

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 502 468**

51 Int. Cl.:

**H03G 3/00**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.09.2011 E 11757498 (8)**

97 Fecha y número de publicación de la concesión europea: **30.07.2014 EP 2619904**

54 Título: **Mezcla de transmisión de audio con normalización de nivel de diálogo**

30 Prioridad:

**22.09.2010 US 385428 P**

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

**03.10.2014**

73 Titular/es:

**DOLBY LABORATORIES LICENSING  
CORPORATION (50.0%)  
100 Potrero Avenue  
San Francisco, CA 94103-4813, US y  
DOLBY INTERNATIONAL AB (50.0%)**

72 Inventor/es:

**GROESCHEL, ALEXANDER;  
WILLIAMS, PHILLIP A.;  
COOPER, JARRET A. y  
SCHILDBACH, WOLFGANG A.**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

**ES 2 502 468 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

Mezcla de transmisión de audio con normalización de nivel de diálogo.

### 5 Campo técnico

Esta solicitud se refiere al campo de las señales de audio. En particular, se refiere a la mezcla de señales de audio.

#### Antecedentes de la invención

10

La inclusión de metadatos junto con señales de audio ha permitido significativas mejoras en la experiencia de audición del usuario. Para una experiencia del usuario agradable, generalmente es deseable para el nivel de sonido general o el volumen sonoro de los diferentes programas que sea consistente. No obstante, las señales de audio de los diferentes programas que usualmente provienen de diferentes fuentes, son controladas por diferentes  
15 productores y pueden contener contenido diverso que va desde diálogo de discurso a música a pistas de audio de películas con efectos de baja frecuencia. Esta posibilidad de diferencia en el nivel de sonido hace que sea un reto mantener el mismo nivel de sonido general a lo largo de tal variedad de programas durante la reproducción. En términos prácticos, no es conveniente para el oyente sentir la necesidad de ajustar el volumen de reproducción cuando cambie de un programa a otro con el fin de ajustar un programa para que esté más alto o más bajo con  
20 respecto a otro programa debido a las diferencias en el nivel de sonido percibido de los diferentes programas. Las técnicas para cambiar las señales de audio con el fin de mantener un nivel de sonido consistente entre los programas se conocen generalmente como nivelación de señal. En el contexto de las pistas de audio de diálogo, una medida relacionada con el nivel de sonido percibido se conoce como el nivel de diálogo, el cual se basa en un nivel ponderado de media de la señal de audio. El nivel de diálogo se especifica con frecuencia mediante el uso de  
25 un parámetro dialnorm, el cual indica un nivel en decibelios (dB) con respecto a la escala digital completa.

En el pasado, las emisoras que funcionaban con señales de audio tenían particulares problemas con las señales de audio tales como las pistas de audio cuyos niveles de audio estaban por encima o por debajo de los de otros programas, particularmente el audio que puede variar sustancialmente con el tiempo, tal como el diálogo. Con el  
30 desarrollo del audio digital, el audio de multicanal y, particularmente, la capacidad de incluir metadatos junto con la señal de audio, los productores e ingenieros de audio tienen ahora una amplia gama de opciones para especificar ajustes, los cuales se pueden incrustar en la señal como metadatos con el fin de especificar de forma precisa los niveles de reproducción para diversos sistemas de reproducción. Estos ajustes se pueden proporcionar incluso en la fase de postproducción, de forma que las emisoras puedan entregar una señal de audio muy consistente y  
35 garantizar que los elementos de audio más importantes lleguen al usuario final.

De manera similar, cuando se mezclan señales de audio, es conveniente para una experiencia de usuario agradable mantener también el mismo nivel de sonido percibido cuando se mezclan señales de entrada de audio en una única  
40 señal. Una técnica para llevar a cabo este objetivo es para las señales de entrada incluir metadatos de mezcla que especifican cómo se debería modificar la escala de la señal cuando se mezcla.

Muchos estándares de audio actuales permiten al productor de contenido incluir señales de audio asociadas acopladas con la señal de audio principal incluyendo metadatos variables en el tiempo junto con las señales de  
45 audio asociadas. Por ejemplo, un productor de contenido podría proporcionar una pista con los comentarios del director con una señal de audio asociada tal. Los metadatos que acompañan a la señal asociada especifican exactamente cómo desea el productor de contenido que la señal de audio de la pista principal se ajuste durante la mezcla para la reproducción combinada. Por ejemplo, E-AC-3 (Dolby Digital Plus) y el Codificador avanzado de audio de alta eficiencia (HE-AAC) son dos ejemplos de estándares que proporcionan tales metadatos de mezcla. Para obtener más detalles, consulte "ETSI TS 102 366 v1.2.1 (2008-08): Estándar de compresión de audio digital  
50 (AC-3, AC-3 mejorado)", el cual describe E-AC-3 (Dolby Digital Plus); o consulte "ETSI TS 101 154 V1.9.1 (2009-09): Difusión de video digital (Digital Video Broadcasting) (DVB); especificación para el uso de codificación de audio y vídeo en las aplicaciones de difusión basada en la trama de transporte MPEG-2", que describe el Codificador avanzado de audio de alta eficiencia (HE-AAC). Ambos incorporados por la presente en su totalidad como referencia.

55

No obstante, un usuario puede desear divergir de los ajustes proporcionados por el productor, los cuales son dictados por los metadatos transmitidos junto con la señal asociada. Por ejemplo, un usuario que activa los comentarios del director mientras ve una película puede en algún momento durante la reproducción decidir que preferiría oír el diálogo original que el productor puede haber indicado en los metadatos que se atenúe en la mezcla

con el fin de que no se anteponga a los comentarios del director.

De este modo, existe una necesidad de proporcionar un ajuste que permita al usuario ajustar la mezcla de las señales de audio de entrada a la vez que se proporciona también una experiencia de usuario agradable y se mantiene también el nivel de sonido percibido de la señal mezclada. Además, existe también una necesidad de proporcionar un ajuste de la mezcla de las señales de audio de entrada a la vez que se mantiene un nivel de sonido percibido consistente para la señal mezclada, incluso si la información de modificación de escala de los metadatos y una entrada de usuario externa pueden cambiar con el tiempo, por lo que no existiría la necesidad de realizar una nivelación adicional en la señal mezclada.

10

### **Resumen**

El uso de metadatos de mezcla permite al productor de contenido controlar y especificar de forma precisa cómo se deberían combinar las señales durante la reproducción. De este modo, el nivel de sonido o nivel de diálogo de las señales de audio de entrada se puede normalizar antes de mezclar por la información de modificación de escala proporcionada por los metadatos, de tal forma que la escala de cada entrada de señal al mezclador se modifique adecuadamente para lograr un nivel de sonido percibido consistente. De esta forma, el nivel relativo de cada señal de entrada que contribuye a la señal mezclada se controla durante la mezcla por medio de la información de modificación de escala contenida en los metadatos, que se proporciona con la señal. En la práctica, esta técnica usualmente conlleva la identificación de una señal principal y una o varias señales asociadas para que se mezclen. Como sugieren los nombres, la señal principal es la señal estándar y una señal asociada es una señal relacionada de alguna forma con su respectiva señal principal. Por consiguiente, los metadatos proporcionan entonces información de modificación de escala para la señal principal relativa a la señal asociada cuando se mezclan.

25 Por ejemplo, con frecuencia como una característica especial, un productor puede incluir una pista de voz de "comentarios del director" adicional como una señal asociada junto con una película, donde el usuario puede escuchar los comentarios del director superpuestos en tiempo real en la parte superior de la película. De este modo, el usuario puede ver la película con su audio original y los comentarios del director al mismo tiempo. Durante la reproducción, una señal asociada tal está destinada a mezclarse con la señal de audio principal de la película original, por ejemplo en un dispositivo de usuario final tal como un convertidor y decodificador. No obstante, otro diálogo, efectos y música de la señal de audio principal de la película pueden estar muy altos y, por consiguiente, enmascarar el audio asociado. De este modo, generalmente es conveniente reducir o atenuar la señal de audio principal de la película, en algunas ocasiones para que los comentarios del director se puedan entender adecuadamente sobre el audio principal. Esta atenuación de la señal de audio principal se puede conseguir, por ejemplo, proporcionando metadatos acompañando a la señal de audio asociada adicional, en la cual los metadatos especifican exactamente en qué forma y en qué medida se debería atenuar la señal principal.

Además, esta atenuación puede necesitar variar con el tiempo. Por ejemplo, en una película de acción, puede ser necesario en la mezcla de las señales atenuar mucho la señal de audio principal durante las escenas de mucho ruido, lo que puede incluir aspectos tales como vehículos al acelerar con chirridos de ruedas o la explosión de aviones y misiles, sobre los cuales sería difícil para el usuario final escuchar suficientemente los comentarios del director de la señal asociada. Por ejemplo, el director podría desear explicar cómo se consiguió la mayor explosión de la escena de acción incluyendo cables de guía y cámaras remotas especiales y cómo se sincronizaron los aspectos, todo al mismo tiempo que la escena se ejecuta en segundo plano. En este caso, el productor de contenido puede incluso desear atenuar mucho más los efectos de baja frecuencia mientras se mantiene todavía el nivel de sonido normal del diálogo. En otras palabras, los niveles de sonido elevados de ciertos tipos de audio principal pueden en ocasiones bloquear los comentarios del director. En otras ocasiones, tales como una escena íntima silenciosa en una película, el director puede desear mantener el nivel de sonido completo de la película, por ejemplo, de forma que el usuario final puede escuchar de forma precisa el susurro de un mensaje importante en una escena crucial de la película.

No obstante, el usuario puede desear tener la capacidad de ajustar los ajustes proporcionados por el productor de contenido. Por ejemplo, un usuario que ve una película con los comentarios del director activados puede en algún momento durante la reproducción decidir que preferiría darle más importancia a escuchar los efectos, sonidos, música y diálogo originales que los comentarios del director. No obstante, como los ajustes de metadatos del productor pueden indicar que la señal principal se debería atenuar mucho en la mezcla con el fin de no anteponerse a los comentarios del director, el usuario no tiene una forma directa de modificar los ajustes de mezcla proporcionados por el productor de contenido (el usuario podría por supuesto subir el volumen con el fin de escuchar mejor los sonidos atenuados, pero con el efecto no deseado de que los demás sonidos no atenuados también

estarán más altos). De este modo, existe una necesidad de proporcionar una forma de personalizar la mezcla de las señales de audio de entrada.

Para lograr esta función, el usuario podría contar con un control de entrada para regular el balance entre las señales principal y asociada. No obstante, dado que la mezcla claramente influye en el nivel de sonido, sin medidas adicionales el nivel de sonido podría desajustarse fácilmente en el proceso, conduciendo a saltos en el volumen sonoro al cambiar de un programa con una señal mezclada regulada por una entrada de usuario a un programa con una señal de audio no mezclada. Además, debido a que los metadatos y la entrada de usuario pueden, ambos, variar con el tiempo, sin un manejo cuidadoso, el nivel de sonido percibido de la señal mezclada puede ser también variable en el tiempo. De este modo, existe una necesidad adicional de modificar la escala de las señales de entrada principal y asociada durante la mezcla de tal forma que el nivel de sonido percibido de la señal mezclada de un programa permanezca constante.

De acuerdo con un aspecto, se revela un método que permite el mantenimiento de un nivel de sonido percibido consistente para la señal mezclada mediante el mantenimiento del nivel de sonido de la señal dominante en la constante de mezcla y el ajuste del nivel de sonido de la señal no dominante en relación con la señal dominante. El método incluye además la recepción de una entrada de balance de mezcla, la cual denota el balance ajustable entre las señales principal y asociada. El método incluye además la identificación de la señal dominante basada en la entrada de balance de mezcla y los metadatos de mezcla. El método puede incluir de forma adicional también la determinación de un factor de escala apropiado para la señal no dominante directamente desde la información de modificación de escala, sin la necesidad de ningún análisis ni medida de las señales de audio que se van a mezclar. Dado que estas técnicas no requieren ningún análisis ni medida de las señales de audio, el hardware requerido para implementar las técnicas puede ser mucho más sencillo comparado con los sistemas de hardware que deben probar la señal o llevar a cabo cálculos computacionalmente intensivos. Asimismo, la mezcla en tiempo real se puede lograr con relativa facilidad ya que todos los datos están disponibles sin tener que probar ni analizar grandes cantidades de datos sobre la marcha.

De acuerdo con el primer aspecto, se proporciona un método de mezcla de dos señales de audio de entrada en una única señal de audio mezclada con un nivel de sonido percibido consistente. Los primeros pasos de este método de mezcla conllevan la recepción de una señal de audio de entrada principal, la recepción de una señal de audio de entrada asociada, la recepción de metadatos de mezcla, los cuales contienen información de modificación de escala y la recepción de una entrada de balance de mezcla, la cual denota un balance ajustable entre las señales principal y asociada. A partir de estas entradas, la señal dominante se identifica basándose en la información de modificación de escala proporcionada por los metadatos de mezcla y la entrada de balance de mezcla. Entonces, la escala de la señal no dominante se modifica en relación con la señal dominante. Finalmente, la señal no dominante con su escala modificada se combina con la señal dominante en una señal mezclada. De este modo, a través del uso de la entrada de balance de mezcla, el método hace posible para el usuario influir en la mezcla permitiendo al usuario escoger la señal dominante así como el nivel de modificación de escala de la señal no dominante relativo a la señal dominante, mientras que el nivel de sonido percibido de la señal mezclada se mantiene constante mediante la modificación de la escala de la señal no dominante relativa a la señal dominante.

De acuerdo con otro aspecto, el nivel de sonido se puede representar por medio de un nivel ponderado de media de las señales. El nivel de sonido puede expresarse además como el nivel de diálogo de la señal. De este modo, el método puede relacionarse con la mezcla de dos señales de audio de entrada en una única señal de audio mezclada con un nivel de diálogo percibido consistente. Además, el nivel de diálogo se puede medir por medio de un valor dialnorm asociado a cada señal. Dialnorm se proporciona con frecuencia como un parámetro de metadatos para controlar la ganancia de descodificador. De acuerdo con el documento de estándares ETSI TS 102 366 v1.2.1, el parámetro de metadatos "dialnorm" es un código de 5 bits que va de 1 a 31 (el valor 0 está reservado). El código se debería interpretar como un nivel que va desde -1 dB a -31 dB con respecto a la escala completa. A partir del parámetro de metadatos dialnorm, se puede determinar un factor de modificación de escala. El factor de modificación de escala es equivalente a  $(31 - \text{dialnorm})$  dB. De este modo, un valor dialnorm de 31 indica un factor de escala de 0 dB, mientras que un valor dialnorm de 1 indica un factor de escala de -30 dB. En resumen, dialnorm es un valor entero que va de 31 a 1, donde un valor 31 indica que la ganancia de descodificador permanece unitaria y un valor de 1 indica que la ganancia de descodificador se va a reducir en 30 dB.

Siempre que el nivel de sonido de la señal dominante permanezca consistente, cualquier variación en los niveles de las señales no dominantes será menos evidente y el nivel de sonido percibido así como el nivel de diálogo percibido de la señal mezclada deberían permanecer consistentes. Preferiblemente, las señales de entrada deberían normalizarse apropiadamente. Además, se debería prestar atención al dialnorm de contenido de las señales de

entrada durante la mezcla para mantener un nivel de diálogo percibido consistente. Para obtener los mejores resultados, el dialnorm de contenido de las señales de entrada se debería establecer apropiadamente, usualmente en un valor dialnorm de 31, indicando la ganancia unitaria. No obstante, la necesidad de dialnorm no siempre se establece en 31, pero la normalización del diálogo se debería haber aplicado ya preferiblemente a ambas señales de entrada. Durante la mezcla, la escala de la señal dominante generalmente no se modifica, mientras que la escala de la señal de entrada no dominante se modifica por medio del factor de escala determinado. Por lo tanto, después de la mezcla del audio de diálogo de acuerdo con los métodos descritos, el nivel de diálogo de la señal dominante permanecerá consistente mientras que el nivel de diálogo de la señal no dominante relativo a la señal dominante será apropiado, dada la información de modificación escala a partir de los metadatos de mezcla y la entrada de balance de mezcla.

De acuerdo con otro aspecto, la entrada de balance de mezcla puede comprender además una entrada de usuario externa proporcionando un valor desde valores negativos muy amplios hasta valores positivos muy amplios, permitiendo de este modo la mezcla para favorecer tanto a la señal asociada como a la señal principal durante el proceso de mezcla en la cantidad deseada. En general, la entrada de balance de mezcla puede ser un número real positivo o negativo. Esta entrada de balance de mezcla permite al usuario ajustar la señal deseada en la que se centra la atención tal como se desee al mezclar dos señales de entrada. Además, la entrada de balance de mezcla permite al usuario especificar de forma precisa la atenuación de la señal no dominante relativa a la señal dominante en el proceso de mezcla, permitiendo típicamente al usuario escoger de entre la gama completa desde sin atenuación hasta atenuación completa. La aplicación de los métodos descritos en este documento permite al usuario cambiar sin problemas la mezcla desde atenuación completa de la señal principal hasta atenuación completa de la señal asociada.

De acuerdo con un aspecto, el método puede comprender además el paso de determinación de un factor de escala con el cual modificar la escala de la señal no dominante, donde el factor de escala se determina directamente a partir de la información de modificación de escala contenida en los metadatos de mezcla y la entrada de balance de mezcla. El método puede incluir entonces además la modificación de escala de la señal no dominante mediante el uso del factor de escala determinado. Además, los metadatos de mezcla pueden comprender un factor de escala de metadatos para la señal principal, denotando un factor de escala para la modificación de escala de la señal principal relativa a la señal asociada y los metadatos de mezcla pueden comprender opcionalmente un factor de escala de metadatos para la señal asociada, denotando un factor de escala para la modificación de escala de la señal asociada relativa a la señal principal. Los factores de escala pueden comprender además valores de dB. Al igual que el nivel de sonido de referencia de las señales se especifica de forma típica como un ajuste de dialnorm de 31, para obtener los mejores resultados, las señales de entrada deberían tener también ajustes de dialnorm de 31.

El método puede incluir también de forma adicional el paso de identificación de la señal dominante mediante la realización de la siguiente comparación, dado que no existe ningún factor de escala de metadatos para la señal asociada: en el caso en que el valor de la entrada de balance de mezcla sea mayor que el factor de escala de metadatos para la señal principal a partir de los metadatos de mezcla, entonces la señal asociada se determinará como la señal dominante; de lo contrario, la señal principal se determinará como la señal dominante. Si la señal asociada se identifica como la señal dominante, el método podrá comprender además la determinación de un factor de escala para la señal principal, que se calcula como la diferencia entre el factor de escala de metadatos para la señal principal a partir de los metadatos de mezcla y la entrada de balance de mezcla. Alternativamente, si la señal principal se identifica como la señal dominante, el método podrá comprender además la determinación de un factor de escala para la señal asociada, el cual se calcula como la diferencia entre la entrada de balance de mezcla y el factor de escala de metadatos a partir de los metadatos de mezcla para la señal principal.

En el caso en que exista también un factor de escala de metadatos para la señal asociada, el método podrá incluir también de forma adicional el paso de la identificación de la señal dominante mediante la realización de la siguiente comparación. El factor de escala de metadatos para la señal principal a partir de los metadatos de mezcla se puede ajustar por medio del factor de escala de metadatos para la señal asociada a partir de los metadatos de mezcla mediante la sustracción del factor de escala de metadatos para la señal asociada a partir del factor de escala de metadatos para la señal principal. En el caso en que el valor de la entrada de balance de mezcla sea mayor que el factor de escala ajustado, entonces la señal asociada se determinará como la señal dominante; de lo contrario, la señal principal se determinará como la señal dominante. Si la señal asociada se identifica como la señal dominante, el método podrá comprender además la determinación de un factor de escala para la señal principal, el cual se calcula como la diferencia entre el factor de escala ajustado y la entrada de balance de mezcla. Alternativamente, si la señal principal se identifica como la señal dominante, el método podrá comprender además la determinación de un factor de escala para la señal asociada, el cual se calcula como la diferencia entre la entrada de balance de

mezcla y el factor de escala ajustado.

Mediante la determinación del factor de escala para la señal no dominante y la mezcla de las señales de entrada basándose en qué señal se determina como la señal dominante, el nivel de sonido percibido de la señal mezclada se podrá mantener constante permitiendo aún para ambas el ajuste de la señal en la que se centra la atención así como la modificación de escala correspondiente de la señal no dominante. En otras palabras, al menos una de las señales es siempre la señal dominante, la cual es la señal en la que se centra la atención. Dado que la escala de la señal dominante generalmente no se modifica, el nivel de sonido de la señal mezclada se mantiene con respecto a otros programas, independientemente de cuánto se atenúe la señal no dominante en cuanto a la señal dominante basándose en la combinación de la entrada de balance de mezcla y los factores de escala de mezcla proporcionados por los metadatos.

En algunos casos, las señales de entrada pueden constar de múltiples canales de datos de audio. En una forma de realización alternativa, los factores de escala independientes se pueden determinar para cada canal de la señal de entrada principal. Durante la mezcla, se puede modificar entonces la escala de cada canal de la señal no dominante por medio de su factor de escala correspondientemente determinado y como en la forma de realización anterior, la escala de la señal dominante permanece sin modificar. De este modo, las señales pueden comprender también múltiples canales, donde los metadatos de mezcla pueden comprender entonces de manera adicional un factor de escala de metadatos primario para la señal principal con respecto a la señal asociada y un factor de escala de metadatos para cada canal de la señal principal con respecto a la señal asociada.

Los múltiples canales de las señales pueden comprender canales de sonido envolvente con canales izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente derecho y de efecto de baja frecuencia (LFE). Los múltiples canales pueden, por ejemplo, comprender señales de multicanal 5.1, señales de multicanal 3.1, señales de multicanal 13.1 u otras señales de multicanal.

Para las señales de multicanal que comprenden un canal de efecto de baja frecuencia (LFE), el método puede comprender además el cálculo del factor de escala para el canal de LFE como el mínimo del factor de escala de metadatos de LFE y el factor de escala de metadatos máximo de los demás canales de la señal principal. Esta función está destinada a evitar que el canal de LFE domine con demasiada fuerza la señal mezclada.

Particularmente útil para las señales de entrada que comprenden múltiples canales, el método puede incluir además la identificación de la señal dominante a partir de la entrada de balance de mezcla y los factores de escala de metadatos para la señal principal y el factor de escala de metadatos para la señal asociada de la manera siguiente. Este método conlleva la determinación del factor de escala de metadatos secundario máximo como el máximo de todos los factores de escala de metadatos para todos los canales de la señal principal excepto para el canal de efectos de baja frecuencia (LFE). De este modo, el canal de LFE, si existe, no se utiliza para determinar el factor de escala de metadatos secundario máximo. El método incluye además el cálculo de la suma de este factor de escala de metadatos secundario máximo y el factor de escala de metadatos primario para la señal principal. En ese caso, si el valor de entrada de balance de mezcla es mayor que esta suma, entonces la señal asociada se determinará como la señal dominante; de lo contrario, la señal principal se determinará como la señal dominante. Alternativamente, en caso de que exista también un factor de escala de metadatos para la señal asociada, la determinación anterior de la señal dominante se determinará mediante la comparación de si el valor de entrada de balance de mezcla es mayor que esta suma, menos el factor de escala de metadatos para la señal asociada.

Además, en el caso en que la señal asociada se determine como la señal dominante, el método podrá incluir además el cálculo de un factor de escala para cada canal de la señal principal, calculado como el factor de escala de metadatos primario para la señal principal más el factor de escala de metadatos para el canal principal correspondiente menos la entrada de balance de mezcla.

En el caso en que la señal principal se determine como la señal dominante y no exista ningún factor de escala de metadatos para la señal asociada, el método podrá incluir además la determinación de un factor de escala para la señal asociada, calculado como el valor de la entrada de balance de mezcla menos la suma del factor de escala de metadatos primario para la señal principal y el factor de escala de metadatos secundario máximo, el cual es el máximo de todos los factores de escala de metadatos para todos los canales de la señal principal, excepto para el canal de efectos de baja frecuencia (LFE), si existe.

En el caso en que la señal principal se determine como la señal dominante y exista un factor de escala de metadatos para la señal asociada, el método podrá incluir además la determinación de un factor de escala para la señal

asociada, calculado como el valor de la entrada de balance de mezcla más el factor de escala de metadatos para la señal asociada menos la suma del factor de escala de metadatos primario para señal principal y el factor de escala de metadatos secundario máximo, el cual es el máximo de todos los factores de escala de metadatos para todos los canales de la señal principal, excepto para el canal de efectos de baja frecuencia (LFE), si existe.

5

Este método puede comprender además la determinación de un factor de escala para cada canal de la señal principal, calculado como el factor de escala primario para el canal principal más el factor de escala para el canal principal menos la suma del factor de escala principal primario y el factor de escala de metadatos secundario máximo, el cual es el máximo de todos los factores de escala para los canales de la señal principal, excepto para el canal de efectos de baja frecuencia (LFE), si existe.

10

De acuerdo con un aspecto adicional, las señales pueden comprender señales Dolby Digital Plus (DD+) o Dolby Pulse. Asimismo, las señales se pueden codificar como E-AC-3, MPEG-4 HE-AAC, aacPlus, AC-3, MPEG-1 Capa 2, MPEG-4 AAC, cualquier derivación de las señales de audio MPEG-4 u otros formatos de codificación de audio similares. De este modo, los metadatos pueden comprender metadatos de mezcla de nivel ES (transmisiones elementales) para mezclar señales DD+, por ejemplo metadatos de mezcla de acuerdo con ETSI TS 102 366 V1.2.1 Ch. E.1.2.2. Estos metadatos de nivel ES pueden comprender además factores de escala para los canales de una señal de multicanal: extpgmscl, extpgmlscl, extpgmrscl, extpgmcscl, extpgmlscl, extpgmrsscl, extpgmlfescl, correspondientes al factor de escala primario y a los factores de escala para el canal izquierdo, el canal derecho, el canal central, el canal de sonido envolvente izquierdo, el canal de sonido envolvente derecho y el canal de efectos de baja frecuencia (LFE). Los metadatos de nivel ES pueden comprender además opcionalmente un valor panmean. Los metadatos de nivel ES pueden comprender además opcionalmente un valor pgmscl correspondiente al factor de escala para la señal asociada. En muchos casos relacionados con las señales DD+, los metadatos de mezcla se pueden codificar con o incrustar en la señal asociada, aunque los diversos estándares de metadatos y las implementaciones varían. Por ejemplo, con MPEG-4 HE-AAC, los metadatos de mezcla se desplazan junto a, pero no estrictamente dentro, los datos de audio. Por consiguiente, la aplicación de las técnicas anteriores no debería limitarse únicamente a casos en los que los metadatos de mezcla estén contenidos en la señal asociada.

15

20

25

De acuerdo con otro aspecto, los metadatos pueden comprender metadatos de mezcla de nivel PES (transmisiones elementales por paquetes) para señales por impulso/HE-AAC de mezcla, así como E-AC-3, MPEG-4 HE-AAC, aacPlus, AC-3, MPEG-1 Capa 2, MPEG-4 AAC o cualquier derivación de las señales de audio MPEG-4, por ejemplo los metadatos de mezcla de acuerdo con ETSI TS 101 154 V1.9.1 Anexo E2. Estos metadatos de nivel PES pueden comprender además factores de escala para los canales de una señal multicanal: AD\_fade\_byte, AD\_gain\_byte\_center, AD\_gain\_byte\_front, AD\_gain\_byte\_surround o AD\_pan\_byte, correspondientes al factor de escala primario y los factores de escala para el central, frontal, de sonido envolvente y de inclinación. En muchos casos, que implican señales por impulso, los metadatos de mezcla pueden estar asociados directamente con la señal asociada. Por ejemplo, los metadatos de mezcla y la señal asociada se pueden transmitir juntos en la misma o en transmisiones relacionadas. No obstante, la aplicación de las técnicas descritas en este documento no debería limitarse en forma alguna a tales casos y las técnicas descritas únicamente requieren la recepción de los metadatos de mezcla que proporcionan información de modificación de escala para la señal principal relativa a la señal asociada.

30

35

40

De acuerdo con otro aspecto, el método puede incluir además la restricción de que la señal asociada no contenga canales que no estén presentes en la señal principal, a menos que la señal asociada sea mono. Una señal mono puede estar indicada, por ejemplo, por un "modo de codificación de audio" o valor acmod de 1, el cual indica que la señal únicamente tiene un canal central. Sin embargo, la señal asociada puede comprender además canales de LFE y mono, donde el canal de LFE de la señal asociada se mezcla con el canal de LFE de la señal principal y el canal mono de la señal asociada se coloca en los canales principales de la señal principal, donde la colocación conlleva la división o distribución de la señal en múltiples canales de acuerdo con un valor que especifica cómo se debería distribuir la señal, por ejemplo, un valor "panmean".

45

50

De acuerdo con otro aspecto, cuando el canal principal no es mono y, de este modo, tiene canales izquierdo y derecho pero la señal asociada es mono (es decir, no tiene componentes estéreo, de sonido envolvente o izquierdo-derecho), los metadatos para la señal asociada pueden comprender además información de metadatos de colocación. Por ejemplo, tal información de metadatos de colocación puede estar proporcionada por la clave de metadatos "panmean" en los metadatos de mezcla. El método incluye entonces el uso del valor de colocación (pan) para colocar la señal asociada mono en los correspondientes canales izquierdo y derecho de la señal mezclada, donde el valor panmean especifica una dirección desde la cual se pueden computar los factores de escala para cada canal. La sección "E.4.3.5 Colocación" del estándar ETSI TS 102 366 v1.2.1 proporciona detalles adicionales para la

55

aplicación del valor panmean. De acuerdo con esta especificación, panmean determina el ángulo efectivo en el cual la señal asociada mono tiene que aparecer dentro del "espacio" de la señal principal descodificada, siendo central con 0 grados, indicando +/- 90 la colocación total izquierda o derecha. La colocación (pan) se puede aplicar a diversas señales de multicanal tales como 3.1, 5.1 y 13.1. El rango exacto para panmean es de 0 a 239, lo cual  
 5 representa de 0 a 358,5 grados en los pasos de 1,5 grados y donde 0 grados es la dirección del altavoz central. Por ejemplo, un valor panmean de 121 se aplicaría con los siguientes factores de escala en los cinco canales que no son de LFE de una señal de multicanal 5.1: 0,078 izquierdo, 0,0997 central, 0 derecho, 0,734 de sonido envolvente izquierdo y 0,679 de sonido envolvente derecho.

- 10 Aún en otra forma de realización alternativa, un factor de escala adicional se puede aplicar a la señal dominante. Este factor de escala se puede determinar a partir de los metadatos, desde un control de entrada adicional o desde un valor fijo en el dispositivo de mezcla. No obstante, este factor de escala se podría determinar también a partir de un análisis en tiempo real de las señales de entrada para mejorar aún más el mantenimiento de un nivel de sonido percibido constante. Otra alternativa es que un factor de escala adicional se podría derivar desde los metadatos de  
 15 flujo de bits, p. ej. mediante el uso de dialnorm. En algunos casos, el factor de escala es unitario, convirtiendo la salida de señal mezclada de esta forma de realización en idéntica a la salida de las formas de realización previas. No obstante, en otros casos, el factor de escala no es unitario y la salida de señal mezclada de esta forma de realización diferirá de la salida de las formas de realización previas. No obstante, independientemente de la forma de realización que se utilice, el nivel de diálogo de la transmisión dominante permanecerá consistente, garantizando  
 20 que el nivel de diálogo percibido de la señal mezclada permanezca consistente. De este modo, de acuerdo con otro aspecto, la escala de la señal dominante se puede modificar también mediante la aplicación de un factor de escala a la señal dominante. Esta modificación de escala adicional de la señal dominante puede servir, por ejemplo, para permitir la integración de la normalización de las señales de audio de entrada dentro del dispositivo de mezcla en el caso en que las señales de entrada no se hayan normalizado antes de la mezcla.

25 De acuerdo con otro aspecto, el método puede incluir además el paso de verificar que el dialnorm de contenido para las señales de entrada se establezca apropiadamente y posiblemente proporcionando una indicación cuando esta condición no se cumpla.

- 30 De acuerdo con otro aspecto, se proporciona también un dispositivo para las señales de mezcla que aplica el método de mezcla de señales descrito anteriormente. De manera similar, se proporciona también un descodificador que aplica el método de mezcla de señales descrito anteriormente.

De acuerdo con otro aspecto, se proporciona también un medio de almacenamiento legible por un procesador que  
 35 almacena instrucciones legibles por ordenador para ejecutar los pasos de cualquiera de los métodos descritos anteriormente. De manera similar, se proporciona también un producto de programa legible por un procesador que comprende instrucciones ejecutables para llevar a cabo cualquiera de los métodos descritos anteriormente cuando se ejecuta en un procesador.

- 40 De acuerdo con otro aspecto, se proporciona también un dispositivo para mezclar señales de audio de entrada en una única señal de audio mezclada con un nivel de sonido percibido consistente. El dispositivo comprende un receptor para recibir una señal de audio de entrada principal, una señal de audio de entrada asociada y metadatos de mezcla con información de modificación de escala. El dispositivo comprende además una entrada de balance de mezcla, la cual denota el balance ajustable entre las señales principal y asociada. El dispositivo comprende además  
 45 un mezclador configurado para identificar la señal dominante, modificar la escala de las señales de entrada y combinar las señales en una señal mezclada. La señal dominante se identifica mediante el análisis de la información de modificación de escala proporcionada por los metadatos de mezcla y la entrada de balance de mezcla y la escala de la señal no dominante se modifica en relación con la señal dominante. Además, se proporciona también un dispositivo para mezclar señales de audio de entrada en una única señal de audio mezclada con un nivel de sonido  
 50 percibido consistente, donde el factor de escala de la señal no dominante se determina directamente desde la información de modificación de escala a partir de los metadatos de mezcla y la entrada de balance de mezcla. La escala de la señal no dominante se modifica mediante el uso del factor de escala determinado y la señal no dominante con su escala modificada se combina con la señal dominante en una señal mezclada.

De acuerdo con otro aspecto, se proporciona un dispositivo para mezclar señales de audio de entrada en una única  
 55 señal de audio mezclada con un nivel de sonido percibido consistente, el cual está configurado además para recibir metadatos de mezcla que están integrados con la señal de audio de entrada asociada. La entrada de balance de mezcla del dispositivo puede comprender además una entrada de usuario externa que proporciona un valor que va desde valores negativos muy amplios hasta valores positivos muy amplios. El receptor del dispositivo se puede

configurar además para recibir señales de multicanal, donde la unidad de mezcla se configura para mezclar las señales de multicanal.

Las técnicas descritas anteriormente están pensadas para mantener un nivel de sonido percibido consistente para la señal mezclada. Como tales técnicas se pueden describir también como la nivelación de señal, las técnicas descritas anteriormente que modifican la escala de las señales se pueden ver también en términos de nivelación y ajuste del nivel de señal. Por consiguiente, para generalizar la materia objeto de los métodos para mezclar señales basados en el resultado que se logra, la escala de las señales de entrada se debería modificar de tal forma que la diferencia del nivel de modificación de escala resultante entre la modificación de escala de la señal principal y la modificación de escala de la señal asociada sea esencialmente la suma del factor de escala de metadatos de mezcla para la señal principal y el valor de entrada de balance de mezcla. Asimismo, si existe un factor de escala de metadatos de mezcla para la señal asociada, entonces la diferencia del nivel de modificación de escala resultante entre la modificación de escala de la señal principal y la modificación de escala de la señal asociada es esencialmente la suma de la diferencia de los factores de escala de metadatos de mezcla para la señal principal y la señal asociada y el valor de entrada de balance de mezcla.

Con los métodos descritos anteriormente, el autor de contenido puede influir en la diferencia del nivel entre las señales principal y asociada en todo momento. Por consiguiente, el usuario final puede influir también en la diferencia de nivel entre las señales principal y asociada en todo momento. Además, cuando la entrada de balance de mezcla está en el ajuste neutral de cero, las señales se mezclan exactamente tal como lo especifica el autor de contenido. Las técnicas descritas en este documento proporcionan un control sin problemas a través de la gama de la entrada de balance de mezcla, desde la señal principal como la señal única (donde la señal asociada se atenúa por completo) hasta el otro extremo donde la señal asociada es la señal únicamente presente en la reproducción (donde la señal principal se atenúa por completo). Sin embargo, la señal dominante se nivela, lo que significa que independientemente de los valores de los metadatos de mezcla o la entrada de balance de mezcla, el oyente no debería percibir un cambio significativo en el nivel de sonido al cambiar a o desde otro programa. De este modo, para el audio de diálogo, el diálogo en el que se centra la atención puede ser seleccionado por el usuario y el diálogo dominante está siempre nivelado, ya que el diálogo dominante está asociado a la señal dominante.

Los métodos y dispositivos que incluyen ejemplos y formas de realización tal como se resume la presente solicitud de patente se pueden utilizar de forma independiente o en combinación con los demás métodos y dispositivos revelados en este documento. Además, todos los aspectos de los métodos y dispositivos resumidos en la presente solicitud de patente se pueden combinar arbitrariamente. En particular, las características de las reivindicaciones se pueden combinar con otra de una manera arbitraria como es evidente para una persona experta en la materia.

### **Breve descripción de las figuras**

La invención se explica a continuación a través del uso de ejemplos con referencia a los dibujos adjuntos, donde los ejemplos se ilustran por medio de las siguientes figuras:

Fig. 1A: muestra los niveles de sonido de los diferentes programas de audio.

Fig. 1B: muestra los niveles de sonido de los diferentes programas de audio después de la nivelación.

Fig. 2: muestra un diagrama de bloques de un convertidor y decodificador con un mezclador.

Fig. 3: muestra un diagrama de proceso para la identificación de la señal dominante y las señales de entrada de mezcla.

Fig. 4: muestra un diagrama de flujos para la identificación de la señal dominante y la mezcla.

Fig. 5A: muestra un gráfico de los factores de escala para las señales principal y asociada, para el caso en que el factor de escala de metadatos de mezcla es -10 dB.

Fig. 5B: muestra un ejemplo para la lectura de los factores de escala de la fig. 5A de las señales principal y asociada donde el factor de escala de metadatos de mezcla es -10 dB para el caso donde la entrada de balance de mezcla es +5 dB.

Fig. 5C: muestra un ejemplo para la lectura de los factores de escala de la fig. 5A de las señales principal y asociada donde el factor de escala de metadatos de mezcla es = -10 dB para el caso donde la entrada de balance de mezcla es -15 dB.

5 Fig. 6: muestra un diagrama de bloques de señal para la mezcla de señales de multicanal 5.1.

Fig. 7A: muestra un diagrama de flujos de un ejemplo para la mezcla de señales de multicanal 5.1.

Fig. 7B: muestra un diagrama de flujos para la determinación del factor de escala para el canal de LFE para el paso 10 703 de la fig. 7.

Fig. 8: muestra un diagrama de bloques de señal para la mezcla de una señal principal de multicanal 5.1 con una señal asociada 1.1 (mono + LFE).

## 15 **Descripción detallada**

Es un concepto común para incrustar metadatos de audio en una transmisión de audio digital, p. ej. en entornos de difusión digital. Tales metadatos son “datos sobre datos”, es decir, datos sobre el audio digital en la transmisión. Los metadatos pueden proporcionar información a un descodificador de audio sobre cómo reproducir el audio. Tales metadatos se transmiten típicamente junto con un flujo de bits de audio digital. Estos metadatos proporcionan una capacidad sin precedentes para los productores de contenido para suministrar la mayor calidad de audio a los consumidores en una gama de entornos de audición. También proporciona opciones que permiten a los consumidores ajustar su configuración de ajustes para la mejor adaptación de sus entornos de audición.

25 Las técnicas descritas en este documento se pueden aplicar a una amplia gama de formatos de señal de audio y esquemas de codificación. Con el fin de aplicar los métodos, las señales únicamente necesitan tener metadatos adjuntos a la señal de audio que proporcionan la información de mezcla. Las señales de entrada pueden ser un canal único pero con frecuencia tendrán múltiples canales, por ejemplo las ampliamente conocidas señales de multicanal 5.1 con 6 canales: izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente 30 derecho y de LFE (efectos de baja frecuencia). Dos ejemplos de un formato de señal de audio tal que permiten tales metadatos de mezcla son Dolby Digital Plus (DD+ o E-AC-3 (AC-3 mejorado)) y el codificador avanzado de audio de alta eficiencia (HE-AAC).

Dolby Digital Plus (DD+ o E-AC-3 (AC-3 mejorado)) es un esquema de compresión de audio digital, el cual es un sistema de codificación mejorado basado en el códec AC-3. E-AC-3 soporta hasta 13 canales de audio de gama completa en una tasa de bits codificada de 6.144 Mbit/s máximo. Por ejemplo, el documento ETSI TS 102 366 v1.2.1 describe los metadatos DD+ de forma detallada, por ejemplo, “E1.3.1.17 extpgmscl: factor de escala de programa externo - 6 bits: en algunas aplicaciones, dos flujos de bits se pueden descodificar y mezclar juntos. Este campo especifica un factor de escala que se debería aplicar al programa externo (es decir, un programa que se transmite 40 en un flujo de bits independiente o un flujo adicional independiente) durante la mezcla. Este campo utiliza la misma escala que pgmscl”. Los metadatos de mezcla más relevantes comprenden las siguientes claves de metadatos: extpgmscl, extpgmlscl, extpgmrscl, extpgmcscl, extpgmlscl, extpgmrsscl, extpgmlfesccl, pgmscl. En algunos casos, por ejemplo cuando la señal de audio asociada es mono, se puede utilizar también la clave de metadatos panmean.

45 El codificador avanzado de audio de alta eficiencia (HE-AAC) es un esquema de compresión de datos con pérdidas para el audio digital definido como un perfil de audio MPEG-4 en ISO/IEC 14496-3. Es una extensión de complejidad baja AAC (AAC LC) optimizada para aplicaciones de tasa de bits baja tales como el audio de transmisión. El perfil de HE-AAC versión 1 (HE-AAC v1) utiliza replicación de la banda espectral (SBR) para mejorar la eficiencia de compresión en el dominio de frecuencia. El perfil de HE-AAC versión 2 (HE-AAC v2) acopla SBR con estéreo 50 paramétrico (PS) para mejorar la eficiencia de compresión de las señales estéreo. Es una versión estandarizada y mejorada del códec AACplus. Los metadatos de mezcla más relevantes comprenden las siguientes claves de metadatos: AD\_fade\_byte, AD\_gain\_byte\_center, AD\_gain\_byte\_front, AD\_gain\_byte\_surround y AD\_pan\_byte. HE-AAC se utiliza en los estándares de radio digital como DAB+ y Digital Radio Mondiale (radio digital mundial).

55 Con el fin de aprovechar por completo sus capacidades, los autores de contenido o productores, deberían tener un conocimiento sólido de los metadatos de audio, los cuales se añaden típicamente durante la producción o la postproducción.

Tal como se describe en la sección Antecedentes, con el fin de mantener niveles de sonido consistentes de los diferentes programas, las señales de audio más modernas incluyen metadatos, tales como un nivel de diálogo, el cual se utiliza para modificar la escala de la señal de audio. Tal modificación de escala permite que cada señal se represente de forma apropiada en una forma permitiendo una elevada calidad y buena comprensión manteniendo al mismo tiempo también el mismo nivel de sonido. La figura 1a muestra un ejemplo de los diferentes programas sin tal nivelación o normalización de nivel de diálogo. Tal como se aprecia en el diagrama, el nivel de sonido de los programas difiere, de tal forma que el oyente puede percibir un nivel de sonido diferente al cambiar de un programa a otro. La figura 1a ilustra los niveles de sonido de cuatro programas diferentes S1, S2, S3 y S4, por ejemplo en casos donde el usuario cambie el canal o se transmita un programa nuevo en tiempos t0, t1, t2 y t3. La progresión del tiempo se muestra en el eje horizontal y el nivel de sonido en el eje vertical. La línea horizontal A1 muestra el nivel de sonido del programa S1, el cual se ejecuta del tiempo t0 a t1. Las líneas horizontales A2, A3 y A4 muestran asimismo los niveles de sonido del programa S2, S3 y S4, respectivamente, el cual se ejecuta de los tiempos t1 a t2, t2 a t3 y t3 a t4. Tal como se puede ver, el nivel de sonido A1 es superior a los niveles de sonido A2, A3 y A4.

15 Para corregir tal diferencia en el nivel de sonido, las señales del programa se nivelan o se modifica su escala a un nivel de referencia común. De este modo, la figura 1b muestra las señales del programa después de ser niveladas a un nivel de sonido de referencia común. El nivel de sonido de referencia para las señales se representa típicamente como un nivel ponderado de media, el cual se especifica con frecuencia en los metadatos como un ajuste dialnorm de 31. Tal como se puede ver en la figura 1b, los niveles de sonido A1'; A2', A3' y A4' son todos los mismos. De este modo, el usuario no percibe ningún cambio significativo en el nivel de sonido al cambiar entre programas.

Muchos formatos de señal de audio nuevos permiten ahora también que una o varias señales asociadas se agrupen con la señal principal. Estas señales de audio se pueden activar basándose en cada ajuste de dispositivo o en las selecciones del usuario, por ejemplo durante la reproducción. Típicamente, tal activación conlleva la combinación de la señal asociada con su correspondiente señal principal. Esta combinación de las señales en una señal se conoce como mezcla. Por ejemplo, un productor puede proporcionar una pista de comentarios del director en una pista de audio, por ejemplo en un vídeo DVD. El oyente puede utilizar entonces el convertidor y descodificador o el reproductor de DVD para activar esta pista asociada, la cual se superpone y se mezcla entonces con la señal principal. Con el fin de mantener el mismo nivel de sonido percibido que otros programas, una señal asociada tal contiene usualmente metadatos que especifican cómo se deberían mezclar las señales. Estos metadatos de mezcla permiten al productor controlar de forma precisa la mezcla de las señales proporcionando detalles específicos sobre cómo mezclar las señales incluyendo con frecuencia también la opción de especificar detalles para la mezcla de señales de multicanal.

Aunque los metadatos de mezcla describen los niveles de mezcla concebidos por el productor, el oyente puede tener en algún momento un deseo diferente relacionado con cómo se deberían mezclar las señales. De este modo, el oyente puede decir si desea que la otra señal sea el "diálogo en el que se centra la atención". De este modo, el "diálogo en el que se centra la atención" es la señal que debería dominar en la señal mezclada. No obstante, este "diálogo en el que se centra la atención" es en cierto modo subjetivo. El autor de contenido tiene una idea de qué señal debería ser el centro de atención, pero es conveniente que el usuario también tenga algún control sobre ella del mismo modo. De acuerdo con las enseñanzas de esta solicitud, la indicación de "enfoque del usuario" se puede lograr ofreciendo una entrada de control de balance de mezcla para ajustar la mezcla de las señales principal y asociada.

De este modo, para el caso en el que al usuario le gustaría desviarse de la ponderación de la señal especificada por los metadatos mediante el ajuste de la señal mezclada para favorecer una señal u otra, una forma de realización incluye el suministro de un control de balance de mezcla para permitir la entrada de la entrada de balance de mezcla. Este control de balance de mezcla permite al usuario indicar que una de las señales debería estar más presente en la señal mezclada y hasta qué punto la señal debería ser más predominante. Adicionalmente, esta entrada de balance de mezcla proporciona al usuario el control inmediato sobre la señal en la que se centra la atención. En la posición estándar, el control de balance de mezcla es neutral, el cual indica que la mezcla se debería producir tal como se especifica en los metadatos. Cuando el control de balance de mezcla se gira en una dirección (en esta forma de realización, en la dirección negativa), la señal principal se vuelve más alta con respecto a la señal asociada. Cuando el control de balance se gira en la otra dirección (en esta forma de realización, en la dirección positiva), la señal asociada se vuelve más alta con respecto a la señal principal.

No obstante, cuando un control de balance de mezcla tal o la entrada de balance de mezcla se utilizan entonces para incrementar o reducir los niveles de las señales principal y asociada, es probable que el nivel de sonido o el nivel de diálogo percibido de la señal mezclada se alteren de este modo. Así, es probable que al cambiar entre

programas con y sin una señal asociada activada se produciría de nuevo una situación no deseada como la que se describe en la figura 1a. De este modo, existe una necesidad de proporcionar una técnica de mezcla de las señales principales y asociadas bajo la influencia de una entrada de balance de mezcla, de tal forma que el nivel de sonido percibido de la señal mezclada se mantenga también constante.

5

Las técnicas descritas en este documento se pueden implementar en muchas formas de realización, tales como un convertidor y descodificador, un reproductor de audio portátil o no portátil, un dispositivo de mezcla, un reproductor de DVD, una televisión, un teléfono inteligente o un dispositivo receptor de TV para un sistema informático. Además, los aspectos individuales se pueden implementar en una combinación de sistemas de hardware y software. Los componentes de una forma de realización concreta se describirán con más detalle basándose en un convertidor descodificador incluyendo un mezclador de acuerdo con las técnicas descritas en este documento. Este convertidor y descodificador de la forma de realización se ilustra en la figura 2. El convertidor y descodificador 200 comprende un receptor 209, un mezclador 204 y un procesador 205. Aunque este ejemplo se muestra con estos componentes, la aplicación de las técnicas descritas en este documento no debería limitarse a esta configuración exacta. Por ejemplo, el receptor 209 se podría integrar en el mezclador 204; además, el convertidor y descodificador 200 no necesita tener un procesador, ya que la funcionalidad podría ser realizada por los circuitos electrónicos tradicionales. En este ejemplo, el receptor 209 del convertidor y descodificador 200 se configura para recibir la señal de entrada principal 201 y la señal de audio de entrada asociada 202, incluyendo los metadatos correspondientes. En algunos casos, el receptor puede necesitar adaptarse además para recibir los metadatos de mezcla que se pueden transmitir o codificar de forma independiente. Si es necesario, el receptor 209 se puede configurar además para descodificar las señales de entrada 201, 202 y los correspondientes metadatos de mezcla 202a de acuerdo con la codificación de las señales. Además, el convertidor y descodificador 200 y el mezclador 204 se configuran también para recibir la entrada de control de balance de mezcla 203. Opcionalmente, el receptor 209 puede configurarse también para realizar una descodificación especial de la entrada de control de balance de mezcla 203. Existen diversas opciones para las características adicionales y la implementación exacta de la entrada de control de balance de mezcla 203. Por ejemplo, el convertidor y descodificador 200 podría vincular esta entrada con un dispositivo de control remoto para permitir al usuario establecer este nivel de balance de mezcla a través de un menú al que se accede a través de un dispositivo de control remoto. Por supuesto, el convertidor y descodificador 200 o su control remoto podrían realizar la entrada de control de balance de mezcla con un elemento dedicado, tal como un dial de entrada, teclas de más y menos o un dispositivo de entrada similar. El convertidor y descodificador 200 y el mezclador 204 podrían por supuesto configurarse también para recibir la entrada de balance de mezcla 203 como una señal de otro sistema o dispositivo. En cualquier caso, la señal dominante 206 es la señal de entrada que se identifica por medio del procesador 204 del mezclador 205 basándose en la señal de entrada principal 201, la señal de audio de entrada asociada 202, los metadatos de mezcla 202a y la entrada de control de balance de mezcla 203. De este modo, la señal no dominante 207 es la otra señal de entrada que no se determina como la señal dominante 206. Entonces, la señal no dominante con su escala modificada 207s es la señal no dominante 207 después de haber modificado su escala. Finalmente, la señal dominante 206 y la señal no dominante con su escala modificada 207s se mezclan como la señal mezclada 208.

Una solución de ejemplo para lograr este objetivo de la mezcla de las señales de audio basándose en la configuración de la figura 2 se ilustra en la figura 3. De este modo, las entradas constan de una señal de entrada principal 201, una señal de entrada asociada 202 y una entrada de control de balance de mezcla 203. Primero, en el paso 301, la señal de entrada principal 201 es recibida por el mezclador 204, el convertidor y descodificador 200 y/o el receptor 209. En el paso 302, se recibe la señal de entrada asociada 202, incluyendo los metadatos de mezcla 202a. Hay que tener en cuenta que la inclusión de los metadatos de mezcla 202a en la señal asociada 202 es únicamente para este ejemplo y no se debería interpretar como limitación de la aplicación de las técnicas ya que otros estándares y formatos de codificación de señal pueden diferir. En el paso 303, se recibe la entrada de balance de mezcla 203. Estos pasos incluyen cualquier descodificación especial que pueda requerir el formato de las señales. A continuación, en el paso 304, la información de modificación de escala a partir de los metadatos de mezcla 202a y la entrada de control de balance de mezcla 203 se utilizan para la identificación de la señal dominante 206 desde las señales de entrada 201, 202 que se van a mezclar. Este paso puede ser realizado por un dispositivo electrónico común o circuito ya que no requiere pruebas, análisis minucioso, medida ni computación de alta potencia de las señales de entrada. Entonces, en el paso opcional 305, un factor de escala de la señal no dominante 207 se podrá determinar directamente a partir de los metadatos de mezcla 202a y la entrada de balance de mezcla 203. En el paso 306, se modifica la escala de la señal no dominante 207. En el caso en que un factor de escala de la señal no dominante 207 se determinara en el paso 305 basándose en los metadatos de mezcla 202a y la entrada de balance de mezcla 203, el paso 306 incluirá la modificación de escala de la señal no dominante 207 de acuerdo con el factor de escala determinado en el paso 305. Finalmente, en el paso 307, la señal no dominante con su escala modificada 207s se combina con la señal dominante 206 como la señal mezclada 208 del último paso 308.

Una forma de realización adicional incluye el paso opcional 305 para la determinación del factor de escala para la señal no dominante 207 basándose directamente en los metadatos de mezcla 202a y en la entrada de balance de mezcla 203. La figura 4 muestra un diagrama de flujos para esta determinación para esta forma de realización  
5 adicional. La identificación de la señal dominante 206 se basa en la entrada de balance de mezcla 203 y los metadatos de mezcla 202a. En algunos casos, los metadatos de mezcla pueden contener también un factor de modificación de escala para la señal asociada. Si este valor de metadatos está presente, entonces se puede utilizar para ajustar el factor de metadatos para mezclar las señales. De este modo, en el paso opcional 401, el factor de escala para la señal principal se puede ajustar mediante la sustracción del factor de escala para la señal asociada a  
10 partir del factor de escala para la señal principal. En el paso 402, el valor de la entrada de balance de mezcla 203 se compara al factor de escala de metadatos ajustado. En el caso en que el factor de escala no se haya ajustado en el paso 401, el valor de la entrada de balance de mezcla 203 se compara al factor de escala para la señal principal a partir de los metadatos de mezcla 202a. Para el caso afirmativo en el que la entrada de balance de mezcla 203 sea superior al factor de escala de metadatos de mezcla 202a, entonces en el paso 403y la señal asociada 202 se  
15 determinará como la señal dominante 206 y la señal principal 201 será la señal no dominante 207. De otra manera, para el caso negativo en el que el valor de la entrada de balance de mezcla 203 sea menor o igual al factor de escala de metadatos de mezcla 202a, entonces en el paso 403n la señal principal 201 se determinará como la señal dominante 207 y la señal asociada 202 será la señal no dominante 207. Ambos pasos 403y y 403n van seguidos del paso 404, donde se modifica entonces la escala de la señal no dominante 207. En el paso 405, la señal dominante se combina con la señal no dominante con su escala modificada 207s, para formar la señal mezclada 208 del paso  
20 406.

La figura 5 muestra un gráfico que ilustra un ejemplo concreto de los factores de escala que se van a aplicar a las señales principal y asociada 201, 202 al mezclar la señal asociada 202 con la señal principal 201. En este ejemplo,  
25 el factor de escala de los metadatos de mezcla 202a de la señal asociada 202 especifica que la escala de la señal principal 201 se debería modificar a -10 dB. Este factor de escala de -10 dB para la modificación de escala de la señal principal 201 cuando se mezcla con la señal asociada 202 indica la relación de mezcla deseada establecida por el productor de contenido durante el control. El eje horizontal de la figura 5 muestra la entrada de balance de mezcla 203 tal como varía de infinidad negativa a infinidad positiva. El eje vertical indica el factor de escala en dB  
30 desde  $-\infty$  dB a 0dB desde la parte inferior a la superior. Los dos gráficos de la figura 5 muestran los factores de escala de las dos señales de entrada dibujadas para todos los valores de la entrada de balance de mezcla desde  $-\infty$  dB a  $+\infty$  dB (los valores de la entrada de balance de mezcla están en el eje horizontal). Los dos gráficos de líneas, uno con una línea discontinua y el otro con una línea de puntos, representan los factores de escala para la señal principal 201 y la señal asociada 202, respectivamente. De este modo, el gráfico con la línea discontinua 501a,  
35 501b muestra la señal principal y el gráfico con la línea de puntos 502a, 502b muestra la señal asociada. Tal como se puede ver, cuando la entrada de balance de mezcla 203 está en su ajuste neutral, es decir, cero dB, la entrada de balance 203 no tiene efecto alguno y el factor de escala es como se especifica en los metadatos 202a: en este ejemplo donde el factor de escala de metadatos especificado es -10 dB, indicando la cantidad de atenuación de la señal principal con respecto a las señales asociadas, la señal asociada 202 es por tanto la señal dominante 206 y su  
40 escala no se modifica y la escala de la señal principal 201 se modifica de acuerdo con el factor de escala de metadatos de -10 dB. No obstante, como se puede ver en la figura, si la entrada de balance 203 se realiza más negativa que el factor de escala de metadatos de mezcla, es decir, menos de -10 dB, entonces la señal principal 201 se convertirá en la señal dominante 206. De forma adicional, la señal de entrada que es la señal dominante se indica además en la figura 5 por las etiquetas 501a, 501b y 502a, 502b, donde la línea horizontal 501a representa la señal  
45 de entrada principal 201 cuando es la señal dominante 206, la línea diagonal 501b representa la señal de entrada principal 201 cuando es la señal no dominante 207, la línea horizontal 502b representa la señal de entrada asociada 202 cuando es la señal dominante 206 y la línea diagonal 502a representa la señal de entrada asociada 202 cuando es la señal no dominante 207.

50 Para un ejemplo específico para la lectura de los factores de escala del caso de la figura 5, la figura 5a muestra los factores de escala para el caso en que la entrada de balance de mezcla 203 tiene un valor de +5 dB y, como en la figura 5, el factor de escala de metadatos de mezcla para la señal principal es -10 dB. En la figura, el factor de escala 501a para la señal principal 201 se puede ver que es -15 dB, lo cual es  $-10 \text{ dB} - 5 \text{ dB} = -15 \text{ dB}$ . El factor de escala 502a para la señal asociada 202 es de 0 dB. Dado que la señal asociada 202 es la señal dominante 206, se espera  
55 que el factor de escala para la señal asociada 202 debiera ser unitario. De este modo, si la entrada de balance 203 se incrementa para favorecer a la señal asociada 202, entonces la señal asociada 202 seguirá siendo la señal dominante 206, la escala de la señal asociada 202 no se modificará y únicamente se modificará la escala de la señal principal 201.

Por otro lado, la figura 5b muestra los factores de escala para el caso de la figura 5 donde la entrada de balance de mezcla 203 tiene un valor de -15 dB y, como en la figura 5, el factor de escala de metadatos de mezcla para la señal principal es de -10 dB. En la figura, se puede ver que el factor de escala para la señal principal 501b es de 0 dB por lo que la escala de la señal principal 201 no se modifica, ya que es la señal dominante 206 y que la escala de la señal asociada 202 se modifica por el factor de escala asociado 502b de -5 dB, el cual es  $-15 \text{ dB} - (-10 \text{ dB}) = -5 \text{ dB}$ .

Además, en las figuras 5A, 5B y 5C, se puede ver que cuando la entrada de balance de mezcla 203 se establece en infinito positivo, la señal principal 201 está completamente atenuada y la escala de la señal asociada 202 no se modifica. Al contrario, cuando la entrada de balance de mezcla 203 se establece en infinito negativo, la escala de la señal principal 201 no se modifica y la señal asociada 202 está completamente atenuada.

Los métodos descritos en este documento se pueden aplicar también para las señales de multicanal de mezcla. Dado que los metadatos de mezcla para señales de multicanal pueden tener muchas formas diferentes, como puede ser de 2 a 14 canales diferentes en señales de multicanal comunes (p. ej. 2 canales para estéreo, 2 canales para mono+LFE, 4 canales para 3.1, 6 canales para 5.1 y 14 canales para 13.1), el método se describirá aquí por medio del uso de claves de metadatos genéricos. En la práctica, la persona experta en la materia tendrá que sustituir los metadatos genéricos enumerados a continuación por aquellos apropiados para la codificación particular utilizada por las señales. De este modo, los términos genéricos "sclmain", "scl[ch]" y "pan" se utilizan a continuación, donde "sclmain" es el factor de escala primario, "scl[ch]" es el factor de escala para cada canal individual donde [ch] es un índice para el canal y "pan" es un valor opcional que indica cómo se deberían aplicar los canales a una mezcla con un número de canales diferente (p. ej. cómo se debería aplicar una señal mono a una señal 5.1, por ejemplo como 50% de la señal mono en el canal izquierdo, 50% en el canal derecho o 0% en los demás canales o 60% en el canal izquierdo, 40% en el canal central y 0% en los canales derecho, de LFE y de sonido envolvente).

La tabla 1 muestra una representación generalizada de los metadatos de mezcla (común para ES y PES), mostrando la columna uno las claves de metadatos genéricos, mostrando la columna dos los metadatos correspondientes para ES (p. ej. para Dolby Digital Plus o DD+) y mostrando la columna tres los metadatos correspondientes para PES (p. ej. para Dolby Pulse o DVB):

Tabla 1

Generalizado	ES (DD+)	PES (DVB)
sclmain	extpgmscl	AD_fade_byte
sclasso	pgmscl	-
scl [ch]	extpgmlscl, extpgmrsccl, extpgmcscl, extpgmlsscl, extpgmrsscl, extpgmfescl	AD_gain_byte_center, AD_gain_byte_front, AD_gain_byte_surround
pan	panmean	AD_pan_byte

Para la siguiente sección y el pseudo código que aparece a continuación, la expresión "pref" denota el valor de entrada de balance de mezcla 203 para el ajuste del balance entre las señales principal 201 y asociada 202. El valor de entrada de control de balance de mezcla 203 tendrá un rango de  $[-\infty \dots +\infty]$ , con  $-\infty$  conduciendo a la atenuación completa de la señal asociada 202, siendo 0 un valor neutral que conduce a la mezcla de acuerdo con los metadatos de mezcla 202a y  $\infty$  conduciendo a la atenuación completa de la señal principal 201.

Para este ejemplo que conlleva una señal principal de multicanal 5.1 201 y una señal asociada de multicanal 5.1 202, la aplicación de las técnicas para una señal de multicanal resulta en el siguiente pseudo código, siendo gainA y gainM[ch] las ganancias efectivas o los factores de escala aplicados en la señal asociada 202 y los canales de la señal principal 201, respectivamente.

El pseudo código que aparece a continuación funciona en el dominio dB.

```

45 maxscl = max (scl[c], scl[1], scl[r], scl[ls], scl[rs]);
   mainscltotal = sclmain + maxscl;
   scl[lfe] = min(scl[lfe], maxscl);
   si (pref > mainscltotal - sclasso) {
40 gainA =0;
   gainM[ch] = sclmain + scl[ch] - pref - sclasso ;
   } else {
   gainA = pref - mainscltotal + sclasso;
   gainM[ch] = sclmain + scl[ch] - mainscltotal;

```

}

El pseudo código se ejecuta para cada canal de las señales de entrada.

5 De este modo, cuando el balance de mezcla está en su valor máximo de  $+\infty$  dB, la señal asociada 201 es la señal dominante 206, por lo que la señal principal 201 está completamente atenuada y la señal mezclada 208 es esencialmente la señal asociada cuya escala no está modificada 202. En cambio, cuando el balance de mezcla 203 está en su valor mínimo de  $-\infty$  dB, la señal principal 201 es la señal dominante 206, por lo que la señal asociada 202 está completamente atenuada y la señal mezclada 208 es esencialmente la señal principal cuya escala no está  
10 modificada 201.

Hay que tener en cuenta que en este ejemplo particular, la escala de la señal dominante se puede modificar también ligeramente en el caso en que la señal principal sea la señal dominante (line gainM[ch] = sclmain + scl[ch] - mainscltotal en la cláusula adicional). Esto permite que el algoritmo evite que la ganancia positiva se aplique a la  
15 señal principal, lo cual es posible con los metadatos DD+. El resultado final debería ser que el canal más alto principal esté limitado a 0 dB, lo cual coincide aproximadamente con el principio general de que la señal dominante debería tener una ganancia unitaria.

Para aplicar este pseudo código a un estándar de metadatos específico, los términos de metadatos apropiados se  
20 deben sustituir por las variables genéricas tales como scl[ch]. De este modo, basándose en la forma específica de las señales, incluyendo el número de canales, el tipo de codificación y los metadatos disponibles en las señales, la persona experta en la materia debería ser capaz de asignar las claves de metadatos generalizados apropiadas para un caso específico. Para algunos casos específicos, no todas las claves de metadatos enumeradas anteriormente pueden ser relevantes. Por ejemplo, para la mezcla de dos señales de multicanal 5.1, la persona experta en la  
25 materia debería ser capaz de determinar que el metadato pan ya no está presente o debería ser ignorado ya que la señal asociada ya contiene los mismos canales exactos que la señal principal.

Para el caso en que la señal asociada 202 sea mono y la señal principal 201 no lo sea, puede que se tengan que  
30 aplicar las ganancias de colocación a la señal asociada 201. En este caso, si la señal asociada 201 tiene un canal de LFE, el canal de LFE se puede mezclar tal como se ha descrito en el pseudo código anterior como scl[lfe].

Para la mezcla de DD+ mediante el uso de los metadatos de mezcla del nivel ES, se deberán utilizar las siguientes claves de metadatos de mezcla del nivel ES:

35 extpgmscl, panmean, extpgmslcl, extpgmrsccl, extpgmcscl, extpgmlssccl, extpgmrssccl, extpgmlfesccl, pgmscl.

En general, todos los demás metadatos de mezcla no son necesarios en la mezcla. La clave de metadatos dmxscl se puede utilizar adicionalmente para la mezcla descendente, donde la mezcla descendente es la asignación de la  
40 señal a un formato con un número diferente de canales. La clave de metadatos pgmscl, la cual representa un factor de escala para la señal asociada, se puede utilizar adicionalmente para ajustar los factores de escala de metadatos para el canal principal durante la mezcla. Para el pseudo código anterior, la variable sclclasso representa pgmscl. En la línea del pseudo código, si (pref > mainscltotal - sclclasso), se puede ver que sclclasso, el cual representa la clave de metadatos pgmscl en DD+, influye entonces si existe comparación donde la señal dominante se determina, porque sclclasso se sustrae de mainscltotal antes de la comparación con pref. El uso de sclclasso es, no obstante, opcional y si  
45 sclclasso no está presente o no se desea utilizar el valor, entonces pref se compara sencillamente con mainscltotal.

Para la mezcla de Dolby Pulse/HE-AAC mediante el uso de los metadatos de mezcla del nivel PES, por ejemplo en DVB, se deberán utilizar las siguientes claves de metadatos de mezcla del nivel ES:

50 AD\_gain\_byte\_center, AD\_gain\_byte\_front,  
AD\_gain\_byte\_surround.

Para obtener los mejores resultados, la ganancia máxima resultante de cada

55 [AD\_gain\_byte\_center + AD\_fade\_byte],  
[AD\_gain\_byte\_front + AD\_fade\_byte],  
[AD\_gain\_byte\_surround + AD\_fade\_byte]

no deberá ser superior a +12 dB y estas ganancias efectivas se deberían normalizar de una forma que la ganancia más elevada no exceda los 12 dB, mientras que las relaciones entre AD\_gain\_byte\_center, AD\_gain\_byte\_front, AD\_gain\_byte\_surround se mantendrán, es decir

$$5 \text{ AD\_fade\_byte} = \min(12 \text{ dB}, \text{AD\_fade\_byte} + \max(\text{AD\_gain\_byte\_center}, \text{AD\_gain\_byte\_front}, \text{AD\_gain\_byte\_surround})).$$

La figura 6 muestra una forma de realización adicional de un mezclador 600 para la mezcla de una señal principal de multicanal 5.1 610 con una señal asociada de multicanal 5.1 620 incluyendo los metadatos de mezcla 640, por ejemplo las señales Dolby Digital Plus (DD+ o E-AC-3 (AC-3 mejorado)). Las entradas incluyen una entrada de balance de mezcla 601, una señal principal 610 y una señal asociada 620 con metadatos de mezcla 640. En este ejemplo, la señal principal de entrada 610 y la señal mezclada de salida 630 son señales de multicanal 5.1 que contienen seis canales: izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente derecho y de LFE (efectos de baja frecuencia). De este modo, la señal de entrada principal 610 comprende los seis canales: izquierdo 611, derecho 612, central 613, de sonido envolvente izquierdo 614, de sonido envolvente derecho 615 y de LFE (efectos de baja frecuencia) 616. La señal asociada 620 comprende también seis canales: izquierdo 621, derecho 622, central 623, de sonido envolvente izquierdo 624, de sonido envolvente derecho 625 y de LFE (efectos de baja frecuencia) 626. Los metadatos de mezcla 640 comprenden también un factor de escala primario 647 y factores de escala para los seis canales: izquierdo 641, derecho 642, central 643, de sonido envolvente izquierdo 644, de sonido envolvente derecho 645 y de LFE (efectos de baja frecuencia) 646. En algunos casos, los metadatos de mezcla pueden comprender además un factor de escala para la señal asociada 648. Si este factor de escala para la señal asociada no está presente, entonces se ignora o se trata como unitario o de 0 dB. El mezclador 600 se configura también para recibir la entrada de la entrada de balance de mezcla 601. Basándose en los metadatos de mezcla 640 y la entrada de balance de mezcla 601, los factores de escala para cada canal 651, 652, 653, 654, 655, 656 se determinan basándose en la lógica en el pseudo código anterior y que se muestra también en la figura 7A, la cual es un diagrama de flujos que describe cómo se determina el factor de escala para cada canal. A partir de los factores de escala de mezcla 651, 652, 653, 654, 655, 656, se modifica la escala de cada canal de las señales de entrada, donde la modificación de escala primaria se produce en la señal no dominante y la escala de la señal dominante solamente se modifica mínimamente o su escala no se modifica en absoluto. Finalmente, los canales correspondientes de las señales de canal con su escala modificada 660, 670 se combinan en canales mezclados y los seis canales mezclados 631, 632, 633, 634, 635, 636 comprenden entonces la señal de salida mezclada 5.1 630. De este modo, la señal 5.1 mezclada de salida 630 comprende también seis canales, izquierdo 631, derecho 632, central 633, de sonido envolvente izquierdo 634, de sonido envolvente derecho 635 y de LFE (efectos de baja frecuencia) 616, los cuales se calculan de acuerdo con la lógica de la figura 7A.

La figura 7A muestra un diagrama de flujos que describe el ejemplo de la figura 6 para la mezcla de una señal principal de entrada de multicanal 5.1 610 con una señal asociada de entrada de multicanal 5.1 620, donde la escala de la señal dominante 206 se puede modificar también ligeramente si la señal principal 610 es la señal dominante 206. Tal como se muestra en la figura 6, las entradas comprenden una entrada de balance de mezcla 601, una señal principal 610 y una señal asociada 620 con metadatos de mezcla 640, donde las señales de entrada y la señal de salida mezclada 630 son señales de multicanal 5.1. De este modo, las señales 610, 620, 630 comprenden seis canales y los metadatos de mezcla 640 comprenden un factor de escala primario 647 y factores de escala para los seis canales: izquierdo 641, derecho 642, central 643, de sonido envolvente izquierdo 644, de sonido envolvente derecho 654 y de LFE (efectos de baja frecuencia) 646. En algunos casos, los metadatos de mezcla 640 pueden comprender además un factor de escala para la señal asociada, sclasso 648. Si sclasso 648 no está presente o no se desea utilizarlo en la mezcla, se utiliza un valor de 0 dB para sclasso 648 y no se realiza ningún ajuste de los demás factores de escala. En el paso 701, el factor de escala máximo de los cinco "canales normales", izquierdo 641, derecho 642, central 643, de sonido envolvente izquierdo 644 y de sonido envolvente derecho 645 se determina como "maxscl". Entonces, en el paso 702, mainscltotal se calcula como la suma de sclmain 647 y maxscl, donde sclmain 647 es el factor de escala primario de los metadatos de mezcla, maxscl es el máximo de los factores de escala de los canales normales 641, 642, 643, 644, 645. El paso 703 en el cual se determina el factor de escala de LFE se describe con más detalle en la figura 7a.

La figura 7B muestra un diagrama de flujos que proporciona detalles para el paso 703 de la figura 7A para la determinación del factor de escala para el canal de LFE con respecto al ejemplo de la figura 7A para la mezcla de dos señales de multicanal 5.1. En el paso 703a, los metadatos de mezcla para el canal de LFE 646 se comparan con el factor de escala máximo para los demás canales, maxscl, determinados en el paso 701 de la figura 7A. Si el factor de escala de LFE 646 es superior o igual a maxscl, entonces en el paso 703y, el factor de escala para el canal

de LFE, gain[lfe] 656, se determinará como maxscl. De otra manera, para el caso negativo, en el paso 703n, el factor de escala para el canal de LFE, gain[lfe] 656, se determina como el factor de escala en los metadatos de mezcla para el canal de LFE 646. Entonces, en el paso 703c, gain [lfe] 656 se determina a partir de los valores calculados en los pasos 703y o 703n. El gain[lfe] 656, el cual es el factor de escala que se va a aplicar al canal de LFE para la  
5 mezcla de las señales, se utilizará entonces con posterioridad en los pasos 707n y 707y de la figura 7A.

Entonces, continuando con la figura 7A, en el paso 704, pref, el cual es el valor de la entrada de control de balance de mezcla 601, se compara a mainscltotal menos sclasso 648, donde sclasso 648 es el factor de escala para la señal asociada. Hay que tener en cuenta que sclasso 648 puede no estar presente o se puede ignorar, en cuyo caso  
10 pref se compara sencillamente a mainscltotal. Por ejemplo, para las señales DD+, la clave de metadatos pgmscl corresponde a sclasso 648. Si pref 602 es superior a mainscltotal menos sclasso 648, entonces en el paso 705y, la señal asociada 620 se determinará como la señal dominante 206. Entonces, en 706y, no existe ganancia para la señal asociada, ya que es la señal dominante, por lo que la ganancia para la señal asociada es unitaria o cero dB, indicando que la escala del canal asociado no se ha modificado. Entonces, en 707y, la ganancia para cada canal de  
15 la señal principal para todos los canales (izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente derecho y de lfe), gainM[ch] 651, 652, 653, 654, 655 se determina como la suma del factor de escala primario, sclmain 647 y los factores de escala para los canales normales scl[ch] 641, 642, 643, 644, 645 menos pref 602 menos sclasso. El cálculo en 707y se repite para los seis canales totales de la señal principal.

20 Si la decisión del paso 704 es negativa al igual que pref 602 no es superior a mainscltotal, entonces en el paso 705n, la señal principal se determinará como la señal dominante. Entonces, en 706n, el factor de ganancia para la señal asociada, gainA, se determinará como pref 602 menos mainscltotal más sclasso 648. En este caso especial de este ejemplo particular, la escala de la señal principal 610 se modifica también incluso si se ha determinado como la señal dominante y los factores de escala 651, 652, 653, 654, 655, 656 para cada canal de la señal principal 620 se  
25 calculan como sclmain 647 más el factor de escala para los canales correspondientes scl[ch] 641, 642, 643, 644, 645, menos mainscltotal. De este modo, en el paso 707n, la ganancia para cada canal de la señal principal para todos los canales (izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente derecho y de lfe), gainM[ch] 651, 652, 653, 654, 655 se determina como la suma de sclmain 647 y scl[ch] 641, 642, 643, 644, 645 menos mainscltotal. El cálculo en 707n se repite para los seis canales totales. Las ganancias de los pasos 706y,  
30 707y, 706n y 707n así como las señales de entrada 620, 610 se incorporan entonces en el paso 708, donde la escala de cada canal de las señales principales 611, 612, 613, 614, 615, 616 y asociadas 621, 622, 623, 624, 625, 626 se modifica de acuerdo con los factores de escala determinados 651, 652, 653, 654, 655, 656. Hay que tener en cuenta que en el caso típico, únicamente se modifica la escala de la señal dominante, pero en este caso especial, la escala de la señal no dominante también se puede modificar. Entonces en el paso 709, el cual se repite para cada  
35 uno de los seis canales, la señal asociada 670 y principal 660 con su escala modificada para cada canal se mezclan en seis canales mezclados 631, 632, 633, 634, 635, 636. Finalmente, en el paso 710 los seis canales mezclados 631, 632, 633, 634, 635, 636 se combinan en una señal única para comprender la señal de salida mezclada 5.1 630 del paso 711.

40 La figura 8 muestra una forma de realización adicional de un mezclador 800 para la mezcla de una señal principal de multicanal 5.1 810 con una señal asociada mono+LFE de dos canales 820 incluyendo los metadatos de mezcla 840, por ejemplo las señales Dolby Digital Plus (DD+ o E-AC-3 (AC-3 mejorado)). La entrada incluye una entrada de balance de mezcla 801, una señal principal 810 y una señal asociada 820 con metadatos de mezcla 840. En este ejemplo, la señal principal de entrada 810 y la señal mezclada de salida 830 son señales de multicanal 5.1 que  
45 contienen seis canales: izquierdo, derecho, central, de sonido envolvente izquierdo, de sonido envolvente derecho y de LFE (efectos de baja frecuencia). De este modo, la señal de entrada principal 810 comprende los seis canales: izquierdo 811, derecho 812, central 813, de sonido envolvente izquierdo 814, de sonido envolvente derecho 815 y de LFE (efectos de baja frecuencia) 816. La señal asociada mono+LFE 820 comprende únicamente dos canales: central 823 y de LFE (efectos de baja frecuencia) 826. Los metadatos de mezcla 840 comprenden también  
50 metadatos con valores de modificación de escala para el factor de escala primario 847 y los seis canales: izquierdo 841, derecho 842, central 843, de sonido envolvente izquierdo 844, de sonido envolvente derecho 845 y de LFE (efectos de baja frecuencia) 846. El mezclador 800 se configura también para recibir la entrada de la entrada de balance de mezcla 801. Basándose en los metadatos de mezcla 840 y la entrada de control de balance de mezcla 801, se determina un factor de escala para cada canal 851, 852, 853, 854, 855, 856 basado en la lógica en el  
55 pseudo código anterior y que se muestra también en la figura 7, la cual es un diagrama de flujos que describe cómo se determina la escala para cada canal. A partir de los factores de escala de mezcla, se modifica la escala de cada canal de la señal de entrada no dominante 207, donde la escala de la señal dominante 206 no se modifica generalmente en absoluto o en casos especiales, su escala únicamente se modifica mínimamente. El valor panmean se utiliza para colocar el canal mono del canal mono+LFE de la señal asociada en los canales 5.1. Panmean es un

- valor que va de 0 a 239 representando 0 a 358,5 grados en los pasos de 1,5 grados, donde 0 grados es la dirección del altavoz central. Para obtener más detalles sobre la aplicación de panmean, consulte la sección "E.4.3.5 Colocación" de ETSI TS 102 366 v1.2.1. Finalmente, los pares de seis canales para las dos señales de entrada con su escala modificada se combinan por pares para cada canal en los seis canales mezclados 831, 832, 833, 834, 835, 836. Los seis canales mezclados 831, 832, 833, 834, 835, 836 comprenden entonces la señal de salida mezclada 5.1 830. De este modo, la señal mezclada de salida 5.1 830 comprende también seis canales, izquierdo 831, derecho 832, central 833, de sonido envolvente izquierdo 834, de sonido envolvente derecho 835 y de LFE (efectos de baja frecuencia) 836, los cuales se calculan cada uno de acuerdo con la lógica de la figura 7.
- 10 En general, para obtener los mejores resultados, el modo de canal de la señal mezclada deberá ser el mismo que el de la señal de audio principal descodificada, aunque se puede cambiar por supuesto en una etapa posterior, p. ej. mezclado descendente. Para la mayoría de los casos, se espera que el valor de la entrada de balance de mezcla sea reducido, generalmente de hasta 10 dB, aunque la técnica soporta valores arbitrariamente amplios (y reducidos).
- 15 En algunos casos, el número de canales de las señales principal y asociada puede diferir. En un caso tal, se pueden seguir aplicando las técnicas descritas en este documento, donde puede que algunos de los canales de entrada tengan que ser asignados a los canales de la señal mezclada que pueden estar o no presentes en la señal de entrada. Para obtener los mejores resultados, la señal asociada debería contener únicamente ubicaciones de canal que también estén presentes en la señal principal, a menos que la señal asociada sea mono (acmod 1) o los canales de la señal asociada se puedan asignar de otra manera, apropiadamente, a los canales de la señal principal. Asimismo, un canal de efectos de baja frecuencia (LFE) debería estar presente preferiblemente sólo en la señal asociada si la señal principal también contiene un canal de LFE. Por otra parte, el canal de LFE de la señal asociada se debería ignorar en el proceso de mezcla. Asimismo, si la señal asociada es mono+LFE, LFE se debería mezclar
- 20 preferiblemente en el LFE de audio principal y mono se debería colocar preferiblemente en los canales principales de la señal de audio principal.
- 25 En general, para el proceso de mezcla es preferible que las señales de entrada se normalicen. La normalización se puede aplicar tanto antes como después de la determinación de la señal dominante, ya que los resultados serán los mismos. En la práctica, es preferible garantizar que el valor dialnorm de contenido de las señales de entrada se establezca correctamente y para ambas señales principal y asociada para que estén en el nivel de diálogo 31 antes de la mezcla. Posteriormente, el nivel de la señal dominante, la cual se considera la "señal en la que se centra la atención", deberá de forma ideal salir también del mezclador en dialnorm 31. Si las señales de entrada no se normalizan, entonces la normalización se debería realizar generalmente antes de la mezcla, aunque las técnicas
- 30 también se puedan aplicar sin normalización y la señal mezclada entonces normalizada, si es necesario.
- 35 Al mezclar señales de multicanal de mezcla, las relaciones entre scl [ch] se mantendrán preferiblemente en la señal saliente. No obstante, si scl[lfe] tiene un valor superior al de los demás scl [ch] (transmitido o cero por defecto), se modificará hacia el valor máximo del otro scl [ch].
- 40 Se debería tener en cuenta que para las señales Dolby Pulse, los metadatos de mezcla se incluyen típicamente en la transmisión de audio secundario, tal como se define en la sección "E.2" de ETSI TS 101 154 V1.9.1 y ETSI TS 102 366 V1.2.1, Anexo E. (consulte también el "libro azul de DVB A0001 r8" (DVB blue book A0001 r8). Independientemente de cómo se transmitan o se empaqueten los metadatos en la señal de transmisión, las técnicas
- 45 descritas en este documento se podrán seguir aplicando mientras que los metadatos de mezcla estén disponibles. De este modo, puede que haya que adaptar las formas de realización y ejemplos descritos apropiadamente para manejar tal caso.
- 50 El ejemplo anterior con una pista asociada que contiene los comentarios del director es únicamente un ejemplo. Las técnicas descritas pueden aplicarse no obstante a cualquier mezcla de señales principales y asociadas, independientemente del contenido incluido en las señales. Por ejemplo, la primera señal podría contener un programa de deportes en directo donde la segunda señal proporciona comentarios locales y/o comentarios en un idioma alternativo. Aquí existen múltiples posibilidades en cuanto a cómo se podrían distribuir las señales. Por ejemplo, la primera señal podría ser la pista de audio general sin comentarios y la segunda señal podría tener
- 55 comentarios locales.
- Un caso de uso adicional para las técnicas descritas en este documento podría referirse a una pista independiente para la dificultad de audición, donde la señal asociada contiene mejoras especiales para mejorar la comprensibilidad

del diálogo. De manera similar, como una ayuda para los espectadores visualmente afectados, la señal asociada podría contener una explicación oral de los contenidos de la escena.

No obstante, las técnicas descritas en este documento no se limitan solamente al diálogo, sino que se pueden aplicar también a todos los tipos de señales de audio tales como la música, ya que las técnicas se basan en el nivel de sonido percibido. Por ejemplo, en referencia a la música, una pista asociada podría contener una voz o una pista de instrumento de superposición adicional para música; en otras palabras, la pista principal podría ser la pista instrumental básica con, digamos, únicamente piano y guitarra acústica, una primera pista asociada podría contener una superposición de voz, una segunda música de rock de pista asociada con bajo y guitarra eléctrica y una tercera pista con una superposición tecno con un ritmo rápido del bajo. En este caso, el autor de contenido tendría la oportunidad de especificar las características de mezcla deseadas especificando cómo hay que mezclar cada pista asociada y las técnicas descritas en este documento permitirían al oyente personalizar estos ajustes mediante el ajuste de la entrada de balance para las señales mezcladas. Como otro ejemplo, las técnicas se podrían aplicar también a una señal de música donde los diferentes canales contengan diferentes aspectos musicales tales como instrumentos como cuerdas o percusión y vocales, donde los diferentes canales se mezclen a través de los metadatos y una entrada de balance de mezcla.

Las técnicas descritas en este documento dependen de la determinación de la señal dominante en la señal mezclada. De este modo, la señal que se determina como la señal dominante recibe la ganancia compuesta más elevada. Además de mantener un nivel de sonido percibido constante de la señal mezclada a través del uso del control de entrada de balance de mezcla, el usuario puede controlar el balance de mezcla en todo momento. Además, los cambios de ganancia son también fluidos a lo largo de toda la gama.

Aunque las técnicas descritas en este documento se han ilustrado con ejemplos específicos, tales como las señales de multicanal de mezcla que contienen pistas de diálogo, estas técnicas no deberían limitarse a estos ejemplos o formas de realización ya que las técnicas se pueden aplicar igualmente a una situación en la que las señales principal y asociada tengan una composición muy diferente, siempre y cuando los metadatos de mezcla y la entrada de balance de mezcla se puedan utilizar para determinar la señal dominante para la señal mezclada.

En el presente documento, se han descrito diversos métodos y disposiciones para la mezcla de señales de audio. Mediante el uso de estos métodos y/o dispositivos tales como los mezcladores de señal, es posible para mantener un nivel de sonido percibido consistente para una señal mezclada mientras se permite el ajuste del balance entre la señal principal y la asociada. Los métodos se pueden realizar sin pruebas ni análisis de las señales de entrada ni procesamiento de alta potencia, ya que los cálculos se pueden realizar fácilmente en tiempo real mediante el uso de los metadatos que acompañan a las señales de entrada por medio de dispositivos electrónicos relativamente sencillos y básicos.

Se debería tener en cuenta que la descripción y los dibujos ilustran meramente los principios de los métodos y los sistemas propuestos. Se apreciará de este modo que los expertos en la materia serán capaces de idear diversas disposiciones que, aunque no se describan o se muestren explícitamente en este documento, incorporan los principios de los métodos y sistemas propuestos y se consideran como parte de la revelación de este documento. Además, todas las exposiciones de este documento que relatan los principios, aspectos y formas de realización de los métodos y dispositivos propuestos, así como los ejemplos específicos del mismo, están destinados a abarcar los equivalentes del mismo.

Además, todos los ejemplos relatados en este documento están destinados principalmente para ser expresa y únicamente con fines pedagógicos para ayudar al lector en la comprensión de los principios de los métodos y dispositivos propuestos y los conceptos aportados por los inventores para impulsar la técnica y se deben interpretar sin limitación a tales ejemplos y condiciones específicamente relatados.

Además, se debería tener en cuenta que cualquier diagrama de bloques de este documento representa vistas conceptuales de dispositivos ilustrativos que incorporan los principios de la invención. De manera similar, se apreciará que cualquier organigrama, diagrama de flujos, diagrama de transición de estado, pseudo código y otros representan diversos procesos los cuales se pueden representar sustancialmente en medios legibles por ordenador y ejecutados de este modo por un ordenador o procesador, tanto si tal ordenador o procesador se muestra explícitamente o no.

**REIVINDICACIONES**

- 1.-Un método para la mezcla de dos señales de audio de entrada en una única señal de audio mezclada mientras se mantiene un nivel de sonido percibido de la señal de audio mezclada, comprendiendo el método:
- 5 la recepción de una señal de audio de entrada principal;
- la recepción de una señal de audio de entrada asociada; donde la señal de audio de entrada asociada se acopla con la señal de audio de entrada principal;
- 10 la recepción de los metadatos de mezcla, los cuales contienen información de modificación de escala para la modificación de la escala de la señal de audio de entrada principal y que especifican cómo se deberían mezclar la señal de audio de entrada principal y la señal de audio de entrada asociada, con el fin de generar una señal de audio mezclada en el nivel de sonido percibido; donde la información de modificación de escala a partir de los metadatos
- 15 de mezcla comprende un factor de escala de metadatos para la señal de audio de entrada principal, para la modificación de la escala de la señal de audio de entrada principal relativa a la señal de audio de entrada asociada;
- la recepción de una entrada de balance de mezcla, la cual denota un balance ajustable entre la señal de audio de entrada principal y la señal de audio de entrada asociada, donde la entrada de balance de mezcla comprende
- 20 información de modificación de escala la cual permite una desviación de una ponderación de la señal de audio de entrada principal y la señal de audio de entrada asociada en la señal de audio mezclada como se especifica en los metadatos de mezcla;
- la identificación de una señal dominante como la señal de audio de entrada principal o la señal de audio de entrada asociada a partir de la información de modificación de escala proporcionada por los metadatos de mezcla y a partir
- 25 de la entrada de balance de mezcla, donde la otra señal de audio de entrada respectiva se identifica entonces como una señal no dominante; y donde la señal dominante se identifica mediante la comparación de la entrada de balance de mezcla con el factor de escala de metadatos para la señal de audio de entrada principal;
- 30 la modificación de escala de la señal no dominante en relación con la señal dominante; y
- la combinación de la señal no dominante con su escala modificada con la señal dominante para producir la señal de audio mezclada.
- 35 2.-El método de la reivindicación 1 comprendiendo además:
- la determinación de un factor de escala para la modificación de escala de la señal no dominante basada en la información de modificación de escala a partir de los metadatos de mezcla y basada en la entrada de balance de
- 40 mezcla;
- donde el factor de escala determinado se utiliza para la modificación de escala de la señal no dominante.
- 3.-El método de las reivindicaciones 1-2, donde
- 45 si el valor de la entrada de balance de mezcla es superior al factor de escala de metadatos para la señal de audio de entrada principal a partir de los metadatos de mezcla, entonces la señal de audio de entrada asociada se determinará como la señal dominante; de lo contrario, la señal de audio de entrada principal se determinará como la señal dominante.
- 50 4.-El método de cualquiera de las reivindicaciones 1-3, donde si la señal de audio de entrada asociada se identifica como la señal dominante, el método comprenderá además:
- el cálculo del factor de escala para la señal de audio de entrada principal como el factor de escala de metadatos para la señal de audio de entrada principal a partir de los metadatos de mezcla menos el valor de la entrada de
- 55 balance de mezcla.
- 5.-El método de cualquiera de las reivindicaciones 1-3, donde si la señal de audio de entrada principal se identifica como la señal dominante, el método comprenderá además:

el cálculo del factor de escala para la señal de audio de entrada asociada como el valor de la entrada de balance de mezcla menos el factor de escala de metadatos a partir de los metadatos de mezcla para la señal de audio de entrada principal.

5 6.-El método de cualquiera de las reivindicaciones 1-5 donde los metadatos de mezcla se contienen en la señal de audio de entrada asociada.

7.-El método de cualquiera de las reivindicaciones 1-6, donde el nivel de sonido percibido se basa en un nivel ponderado de media de la señal de audio de entrada principal y la señal de audio de entrada asociada.

10

8.-El método de cualquiera de las reivindicaciones 1-7, donde la entrada de balance de mezcla comprende una entrada de usuario externa que proporciona un valor de valores negativos a valores positivos.

9.-Un dispositivo para la mezcla de señales que aplica cualquiera de los métodos de mezcla de señales de las reivindicaciones 1-8.

15

10.-Un medio de almacenamiento legible por un procesador que almacena instrucciones legibles por ordenador para ejecutar el método de cualquiera de las reivindicaciones 1-8.

20 11.-Un dispositivo para la mezcla de señales de audio de entrada en una única señal de audio mezclada mientras se mantiene un nivel de sonido percibido de la señal de audio mezclada, comprendiendo:

un receptor para recibir una señal de audio de entrada principal, una señal de audio de entrada asociada y metadatos de mezcla; donde la señal de audio de entrada asociada se acopla con la señal de audio de entrada principal; donde los metadatos de mezcla contienen información de modificación de escala para la modificación de la escala de la señal de audio de entrada principal; donde la información de modificación de escala especifica cómo se deberían mezclar la señal de audio de entrada principal y la señal de audio de entrada asociada, con el fin de generar una señal de audio mezclada en el nivel de sonido percibido; donde la información de modificación de escala a partir de los metadatos de mezcla comprende un factor de escala de metadatos para la señal de audio de entrada principal, para la modificación de la escala de la señal de audio de entrada principal relativa a la señal de audio de entrada asociada;

25

30

una entrada de balance de mezcla, la cual denota el balance ajustable entre las señales de audio de entrada principal y asociada; donde la entrada de balance de mezcla comprende la información de modificación de escala la cual permite una desviación de una ponderación de la señal de audio de entrada principal y la señal de audio de entrada asociada en la señal de audio mezclada como se especifica en los metadatos de mezcla; y

35

una unidad de mezcla configurada para identificar una señal dominante como la señal de audio de entrada principal o la señal de audio de entrada asociada a partir de la información de modificación de escala proporcionada por los metadatos de mezcla y a partir de la entrada de balance de mezcla, donde la otra señal de audio de entrada respectiva se identifica entonces como una señal no dominante; y donde la señal dominante se identifica mediante la comparación de la entrada de balance de mezcla con el factor de escala de metadatos para la señal de audio de entrada principal;

40

45 donde la unidad de mezcla se configura además para modificar la escala de la señal no dominante en relación con la señal dominante; y para combinar la señal no dominante con su escala modificada con la señal dominante para producir la señal de audio mezclada.

12.-El dispositivo de la reivindicación 11, donde el factor de escala para la señal no dominante se determina directamente a partir de la información de modificación de escala a partir de los metadatos de mezcla y la entrada de balance de mezcla; la escala de la señal no dominante se modifica utilizando el factor de escala determinado; y la señal no dominante con su escala modificada se combina con la señal dominante en la señal mezclada.

50

13.-El dispositivo de la reivindicación 11 ó 12, donde los metadatos de mezcla se integran con la señal de audio de entrada asociada.

55

14.-El dispositivo de cualquiera de las reivindicaciones 11-13, donde la entrada de balance de mezcla comprende además una entrada de usuario externa que proporciona un valor desde valores negativos a valores positivos.

15.-El dispositivo de cualquiera de las reivindicaciones 11-14, donde el receptor se configura además para recibir señales de audio de entrada principal de multicanal y señales de audio de entrada asociadas y la unidad de mezcla se configura para mezclar las señales de audio de entrada principal de multicanal y las señales de audio de entrada asociadas.

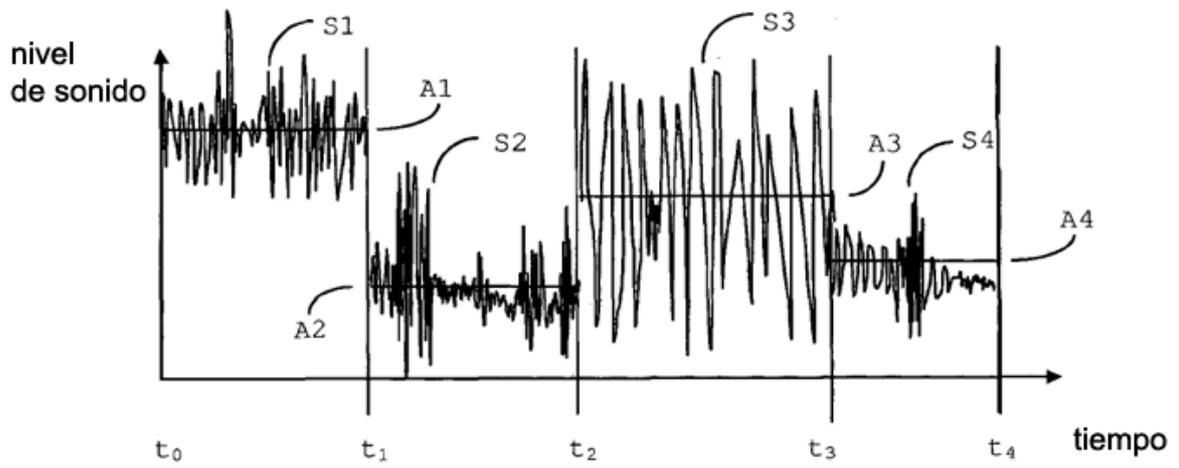


FIGURA 1A

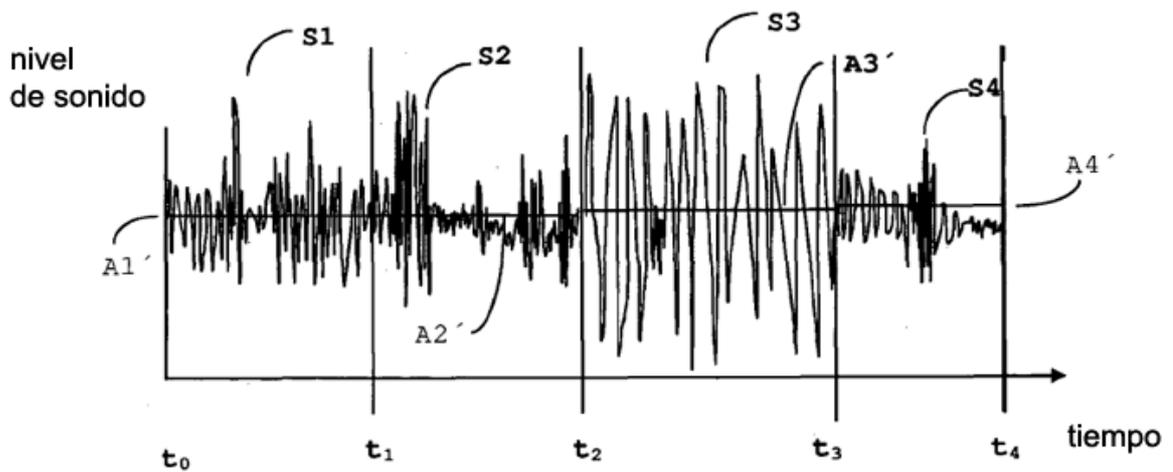


FIGURA 1B

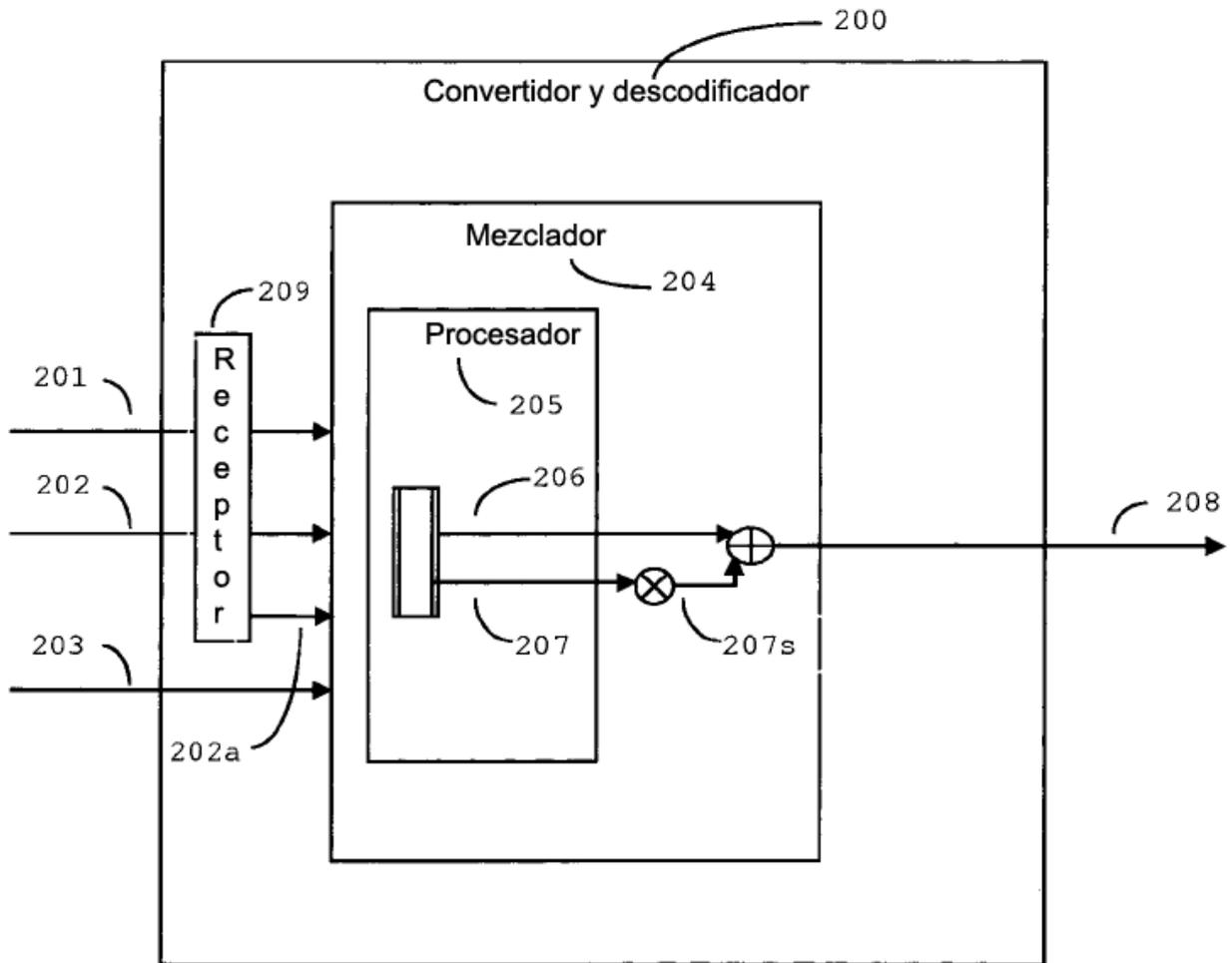


FIGURA 2

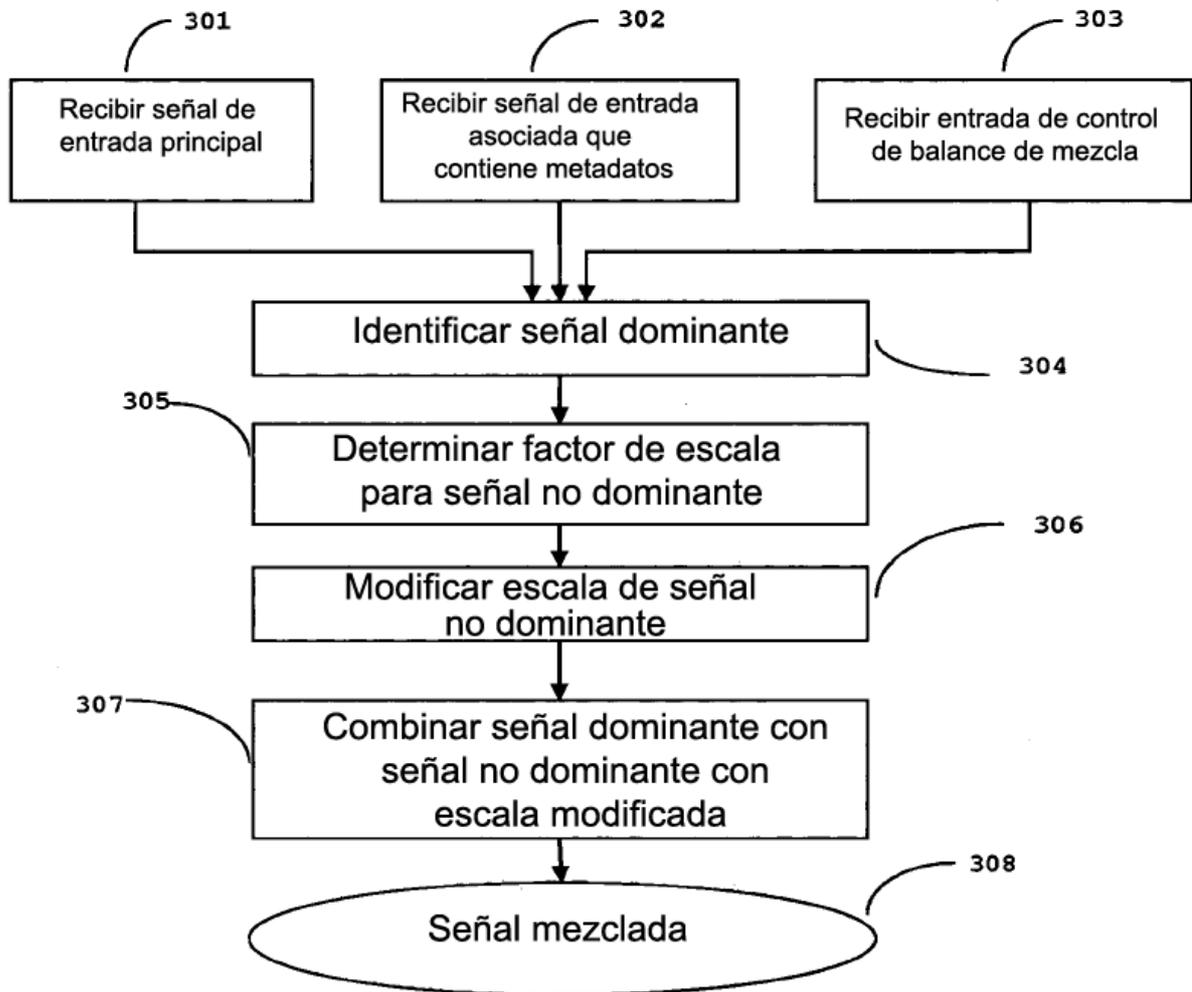


FIGURA 3

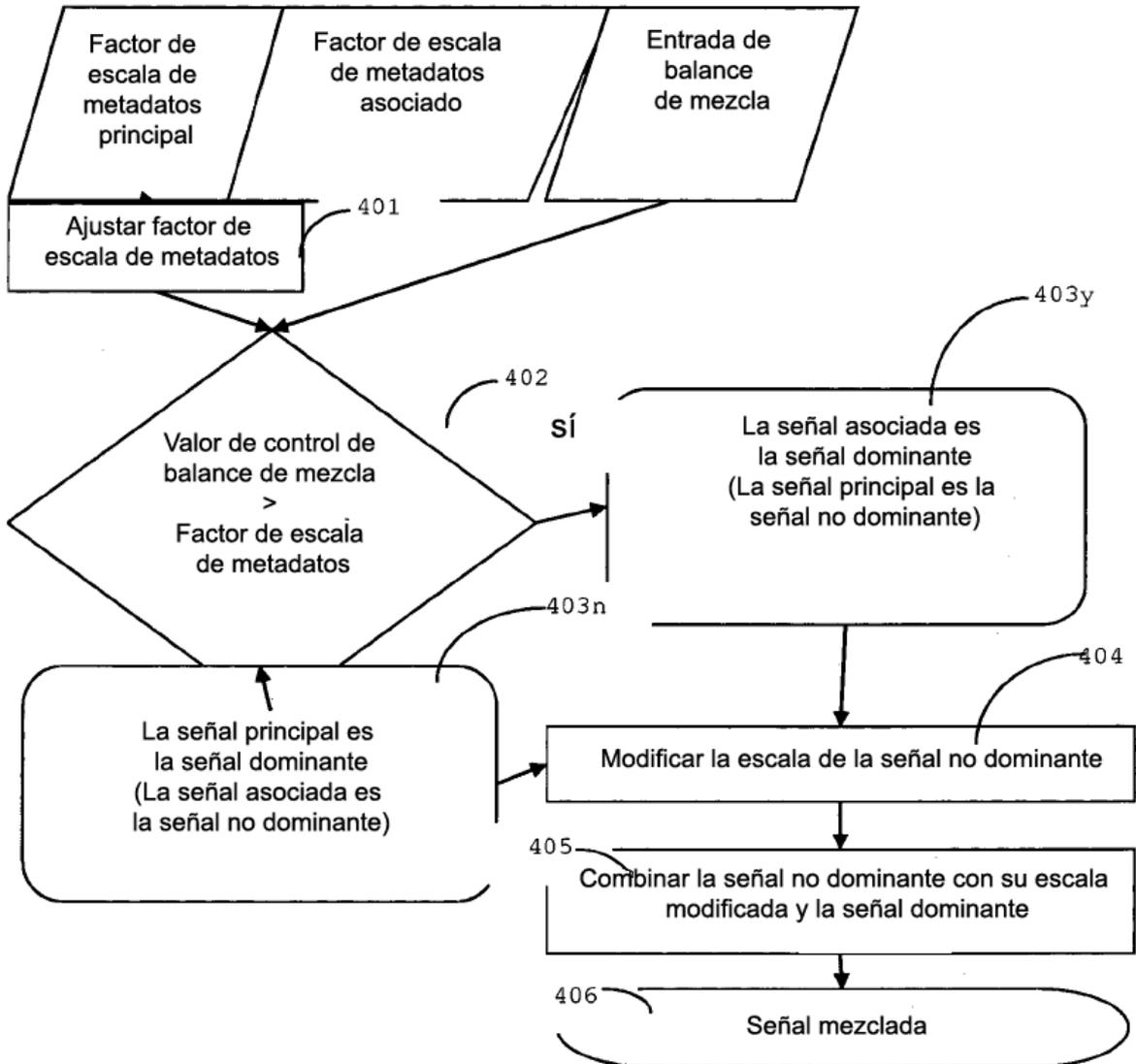


FIGURA 4

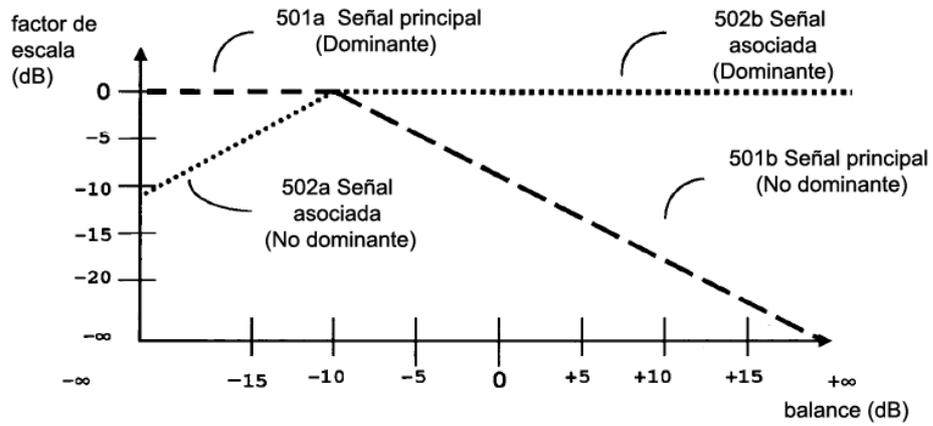


FIGURA 5A

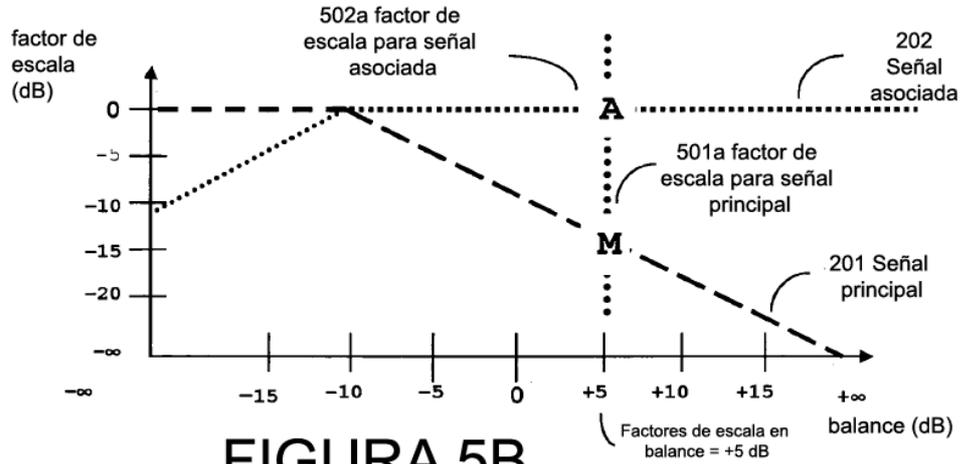


FIGURA 5B

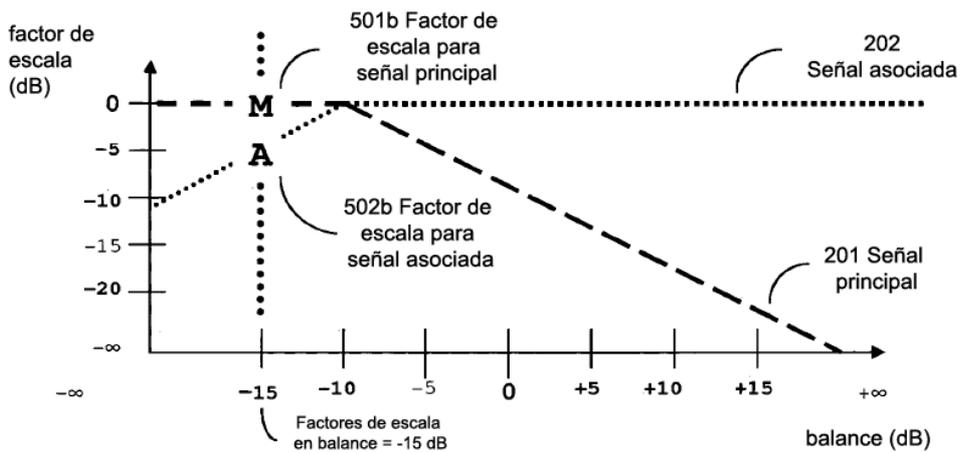


FIGURA 5C

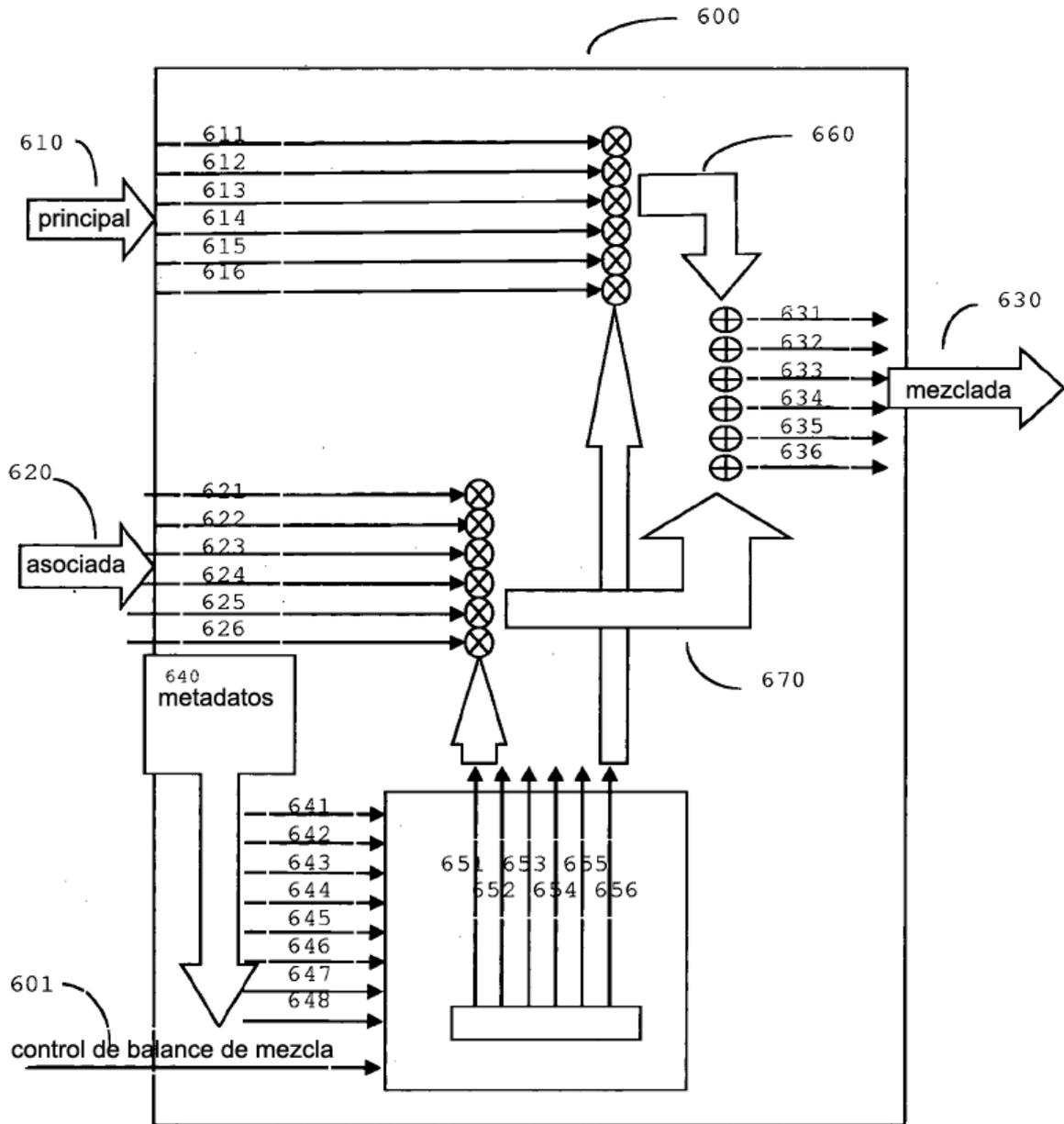


FIGURA 6

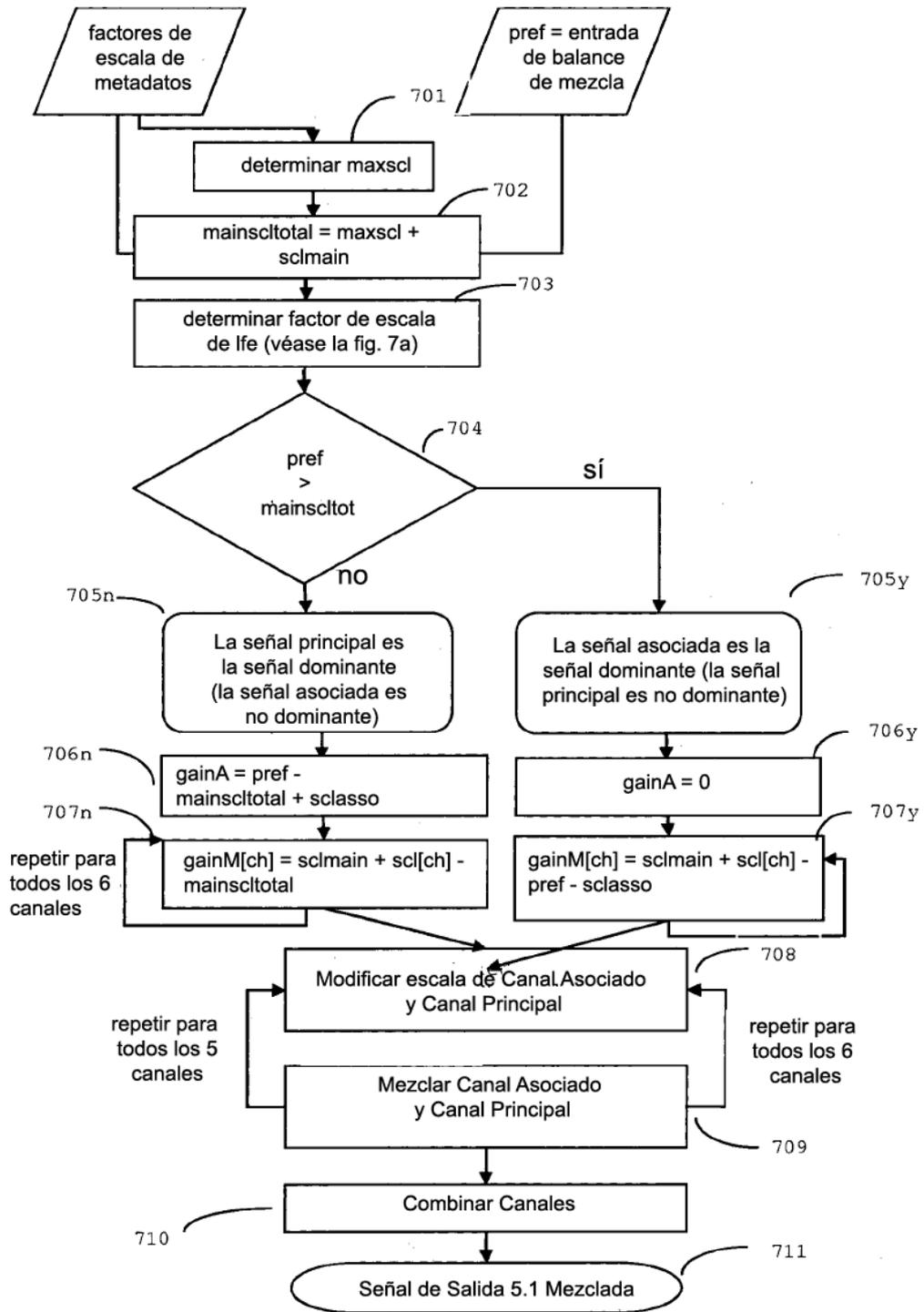


FIGURA 7A

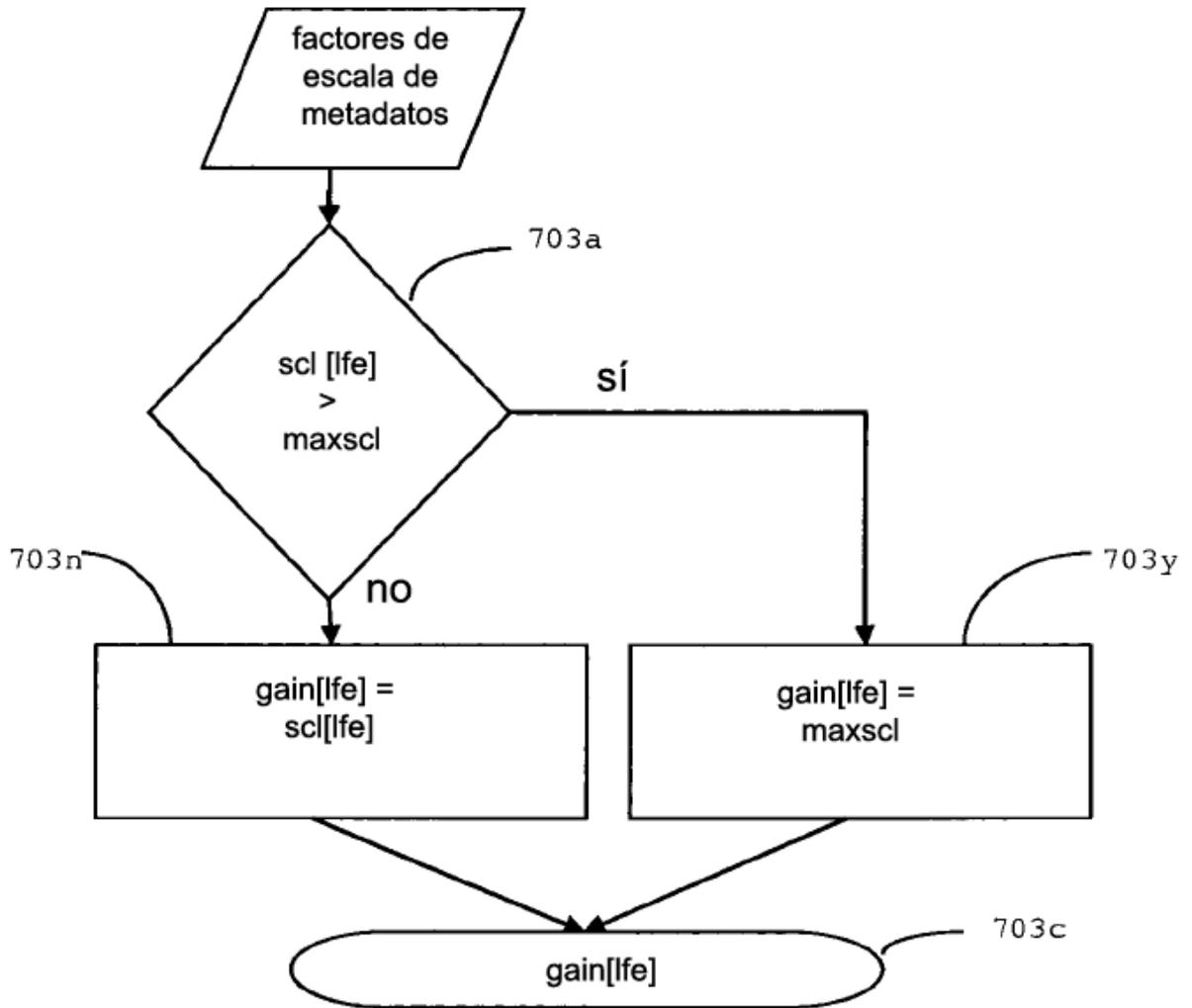


FIGURA 7B

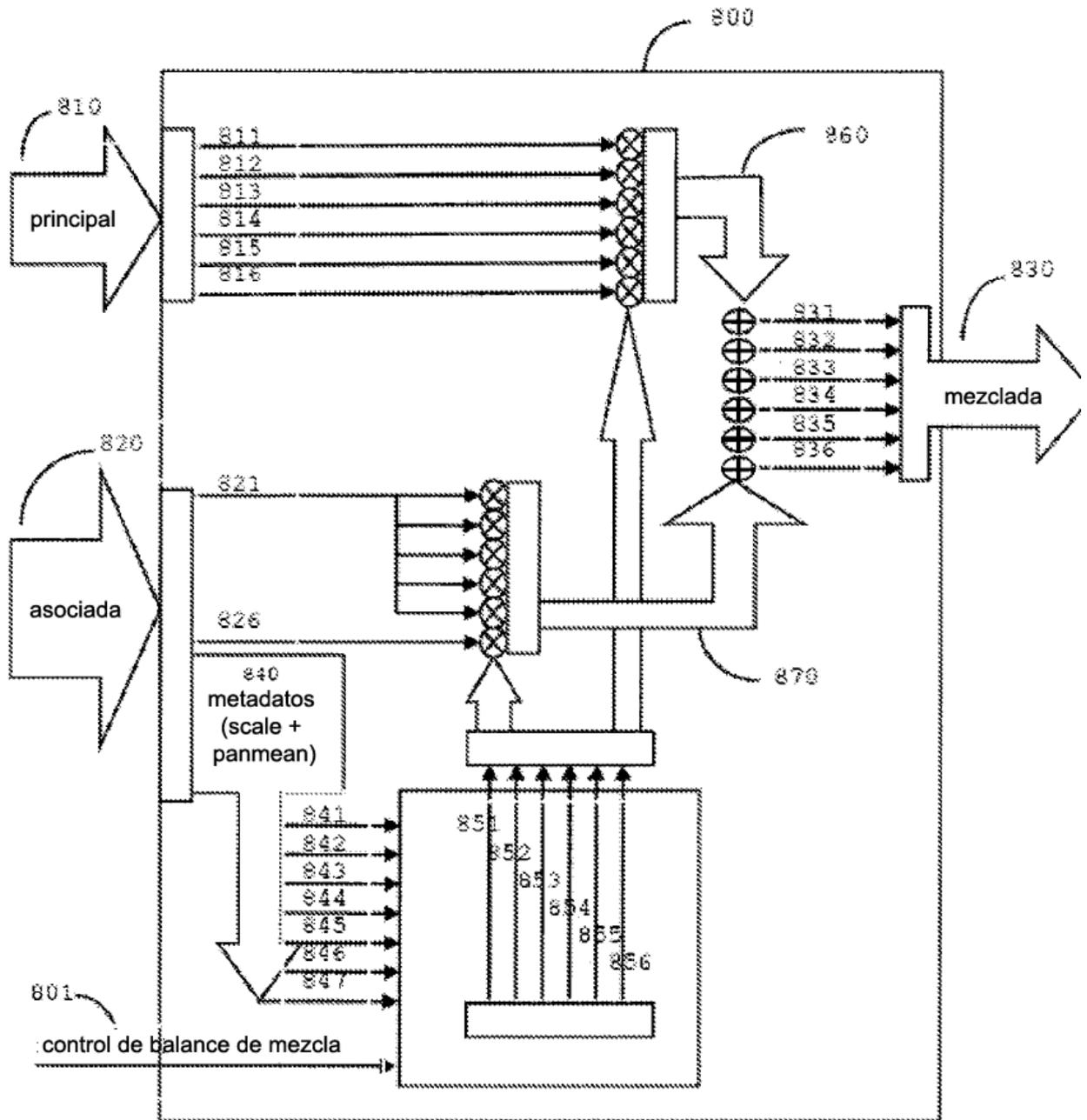


FIGURA 8