

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 692 831**

51 Int. Cl.:

G10L 19/26 (2013.01)

G10L 25/21 (2013.01)

G10L 25/51 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.09.2015 PCT/FR2015/052433**

87 Fecha y número de publicación internacional: **17.03.2016 WO16038316**

96 Fecha de presentación y número de la solicitud europea: **11.09.2015 E 15771686 (1)**

97 Fecha y número de publicación de la concesión europea: **01.08.2018 EP 3192073**

54 Título: **Discriminación y atenuación de pre-ecos en una señal de audio digital**

30 Prioridad:
12.09.2014 FR 1458608

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.12.2018

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 2 692 831 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

Discriminación y atenuación de pre-ecos en una señal de audio digital

5 La invención se refiere a un procedimiento y a un dispositivo de discriminación y de tratamiento de atenuación de los pre-ecos durante una decodificación de una señal de audio digital.

10 Para la transmisión de las señales de audio digitales en unas redes de telecomunicaciones, tanto si se trata por ejemplo de redes fijas como de móviles, o para almacenamiento de las señales, se recurre a unos procedimientos de compresión (o de codificación del origen) que implementan unos sistemas de codificación que son en general del tipo codificación temporal por predicción o codificación de frecuencia por transformada.

15 El procedimiento y el dispositivo, objeto de la invención, tienen también como campo de aplicación la compresión de las señales de sonido, 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 incluye un análisis-síntesis por adición/recubrimiento según la técnica anterior.

20 Ciertas secuencias musicales, tales como las percusiones y ciertos aspectos del habla como las plosivas (/k/, /t/, ...), se caracterizan por unos ataques extremadamente bruscos que se traducen en transiciones muy rápidas y una variación muy grande 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 tratamiento de la codificación/decodificación, la señal de entrada se trocea en bloques de muestras de longitud L cuyas fronteras se representan en la figura 1 por unos trazos verticales de puntos. La señal de entrada se indica por $x(n)$, en la que n es el índice de la muestra. El troceado en bloques sucesivos (o tramas) conduce a definir los bloques $X_N(n) = [x(N.L) \dots x(N.L+L-1)] = [x_N(0) \dots x_N(L-1)]$, en la que N es el índice del bloque (o 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 discreta del coseno modificada MDCT (por "Modified Discrete Cosine Transform" en inglés), se analizan conjuntamente dos bloques $X_N(n)$ y $X_{N+1}(n)$ para dar un bloque de coeficientes transformados asociados a la trama de índice N y la ventana de análisis es sinusoidal.

35 La división en bloques, también llamados tramas, realizada mediante la codificación por transformada es totalmente independiente de la señal sonora y las transiciones pueden aparecer por tanto en un punto cualquiera de la ventana de análisis. Ahora bien después de la decodificación por transformada, la señal reconstruida está manchada de "ruido" (o distorsión) generado por la operación de cuantificación (Q) - cuantificación inversa (Q^{-1}). Este ruido de codificación se reparte temporalmente de manera 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 tasa 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, el ruido es por tanto igualmente de nivel elevado.

45 En la codificación por transformada, el nivel de ruido de codificación es típicamente inferior al de la señal para los segmentos de gran energía que siguen inmediatamente a la transición, pero el nivel es superior al de la señal para los segmentos de energía más reducida, principalmente para la parte que precede a la transición (muestras 160-410 de la figura 1). Para la parte antes mencionada, la relación señal a ruido es negativa y la degradación resultante puede aparecer muy molesta para el oyente. Se denomina pre-eco al ruido de codificación anterior a la transición y pos-eco al ruido posterior a la transición.

50 Se puede observar en la figura 1 que el pre-eco afecta a la trama precedente a la transición así como a la trama en la que se produce la transición.

55 Los experimentos psico-acústicos han demostrado que el oído humano efectúa un pre-enmascarado temporal de los sonidos bastante limitado, del orden de algunos milisegundos. El ruido precedente al ataque, o pre-eco, es audible cuando la duración del pre-eco es superior a la duración del pre-enmascarado.

60 El oído humano efectúa igualmente un post-enmascarado de una duración más larga, de 5 a 60 milisegundos, durante el paso de las secuencias de gran energía a secuencias de energía reducida. La tasa o nivel de molestia aceptable para los pos-ecos es por tanto mayor que para los pre-ecos.

65 El fenómeno de los pre-ecos, más crítico, es tanto más molesto cuanto mayor es la longitud de los bloques en número de muestras. Ahora bien, en codificación por transformada, es bien conocido que para las señales fijas cuanto más aumenta la longitud de la transformada, mayor es la ganancia de la codificación. A frecuencia de

muestreo fija y a velocidad fija, si se aumenta el número de puntos de la ventana (por tanto la longitud de la transformada) se dispondrá de más bits por trama para codificar las rayas de frecuencia juzgadas útiles por el modelo psico-acústico, de ahí la ventaja de utilizar unos bloques de gran longitud. La codificación MPEG AAC (Advance Audio Coding), por ejemplo, utiliza una ventana de gran longitud que contiene un número fijo de muestras, 2048, es decir sobre una duración de 64 ms si la frecuencia de muestreo es de 32 kHz; el problema de los pre-ecos se gestiona aquí permitiendo conmutar estas ventanas largas a 8 ventanas cortas por medio de ventanas intermedias (denominadas de transición), lo que necesita cierto retardo en la codificación para detectar la presencia de una transición y adaptar las ventanas. La longitud de estas ventanas cortas es por tanto de 256 muestras (8 ms a 32 kHz). A baja velocidad se puede siempre tener un pre-eco audible de algunos milisegundos. La conmutación de las ventanas permite atenuar el pre-eco pero no suprimirlo. Los codificadores por transformada utilizados para las aplicaciones de conversación, como UIT-T G.722.1, G.722.1C o G.719, utilizan frecuentemente una longitud de trama de 20 ms y la ventana de duración de 40 ms a 16, 32 o 48 kHz (respectivamente). Se puede observar que el decodificador UIT-T G.719 integra un mecanismo de conmutación de ventanas con detección de transitorio, sin embargo el pre-eco no se reduce completamente a baja velocidad (típicamente a 32 kbits/s).

Con el objetivo de reducir el molesto efecto antes mencionado del fenómeno de los pre-ecos, se han propuesto diferentes soluciones a nivel del codificador y/o del decodificador.

La conmutación de ventanas se ha citado anteriormente; necesita transmitir una información auxiliar para identificar el tipo de ventanas utilizadas en la trama actual. Otra solución consiste en aplicar un filtrado adaptativo. En la zona precedente 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 titulado High Quality Audio Transform Coding at 64 kbits, IEEE Trans. on Communications Vol. 42, N.º 11, noviembre de 1994, publicada por Y. Mahieux y J. P. Petit.

La implementación de un filtrado de ese tipo necesita el conocimiento de parámetros de los que algunos, como los coeficientes de predicción y la variación de la señal corrompida por el pre-eco, se estiman en el decodificador a partir de las muestras con ruido. Por el contrario, informaciones tales como la energía de la señal de origen no pueden ser conocidas en el codificador y en consecuencia deben transmitirse. Esto hace preciso transmitir informaciones suplementarias, lo que con limitación de velocidad disminuye el fondo relativo asignado a la codificación por transformada. Cuando el bloque recibido contiene una variación brusca de dinámica, se le aplica el tratamiento de filtrado.

El proceso de filtrado antes citado no permite retomar la señal de origen, pero proporciona una fuerte reducción de los pre-ecos. Necesita sin embargo transmitir los parámetros suplementarios al decodificador.

A diferencia de las soluciones anteriores, se han propuesto diferentes técnicas de reducción del pre-eco sin transmisión específica de la información. Por ejemplo, una revisión de la reducción de los pre-ecos en el contexto de la codificación jerárquica se presenta en el artículo de B.Kövesi, S. Ragot, M. Gartner, H. Taddei, "Pre-echo reduction in the ITU-T G.729.1 embedded coder", EUSIPCO, Lausana, Suiza, agosto 2008.

Un ejemplo típico de procedimiento de atenuación de pre-ecos sin información auxiliar se describe en la solicitud de patente francesa FR 08 56248. En este ejemplo, se determinan unos factores de atenuación por sub-bloque, en los sub-bloques de reducida energía que preceden a un sub-bloque en el que se ha detectado una transición o ataque.

El factor de atenuación $g(k)$ en el k -ésimo sub-bloque se calcula por ejemplo en función de la relación $R(k)$ entre la energía del sub-bloque de mayor energía y la energía del k -ésimo sub-bloque en cuestión:

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

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

Si la energía de los sub-bloques varía poco con relación a la energía máxima en los sub-bloques considerados en la trama actual, no es necesaria entonces ninguna atenuación; el factor $g(k)$ es fijo con 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 pre-eco es molesto, la trama que precede a la trama de pre-eco tiene una energía homogénea que corresponde a la energía de un segmento de reducida energía (típicamente un ruido de fondo). Según la experiencia no es útil ni incluso deseable que después del tratamiento de atenuación del pre-eco la energía de la señal sea inferior a la energía media (por sub-bloque) de la señal que precede a la zona de tratamiento, típicamente la de la trama precedente, indicada \overline{En} , o la de la segunda mitad de la trama precedente, indicada por $\overline{En'}$.

Para el sub-bloque de índice k a tratar se puede calcular el valor límite, indicado por $\lim_g(k)$, del factor de atenuación con el fin de obtener exactamente la misma energía que la energía media por sub-bloque del segmento anterior al sub-bloque a tratar. Este valor está por supuesto limitado a un máximo de 1 puesto que se trata aquí de los valores de atenuación. Más precisamente se define en este caso:

5

$$\lim_g(k) = \min \left(\sqrt{\frac{\max(\overline{En}, \overline{En}')}{En(k)}}, 1 \right)$$

en la que la energía media del segmento anterior se aproxima por el valor $\max(\overline{En}, \overline{En}')$.

10 El valor $\lim_g(k)$ así obtenido sirve de límite inferior en el cálculo final del factor de atenuación del sub-bloque, se utiliza por tanto como sigue:

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

15 Los factores de atenuación (o ganancias) $g(k)$ determinadas por sub-bloque pueden alisarse a continuación mediante una función de alisado aplicada muestra por muestra para evitar variaciones bruscas del factor de atenuación en las fronteras de los bloques.

Por ejemplo, se puede definir inicialmente la ganancia por muestra como una función constante por fragmento:

20

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

en la que L' representa la longitud de un sub-bloque.

La función se alisa a continuación según la ecuación siguiente:

25

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

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

30

Son igualmente posibles otras funciones de alisado como por ejemplo el fundido encadenado lineal sobre u muestras:

$$g_{pre}(n) = \frac{1}{u} \sum_{i=0}^{u-1} g_{pre}'(n-i), n = 0, \dots, L-1$$

35

en la que $g_{pre}'(n)$ es la atenuación no alisada y $g_{pre}(n)$ es la atenuación alisada, $g_{pre}'(n)$ con $n = -(u-1), \dots, -1$ son los últimos $u-1$ factores de atenuación obtenidos para las últimas muestras del sub-bloque precedente. Se puede tomar por ejemplo $u = 5$.

40 Una vez así calculados los factores $g_{pre}(n)$, la atenuación de los pre-ecos se realiza sobre la señal reconstruida en la trama actual, $x_{rec}(n)$, multiplicando cada muestra por el factor correspondiente:

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

45

en la que $x_{rec,g}(n)$ es la señal decodificada y pos-tratada mediante la reducción del pre-eco.

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, antes mencionada, y resumida anteriormente.

50 En estos ejemplos la señal se muestrea a 32 kHz. La longitud de la trama es $L = 640$ muestras y cada trama se divide en 8 sub-bloques de $K = 80$ muestras.

En la parte a) de la figura 2, se representa una trama de una señal original muestreada a 32 kHz. Un ataque (o transición) en la señal se sitúa en el sub-bloque que comienza en el índice 320. Esta señal se ha codificado mediante un codificador por transformada de tipo MDCT de baja velocidad (24 kbits/s).

55

En la parte b) de la figura 2, se ilustra el resultado de la decodificación sin tratamiento de pre-eco. Se puede observar el pre-eco a partir de la muestra 160, en los sub-bloques precedentes al que contiene el ataque.

La parte c) muestra la evolución del factor de atenuación de pre-eco (línea continua) obtenido mediante el procedimiento descrito en la solicitud de patente del estado de la técnica antes citada. La línea de puntos representa el factor antes del alisado. Se remarca en este caso 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 la aplicación del tratamiento de pre-eco (multiplicación de la señal b) con la señal c)). Se ve que el pre-eco se ha atenuado correctamente. La figura 2 muestra igualmente que el factor de alisado no se eleva 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 reducido pero puede sin embargo evitarse. La figura 3 ilustra el mismo ejemplo que la figura 2, en el que, antes del alisado, el valor del factor de atenuación se fuerza a 1 para algunas muestras del sub-bloque precedente al sub-bloque en el que se sitúa el ataque. La parte c) de la figura 3 da un ejemplo de una corrección de ese tipo.

En este ejemplo se ha afectado por el valor de factor 1 a las 16 últimas muestras del sub-bloque precedente al ataque, a partir del índice 364. De ese modo la función de alisado incrementa progresivamente el factor para tener un valor próximo a 1 en el momento del ataque. Se preserva entonces la amplitud del ataque, como se ilustra en la parte d) de la figura 3, por el contrario algunas muestras de pre-eco no se atenúan.

En el ejemplo de la figura 3 la reducción del pre-eco por atenuación no permite reducir el pre-eco hasta el nivel del ataque, a causa del alisado de la ganancia.

Esta técnica de reducción de los pre-ecos es sin embargo perfectible para ciertos tipos de señales como unas señales de música moderna por ejemplo. En efecto, en ciertos casos, puede tener lugar una falsa detección de pre-eco. La figura 4 ilustra un ejemplo de una señal original de ese tipo, no codificada por tanto sin pre-eco. Se trata de la batida de un instrumento de percusión electrónico/sintético. Se puede observar aquí que antes del ataque neto hacia el índice 1600 hay un ruido sintético que comienza con el índice 1250. Este ruido sintético que forma parte por tanto de la señal se detectaría como un pre-eco por el algoritmo de detección de pre-eco descrito anteriormente, suponiendo una codificación/decodificación perfecta de la señal. El tratamiento de la atenuación de pre-ecos suprimiría por tanto esta componente de la señal. Esto desnaturizaría la señal decodificada (cuando la codificación/decodificación es perfecta), lo que no es deseable.

Existe por tanto una necesidad de una técnica mejorada de discriminación y de atenuación de pre-ecos en la decodificación, que permita hacer más fiable la detección de los pre-ecos y evitar falsas detecciones cuando no se transmite ninguna información auxiliar por parte del codificador.

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

Con este fin, la presente invención se refiere a un procedimiento de discriminación y de atenuación del pre-eco en una señal de audio digital generada a partir de una codificación por transformada, en la que, para una trama actual descompuesta en sub-bloques, los sub-bloques de baja energía débil que preceden a un sub-bloque en el que se detecta la transición o ataque determinan una zona de pre-eco en la que se efectúa un tratamiento de atenuación del pre-eco. El procedimiento es tal que, en el caso de que se detecte un ataque a partir del tercer sub-bloque de la trama actual, incluye las siguientes etapas:

- cálculo de un coeficiente director de las energías para al menos dos sub-bloques de la trama actual que preceden al sub-bloque en el que se detecta un ataque;
- comparación del coeficiente director con un umbral predefinido; y
- inhibición del tratamiento de atenuación del pre-eco en la zona de pre-eco en el caso de que el coeficiente director calculado sea inferior al umbral predefinido.

El coeficiente director de las energías calculado para los sub-bloques que preceden a la posición del ataque, permiten verificar la tendencia de aumento de la energía de la señal en la zona de pre-eco. Esto permite hacer más fiable la detección de los pre-ecos evitando falsas detecciones de pre-ecos. En efecto, observando la figura 1 se puede ver que el pre-eco tiene una característica típica: su energía tiene una tendencia creciente al aproximarse al ataque de origen del pre-eco. La forma de las ventanas de ponderación de la adición-recubrimiento explica esto. Igualmente si el pre-eco tiene una energía casi constante con la adición-recubrimiento, las señales en la entrada del módulo de adición-recubrimiento se multiplican por unas ventanas de ponderación cuyos pesos disminuyen hacia el pasado. En el caso de la señal del ejemplo de la figura 4, la energía de la señal antes del ataque es aproximadamente constante lo que permite diferenciarla de un pre-eco. De ese modo, la verificación de una energía creciente de la señal en la zona de pre-eco permite aumentar la fiabilidad de la detección del pre-eco.

En un modo de realización particular, el procedimiento incluye además una etapa de descomposición de la señal de audio digital en al menos dos sub-señales en función de un criterio de frecuencia y porque las etapas de cálculo de comparación se efectúan para al menos una de las sub-señales.

5 Cuando la posición del ataque se detecta en el tercer sub-bloque de la trama actual, la energía de los dos sub-bloques se utiliza en la zona de pre-eco para calcular un coeficiente director y compararlo con un umbral. Solamente con dos puntos, solo la verificación para la sub-señal de alta frecuencia en el caso de una descomposición en dos sub-señales es suficiente para detectar una falsa detección de pre-eco.

10 En el caso en el que es suficiente el número de sub-bloques que preceden al sub-bloque en el que se ha detectado una posición de ataque, el procedimiento incluye además una etapa de descomposición de la señal de audio digital en al menos dos sub-señales en función de un criterio de frecuencia y porque las etapas de cálculo y de comparación se efectúan para cada una de las sub-señales, la inhibición del tratamiento de atenuación de pre-eco en la zona de pre-eco de todas la sub-señales, se efectúa cuando un coeficiente director calculado es inferior al umbral predefinido para al menos una sub-señal.

15 La división en sub-señales permite de ese modo efectuar una atenuación del pre-eco de manera independiente y adaptada a las sub-señales. La fiabilidad de detección de la zona de pre-eco se refuerza para cada una de las sub-señales mediante la verificación del valor de los coeficientes directores respectivos.

Según un modo de realización particular, se define un umbral diferente por sub-señal.

20 Esto permite adaptar la verificación a las características espectrales de las sub-señales.

En un modo de realización, el coeficiente director se calcula según un método de estimación de mínimos cuadrados.

Este método de cálculo es de reducida complejidad.

25 En un modo posible de realización, el coeficiente director se normaliza.

De ese modo el coeficiente director es más fácilmente comparable con un umbral cuando este es diferente de 0.

30 En un modo de realización posible, en el caso de que se detecte un ataque en el primer o segundo sub-bloque de la trama actual, se utiliza para la etapa de comparación un coeficiente director calculado para la trama precedente.

35 La presente invención se refiere igualmente a un dispositivo de discriminación y de atenuación de pre-eco en una señal de audio digital generada a partir de una codificación por transformada, que incluye un módulo de detección de transición o ataque, un módulo de discriminación de zona de pre-eco y un módulo de tratamiento de atenuación del pre-eco, siendo efectuado un tratamiento de atenuación del pre-eco para una trama actual descompuesta en sub-bloques, en los sub-bloques de reducida energía que preceden a un sub-bloque en el que se detecta una transición o ataque que determina una zona de pre-eco. El dispositivo es tal que, en el caso en el que se detecte un ataque a partir del tercer sub-bloque de la trama actual, incluye además:

- 40
- un módulo de cálculo que calcula el coeficiente director de las energías para al menos dos sub-bloques de la trama actual que preceden al sub-bloque en el que se detecta un ataque;
 - un comparador adecuado para efectuar una comparación del coeficiente director con un umbral predefinido; y
 - un módulo de discriminación adecuado para inhibir el tratamiento de la atenuación de pre-eco en la zona de pre-eco en el caso de que el coeficiente director calculado sea inferior al umbral predefinido.
- 45

Las ventajas de este dispositivo son las mismas que las descritas para el procedimiento de discriminación y de tratamiento de atenuación que implementa.

50 La invención se dirige a un decodificador de la señal de audio digital que incluye un dispositivo tal como el descrito anteriormente.

La invención se dirige también a un programa informático que incluye unas instrucciones de códigos para la implementación de las etapas del procedimiento tal como se ha descrito anteriormente, cuando estas instrucciones se ejecutan por un procesador.

55 Finalmente, la invención se refiere a un soporte de almacenamiento, legible por un procesador, integrado o no en el dispositivo de tratamiento, eventualmente extraíble, que memoriza un programa informático que implementa un procedimiento de tratamiento tal como se ha descrito anteriormente.

60 Aparecerán más claramente otras características y ventajas de la invención con la lectura de la descripción que sigue, dada únicamente a título de ejemplo no limitativo, y realizada con referencia a los dibujos adjuntos, en los que:

- 65
- la figura 1 descrita anteriormente ilustra un sistema de codificación-decodificación por transformada según el estado de la técnica;

- la figura 2 descrita anteriormente ilustra un ejemplo de señal de audio digital para el que se efectúa un método de atenuación según el estado de la técnica;
- la figura 3 ilustra otro ejemplo de señal de audio digital para el que se efectúa un método de atenuación según el estado de la técnica;
- 5 - la figura 4 descrita anteriormente ilustra un ejemplo de una señal para el que la técnica del estado de la técnica detectaría erróneamente un pre-eco;
- la figura 5 ilustra un modo de realización de un procedimiento y un dispositivo de discriminación y de tratamiento de atenuación de pre-eco comprendido en un decodificador según la invención;
- la figura 6 ilustra un ejemplo de ventanas de análisis y de ventanas de síntesis de reducido retardo para la codificación y la decodificación por transformada susceptible de crear el fenómeno de pre-eco;
- 10 - la figura 7 ilustra un ejemplo de señal de audio digital para el que se implementa el método de atenuación de pre-eco según un modo de realización de la invención;
- la figura 8 ilustra un ejemplo material de dispositivo de discriminación y de tratamiento de la atenuación según la invención.

15 Con referencia a la figura 5, se describe un dispositivo 600 de discriminación y de tratamiento de atenuación del pre-eco. El dispositivo de tratamiento de la atenuación 600 tal como se describe a continuación está comprendido en un decodificador que incluye un módulo 610 de cuantificación inversa (Q^{-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 (Ad/Rec) como se describe con referencia a la figura 1 y que proporciona una señal reconstruida $x_{rec}(n)$ al dispositivo de discriminación y de tratamiento de atenuación según la invención. Se puede observar que se toma en este caso el ejemplo de la transformada MDCT que es la más corriente en codificación de la voz y audio, sin embargo el dispositivo 600 se aplica igualmente a cualquier otro tipo de transformada (FFT, DCT, etc.).

20 A la salida del dispositivo 600, se proporciona la señal tratada Sa en la que se ha efectuado una atenuación de pre-eco.

El dispositivo 600 implementa un método de discriminación y de tratamiento de atenuación de los pre-ecos en la señal decodificada $x_{rec}(n)$.

30 En un modo de realización de la invención, el procedimiento de discriminación y de tratamiento de atenuación incluye una etapa de detección (E601) de los ataques que pueden generar un pre-eco, en la señal decodificada $x_{rec}(n)$.

35 De ese modo, el dispositivo 600 incluye un módulo de detección 601 adecuado para implementar una etapa de detección (E601) de la posición de un ataque en una señal de audio decodificada.

40 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 "transitorio". En lo que sigue y sin pérdida de generalidad, se utilizarán únicamente los términos de ataque o de transición para designar igualmente unos transitorios.

45 Cada trama actual de L muestras de la señal decodificada $x_{rec}(n)$ se divide en K sub-bloques de longitud L', con, por ejemplo, L = 640 muestras (20 ms) a 32 kHz, L' = 80 muestras (2,5 ms) y K = 8. Preferentemente el tamaño de estos sub-bloques es por tanto idéntico pero la invención continúa siendo válida y fácilmente generalizable cuando los sub-bloques tienen un tamaño variable. Este puede ser el caso por ejemplo cuando la longitud de la trama L no es divisible por el número de sub-bloques K o si la longitud de la trama es variable.

50 Se utilizan unas ventanas de análisis - síntesis especiales de bajo retardo similares a las descritas en la norma UIT-T G.718 para la parte de análisis y para la parte de síntesis de la transformación MDCT. Un ejemplo de dichas ventanas se ilustra con referencia a la figura 6. El retardo generado por la transformación es de solamente 280 muestras contrariamente al retardo de 640 muestras en caso de la utilización de ventanas sinusoidales clásicas. De ese modo la memoria MDCT con unas ventanas de análisis - síntesis especiales de bajo retardo no contiene más que 140 muestras independientes (no replegadas con la trama actual) contrariamente a las 320 muestras en el caso de reutilización de las ventanas sinusoidales clásicas.

55 Se puede remarcar en efecto en la figura 6 para las ventanas de análisis (Ana.), que la zona de repliegue está limitada por los trazos de puntos entre las muestras 820 y 1100. La línea de repliegue se representa en trazo mixto en la muestra 960.

60 Para la síntesis (Sínt.), solo son necesarias las muestras representadas por el intervalo M (140 muestras) para obtener la información sobre la zona de repliegue del análisis, aprovechando la simetría. Estas muestras contenidas en memoria son útiles entonces para decodificar esta zona de repliegue utilizando igualmente las muestras replegadas de la ventana de la trama siguiente. En el caso de un ataque en esta zona entre las muestras 820 y 1100 la energía media de las muestras representadas por el intervalo M es netamente superior a la energía de las

subtramas que preceden a la muestra 820. El aumento brusco de la energía del intervalo M contenido en la memoria MDCT puede por tanto señalar un ataque en la trama siguiente que puede generar un pre-eco en la trama actual.

5 Se utiliza la memoria MDCT $x_{MDCT}(n)$ que da una versión con repliegue temporal ("folding" en inglés) de la señal futura. Con las ventanas de análisis - síntesis especiales de bajo retardo como se ilustran en la figura 6, no se retiene más que un ($K'=1$) bloque de longitud $L_m(0)=140$ que contiene todas las muestras independientes de la memoria MDCT. A pesar del número de muestras superior en este sub-bloque, su energía sigue siendo comparable con la de los sub-bloques de la trama actual (si la señal continúa siendo estable) porque la parte de la memoria se ha enmarcado (por tanto atenuado) por la ventana de análisis.

10 En efecto, la figura 1 muestra que el pre-eco influye en la trama que precede a la trama en la que se sitúa el ataque, y es deseable detectar un ataque en la trama futura que está en parte contenida en la memoria MDCT.

15 La trama actual y la memoria MDCT pueden verse como unas señales concatenadas que forman una señal troceada en $(K+K')$ sub-bloques consecutivos. En estas condiciones, se define la energía en el k-ésimo sub-bloque como:

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

20 cuando el k-ésimo sub-bloque se sitúa en la trama actual y, como:

$$En(k) = \sum_{n=0}^{L_{mem}-1} x_{MDCT}(n)^2$$

25 cuando el sub-bloque está en la memoria MDCT (que representa la señal disponible para la trama futura) y L_{mem} es la longitud del sub-bloque de la parte de memoria:

La energía media de los sub-bloques en la trama actual se obtiene por tanto como:

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

30 Se define igualmente la energía media de los sub-bloques en la segunda parte de la trama actual como (suponiendo que K sea un número par):

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

Un ataque asociado a un pre-eco se detecta si la relación $R(k) = \frac{\max_{n=0, K+K'-1} (En(n))}{En(k)}$ sobrepasa un umbral

35 predefinido, en uno de los sub-bloques considerados. Son posibles otros criterios de detección del pre-eco sin cambiar la naturaleza de la invención.

Por otro lado, se considera que la posición del ataque se define como

$$pos = \min \left(L', \left(\arg \max_{k=0, K+K'-1} (En(k)) \right), L \right)$$

40 en la que la limitación a L asegura que la memoria MDCT no se modifica jamás. Son igualmente posibles otros métodos de estimación más precisa de la posición del ataque.

45 El dispositivo 600 incluye igualmente un módulo 602 de discriminación de la zona de pre-eco que implementa una etapa de determinación (E602) en la zona de pre-eco (ZPE) que precede a la posición de ataque detectada. Se llama en este caso zona de pre-eco a la zona que cubre las muestras antes de la posición estimada del ataque que están perturbadas por el pre-eco generado por el ataque y en el que la atenuación del este pre-eco es deseable. En el modo de realización presentado la zona de pre-eco puede determinarse sobre la señal decodificada.

50 En un modo de realización de obtención de las zonas de pre-ecos, las energías $En(k)$ se concatenan en orden cronológico, inicialmente con la envolvente temporal de la señal decodificada, posteriormente 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, la presencia de pre-eco se detecta por ejemplo si la relación $R(k)$ sobrepasa un umbral, típicamente este umbral es 16.

Los sub-bloques en los que un pre-eco se ha detectado constituyen de ese modo una zona de pre-eco, que en general cubren las muestras $n = 0, \dots, pos-1$, es decir del comienzo de la trama actual a la posición del ataque (pos). Se puede también anotar que la zona de pre-eco puede extenderse en toda la trama actual si el ataque se ha detectado en la trama futura.

5 El dispositivo 600 incluye un módulo 603 de cálculo adecuado para implementar una etapa de cálculo del coeficiente director (o indicador de tendencia de variación) de las energías de los sub-bloques que preceden al sub-bloque en el que se ha detectado un ataque.

10 Se define el modelo lineal que representa un conjunto de n realizaciones (t_i, e_i) , $0 \leq i < n$ en el que t_i son los índices temporales de los sub-bloques y e_i sus energías, con la ecuación

$$e = b_0 + b_1 t \quad (1)$$

15 En la que b_0 es el valor en el instante $t=0$ y b_1 es el coeficiente director. El coeficiente director da la información sobre la tendencia (media) de la variación de la energía. Un coeficiente director positivo indica un aumento de las energías. Un valor próximo a cero indica una energía constante.

20 Se puede determinar el valor de b_1 por ejemplo mediante regresión lineal según los mínimos cuadrados:

$$b_1 = \frac{\sum(t_i - \bar{t})(e_i - \bar{e})}{\sum(t_i - \bar{t})^2} \quad (2)$$

En el que la suma se efectúa sobre unos índices i predeterminados.

25 El valor de b_1 depende igualmente de la magnitud (en valor absoluto) de las energías; es en efecto homogénea con la energía en el tiempo. Para poder comparar mejor el valor de b_1 con un umbral (por ejemplo fijo) se puede suprimir esta dependencia. Por ejemplo, el valor de b_1 puede dividirse por el valor medio de las energías para obtener el coeficiente director normalizado:

30
$$b_{1n} = \frac{b_1 n}{\sum e_i} \quad (3)$$

De manera alternativa se podrá tomar el coeficiente de correlación.

$$b_{1n,alt} = \frac{\sum(t_i - \bar{t})(e_i - \bar{e})}{\sqrt{\sum(t_i - \bar{t})^2 \sum(e_i - \bar{e})^2}} \quad (4)$$

35 Esta solución alternativa tiene una complejidad de cálculo más elevada porque necesita calcular una raíz cuadrada.

Son igualmente posibles otros métodos de estimación del coeficiente director como por ejemplo el método mediana - mediana de Tukey.

40 Se puede observar igualmente que cuando el coeficiente director debe compararse con un umbral de valor nulo —lo que se convierte en verificar el signo de este coeficiente— no es necesario normalizar este coeficiente.

45 Por otro lado, en lugar de normalizar el coeficiente director, será posible hacer el umbral variable porque las relaciones siguientes son equivalentes:

$$b_{1n} = \frac{b_1 n}{\sum e_i} < umbral$$

$$b_1 < umbral \cdot \frac{\sum e_i}{n}$$

50 Si el ataque se detecta en el primer o segundo sub-bloque es posible la verificación según la invención. Si el ataque se detecta en el tercer sub-bloque se dispone de la energía de 2 sub-bloques en la zona de pre-eco, e_0 y e_1 para hacer esta verificación (siendo e_1 el más próximo al ataque). Con 2 puntos la ecuación (3) se simplifica así:

55
$$b_{1n} = \frac{2(e_1 - e_0)}{e_1 + e_0} \quad (5)$$

Si el ataque se detecta en el cuarto sub-bloque se dispone de la energía de 3 sub-bloques en la zona de pre-eco, e_0 , e_1 y e_2 para realizar esta verificación (siendo e_2 el más próximo al ataque). Con 3 putos la ecuación (3) se simplifica así:

$$b_{1n} = \frac{3(e_2 - e_0)}{2(e_2 + e_1 + e_0)} \quad (6)$$

Si se dispone de 4 o más sub-bloques, se puede calcular el coeficiente director sobre los 4 o más de los sub-bloques. La experiencia muestra que la verificación del coeficiente director calculado sobre los 3 sub-bloques precedentes al sub-bloque en el que se ha detectado el ataque es suficiente para evitar las falsas detecciones de los pre-ecos —esta conclusión se aplica sobre el caso de 8 sub-bloques en cada trama de 20 ms y puede adaptarse según el tamaño de los sub-bloques y de la trama—.

Así en el modo de realización preferido, el coeficiente director se calcula con al menos 3 sub-bloques. Esto permite limitar la complejidad máxima del cálculo del coeficiente director.

Según la invención, el coeficiente director normalizado b_{1n} así obtenido se compara a continuación en la etapa E604 mediante un módulo comparador 604 con un umbral predefinido. El umbral puede predefinirse con un valor fijo o puede ser variable en función por ejemplo de la clasificación de la señal según un criterio de la voz o música. Típicamente este umbral es igual a 0 si se verifica únicamente que la energía no disminuye o es igual a 0,2 si se impone un ligero aumento de la energía en la zona de pre-eco. Si el coeficiente director normalizado b_{1n} es inferior a este umbral se concluye que la señal en la zona de pre-eco no corresponde a un pre-eco típico y se inhibe la atenuación de los pre-ecos en esta zona en la etapa E602. De ese modo se evita que una señal decodificada cuya señal de entrada original contiene un componente de reducida energía antes del ataque quede modificada/alterada por error por el módulo de atenuación de los pre-ecos que detecta este componente como un pre-eco.

Una atenuación de pre-eco se implementa en la etapa E607 por el módulo de atenuación 607 para la zona de pre-eco discriminada. El factor de atenuación se calcula por ejemplo como en la solicitud FR 08 56248. En el caso en el que el módulo 604 ha detectado una falsa detección de pre-eco, el factor de atenuación puede forzarse a 1 inhibiendo así la atenuación o bien el módulo de discriminación 602 no discrimina esta zona como una zona de pre-eco, no siendo solicitado entonces el módulo de atenuación.

En un modo de realización particular, el dispositivo 600 incluye además un módulo 605 de descomposición de la señal, adecuado para efectuar una etapa E605 de descomposición de la señal decodificada en al menos dos sub-señales según un criterio predeterminado. Este método se describe principalmente en la solicitud FR 12 62598 de la que se recuerdan aquí algunos elementos.

En una realización particular de la invención, la señal decodificada $x_{rec}(n)$ se descompone en la etapa E605 en dos sub-señales de la manera siguiente:

- la primera sub-señal $x_{rec,ss1}(n)$ se obtiene por filtrado paso bajo utilizando un filtro FIR (filtro de respuesta a impulsos finita) de 3 coeficientes y de fase nula de función de transferencia $c(n)z^{-1} + (1-2c(n)) + c(n)z$ siendo $c(n)$ un valor comprendido entre 0 y 0,25, en la que $[c(n), 1-2c(n), c(n)]$ son los coeficientes del filtro paso bajo; este filtro se implementa con la ecuación en las diferencias:

$$x_{rec,ss1}(n) = c(n) x_{rec}(n-1) + (1-2c(n)) x_{rec}(n) + c(n) x(n+1)$$

En un modo particular de realización, se utiliza un valor constante $c(n) = 0,25$.

Se puede remarcar que la sub-señal $x_{rec,ss1}(n)$ resultante de este filtrado, contiene por tanto unas componentes más bien de baja frecuencia de la señal decodificada.

- La segunda sub señal $x_{rec,ss2}(n)$ se obtiene por filtrado paso alto utilizando un filtro FIR de 3 coeficientes y de fase nula de función de transferencia $-c(n)z^{-1} + 2c(n) - c(n)z$, en la que $[-c(n), 2c(n), -c(n)]$ son los coeficientes del filtro paso alto; este filtro se implementa con la ecuación en las diferencias:

$$x_{rec,ss2}(n) = -c(n) x_{rec}(n-1) + 2c(n) x_{rec}(n) - c(n) x(n+1). \text{ La sub-señal } x_{rec,ss2}(n) \text{ resultante de este filtrado, contiene por tanto unas componentes más bien de alta frecuencia de la señal decodificada.}$$

A observar que $x_{rec,ss1}(n) + x_{rec,ss2}(n) = x_{rec}(n)$.

Es por tanto igualmente posible obtener $x_{rec,ss2}(n)$ restando $x_{rec,ss1}(n)$ de $x_{rec}(n)$ lo que reduce la complejidad de los cálculos: $x_{rec,ss2}(n) = x_{rec}(n) - x_{rec,ss1}(n)$.

La combinación de las sub-señales atenuadas para la obtención de la señal atenuada Sa se realiza por simple adición de las sub-señales atenuadas en la etapa E608 descrita posteriormente.

Para no utilizar señal futura para estos filtrados, se puede completar por ejemplo la señal decodificada mediante una muestra a 0 al final del bloque. En el caso de la señal decodificada completada por una muestra a 0 al final del bloque para $n=L-1$ la sub-señal $x_{rec,ss1}(n)$ se obtiene mediante:

$$x_{rec,ss1}(L-1) = c(L-1) x_{rec}(L-2) + (1-2c(L-1)) x_{rec}(L-1),$$

$x_{rec,ss2}(n)$ se calcula siempre como $x_{rec,ss2}(n) = x_{rec}(n) - x_{rec,ss1}(n)$.

Se puede observar que las dos sub-señales permanecen en este caso con la misma frecuencia de muestreo que la señal decodificada.

Se implementa una etapa E606 de cálculo de factores de atenuación de pre-eco en el módulo de cálculo 606. Este cálculo se realiza separadamente para las dos sub-señales.

Estos factores de atenuación se obtienen por muestreo en la zona de pre-eco determinada en E602 en función de la trama en la que se ha detectado el ataque y de la trama precedente.

Se obtienen entonces los factores $g_{pre,ss1}'(n)$ y $g_{pre,ss2}'(n)$ en los que n es el índice de la muestra correspondiente. Estos factores se alisarán eventualmente para obtener los factores $g_{pre,ss1}(n)$ y $g_{pre,ss2}(n)$ respectivamente. Este alisado es importante sobre todo para las sub-señales que contienen las componentes de baja frecuencia (por tanto para $g_{pre,ss1}'(n)$ en este ejemplo).

Un ejemplo de realización de cálculo de atenuación se describe en la solicitud de patente FR 08 56248. Los factores de atenuación se calculan por sub-bloque. En el procedimiento descrito aquí, se calculan además por separado para cada sub-señal. Para las muestras que preceden al ataque detectado, se calculan por tanto los factores de atenuación $g_{pre,ss1}'(n)$ y $g_{pre,ss2}'(n)$. A continuación estos valores de atenuación se alisan eventualmente para obtener los valores de atenuación por muestra.

El cálculo de factor de atenuación de una sub-señal (por ejemplo $g_{pre,ss2}'(n)$) puede ser similar al descrito en la solicitud de patente FR 08 56248 para la señal decodificada en función de la relación $R(k)$ (utilizada igualmente para la detección del ataque) entre la energía del sub-bloque de mayor energía y la energía del k -ésimo sub-bloque de la señal decodificada. Se inicializa $g_{pre,ss2}'(n)$ como:

$$g_{pre,ss2}'(n) = g(k) = f(R(k)), n = kL', \dots, (k+1)L'-1; k=0, \dots, K-1$$

en la que f es una función decreciente de valores entre 0 y 1, por ejemplo $f = 0$ si $R(k) \leq 16$, $f = 0,1$ si $16 > R(k) \geq 32$ y $f = 0,01$ si $R(k) > 32$.

Si la variación de la energía con relación a la energía máxima es reducida, no es necesaria entonces ninguna atenuación. El valor se fija entonces a un valor de atenuación que inhibe la atenuación, es decir a 1. Si no, el factor de atenuación está comprendido entre 0 y 1. Esta inicialización puede ser común para todas la sub-señales.

Los valores de atenuación se afinan a continuación por sub-señal para poder regular el nivel de atenuación óptimo por sub-señal en función de las características de la señal decodificada. Por ejemplo las atenuaciones pueden limitarse en función de la energía media de la sub-señal de la trama precedente porque no es deseable que después del tratamiento de atenuación del pre-eco, la energía de la señal se convierta en inferior a la energía media por sub-bloque de la señal precedente a la zona de tratamiento (típicamente la de la trama precedente o la de la segunda mitad de la trama precedente).

Esta limitación puede hacerse de manera similar a la descrita en la solicitud de patente FR 08 56248. Por ejemplo para la segunda sub-señal $x_{rec,ss2}(n)$ se calcula inicialmente la energía en los K sub-bloques de la trama actual como:

$$En_{ss2}(k) = \sum_{n=kL'}^{(k+1)L'-1} x_{rec,ss2}(n)^2, k=0, \dots, K-1$$

Se conoce igualmente por memorización la energía media de la trama precedente $\overline{En_{ss2}}$ y la de la segunda mitad de la trama precedente $\overline{En_{ss2}'}$ que pueden calcularse (en la trama precedente) como:

$$\overline{En_{ss2}} = \frac{1}{K} \sum_{k=0}^{K-1} En_{ss2}(k)$$

y

$$\overline{En}_{ss2}' = \frac{2}{K} \sum_{k=K/2}^{K-1} En_{ss2}(k)$$

en la que los índices del sub-bloque de 0 a K corresponden a la trama actual.

- 5 Para el sub-bloque k a tratar se puede calcular el valor límite del factor $\lim_{g,ss2}(k)$ con el fin de obtener exactamente la misma energía que la energía media por sub-bloque del segmento precedente al sub-bloque a tratar. Este valor está limitado por supuesto a un máximo de 1 puesto que se trata aquí de los valores de atenuación. Más precisamente:

$$\lim_{g,ss2}(k) = \min \left(\sqrt{\frac{\max(\overline{En}_{ss2}, \overline{En}_{ss2}')} {En_{ss2}(k)}}, 1 \right)$$

en la que la energía media del segmento precedente se aproxima mediante $\max(\overline{En}_{ss2}, \overline{En}_{ss2}')$.

El valor $\lim_{g,ss2}(k)$ así obtenido sirve de límite inferior en el cálculo final del factor de atenuación del sub-bloque:

$$g_{pre,ss2}'(n) = \max \left(g_{pre,ss2}'(n), \lim_{g,ss2}(k) \right), n = kL', \dots, (k+1)L'-1; k = 0, \dots, K-1$$

En una primera variante de realización la zona de pre-eco en la que se extiende la atenuación desde el inicio de la trama actual hasta el inicio del sub-bloque en el que se ha detectado el ataque — hasta el índice *pos* en donde

20 $pos = \min \left(L' \cdot \left(\arg \max_{k=0, K'+K'-1} (En(k)) \right), L' \right)$. Las atenuaciones asociadas a las muestras del sub-bloque del ataque son todas puestas a 1 incluso si el ataque se sitúa hacia el final de este sub-bloque.

En otra variante de realización la posición de inicio del ataque *pos* se afina en el sub-bloque del ataque, por ejemplo troceando el sub-bloque en sub-sub-bloques y observando la evolución de la energía de estos sub-sub-bloques. Supóngase que la posición de inicio del ataque se detecta en el sub-bloque k, k>0 y el inicio del ataque afinado *pos* se encuentra en este sub-bloque, los valores de atenuación para las muestras de este sub-bloque que se encuentran antes del índice *pos* pueden inicializarse en función del valor de atenuación correspondiente a la última muestra del sub-bloque precedente:

$$g_{pre,ss2}'(n) = g_{pre,ss2}'(kL'-1), n = kL', \dots, pos-1$$

Todas las atenuaciones a partir del índice *pos* se ponen a 1.

Para la primera sub-señal que contiene las componentes de baja frecuencia de la señal decodificada, el cálculo de los valores de atenuación que se basa en la sub-señal $x_{rec,ss1}(n)$ puede ser similar al cálculo de los valores de atenuación que se basan en la señal decodificada $x_{rec}(n)$. De ese modo, en una variante de realización, por razones de reducción de complejidad de cálculo, los valores de atenuación pueden determinarse basándose en la señal decodificada $x_{rec}(n)$. En el caso en el que la detección de los ataques se realiza sobre la señal decodificada no es por tanto ya necesario recalcular unas energías de los sub-bloques porque para esta señal los valores de energía por sub-bloque ya se han calculado para detectar los ataques. Como para la gran mayoría de las señales las bajas frecuencias son mucho más energéticas que las altas frecuencias, las energías por sub-bloque de la señal decodificada $x_{rec}(n)$ y de la sub-señal $x_{rec,ss1}(n)$ están muy próximas, dando esta aproximación un resultado muy satisfactorio.

Los factores de atenuación $g_{pre,ss1}(n)$ y $g_{pre,ss2}(n)$ determinados por sub-bloques pueden alisarse a continuación mediante una función de alisado aplicada muestra por muestra para evitar variaciones bruscas del factor de atenuación en las fronteras de los bloques. Esto es particularmente importante para la sub-señales que contienen componentes de baja frecuencia como la sub-señal $x_{rec,ss1}(n)$ pero no necesariamente para las sub-señales que no contienen más que componentes de alta frecuencia como la sub-señal $x_{rec,ss2}(n)$.

La figura 7 ilustra un ejemplo de aplicación de una ganancia de atenuación con unas funciones de alisado representadas por las flechas L.

Esta figura ilustra en a), un ejemplo de señal original, en b), la señal decodificada sin atenuación de pre-eco, en c), las ganancias de atenuación para las dos sub-señales obtenidas según la etapa de descomposición E605 y en d), la

señal decodificada con atenuación de pre-eco de las etapas E607 y E608 (es decir después de la combinación de las dos sub-señales atenuadas).

5 Se puede remarcar en esta figura que la ganancia de atenuación representada en trazos de puntos y correspondiente a la ganancia calculada para la primera sub-señal que incluye unos componentes de baja frecuencia, incluye funciones de alisado como se ha descrito anteriormente. La ganancia de atenuación representada en trazo continuo y calculada para la segunda sub-señal que incluye unas componentes de alta frecuencia, no incluye ganancia de alisado.

10 La señal representada en d) muestra claramente que el pre-eco se ha atenuado de manera eficaz para el tratamiento de atenuación implementado.

La función de alisado se define por ejemplo preferentemente por las ecuaciones siguientes:

$$15 \quad g_{pre,ss1}(n) = \frac{1}{u} \sum_{i=0}^{u-1} g_{pre,ss1}'(n-i), \quad n = 0, \dots, L-1$$

20 con la convención de que $g_{pre,ss1}'(n) \quad n = -(u-1), \dots, -1$ son los últimos $u-1$ factores de atenuación obtenidos para las últimas muestras del sub-bloque precedente de la sub-señal $x_{rec,ss1}(n)$. Típicamente $u = 5$ pero podría utilizarse otro valor. En función del alisado utilizado, la zona de pre-eco (el número de muestras atenuadas) puede por tanto ser diferente para las 2 sub-señales tratadas por separado, incluso si la detección del ataque se realiza en común sobre la base de la señal decodificada.

25 El factor de atenuación alisado no se eleva 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 reducido pero debe en cualquier caso evitarse. Para paliar este problema el valor del factor de atenuación puede forzarse a 1 para las $u-1$ muestras que preceden al índice pos en el que se sitúa el inicio del ataque. Esto es equivalente a avanzar el marcador pos en $u-1$ muestras para la sub-señal en la que se aplica el alisado. De ese modo la función de alisado incrementa progresivamente el factor para tener un valor 1 en el momento del ataque. Se preserva entonces la amplitud del ataque.

30 En este modo de realización con descomposición de la señal, la verificación del aumento de la energía de la zona de pre-eco según la invención se realiza para al menos una sub-señal o para cada una de estas sub-señales.

35 El umbral de comparación utilizado puede ser diferente según las sub-señales y según el número de sub-bloques disponibles antes del ataque.

Si en al menos una sub-señal el coeficiente director normalizado b_{1n} es inferior al umbral de esta sub-señal, se inhibe la atenuación de los pre-ecos para el conjunto de las sub-señales.

40 En caso de pre-eco en una señal procedente de una transformada MDCT inversa la energía del componente de pre-eco aumenta o es al menos estable en todas la sub-señales. La inhibición del tratamiento de pre-eco puede realizarse por ejemplo poniendo los factores de atenuación a 1 o no discriminando la zona como zona de pre-eco, no siendo solicitado entonces el módulo de tratamiento de atenuación del pre-eco como se ilustra a título de ejemplo en el modo de realización de la figura 5 por el enlace entre el bloque 604 y 602.

45 En unas variantes, la atenuación se inhibirá por separado para cada sub-señal a partir de que el coeficiente director normalizado b_{1n} sea inferior al umbral de esta sub-señal. La inhibición podrá ser implementada por ejemplo poniendo los factores de atenuación a 1 o no solicitando el módulo de pre-eco para la sub-señal considerada.

50 De ese modo, en el modo de realización particular descrito anteriormente con descomposición en dos sub-señales, si el número de sub-bloques antes del ataque permite hacer esta verificación, se verifica en las dos sub-señales la evolución de la energía de los sub-bloques que preceden al sub-bloque en el que se detecta el ataque, por regresión lineal. Esta verificación puede hacerse según las etapas E603 y E604, en no importa qué momento después de la división de la señal decodificada en sub-señales (E605) y antes de la aplicación de los factores de atenuación de los pre-ecos (E607). La verificación es posible si al menos dos sub-bloques preceden al sub-bloque en el que se ha detectado el ataque. Si el ataque se detecta en el primer o segundo sub-bloque la verificación según la invención no es posible.

60 En unas variantes, se podrá(n) reutilizar el (los) coeficiente(s) director(es) eventualmente calculado(s) en la trama precedente si el ataque se detecta en el primer o segundo sub-bloque de la trama actual.

Si el ataque se detecta en el tercer sub-bloque se dispone entonces de la energía de dos sub-bloques en la zona de pre-eco para hacer esta verificación. Por experiencia, con dos puntos, la verificación no es suficientemente fiable en la sub-señal de baja frecuencia $x_{rec,ss1}(n)$. Se verifica entonces únicamente la sub-señal de alta frecuencia $x_{rec,ss2}(n)$ y

únicamente que la energía no disminuye. El coeficiente director de la sub-señal de alta frecuencia $x_{rec,ss2}(n)$ se compara con el umbral de valor 0. Solo su signo es importante aquí, no es necesario hacer una normalización. Es suficiente por tanto calcular en la etapa E603 un coeficiente director simple (sin normalización) como:

$$b_{1ss2} = En_{ss2}(1) - En_{ss2}(0)$$

Si b_{1ss2} es inferior a 0, la atenuación de los pre-ecos para esta zona de pre-eco se inhibe para el conjunto de las sub-señales.

Si el ataque se detecta en el cuarto sub-bloque o en un sub-bloque de índice superior a 4, se verifica la evolución de la energía de los últimos 3 sub-bloques en la zona de pre-eco que precede al sub-bloque en el que se detecta el ataque. El coeficiente director de la sub-señal de baja frecuencia $x_{rec,ss1}(n)$ se compara con 0, solo su signo es importante y no es necesario normalizar este coeficiente. Es suficiente por tanto calcular un coeficiente director simple. Si el ataque se ha detectado en el sub-bloque de índice id con $id \geq 3$ se determina este coeficiente como:

$$b_{1ss1} = En(id - 1) - En_{ss2}(id - 3)$$

Si b_{1ss1} es inferior a 0 se inhibe la atenuación de los pre-ecos para esta zona de pre-eco, y para el conjunto de las sub-señales.

El coeficiente director de la sub-señal de alta frecuencia $x_{rec,ss2}(n)$ se compara con un umbral de valor 0,2. Se calcula el coeficiente director normalizado. Si el ataque se ha detectado en el sub-bloque del índice id con $id \geq 3$ se determina este coeficiente como:

$$b_{1nss2} = \frac{3(En_{ss2}(id - 1) - En_{ss2}(id - 2))}{2(En_{ss2}(id - 1) + En_{ss2}(id - 2) + En_{ss2}(id - 3))}$$

Si b_{1nss2} es inferior a 0,2 se inhibe la atenuación de los pre-ecos para esta zona de pre-eco, y para el conjunto de las sub-señales.

A observar que la condición

$$\frac{3(En_{ss2}(id - 1) - En_{ss2}(id - 2))}{2(En_{ss2}(id - 1) + En_{ss2}(id - 2) + En_{ss2}(id - 3))} < 0,2$$

es equivalente a

$$En_{ss2}(id - 1) - En_{ss2}(id - 2) < \frac{1}{7,5} (En_{ss2}(id - 1) + En_{ss2}(id - 2) + En_{ss2}(id - 3))$$

evitando así una operación de división para reducir la complejidad y para facilitar la implementación sobre un procesador DSP (por "Digital Signal Processor") de aritmética en coma fija.

El módulo 607 del dispositivo 600 de la figura 5 implementa la etapa E607 de atenuación de pre-eco en la zona de pre-eco de cada una de las sub-señales para aplicación a las sub-señales de los factores de atenuación así calculados.

La atenuación de pre-eco se realiza por tanto de manera independiente en las sub-señales. De ese modo, en las sub-señales que representan diferentes bandas de frecuencia, la atenuación puede elegirse en función del reparto espectral del pre-eco.

Finalmente, una etapa E608 del módulo de obtención 608 permite obtener la señal de salida atenuada (la señal decodificada después de la atenuación de pre-eco) por combinación (en este ejemplo por simple adición) de las sub-señales atenuadas, según la ecuación:

$$x_{rec,f}(n) = g_{pre,ss1}(n) x_{rec,ss1}(n) + g_{pre,ss2}(n) x_{rec,ss2}(n), n = 0, \dots, L-1$$

Contrariamente a una descomposición en sub-bandas clásica, se puede observar aquí que los filtrados utilizados no se asocian a operaciones de diezmado de las sub-señales y la complejidad y el retardo ("lookahead" o trama futura) se reducen al mínimo.

Se describe ahora con referencia a la figura 8 un ejemplo de realización de un dispositivo de discriminación y de tratamiento de atenuación según la invención.

5 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 incluye una memoria de almacenamiento y/o de trabajo, así como una memoria tampón MEM antes citada en tanto que medios para memorizar todos los datos necesarios para la implementación del procedimiento de discriminación y de tratamiento de atenuación tal como se describen con referencia a la figura 5. Este dispositivo recibe en la entrada unas tramas sucesivas de la señal digital Se y entrega la señal Sa reconstruida con atenuación de pre-eco en las zonas de pre-eco discriminadas con, si es necesario, reconstrucción de la señal atenuada por combinación de sub-señales atenuadas.

10 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 principalmente las etapas de cálculo de un coeficiente director de las energías para al menos dos sub-bloques que preceden al sub-bloque en el que se detecta un ataque, de comparación del coeficiente director con un umbral predefinido y de inhibición del tratamiento de atenuación de pre-eco en la zona de pre-eco en el caso de que el coeficiente director calculado sea inferior al umbral predefinido.

15 La figura 5 puede ilustrar el algoritmo de un programa informático de ese tipo.

Este dispositivo de discriminación y de tratamiento de atenuación según la invención puede ser independiente o estar integrado en un decodificador de señal digital. Un decodificador de ese tipo puede integrarse en unos equipos de almacenamiento o de transmisión de señales de audio digitales tales como pasarelas de comunicación, terminales de comunicación o servidores de una red de comunicación.

20

REIVINDICACIONES

1. Procedimiento de discriminación y de atenuación de pre-eco en una señal de audio digital generada a partir de una codificación por transformada, en el que, en la decodificación, para una trama actual descompuesta en sub-bloques, los sub-bloques de baja energía que preceden a un sub-bloque en el que se detecta una transición o ataque (E601) determinan una zona de pre-eco (E602) en la que se efectúa un tratamiento de atenuación de pre-eco (E607), estando el procedimiento caracterizado por que, en el caso en el que se detecta un ataque a partir del tercer sub-bloque de la trama actual, incluye las siguientes etapas:
- cálculo (E603) de un coeficiente director de las energías para al menos dos sub-bloques de la trama actual que preceden al sub-bloque en el que se detecta un ataque;
 - comparación (E604) del coeficiente director con un umbral predefinido; y
 - inhibición (E602) del tratamiento de atenuación de pre-eco en la zona de pre-eco en el caso de que el coeficiente director calculado sea inferior al umbral predefinido.
2. Procedimiento según la reivindicación 1, caracterizado por que incluye además una etapa de descomposición de la señal de audio digital en al menos dos sub-señales en función de un criterio de frecuencia y en el que las etapas de cálculo de comparación se efectúan para al menos una de las sub-señales.
3. Procedimiento según la reivindicación 1, caracterizado por que incluye además una etapa de descomposición de la señal de audio digital en al menos dos sub-señales en función de un criterio de frecuencia en el que las etapas de cálculo y de comparación se efectúan para cada una de las sub-señales, efectuándose la inhibición del tratamiento de atenuación del pre-eco en la zona de pre-eco de todas la sub-señales cuando un coeficiente director calculado es inferior al umbral predefinido para al menos una sub-señal.
4. Procedimiento según la reivindicación 3, caracterizado por que se define un umbral diferente por sub-señal.
5. Procedimiento según una de las reivindicaciones 1 a 4, caracterizado por que el coeficiente director se calcula según un método de estimación de mínimos cuadrados.
6. Procedimiento según una de las reivindicaciones 1 a 5, caracterizado por que el coeficiente director se normaliza.
7. Procedimiento según la reivindicación 1, caracterizado por que en el caso de que se detecte un ataque en el primer o segundo sub-bloque de la trama actual, se utiliza un coeficiente director calculado para la trama precedente para la etapa de comparación.
8. Dispositivo de discriminación y de atenuación de pre-eco en una señal de audio digital generada por un codificador por transformada, asociándose el dispositivo a un decodificador e incluyendo un módulo de detección (601) de transición o ataque, un módulo de discriminación de la zona de pre-eco (602) y un módulo de tratamiento de atenuación del pre-eco (607), efectuándose un tratamiento de atenuación del eco para una trama actual descompuesta en sub-bloques, en los sub-bloques de reducida energía que preceden a un sub-bloque en el que se detecta una transición o ataque que determinan una zona de pre-eco, estando el dispositivo caracterizado por que incluye además:
- un módulo de cálculo (603) que calcula un coeficiente director de las energías para al menos dos sub-bloques de la trama actual que preceden al sub-bloque en el que se detecta un ataque, en el caso de que se detecte un ataque a partir del tercer sub-bloque de la trama actual;
 - un comparador (604) adecuado para efectuar una comparación del coeficiente director con un umbral predefinido; y
 - un módulo de discriminación (602) adecuado para inhibir el tratamiento de atenuación de pre-eco en la zona de pre-eco en el caso de que el coeficiente director calculado sea inferior al umbral predefinido.
9. Decodificador de señal de audio digital que incluye un dispositivo de discriminación y de atenuación de pre-eco según la reivindicación 8.
10. Programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento según una de las reivindicaciones 1 a 7, cuando estas instrucciones se ejecutan por un procesador.
11. Soporte de almacenamiento legible por un dispositivo de discriminación y de tratamiento de atenuación de pre-eco en el que está registrado un programa informático que comprende unas instrucciones de código para la ejecución de las etapas del procedimiento de discriminación y de tratamiento de atenuación de pre-eco según una de las reivindicaciones 1 a 7 cuando dicho soporte de almacenamiento funciona en un ordenador.

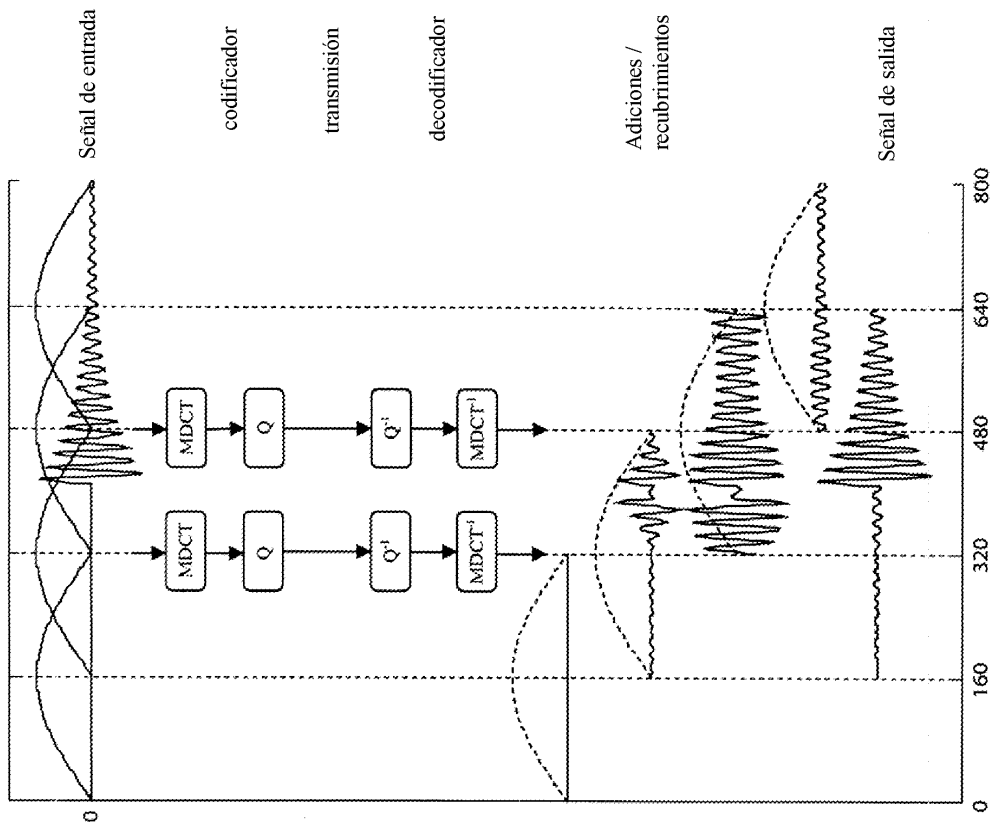


FIG.1

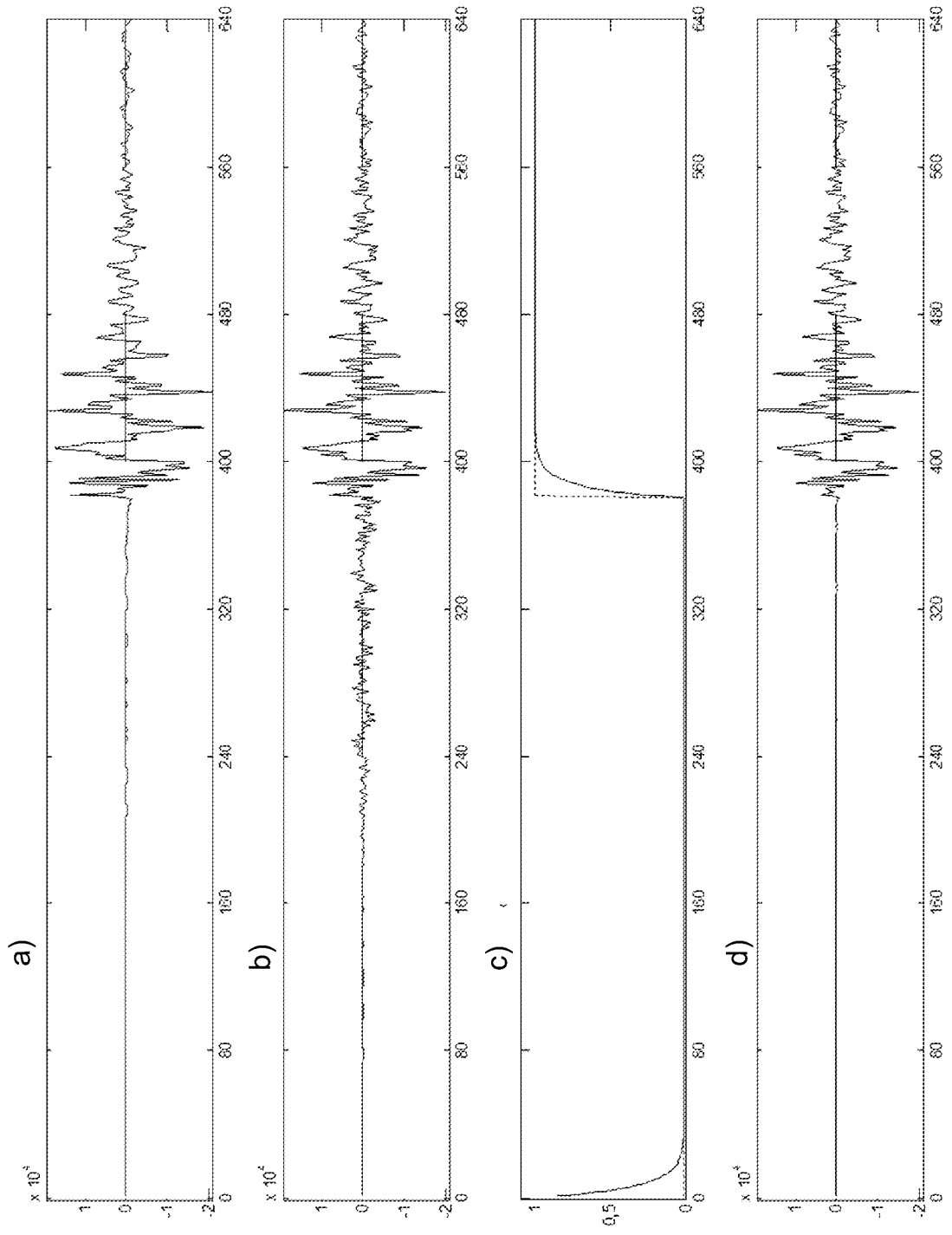


FIG.2

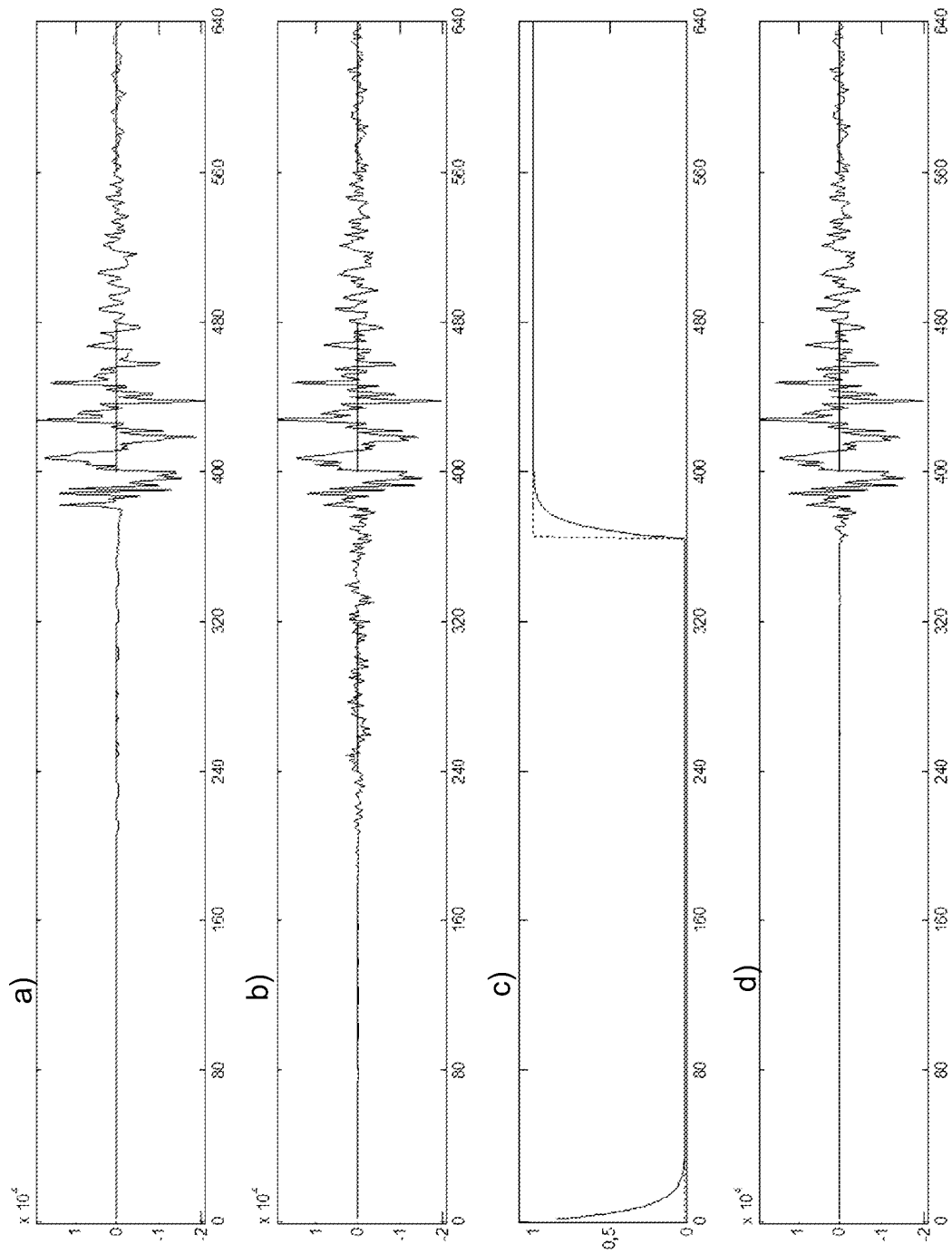


FIG.3

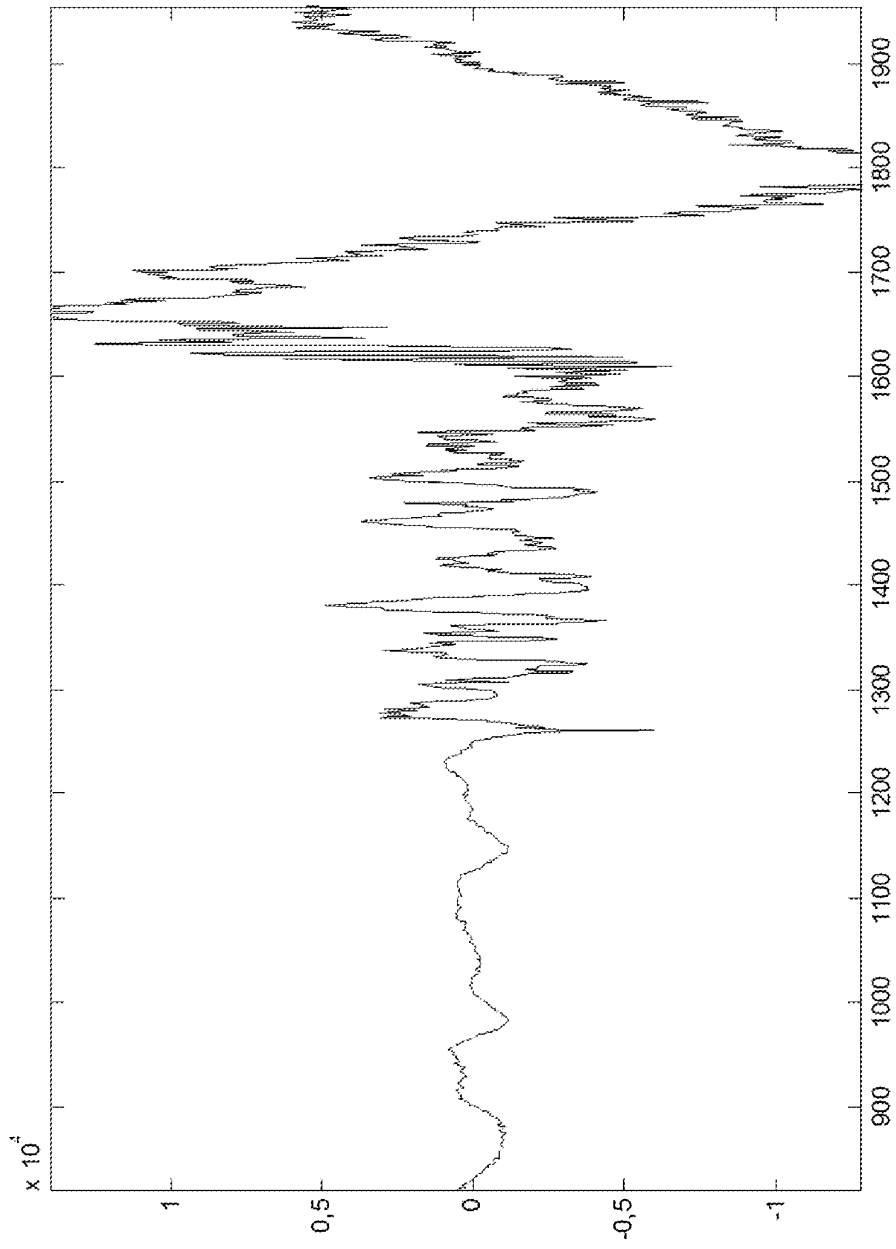


FIG.4

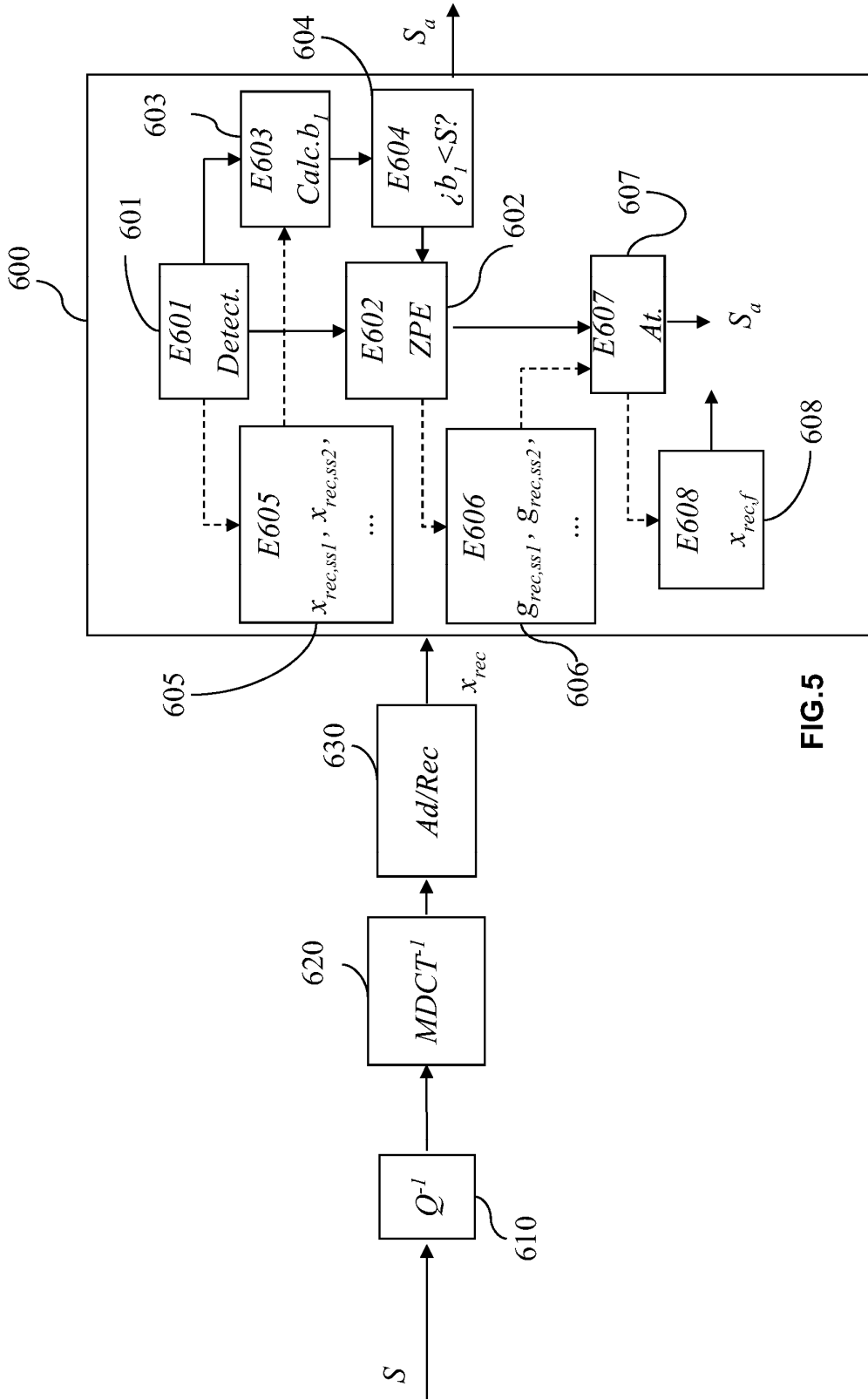


FIG. 5

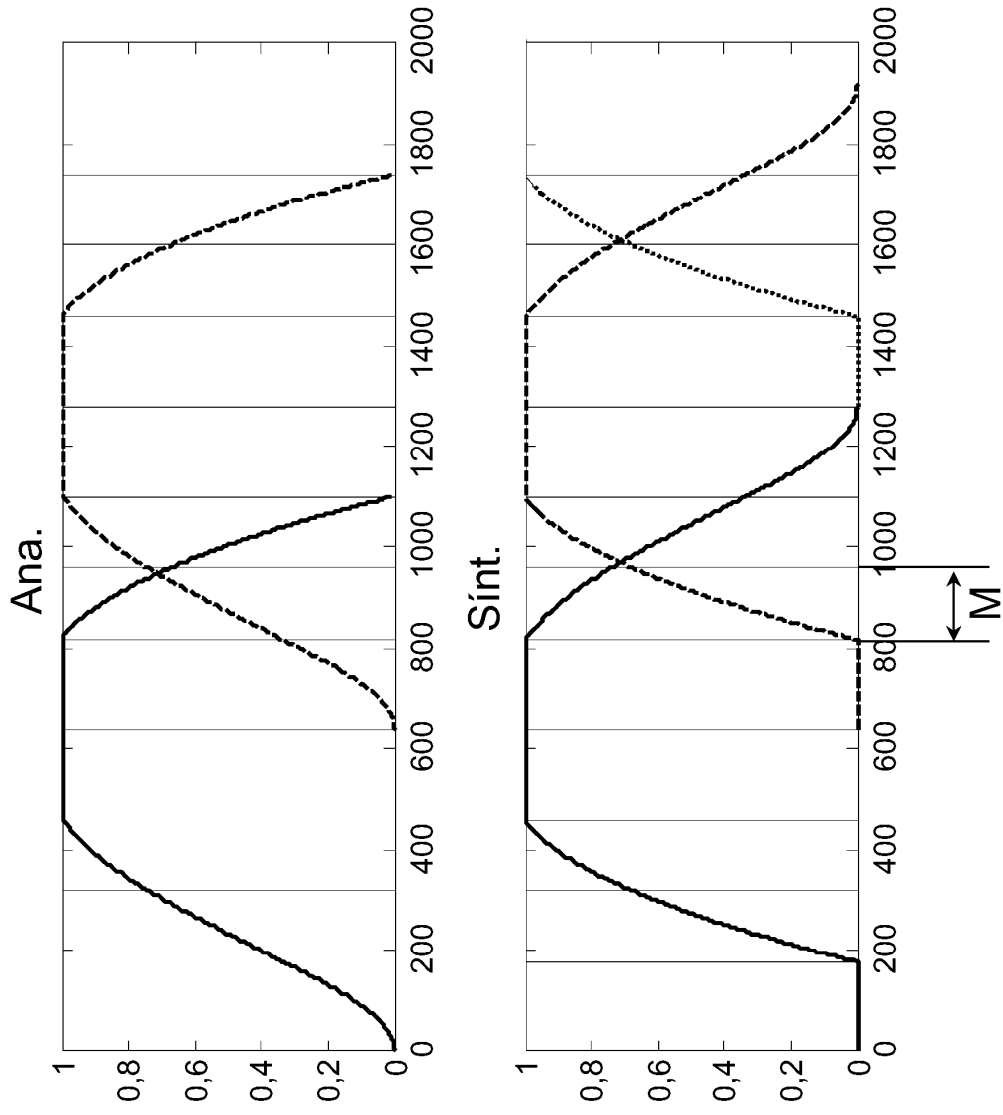


FIG.6

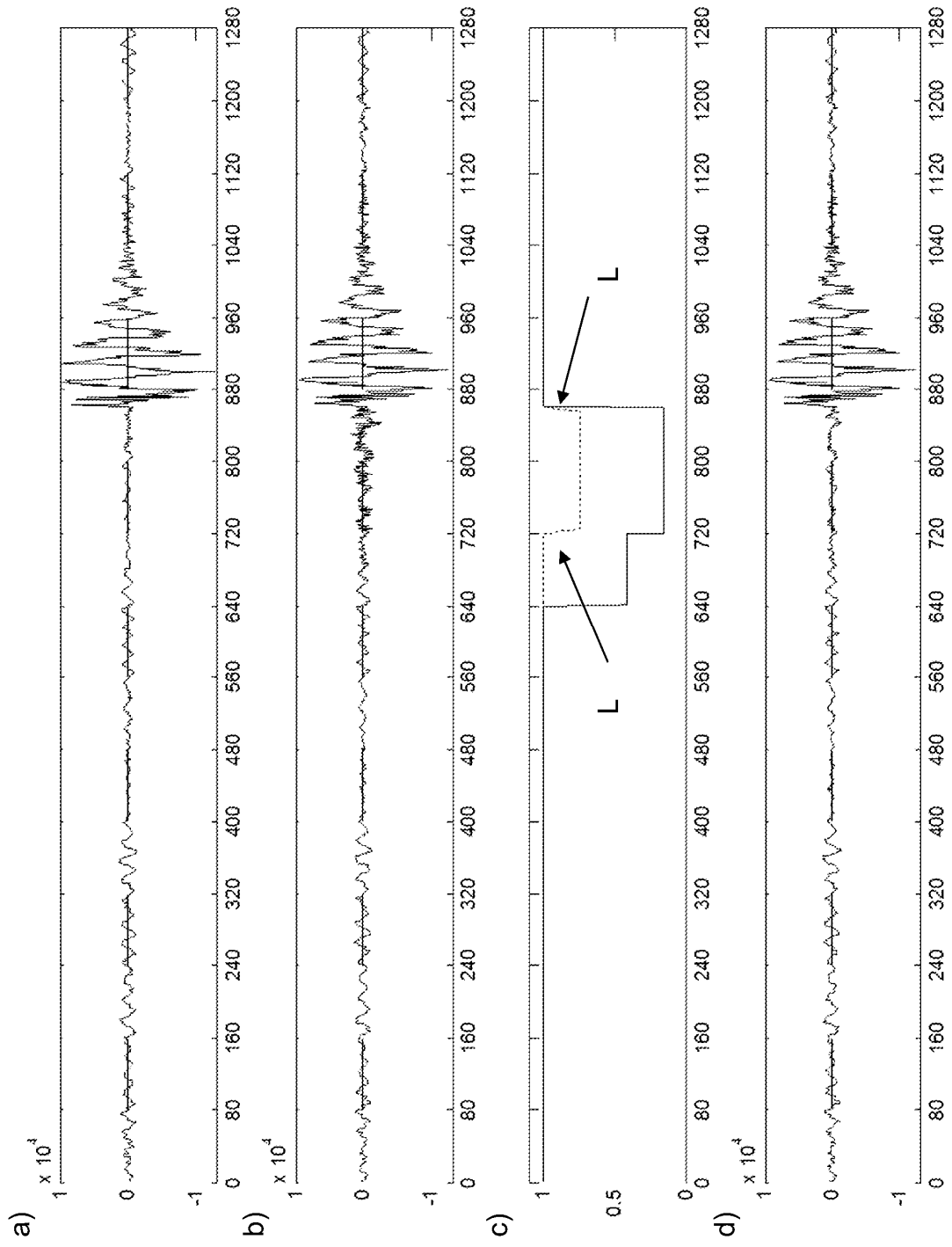


FIG.7

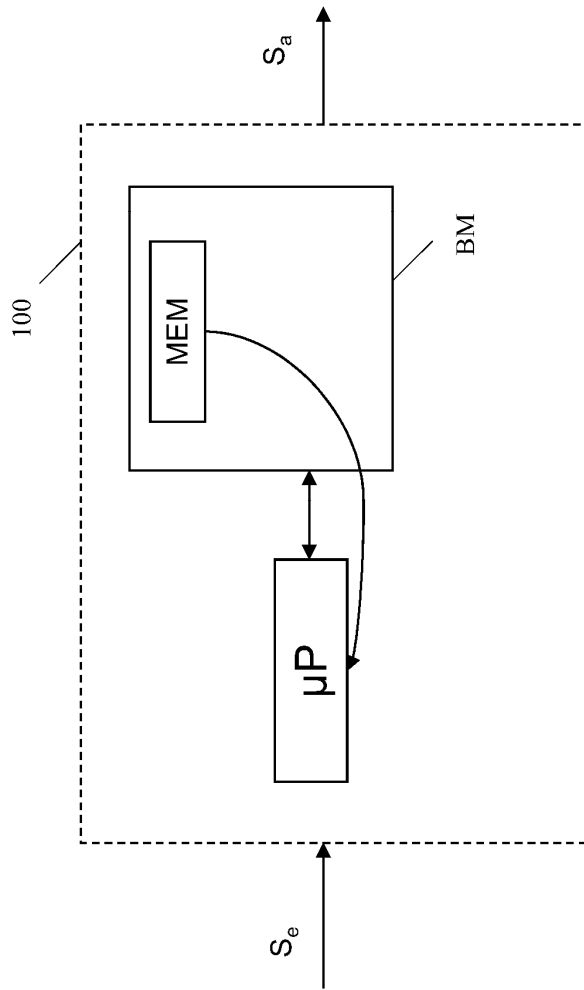


FIG.8