

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 612 385**

51 Int. Cl.:

G10L 19/025 (2013.01)

G10L 21/02 (2013.01)

H04B 3/21 (2006.01)

G10L 19/26 (2013.01)

G10L 21/0364 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.12.2013** **PCT/FR2013/053216**

87 Fecha y número de publicación internacional: **26.06.2014** **WO14096733**

96 Fecha de presentación y número de la solicitud europea: **20.12.2013** **E 13820844 (2)**

97 Fecha y número de publicación de la concesión europea: **09.11.2016** **EP 2936488**

54 Título: **Atenuación eficaz de pre-ecos en una señal de audio digital**

30 Prioridad:

21.12.2012 FR 1262598

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

16.05.2017

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 2 612 385 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 pre-ecos en una señal de audio digital

- 5 La invención se refiere a un procedimiento y un dispositivo de tratamiento de atenuación de los pre-ecos durante una decodificación de la 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 el almacenamiento de las señales, se recurre a unos procedimientos de compresión (o codificación de la fuente) que implementan unos sistemas de codificación que son en general de tipo de codificación temporal por predicción lineal o codificación en frecuencia por transformada.

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

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

20 Ciertas secuencias musicales, tales como las percusiones y ciertos segmentos de la voz, como las plosivas (/k/, /t/,...), se caracterizan 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 tratamiento de codificación/decodificación, la señal de entrada se recorta en bloques de muestras de longitud L cuyos límites se representan en la figura 1 por unos trazados verticales discontinuos. La señal de entrada se indica por $x(n)$, en la que n es el índice de la muestra. El recorte 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 las que N es el índice del bloque (o de la trama), L es la longitud de la trama. En la figura 1 se tiene una $L = 160$ 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 por 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 enturbiada con "ruido" (o distorsión) engendrado 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 del ruido de codificación es típicamente inferior al de la señal para los segmentos de fuerte energía que siguen inmediatamente a la transición, pero nivel es superior al de la señal para los segmentos de energía más reducida, particularmente en 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 como muy molesta para la escucha. Se denomina pre-eco al ruido de codificación anterior a la transición y post-eco al ruido posterior a la transición.

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

55 Las experiencias psico-acústicas han demostrado que el oído humano efectúa un pre-enmascarado temporal de los sonidos bastante limitado, del orden de algunos milisegundos. El ruido que precede 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 secuencias de fuerte energía a unas secuencias de energía reducida. La tasa o nivel de molestia aceptable para los post-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 sea la longitud de los bloques en número de muestras. Ahora bien, en codificación por transformada, es bien conocido que para las señales estacionarias cuanto más aumenta la longitud de la transformada, mayor es la ganancia de codificación. A

frecuencia de muestreo fija y a tasa fija, si se incrementa 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 que se juzgan útiles por el modelo psico-acústico, de ahí 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, 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 en este caso permitiendo conmutar desde estas ventanas largas a 8 ventanas cortas por medio de ventanas intermedias (denominadas de transición), lo que necesita de 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 tanto de 256 muestras (8 ms a 32 kHz). Con baja tasa se puede tener siempre 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 una ventana de duración de 40 ms a 16, 32 o 48 kHz (respectivamente). Se puede observar que el codificador UIT-T G.719 integra un mecanismo de conmutación de ventanas con detección de transitorios, sin embargo el pre-eco no se reduce completamente a baja velocidad (típicamente a 32 kb/s).

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

Se ha citado anteriormente la conmutación de ventanas; 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 que precede al ataque, la señal reconstruida se ve como la suma de la señal original y el 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, publicado 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 varianza de la señal corrompida por el pre-eco, 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 ser conocidas más que en el codificador y deben transmitirse en consecuencia. Esto hace necesario transmitir unas informaciones suplementarias, lo que a una velocidad limitada disminuye los recursos relativos asignados 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 mencionado no permite recuperar la señal de origen, sino que 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 de 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 los 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 sub-bloque k -ésimo se calcula por ejemplo en función de la relación $R(k)$ entre la energía del sub-bloque de energía mayor y la energía del k -ésimo sub-bloque en cuestión:

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

en la que f es una función decreciente a 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 necesario entonces ninguna atenuación; el factor $g(k)$ se fija 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 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 tras el tratamiento de atenuación del pre-eco la energía de la señal se haga 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 por $\overline{En}$, 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 precedente al sub-bloque a tratar. Este valor está por supuesto limitado a un máximo de 1 puesto que se trata en este caso de los valores de atenuación. Más precisamente se define aquí:

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

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

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:

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

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

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

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

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

La función se alisa a continuación según la ecuación siguiente:

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

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 $\alpha = 0,85$.

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

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

En donde $g_{pre}'(n)$ 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 por ejemplo tomar $u = 5$.

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:

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

donde $x_{rec,g}(n)$ es la señal decodificada y post-tratada por la reducción del pre-eco.

Las figuras 2 y 3 ilustran la implementación del procedimiento de atenuación tal como se ha descrito en la solicitud de patente del estado de la técnica, antes mencionada, y resumida anteriormente.

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. Se sitúa un ataque (o transición) en la señal en el sub-bloque que comienza en el índice 320. Esta señal se ha codificado mediante un codificador por transformada de tipo MDCT con baja tasa (24 kb/s).

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 del pre-eco (línea continua) obtenido mediante el procedimiento descrito en la solicitud de patente del estado de la técnica antes mencionada. La línea de puntos representa el factor antes de alisado. Se observa 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) por 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 las algunas muestras del sub-bloque precedente al sub-bloque en donde 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 del 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. La amplitud del ataque es entonces preservada, como se ilustra en la parte d) de la figura 3, por el contrario algunas muestras del pre-eco no se atenúan.

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

Otro ejemplo con el mismo ajuste que el de la figura 3 se ilustra en la figura 4. Esta figura representa 2 tramas para mostrar mejor la naturaleza de la señal antes del ataque. En este caso, la energía de la señal original antes del ataque es mayor (parte a)) que en el caso ilustrado en la figura 3, y la señal antes del ataque es audible (muestras 0-850). En la parte b) puede observarse el pre-eco en la señal decodificada sin tratamiento de pre-eco en la zona 700-850. Según el procedimiento de limitación de la atenuación explicado anteriormente se atenúa la energía de la señal de la zona de pre-eco hasta la energía media de la señal precedente a la zona de tratamiento. Se observa en la parte c) que el factor de atenuación calculado teniendo en cuenta la limitación de energía es próximo a 1 y que el pre-eco está siempre presente sobre la parte d) después de la aplicación del tratamiento de pre-eco (multiplicación de la señal b) con la señal c)), a pesar de la buena nivelación de la señal en la zona de pre-eco. Se puede distinguir bien en efecto este pre-eco en la forma de onda en donde se remarca que en esta zona se superpone a la señal una componente de alta frecuencia.

Este componente de alta frecuencia es bien audible y molesto, y el ataque es menos neto (parte d) de la figura 4).

La explicación de este fenómeno es la siguiente: en el caso de un ataque muy brusco, impulsivo (como el ilustrado en la figura 4) el espectro de la señal (en la trama que contiene el ataque) es más bien blanco y por tanto contiene igualmente muchas altas frecuencias. De ese modo el ruido de cuantificación está igualmente repartido y relativamente plano en frecuencias (blanco) y compuesto de altas frecuencias, lo que no es el caso de la señal precedente a la zona de pre-eco. Hay por tanto un cambio brusco en el espectro de una trama a otra, lo que da como resultado un pre-eco audible a pesar del hecho de que la energía se ha nivelado correctamente.

Este fenómeno se ha representado de nuevo en las figuras 5a y 5b que muestran respectivamente los espectrogramas de la señal original en 5a, que corresponden a la señal representada en la parte a) de la figura 4 y el espectrograma de la señal con atenuación de pre-ecos según el estado de la técnica, en 5b, que corresponde a la señal representada en la parte d) de la figura 4.

Se observa bien un pre-eco aún audible en la parte recuadrada en la figura 5b.

Otros ejemplos de estos procedimientos de atenuación de pre-ecos se describen en la solicitud de patente FR2897733 y FR2888704, así como en la ITU-T Recommendation G.729.1, "G.729 based Embedded Variable bit-rate coder: An 8-32 kb/s scalable wideband coder bitstream interoperable with G.729" (05/2006).

Existe por tanto una necesidad de una técnica mejorada de atenuación de pre-ecos en la decodificación, que permita atenuar las altas frecuencias indeseables y más generalmente los pre-ecos parásitos de manera precisa y universal y sin que sea transmitida por el codificador ninguna información auxiliar.

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

Con este fin, la presente invención se refiere a un procedimiento de tratamiento de la atenuación de pre-eco en una señal de audio digital decodificada según una decodificación por transformada. El procedimiento es tal que incluye las siguientes etapas:

- descomponer la señal decodificada en al menos dos sub-señales según un criterio de descomposición predeterminado;
- calcular factores de atenuación por sub-señal y por muestra de una zona de pre-eco previamente determinada;

- atenuar el pre-eco en la zona de pre-eco de cada una de las sub-señales mediante aplicación de los factores de atenuación a las sub-señales; y
- obtener la señal atenuada por combinación de las sub-señales atenuadas.

5 De ese modo, el procedimiento permite controlar de manera precisa, la atenuación a aplicar a cada una de las sub-señales. Estas sub-señales tienen unas componentes representativas de características particulares de la señal decodificada, según el criterio de descomposición elegido. De ese modo, puede adaptarse la cantidad de atenuación a aplicar para estas diferentes características. La atenuación de pre-eco en esta señal decodificada es entonces más precisa y más eficaz.

10 Los diferentes modos particulares de realización mencionados en el presente documento a continuación pueden añadirse independientemente o en combinación unos con los otros, a las etapas del procedimiento definido anteriormente en el presente documento.

15 En un primer modo de realización, el criterio de descomposición predeterminado es un criterio de frecuencia.

De ese modo, la atenuación se adapta de manera precisa a las características de frecuencia de la señal decodificada.

20 En un modo particular de realización, la descomposición de la señal decodificada se efectúa mediante un primer filtrado paso bajo o paso alto para obtener una primera sub-señal.

25 Una primera sub-señal incluye por tanto unas componentes de baja frecuencia en el caso de filtrado paso bajo o unas componentes de alta frecuencia en el caso de un filtrado paso alto. La atenuación para esta primera sub-señal está adaptada a sus componentes de frecuencia.

Según un modo posible de realización, la descomposición de la señal se efectúa además mediante un segundo filtrado paso alto o paso bajo, complementario al primer filtrado, para obtener una segunda sub-señal.

30 Una segunda sub-señal incluye por tanto unas componentes de alta frecuencia en el caso de un filtro paso alto complementario o unas componentes de baja frecuencia en el caso de filtrado paso bajo complementario. La atenuación para esta segunda sub-señal se adapta también a sus componentes de frecuencia.

35 Según otro modo posible de realización, se obtiene una segunda sub-señal por resta de la primera sub-señal a la señal decodificada, de manera que se evite un segundo filtrado para obtener la segunda sub-señal. Esto reduce por tanto la complejidad de la etapa de descomposición del procedimiento.

De manera particular, el filtrado es una respuesta a impulsos finita de fase nula de la función de transferencia:

$$40 \quad c(n)z^{-1} + (1-2c(n)) + c(n)z$$

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

Este tipo de filtrado es de reducida complejidad.

45 En una variante de realización, la descomposición de la señal decodificada se efectúa por filtrado QMF y PQMF para obtener unas sub-señales en sub-bandas.

50 Se obtienen así varias sub-señales, estando representada cada una de las sub-señales en una banda de frecuencia diferente. La atenuación resultante permite tener en cuenta de ese modo el reparto espectral del pre-eco. La atenuación está adaptada entonces a estas características espectrales.

En un segundo modo de realización, el criterio de descomposición predeterminado es un criterio de periodicidad de la señal.

55 En este modo de realización, la atenuación se adapta a las características de periodicidad de la señal. Una primera sub-señal que incluye por ejemplo unas componentes periódicas de tipo sinusoidal y una segunda sub-señal que incluye unas componentes de ruido.

60 Según un modo posible de realización, la zona de pre-eco se determina según las siguientes etapas:

- detectar una posición de ataque en la señal decodificada, antes de la etapa de descomposición;
- determinar una zona de pre-eco que precede a la posición de ataque detectada en la señal decodificada antes de la etapa de descomposición o en las sub-señales después de la etapa de descomposición.

65

De ese modo, la etapa de detección de la posición de ataque se uniformiza para todas las sub-señales, permitiendo de ese modo reducir la complejidad de tratamiento. La etapa de determinación de la zona de pre-eco puede uniformizarse igualmente en interés de la reducida complejidad o puede realizarse en las sub-señales para un buen compromiso complejidad reducida/adaptación a la señal.

Según otro modo posible de realización, la zona de pre-eco se determina según las siguientes etapas:

- detectar una posición de ataque en cada una de las sub-señales obtenidas después de la etapa de descomposición;
- determinar una zona de pre-eco que precede a la posición de ataque detectada en cada una de las sub-señales.

La detección de la posición de ataque puede efectuarse de diferente manera según la sub-señal a la que se aplica. Esto permite adaptar mejor la detección en función de la sub-señal, en detrimento de una complejidad de tratamiento mayor.

De manera que se simplifique aún más el procedimiento según un modo de realización de la invención, al menos para una sub-señal, el cálculo del factor de atenuación se efectúa utilizando al menos un parámetro de la etapa de detección de una posición de ataque en la señal decodificada.

Este puede ser en efecto el caso para la sub-señal que incluye unas componentes de baja frecuencia que son generalmente mucho más energéticas de las componentes de alta frecuencia, estando entonces muy próximas las energías por sub-bloque de la señal decodificada $x_{rec}(n)$ y de la sub-señal que incluye mayoritariamente unas componentes de baja frecuencia $x_{rec,ss1}(n)$. No es por tanto necesario recalcular unas energías de los sub-bloques para calcular el factor de atenuación de la sub-señal porque para esta señal los valores de energía por sub-bloque ya se han calculado para detectar los ataques.

En un modo particular de realización, se aplica una etapa de alisado del factor de atenuación al menos para una sub-señal.

Este alisado permite evitar las variaciones bruscas del factor de atenuación en los límites de los bloques de muestras.

En un modo ventajoso de realización, el factor de atenuación se fuerza a 1 para un número predeterminado de muestras que preceden al inicio del ataque.

De ese modo, el factor de atenuación tiene un valor de 1 en el momento del ataque, lo que permite preservar la amplitud del ataque.

La invención se refiere igualmente a un dispositivo de tratamiento de atenuación de pre-eco en una señal de audio digital decodificada a partir de un decodificador por transformada. Este dispositivo es tal que comprende:

- un módulo de descomposición de la señal decodificada en al menos dos sub-señales según un criterio de descomposición predeterminado;
- un módulo de cálculo de los factores de atenuación por sub-señal y por muestra de una zona de pre-eco previamente determinada;
- un módulo de atenuación de pre-eco en la zona de pre-eco de cada una de las sub-señales por aplicación de los factores de atenuación a las sub-señales;
- un módulo de obtención de la señal atenuada por combinación de las sub-señales atenuadas.

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

La invención se refiere a un decodificador de una señal de audio digital que incluye un dispositivo tal como el descrito anteriormente.

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

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 como el descrito anteriormente.

Surgirán 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 en referencia a los dibujos adjuntos, en los que:

- 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;
- 5 - la figura 3 descrita anteriormente ilustra otro ejemplo de señal de audio digital para el que se efectúa un método de atenuación según el estado de la técnica;
- la figura 4 descrita anteriormente ilustra otro ejemplo más de señal de audio digital para el que se aplica 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 pre-ecos según el estado de la técnica (correspondiente respectivamente las partes a) y d) de la figura 4);
- 10 - la figura 6 ilustra un dispositivo de tratamiento de atenuación de pre-ecos en un decodificador de señal de audio digital, así como las etapas implementadas por el procedimiento de tratamiento según un modo de realización de la invención;
- 15 - la figura 7a ilustra la respuesta en frecuencia de un filtro paso bajo implementado según un modo de realización de la invención en la etapa de descomposición de la señal, para diferentes valores de parámetros del filtro;
- la figura 7b ilustra la respuesta en frecuencia de un filtro paso alto implementado según un modo de realización de la invención en la etapa de descomposición de la señal, para diferentes valores de parámetros del filtro;
- 20 - la figura 8 ilustra un ejemplo de ventanas de análisis y de ventanas de síntesis con reducido retardo para la codificación y la decodificación por transformada susceptible de crear el fenómeno de pre-eco;
- la figura 9 ilustra un ejemplo de señal de audio digital para el que se implementa el método de atenuación de pre-ecos según la invención;
- la figura 10 ilustra un ejemplo de descomposición de una señal de audio digital según la invención para la implementación del tratamiento de atenuación de pre-eco;
- 25 - la figura 11 retoma las señales de la figura 10 ampliándola sobre las muestras próximas a la zona de ataque;
- las figuras 12a, 12b y 12c ilustran los espectrogramas respectivos de una señal de audio digital de origen, de la señal decodificada sin tratamiento de atenuación de pre-eco y de la señal decodificada con tratamiento de atenuación de pre-eco según la invención;
- la figura 13 ilustra un ejemplo material de dispositivo de tratamiento de atenuación según la invención.

En referencia a la figura 6, se describe un dispositivo 600 de tratamiento de atenuación de pre-eco. El dispositivo de tratamiento de atenuación 600, tal como se describe en el presente documento 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 (add/rec) tal como se ha descrito en referencia a la figura 1 y que suministra una señal reconstruida $x_{rec}(n)$ al dispositivo 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 voz y de audio, sin embargo el dispositivo 600 se aplica igualmente a cualquier otro tipo de transformada (FFT, DCT, etc.).

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

El dispositivo 600 implementa un método de atenuación de los pre-ecos en la señal decodificada $x_{rec}(n)$ que se descompone en sub-señales —esta descomposición es propia del tratamiento de los pre-ecos—.

En un modo de realización de la invención, el procedimiento de tratamiento de atenuación incluye una etapa de detección (E601) de los ataques que pueden engendrar un pre-eco, en la señal decodificada $x_{rec}(n)$. Se puede observar que el dispositivo de atenuación 600 puede aplicarse también tanto en el caso de que la división en ventanas MDCT de síntesis sea fija como en aquellos en los que la división en ventanas sea adaptativa.

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.

Un ataque (u *onset* en inglés) es una transición rápida y una variación brusca de la dinámica (o amplitud) de la señal. Se pueden designar este tipo de señales mediante 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.

En este modo de realización, ilustrado por las flechas en trazos mixtos “a”, la detección del pre-eco se realiza sobre la señal decodificada antes de la etapa de descomposición (E603) en sub-señales del módulo 603 descrito posteriormente.

Cada trama actual de L muestras de señal decodificada $x_{rec}(n)$ se divide en K sub-bloques de longitud L' , por ejemplo con $L = 640$ 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.

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 tales ventanas se ilustra en referencia a la figura 8. El retardo engendrado por la transformación es de únicamente 192 muestras contrariamente al retardo de 640 muestras en el caso de utilización de las ventanas sinusoidales clásicas. De ese modo la memoria MDCT con unas ventanas de análisis-síntesis especiales de bajo retardo no contiene más que 96 muestras independientes (no replicadas con la trama actual) contrariamente a las 320 muestras en el caso de utilización de las ventanas sinusoidales clásicas.

Se puede en efecto observar en la figura 8 para las ventanas de análisis (Ana.), que la zona de repliegue está limitada por los trazos discontinuos entre las muestras 864 y 1055. La línea de repliegue se representa en trazo mixto en la muestra 960.

Para la síntesis (Sínt.), solo son necesarias las muestras representadas por el intervalo M (96 muestras) para obtener la información sobre la zona de repliegue del análisis, aprovechando la simetría. Estas muestras contenidas en la memoria son útiles entonces para decodificar esta zona de repliegue utilizando igualmente las muestras repliegadas de la ventana de la trama siguiente. En el caso de un ataque en esta zona entre las muestras 864 y 1055 la energía media de las muestras representadas por el intervalo M es netamente superior a la energía de las sub-tramas que preceden a la muestra 864. El incremento brusco de la energía del intervalo M contenido en la memoria MDCT puede por tanto señalar un ataque en la trama siguiente que puede engendrar un pre-eco en la trama actual.

En una variante de este modo de realización, pueden utilizarse otras ventanas de análisis/síntesis, o pueden utilizarse unas comunicaciones entre ventanas largas y cortas.

Se utiliza la memoria MDCT $x_{MDCT}(n)$ que da una versión con repliegue temporal ("folding" en inglés) de la señal futura. Esta memoria o una parte de esta memoria (a causa de la redundancia) se divide también en K' sub-bloques de longitud $L_m(n)$, $n = 0, \dots, K'$ en donde $L_m(n)$ es comparable pero no forzosamente idéntica a L' . En el caso de utilización de ventanas sinusoidales clásicas, no se requieren más que los $K' = 4$ primeros sub-bloques de longitud $L_m(n) = L' = 80$ de la señal $x_{MDCT}(n)$ porque los 4 últimos sub-bloques son simétricos con relación a los 4 primeros sub-bloques y no transportan así ninguna información suplementaria para la detección de un ataque en la trama siguiente. Por el contrario, si se utilizan unas ventanas de análisis-síntesis especiales de bajo retardo como se ha ilustrado en la figura 8, no se retiene más que un ($K' = 1$) bloque de longitud $L_m(0) = 96$ que contiene todas las muestras independientes de la memoria MDCT. A pesar del número de muestras superior en este sub-bloque, su energía permanece comparable con la de los sub-bloques de la trama actual porque la parte de memoria se ha enmarcado (por tanto atenuado) por la ventana de análisis. Se puede observar que el número de muestras es de 96 (en lugar de 80) y que la división en ventanas podría compensarse; la invención se aplica paralelamente en el caso en el que se implementa una compensación de ese tipo.

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

Puede observarse que la señal contenida en la memoria MDCT incluye un repliegue temporal (que se compensa cuando se recibe la trama siguiente). Como se explica en la presente memoria a continuación, la señal $x_{MDCT}(n)$ en la memoria MDCT sirve aquí esencialmente para estimar la energía por sub-bloque de la señal en la trama siguiente (futura) y se considera que esta estimación es suficientemente precisa para las necesidades de la detección y de la atenuación de pre-eco cuando se realiza a partir de la memoria MDCT disponible en la trama actual en lugar de la señal completamente decodificada en la trama futura.

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

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

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

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

cuando el sub-bloque está en la memoria MDCT (que representa la señal disponible para la trama futura) y $L_{cm}(i)$ contiene los límites de los sub-bloques de la parte en memoria:

$$L_{cm}(i) = \begin{cases} 0 & \text{si } i = 0 \\ \sum_{k=0}^{i-1} L_m(k) & \text{si } i > 0 \end{cases}$$

5

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

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

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

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

15 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 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

20

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

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

25

En las variantes de realización con conmutación de las ventanas —dicho de otra manera con una división de ventanas MDCT adaptativa—, pueden utilizarse otros métodos que dan la posición del ataque con una precisión que va desde la escala de un sub-bloque hasta una posición próxima a la muestra.

30 En el ejemplo dado en el presente documento anteriormente la detección de los ataques que pueden engendrar un pre-eco se realiza sobre la señal decodificada, común para todas las sub-señales. En una variante de realización, esta detección se puede realizar por separado en las sub-señales obtenidas a continuación de la etapa E603 descrita posteriormente. Este caso se ilustra por las flechas en trazos discontinuos "b".

35 El dispositivo 600 incluye igualmente un módulo 602 de determinación que implementa una etapa de determinación (E602) de una zona de pre-eco (ZPE) que precede a la posición de ataque detectada. Se denomina en este caso zona de pre-eco a la zona que cubre las muestras antes de la posición estimada del ataque que son perturbadas por el pre-eco engendrado por el ataque y donde es deseable la atenuación de este pre-eco. Según la invención, la zona de pre-eco puede determinarse sobre la señal decodificada (flechas a de la figura 6). En una variante de realización, puede determinarse por separado para cada sub-señal.

40

En esta variante de realización, la zona de pre-eco se determina independientemente para las sub-señales obtenidas por la etapa E603 descrita posteriormente, según el modo de realización descrito por las flechas b de la figura 6. De ese modo la zona de pre-eco puede ser diferente para las diferentes sub-señales.

45

Esta diferencia se ilustra por ejemplo en las figuras 10 y 11 que representan en a) un ejemplo de señal decodificada (antes de la atenuación de pre-eco), en b), una primera sub-señal con unas componentes mayoritariamente de baja frecuencia (sin atenuación de pre-eco), en c), una segunda sub-señal con unos componentes mayoritariamente de alta frecuencia (sin atenuación de pre-eco) y en d), la segunda sub-señal después de la atenuación de pre-eco según la invención. La figura 11 retoma las mismas señales pero con una ampliación sobre las muestras que van de la 560 a 1040.

50

Se puede observar en estas figuras que el pre-eco es más importante en la segunda sub-señal que en la primera. La zona de pre-eco de las dos sub-señales ilustradas en este caso, es por tanto muy diferente. Como resultado de ello, en la figura 9 c) que ilustra las ganancias de atenuación obtenidas para las dos sub-señales, se puede observar que las ganancias de atenuación (en puntos para la primera sub-señal y en trazo continuo para la segunda) son tales

55

que para la segunda sub-señal con unos componentes de alta frecuencia, se atenúan más muestras (muestras 640 a 860) que para la primera sub-señal con unas componentes de baja frecuencia (muestras 720 a 860).

En un modo de realización de obtención de las zonas de pre-ecos, las energías $En(k)$ están concatenadas en orden cronológico, inicialmente con la envolvente temporal de la señal decodificada, posteriormente la envolvente de la señal de la trama siguiente estimada 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 pre-eco por ejemplo si la relación $R(k)$ sobrepasa un umbral, típicamente este umbral es 32.

Los sub-bloques en los que se ha detectado un pre-eco constituyen así una zona de pre-eco, que en general cubre las muestras $n = 0, \dots, pos-1$, es decir desde el inicio de la trama actual a la posición del ataque (pos).

En las variantes de realización, la zona de pre-eco no comienza necesariamente al inicio de la trama, y puede hacer intervenir una estimación de la longitud del pre-eco. Si se utiliza una conmutación de ventanas, la zona de pre-eco deberá definirse para tener en cuenta las ventanas utilizadas. Se puede observar así que la zona de pre-eco puede extenderse perfectamente en toda la trama actual si el ataque se ha detectado en la trama futura.

El dispositivo 600 incluye un módulo 603 de descomposición de la señal, adecuado para efectuar una etapa E603 de descomposición de la señal decodificada en al menos dos sub-señales según un criterio predeterminado.

Este criterio puede ser por ejemplo un criterio de frecuencia que permita así obtener unas sub-señales con unos componentes de frecuencias diferentes como se ha ilustrado en las figuras 10 y 11 descritas anteriormente.

El criterio puede ser por ejemplo un criterio de periodicidad de la señal que da así unas sub-señales con unos componentes representativos de la periodicidad de la señal o por el contrario de su ruido.

En una realización particular de la invención, la señal decodificada $x_{rec}(n)$ se descompone en la etapa E603 en dos sub-señales de la siguiente manera:

- la primera sub-señal $x_{rec,ss1}(n)$ se obtiene mediante 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, donde $[c(n), 1-2c(n), c(n)]$ son los coeficientes del filtro paso bajo; este filtro se implementa con la ecuación de 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 $c(n) = 0,25$. La respuesta en frecuencia de este filtro se ilustra en la figura 7a, en función de los coeficientes $[c(n), 1-2c(n), c(n)]$, para $c(n) = 0,05, 0,1, 0,15, 0,2$ y $0,25$.

Se puede observar 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 mediante 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 donde $[-c(n), 2c(n), -c(n)]$ son los coeficientes del filtro paso alto; este filtro se implementa con la ecuación de las diferencias:

$$x_{rec,ss2}(n) = -c(n) x_{rec}(n-1) + 2c(n) x_{rec}(n) - c(n) x(n+1)$$

La respuesta en frecuencia de este filtro se ilustra en la figura 7b, en función de los coeficientes $[-c(n), 2c(n), -c(n)]$, para $c(n) = 0,05, 0,1, 0,15, 0,2$ y $0,25$. La sub-señal $x_{rec,ss2}(n)$ resultante de este filtrado, contiene por tanto unas componentes más bien de alta frecuencia de la señal decodificada.

La motivación para utilizar estos filtros es su muy reducida complejidad de implementación (en términos de número de cálculos por muestra), su fase nula y el reducido retardo asociado (una muestra) —a observar que estos filtros son de hecho utilizados en retardos, debido a que la señal futura ("lookahead" en inglés) se reduce a una muestra y que se predice a partir de la señal decodificada, siendo esta previsión descrita más adelante—. Sus respuestas en frecuencia corresponden mucho a las características de los filtros paso bajo y paso alto y permiten descomponer la señal en 2 sub-señales según su contenido de frecuencia. 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: $x_{rec,ss2}(n) = x_{rec}(n) - x_{rec,ss1}(n)$.

En los dos casos, la combinación de las sub-señales atenuadas para la obtención de la señal atenuada se realiza mediante simple adición de las sub-señales atenuadas en la etapa E606 descrita posteriormente.

Para no utilizar una señal futura para estos filtrados, se puede completar por ejemplo la señal decodificada por una muestra a 0 al final del bloque. Son posibles otros valores para completar (predecir) la señal futura: por ejemplo se puede repetir la última muestra (el último valor del bloque o predecir la muestra futura, por ejemplo por simple extrapolación lineal. 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:

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

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

Se puede observar que las dos sub-señales permanecen en este caso en la misma frecuencia de muestreo que la señal decodificada. En unas variantes del modo de realización se podrá extender el orden de los filtros paso bajo y paso alto, por ejemplo pasando al orden 4 (en lugar de 2), manteniendo sus características de fase nula.

Se implementa una etapa E604 de cálculo de los factores de atenuación del pre-eco en el módulo de cálculo 604. Este cálculo se realiza por separado para las dos sub-señales.

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

Por razones de simplicidad detallamos en este caso el ejemplo de la descomposición en dos sub-señales de la señal decodificada por filtrado FIR de 3 coeficientes, en donde estas dos sub-señales contienen respectivamente las componentes mayoritariamente de baja frecuencia y de alta frecuencia de la señal decodificada.

De ese modo el factor de atenuación se puede elegir totalmente independiente para las sub-señales y el nivel de atenuación puede ajustarse precisamente, en función del pasado y del futuro de cada sub-señal.

Un ejemplo de realización de cálculos de atenuación se describe en la solicitud de patente FR 08 56248. Los factores de las atenuaciones se calculan por sub-bloque. En el procedimiento descrito en este caso, se calculan además separadamente 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 se alisan eventualmente estos valores de atenuaciones para obtener los valores de atenuación por muestra.

El cálculo del 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 donde f es una función decreciente en 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 relación a la energía máxima es reducida, no es necesaria entonces ninguna atenuación. El factor se fija 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. Esta inicialización puede ser común para todas las sub-señales.

En una variante de realización, la inicialización puede ser diferente para las sub-señales.

En otra variante de realización, en lugar de utilizar la misma relación $R(k)$ para la detección del ataque y el tratamiento en el campo de la señal decodificada, se puede recalcular el mismo tipo de relación para la sub-señal en cuestión.

En otra variante más de realización $f(R(k))$ puede ser constante para toda la zona de pre-eco, por ejemplo $f(R(k)) = 0,01$.

Los valores de atenuación se afinan a continuación por sub-señal para poder ajustar 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 de pre-eco, la energía de la señal se haga inferior a la energía media por sub-bloque

de la señal que precede a la zona de tratamiento (típicamente la de la trama precedente o la de la segunda mitad de la trama precedente).

Esta limitación puede realizarse 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:

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

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

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

donde los índices de sub-bloque de 0 a K corresponden a la trama actual.

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 se limita por supuesto a un máximo de 1 puesto que se afecta en este caso a los valores de atenuación. Más precisamente:

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

donde 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:

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

En una primera variante de realización la zona de pre-eco en la que la atenuación se extiende 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 $pos =$

$\min \left(L' \left(\arg \max_{k=0, K+K'-1} (En(k)) \right), L \right)$ —. 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 recortando el sub-bloque en sub-sub-bloques y observando la evolución de la energía en estos sub-sub-bloques. Supongamos que la posición de inicio del ataque se detecta en el sub-bloque k, $k > 0$ y el inicio del ataque afinado de nuevo 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:

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

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 basa 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 necesario ya por tanto recalculer 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, esta aproximación da un resultado muy satisfactorio. Esto puede observarse en las figuras 10 y 11 en donde la amplitud de la señal decodificada en a) y de la sub-señal que incluye mayoritariamente unas componentes de baja frecuencia en b) están muy próximas, contrariamente a la de la sub-señal que incluye mayoritariamente unas componentes de alta frecuencia en c) mucho menos energética. Así en esta variante, al menos para una sub-señal, el cálculo del factor de atenuación se efectúa utilizando al menos un parámetro de la etapa de detección de una posición de ataque en la señal decodificada, lo que reduce de nuevo la complejidad.

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

La figura 9 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 el procedimiento de la invención en la etapa de descomposición y en d), la señal decodificada con atenuación de eco según la invención (es decir después de la combinación de las dos sub-señales atenuadas).

Se puede observar 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 unas componentes de baja frecuencia, incluye unas funciones de alisado como se ha descrito en el presente documento 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 una ganancia de alisado.

La señal representada en d) muestra correctamente que el pre-eco se ha atenuado de manera eficaz mediante el procedimiento según la invención.

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

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

con la convención de que $g_{pre,ss1}'(n)$, $n = -(u-1), \dots, -1$ 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 ser por tanto diferente para las 2 sub-señales tratadas por separado, incluso si la detección del ataque se realiza en común basándose en la señal decodificada.

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 sin embargo 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 donde se sitúa el inicio del ataque. Esto es equivalente a avanzar el marcador pos en $u-1$ muestras para la sub-señal donde 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. Entonces se preserva la amplitud del ataque.

Son posibles otras funciones de alisado como por ejemplo:

$$g_{pre,ss1}(n) = \alpha g_{pre,ss1}(n-1) + (1-\alpha) g_{pre,ss1}'(n) \text{ siendo típicamente } \alpha = 0,85.$$

Si no se aplica el alisado:

$$g_{pre,ss2}(n) = g_{pre,ss2}'(n)$$

El módulo 605 del dispositivo 600 de la figura 6 implementa la etapa E605 de atenuación de pre-eco en la zona de pre-eco de cada una de las sub-señales por 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, en una etapa E606 del módulo de obtención 606 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:

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

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

Pueden aplicarse evidentemente otros ejemplos de descomposición de la señal decodificada en el marco de la presente invención.

Por ejemplo, la descomposición de la señal decodificada puede efectuarse utilizando unos bancos de filtros de diezmado crítico, por ejemplo de tipo QMF (por "Quadrature Mirror Filter" en inglés), PWMF (por "Pseudo Quadrature Mirror Filter" en inglés), para obtener unas señales en sub-bandas, teniendo estas señales en sub-bandas una frecuencia de muestreo inferior a la de la señal decodificada.

Una vez que se realiza la atenuación de pre-eco en las sub-bandas, se obtiene la señal de salida mediante filtrado QMF de síntesis.

La señal decodificada a tratar puede descomponerse en varias sub-señales, por ejemplo utilizando un filtro QMF o PQMF de 4 sub-bandas. Este enfoque tiene sin embargo el inconveniente de ser más complejo y de ocasionar un retardo de varias muestras.

De ese modo, más generalmente, la descomposición puede permitir obtener más de dos sub-señales.

Además, la descomposición puede efectuarse según un criterio diferente del criterio de frecuencia.

Puede utilizarse por ejemplo un criterio del tipo periodicidad de la señal. En este ejemplo de criterio, la descomposición de la señal decodificada da así una primera sub-señal con unas componentes sinusoidales y una segunda sub-señal con unas componentes de tipo ruido cuando el modelo de señal este tipo seno + ruido. También aquí, esta descomposición tiene el inconveniente de ser más compleja.

Cualquiera que sea el criterio de descomposición utilizado, la descomposición en sub-señales, en el marco de la presente invención, se aplica específicamente para el tratamiento de los pre-ecos sin que estén naturalmente disponibles en el decodificador.

De ese modo, los factores de atenuación calculados se adaptan específicamente a las componentes de las sub-señales obtenidas según el criterio de descomposición. Esto permite hacer la atenuación del pre-eco más precisa y mejor adaptada a la señal.

La utilización de un criterio de frecuencia permite controlar con precisión la cantidad de atenuación aplicada en las diferentes zonas de frecuencia y de ese modo atenuar más eficazmente el pre-eco.

Este resultado es particularmente visible en las figuras 9 a 11 descritas anteriormente en donde se puede ver que la atenuación de pre-eco de la señal (particularmente de la segunda sub-señal) (representada en d)) se efectúa de manera precisa.

Las figuras 12a a 12c muestran igualmente este resultado. En efecto, la figura 12a ilustra el espectrograma de una señal original donde puede verse netamente el ataque. La figura 12b ilustra la señal decodificada sin tratamiento de atenuación de pre-eco. El ataque es entonces más visible. La figura 12c ilustra el espectrograma de la señal decodificada con tratamiento de atenuación de pre-eco según la invención. Se puede distinguir de nuevo precisamente el ataque tanto si es en la parte de alta frecuencia como en la de baja frecuencia.

Se describe ahora en referencia a la figura 13 un ejemplo de realización de un dispositivo de tratamiento de atenuación según la invención.

Materialmente, este dispositivo 100 en el sentido de la invención incluye típicamente, un procesador μP 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 mencionada en tanto que medio para memorizar todos los datos necesarios para la implementación del procedimiento de tratamiento de atenuación tal como se describe en referencia a la figura 6. Este dispositivo recibe en la entrada unas tramas sucesivas de la señal digital Se y suministra la señal Sa reconstruida con atenuación de pre-eco en cada una de las sub-señales y reconstrucción de la señal atenuada por combinación de las sub-señales atenuadas.

El bloque de memoria BM puede incluir un programa informático que incluye las instrucciones de códigos para la implementación de las etapas del procedimiento según la invención cuando estas instrucciones se ejecutan por un procesador μ P del dispositivo y particularmente una etapa de descomposición de la señal decodificada en al menos dos sub-señales según un criterio de descomposición predeterminado, de cálculo de factores de atenuación por sub-señal y por muestra de una zona de pre-eco previamente determinada, de atenuación de pre-eco en la zona de pre-eco de cada una de las sub-señales mediante aplicación de los factores de atenuación a las sub-señales y de obtención de la señal atenuada por combinación de las sub-señales atenuadas.

La figura 6 puede ilustrar el algoritmo de un programa informático de ese tipo.

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

REIVINDICACIONES

1. Procedimiento de tratamiento de atenuación de pre-eco en una señal de audio digital decodificada según una decodificación por transformada, caracterizada por que incluye las siguientes etapas:
5
 - descomponer (E603) la señal decodificada en al menos dos sub-señales según un criterio de descomposición predeterminado;
 - calcular (E604) factores de atenuación por sub-señal y por muestra de una zona de pre-eco previamente determinada;
 - 10
 - atenuar (E605) el pre-eco en la zona de pre-eco de cada una de las sub-señales mediante aplicación de los factores de atenuación a las sub-señales; y
 - obtener (E606) la señal atenuada por combinación de las sub-señales atenuadas.
152. Procedimiento según la reivindicación 1, caracterizado por que el criterio de descomposición predeterminado es un criterio de frecuencia.
3. Procedimiento según la reivindicación 2, caracterizado por que la descomposición de la señal decodificada se efectúa mediante un primer filtrado paso bajo o paso alto para obtener una primera sub-señal.
204. Procedimiento según la reivindicación 3, caracterizado por que la descomposición de la señal se efectúa además mediante un segundo filtrado paso alto o paso bajo, complementario del primer filtrado, para obtener una segunda sub-señal.
255. Procedimiento según la reivindicación 3, caracterizado por que se obtiene una segunda sub-señal por sustracción de la primera sub-señal a la señal decodificada.
6. Procedimiento según la reivindicación 1, caracterizado por que la descomposición de la señal decodificada se efectúa por filtrado QMF y PQMF para obtener unas sub-señales en sub-bandas.
307. Procedimiento según la reivindicación 1, caracterizado por que el criterio de descomposición predeterminado es un criterio de periodicidad de la señal.
8. Procedimiento según la reivindicación 1, caracterizado por que la zona de pre-eco se determina según las siguientes etapas:
35
 - detectar una posición de ataque en la señal decodificada, antes de la etapa de descomposición;
 - determinar una zona de pre-eco que precede a la posición de ataque detectada en la señal decodificada antes de la etapa de descomposición o en las sub-señales después de la etapa de descomposición.
409. Procedimiento según la reivindicación 1, caracterizado por que la zona de pre-eco se determina según las siguientes etapas:
45
 - detectar una posición de ataque en cada una de las sub-señales obtenidas después de la etapa de descomposición;
 - determinar una zona de pre-eco que precede a la posición de ataque detectada en cada una de las sub-señales.
5010. Procedimiento según la reivindicación 8, caracterizado por que al menos para una sub-señal, el cálculo del factor de atenuación se efectúa utilizando al menos un parámetro de la etapa de detección de una posición de ataque en la señal decodificada.
11. Procedimiento según la reivindicación 1, caracterizado por que se aplica una etapa de alisado del factor de atenuación al menos para una sub-señal.
5512. Procedimiento según la reivindicación 1, caracterizado por que el factor de atenuación se fuerza a 1 para un número predeterminado de muestras que preceden al inicio del ataque.
13. Dispositivo de tratamiento de atenuación de pre-eco en una señal de audio digital decodificada a partir de un decodificador por transformada, caracterizado por que comprende:
60
 - un módulo (603) de descomposición de la señal decodificada en al menos dos sub-señales según un criterio de descomposición predeterminado;
 - un módulo (604) de cálculo de los factores de atenuación por sub-señal y por muestra de una zona de pre-eco previamente determinada;
 - 65
 - un módulo (605) de atenuación de pre-eco en la zona de pre-eco de cada una de las sub-señales por aplicación de los factores de atenuación a las sub-señales;

- un módulo (606) de obtención de la señal atenuada por adición de las sub-señales atenuadas.

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

- 5 15. Programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento según una de las reivindicaciones 1 a 10, cuando se ejecutan estas instrucciones mediante un procesador.

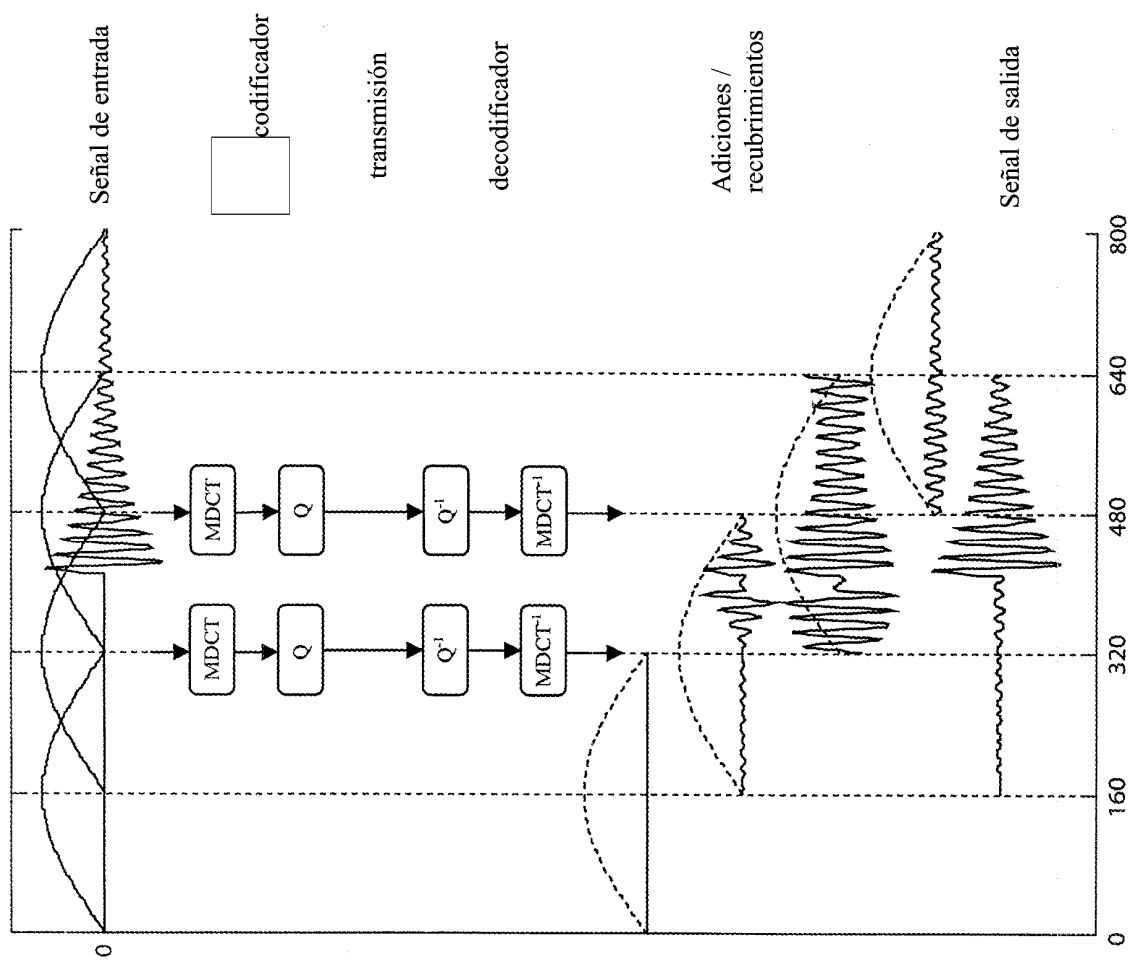


FIG.1

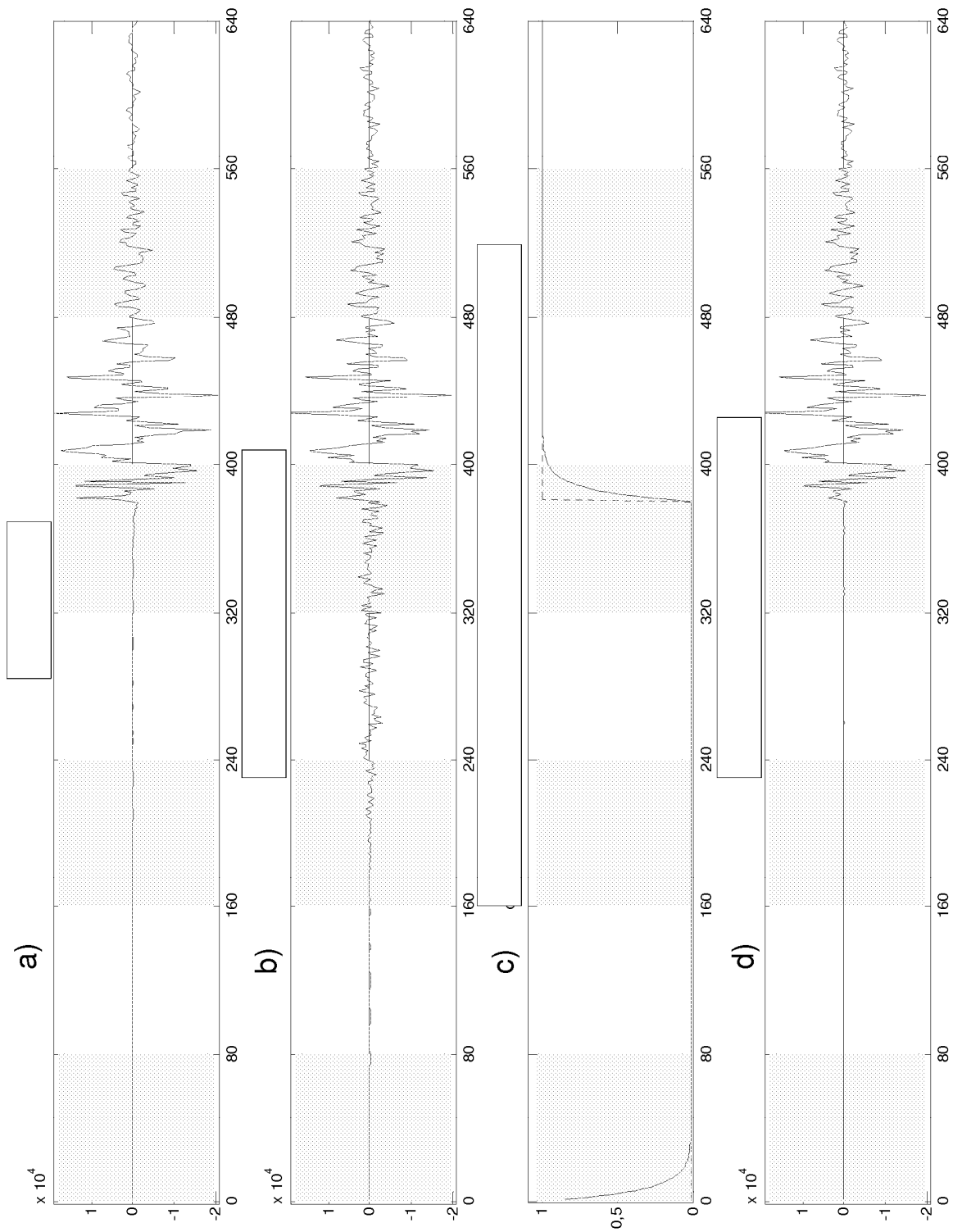


FIG.2

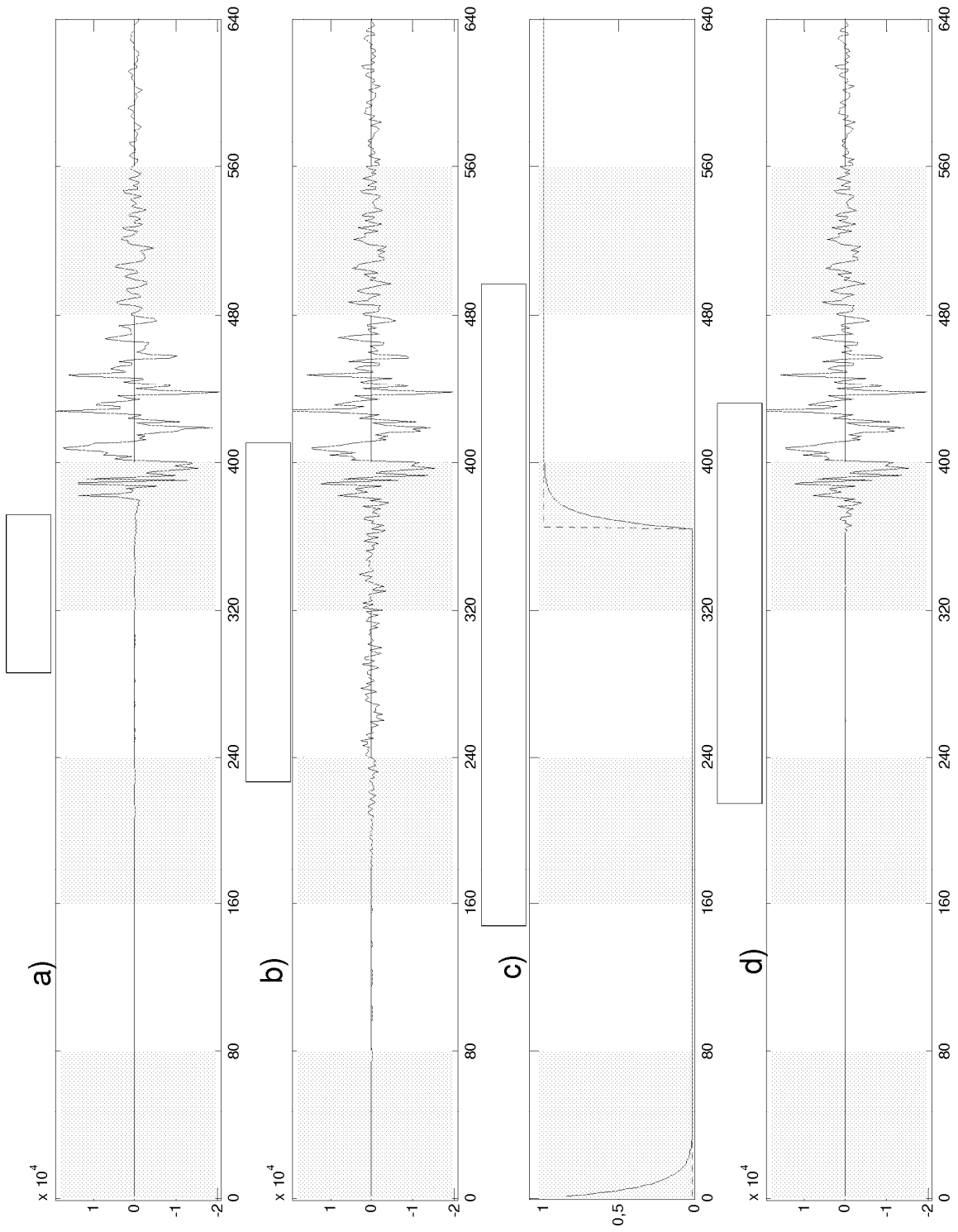


FIG.3

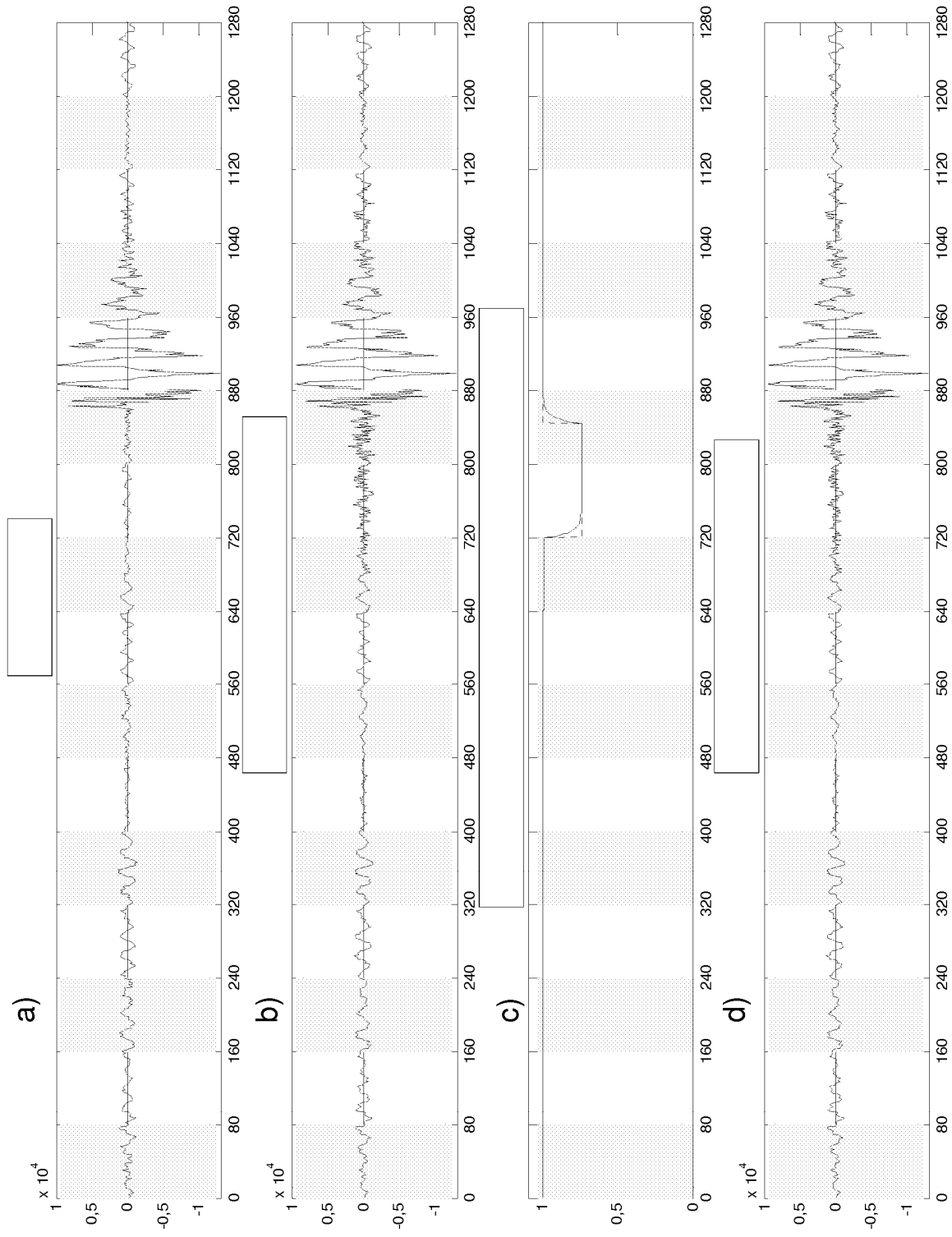


FIG.4

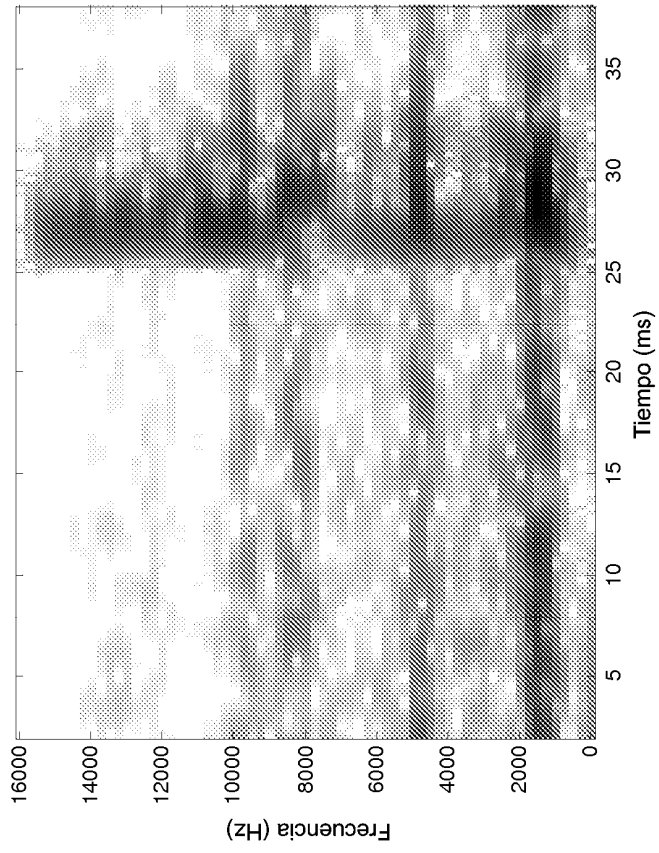


FIG.5a

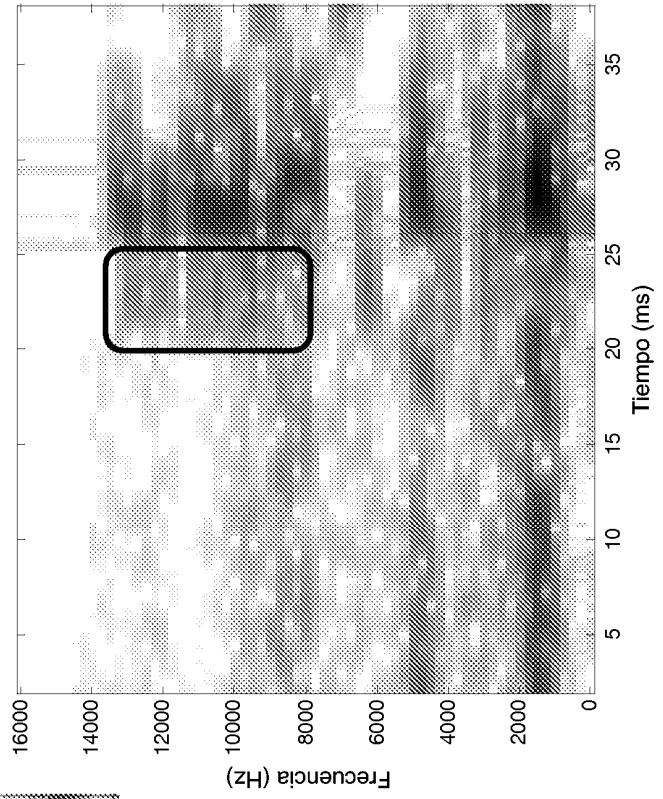


FIG.5b

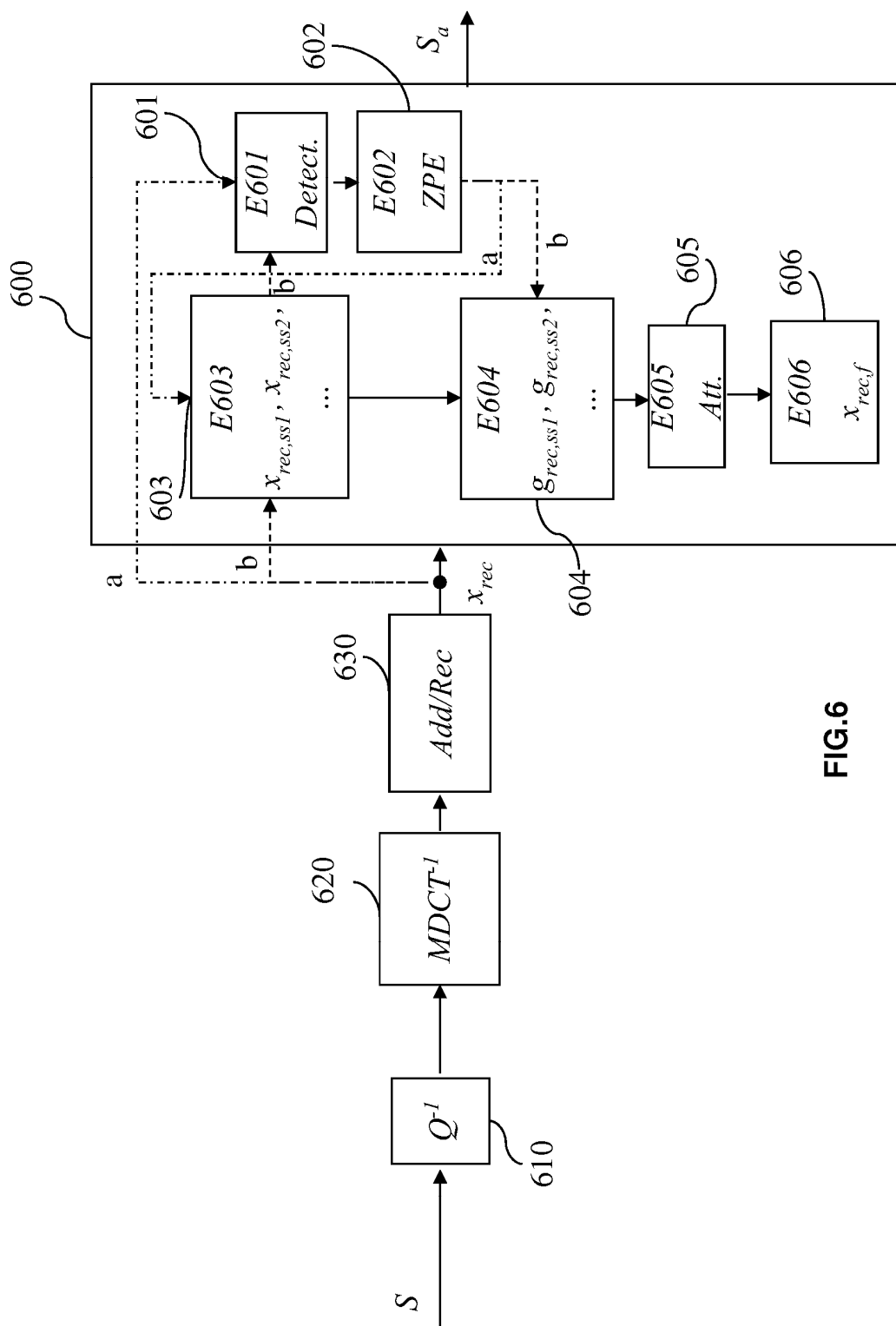


FIG.6

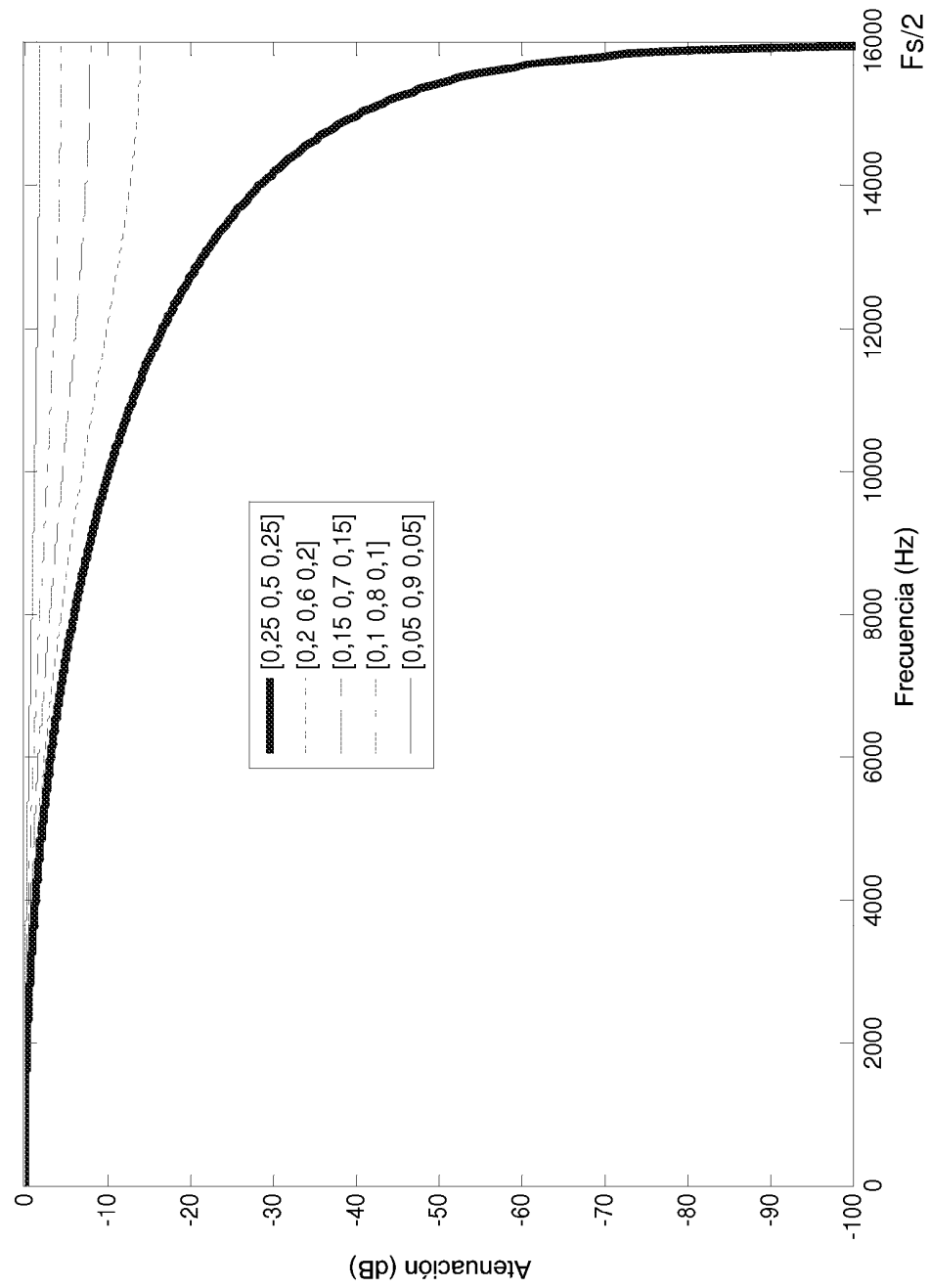


FIG.7a

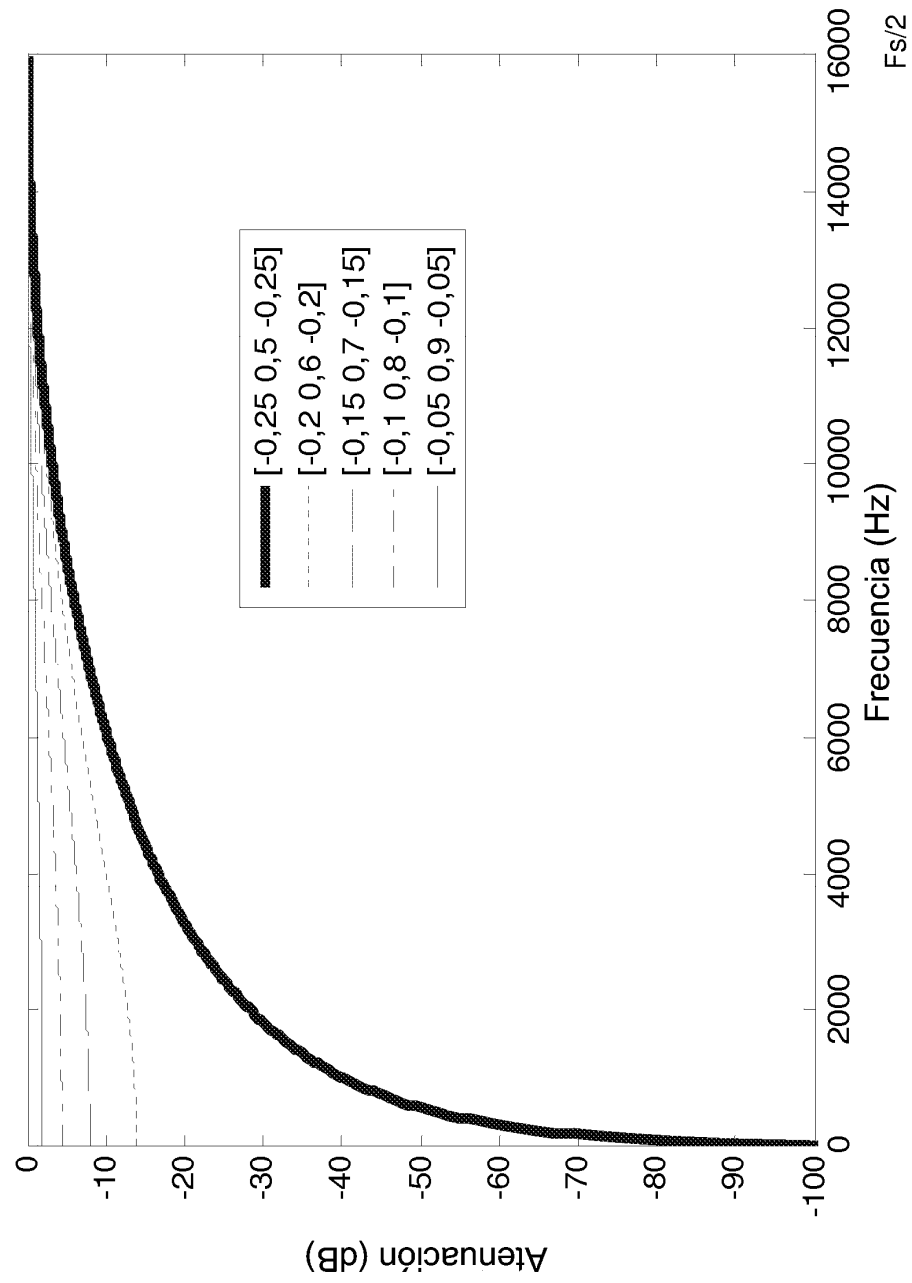


FIG.7b

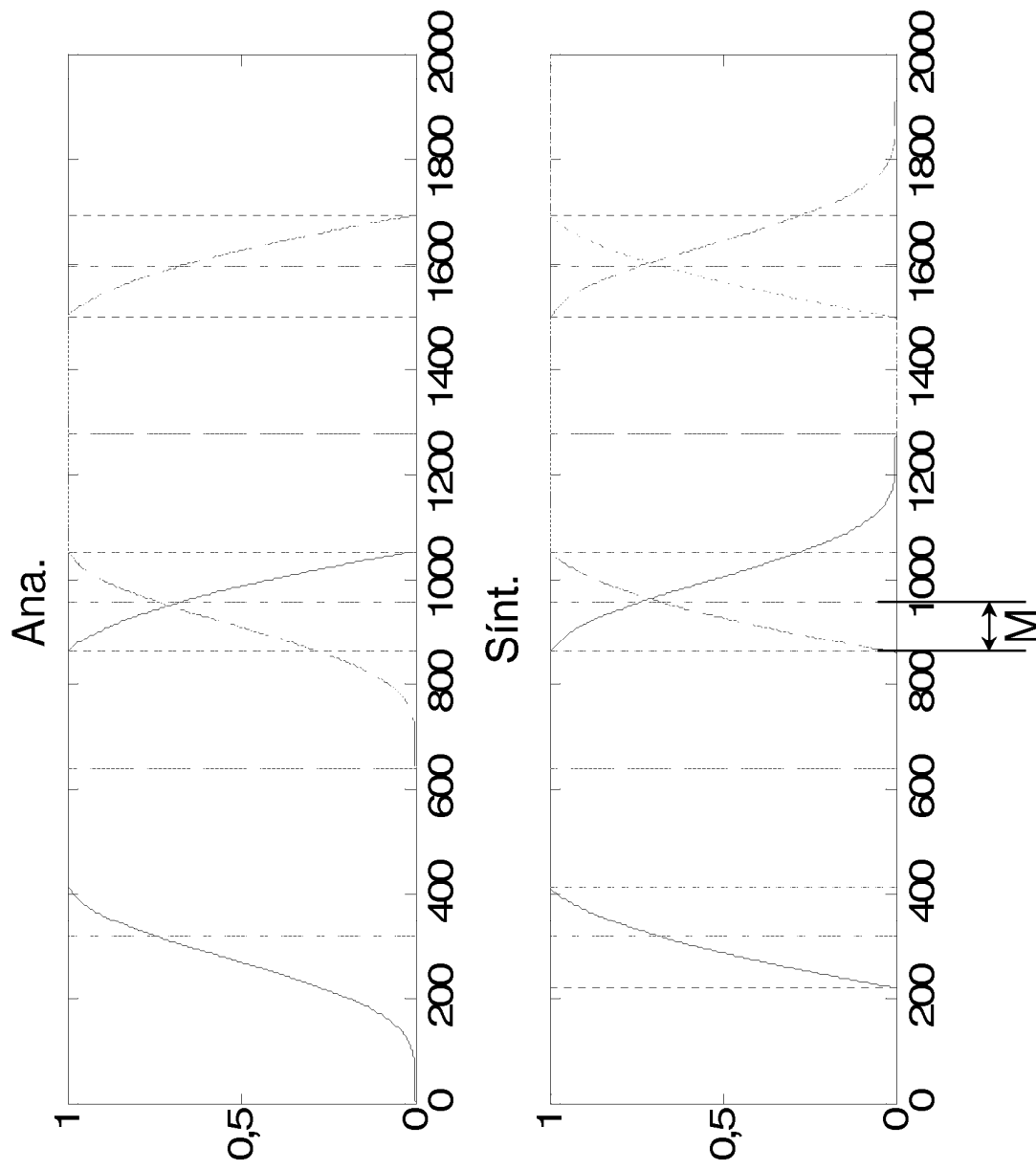


FIG.8

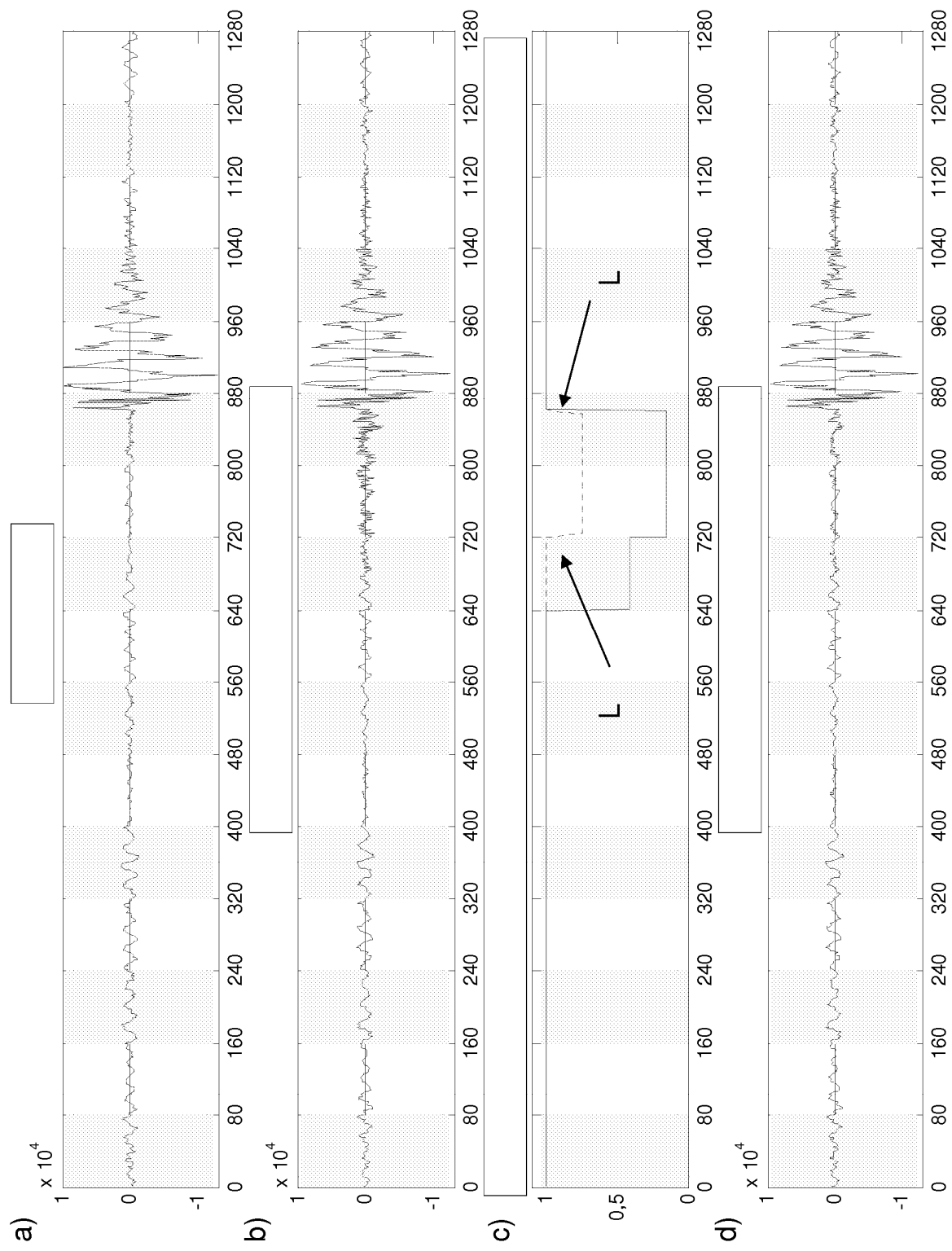


FIG.9

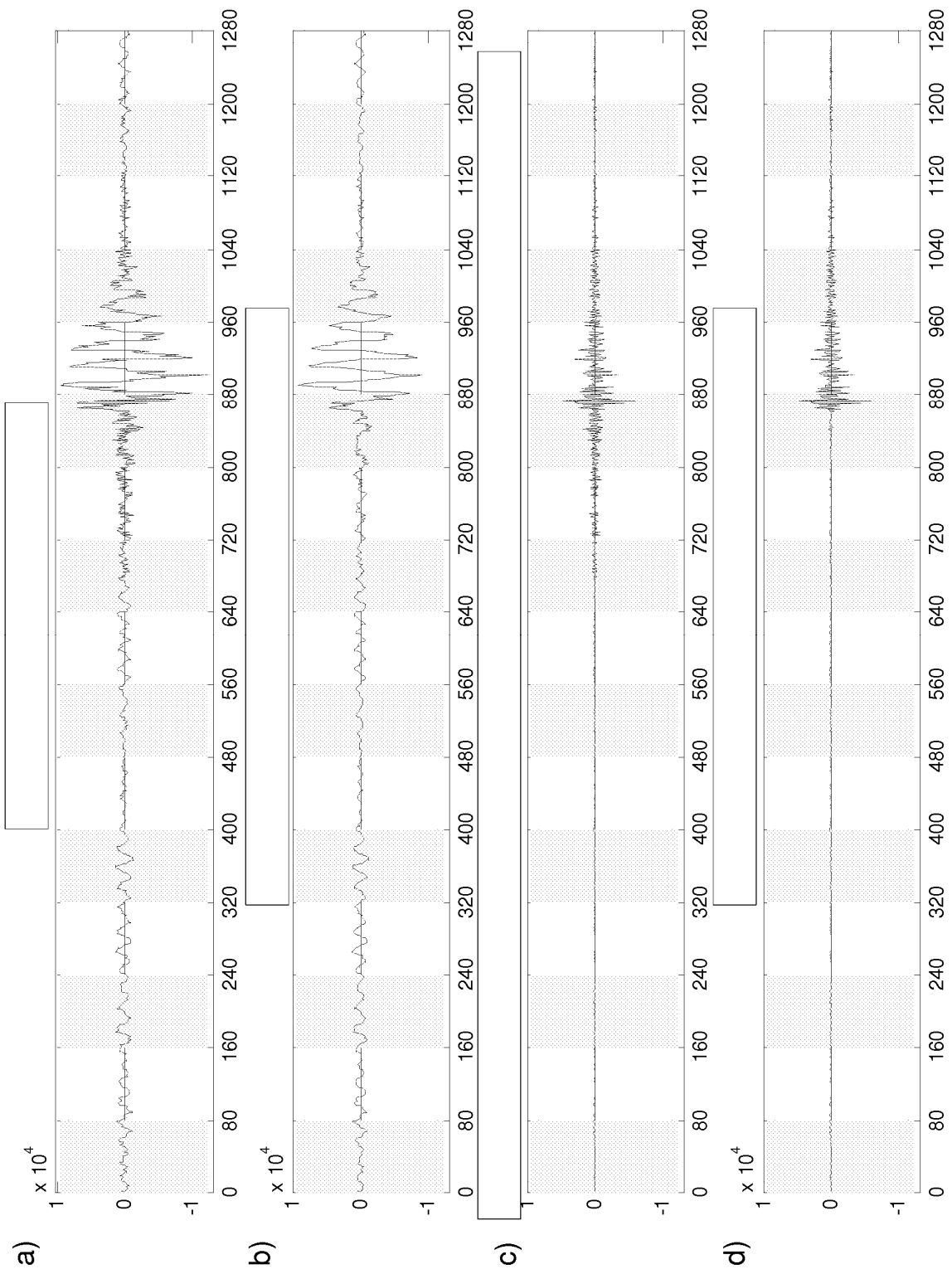


FIG.10

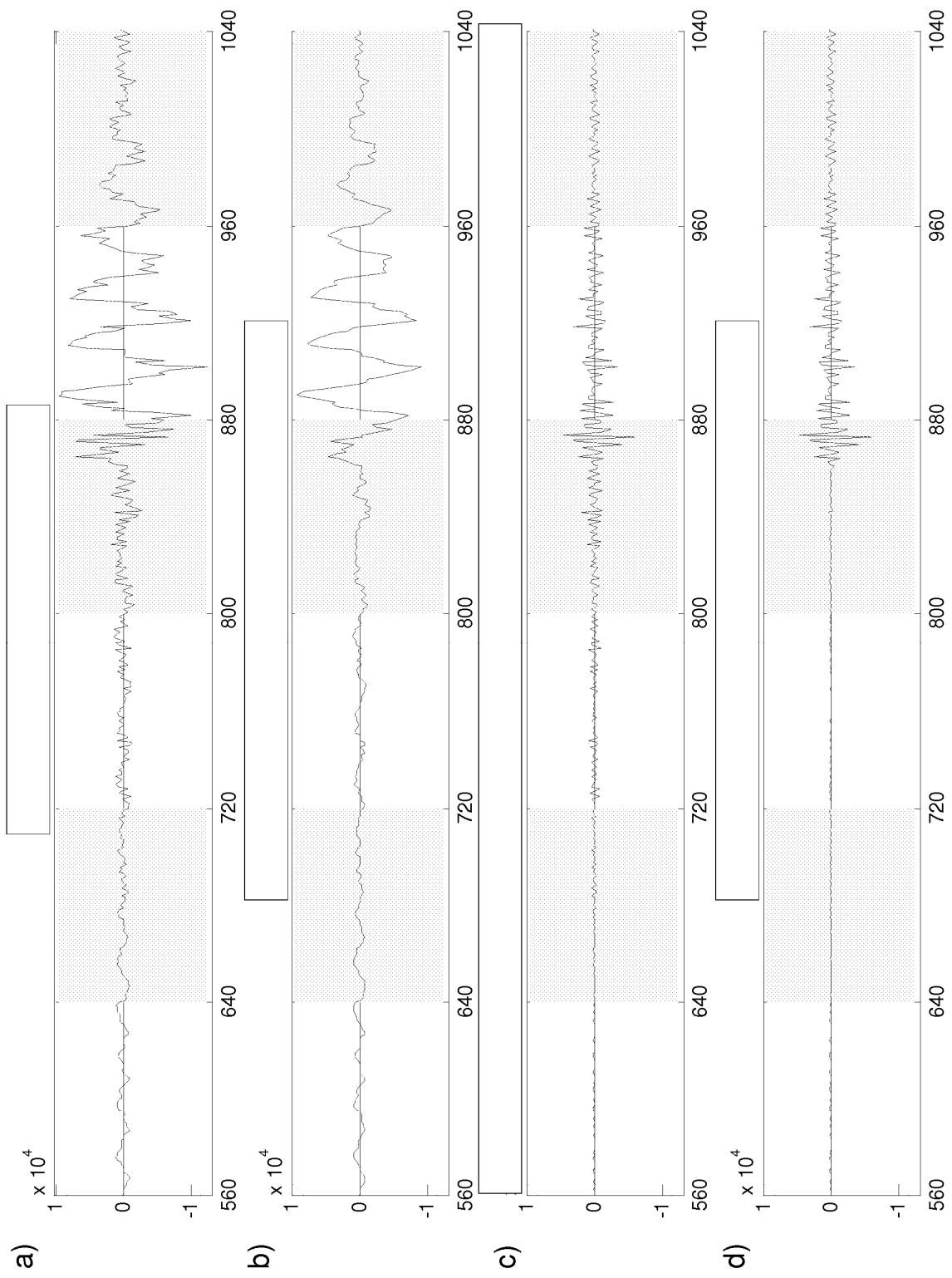


FIG.11

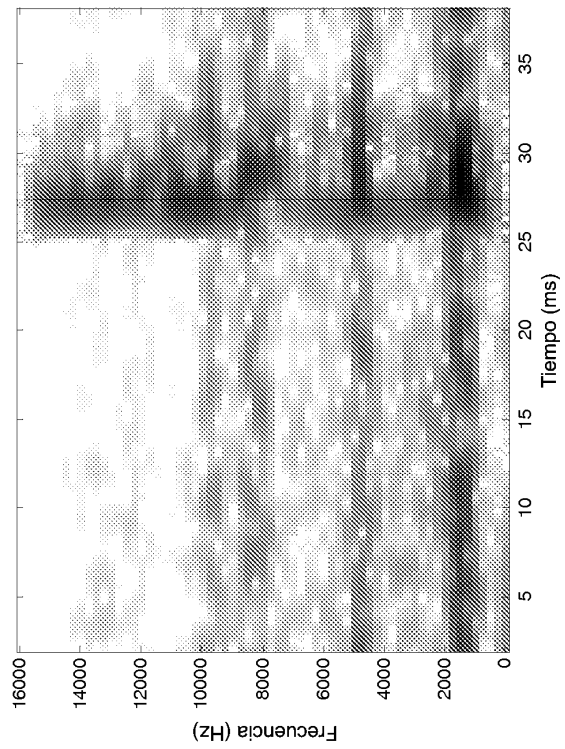


FIG. 12a

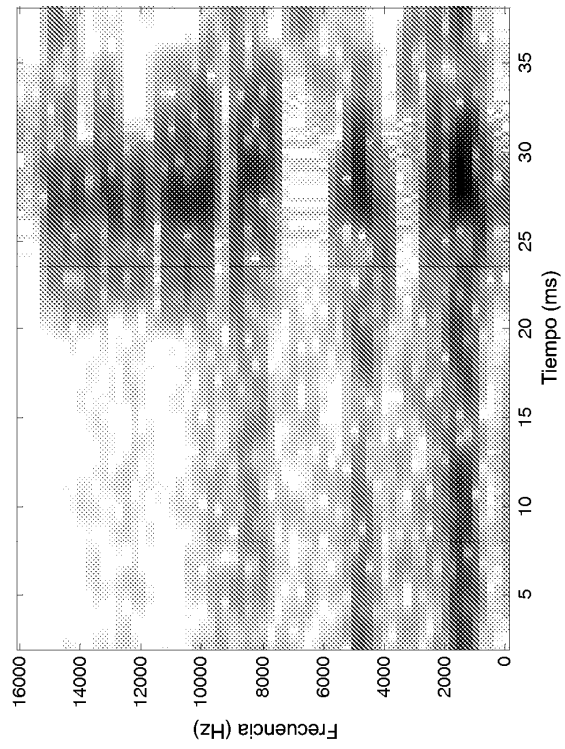


FIG. 12b

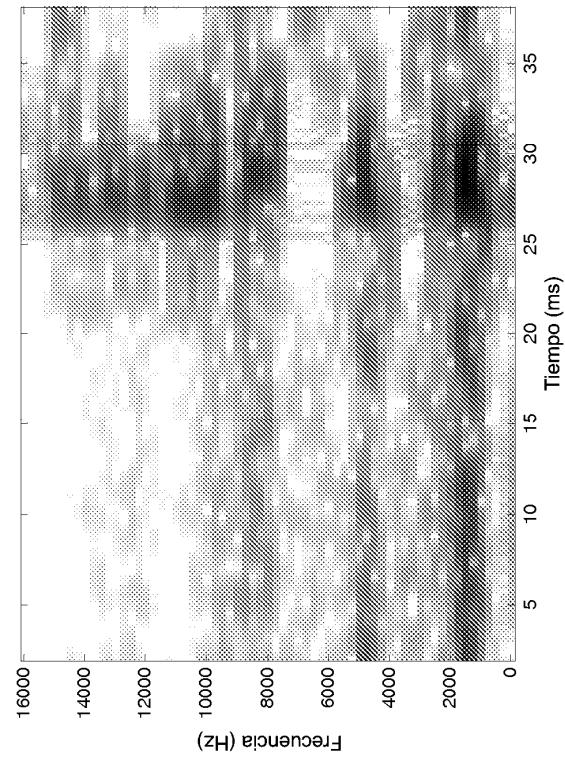


FIG. 12c

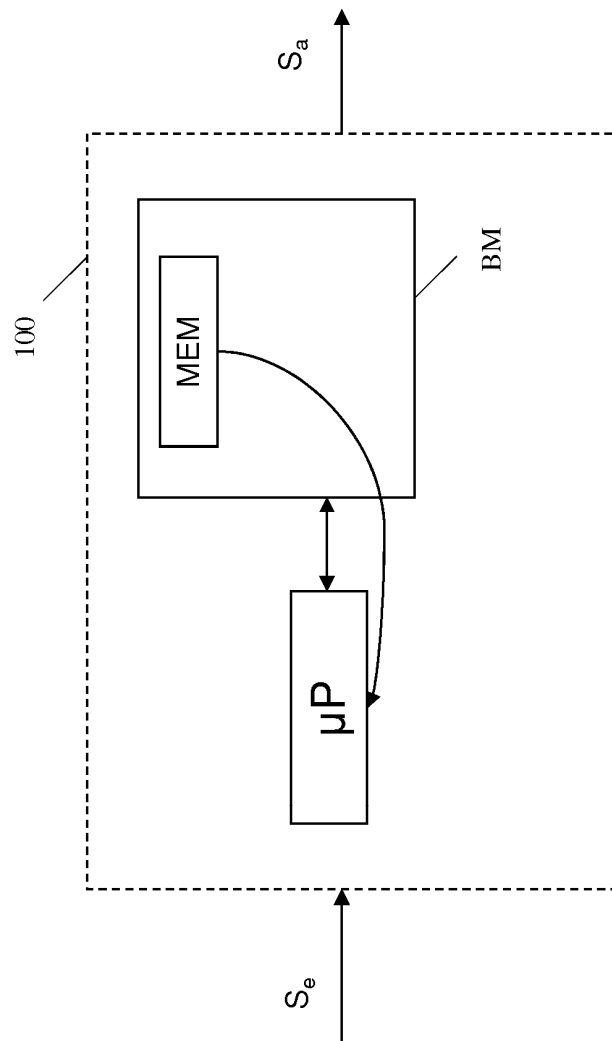


FIG.13