede definir la ganancia por muestra como una función constante por trozos:

$$g_{pre}(n) = g(k), n = kL', \dots, (k+1)L'-1$$

donde L' representa la longitud de un subbloque.

La función se normaliza, a continuación, según la siguiente ecuación:

$$g_{pre}(n) := \alpha g_{pre}(n-1) + (1-\alpha)g_{pre}(n), n=0, \dots, L-1$$

con la convención de que $g_{pre}(-1)$ es el último factor de atenuación obtenido para la última muestra del subbloque precedente, α es el coeficiente de normalización, típicamente, $\alpha=0,85$.

Son posibles, igualmente, otras funciones de normalización. Una vez calculados de este modo los factores $g_{pre}(n)$, la atenuación de preeco se hace sobre la señal reconstruida de la trama corriente, $x_{rec}(n)$, multiplicando cada muestra por el factor correspondiente:

$$x_{rec,g}(n) = g_{pre}(n)x_{rec}(n), n = 0, \dots, L-1$$

donde $x_{rec,g}(n)$ es la señal decodificada y postprocesada por la reducción de preeco.

Las figuras 2 y 3 ilustran la implementación del procedimiento de atenuación tal como se describe en la solicitud de patente del estado de la técnica, anteriormente citada, y resumido con anterioridad.

En estos ejemplos, la señal se muestrea a 32 kHz, la longitud de la trama es $L=640$ muestras y cada trama está dividida en 8 subbloques de $K=80$ muestras.

En la parte a) de la figura 2, está representada una trama de una señal original muestreada a 32 kHz. Un ataque (o transición) en la señal está situada en el subbloque que comienza en el índice 320. Esta señal se ha codificado por un codificador por transformada de tipo MDCT de baja velocidad (24 kbit/s).

En la parte b) de la figura 2, se ilustra el resultado de la decodificación sin procesamiento de preeco. Se puede observar el preeco a partir de la muestra 160, en los subbloques precedentes al que contiene el ataque.

La parte c) muestra la evolución del factor de atenuación de preeco (línea continua) obtenido por el procedimiento descrito en la solicitud de patente del estado de la técnica anteriormente citada. La línea en punteado representa el factor antes de normalización. Se señala, en el presente documento, que la posición del ataque se estima alrededor de la muestra 380 (en el bloque delimitado por las muestras 320 y 400).

La parte d) ilustra el resultado de la decodificación después de aplicación del procesamiento de preeco (multiplicación de la señal b) con la señal c)). Se ve que el preeco se ha atenuado bien. La figura 2 muestra, igualmente, que el factor normalizado no vuelve a subir a 1 en el momento del ataque, lo que implica una disminución de la amplitud del ataque. El impacto perceptible de esta disminución es muy escaso, pero, no obstante, puede evitarse. La figura 3 ilustra el mismo ejemplo que la figura 2, en el que, antes de normalización, el valor de factor de atenuación se fuerza a 1 para las algunas muestras del subbloque que precede al subbloque donde se sitúa el ataque. La parte c) de la figura 3 da un ejemplo de una corrección de este tipo.

En este ejemplo, se ha asignado el valor de factor 1 a las 16 últimas muestras del subbloque que precede al ataque, a partir del índice 364. De este modo, la función de normalización incrementa progresivamente el factor para tener un

valor cercano a 1 en el momento del ataque. Entonces, se preserva la amplitud del ataque, como se ilustra en la parte d) de la figura 3, por el contrario, algunas muestras de preeco no se atenúan.

5 En el ejemplo de la figura 3, la reducción de preeco por atenuación no permite reducir el preeco hasta el nivel del ataque, debido a la normalización de la ganancia.

10 En la figura 4 se ilustra otro ejemplo con el mismo ajuste que la de la figura 3. Esta figura representa 2 tramas para mostrar mejor la naturaleza de la señal antes del ataque. En el presente documento, la energía de la señal original antes del ataque es más fuerte (parte a)) que en el caso ilustrado por la figura 3 y la señal antes del ataque es audible (muestras 0 - 850). En la parte b) se puede observar el preeco sobre la señal decodificada sin procesamiento de preeco en la zona 700-850. Según el modo de proceder de limitación de la atenuación explicado con anterioridad, se atenúa la energía de la señal de la zona de preeco hasta la energía media de la señal que precede a la zona de procesamiento. Se observa en la parte c) que el factor de atenuación calculado teniendo en cuenta la limitación de energía está cercano a 1 y que el preeco todavía está presente en la parte d) después de aplicación del procesamiento de preeco (multiplicación de la señal b) con la señal c)), a pesar de la buena puesta a nivel de la señal en la zona de preeco. En efecto, se puede distinguir bien este preeco sobre la forma de onda donde se señala que una componente de alta frecuencia está superpuesta a la señal en esta zona.

20 Esta componente de alta frecuencia es bien audible y molesta y el ataque es menos claro (parte d) figura 4).

25 La explicación de este fenómeno es la siguiente: en el caso de un ataque muy brusco, impulsivo (como se ilustra en la figura 4) el espectro de la señal (en la trama que contiene el ataque) es más bien blanco y, por lo tanto, contiene, igualmente, muchas altas frecuencias. De este modo, el ruido de cuantificación es blanco, igualmente, y está compuesto por altas frecuencias, lo que no es el caso de la señal que precede a la zona de preeco. Por lo tanto, hay un cambio brusco en el espectro de una trama a la otra, que tiene como resultado un preeco audible, a pesar del hecho de que la energía se ha puesto al buen nivel.

30 Este fenómeno está representando de nuevo en las figuras 5a y 5b que muestran respectivamente los espectrogramas de la señal original en 5a, correspondiente a la señal representada en la parte a) de la figura 4 y el espectrograma de la señal con atenuación de preecos según el estado de la técnica, en 5b, correspondiente a la señal representada en la parte d) de la figura 4.

Se señala bien un preeco también audible en la parte enmarcada en la figura 5b.

35 Por lo tanto, existe una necesidad de una tecnología de atenuación mejorada de preecos en la decodificación, que permita atenuar, igualmente, las altas frecuencias no deseadas o preecos parásitos y sin que se transmita por el codificador ninguna información auxiliar.

40 La presente invención mejora la situación del estado de la técnica.

45 Para tal efecto, la presente invención trata sobre un procedimiento de procesamiento de atenuación de preeco en una señal de audio digital generada a partir de una codificación por transformada, en el que, en la decodificación, el procedimiento incluye las siguientes etapas:

- detección de una posición de ataque en la señal decodificada;
- determinación de una zona de preeco que precede a la posición de ataque detectado en la señal decodificada;
- cálculo de factores de atenuaciones por subbloque de la zona de preeco, en función al menos de la trama en la que se ha detectado el ataque y de la trama precedente;
- atenuación de preeco en los subbloques de la zona de preeco por los factores de atenuación correspondientes. El procedimiento es tal que incluye, además:
- la aplicación de un filtrado adaptativo de conformación espectral de la zona de preeco sobre la trama corriente hasta la posición detectada del ataque, siendo el filtrado de respuesta de impulso finita de fase nula de función de transferencia:

$$55 \quad c(n)z^{-1} + (1-2c(n)) + c(n)z$$

60 con $c(n)$ un coeficiente comprendido entre 0 y 0,25. Este tipo de filtrado es de escasa complejidad y permite, además, un procesamiento sin retardo (parándose el procesamiento antes del final de la trama corriente). Gracias a su retardo nulo, el filtrado puede atenuar las altas frecuencias antes del ataque sin modificar el propio ataque. Este tipo de filtrado permite evitar las discontinuidades y permite pasar de una señal no filtrada a una señal filtrada de forma progresiva.

65 De este modo, la conformación espectral aplicada, permite mejorar la atenuación de preeco. El procesamiento permite atenuar las componentes de preeco que podrían subsistir en la implementación de la atenuación de preeco tal como se describe en el estado de la técnica.

Estando el filtrado aplicado hasta la posición detectada del ataque, permite procesar la atenuación del preeco hasta lo más cerca del ataque. Esto compensa, por lo tanto, la desventaja de la reducción de eco por atenuación temporal que está limitada a una zona que no va hasta la posición del ataque (margen de 16 muestras, por ejemplo).

5 Este filtrado no necesita informaciones que provengan del codificador.

Esta técnica de procesamiento de atenuación de preeco puede implementarse con o sin conocimiento de una señal procedente de una decodificación temporal y para la codificación de una señal monofónica o de una señal estereofónica.

10 La adaptación del filtrado permite adaptarse a la señal y no quitar más que las componentes parásitas molestas.

Los diferentes modos particulares de realización mencionados a continuación pueden añadirse de manera independiente o en combinación los unos con los otros, a las etapas del procedimiento definido más arriba.

15 En un modo de realización particular, el procedimiento incluye, además, el cálculo de al menos un parámetro de decisión sobre el filtrado que hay que aplicar a la zona de preeco y la adaptación de los coeficientes del filtrado en función de dicho al menos un parámetro de decisión.

20 De este modo, el procesamiento no se aplica, entonces, más que cuando esto es necesario a un nivel de filtrado adaptado.

En un modo de realización, dicho al menos un parámetro de decisión es una medición de la fuerza del ataque detectado.

25 La fuerza del ataque determina, en efecto, la presencia de componentes de altas frecuencias audibles en la zona de preeco. Cuando el ataque es brusco, el riesgo de tener una componente parásita molesta en la zona de preeco es grande y, entonces, hay que prever el filtrado que hay que implementar según la invención.

30 En un modo de cálculo posible de este parámetro, la medición de la fuerza del ataque detectado es de la forma: $P = \max(EN(k), EN(k+1)/\min(EN(k-1), EN(k-2)))$ con k , el número del subbloque en el que se ha detectado el ataque y $EN(k)$ la energía del $k^{\text{ésimo}}$ subbloque.

Este cálculo es de menor complejidad y permite definir bien la fuerza del ataque detectado.

35 Dicho al menos un parámetro de decisión también puede ser el valor del factor de atenuación en el subbloque que precede al que contiene la posición del ataque.

En efecto, un ataque puede considerarse como brusco si esta atenuación es significativa.

40 En otro modo de realización, dicho al menos un parámetro de decisión está basado en un análisis de reparto espectral de la señal de la zona de preeco y/o de la señal que precede a la zona de preeco.

45 Esto permite, por ejemplo, determinar la importancia de las componentes de altas frecuencias en la señal de preeco y saber, igualmente, si estas componentes de altas frecuencias ya estaban presentes en la señal antes de la zona de preeco.

50 De este modo, en el caso donde unas componentes de altas frecuencias estuvieran ya presentes antes de la zona de preeco, entonces, no es necesario efectuar un filtrado para atenuar estas componentes de altas frecuencias, entonces, la adaptación de los coeficientes de filtrado se efectúa por la puesta a 0 o a un valor cercano a 0 de los coeficientes de filtrado.

De este modo, la adaptación de los coeficientes del filtrado puede efectuarse de forma discreta en función de la comparación de al menos un parámetro de decisión con un umbral predeterminado.

55 Los coeficientes de filtrado pueden tomar unos valores predeterminados según un juego de valores. Siendo el juego de valores más pequeño aquel donde solamente son posibles dos valores, es decir, por ejemplo, la elección entre un filtrado y sin filtrado.

60 En una variante de realización, la adaptación de los coeficientes del filtrado se efectúa de forma continua en función de dicho al menos un parámetro de decisión.

Entonces, la adaptación es más precisa y más progresiva.

65 Según un modo de realización, la etapa de atenuación se efectúa al mismo tiempo que el filtrado de conformación espectral integrando los factores de atenuación en los coeficientes que definen el filtrado.

La presente invención tiene como propósito, igualmente, un dispositivo de procesamiento de atenuación de preecos en una señal de audio digital generada a partir de un codificador por transformada, en el que, el dispositivo asociado a un decodificador comprende:

- 5 - un módulo de detección para detectar una posición de ataque en la señal decodificada;
- un módulo de determinación para determinar una zona de preeco que precede a la posición de ataque detectado en la señal decodificada;
- un módulo cálculo de factores de atenuaciones por subbloque de la zona de preeco, en función al menos de la trama en la que se ha detectado el ataque y de la trama precedente;
- 10 - un módulo de atenuación para atenuar los preecos en los subbloques de la zona de preeco por los factores de atenuación correspondientes. El dispositivo es tal que comprende, además:
- un módulo de filtrado adaptativo para efectuar una conformación espectral de la zona de preeco sobre la trama corriente hasta la posición detectada del ataque, siendo el filtrado de respuesta de impulso finita de fase nula de función de transferencia:

$$c(n)z^{-1} + (1-2c(n)) + c(n)z$$

con $c(n)$ un coeficiente comprendido entre 0 y 0,25.

20 La invención tiene como propósito un decodificador de una señal de audio digital que incluye un dispositivo tal como se ha descrito con anterioridad.

Finalmente, la invención tiene como propósito un programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento de procesamiento de atenuación tal como se describe, cuando estas instrucciones se ejecutan por un procesador.

Finalmente, la invención hace referencia a un soporte de almacenamiento, legible por un procesador, integrado o no en el dispositivo de procesamiento, eventualmente amovible, que memoriza un programa informático que implementa un procedimiento de procesamiento tal como se ha descrito con anterioridad.

30 Otras características y ventajas de la invención se pondrán de manifiesto de manera más clara con la lectura de la siguiente descripción, dada únicamente a título de ejemplo no limitativo y hecha con referencia a los dibujos adjuntos, en los que:

- 35 - la figura 1 descrita con anterioridad ilustra un sistema de codificación-decodificación por transformada según el estado de la técnica;
- la figura 2 descrita con anterioridad ilustra un ejemplo de señal de audio digital para la que se efectúa un método de atenuación según el estado de la técnica;
- la figura 3 descrita con anterioridad ilustra otro ejemplo de señal de audio digital para la que se efectúa un método de atenuación según el estado de la técnica;
- 40 - la figura 4 descrita con anterioridad también ilustra otro ejemplo de señal de audio digital para la que se efectúa un método de atenuación según el estado de la técnica;
- las figuras 5a y 5b ilustran respectivamente el espectrograma de la señal original y el espectrograma de la señal con atenuación de preecos según el estado de la técnica (correspondiente respectivamente a las partes a) y d) de la figura 4);
- 45 - la figura 6 ilustra un dispositivo de procesamiento de atenuación de preecos en un decodificador de señal de audio digital, así como las etapas implementadas por el procedimiento de procesamiento según un modo de realización de la invención;
- la figura 7 ilustra la respuesta de frecuencia de un filtro de conformación espectral implementado según un modo de realización de la invención, en función del parámetro del filtro;
- 50 - la figura 8 ilustra un ejemplo de señal de audio digital para la que se ha implementado el procesamiento según la invención;
- la figura 9 ilustra el espectrograma de la señal correspondiente a la señal d) de la figura 4, para la que se implementa el procesamiento según la invención;
- 55 - la figura 10 ilustra un ejemplo de señal que presenta unas componentes de altas frecuencias en el origen para la que se implementa un método de atenuación de los preecos según el estado de la técnica;
- La figura 11 ilustra la misma señal que la figura 11, que presenta unas componentes de altas frecuencias en el origen para la que se ha implementado el procesamiento según la invención sin la toma en cuenta de un criterio de decisión del nivel de filtrado que hay que aplicar;
- 60 - la figura 12 ilustra un ejemplo material de dispositivo de procesamiento de atenuación según la invención.

Con referencia a la figura 6, se describe un dispositivo 600 de procesamiento de atenuación de preeco. En un modo de realización, este dispositivo implementa un método de atenuación de los preecos en la señal decodificada como, por ejemplo, el descrito en la solicitud de patente FR 08 56248. Implementa, además, un filtrado de conformación espectral de la zona de preeco.

De este modo, el dispositivo 600 incluye un módulo de detección 601 adecuado para implementar una etapa de detección (Detecc.) de la posición de un ataque en una señal de audio decodificada.

5 Un ataque (u *onset* en inglés) es una transición rápida y una variación brusca de la dinámica (o amplitud) de la señal. Se puede designar este tipo de señales por el término más general de "transitoria". En lo que sigue y sin pérdida de generalidad, se utilizarán únicamente los términos de ataque o de transición para designar, igualmente, unas transitorias.

10 En un modo de realización, cada trama de L muestras de la señal decodificada $x_{rec}(n)$ está dividida en K subbloques de longitud L' , con, por ejemplo, $L = 640$ muestras (20 ms) a 32 kHz, $L' = 80$ muestras (2,5 ms) y $K = 8$.

15 Unas ventanas de análisis-síntesis especiales de bajo retardo similares a las descritas en la norma UIT-T G.718 se utilizan para la parte de análisis y para la parte de síntesis de la transformación MDCT. De este modo, la ventana de síntesis MDCT no contiene más que 415 muestras no nulas contrariamente a las 640 muestras en el caso de utilización de las ventanas sinusoidales convencionales. En una variante de este modo de realización, pueden utilizarse otras ventanas de análisis/síntesis o pueden utilizarse unas conmutaciones entre ventanas largas y cortas.

20 Por otra parte, se utiliza la memoria MDCT $x_{MDCT}(n)$ que da una versión con repliegue temporal ("folding" en inglés) de la señal futura. Esta memoria también está dividida en subbloques de longitud L' y no se retienen - en función de la ventana MDCT utilizada - más que los K' primeros subbloques, donde K' depende de la ventana utilizada - por ejemplo, $K' = 4$ para una ventana sinusoidal. En efecto, la figura 1 muestra que el preeco influye en la trama que precede a aquella donde se sitúa el ataque, y es deseable detectar un ataque en la trama futura que está en parte contenida en la memoria MDCT.

25 La reducción de preecos depende, en el presente documento, de varios parámetros:

- La señal decodificada en la trama corriente (que contiene potencialmente unos preecos) de longitud L,
- La memoria de la transformación inversa MDCT que corresponde a la señal parcialmente decodificada en la trama siguiente antes de adición-recubrimiento.
- 30 ◦ El nivel medio de energía en la trama (o semitrama) precedente.

35 Se puede anotar que la señal contenida en la memoria MDCT consta de un repliegue temporal (que está compensado cuando se recibe la trama siguiente). Como se explica más abajo, la memoria MDCT sirve, en el presente documento, sustancialmente para estimar la energía por subbloques de la señal en la trama siguiente (futura) y se considera que esta estimación es suficientemente precisa para las necesidades de la detección y reducción de preeco cuando está realizada con la memoria MDCT disponible en la trama corriente, en lugar de la señal completamente decodificada en la trama futura.

40 La trama corriente y la memoria MDCT pueden verse como unas señales concatenadas que forman una señal de longitud $(K + K')L'$ recortada en $(K+K')$ subbloques consecutivos. En estas condiciones, se define la energía en el k-ésimo subbloque como:

$$En(k) = \sum_{n=kL'}^{(k+1)L'-1} x_{rec}(n)^2, \quad k=0, \dots, K-1$$

45 cuando el k-ésimo subbloque se sitúa en la trama corriente y, como:

$$En(k) = \sum_{n=(k-K)L'}^{(k-K+1)L'-1} x_{MDCT}(n)^2, \quad k=K, \dots, K+K'$$

50 cuando el subbloque está en la memoria MDCT (que representa la señal disponible para la trama futura). La energía media de los subbloques en la trama corriente se obtiene, por lo tanto, como:

$$\overline{En} = \frac{1}{K} \sum_{k=0}^{K-1} En(k)$$

Se define, igualmente, la energía media de los subbloques en la segunda parte de la trama corriente como:

55

$$\overline{En}' = \frac{2}{K} \sum_{k=K/2}^{K-1} En(k)$$

Una transición asociada a un preeco se detecta si la relación $R(k) = \frac{\max_{k=0, K+K'}(En(k))}{En(k)}$ rebasa un umbral predefinido, en uno de los subbloques considerados. Son posibles otros criterios de detección de preeco sin cambiar la naturaleza de la invención.

Por otra parte, se considera que la posición del ataque está definida como

$$pos = \min \left(L', \left(\arg \max_{k=0, K+K'} (En(k)) \right), L \right)$$

donde la limitación a L asegura que la memoria MDCT no se ha modificado nunca. Son posibles, igualmente, otros métodos de estimación más precisa de la posición del ataque.

En unas variantes de realización con conmutación de las ventanas, pueden utilizarse otros métodos que dan la posición del ataque con una precisión que va de la escala de un subbloque hasta una posición con una muestra de aproximación.

El dispositivo 600 incluye, igualmente, un módulo 602 de determinación que implementa una etapa de determinación (ZPE) de una zona de preeco que precede a la posición de ataque detectado.

Las energías $En(k)$ están concatenadas en orden cronológico, con, en primer lugar, la envolvente temporal de la señal decodificada, luego, la envolvente de la señal de la trama siguiente estimada a partir de la memoria de la transformada MDCT. En función de esta envolvente temporal concatenada y de las energías medias \overline{En} y \overline{En}' de la trama precedente, se detecta la presencia de preeco si la relación $R(k)$ es suficientemente fuerte.

Los subbloques en los que se ha detectado un preeco constituyen, de este modo, una zona de preeco, que, en general, cubre las muestras $n = 0, \dots, pos-1$, esto es, del inicio de la trama corriente a la posición del ataque (pos).

En unas variantes de realización, la zona de preeco no comienza necesariamente al inicio de la trama y puede hacer intervenir una estimación de la longitud del preeco. Si se utiliza una conmutación de ventanas, la zona de preeco deberá estar definida para tomar en cuenta las ventanas utilizadas.

Un módulo 603 del dispositivo 600 implementa una etapa de cálculo de factores de atenuación por subbloques de la zona de preeco determinada, en función de la trama en la que se ha detectado el ataque y de la trama precedente.

De conformidad con la descripción de la solicitud de patente FR 08 56248, las atenuaciones $g(k)$ se estiman por subbloque.

El factor de atenuación por subbloque $g(k)$ se calcula, por ejemplo, en función de la relación $R(k)$ entre la energía del subbloque de energía más fuerte y la energía del k-ésimo subbloque en cuestión:

$$g(k) = f(R(k))$$

donde f es una función decreciente de valores entre 0 y 1. Son posibles otras definiciones del factor $g(k)$, por ejemplo, en función de $En(k)$ y de $En(k-1)$.

Si la variación de la energía con respecto a la energía máxima es escasa, entonces, no es necesaria ninguna atenuación. El factor está fijado, entonces, a un valor de atenuación que inhibe la atenuación, es decir, 1. Si no, el factor de atenuación está comprendido entre 0 y 1.

Estas atenuaciones son limitadas en función de la energía media de la trama precedente.

Para el subbloque que hay que procesar, se puede calcular el valor límite del factor $\lim_g(k)$, con el fin de obtener exactamente la misma energía que la energía media del segmento que precede al subbloque que hay que procesar. Este valor está, por supuesto, limitado a un máximo de 1, puesto que, en el presente documento, se presta interés a los valores de atenuación. De manera más precisa:

$$\lim_g(k) = \min\left(\sqrt{\frac{\max(E\bar{n}, E\bar{n}')}{E\bar{n}(k)}}, 1\right)$$

El valor $\lim_g(k)$ obtenido de este modo sirve de límite inferior en el cálculo final del factor de atenuación del subbloque:

$$g(k) = \max(g(k), \lim_g(k))$$

Los factores de atenuación $g(k)$ determinados por subbloques se normalizan, a continuación, por una función de normalización aplicada muestra por muestra para evitar unas variaciones bruscas del factor de atenuación en las fronteras de los bloques.

La ganancia por muestra se define, en primer lugar, como una función constante por trozos:

$$g_{pre}(n) = g(k), \quad n = kL', \dots, (k+1)L'-1$$

La función de normalización está definida, por ejemplo, por las siguientes ecuaciones:

$$g_{pre}(n) := \alpha g_{pre}(n-1) + (1-\alpha)g_{pre}(n), \quad n = 0, \dots, L-1$$

con la convención de que $g_{pre}(-1)$ es el último factor de atenuación obtenido para la última muestra del subbloque precedente, α es el coeficiente de normalización, típicamente, $\alpha=0,85$.

Son posibles otras funciones de normalización.

El módulo 604 del dispositivo 600 de la figura 6 implementa la atenuación (At.) en los subbloques de la zona de preeco, por los factores de atenuación obtenidos.

De este modo, una vez calculados los factores $g_{pre}(n)$, la atenuación de preeco se hace sobre la señal reconstruida de la trama corriente, $x_{rec}(n)$, multiplicando cada muestra por el factor correspondiente:

$$x_{rec,g}(n) = g_{pre}(n)x_{rec}(n), \quad n = 0, \dots, L-1$$

donde $x_{rec,g}(n)$ es la señal decodificada y postprocesada para la reducción de preeco.

El dispositivo 600 incluye un módulo de filtrado 606 adecuado para efectuar la etapa (F) de aplicación de un filtrado de conformación espectral de la zona de preeco sobre la trama corriente de la señal decodificada, hasta la posición detectada del ataque.

Típicamente, el filtro de conformación espectral utilizado es un filtro lineal. Como la operación de multiplicación por una ganancia es, igualmente, una operación lineal, su orden puede invertirse: se puede, igualmente, hacer, en primer lugar, el filtrado de conformación espectral de la zona de preeco, luego, la atenuación de preeco multiplicando cada muestra de la zona de preeco por el factor correspondiente.

El filtro utilizado para atenuar las altas frecuencias en la zona de preeco es un filtro FIR (filtro de respuesta de impulso finita) de 3 coeficientes y de fase nula de función de transferencia $c(n)z^{-1} + (1-2c(n)) + c(n)z$ con $c(n)$ un valor comprendido entre 0 y 0,25, donde $[c(n), 1-2c(n), c(n)]$ son los coeficientes del filtro de conformación espectral; este filtro se implementa con la ecuación de las diferencias:

$$x_{rec,f}(n) = c(n)x_{rec,g}(n-1) + (1-2c(n))x_{rec,g}(n) + c(n)x_{rec,g}(n+1)$$

con, por ejemplo, $c(n) = 0,25$ sobre la zona $n = 5, \dots, pos - 5$.

La respuesta de frecuencia de este filtro se ilustra en la figura 7, en función del coeficiente $c(n)$, para $c(n) = 0,05, 0,1, 0,15, 0,2$ y $0,25$. La motivación para utilizar este filtro es su escasa complejidad, su fase nula y, por lo tanto, su retardo nulo (posible, ya que el procesamiento se para antes del final de trama corriente), pero, igualmente, su respuesta de frecuencia que corresponde bien a las características de paso bajo deseadas para este filtro.

La aplicación de este filtro puede compensar el hecho de que la atenuación temporal del preeco está típicamente limitada a una zona que no va hasta la posición del ataque (con un margen de, por ejemplo, 16 muestras), mientras que el filtrado de conformación espectral tal como se define por la función de transferencia $c(n)z^{-1} + (1-2c(n)) + c(n)z$ puede aplicarse hasta la posición del ataque, con eventualmente algunas muestras de interpolación de los coeficientes del filtro.

Para pasar de una señal no filtrada a una señal filtrada y evitar unas discontinuidades, es preferible introducir el filtrado de forma progresiva. El filtro FIR propuesto permite pasar fácilmente de manera suave del campo no filtrado hacia el

campo filtrado y viceversa, por interpolación o variación lenta de sus coeficientes. Por ejemplo, si la posición del ataque es $pos = 16$, el filtrado de las 16 muestras en la zona de preeco $n = 0, \dots, pos^{-1}$ puede efectuarse de la siguiente forma:

$$x_{rec,f}(0) = x_{rec}(0)$$

$$x_{rec,f}(1) = 0,1x_{rec}(0) + 0,8x_{rec}(1) + 0,1x_{rec}(2)$$

$$x_{rec,f}(2) = 0,1x_{rec}(1) + 0,8x_{rec}(2) + 0,1x_{rec}(3)$$

$$x_{rec,f}(3) = 0,15x_{rec}(2) + 0,7x_{rec}(3) + 0,15x_{rec}(4)$$

$$x_{rec,f}(4) = 0,2x_{rec}(3) + 0,6x_{rec}(4) + 0,2x_{rec}(5) =$$

$$x_{rec,f}(n) = 0,25x_{rec}(n-1) + 0,5x_{rec}(n) + 0,25x_{rec}(n+1), n = 5, \dots, 11$$

$$x_{rec,f}(12) = 0,2x_{rec}(11) + 0,6x_{rec}(12) + 0,2x_{rec}(13)$$

$$x_{rec,f}(13) = 0,15x_{rec}(12) + 0,7x_{rec}(13) + 0,15x_{rec}(14)$$

$$x_{rec,f}(14) = 0,1x_{rec}(13) + 0,8x_{rec}(14) + 0,1x_{rec}(15)$$

$$x_{rec,f}(15) = 0,05x_{rec}(14) + 0,9x_{rec}(15) + 0,05x_{rec}(16)$$

Se observa, que, gracias a su retardo nulo, el filtro $c(n)z^{-1} + (1-2c(n)) + c(n)z$ puede atenuar las altas frecuencias antes del ataque sin modificar el propio ataque.

Un ejemplo de señal de audio digital, para la que se efectúa el procesamiento tal como se describe en el presente documento, se ilustra en la parte d) de la figura 8. Las partes a), b) y c) de esta figura recogen las mismas señales que las descritas con referencia a la figura 4 con anterioridad. La parte d) difiere por la implementación del filtrado según la invención. De este modo, se puede señalar que la componente de alta frecuencia molesta está fuertemente disminuida, con lo que la señal decodificada después de filtrado tiene una mejor calidad que la descrita en la parte d) de la figura 4.

El espectrograma que representa esta señal filtrada está representado en la figura 9. Se observa bien con respecto a la figura 5b que representa la misma señal sin filtrado de conformación, la atenuación de las altas frecuencias molestas antes del ataque. Entonces, el ataque se hace más claro en la decodificación.

Por supuesto, pueden considerarse otros tipos de filtro de conformación espectral para sustituir al filtro $c(n)z^{-1} + (1-2c(n)) + c(n)z$. Por ejemplo, es posible utilizar un filtro FIR de orden diferente o con unos coeficientes diferentes. Alternativamente, el filtro de conformación espectral puede ser de respuesta de impulso infinita (IIR). Además, la conformación espectral puede ser diferente de un filtrado de paso bajo, por ejemplo, podría implementarse un filtro de pasa banda. Estos modos de realización no forman parte de la invención que está definida por las reivindicaciones.

Un filtro de orden 1, de la forma $c(n)z^{-1} + (1-c(n))$ puede utilizarse, igualmente, en un modo de realización que no forma parte de la invención.

En un modo de realización particular, el filtrado implementado según el procedimiento descrito, es un filtrado adaptativo. De este modo, puede adaptarse a las características de la señal de audio decodificada.

En este modo de realización, se implementa una etapa de cálculo de un parámetro (P) de decisión sobre el filtrado que hay que aplicar a la zona de preeco en el módulo de cálculo 605 de la figura 6.

En efecto, existen unos casos como el que se ilustra, por ejemplo, en la figura 10 donde es preferible no aplicar un filtrado de este tipo en la zona de preeco.

En efecto, en el caso, más raro, ilustrado en la figura 10, parte a) las altas frecuencias ya están presentes en la señal que hay que codificar. En este caso, la atenuación de las altas frecuencias podría causar una degradación audible que, por lo tanto, hay que evitar. En este ejemplo de señal, se observa que el ataque es menos brusco que en los ejemplos precedentes.

Entonces, es interesante determinar al menos un parámetro que permita decidir si hay que conformar espectralmente la zona de la señal que contiene un preeco, atenuando (o no) las altas frecuencias.

En un ejemplo de realización, este parámetro de decisión es representativo de la presencia de componentes de altas frecuencias en la zona de preeco.

Este parámetro puede ser, por ejemplo, una medición de la fuerza del ataque (brusco o no). Si el ataque está localizado en el subbloque número k , el parámetro puede calcularse como:

$$P = \frac{\max(En(k), En(k+1))}{\min(En(k-1), En(k-2))}$$

5 donde k el número del subbloque y $En(k)$ la energía en el k -ésimo subbloque.

Según un ajuste experimental, en este ejemplo de realización, $P \geq 32$ indica un ataque brusco (muy impulsivo).

10 La medición de fuerza del ataque puede completarse teniendo en cuenta, igualmente, la atenuación determinada para el subbloque que precede al ataque $g(k-1)$. Un ataque puede considerarse como brusco si esta atenuación es significativa, por ejemplo, si $g(k-1) \leq 0,5$. Esto muestra que la energía en la zona de preeco está considerablemente aumentada (más que duplicada) debido al preeco, lo que señala, igualmente, un ataque brusco.

15 Si $P < 32$ y $g(k-1) > 0,5$, donde k es el índice del subbloque que contiene el inicio del ataque, no es necesario el filtrado. En efecto, si $g(k-1) > 0,5$, $\lim_g(k) > 0,5$, lo que significa que la zona de preeco tiene una energía comparable con la de la trama precedente y como el ataque que genera el preeco no es brusco, el riesgo de tener una componente parásita molesta es escaso.

20 De este modo, en este modo de realización con las condiciones ($P < 32$ y $g(k-1) > 0,5$), no se hará ningún filtrado sobre la zona de preeco.

En los otros casos ($g(k-1) \leq 0,5$ o $P > 32$), se aplica el filtro de conformación espectral, según la invención, del inicio de la trama corriente hasta la posición pos de posición del ataque.

25 En el ejemplo de realización descrito más arriba, la conformación espectral de la zona de preeco por filtrado según la invención es adaptativa en función del parámetro P y de los valores de atenuación. De este modo, el filtrado se aplica o bien con unos coeficientes $[0,25, 0,5, 0,25]$, o bien se desactiva con unos coeficientes $[0, 1, 0]$.

30 La adaptación de los coeficientes de filtrado se efectúa, entonces, de forma discreta limitada a un juego de valores predefinido.

La adaptación de los coeficientes de filtrado (que permite adaptar el nivel de atenuación de las altas frecuencias) está determinado, por lo tanto, por unos parámetros de decisión que miden la fuerza del ataque como los parámetros P y $g(k-1)$.

35 Se trata, en este caso, de una adaptación de los coeficientes del filtro de forma discreta según dos juegos de valores posibles ($[0,25, 0,5, 0,25]$ o $[0, 1, 0]$). Se puede anotar que el juego de coeficientes $[0, 1, 0]$ corresponde a una desactivación del filtrado.

40 Una transición progresiva entre estos dos filtros puede efectuarse utilizando, igualmente, por ejemplo, los filtros intermedios de coeficiente $[0,05, 0,9, 0,05]$, $[0,1, 0,8, 0,1]$, $[0,15, 0,7, 0,15]$ y $[0,2, 0,6, 0,2]$.

45 Se trata, en este caso, de una adaptación de los coeficientes del filtro de forma discreta según varios juegos de valores posibles, si se tiene en cuenta la variación lenta (o interpolación).

En unas variantes de realizaciones, pueden utilizarse otros métodos de interpolación.

50 Por ejemplo, el filtrado puede ser todavía más finamente adaptativo con $c(n) = f(P)$, por ejemplo, utilizando un filtro intermedio con $c(n) = [0,15, 0,7, 0,15]$ si $16 < P < 32$. $c(n)$ puede calcularse, igualmente, de forma continua en función de P , por ejemplo, con la fórmula $c(n) = \frac{\arctan(P/10)}{2\pi}$.

Se trata, en este caso, de una adaptación de los coeficientes del filtro de forma continua según unos valores posibles donde $c(n)$ está en el intervalo $[0, 0,25]$.

55 Pueden utilizarse, igualmente, otros parámetros de decisión en la decisión de la elección y de la adaptación del filtro, como, por ejemplo, la tasa de paso a cero ("zero crossing rate" en inglés) de la señal decodificada de la zona de preeco de la trama corriente y/o de la trama precedente. La tasa de paso por cero puede calcularse de la siguiente forma si se considera la zona $n = 0, \dots, L-1$ a título de ejemplo:

$$60 \quad zc = \frac{1}{2} \sum_{n=0}^{L-1} \left| \text{señ} [x_{rec,g}(n-1)] - \text{señ} [x_{rec,g}(n)] \right|$$

donde

$$\text{señ}(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$$

5 En efecto, una tasa elevada de paso a cero z_c en la trama precedente (por lo tanto, sin preeco) señala la presencia de altas frecuencias en la señal. En este caso, por ejemplo, cuando $z_c > L/2$ sobre la trama precedente, es preferible no aplicar el filtrado

$$c(n)z^{-1} + (1-2c(n)) + c(n)z.$$

10 Con el fin de eliminar el sesgo de la componente continua, es posible, igualmente, un prefiltrado de la señal decodificada antes de cálculo de la tasa de paso por cero o bien puede utilizarse el número de paso por cero de la derivada estimada $x_{rec,g}(n) - x_{rec,g}(n-1)$.

15 En una variante, puede hacerse, igualmente, un análisis espectral de la señal para ayudar a la decisión. Por ejemplo, puede explotarse la envolvente espectral en el campo MDCT procedente de la codificación/decodificación MDCT en la elección del filtro que hay que utilizar, sin embargo, esta variante supone que las ventanas de análisis/síntesis MDCT son suficientemente cortas para que las estadísticas locales de la señal antes del ataque permanezcan estables sobre la longitud de una ventana.

20 De manera alternativa, se podrá filtrar la señal en la zona de preeco y en la trama pasada por un filtro complementario de paso alto como $-c(n)z^{-1} + (1-2c(n)) - c(n)z$, con, por ejemplo, $c(n) = 0,25$ y, a continuación, se elegirá el valor de $c(n)$ de modo que la energía media de las señales filtradas en la zona de preeco y sobre la trama pasada sean lo más cercanas posibles; la elección de $c(n)$ podrá hacerse sobre un juego limitado de valores posibles mostrados en la figura 7 o a partir de la ratio de energía (o de una cantidad equivalente, como la raíz cuadrada de la energía) de la señal después de filtrado de paso alto en la zona de preeco y en la trama pasada.

25 Debe anotarse que el filtrado de paso alto puede implementarse, igualmente, de forma alternativa calculando la diferencia entre la señal $x_{rec,g}(n)$ y la señal filtrada por el filtro de paso bajo $c(n)z^{-1} + (1-2c(n)) + c(n)z$ cuando $c(n) = 0,25$.

30 En otra variante, cuando el filtrado de conformación es de tipo $c(n)z^{-1} + (1-c(n))$, se podrá fijar el valor de $c(n)$ en función del coeficiente de predicción $-r(1)/r(0)$ procedente de un análisis por predicción lineal (LPC para "Linear Predictive Coding" en inglés) del orden 1 de la señal en la zona de preeco y de la señal en la trama pasada.

35 En todas estas últimas variantes (tasa de paso a cero, envolvente espectral MDCT, filtrado de paso alto, análisis LPC), el parámetro de decisión sobre el filtrado que hay que aplicar a la zona de preeco está basado en un análisis de reparto espectral de la señal de la zona de preeco y/o de la señal que precede de la zona de preeco; si la señal que precede a la zona de preeco ya contiene muchas altas frecuencias o si la cantidad de las altas frecuencias de la señal en la zona de preeco y de la señal que precede a la zona de preeco es sustancialmente idéntica, no es necesario el filtrado según la invención y puede incluso causar una ligera degradación. En estos casos, hay que desactivar o atenuar el filtrado según la invención fijando $c(n)$ a 0 o a un escaso valor cercano a 0.

En una variante de la invención, el orden entre la etapa de atenuación y de filtrado podrá invertirse.

45 En efecto, puede que el filtrado (F) de conformación espectral se haga antes de la atenuación (At.). De este modo, después de haber efectuado el filtrado adaptativo de las muestras de la zona de preeco de la señal reconstruida de la trama corriente, entonces, se ponderan estas muestras multiplicando cada muestra por el factor de atenuación correspondiente calculado con anterioridad:

$$50 \quad x_{rec,f,g}(n) = g_{pre}(n)x_{rec,f}(n), \quad n=0, \dots, L-1$$

La atenuación de las amplitudes puede, igualmente, combinarse (o integrarse) definiendo un juego de coeficientes de filtro "conjunto", por ejemplo, si para la muestra n el filtro tiene unos coeficientes $[c(n), 1-2c(n), c(n)]$ y el factor de atenuación es $g(n)$, se puede utilizar directamente el filtro $[g_{pre}(n)c(n), g_{pre}(n)2g_{pre}(n)c(n), g_{pre}(n)c(n)]$.

55 La figura 11 ilustra la ventaja de hacer el filtrado adaptativo. Recoge las mismas señales en las partes a), b) y c) que la figura 10 e ilustra el hecho de que la implementación del filtrado no adaptativo representado en la parte d), modifica inútilmente la señal en el caso donde las componentes de altas frecuencias ya están presentes en la señal que hay que codificar. Se observa que a partir de la muestra 640 las altas frecuencias se atenúan inútilmente, lo que podría plantear una ligera degradación de calidad. La utilización de un filtrado adaptativo como se ha descrito más arriba permite inhibir o atenuar el filtrado en estas condiciones, no quitar unas altas frecuencias ya presentes en la señal que hay que codificar y, de este modo, evitar una eventual degradación debida al filtrado.

60 Para volver a la figura 6, el dispositivo de procesamiento de atenuación 600 tal como se describe está comprendido, en el presente documento, en un decodificador que incluye un módulo 610 de cuantificación inversa (C^{-1}) que recibe

una señal S, un módulo 620 de transformada inversa (MDCT⁻¹), un módulo 630 de reconstrucción de la señal por adición/recubrimiento (adi/rec) como se ha descrito con referencia a la figura 1 y que suministra una señal reconstruida al dispositivo de procesamiento de atenuación según la invención.

5 A la salida del dispositivo 600, se proporciona una señal procesada Sa en la que se ha efectuado una atenuación de preeco. El procesamiento efectuado ha permitido mejorar la atenuación de preeco por la atenuación, llegado el caso, de las componentes de altas frecuencias, en la zona de preeco.

10 En este momento, se describe un ejemplo de realización de un dispositivo de procesamiento de atenuación según la invención con referencia a la figura 12.

15 Materialmente, este dispositivo 100 en el sentido de la invención incluye típicamente, un procesador μ P que coopera con un bloque de memoria BM que consta de una memoria de almacenamiento y/o de trabajo, así como de una memoria tampón MEM anteriormente citada en calidad de medio para memorizar cualesquiera datos necesarios para la implementación del procedimiento de procesamiento de atenuación tal como se ha descrito con referencia a la figura 6. Este dispositivo recibe a la entrada unas tramas sucesivas de la señal digital Se y suministra la señal Sa reconstruida con atenuación de preeco y filtrado de conformación espectral, llegado el caso.

20 El bloque de memoria BM puede incluir un programa informático que incluye las instrucciones de código para la implementación de las etapas del procedimiento según la invención cuando estas instrucciones se ejecutan por un procesador μ P del dispositivo y, en concreto, una etapa de detección de una posición de ataque en la señal decodificada, de determinación de una zona de preeco que precede a la posición de ataque detectado en la señal decodificada, de cálculo de factores de atenuaciones por subbloque de la zona de preeco, en función de la trama en la que se ha detectado el ataque y de la trama precedente, de atenuación de preeco en los subbloques de la zona de preeco por los factores de atenuación correspondientes y, además, una etapa de aplicación de un filtrado de conformación espectral de la zona de preeco sobre la trama corriente hasta la posición detectada del ataque, siendo el filtrado de respuesta de impulso finita de fase nula de función de transferencia:

$$c(n)z^{-1} + (1-2c(n)) + c(n)z$$

30 con c(n) un coeficiente comprendido entre 0 y 0,25. La figura 6 puede ilustrar el algoritmo de un programa informático de este tipo.

35 Este dispositivo de atenuación según la invención puede ser independiente o estar integrado en un decodificador de señal digital.

REIVINDICACIONES

1. Procedimiento de procesamiento de atenuación de preeco en una señal de audio digital generada a partir de una codificación por transformada, en el que, en la decodificación, el procedimiento incluye las siguientes etapas:

- 5 - detección (601, Detecc.) de una posición de ataque en la señal decodificada;
- determinación (602, ZPE) de una zona de preeco que precede a la posición de ataque detectado en la señal decodificada;
- 10 - cálculo (603, F. At.) de factores de atenuaciones por subbloque de la zona de preeco, en función al menos de la trama en la que se ha detectado el ataque y de la trama precedente;
- atenuación (604, At.) de preeco en los subbloques de la zona de preeco por los factores de atenuación correspondientes;

estando el procedimiento caracterizado por que incluye, además:

- 15 - la aplicación de un filtrado adaptativo (606, F) de conformación espectral de la zona de preeco sobre la trama corriente hasta la posición detectada del ataque, siendo el filtrado de respuesta de impulso finita de fase nula de función de transferencia:

$$20 \quad c(n)z^{-1} + (1-2c(n)) + c(n)z$$

con $c(n)$ un coeficiente comprendido entre 0 y 0,25.

2. Procedimiento según la reivindicación 1, caracterizado por que incluye, además, el cálculo de al menos un parámetro de decisión sobre el filtrado que hay que aplicar a la zona de preeco y la adaptación de los coeficientes del filtrado en función de dicho al menos un parámetro de decisión.

3. Procedimiento según la reivindicación 2, caracterizado por que dicho al menos un parámetro de decisión es una medición de la fuerza del ataque detectado.

4. Procedimiento según la reivindicación 2, caracterizado por que dicho al menos un parámetro de decisión es el valor del factor de atenuación en el subbloque que precede al que contiene la posición del ataque.

5. Procedimiento según la reivindicación 2, caracterizado por que dicho al menos un parámetro de decisión está basado en un análisis de reparto espectral de la señal de la zona de preeco y/o de la señal que precede a la zona de preeco.

6. Procedimiento según la reivindicación 3, caracterizado por que la medición de la fuerza del ataque detectado es de la forma: $P = \max(EN(k), EN(k+1)) / \min(EN(k-1), EN(k-2))$ con k , el número del subbloque en el que se ha detectado el ataque y $EN(k)$ la energía del $k^{\text{ésimo}}$ subbloque.

7. Procedimiento según la reivindicación 2, caracterizado por que la adaptación de los coeficientes del filtrado se efectúa de forma discreta en función de la comparación de al menos un parámetro de decisión con un umbral predeterminado.

8. Procedimiento según la reivindicación 2, caracterizado por que la adaptación de los coeficientes del filtrado se efectúa de forma continua en función de dicho al menos un parámetro de decisión.

9. Procedimiento según la reivindicación 1, caracterizado por que la etapa de atenuación se efectúa al mismo tiempo que el filtrado de conformación espectral integrando los factores de atenuación en los coeficientes que definen el filtrado.

10. Dispositivo de procesamiento de atenuación de preeco en una señal de audio digital generada a partir de un codificador por transformada, en el que, el dispositivo asociado a un decodificador comprende:

- 55 - un módulo de detección (601) para detectar una posición de ataque en la señal decodificada;
- un módulo de determinación (602) para determinar una zona de preeco que precede a la posición de ataque detectado en la señal decodificada;
- 60 - un módulo cálculo (603) de factores de atenuaciones por subbloque de la zona de preeco, en función al menos de la trama en la que se ha detectado el ataque y de la trama precedente;
- un módulo de atenuación (604) para atenuar los preecos en los subbloques de la zona de preeco por los factores de atenuación correspondientes;

estando el dispositivo caracterizado como que comprende, además:

65

- un módulo de filtrado (606) adaptativo para efectuar una conformación espectral de la zona de preeco sobre la trama corriente hasta la posición detectada del ataque, siendo el filtrado de respuesta de impulso finita de fase nula de función de transferencia:

5

$$c(n)z^{-1} + (1-2c(n)) + c(n)z$$

con $c(n)$ un coeficiente comprendido entre 0 y 0,25.

10

11. Decodificador de una señal de audio digital que incluye un dispositivo según la reivindicación 10.

12. Programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento según una de las reivindicaciones 1 a 9, cuando estas instrucciones se ejecutan por un procesador.

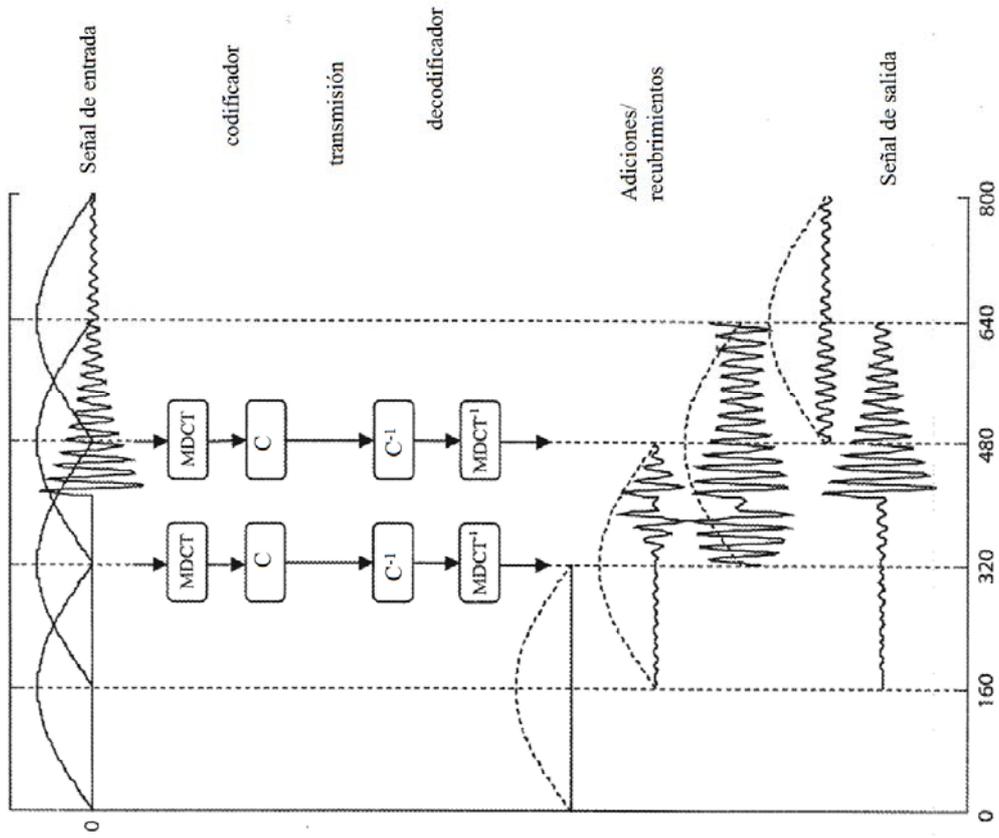


FIG.1

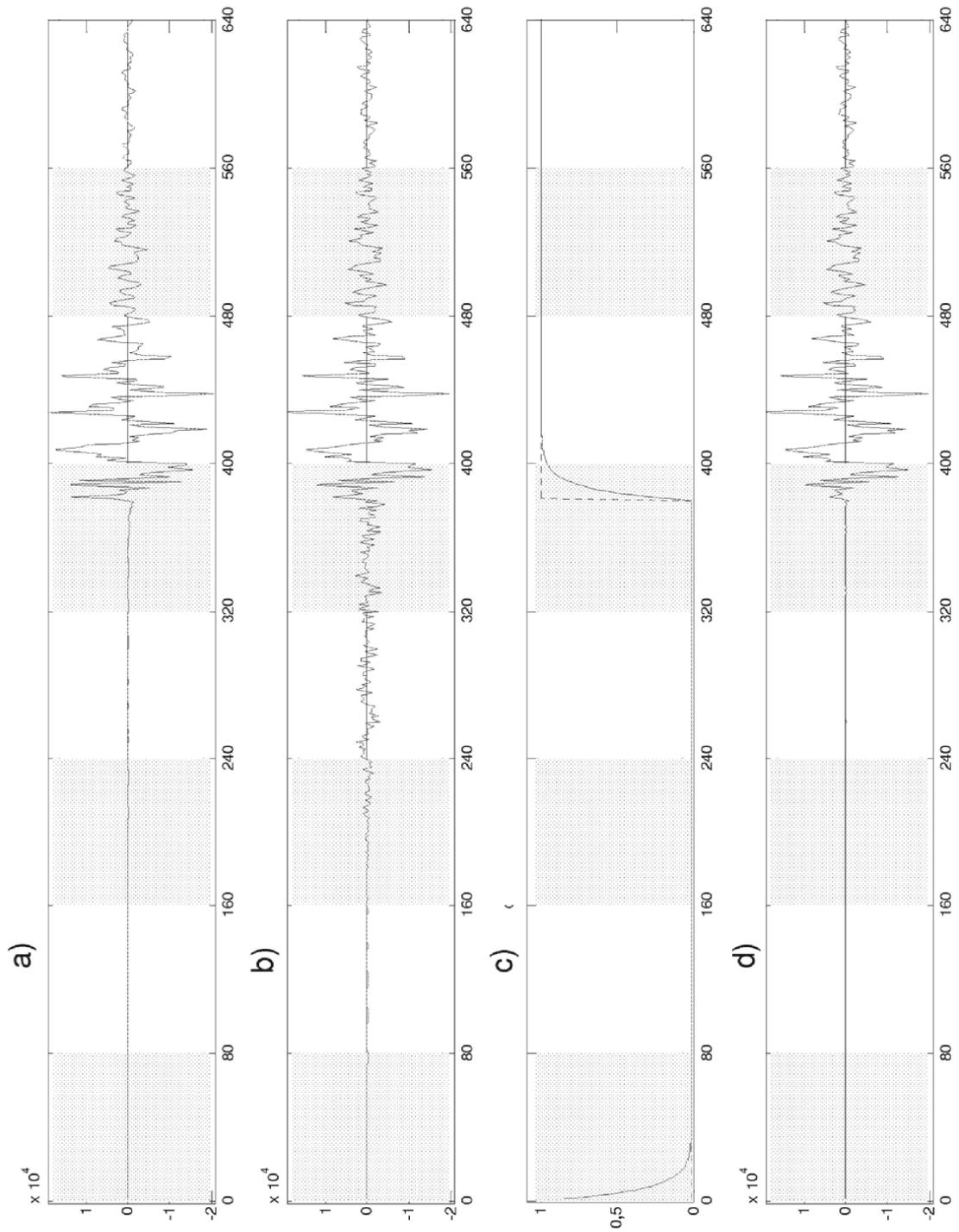


FIG.2

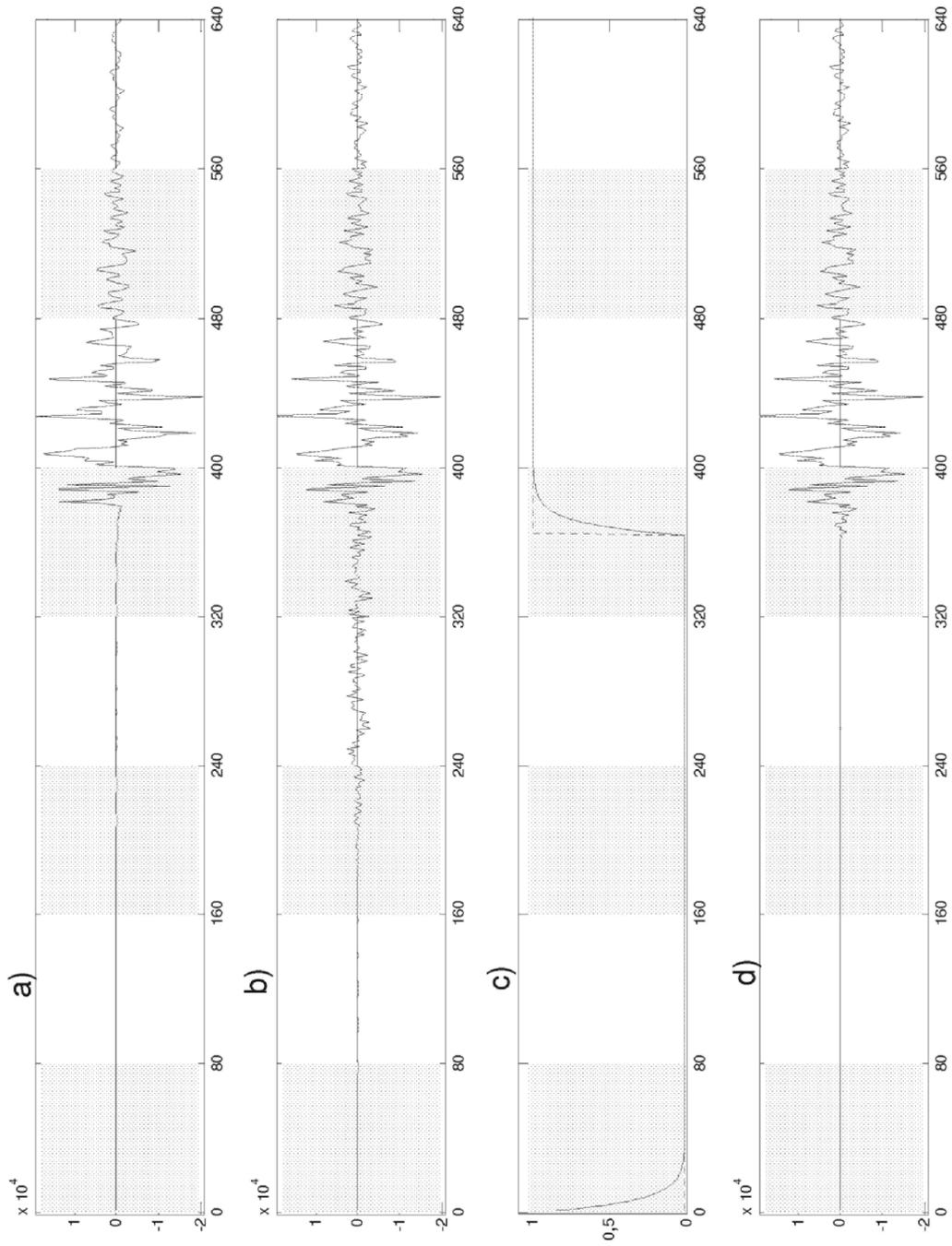


FIG.3

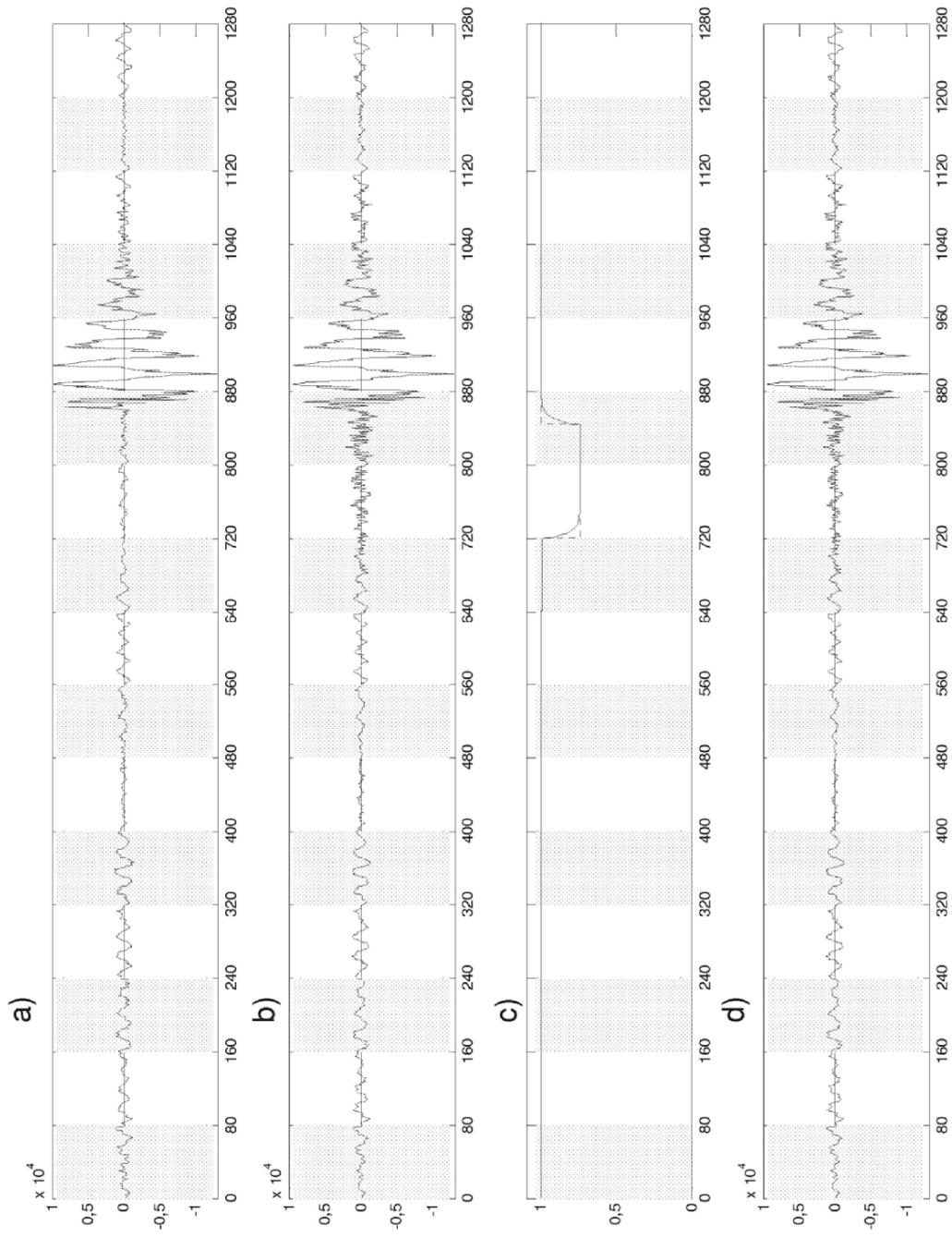


FIG.4

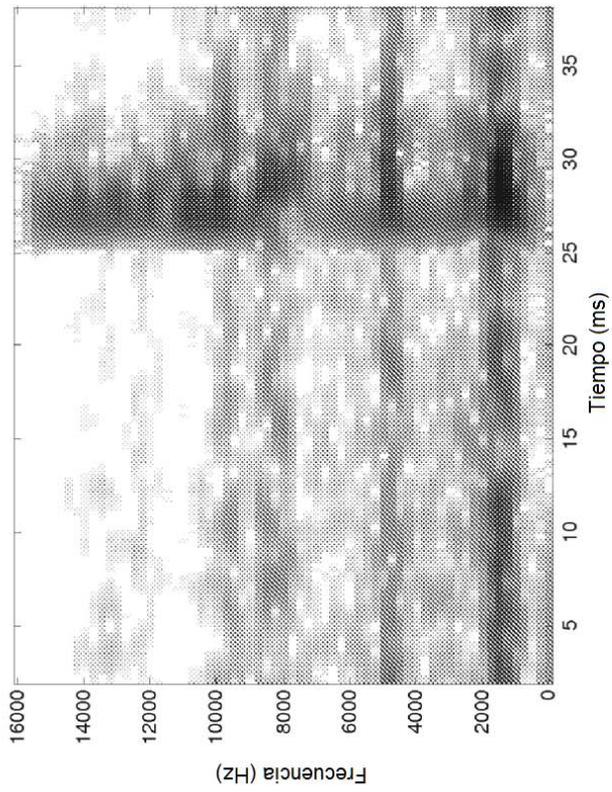


FIG.5a

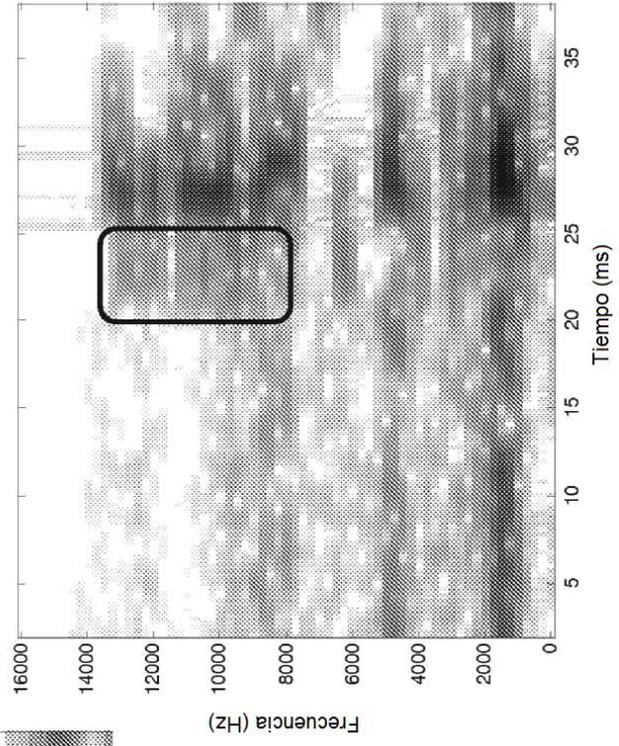


FIG.5b

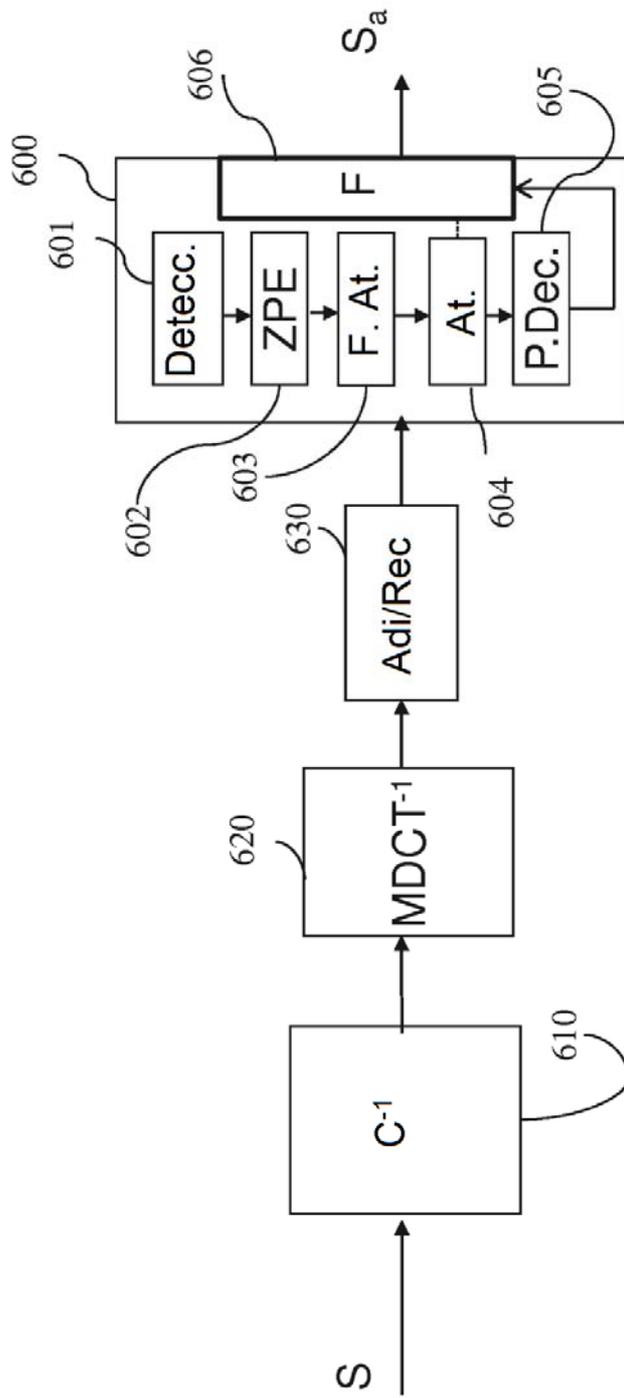


FIG. 6

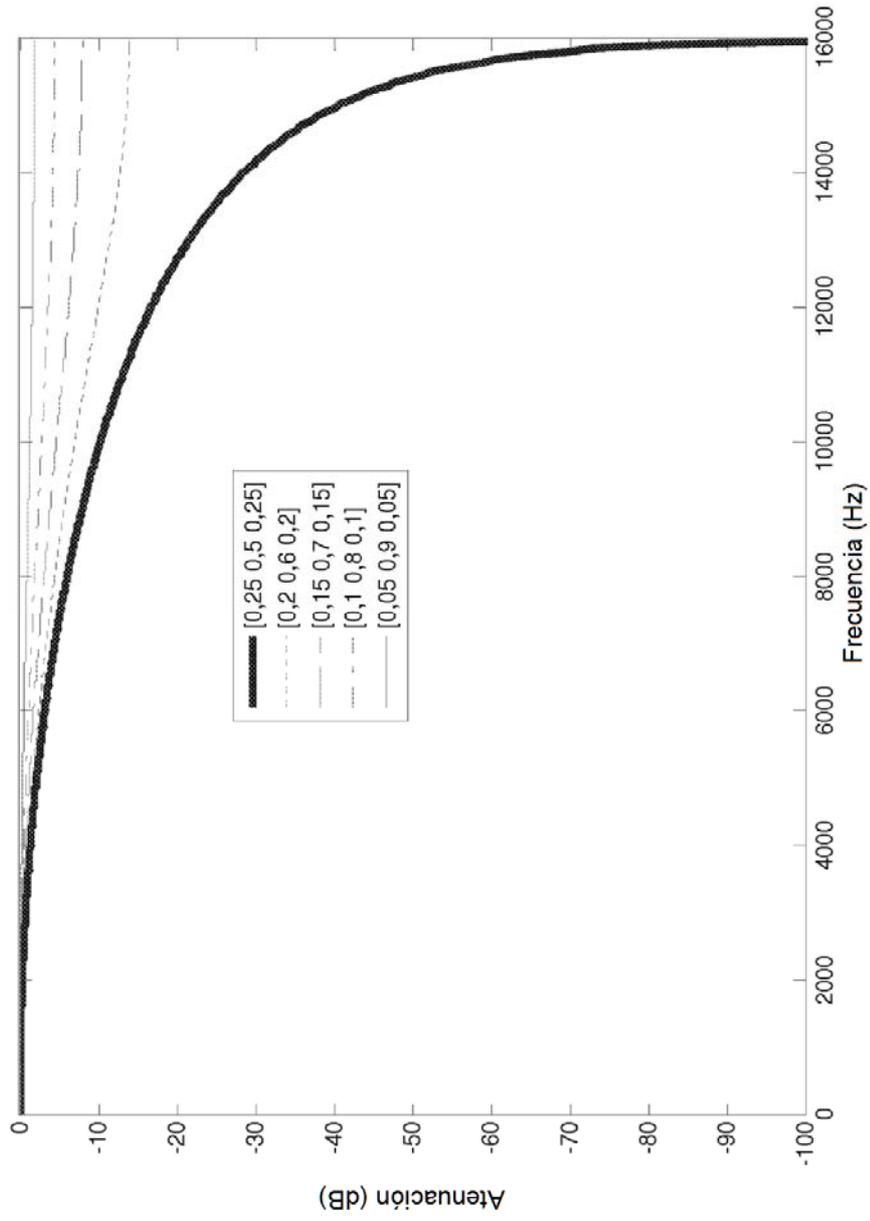


FIG.7

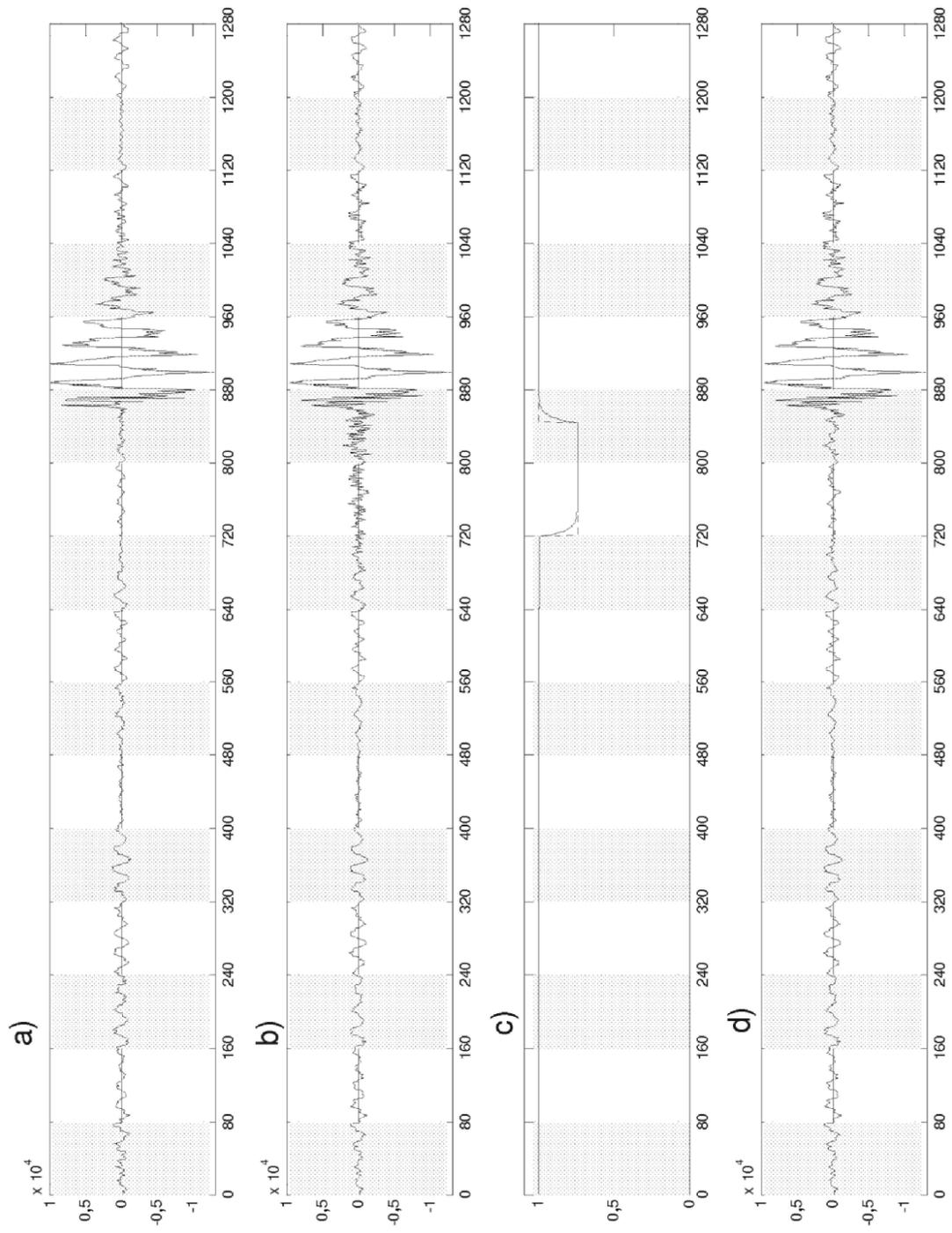


FIG.8

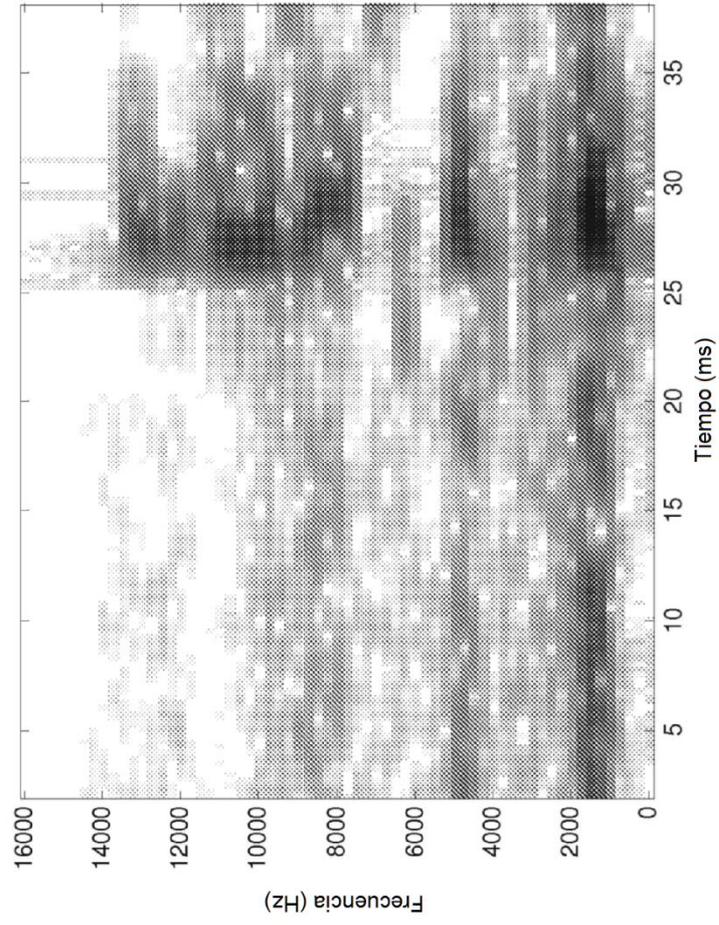


FIG.9

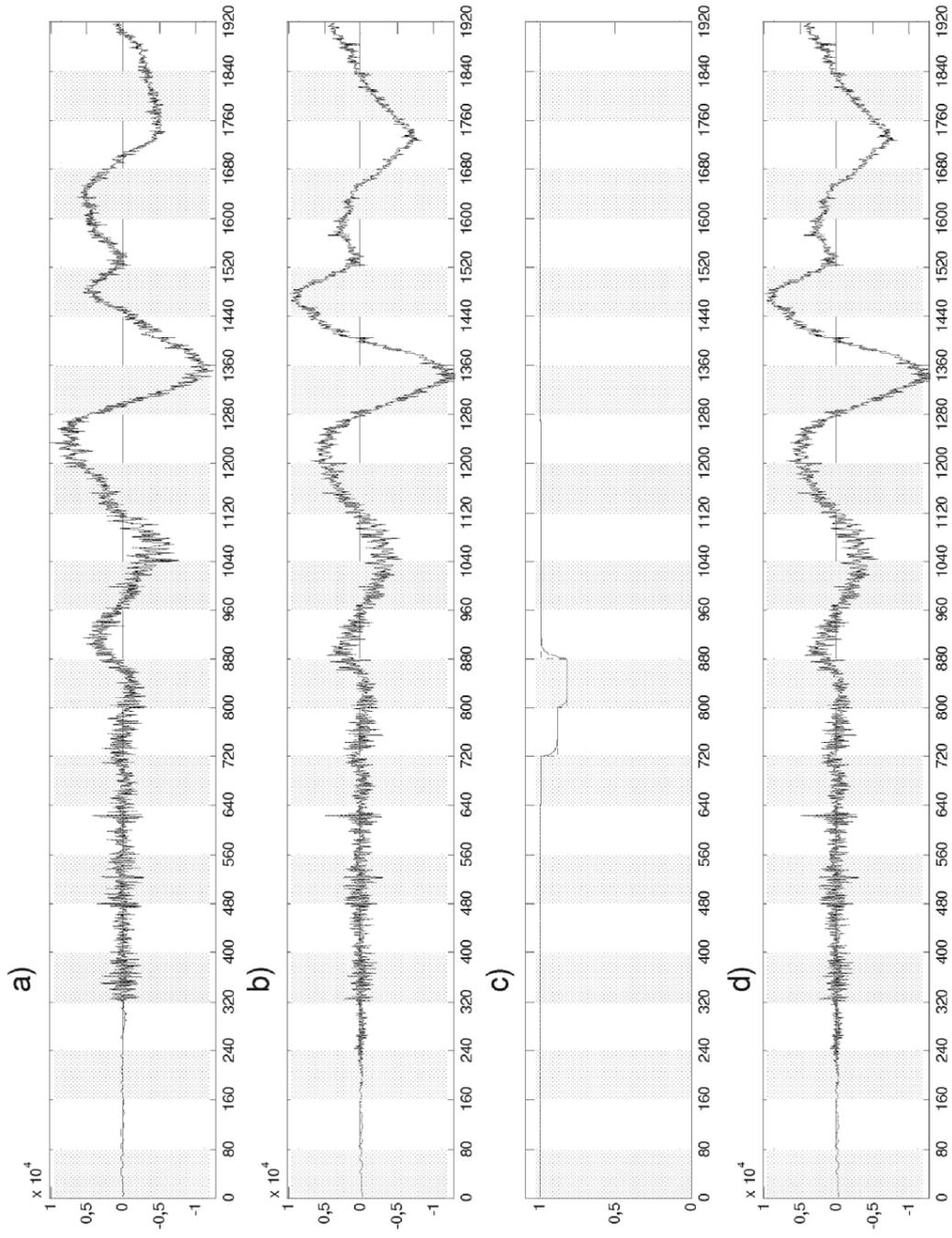


FIG.10

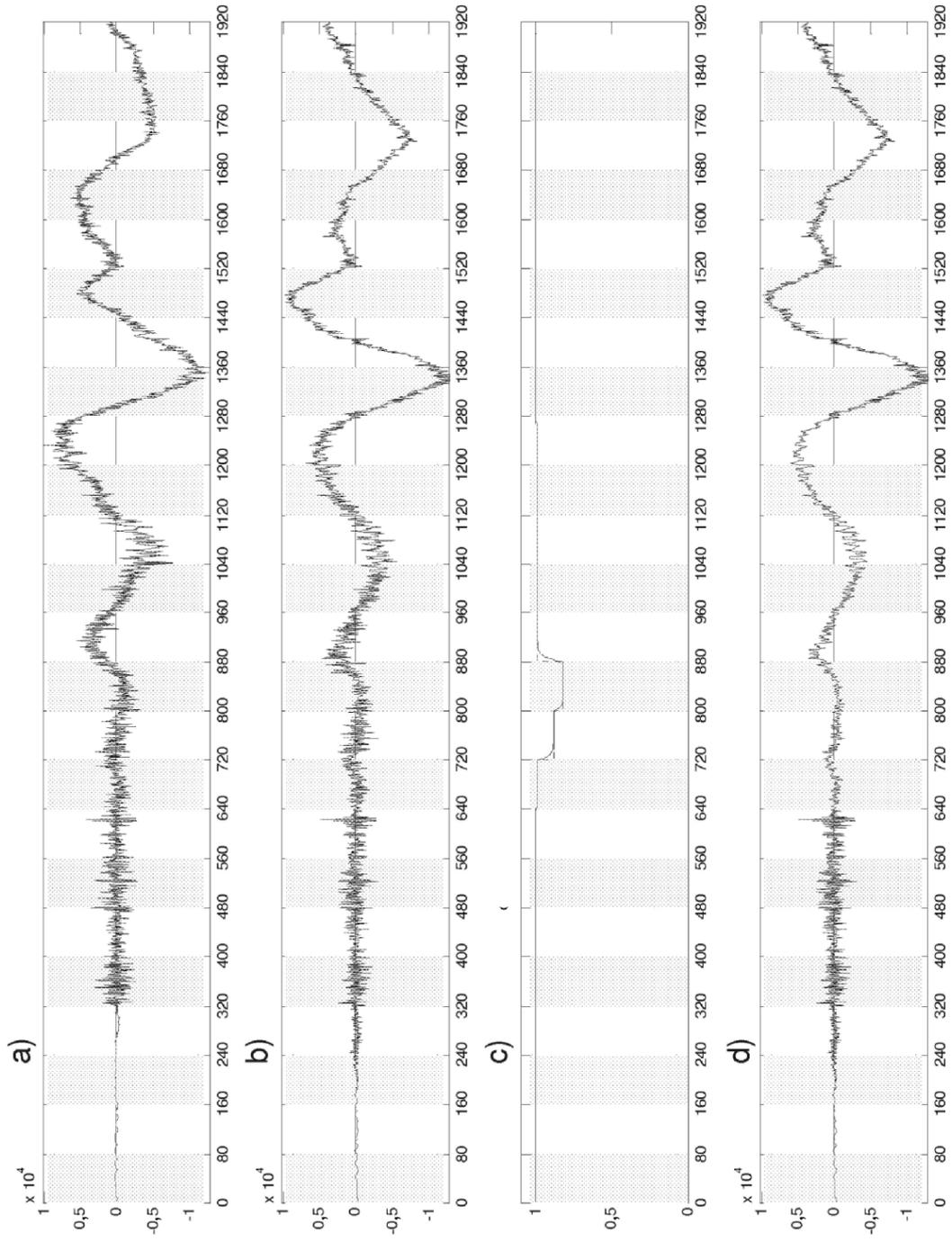


FIG.11

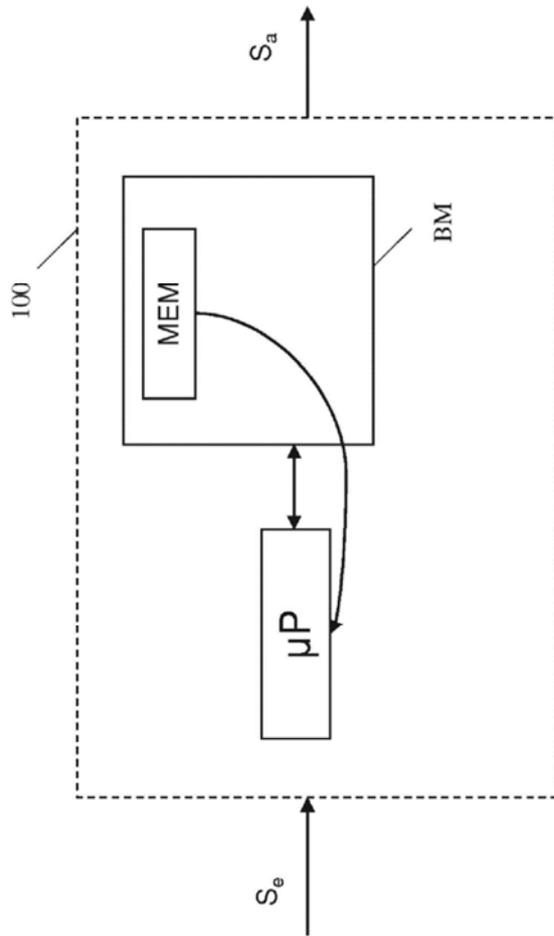


FIG.12