

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 016 645**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/008 (2013.01)

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.10.2009 E 23202859 (7)**

97 Fecha y número de publicación de la concesión europea: **22.01.2025 EP 4293665**

54 Título: **Protección contra el recorte de señal usando metadatos de ganancia de audio preexistentes**

30 Prioridad:

29.10.2008 US 109433 P

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

09.05.2025

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.00%)
77 Sir John Rogerson's Quay Block C Grand
Canal Docklands
Dublin, D02 VK60, IE**

72 Inventor/es:

**SCHILDBACH, WOLFGANG A y
GROESCHEL, ALEXANDER**

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 016 645 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

Protección contra el recorte de señal usando metadatos de ganancia de audio preexistentes

Referencia cruzada a solicitudes relacionadas

5 Esta solicitud reivindica la prioridad de la solicitud provisional de patente de EE. UU. n.º 61/109433, presentada el 29 de octubre de 2008.

Campo de la invención

La solicitud de patente se refiere a la protección contra el recorte de una señal de audio utilizando metadatos de audio preexistentes integrados en una corriente de audio digital. En particular, la solicitud se refiere a la protección contra el recorte cuando se mezcla de manera estereofónica una señal de audio de multicanal en menos canales.

10 Antecedentes de la invención

Es un concepto común integrar metadatos de audio en una corriente de audio digital; por ejemplo, en entornos de difusión digital. Tales metadatos son "datos acerca de datos", es decir, datos acerca del audio digital en la corriente. Los metadatos pueden proporcionar información a un decodificador de audio sobre cómo reproducir el audio. Un tipo de metadatos es la información de control de intervalo dinámico que representa una envolvente de ganancia que varía en el tiempo. Tales metadatos de control de intervalo dinámico pueden servir para múltiples fines:

(1) Controlar el intervalo dinámico del audio reproducido: la transmisión digital permite un alto intervalo dinámico, pero las condiciones de escucha no siempre permiten aprovechar esa ventaja. Aunque un alto intervalo dinámico es deseable en condiciones tranquilas de una sala de estar, puede no ser apropiado para otras condiciones, por ejemplo para una radio de coche, debido al alto nivel de ruido de fondo. Para adaptarse a una amplia variedad de condiciones de escucha, se pueden insertar metadatos en la corriente de audio digital que indiquen a un receptor cómo reducir el intervalo dinámico del audio reproducido en lugar de reducir el intervalo dinámico del audio antes de la transmisión. Este último enfoque no es preferible, ya que imposibilita que un receptor reproduzca el audio con un intervalo dinámico completo. Se prefiere, en cambio, el primer enfoque, ya que permite al oyente decidir si se aplicará o no el control de intervalo dinámico dependiendo del entorno de escucha. Tales metadatos de control de intervalo dinámico hacen que la compresión de intervalo dinámico artístico de alta calidad de una señal decodificada esté disponible para los oyentes a su discreción

(2) Impedir el recorte en caso de hacer una acción de mezcla estereofónica: Cuando se mezcla de manera estereofónica una señal multicanal (por ejemplo, una señal de audio de canales 5.1), el número de canales se reduce, típicamente a dos canales. En el caso de reproducir una señal de audio multicanal que comprende más de dos canales (por ejemplo, una señal de audio de canales 5.1, que tiene 5 canales principales y 1 canal de efecto de baja frecuencia) mediante altavoces estéreo, se realiza típicamente una acción de mezcla estereofónica del lado del receptor, siendo, la señal multicanal, mezclada en dos canales. La acción de mezcla se puede describir mediante una matriz de mezcla estereofónica, por ejemplo mediante una matriz de 2-5 que tiene 2 filas y 5 columnas en caso de mezclar de manera estereofónica una señal de 5 canales en una señal (estéreo) de 2 canales (el canal de efecto de baja frecuencia típicamente no se considera durante la mezcla estereofónica).

Se conocen diferentes esquemas de mezcla de manera estereofónica para mezclar los 5 canales principales de una señal de canales 5.1 en dos canales, por ejemplo Lo/Ro (sólo izquierda, sólo derecha) o Lt/Rt (izquierda total, derecha total).

La operación de mezcla estereofónica conlleva el riesgo de una sobrecarga ocasional de la señal estéreo digital, generando por ello artefactos de recorte no deseados. Tal recorte puede ocurrir cuando la amplitud de una señal digital mezclada de manera estereofónica que excedería el valor máximo (o el mínimo) representable esté limitada al valor máximo (o mínimo) representable. Por ejemplo en caso de una representación binaria simple de punto fijo sin signo, el recorte se produce cuando la amplitud mezclada de manera estereofónica calculada está limitada a la palabra de valor máximo en la que todos los bits corresponden a 1. En el caso de una representación con signo de 16 bits, el valor máximo puede, por ejemplo, corresponder a la palabra "01111111 11111111".

Como las matrices de mezcla estereofónica para los diversos esquemas de mezcla de manera estereofónica son conocidas en el lado del extremo de cabecera, del remitente o de la generación de contenido, para señales que pueden dar lugar a recortes cuando se mezclan de manera estereofónica, se pueden añadir, a la corriente de audio, metadatos de control de intervalo dinámico que indiquen al receptor para que atenúe las señales que se han de mezclar de manera estereofónica antes de mezclar, e impedir así dinámicamente el recorte.

(3) Impedir el recorte en caso de salida intensificada: Para la retransmisión a través de canales dinámicamente muy limitados (por ejemplo, desde un decodificador a través de un enlace de RF analógico a la entrada de RF de un televisor), la señal se intensifica, típicamente en 11 dB, para conseguir una mejor relación señal-ruido en este trayecto. En tales aplicaciones, para señales que pueden resultar en recorte cuando se amplifican en 11 dB, se pueden añadir metadatos de control de intervalo dinámico que indican a un receptor que atenúe las señales antes de aplicar la amplificación de 11 dB a la corriente de audio para evitar dinámicamente el recorte.

Desde la perspectiva del dispositivo que recibe la corriente de audio, no está claro si los metadatos de control de intervalo dinámico entrantes cumplen el fin del punto (1), es decir, el control de intervalo dinámico, el fin del punto (2), es decir, la protección contra el recorte de mezcla estereofónica, o los fines establecidos en ambos puntos (1) y (2). A menudo, los metadatos realizan ambas tareas, pero no siempre es así, por lo que en algunos casos es posible que los metadatos no incluyan protección contra el recorte de mezcla estereofónica. Además, en caso de que los metadatos (típicamente, se utiliza un parámetro de ganancia diferente para el modo RF) estén asociados con el modo RF en el punto (3), los metadatos se pueden utilizar para impedir el recorte en caso de una amplificación extraordinaria (tanto en el caso de mezclar de manera estereofónica como en el caso de no mezclar de manera estereofónica).

Además, la corriente de audio entrante puede no incluir metadatos de control de intervalo dinámico en absoluto, debido al hecho de que para algunos formatos de codificación de audio los metadatos son opcionales.

Si los metadatos de control de intervalo dinámico no se incluyen con la corriente de audio comprimido o se incluyen pero no incluyen protección contra el recorte de mezcla estereofónica, pueden estar presentes artefactos de recorte no deseados en la señal decodificada si una señal de multicanal se mezcla de manera estereofónica en menos canales.

El documento WO 2008/100098 describe un método y un aparato de codificación/decodificación de audio para procesar señales de audio basadas en objetos.

Compendio

La invención está definida por las reivindicaciones independientes 1 y 4. Las realizaciones ventajosas se proporcionan en las reivindicaciones dependientes. La presente invención describe un método y un aparato para impedir el recorte de una señal de audio cuando no se garantiza la protección contra el recorte mediante metadatos de audio.

Un primer aspecto de la solicitud se refiere a un método para proporcionar protección contra el recorte de señal de una señal de audio, por ejemplo una señal de audio digital mezclada de manera estereofónica, que se infiere a partir de datos de audio digitales. Según el método, se determina si los primeros valores de ganancia basados en los metadatos de audio recibidos son suficientes para protección contra recorte de la señal de audio. Los metadatos de audio están integrados en una primera corriente de audio. Por ejemplo, se determina si los metadatos de envolvente de ganancia variable en el tiempo incluidos con una corriente de audio comprimido son suficientes o no, para impedir el recorte de mezcla estereofónica. En caso de que un primer valor de ganancia no sea suficiente para la protección, el primer valor de ganancia respectivo se reemplaza con un valor de ganancia suficiente para la protección contra el recorte de la señal de audio. Preferiblemente, en caso de que no haya metadatos relacionados con el control de intervalo dinámico en la primera corriente de audio, el método puede añadir valores de ganancia suficientes para la protección contra el recorte de señal. Por ejemplo en el caso en que los metadatos de envolvente de ganancia variable en el tiempo no proporcionen suficiente protección contra el recorte de mezcla estereofónica, o no estén presentes en absoluto, los metadatos de envolvente de ganancia variable en el tiempo se modifican o añaden, de modo que proporcionen suficiente protección contra el recorte de mezcla estereofónica.

El método permite la protección contra el recorte, en particular protección contra el recorte en caso de mezcla estereofónica, independientemente de si los valores de ganancia suficientes para la protección contra el recorte son recibidos o no.

Según el método, las palabras de ganancia de audio recibidas (si se proporcionan) se pueden aplicar de la manera más veraz posible, pero se pueden anular cuando las palabras de ganancia entrantes no proporcionen suficiente atenuación para impedir el recorte, por ejemplo en una mezcla estereofónica.

Dado que los datos de control de intervalo dinámico que sirven al fin del punto (1) tienen aspectos artísticos, típicamente no es responsabilidad del dispositivo de recepción (por ejemplo, un decodificador) introducir estos en caso de que los metadatos entrantes no lo proporcionen. Sin embargo, las propiedades como las del punto (2) pueden y, por lo tanto, deben ser proporcionadas por la instancia receptora. Esto significa que el dispositivo de recepción intentará conservar en la medida de lo posible los datos de control de intervalo dinámico previstos para el control de intervalo dinámico de acuerdo con el punto (1), añadiendo al mismo tiempo protección contra el recorte.

Hay diversas maneras de determinar si los primeros valores de ganancia basados en los metadatos de audio recibidos son suficientes para la protección contra el recorte de señal.

Según un enfoque preferido, los segundos valores de ganancia se calculan basándose en los datos de audio digital, donde los segundos valores de ganancia son suficientes para la protección contra el recorte de la señal de audio. Los segundos valores de ganancia pueden ser los valores de ganancia máximos permitidos que no den el recorte como resultado.

5 Preferiblemente, el método determina si los primeros valores de ganancia son suficientes, de tal manera que compara los primeros valores de ganancia basándose en los metadatos de audio recibidos y los segundos valores de ganancia calculados. El método puede comparar un primer valor asociado con un segmento de los datos de audio con el segundo valor de ganancia respectivo asociado con el mismo segmento de datos de audio.

10 Dependiendo de esto, se puede generar una corriente de valores de ganancia compatible con la protección contra el recorte a partir de los valores de ganancia primero y segundo. Preferiblemente, tales valores de ganancia se seleccionan a partir de los primeros valores de ganancia y de los segundos valores de ganancia calculados dependiendo de las acciones de comparación. Al seleccionar un segundo valor de ganancia calculado en lugar del primer valor de ganancia, el primer valor de ganancia se reemplaza por el segundo valor de ganancia seleccionado.

15 Preferiblemente, se selecciona el mínimo de un par de primer y segundo valores de ganancia. Si el primer valor de ganancia es mayor que el segundo valor de ganancia calculado suficiente para la protección, esto indica que existe el riesgo de que el primer valor de ganancia no sea suficiente para la protección contra el recorte, y, de este modo, debería reemplazarse con el segundo valor de ganancia respectivo. De lo contrario, si el primer valor de ganancia es menor que el segundo valor de ganancia calculado suficiente para la protección, esto indica que no hay riesgo de recorte de señal y que se debería conservar el primer valor de ganancia.

20 La selección de valores de ganancia procedentes del primer y segundo valores de ganancia se puede realizar como se explica a continuación:

En caso de que tanto el primer valor de ganancia como el segundo valor de ganancia proporcionen una ganancia menor o igual a 1, se toma el mínimo de ambos. Esto significa que el primer valor de ganancia ya garantiza protección contra el recorte, o, si no, será reemplazado por el segundo valor de ganancia.

25 En caso de que la ganancia del segundo valor de ganancia sea mayor que 1 y el primer valor de ganancia proporcione una ganancia menor o igual a 1, la señal podría amplificarse y aún no se recortaría. Sin embargo, la corriente de audio entrante solicita atenuación, por ejemplo para cumplir fines de limitación de intervalo dinámico, y, de este modo, se conserva.

30 En caso de que el primer valor de ganancia proporcione una ganancia mayor que 1 y el segundo valor de ganancia proporcione una ganancia menor o igual a 1, el primer valor de ganancia entrante violaría la protección contra el recorte, y, por lo tanto, se toma el segundo valor de ganancia.

En caso de que tanto el primer valor de ganancia como el segundo valor de ganancia proporcionen una ganancia mayor que 1, la entrada se amplificará. Esta amplificación está permitida siempre que no se produzca un recorte, y, de este modo, se utiliza el menor de entre el primer valor de ganancia y el segundo valor de ganancia.

35 Un enfoque alternativo para determinar si los primeros valores de ganancia son suficientes para la protección es aplicar los primeros valores de ganancia a los datos de audio y determinar si la señal de audio digital resultante (por ejemplo, la señal mezclada de manera estereofónica) se recorta.

40 En caso de que los primeros valores de ganancia no sean suficientes para la protección, se pueden determinar iterativamente valores de ganancia que sean suficientes para la protección contra el recorte a partir de los primeros valores de ganancia como valores de ganancia iniciales. Por ejemplo, se puede determinar si la señal de audio se recorta con un valor de ganancia que es el menor valor de ganancia más cercano que el primer valor de ganancia según la resolución de los valores de ganancia (por ejemplo, en caso de que el primer valor de ganancia sea 0,8 y la resolución de valor de ganancia sea 0,1, el menor valor de ganancia más cercano sería 0,7). Si la señal se recorta todavía, se puede determinar si la señal de audio se recorta con el siguiente valor de ganancia menor (por ejemplo, un valor de ganancia de 0,6). Esto se repite hasta que se encuentre un valor de ganancia que no dé como resultado un recorte de señal.

45 Preferiblemente, el método se realiza como parte de un proceso de transcodificación, donde la primera corriente de audio en un primer formato de codificación de audio (por ejemplo, el formato AAC o el formato AAC de Alta Eficiencia (HE-AAC), también conocido como aacPlus) es transcodificada en una segunda corriente de audio codificada en un segundo formato de codificación de audio (por ejemplo, el formato Dolby Digital o el formato Dolby Digital Plus). La segunda corriente de audio comprende los valores de ganancia reemplazados suficientes para el recorte o tiene valores de ganancia derivados de éstos.

50 A menudo, la transcodificación de audio es necesaria, ya que el formato de compresión digital para transportar los datos de audio no se puede mantener a lo largo de toda la cadena de transmisión hasta el decodificador de audio final en la cadena de transmisión (por ejemplo, hasta el decodificador del receptor de audio/vídeo - AVR). En el caso de difusión, esto se debe a que, por ejemplo, se pueden usar diferentes esquemas de codificación para la difusión por

55

aire (o difusión al consumidor mediante cable) y la transmisión del audio entre el dispositivo de recepción (por ejemplo, un decodificador STB) y el decodificador final de la cadena de transmisión (por ejemplo, el decodificador del AVR o el decodificador de audio del televisor). Por ejemplo, los datos de audio pueden transmitirse por aire a través del formato AAC o del formato HE-AAC, y luego los datos de audio pueden transcodificarse al formato Dolby Digital o al formato Dolby Digital Plus para su transmisión desde el STB al AVR. En consecuencia, se puede realizar una etapa de transcodificación, por ejemplo en el STB, para pasar de un formato a otro. Tal etapa de transcodificación comprende la transcodificación de los propios datos de audio, pero asimismo, idealmente, la transcodificación de los metadatos que los acompañan, así como, en particular los datos de control de intervalo dinámico. Según una realización preferida, el método proporciona metadatos de ganancia de audio transcodificados en la segunda corriente de audio, siendo los metadatos de ganancia suficientes para la protección contra el recorte de señal.

El método puede ser muy útil en cualquier dispositivo que transcodifica una señal procedente de un formato de corriente de audio comprimido a otro, donde no se sabe de antemano si los metadatos de control de ganancia que varían en el tiempo, si los hay, llevados por el primer formato, incluyen protección contra el recorte de mezcla estereofónica (por ejemplo, en un transcodificador de AAC/HE-AAC a Dolby Digital, un transcodificador de Dolby E a AAC/HE-AAC, o un transcodificador de Dolby Digital a AAC/HE-AAC).

Preferiblemente, para determinar si los primeros valores de ganancia son suficientes para la protección, los datos de audio digital se mezclan de manera estereofónica según al menos un esquema de mezclado de manera estereofónica, por ejemplo, según un esquema de mezclado de manera estereofónica Lt/Rt. El mezclado de manera estereofónica da como resultado una o más señales, por ejemplo una señal asociada al canal derecho y una señal asociada al canal izquierdo. Además, se puede considerar una pluralidad de esquemas de mezclado de manera estereofónica, y los datos de audio digital se mezclan de manera estereofónica según más de un esquema de mezclado de manera estereofónica.

Preferiblemente, se determina continuamente un valor de pico real de diversas señales derivadas desde la señal de audio, es decir, que en un momento dado se determina cuál de las diversas señales tiene el valor de señal más alto. Para calcular un valor de pico, el método puede determinar el máximo de los valores absolutos de dos o más señales en un tiempo dado. Las dos o más señales pueden incluir una o más señales después de mezclarse de manera estereofónica según un primer esquema de mezclado de manera estereofónica, por ejemplo, el valor absoluto de una muestra de la señal del canal derecho mezclada de manera estereofónica y el valor absoluto de una muestra simultánea de la señal del canal izquierdo mezclada de manera estereofónica. Además, para calcular el valor de pico, el método puede también considerar el valor absoluto de una o más señales después de mezclar de manera estereofónica según un segundo (e incluso un tercer) esquema de mezclado de manera estereofónica. Además, la determinación del valor de pico puede considerar el valor absoluto de una o más señales de audio antes del mezclado de manera estereofónica, por ejemplo el valor absoluto de cada uno de los 5 canales principales de una señal de canales 5.1 al mismo tiempo. Cabe señalar que, en el caso de la transcodificación, típicamente no se sabe si la señal multicanal se reproduce más tarde en canales discretos o si se realiza el mezclado de manera estereofónica según un esquema de mezclado de manera estereofónica.

Un valor de pico corresponde al máximo de estos valores de muestra de señal simultánea, indicando por ello la amplitud máxima que la señal puede tener para todos los casos posibles en una instancia de tiempo particular, y este es el peor caso que el algoritmo de protección contra el recorte debe tener en cuenta.

Los datos de control de intervalo dinámico típicamente varían en el tiempo en una cierta granularidad que generalmente se relaciona con la longitud del segmento de datos (por ejemplo, bloque) del formato de codificación de audio respectivo o de partes enteras del mismo. De este modo, se calcula también preferiblemente un segundo valor de ganancia por segmento de datos.

De este modo, la velocidad de muestreo de los valores pico o de los valores pico consecutivos se reduce preferiblemente (muestreo descendente). Esto se puede hacer determinando el valor máximo de una pluralidad de valores pico consecutivos o de valores de pico filtrados consecutivos. En particular, el método puede determinar el valor máximo de una pluralidad de valores de pico consecutivos (filtrados) asociados con un segmento de datos, por ejemplo de un bloque o de una trama. En caso de transcodificación, el método puede determinar el valor de pico más alto de una pluralidad de valores de pico consecutivos (filtrados) asociados con un segmento de datos del segundo flujo de datos (saliente). Cabe señalar que para determinar el máximo se consideran preferiblemente no sólo los valores de pico consecutivos basándose en muestras de señal que están en un segmento de salida, sino también valores de pico adicionales (anteriores y posteriores) que influirían en la decodificación del segmento de datos, es decir, valores de pico que se relacionan con muestras de señal al comienzo y al final de una ventana de decodificación. Estos valores de pico están también asociados con el segmento de datos.

En lugar de elegir el valor de pico más alto, se puede calcular un valor diferente por segmento de datos para reducir la tasa de muestreo.

Cabe señalar que las muestras derivadas de los datos de audio distintos de los valores de pico pueden ser muestreadas en sentido descendente. Por ejemplo, los datos de audio pueden mezclarse de manera estereofónica en un solo canal (mono) y sólo se determina el máximo de muestras consecutivas mezcladas de manera estereofónica por segmento de datos salientes. Según un ejemplo diferente, se calcula primero cada máximo para cada señal de canal mezclada de manera estereofónica por segmento de datos salientes (muestreo en sentido descendente) y luego se determina el valor de pico de estos máximos.

Basándose en el máximo determinado, se puede calcular un valor de ganancia invirtiendo el máximo determinado. Si 1 es el valor de señal máximo que se puede representar, invertir el máximo determinado produce directamente un factor de ganancia. Cuando el factor de ganancia se aplica al máximo de los valores de pico (filtrados), el valor resultante es igual a 1, es decir, el valor máximo de la señal. Esto significa que cada muestra de audio a la que se aplica la ganancia se mantiene por debajo de 1 o igual a 1, evitando, de este modo, el recorte de este segmento de datos. En el caso de que 1 sea el nivel máximo de señal, 1 corresponderá a 0 dBFS -decibelios relativos a la escala completa; generalmente se asigna 0 dBFS al nivel máximo posible.

En lugar de simplemente invertir el máximo determinado, se puede calcular un valor de ganancia dividiendo un valor de señal máximo (que corresponde a 0 dBFS) por el máximo determinado asociado con un segmento de datos. Sin embargo, los costes computacionales son mayores en comparación con una inversión simple.

En el caso de la transcodificación, las longitudes de los segmentos de datos (por ejemplo, bloque o trama) suelen ser diferentes para el primer formato de codificación de audio (formato de corriente de entrada) y el segundo formato de codificación de audio (formato de corriente de salida). Por ejemplo, en AAC, un bloque típicamente contiene 128 muestras (en HE-AAC: 256 muestras por trama), mientras que en Dolby Digital un bloque típicamente contiene 256 muestras. De este modo, el número de muestras por bloque aumenta al transcodificar de AAC a Dolby Digital. En AAC, una trama comprende típicamente 1024 muestras (en HE-AAC: 2048 muestras por trama), mientras que en Dolby Digital una trama típicamente comprende 1536 muestras (6 bloques). De este modo, el número de muestras por trama aumenta también cuando se transcodifica de AAC a Dolby Digital. La granularidad de los datos de control de intervalo dinámico es principalmente o bien el tamaño de bloque o bien el tamaño de trama. Por ejemplo, la granularidad de los metadatos de control de intervalo dinámico "DRC" en MPEG para la corriente HE-AAC y de los metadatos de ganancia "dynrng" en Dolby Digital es el tamaño de bloque. Por el contrario, la granularidad de los metadatos de ganancia "compr" en Dolby Digital y de los metadatos de ganancia "compresión fuerte" en DVB (difusión de vídeo digital) para la corriente HE-AAC es el tamaño de trama.

Además, las tasas de muestreo pueden ser diferentes para la corriente de entrada (por ejemplo, 32 KHz, o 44,1 KHz) y la corriente de salida (por ejemplo, 48 KHz), es decir, el audio se muestrea de nuevo. Esto también altera las relaciones de longitud entre los segmentos de datos entrantes y los segmentos de datos salientes. Además, es posible que los segmentos de datos entrantes y salientes no estén alineados. Además, cabe señalar que los metadatos transmitidos en un segmento de datos de entrada (por ejemplo, en un bloque o trama) tienen un área de impacto de control de intervalo dinámico (es decir, el intervalo en la corriente en el que tiene efecto la aplicación del valor de ganancia) que a menudo no es exactamente tan grande como el segmento de datos sino más grande. Esto se debe a las características de superposición y adición de la transformada utilizada y al hecho de que el control de intervalo dinámico se aplica a menudo en el dominio espectral. Lo mismo sucede a menudo con los datos de control de intervalo dinámico de la corriente de audio saliente. Por lo tanto, para determinar qué valores de ganancia de entrada influyen en un segmento de datos de salida dado, se puede observar la superposición de las longitudes de impacto de entrada y salida (en lugar de considerar la superposición de los segmentos de datos de entrada y salida) como se explicará en detalle más adelante.

Debido a las razones analizadas anteriormente, la transcodificación de los datos de control de intervalo dinámico debe tener en cuenta que un valor de control de intervalo dinámico saliente puede verse influido por más de un valor de control de intervalo dinámico entrante. En este caso, se puede realizar un nuevo muestreo (re-entramado) de los datos de control de intervalo dinámico al transcodificar la corriente de datos.

Por lo tanto, el método puede comprender la etapa de volver a muestrear los valores de ganancia derivados de los metadatos de audio recibidos de la primera corriente de audio. Cuando un segmento de datos de la primera corriente de audio cubre una longitud de tiempo más corta que un segmento de datos de la segunda corriente de audio, los valores de ganancia son muestreados en sentido descendente.

Un valor de ganancia muestreado de nuevo puede determinarse calculando el mínimo de entre una pluralidad de valores de ganancia consecutivos. En otras palabras: de entre varias ganancias de control de intervalo dinámico de entrada (que son relevantes para un segmento de datos salientes), se elige la más pequeña. La motivación para esto es conservar los valores entrantes tanto como sea posible (en caso de que los valores no den como resultado un recorte de señal). Sin embargo, esto a menudo no es posible ya que es necesario muestrear de nuevo los valores de ganancia. Por lo tanto, se elige el valor de ganancia más pequeño, que tiende a reducir la amplitud de la señal. Sin embargo, esta reducción de la amplitud de señal se considera menos perceptible o molesta. Preferiblemente, tal mínimo se determina por segmento de datos de salida.

5 En caso de que no haya metadatos de ganancia relacionados con el control de intervalo dinámico presentes en la primera corriente de audio, el método añade preferiblemente valores de ganancia suficientes para la protección contra el recorte en la segunda corriente de audio (corriente saliente). Estos valores de ganancia deberían preferiblemente limitarse de modo que no superen una ganancia de 1. La razón para evitar que los valores de ganancia superen 1 es que la señal no debería amplificarse innecesariamente para acercarse al límite de recorte.

De este modo, en el caso de que un segundo valor de ganancia calculado respectivo tenga una ganancia inferior a 1, el valor de ganancia añadido respectivo corresponderá al segundo valor de ganancia calculado. En caso de que un segundo valor de ganancia calculado respectivo sea superior a 1, el valor de ganancia añadido respectivo se establece en una ganancia de 1.

10 Un segundo aspecto de la solicitud se refiere a un aparato para proporcionar protección contra el recorte de señal de una señal de audio inferida a partir de datos de audio digital. El aparato está configurado para realizar el método descrito anteriormente. Las características del aparato corresponden a las características del método analizado anteriormente. Por consiguiente, el aparato comprende medios para determinar si los primeros valores de ganancia basados en los metadatos de audio recibidos son suficientes para la protección contra el recorte de señal de audio.
15 Además, el aparato comprende medios para sustituir un primer valor de ganancia por un valor de ganancia suficiente para protección contra recorte de señal de audio en caso de que el primer valor de ganancia no sea suficiente.

Preferiblemente, los medios de determinación comprenden medios para calcular segundos valores de ganancia basados en los datos de audio digital, siendo los segundos valores de ganancia suficientes para protección contra el recorte de la señal de audio. Más preferiblemente, los medios de determinación comprenden también medios de
20 comparación para comparar los primeros valores de ganancia basándose en los metadatos de audio recibidos y en los segundos valores de ganancia calculados. En dependencia con esto, se seleccionan los valores de ganancia a partir de los primeros valores de ganancia y de los segundos valores de ganancia calculados.

Las observaciones anteriores relacionadas con el primer aspecto de la solicitud son también aplicables al segundo aspecto de la solicitud.

25 Un tercer aspecto de la solicitud se refiere a un transcodificador, donde el transcodificador está configurado para transcodificar una corriente de audio desde un primer formato de codificación de audio a un segundo formato de codificación de audio. El transcodificador comprende el aparato según el segundo aspecto de la solicitud. Preferiblemente, el transcodificador es parte de un dispositivo de recepción que recibe la primera corriente de audio, donde la primera corriente de audio es una señal de difusión digital, por ejemplo una corriente de audio de una señal
30 de televisión digital (por ejemplo, DVB-T, DVB-S, DVB-C) o una señal de radio digital (por ejemplo, una señal DAB). Por ejemplo, el dispositivo de recepción es un decodificador. La corriente de audio se puede también transmitir a través de Internet (por ejemplo, TV por Internet o radio por Internet). Alternativamente, la primera corriente de audio puede leerse desde un medio de almacenamiento de datos digitales, por ejemplo un DVD (Disco Versátil Digital) o un disco Blu-ray.

35 Las observaciones anteriores relacionadas con los aspectos primero y segundo de la solicitud son también aplicables al tercer aspecto de la solicitud.

Breve descripción de los dibujos

La invención se explica a continuación a modo de ejemplo con referencia a los dibujos adjuntos, en los que

- la figura 1 ilustra una realización de un transcodificador que proporciona protección contra el recorte;
- 40 la figura 2 ilustra un enfoque preferido para reformular metadatos;
- la figura 3 ilustra una realización para determinar valores de pico basándose en datos de audio recibidos;
- la figura 4 ilustra una realización para fusionar datos de control de intervalo dinámico entrantes con valores de ganancia calculados suficientes para protección contra el recorte;
- la figura 5 ilustra la selección de los valores de ganancia salientes;
- 45 la figura 6 ilustra una realización alternativa para fusionar datos de control de intervalo dinámico entrantes con valores de ganancia calculados suficientes para protección contra el recorte;
- la figura 7 ilustra una realización de una etapa de filtro suavizante;
- la figura 8 ilustra otra realización para proporcionar protección contra el recorte;
- la figura 9 ilustra otra realización más para proporcionar protección contra el recorte; y
- 50 la figura 10 ilustra un dispositivo de recepción que recibe la corriente de audio transcodificada.

Descripción detallada

5 AAC/HE-AAC y Dolby Digital/Dolby Digital Plus soportan el concepto de metadatos, más específicamente palabras de ganancia que llevan una ganancia que varía en el tiempo para ser aplicada opcionalmente a los datos de audio tras la decodificación. Con el fin de reducir los datos, estas palabras de ganancia solo se envían típicamente una vez por segmento de datos, por ejemplo, por bloque o por trama. En dichos formatos de audio estas palabras de ganancia son opcionales, es decir, que es técnicamente posible no enviar los datos. Los codificadores Dolby Digital y Dolby Digital Plus envían típicamente las palabras de ganancia, mientras que los codificadores AAC y HE-AAC no envían a menudo las palabras de ganancia. Sin embargo, el número de codificadores AAC y HE-AAC que envían las palabras de ganancia está aumentando. La aplicación permite que los decodificadores o transcodificadores que reciben una corriente de audio hagan "lo correcto" en ambas situaciones. Si se proporcionan palabras de ganancia de audio, "lo correcto" sería procesar las palabras de ganancia de audio recibidas de la manera más veraz posible, pero anularlas cuando las palabras de ganancia entrantes no proporcionen suficiente atenuación para impedir el recorte de señal, por ejemplo en caso de una mezcla estereofónica. Si no se proporcionan valores de ganancia, "lo correcto" sería calcular y proporcionar valores de ganancia que impidan el recorte de señal.

15 La figura 1 muestra una realización de un transcodificador, en la que el transcodificador proporciona protección contra el recorte de señal, en particular protección contra el recorte en caso de mezclar de manera estereofónica (por ejemplo, mezclar de manera estereofónica una señal de 5,1 canales con una señal de 2 canales). El transcodificador recibe una corriente 1 de audio digital que comprende metadatos de audio. Por ejemplo, la corriente de audio digital es una corriente de audio digital AAC o HE-AAC (HE-AAC versión 1 o HE-AAC versión 2). La corriente de audio digital puede ser parte de una corriente de vídeo/audio DVB, por ejemplo una corriente DVB-T, DVB-S o DVB-C. El transcodificador, transcodifica la corriente 1 de audio recibida en una corriente 14 de audio de salida que está codificada en un formato diferente, por ejemplo Dolby Digital o Dolby Digital Plus. Típicamente, los decodificadores Dolby Digital soportan el mezclado de manera estereofónica de señales multicanal y asumen que las envolventes de ganancia que varían en el tiempo incluidas en los metadatos Dolby Digital recibidos incluyen protección contra el recorte de mezcla estereofónica. Desafortunadamente, la corriente 1 de bits (por ejemplo, una corriente de bits AAC/HE-AAC) no contiene necesariamente metadatos de envolvente de ganancia que varían en el tiempo, e incluso en caso de llevar tales datos no está claro si los datos incluyen protección contra el recorte. El transcodificador impide que un decodificador (por ejemplo, un decodificador Dolby Digital) en un dispositivo de recepción (aguas abajo del transcodificador) produzca señales de salida que contengan artefactos de recorte al mezclar la señal de manera estereofónica. El transcodificador garantiza que la corriente 14 de audio de salida contenga metadatos de envolvente de ganancia que varían en el tiempo, incluyendo protección contra el recorte de mezcla estereofónica.

En la figura 1, la unidad 2 lee los valores 3 de ganancia de control de intervalo dinámico contenidos en los metadatos de audio de la corriente 1 de audio. Opcionalmente, los valores 3 de ganancia se procesan adicionalmente en la unidad 5, por ejemplo, los valores 3 de ganancia se muestrean de nuevo y transcodifican según la temporización del segmento de datos de la corriente 14 de audio de salida transcodificada. El nuevo muestreo y la transcodificación de valores de ganancia de metadatos se analizan en el documento "Transcoding of dynamic range control coefficients and other metadata into MPEG-4 HE AAC", Wolfgang Schildbach et al., Audio Engineering Society Convention Paper, presentado en la 123ª Convención del 5 al 8 de octubre de 2007, Nueva York. La descripción de este documento, en particular los conceptos para el nuevo muestreo y la transcodificación de valores de ganancia de metadatos, se incorpora a la presente memoria como referencia. Además, el 30 de septiembre de 2008, el solicitante presentó la solicitud provisional de EE. UU., 61/101497 con el título "Transcoding of Audio Metadata", con la solicitud provisional de EE. UU. relacionada con el nuevo muestreo y con la transcodificación de valores de ganancia de metadatos. La descripción de esta solicitud, en particular los conceptos para nuevo muestreo y transcodificación de valores de ganancia de metadatos, se incorpora a la presente memoria como referencia.

45 En paralelo al nuevo muestreo, los datos de audio en la corriente 1 de audio se decodifican mediante un decodificador 6, típicamente a datos de audio de PCM (modulación de código de pulsos). Los datos 7 de audio decodificados comprenden una pluralidad de canales de señal paralelos, por ejemplo 6 canales de señal en el caso de una señal de canales 5.1, u 8 canales de señal en el caso de una señal de canales 7.1.

Una unidad 8 de cálculo determina los valores 9 de ganancia calculados basándose en los datos 7 de audio. Los valores 9 de ganancia calculados son suficientes para la protección contra el recorte de señal en un dispositivo de recepción aguas abajo del transcodificador que recibe la corriente de audio transcodificada, en particular cuando se mezcla la señal de manera estereofónica en el dispositivo de recepción. Tal dispositivo puede ser un AVR o un televisor. Los valores de ganancia calculados deben garantizar que la señal mezclada de manera estereofónica alcance como máximo 0 dBFS o menos. Los valores 4 de ganancia derivados de los metadatos en la corriente 1 de audio y los valores 9 de ganancia calculados se comparan entre sí en la unidad 10. La unidad 10 emite los valores 11 de ganancia, donde un valor de ganancia de la corriente 4 de valor de ganancia se reemplaza por un valor de ganancia inferido de la corriente 9 de valores de ganancia en caso de que el valor de ganancia respectivo de la corriente 4 de valor de ganancia no sea suficiente para impedir el recorte de señal en el dispositivo de recepción. En paralelo, los datos 7 de audio son codificados por el codificador 12 a un formato de codificación de audio de salida, como, por ejemplo a Dolby Digital o a Dolby Digital Plus. Los datos de audio codificados y los valores 11 de ganancia se combinan en la unidad 13. La corriente de audio resultante proporciona metadatos de ganancia de audio que impiden el recorte de señal, en particular para el caso de mezcla estereofónica de señal.

En general, los metadatos de ganancia de audio entrantes deben conservarse tanto como sea posible siempre que los metadatos de ganancia proporcionen protección contra recorte de señal. En la mayoría de los casos, la longitud de un segmento de datos (por ejemplo, bloque o trama) de la corriente de audio de entrada (véase 1 en la figura 1) y la longitud de un segmento de datos (por ejemplo, bloque o trama) de la corriente de audio de salida (véase 14 en la figura 1) son diferentes. Además, típicamente, el comienzo de un segmento de datos de la corriente de audio entrante y el comienzo de un segmento de datos de la corriente de audio saliente no están alineados (incluso si las longitudes de los segmentos de datos son idénticas). De este modo, típicamente es necesaria una correspondencia de los metadatos entrantes con los metadatos salientes.

La figura 2 ilustra un enfoque preferido para hacer corresponder metadatos entrantes con metadatos salientes. Como se analizó anteriormente, típicamente cada segmento de datos (por ejemplo, bloque o trama) tiene un valor de ganancia de datos de control de intervalo dinámico (o una pluralidad de valores de ganancia, por ejemplo, 8 valores de ganancia). Sin embargo, los metadatos transmitidos junto con un segmento de datos de entrada (por ejemplo, bloque o trama) tienen un área de impacto de control de intervalo dinámico (es decir, un intervalo en la corriente donde tiene efecto la aplicación del valor de ganancia) que a menudo no es exactamente tan grande como el segmento de datos, sino mayor. Esto se debe a las características de superposición y adición de la transformada utilizada (es decir, se utilizan ventanas que son mayores que el segmento de datos y las ventanas se superponen) y al hecho de que el control de intervalo dinámico se aplica a menudo en el dominio espectral. Lo mismo sucede a menudo con los datos de control de intervalo dinámico de la corriente de bits de audio saliente. En la figura 2, las líneas continuas marcan el comienzo y el final de un segmento 20-23 de datos en la corriente de entrada, y el comienzo y el final de un segmento 24-26 de datos en la corriente de salida. En la figura 2, cada área del impacto 30-33 y 34-36 de control de intervalo dinámico de un valor de ganancia se extiende más allá del final y del comienzo del segmento de datos respectivo. Cada área de impacto 30-33 y 34-36 está indicada por líneas de puntos y trazos.

Por ejemplo en HE-AAC, el tamaño de bloque es de 256 muestras, mientras que una ventana para decodificar tiene 512 muestras. La ventana total de 512 muestras puede considerarse como un área de impacto; sin embargo, el impacto del valor de ganancia en los bordes exteriores de las ventanas es menor en comparación con el impacto en la mitad de la ventana. Por tanto, el área de impacto puede también considerarse como una porción de la ventana. El área de impacto puede ser un número de muestras seleccionadas a partir del tamaño de bloque/trama (aquí: 256 muestras) hasta el tamaño de ventana (aquí: 512 muestras). Preferiblemente, el área de impacto utilizada es mayor que el tamaño del segmento de datos (bloque o trama).

Para determinar qué valores de control de intervalo dinámico de entrada influyen en un segmento de datos de salida dado, se prefiere observar la superposición de las áreas de impacto de entrada y salida (en lugar de observar la superposición de los segmentos de datos de entrada y salida). En la figura 2, se determina qué áreas de impacto 30-33 en la corriente de entrada se superponen con un área de impacto 34-36 de un segmento 24-26 de datos de salida dado. Por ejemplo, el área de impacto 34 del segmento 24 de datos en la corriente de salida se superpone con las áreas 30, 31, 32 y 33. Por lo tanto, preferiblemente, se consideran valores de ganancia asociados con cuatro segmentos 20, 21, 22 y 23 de datos al determinar el valor de ganancia del primer segmento 24 de datos en la corriente de salida ilustrada. El primer segmento 24 de datos está influido por los 4 segmentos 20-23 de datos de entrada. Alternativamente, el método puede observar la superposición de las áreas de impacto de entrada y el segmento de señal de salida, o la superposición de los segmentos de datos de entrada y el segmento de datos de salida.

Tal proceso de correspondencia o muestreo de nuevo se puede realizar en la unidad 5 de la figura 1, que recibe valores 3 de ganancia de la corriente 1 de entrada y hace corresponder uno o más de los valores 3 de ganancia a un valor 4 de ganancia.

La figura 3 ilustra una realización del bloque 50 para determinar valores de pico basándose en datos de audio recibidos. Tal bloque 50 de determinación de pico puede ser parte del bloque 8 de la figura 1. Basándose en los datos 7 de audio multicanal decodificados que comprenden una pluralidad de canales (aquí 5 canales de una señal de canales 5.1, el canal de efecto de baja frecuencia no se considera), el mezclar de manera estereofónica se realiza según uno o más esquemas de mezclar de manera estereofónica (es decir, según una o más matrices de mezcla estereofónica). Cabe señalar que el transcodificador no sabe en absoluto si se realiza el mezclado de manera estereofónica en el dispositivo de recepción ni qué esquema de mezclado de manera estereofónica se utiliza luego en el dispositivo de recepción. De este modo, se desconoce si una señal multicanal se reproduce en canales discretos o si se realiza el mezclado de manera estereofónica según uno de los varios esquemas. El transcodificador simula todos los casos y determina el peor caso.

En el ejemplo de la figura 3, el mezclado de manera estereofónica según el esquema de mezclado de manera estereofónica Lo/Ro se realiza en el bloque 41, el mezclado de manera estereofónica según el esquema de mezclado de manera estereofónica Pro Logic (PL) se realiza en el bloque 42, y el mezclado de manera estereofónica según el esquema de mezclado de manera estereofónica Pro Logic II (PL II) se realiza en el bloque 43. El esquema de mezclado de manera estereofónica PL y el esquema de mezclado de manera estereofónica PL II son dos variantes del esquema de mezclado de manera estereofónica Lt/Rt como se analizó anteriormente. Cada esquema de mezclado de manera estereofónica emite una señal de canal derecho y una señal de canal izquierdo. Luego, se calculan los valores absolutos de las señales después de la mezcla estereofónica (véanse los bloques 44 en la figura 3). Preferiblemente, se calculan también los valores de muestra absolutos de los diversos canales de la señal 7 de audio multicanal (véanse

los bloques 40 para determinar los valores absolutos). Obsérvese también que los valores absolutos de los canales (sin mezclar de manera estereofónica) son útiles para impedir el recorte de señal en otros casos distintos al del mezclar de manera estereofónica, por ejemplo en caso de que la señal se amplifique posteriormente con una ganancia adicional (por ejemplo, ganancia de 11 dB en el caso del modo RF, como se analiza más adelante).

5 El máximo (= valor pico) de los valores absolutos en cada tiempo se calcula en el bloque 45. El cálculo del máximo se realiza continuamente, generando por ello una corriente de valores 46 de pico. Puede ser posible que las diversas muestras tengan diferentes retardos de señal debido a procesamiento de señal diferente. Tales diferentes retardos de señal pueden estar alineados (no se muestra). El máximo de los valores de muestra indica la amplitud máxima que una señal puede tener para todos los casos, por lo que éste es el peor caso que tiene en cuenta el algoritmo de protección contra el recorte. De este modo, el transcodificador simula la amplitud de la señal del peor caso en el dispositivo de recepción. Un valor de control de intervalo dinámico que consiga protección contra el recorte debería atenuar (o amplificar) la señal de manera que alcance 0 dBFS como máximo.

10 Cabe señalar que el bloque 50 puede determinar un valor de pico basándose en menos valores absolutos que los ilustrados en la figura 3 (por ejemplo, sin considerar los valores absolutos de los canales no mezclados de manera estereofónica) o basándose en valores absolutos adicionales no mostrados en la figura 3 (por ejemplo, valores absolutos de otros esquemas de mezclado de manera estereofónica). Alternativamente, es posible mezclar de manera estereofónica los canales 7 sin determinar un valor de pico: por ejemplo, los dos canales resultantes pueden combinarse y la señal combinada se procesa adicionalmente (en lugar de usar los valores 46 de pico emitidos por el bloque 45).

15 El procesamiento adicional de los valores 46 de pico se indica en la figura 4. Los elementos figurativos de las figuras 1 y 4, indicados con los mismos signos de referencia, son básicamente iguales. Los valores 46 de pico se someten a una etapa de bloqueo y acumulación máxima en la unidad 60. Aquí, el valor de pico más alto se determina para un segmento de datos de salida dado (por ejemplo, un bloque). En otras palabras: los valores de pico se sub-muestran al seleccionar el valor de pico más alto (que es el más crítico) para un segmento de datos de salida entre una pluralidad de valores de pico. Cabe señalar que, preferiblemente, no sólo se consideran valores de pico consecutivos correspondientes a las muestras de señal en un segmento de salida para determinar el máximo. En su lugar, también se consideran valores de pico adicionales (anteriores y posteriores) que influyan en un segmento de datos dado, es decir, valores de pico que se refieren a muestras de señal al comienzo y al final de una ventana de decodificación. Preferiblemente, se consideran todas las muestras de la ventana.

20 El resultado de este muestreo se invierte en el bloque 61 según la fórmula $C = 1/X$, donde C se refiere a un valor 9 de ganancia calculado y X se refiere al pico más alto respectivo para el bloque de la corriente 14 de salida. El resultado C es un factor (ganancia) que garantiza que cada muestra de audio del segmento de datos (por ejemplo, bloque) esté por debajo o en igualdad al nivel 1 máximo de señal (correspondiente a 0 dBFS) cuando se aplica la ganancia a la muestra de audio respectiva. Esto evita el recorte de este segmento de datos. Cabe señalar que el nivel máximo de señal significa el nivel máximo de señal de una señal en el receptor de la corriente de audio transcodificada; de este modo, en la salida del bloque 60, la amplitud puede ser superior a 1 (cuando $C < 1$).

25 La ganancia C calculada es la ganancia máxima permitida que impide el recorte; También se puede utilizar un valor de ganancia menor que la ganancia C calculada (en este caso, la señal resultante es incluso menor). Cabe señalar que, en caso de que la ganancia C sea inferior a 1, se tiene que aplicar la ganancia C (o una ganancia menor); de lo contrario, la señal se recortaría al menos en el peor de los casos.

30 En el bloque 5, los valores 3 de ganancia entrantes de los metadatos se someten también a un nuevo muestreo. A partir de una serie de ganancias entrantes relevantes para un segmento de datos de salida, se elige la ganancia más pequeña y se utiliza para un procesamiento adicional. Preferiblemente, el nuevo muestreo se realiza como se analiza en relación con la figura 2: Para determinar qué valores de ganancia entrantes son relevantes para un segmento de datos de salida, se considera la superposición de las áreas de impacto de entrada y salida. Si el área de impacto de un segmento de datos entrantes se superpone con el área de impacto de un segmento de datos de salida dado, se considera el segmento de datos entrantes (y, de este modo, su valor de ganancia) al determinar el menor valor de ganancia. En su lugar, se pueden también utilizar los dos enfoques alternativos analizados en relación con la figura 2.

35 La motivación para esto es conservar los valores entrantes. Sin embargo, esto no es posible, ya que los valores de ganancia tienen que volver a muestrearse según la temporización de la corriente de salida. El uso del menor valor de ganancia de una pluralidad de valores de ganancia consecutivos tiende a reducir la amplitud de señal, que tiende a considerarse menos perceptible o molesta.

40 En caso de que estén presentes datos de control de intervalo dinámico relevantes en la corriente 1 de datos entrante, se realiza en el bloque 10 una comparación entre esta ganancia (preferiblemente después de volver a muestrear en el bloque 5) y los valores 9 de ganancia calculados suficientes para la protección contra el recorte. El bloque 62 determina el mínimo de entre un valor 4 de ganancia muestreado de nuevo y un valor 9 de ganancia calculado, utilizándose el valor de ganancia más pequeño como valor de ganancia saliente (el bloque 62 forma un selector mínimo).

5 En caso de que no haya presentes valores de ganancia entrantes, el conmutador 63 que aparece en la figura 4 conmutará a la posición superior, determinando luego el bloque 62 el valor mínimo de entre una ganancia de 1 y el valor de ganancia calculado, usándose el valor de ganancia más pequeño como valor de ganancia saliente. De este modo, en caso de que no haya presente ningún valor de ganancia entrante, el valor de ganancia saliente se limita a una ganancia máxima de 1.

La siguiente tabla ilustra el funcionamiento del bloque 10 de comparación. Aquí, el término "I" denota la ganancia 4 de control de intervalo dinámico entrante (después del nuevo muestreo), y el término "C" denota la ganancia 9 calculada.

	$I \leq 1$	$I > 1$	I no presente
$C \leq 1$	$\min(I, C)$	$\min(I, C) = C$	C
$C > 1$	$\min(I, C) = I$	$\min(I, C)$	1

10 En caso de que tanto I como C sean menores que, o iguales a 1, se toma el valor mínimo. Esto significa que, o bien I ya garantiza protección contra el recorte, o bien que, en caso contrario, será reemplazada por C.

En caso de que $C > 1$ e $I \leq 1$, la señal podría amplificarse y aún por ello no se recortaría. Sin embargo, la corriente entrante solicita atenuación, para, por ejemplo, cumplir fines de limitación de intervalo dinámico, y, de este modo, se conserva I (I es el valor mínimo de I y C en este caso).

15 En caso de $I > 1$ y $C \leq 1$, el valor entrante violaría la protección contra el recorte, por lo que se toma C (C es el valor mínimo de I y de C en este caso).

En caso de que tanto I como C sean mayores que 1, la entrada se amplificará. Esta amplificación está permitida siempre que no se produzca aún recorte, y, de este modo, se utiliza el menor valor de entre I y C.

20 En caso de que no esté presente ningún valor de intervalo dinámico entrante, la protección contra el recorte se garantiza utilizando C siempre que $C \leq 1$. En caso de $C > 1$, la señal no se modificará (es decir, la señal no se debe amplificar innecesariamente para acercarse al límite de recorte). Por lo tanto, la unidad se toma como la ganancia saliente. En ambos casos, cuando no hay valores de ganancia entrantes, se utiliza el mínimo valor de entre 1 y C (en lugar del valor mínimo de entre I y C).

25 La figura 5 ilustra la selección de los valores 11 de ganancia salientes en forma de diagrama de flujo. Se determina si está presente un valor I de ganancia (véase la referencia 130 de la figura 5). Si está presente actualmente un valor I de ganancia, el valor de ganancia saliente depende de los valores del valor I de ganancia entrante y del valor C de ganancia calculado. Si $I \leq 1$ y $C \leq 1$, el valor de ganancia seleccionado corresponde al mínimo de entre I y C (véase referencia 131). Si $I \leq 1$ y $C > 1$, el valor de ganancia seleccionado corresponde a I (véase la referencia 132). Si $I > 1$ y $C \leq 1$, el valor de ganancia seleccionado corresponde a C (véase la referencia 133). Si $I > 1$ y $C > 1$, el valor de ganancia seleccionado corresponde al mínimo de entre I y C (véase referencia 134). Cabe señalar que, en todos estos
30 cuatro casos, el valor saliente todavía corresponde al mínimo de entre I y C. De este modo, no es necesario determinar si I y C son ≤ 1 o no.

35 Si no está presente actualmente ningún valor I de ganancia, el valor de ganancia saliente depende del valor del valor C de ganancia calculado. Si $C \leq 1$, el valor de ganancia saliente corresponde a C (véase la referencia 135). Si $C > 1$, el valor de ganancia de salida corresponde a 1 (véase la referencia 136). Cabe señalar que en ambos casos el valor saliente aún corresponde al mínimo de entre 1 y C. De este modo, no es necesario determinar si C es ≤ 1 o no.

40 La realización analizada anteriormente consigue que la dinámica entrante se conserve, y, sólo en caso de que se produzca un recorte, la dinámica se modifica para impedir el recorte. En caso de que no haya valores de control de intervalo dinámico presentes, se añaden suficientes valores de control de intervalo dinámico a la corriente para impedir el recorte. La conmutación entre modos funciona de manera instantánea y fluida, mitigando por ello cualesquiera artefactos.

45 La figura 6 ilustra una alternativa a la realización de la figura 4 según la invención. Los elementos figurativos de las figuras 4 y 6, indicados por los mismos signos de referencia, son básicamente iguales. En la figura 6, se reciben y transcodifican metadatos de ganancia separados para dos modos diferentes, el modo de línea y el modo RF. En la realización de la figura 6 se calculan diferentes palabras de ganancia para el modo RF y para el modo de línea, porque utilizan dos tipos diferentes de metadatos. Los metadatos de modo de línea cubren un intervalo menor de valores y se envían con más frecuencia (típicamente, una vez por bloque), mientras que los metadatos de modo RF cubren un intervalo mayor de valores y se envían con menos frecuencia (típicamente, una vez por trama). En el modo RF, la señal se intensifica con una ganancia extraordinaria de 11 dB, lo que permite una mayor relación señal-ruido cuando se transmite la señal a través de un canal dinámicamente muy limitado (por ejemplo, desde un decodificador a la

entrada de RF de un televisor mediante un enlace de antena de RF analógica). Además, dado que los metadatos de ganancia del modo RF cubren una gama más amplia de valores que los metadatos de ganancia del modo de línea, el modo RF permite una mayor compresión de intervalo dinámico. Los metadatos de ganancia para el modo de línea se indican como "DRC" (véase el signo 3 de referencia), mientras que los metadatos de ganancia para el modo RF se indican como "compr" (véase el signo 3' de referencia). Obsérvese que en DVB los metadatos de ganancia para el modo RF se indican como "compresión" o "compresión fuerte". Además, la realización de la figura 6 considera también un nivel de referencia de programa (PRL), que según la invención se transmite como parte de los metadatos. El PRL indica un volumen de referencia del contenido de audio (por ejemplo, en HE-AAC, el PRL puede variar entre 0 dB y -31,75 dB). La aplicación de PRL reduce el volumen del audio a un nivel de referencia objetivo definido. En dependencia del formato de codificación de audio, son comunes otros términos para la referencia, por ejemplo nivel de diálogo, normalización del diálogo o dialnorm.

En la figura 6, el valor de pico más alto para un bloque de datos (como generado por la unidad 60) se ajusta el nivel en la unidad 70 en dependencia con el PRL recibido (típicamente, el PRL reduce el nivel). Para calcular los valores de ganancia asociados con el modo de línea, las muestras de nivel ajustado se invierten en el bloque 61, generando por ello valores de ganancia calculados que garantizan que cada muestra de audio del bloque está por debajo o es igual que el nivel máximo 1 de señal en caso de que la señal de audio sea ajustada en el receptor por el PRL. El nuevo muestreo de los datos 3 de DRC entrantes en el bloque 5, y la comparación de los valores 4 de ganancia muestreados de nuevo y los valores de ganancia calculados son idénticos a los de la figura 4.

Para calcular los valores de ganancia asociados con el modo RF, las muestras de nivel ajustado se amplifican en 11 dB en el bloque 71, ya que, en el receptor, la señal se amplifica también en 11 dB en caso de utilizar el modo RF. De este modo, el transcodificador simula el peor caso de amplitud de la señal en el dispositivo de recepción. Las muestras intensificadas se invierten en el bloque 61', generando por ello valores de ganancia calculados para el modo RF que garantizan que cada muestra de audio del bloque es inferior o igual a 1 (= amplitud máxima de señal) en caso de que la señal de audio sea ajustada en el receptor por el PRL e intensificada en 11 dB.

La realización de la figura 6 se utiliza preferiblemente para un transcodificador que emite una corriente de audio Dolby Digital (por ejemplo, un transcodificador HE-AAC a Dolby Digital o un transcodificador AAC a Dolby Digital). Según Dolby Digital, en el modo de línea, cada bloque de codificación tiene un valor de ganancia "DRC" (control de intervalo dinámico), mientras que en el modo RF cada trama (que comprende 6 bloques) tiene un valor de ganancia "compr". Sin embargo, ambos tipos de valores de ganancia se relacionan con el control de intervalo dinámico. El valor de ganancia calculado para el modo RF es sub-muestreado desde la tasa de bloque a la tasa de trama en el bloque 73. El bloque 73 determina el mínimo de los valores de ganancia calculados para un número total de 6 bloques consecutivos, asignando cada mínimo al valor 72 de ganancia calculado para toda la trama. El nuevo muestreo de los valores 3' de ganancia de compresión entrantes en el bloque 5' difiere del nuevo muestreo en el bloque 5, de tal manera que se determina el mínimo para una trama de salida. La comparación de los valores 4' de ganancia muestreados de nuevo y de los valores 72 de ganancia basados en trama calculados es la misma que se analizó anteriormente.

La realización en la figura 6 proporciona protección no sólo contra el recorte en caso del mezclado de manera estereoscópica, sino también contra el recorte de señal cuando se aplica una ganancia extraordinaria de 11 dB en el modo RF (de lo contrario, la señal intensificada de 11 dB puede recortarse incluso cuando no se use el mezclado de manera estereoscópica). Por lo tanto, es ventajoso considerar también en el bloque 50 los valores absolutos de los canales sin mezcla estereofónica.

Para calcular los valores de ganancia, se puede utilizar una etapa de suavizado. La figura 7 muestra una realización de una etapa 80 de suavizado que puede colocarse en cualquier lugar de la trayectoria existente entre la salida del bloque 50 y la entrada de los bloques 61 y 61'. Preferiblemente, la etapa 80 de suavizado se coloca a la salida del bloque 50, generando por ello valores 46 de pico suavizados basándose en los valores 46 de pico. La etapa 80 de suavizado implementa un filtro de paso bajo para la señal de entrada de la etapa de suavizado, por ejemplo, la señal de valor de pico. Su fin es mejorar la impresión audible después de haberse lanzado la protección contra el recorte en: una liberación inmediata de una ganancia de inmersión después de un período de protección contra el recorte sonaría como molesta. De este modo, como se hace ampliamente en implementaciones de limitadores, la señal de valor de pico (y por eso la señal de ganancia inferida; véase más adelante) se filtra con un filtro de paso bajo de primer orden, que preferiblemente funciona a una constante τ de tiempo de 200 ms. En caso de que un nuevo valor de entrada demande protección contra el recorte en un grado mayor del que conseguiría la señal suavizada (dado que el nuevo valor de entrada es mayor que la señal suavizada), se puentearía la etapa de suavizado y se llevaría a efecto inmediatamente. En este caso, la entrada superior es mayor que la entrada inferior del bloque 81 de cálculo máximo en la figura 7.

Preferiblemente, la realización de las figuras 3-7 es parte de un transcodificador de audio, por ejemplo de AAC y/o HE-AAC a Dolby Digital, o de Dolby E o Dolby Digital a AAC y/o HE-AAC. Sin embargo, cabe señalar que las realizaciones de las figuras 3-7 no son necesariamente parte de un transcodificador de audio. Estas realizaciones pueden ser parte del dispositivo que recibe la corriente 1 de audio entrante y aplica los valores de ganancia modificados (sin transcodificación). Los valores de ganancia modificados pueden usarse directamente para ajustar la ganancia de la corriente de audio recibida. Por ejemplo, las realizaciones de las figuras 3-7 pueden ser parte de un AVR o un televisor.

La figura 8 ilustra una realización alternativa para proporcionar protección de mezcla estereofónica. El aparato recibe palabras 90 de ganancia entrantes contenidas en o inferidas a partir de metadatos de audio. Las palabras 90 de ganancia pueden corresponder a los valores 3 o 4 de ganancia de las figuras 1 y 4. Además, el aparato recibe muestras 91 de audio (por ejemplo, muestras de audio de PCM). Por ejemplo, las muestras 91 de audio pueden ser valores de pico generados por el bloque 50 en la figura 3. Si las muestras 91 de audio no son valores absolutos, el valor absoluto de las muestras 91 de audio puede determinarse antes. En el bloque 92, los valores de ganancia máximos permitidos, $ganancia_{m\acute{a}x}(t)$, se calculan mediante una división según la siguiente ecuación:

$$ganancia_{m\acute{a}x}(t) = \frac{se\tilde{n}al_{m\acute{a}x,permitida}}{se\tilde{n}al(t)}$$

Aquí, el término $se\tilde{n}al_{m\acute{a}x,permitida}$ denota la amplitud de señal máxima permitida, por ejemplo $se\tilde{n}al_{m\acute{a}x,permitida} = 1$. El término $se\tilde{n}al(t)$ denota la muestra 91 de audio actual.

En el bloque 93, los valores de ganancia máximos permitidos $ganancia_{m\acute{a}x}(t)$ se limitan a una ganancia máxima de 1: si un valor $ganancia_{m\acute{a}x}(t)$ está por encima de 1, entonces $ganancia_{m\acute{a}x}(t)$ se establecerá en 1. Sin embargo, si un valor $ganancia_{m\acute{a}x}(t)$ es inferior a 1 o igual a 1, el valor no se modificará.

La salida del bloque 93 se alimenta a una etapa 94 de filtro de suavizado. La etapa 94 de filtro de suavizado contiene un filtro de paso bajo y un selector 95 de mínimo que selecciona el valor mínimo de sus dos entradas. La operación es similar a la de la etapa 80 de filtro de suavizado de la figura 7. Sin embargo, aquí, se usa un selector 95 de mínimo en lugar de un selector 81 de máximo, ya que la etapa 94 de filtro suaviza los valores de ganancia en lugar de muestras de audio (los valores de ganancia se infieren invirtiendo muestras de audio). En su lugar, se puede utilizar una etapa 80 de filtro de suavizado cuando se coloca aguas arriba del bloque 92 (que determina los valores de ganancia mediante inversión). De manera análoga, se puede utilizar la etapa 94 de filtro de suavizado en las figuras 4 y 5 cuando se colocan aguas abajo de los bloques 61 y/o 61' (ya que se procesan señales de ganancia aguas abajo de los bloques 61 y/o 61'). La etapa 94 de filtro de suavizado suaviza la pendiente de señal en caso de un aumento brusco del valor de ganancia en el bloque 93 (de otro modo, el audio puede sonar molesto). Por el contrario, la etapa 94 del filtro de suavizado permite que la señal de ganancia pase sin suavizarse en caso de una disminución brusca del valor de ganancia (de otro modo, la señal sería recortada). La señal 96 de ganancia calculada a la salida de la etapa 95 de filtro de suavizado se compara con las palabras 90 de ganancia entrantes en el selector 97 de mínimos. El mínimo del valor 96 de ganancia calculado real y la palabra 90 de ganancia entrante real se pasan a la salida del selector 97 de mínimos. Los valores 98 de ganancia a la salida del selector 97 de mínimos proporcionan protección de mezcla estereofónica y pueden integrarse en una corriente de audio transcodificada como se analizó anteriormente.

Cabe señalar que la realización de la figura 8 no forma necesariamente parte de un transcodificador de audio. Los valores de ganancia de salida se pueden utilizar directamente para ajustar el nivel de la corriente de audio recibida. En este caso, el aparato de la figura 8 puede ser parte de un AVR o de un televisor.

Además, la realización de la figura 8 se puede utilizar para impedir el recorte de señal sin considerar el mezclado de manera estereoscópica. Por ejemplo, la realización de la figura 8 puede recibir muestras 91 de audio de PCM convencionales sin procesamiento previo adicional en el bloque 50. En este caso, la realización de la figura 8 impide el recorte cuando las muestras 91 de PCM son amplificadas por los valores de ganancia de salida.

La figura 9 ilustra otra realización alternativa. Elementos figurativos en las figuras 8 y 9, indicados con los mismos signos de referencia, son básicamente iguales. En contraste con la realización de la figura 8, la realización de la figura 9 es una versión operativa a modo de bloque como las realizaciones de las figuras 4 y 6, donde sólo se realiza una división por bloque de señal (o cualquier otro segmento de datos como trama). Esto reduce el número de divisiones por tiempo. Como ya se analizó en relación con la figura 8, las muestras 91 de audio pueden ser generadas por el bloque 50 de la figura 3. Si las muestras 91 de audio no son valores absolutos, los valores absolutos de las muestras 91 de audio pueden determinarse antes (no se muestran en la figura 9). Las muestras 91 de audio son alimentadas luego a una etapa 80 de filtro de suavizado que corresponde a la etapa 80 de filtro de suavizado de la figura 7. A diferencia de la figura 8, la etapa 80 de filtro de suavizado procesa muestras de audio en lugar de muestras de ganancia. De este modo, la etapa 80 de filtro de suavizado utiliza un selector 81 de máximos en lugar de un selector 95 de mínimos. Después del suavizado, el máximo de muestras por bloque de audio se determina en la unidad 100. Luego, el valor máximo se invierte en el bloque 101, calculándose por ello la máxima ganancia permitida por bloque. Este valor de ganancia se compara con el valor 90 de ganancia actual en el selector 97 de mínimos, siendo hecho pasar el mínimo de ambos valores a la salida del selector 97 de mínimos. Los valores 98 de ganancia en la salida del selector 97 de mínimos proporcionan protección contra el recorte de mezcla estereofónica y pueden ser integrados en una corriente de audio transcodificada como se analizó anteriormente. La realización de la figura 9 puede modificarse para generar un valor 98 de ganancia de una manera similar cuando no hay presente ningún valor 90 de ganancia entrante: Si no hay presente ningún valor 90 de ganancia entrante y la ganancia calculada es menor o igual a 1, se emite el valor de ganancia calculado. En caso de que el valor de ganancia calculado sea mayor que 1 (y no haya presente ningún valor 90 de ganancia entrante), se emite un valor de ganancia con una ganancia de 1. Esto se puede realizar mediante el conmutador adicional 63 de la figura 6, conmutando el conmutador entre el valor 90 de ganancia entrante y una ganancia de 1 en dependencia de la presencia del valor 90 de ganancia entrante.

Cabe señalar que las realizaciones analizadas anteriormente corresponden a un limitador que respeta los valores de ganancia provenientes de una instancia de compresor diferente.

5 La figura 10 ilustra un dispositivo de recepción que recibe la corriente 14 de audio transcodificada generada por el transcodificador de la figura 1. El bloque 121 separa los valores 11 de ganancia de la corriente 14 de audio. El dispositivo de recepción comprende adicionalmente un decodificador 110 que genera una señal 120 de audio decodificada. La amplitud de la señal 120 de audio decodificada es ajustada en el bloque 112 por los valores 11 de ganancia como se infieren en la figura 1. En caso de que se realice una mezcla estereofónica opcional en el bloque 113, la señal 114 de salida no se recorta, ya que los valores 11 de ganancia son suficientes para impedir el recorte de señal en caso de una mezcla estereofónica. La amplitud de la señal 120 de audio decodificada puede ajustarse
10 adicionalmente mediante el PRL (no mostrado). En caso de que los valores 11 de ganancia consideren también una intensificación de 11 dB en el modo RF como se analizó en relación con la figura 6, la señal 120 de audio puede también intensificarse en 11 dB sin recorte (tanto en el caso de que se produzca una mezcla estereofónica de señal como en el caso de que no se produzca mezcla estereofónica de señal).

REIVINDICACIONES

1. Un método para proporcionar protección contra el recorte en una señal de audio mezclada de manera estereofónica, comprendiendo el método:
- 5 recibir datos (7) de audio que comprenden una pluralidad de canales y sus metadatos relevantes, en donde los metadatos relevantes incluyen metadatos que indican un volumen de datos de audio;
- mezclar de manera estereofónica los datos (7) de audio que comprenden la pluralidad de canales en una señal de audio mezclada de manera estereofónica;
- determinar valores (4) de ganancia basándose en los metadatos relevantes recibidos;
- 10 determinar si los valores (4) de ganancia son suficientes para la protección contra el recorte de la señal de audio mezclada de manera estereofónica; en donde los valores (4) de ganancia son suficientes para la protección contra el recorte, si la aplicación de los valores (4) de ganancia impide que una amplitud de la señal de audio mezclada de manera estereofónica supere un valor máximo representable;
- en respuesta a determinar que los valores (4) de ganancia son suficientes para la protección contra el recorte, aplicar los valores (4) de ganancia a la señal de audio mezclada de manera estereofónica,
- 15 en caso de que los valores (4) de ganancia no sean suficientes para la protección contra el recorte, reemplazar los valores (4) de ganancia con nuevos valores (9) de ganancia suficientes para la protección contra el recorte, de modo que la aplicación de los nuevos valores (9) de ganancia impida que la amplitud de la señal de audio mezclada de manera estereofónica supere el valor máximo representable; y aplicar los nuevos valores (9) de ganancia a la señal de audio mezclada de manera estereofónica.
- 20 2. El método de la reivindicación 1, en donde los nuevos valores (9) de ganancia se calculan basándose en los datos de audio que comprenden la pluralidad de canales.
3. El método de la reivindicación 2, en donde la etapa de determinar si los valores (4) de ganancia son suficientes para la protección contra el recorte comprende la etapa de determinar el máximo de una pluralidad de valores de señal consecutivos derivados de los datos de audio que comprenden la pluralidad de canales.
- 25 4. Un aparato de procesamiento de audio para proporcionar protección contra el recorte en una señal de audio mezclada de manera estereofónica, en donde el aparato de procesamiento de audio está configurado para:
- recibir datos (7) de audio que comprenden una pluralidad de canales y sus metadatos relevantes, en donde los metadatos relevantes incluyen metadatos que indican un volumen de los datos de audio;
- 30 mezclar de manera estereofónica los datos (7) de audio que comprenden la pluralidad de canales en una señal de audio mezclada de manera estereofónica;
- determinar valores (4) de ganancia basándose en el volumen de los datos de audio;
- determinar si los valores (4) de ganancia son suficientes para la protección contra el recorte de señal de audio mezclada de manera estereofónica; en donde los valores (4) de ganancia son suficientes para la protección contra el recorte, si la aplicación de los valores (4) de ganancia impide que una amplitud de la señal de audio mezclada de manera estereofónica supere un valor máximo representable;
- 35 en respuesta a determinar que los valores (4) de ganancia son suficientes para la protección contra el recorte, aplicar los valores (4) de ganancia a la señal de audio mezclada de manera estereofónica,
- en caso de que los valores (4) de ganancia no sean suficientes para la protección contra el recorte, reemplazar los valores (4) de ganancia con nuevos valores (9) de ganancia suficientes para la protección contra el recorte, de tal modo que la aplicación de los nuevos valores (9) de ganancia impida que la amplitud de la señal de audio mezclada de manera estereofónica supere el valor máximo representable; y aplicar los nuevos valores (9) de ganancia a la señal de audio mezclada de manera estereofónica.
- 40 5. El aparato de procesamiento de audio de la reivindicación 4, en donde los nuevos valores (9) de ganancia están calculados basándose en los datos de audio que comprenden la pluralidad de canales.
- 45 6. El aparato de procesamiento de audio de la reivindicación 5, en donde la etapa de determinar si los valores (4) de ganancia son suficientes para la protección comprende la etapa de determinar el valor máximo de entre una pluralidad de valores de señal consecutivos inferidos de los datos de audio que comprenden la pluralidad de canales.

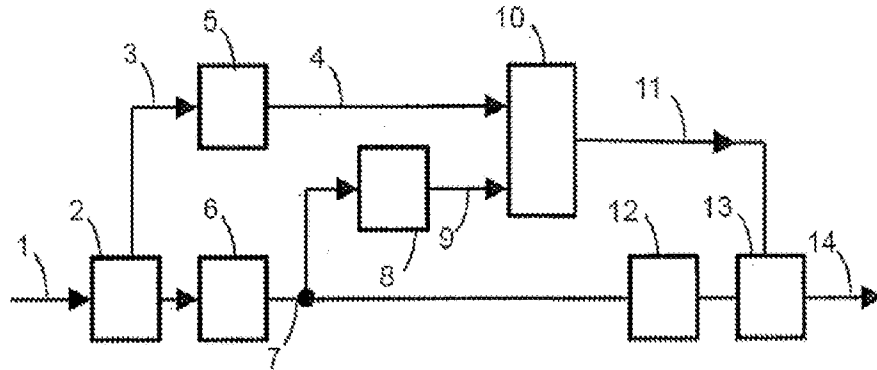


FIG. 1

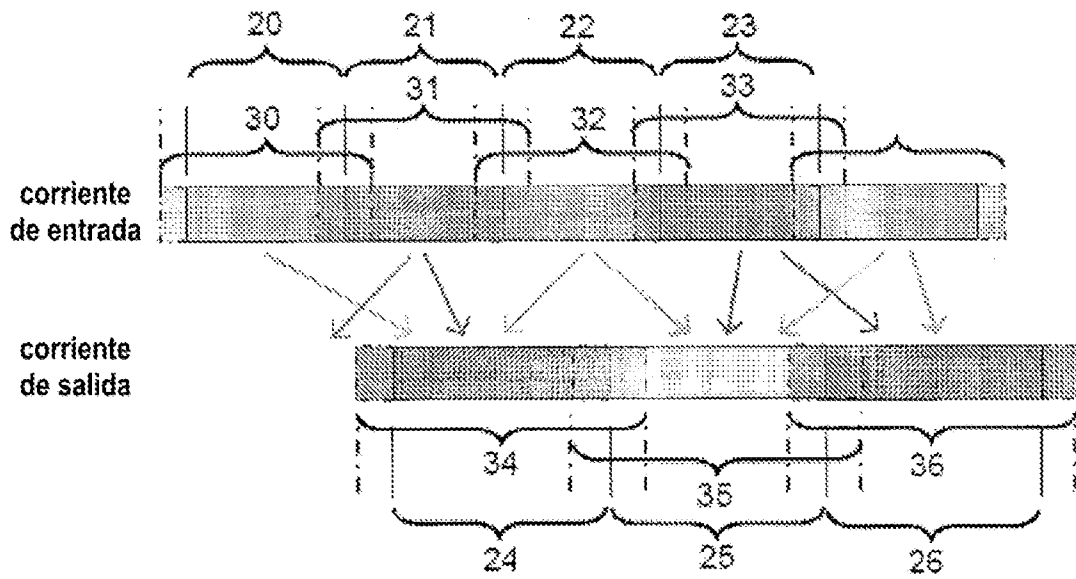


FIG. 2

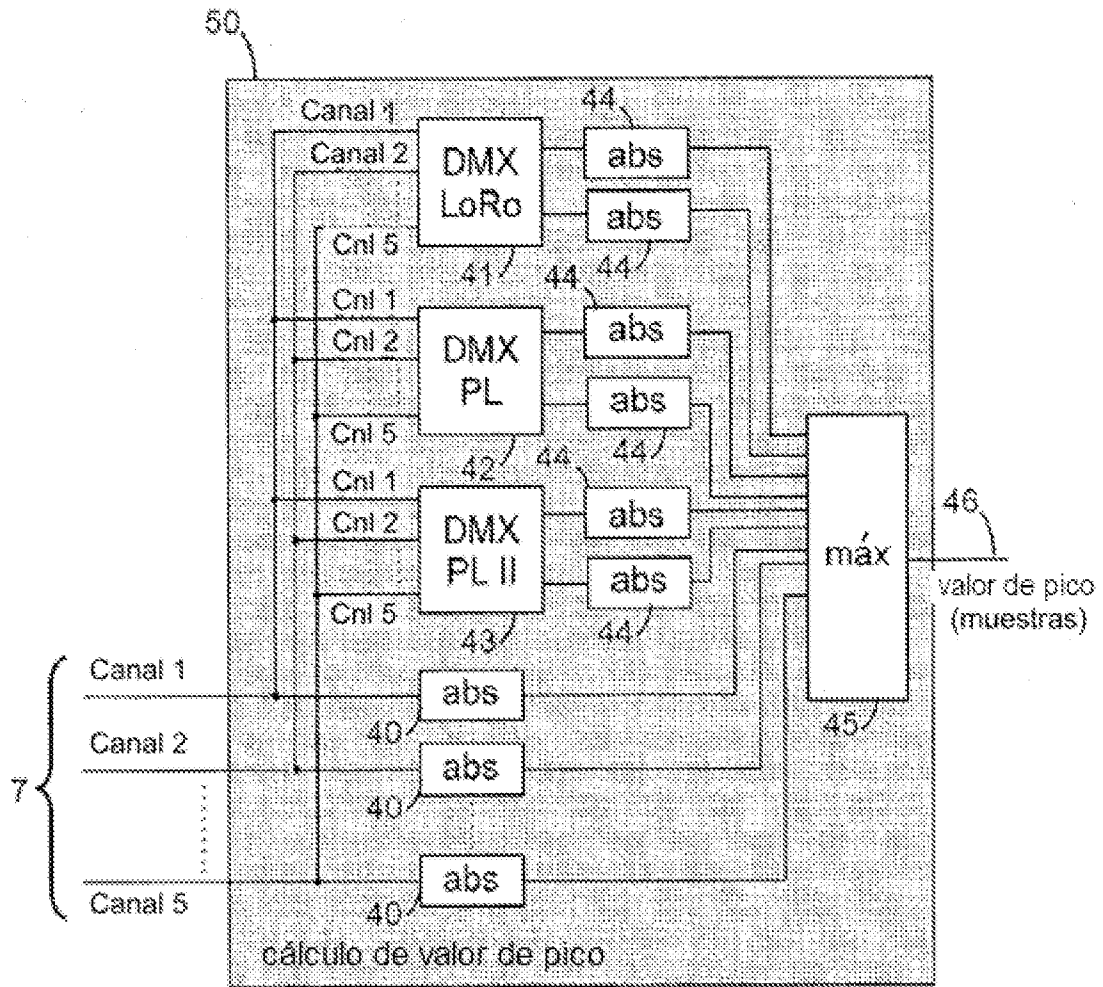


FIG. 3

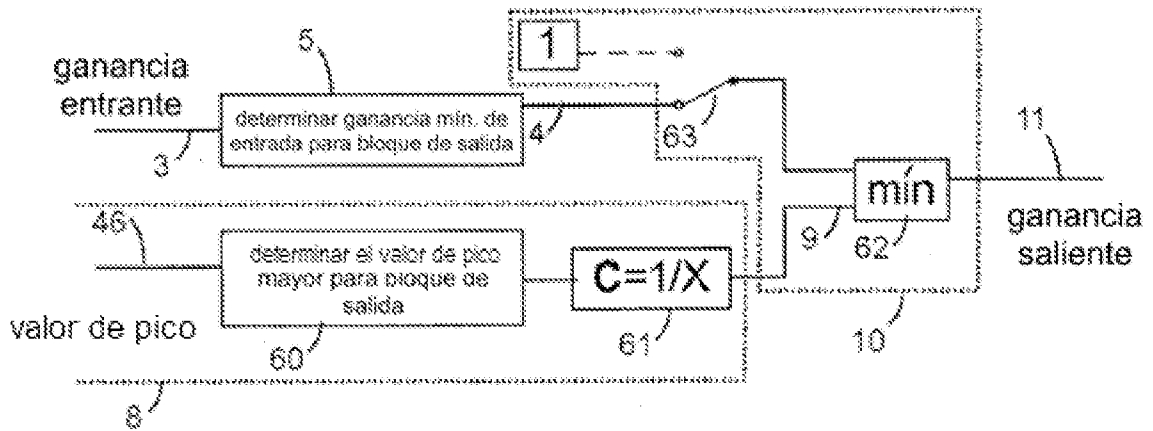


FIG. 4

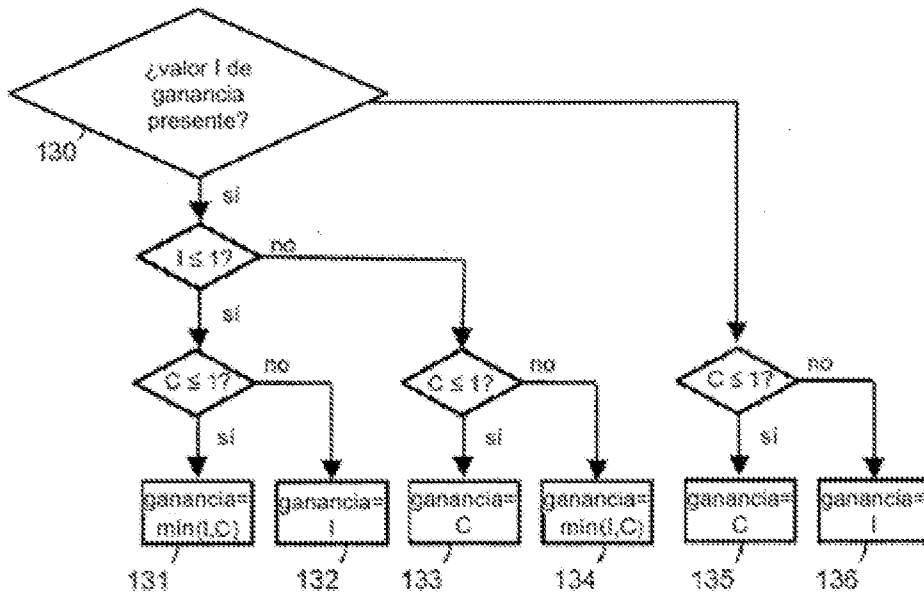


FIG. 5

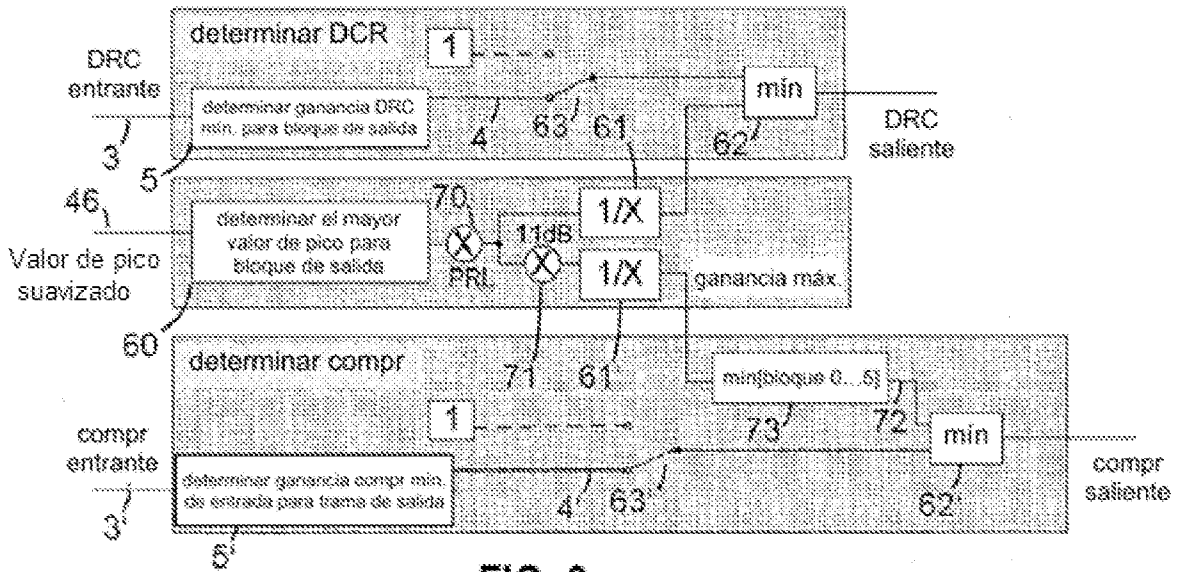


FIG. 6

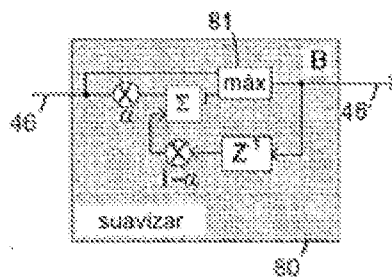


FIG. 7

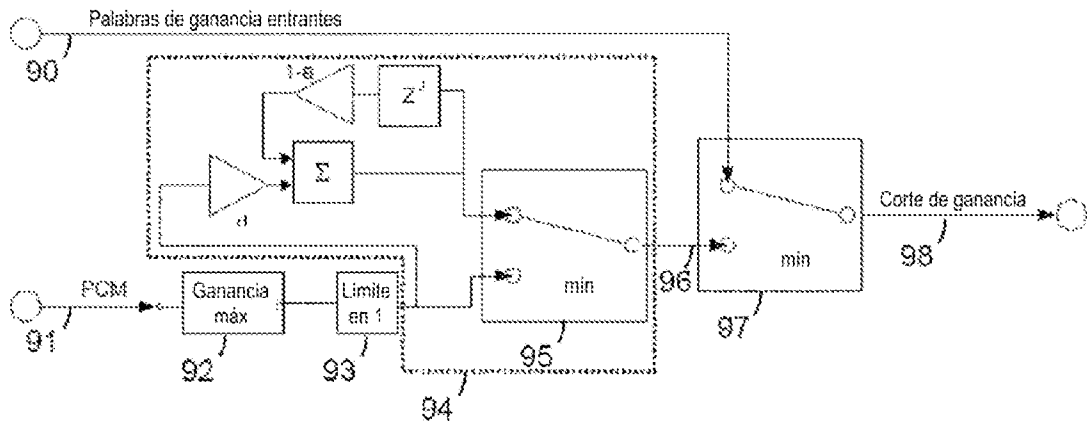


FIG. 8

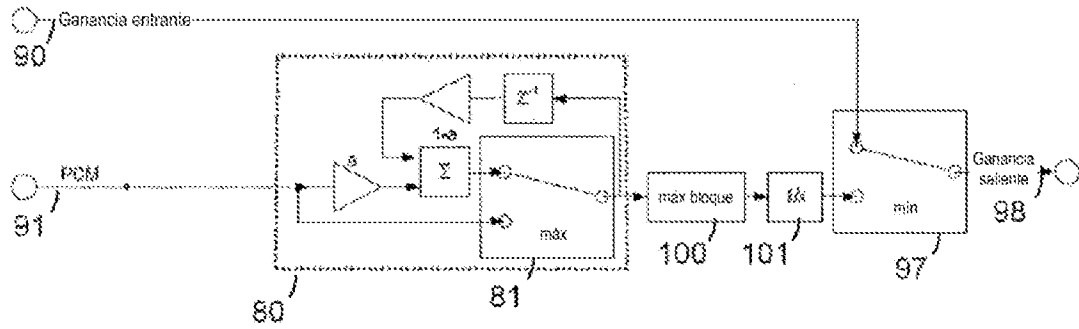


FIG. 9

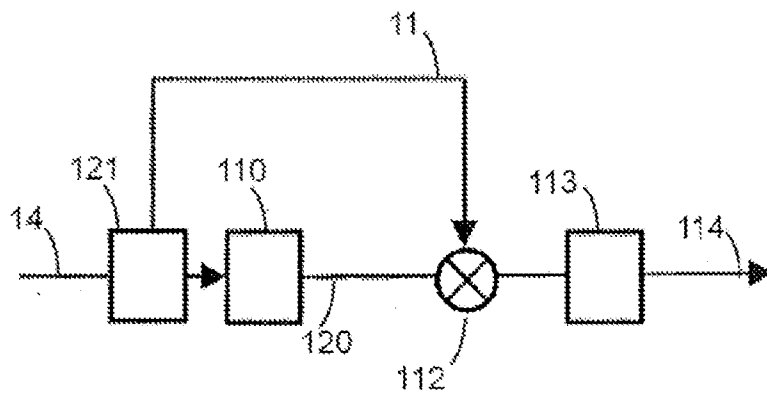


FIG. 10