

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 327**

51 Int. Cl.:

G10L 21/0316 (2013.01)

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.09.2015 PCT/EP2015/072578**

87 Fecha y número de publicación internacional: **07.04.2016 WO16050854**

96 Fecha de presentación y número de la solicitud europea: **30.09.2015 E 15770958 (5)**

97 Fecha y número de publicación de la concesión europea: **12.12.2018 EP 3201918**

54 Título: **Método de descodificación y descodificador para la mejora del diálogo**

30 Prioridad:

02.10.2014 US 201462059015 P
04.03.2015 US 201562128331 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
16.04.2019

73 Titular/es:

DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost, NL

72 Inventor/es:

KOPPENS, JEROEN y
EKSTRAND, PER

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 709 327 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

Método de descodificación y descodificador para la mejora del diálogo

Campo técnico

5 La invención divulgada en esta memoria se refiere generalmente a la codificación de audio. En particular, se refiere a métodos y dispositivos para la mejora del diálogo en sistemas de audio basados en canales.

Antecedentes

10 La mejora del diálogo se refiere a potenciación del diálogo en relación con otros contenidos de audio. Esto puede aplicarse, por ejemplo, con el fin de permitir a personas auditivamente impedidas seguir los diálogos de una película. Para el contenido de audio basado en canales, el diálogo está presente, por lo común, en varios canales y también está mezclado con otros contenidos de audio. Es, por lo tanto, una tarea no baladí mejorar el diálogo.

15 Existen diversos métodos conocidos para llevar a cabo una mejora del diálogo en un descodificador. De acuerdo con algunos de estos métodos, se descodifica en primer lugar el contenido completo de canales, esto es, la totalidad de la configuración de los canales, y, a continuación, se utilizan parámetros de mejora del diálogo recibidos para predecir el diálogo basándose en el contenido completo de canales. El diálogo predicho se utiliza entonces para mejorar el diálogo en los canales relevantes. A lo largo de estas líneas, la Norma de Compresión de Audio Digital (AC-4) (ETSI TS [Especificación Técnica –“Technical Specification”– del Instituto Europeo de Normas de Telecomunicaciones – “European Telecommunications Standards Institute”–] 103 190 v1.1.1) sugiere extraer todo el contenido de canales de las señales de mezcla descendente recibidas y proporcionar un subconjunto del contenido de canales como entrada para una mejora del diálogo. Sin embargo, tales métodos de descodificación se fundamentan en un descodificador que es capaz de descodificar la configuración completa de canales.

20 Sin embargo, los descodificadores de baja complejidad no están diseñados, por lo común, para descodificar la configuración completa de canales. En lugar de esto, un descodificador de baja complejidad puede descodificar y suministrar como salida un bajo número de canales que representan una versión mezclada en sentido descendente de la configuración completa de canales. De acuerdo con esto, la configuración completa de canales no está disponible en el descodificador de baja complejidad. Puesto que los parámetros de mejora del diálogo se han definido con respecto a los canales de la configuración de canales completa (o al menos con respecto a algunos de los canales de la configuración completa de canales), los métodos de mejora de diálogo conocidos no pueden ser aplicados directamente por un descodificador de baja complejidad. En particular, este es el caso, puesto que los canales con respecto a los cuales se aplican los parámetros de mejora del diálogo pueden aún seguir estando mezclados con otros canales.

25 Hay, por tanto, margen para mejoras que permitan a un descodificador de baja complejidad aplicar una mejora del diálogo sin tener que descodificar la configuración de canales completa.

Breve descripción de los dibujos

35 En lo que sigue, se describirán realizaciones proporcionadas a modo de ejemplo con mayor detalle y con referencia a los dibujos que se acompañan, en los cuales:

La Figura 1a es una ilustración esquemática de una configuración de canales de 7.1+4 que se ha mezclado en sentido descendente para obtener una mezcla descendente de 5.1, de acuerdo con un primer esquema de mezcla en sentido descendente.

40 La Figura 1b es una ilustración esquemática de una configuración de canales de 7.1+4 que se ha mezclado en sentido descendente para obtener una mezcla descendente de 5.1, de acuerdo con un segundo esquema de mezcla en sentido descendente.

La Figura 2 es una ilustración esquemática de un descodificador de la técnica anterior para llevar a cabo una mejora del diálogo en una configuración completa de canales descodificados.

La Figura 3 es una ilustración esquemática de la mejora del diálogo de acuerdo con un primer modo.

45 La Figura 4 es una ilustración esquemática de la mejora del diálogo de acuerdo con un segundo modo.

La Figura 5 es una ilustración esquemática de un descodificador de acuerdo con realizaciones proporcionadas a modo de ejemplo.

La Figura 6 es una ilustración esquemática de un descodificador de acuerdo con realizaciones proporcionadas a modo de ejemplo.

50 La Figura 7 es una ilustración esquemática de un descodificador de acuerdo con realizaciones proporcionadas a modo de ejemplo.

La Figura 8 es una ilustración esquemática de un codificador correspondiente a uno cualquiera de los descodificadores de la Figura 2, la Figura 5, la Figura 6 y la Figura 7.

5 La Figura 9 ilustra métodos para computar una operación BA de tratamiento combinada, compuesta por dos suboperaciones, u operaciones subordinadas, A y B, basándose en parámetros que controlan cada una de las suboperaciones.

Todas las figuras son esquemáticas y, generalmente, tan solo muestran los elementos que son necesarios para ilustrar la invención, mientras que otros elementos pueden haberse omitido o ser meramente sugeridos.

Descripción detallada

10 A la vista de lo anterior, es un propósito proporcionar un descodificador y métodos asociados que permitan la aplicación de la mejora del diálogo sin tener que descodificar la configuración completa de canales.

I. Vista general

De acuerdo con un primer aspecto, realizaciones proporcionadas a modo de ejemplo proporcionan un método para mejorar el diálogo en un descodificador de un sistema de audio. El método comprende las etapas de:

15 recibir una pluralidad de señales de mezcla descendente que constituyen una mezcla en sentido descendente de una pluralidad mayor de canales;

recibir parámetros para la mejora del diálogo, de tal manera que los parámetros son definidos con respecto a un subconjunto de la pluralidad de canales que incluye canales que comprenden diálogo, de tal modo que el subconjunto de la pluralidad de canales se mezcla en sentido descendente para obtener un subconjunto de la pluralidad de señales de mezcla descendente;

20 recibir parámetros de reconstrucción que permiten la reconstrucción paramétrica de canales que se han mezclado en sentido descendente para obtener el subconjunto de la pluralidad de señales de mezcla descendente;

25 mezclar en sentido ascendente únicamente el subconjunto de la pluralidad de señales de mezcla descendente paramétricamente, basándose en los parámetros de reconstrucción, a fin de reconstruir únicamente un subconjunto de la pluralidad de canales que incluye el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo;

aplicar la mejora del diálogo al subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, utilizando los parámetros para la mejora del diálogo con el fin de proporcionar al menos una señal mejorada en el diálogo; y

30 proporcionar versiones mejoradas en diálogo del subconjunto de la pluralidad de señales de mezcla descendente, mezclando la al menos una señal mejorada en diálogo con al menos otra de dicho subconjunto de la pluralidad de señales de mezcla descendente.

35 Con esta disposición, el descodificador no tiene que reconstruir la configuración de canales completa para llevar a cabo la mejora del diálogo, por lo que se reduce la complejidad. En lugar de ello, el descodificador reconstruye los canales que se requieren para la aplicación de la mejora del diálogo. Esto incluye, en particular, un subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros recibidos para la mejora del diálogo. Una vez que se ha llevado a cabo la mejora del diálogo, esto es, cuando se ha determinado al menos una señal mejorada en diálogo basándose en los parámetros para la mejora del diálogo y en el subconjunto de la pluralidad de canales con respecto al cual se han definido estos parámetros, se determinan versiones mejoradas en diálogo de las señales de mezcla descendente recibidas, al someter la(s) señal(es) mejorada(s) en diálogo a un procedimiento de mezcla.
40 Como resultado de ello, se producen versiones mejoradas en diálogo de las señales de mezcla descendente para su subsiguiente reproducción por parte del sistema de audio.

En realizaciones proporcionadas a modo de ejemplo, una operación de mezcla en sentido ascendente puede ser completa (reconstrucción del conjunto completo de canales codificados) o parcial (reconstrucción de un subconjunto de los canales).

45 Tal y como se utiliza en esta memoria, una señal de mezcla descendente se refiere a una señal que es una combinación de una o más señales / canales.

50 Tal y como se utiliza en esta memoria, mezclar en sentido ascendente paramétricamente se refiere a la reconstrucción de una o más señales / canales a partir de una señal de mezcla descendente por medio de técnicas paramétricas. Se hace énfasis en que las realizaciones proporcionadas a modo de ejemplo que se divulgan en esta memoria no están limitadas por el contenido basado en canales (en el sentido de señales de audio asociadas con direcciones, ángulos y/o posiciones en el espacio invariables o predefinidos), sino que también se extienden a contenido basado en objetos.

De acuerdo con realizaciones proporcionadas a modo de ejemplo, en la etapa de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente paramétricamente, se utilizan señales no descorrelacionadas con el fin de reconstruir el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo.

5 Esto es ventajoso en tanto en cuanto reduce la complejidad computacional, al mismo tiempo que mejora la calidad de las versiones mejoradas en diálogo resultantes de las señales de mezcla descendente (esto es, la calidad de la salida). Con más detalle, las ventajas obtenidas mediante el uso de señales descorrelacionadas a la hora de mezclar en sentido ascendente, se ven reducidas por la subsiguiente mezcla a la que se somete la señal mejorada en diálogo. Por lo tanto, el uso de señales descorrelacionadas puede, ventajosamente, omitirse, con lo que se ahorra complejidad de computación. De hecho, el uso de señales descorrelacionadas en la mezcla en sentido ascendente puede, en combinación con la mejora del diálogo, dar lugar a una peor calidad, puesto que puede tener como resultado una reverberación descorrelacionadora en el diálogo mejorado.

15 De acuerdo con realizaciones proporcionadas a modo de ejemplo, la mezcla se realiza con arreglo a parámetros de mezcla que describen una contribución de la al menos una señal mejorada en diálogo a las versiones mejoradas en diálogo del subconjunto de la pluralidad de señales de mezcla descendente. Pueden existir, por tanto, algunos parámetros de mezcla que describen el modo como mezclar la al menos una señal mejorada en diálogo al objeto de proporcionar versiones mejoradas en diálogo del subconjunto de la pluralidad de señales de mezcla descendente. Por ejemplo, los parámetros de mezcla pueden darse en la forma de factores de ponderación, o pesos, que describen cuánto de la al menos una señal mejorada en diálogo ha de mezclarse en cada una de las señales de mezcla descendente del subconjunto de la pluralidad de señales de mezcla descendente, para obtener las versiones mejoradas en diálogo del subconjunto de la pluralidad de señales de mezcla descendente. Tales pesos pueden darse, por ejemplo, en la forma de parámetros de representación que son indicativos de las posiciones espaciales asociadas con la al menos una señal mejorada en diálogo, en relación con posiciones espaciales asociadas con la pluralidad de canales, y, por tanto, el subconjunto correspondiente de señales de mezcla descendente. De acuerdo con otros ejemplos, los parámetros de mezcla pueden indicar si la al menos una señal mejorada en diálogo ha de contribuir, o no, tal como estar incluida en, una particular de la versión mejorada en diálogo del subconjunto de señales de mezcla descendente. Por ejemplo, un '1' puede indicar que una señal mejorada en diálogo ha de incluirse a la hora de formar una particular de la versión mejorada en diálogo de las señales de mezcla descendente, y un '0' puede indicar que esta no debe ser incluida.

20 En la etapa de someter la al menos una señal mejorada en diálogo a mezcla de tal manera que se proporcionen versiones mejoradas en diálogo del subconjunto de señales de mezcla descendente, las señales mejoradas en diálogo pueden ser mezcladas con otras señales / canales.

25 De acuerdo con realizaciones proporcionadas a modo de ejemplo, la al menos una señal mejorada en diálogo se mezcla con canales que se han reconstruido en la etapa de mezcla ascendente, pero que no han sido sometidos a mejora del diálogo. Más detalladamente, la etapa de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente paramétricamente, puede comprender reconstruir al menos un canal adicional, además de la pluralidad de canales con respecto a los cuales se han definido los parámetros para la mejora del diálogo, y de tal manera que la mezcla comprende mezclar el al menos un canal adicional conjuntamente con la al menos una señal mejorada en diálogo. Por ejemplo, todos los canales que se mezclan en sentido descendente para formar el subconjunto de la pluralidad de señales de mezcla descendente pueden ser reconstruidos e incluidos en la mezcla. En tales realizaciones, existe, por lo común, una correspondencia directa entre cada señal mejorada en diálogo y el canal.

30 De acuerdo con otras realizaciones proporcionadas a modo de ejemplo, la al menos una señal mejorada en diálogo se mezcla con el subconjunto de la pluralidad de señales de mezcla descendente. Con mayor detalle, la etapa de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente paramétricamente, puede comprender reconstruir únicamente el subconjunto de la pluralidad de canales con respecto al cual se definen los parámetros para la mejora del diálogo, y la etapa de aplicar la mejora del diálogo puede comprender predecir y mejorar un componente de diálogo a partir del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, utilizando los parámetros para la mejora del diálogo, a fin de proporcionar la al menos una señal mejorada en diálogo, y la mezcla puede comprender mezclar la al menos una señal mejorada en diálogo con el subconjunto de la pluralidad de señales de mezcla descendente. Tales realizaciones sirven, por lo tanto, para predecir y mejorar el contenido de diálogo y mezclarlo para formar el subconjunto de la pluralidad de señales de mezcla descendente.

35 Generalmente, debe apreciarse que un canal puede comprender contenido de diálogo que se mezcla con contenido que no es de diálogo. Por otra parte, el contenido de diálogo correspondiente a un único diálogo puede ser mezclado en varios canales. Con la predicción de un componente de diálogo a partir del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, quiere decirse generalmente que el contenido de diálogo es extraído, es decir, separado de los canales y combinado con el fin de reconstruir el diálogo.

40 La calidad de la mejora del diálogo puede, adicionalmente, ser mejorada recibiendo y utilizando una señal de audio que representa diálogo. Por ejemplo, la señal de audio que representa diálogo puede ser codificada a una baja

velocidad de transferencia de bits, lo que provoca señales espurias que son bien audibles cuando se escuchan por separado. Sin embargo, cuando se utilizan conjuntamente con la mejora paramétrica del diálogo, es decir, en la etapa de aplicar la mejora del diálogo al subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, utilizando los parámetros para la mejora del diálogo, es posible perfeccionar la mejora del diálogo, por ejemplo, en términos de calidad de audio. Más particularmente, el método puede comprender, adicionalmente: recibir una señal de audio que representa diálogo, de tal manera que la etapa de aplicar la mejora del diálogo comprende aplicar la mejora del diálogo al subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, adicionalmente utilizando la señal de audio que representa diálogo.

En algunas realizaciones, los parámetros de mezcla pueden estar ya disponibles en el descodificador, es decir, pueden estar codificados en hardware. Este será el caso, en particular, si la al menos una señal mejorada en diálogo es siempre mezclada de la misma manera, por ejemplo, si esta es siempre mezclada con los mismos canales reconstruidos. En otras realizaciones, el método comprende recibir parámetros de mezcla para la etapa de someter a mezcla la al menos una señal mejorada en diálogo. Por ejemplo, los parámetros de mezcla pueden formar parte de los parámetros de mejora del diálogo.

De acuerdo con realizaciones proporcionadas a modo de ejemplo, el método comprende recibir parámetros de mezcla que describen un esquema de mezcla en sentido descendente que describe en qué señal de mezcla descendente es mezclado cada uno de la pluralidad de canales. Por ejemplo, si cada señal mejorada en diálogo corresponde a un canal, que, a su vez, es mezclado con otros canales reconstruidos, la mezcla se lleva a cabo de acuerdo con el esquema de mezcla en sentido descendente, de tal modo que cada canal es mezclado en la señal de mezcla descendente correcta.

El esquema de mezcla en sentido descendente puede variar con el tiempo, es decir, puede ser dinámico, con lo que se incrementa la flexibilidad del sistema.

El método puede comprender, de manera adicional, recibir datos que identifican el subconjunto de la pluralidad de canales con respecto al cual se han definido para la mejora del diálogo. Por ejemplo, los datos que identifican el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, pueden ser incluidos en los parámetros para la mejora del diálogo. De esta manera, puede señalarse al descodificador con respecto a qué canales ha de llevarse a cabo la mejora del diálogo. Alternativamente, tal información puede hacerse disponible en el descodificador, por ejemplo, al estar codificado en hardware, lo que significa que los parámetros para la mejora del diálogo son siempre definidos con respecto a los mismos canales. En particular, el método puede incluir, adicionalmente, recibir información que indica qué señales de entre las señales mejoradas en diálogo han de ser sometidas a mezcla. Por ejemplo, el método de acuerdo con esta variante puede llevarse a cabo por medio de un sistema de descodificación que funciona en un modo particular, de tal manera que las señales mejoradas en diálogo no se mezclan de vuelta hasta formar un conjunto completamente idéntico de señales de mezcla descendente, como era el uso para proporcionar las señales mejoradas en diálogo. De esta manera, la operación de mezcla puede, en la práctica, restringirse a una selección incompleta (una o más señales) del subconjunto de la pluralidad de señales de mezcla descendente. Las demás señales mejoradas en diálogo se añaden a señales de mezcla descendente ligeramente diferentes, tales como señales de mezcla descendente que se han sometido a una conversión de formato. Una vez que los datos que identifican el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, y el esquema de mezcla en sentido descendente son conocidos, es posible encontrar el subconjunto de la pluralidad de señales de mezcla descendente en el que es mezclado en sentido descendente el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo. Con mayor detalle, los datos que identifican el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, conjuntamente con el esquema de mezcla en sentido descendente, pueden ser utilizados para encontrar el subconjunto de la pluralidad de señales de mezcla descendente en el que se mezcla en sentido descendente el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo.

Las etapas de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente, aplicar la mejora del diálogo y mezclar pueden llevarse a cabo como operaciones matriciales definidas por los parámetros de reconstrucción, los parámetros para la mejora del diálogo y los parámetros de mezcla, respectivamente. Esto es ventajoso por cuanto el método puede ser implementado de una manera eficiente llevando a cabo una multiplicación matricial.

Es más, el método puede comprender combinar, por multiplicación matricial, las operaciones matriciales correspondientes a las etapas de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente, aplicar la mejora del diálogo y mezclar, en una única operación matricial, antes de su aplicación al subconjunto de la pluralidad de señales de mezcla descendente. De esta forma, las diferentes operaciones matriciales pueden ser combinadas en una única operación matricial, con lo que se mejora adicionalmente la eficiencia y se reduce la complejidad computacional del método.

Los parámetros de mejora del diálogo y/o los parámetros de reconstrucción pueden ser dependientes de la frecuencia, con lo que se permite que los parámetros difieran entre bandas de frecuencias diferentes. De este modo, la mejora del diálogo y la reconstrucción pueden ser optimizadas en las diferentes bandas de frecuencias, con lo que se mejora

la calidad del audio de salida.

Con mayor detalle, los parámetros para la mejora del diálogo pueden ser definidos con respecto a un primer conjunto de bandas de frecuencias, y los parámetros de reconstrucción pueden ser definidos con respecto a un segundo conjunto de bandas de frecuencias, de tal manera que el segundo conjunto de bandas de frecuencias es diferente del primer conjunto de bandas de frecuencias. Esto puede ser ventajoso a la hora de reducir la velocidad de transferencia de bits para transmitir los parámetros para la mejora del diálogo y los parámetros de reconstrucción en una corriente de bits, cuando, por ejemplo, el procedimiento de reconstrucción requiere parámetros con una resolución en frecuencia más alta que el procedimiento de mejora del diálogo, y/o cuando, por ejemplo, el procedimiento de mejora del diálogo se lleva a cabo en una anchura de banda más pequeña que la del procedimiento de reconstrucción.

De acuerdo con realizaciones proporcionadas a modo de ejemplo, los valores (preferiblemente discretos) de los parámetros para la mejora del diálogo pueden ser recibidos repetidamente y asociados con un primer conjunto de instantes de tiempo, en el que los respectivos valores se aplican exactamente. En la presente divulgación, es la intención que la afirmación del efecto de que un valor se aplica, o es conocido, "exactamente" en un cierto instante de tiempo, signifique que el valor ha sido recibido por el descodificador, por lo común, conjuntamente con una indicación explícita o implícita de un instante de tiempo en el que este se aplica. En contraposición, un valor que se ha interpolado o predicho para un cierto instante de tiempo no se aplica "exactamente" en ese instante de tiempo en este sentido, sino que es una estimación por parte del descodificador. "Exactamente" no implica que el valor consiga una reconstrucción exacta de una señal de audio. Entre instantes de tiempo consecutivos del conjunto, puede prescribirse una primera configuración de interpolación predefinida. Una configuración de interpolación, que define el modo como estimar un valor aproximado de un parámetro en un instante de tiempo situado entre dos instantes de tiempo límite del conjunto en los cuales se conocen los valores del parámetro, puede ser, por ejemplo, una interpolación lineal o constante por tramos. Si el instante de tiempo de predicción está alejado a una cierta distancia de uno de los instantes de tiempo límite, una configuración de interpolación lineal se basa en la suposición de que el valor del parámetro en el instante de tiempo de predicción depende linealmente de dicha distancia, en tanto que una configuración de interpolación constante por tramos garantiza que el valor del parámetro no cambia entre cada valor conocido y el siguiente. Pueden existir también otras configuraciones de interpolación posibles, incluyendo, por ejemplo, configuraciones que se sirven de polinomios de grados mayores que uno, curvas de aproximación, funciones racionales, procedimientos gaussianos, polinomios trigonométricos, aproximación de ondas pequeñas, o una combinación de los mismos, a fin de estimar el valor del parámetro en un instante de tiempo de predicción dado. El conjunto de instantes de tiempo puede no ser explícitamente transmitido o establecido, sino, en lugar de ello, inferido a partir de la configuración de interpolación, por ejemplo, el punto de partida o punto final de un intervalo de interpolación lineal, el cual puede ser implícitamente fijado en los límites de trama de un algoritmo de tratamiento de audio. Los parámetros de reconstrucción pueden ser recibidos de una forma similar: los valores (preferiblemente discretos) de los parámetros de reconstrucción pueden estar asociados con un segundo conjunto de instantes de tiempo, y puede llevarse a cabo una segunda configuración de interpolación entre instantes de tiempo consecutivos.

El método puede incluir, de manera adicional, seleccionar un tipo de parámetro, consistiendo el tipo bien en parámetros para la mejora del diálogo, o bien en parámetros de reconstrucción, de tal manera que el conjunto de instantes de tiempo asociado con el tipo seleccionado incluye al menos un instante de predicción que es un instante de tiempo que está fuera del conjunto asociado con el tipo no seleccionado. Por ejemplo, si el conjunto de instantes de tiempo con los que están asociados los parámetros de reconstrucción incluye un cierto instante de tiempo que está fuera del conjunto de instantes de tiempo con los que están asociados los parámetros para la mejora del diálogo, ese cierto instante de tiempo será un instante de predicción si el tipo seleccionado de parámetro son los parámetros de reconstrucción y el tipo no seleccionado de parámetro son los parámetros para la mejora del diálogo. De una manera similar, en otra situación, el instante de predicción puede, en lugar de ello, ser encontrado en el conjunto de instantes de tiempo con los que están asociados los parámetros para la mejora del diálogo, y se conmutarán los tipos seleccionado y no seleccionado. Preferiblemente, el tipo de parámetro seleccionado es el tipo que tenga la más alta densidad de instantes de tiempo con valores de parámetro asociados; en un caso de uso dado, esto puede reducir la cantidad total de operaciones de predicción necesarias.

El valor de los parámetros del tipo no seleccionado, en el instante de predicción, puede ser predicho. La predicción puede llevarse a cabo utilizando un método de predicción adecuado, tal como la interpolación o la extrapolación, y a la vista de la configuración de interpolación predefinida para los tipos de parámetros.

El método puede incluir la etapa de computar, basándose en al menos el valor predicho de los parámetros del tipo no seleccionado y en un valor recibido de los parámetros del tipo seleccionado, una operación de tratamiento compuesta que representa al menos la mezcla en sentido ascendente del subconjunto de las señales de mezcla descendente, seguida de la mejora del diálogo en el instante de predicción. Además de en valores de los parámetros de reconstrucción y de los parámetros para la mejora del diálogo, la computación puede estar basada en otros valores, tales como valores de parámetros para la mezcla, y la operación de tratamiento compuesta puede representar también la etapa de mezclar una señal de mejora de diálogo de vuelta hasta formar una señal de mezcla descendente.

El método puede incluir la etapa de computar, basándose en al menos un valor (recibido o predicho) de los parámetros del tipo seleccionado y en al menos un valor (recibido o predicho) de los parámetros del tipo no seleccionado, de tal manera que al menos alguno de los valores es un valor recibido, la operación de tratamiento compuesta en un instante

de tiempo adyacente del conjunto asociado con el tipo seleccionado o el no seleccionado. El instante de tiempo adyacente puede ser bien más temprano o bien más tardío que el instante de predicción, y no es esencial exigir que el instante de tiempo adyacente sea el vecino más próximo en términos de distancia.

5 En el método, las etapas de mezclar en sentido ascendente el subconjunto de la pluralidad de señales de mezcla descendente y aplicar la mejora del diálogo pueden llevarse a cabo entre el instante de predicción y el instante de tiempo adyacente, por medio de un valor interpolado de la operación de tratamiento compuesta computada. Interpolando la operación de tratamiento compuesta computada, es posible conseguir una complejidad computacional reducida. Al no interpolar los dos tipos de parámetros por separado, y al no formar un producto (por ejemplo, una operación de tratamiento compuesta) en cada punto de interpolación, pueden requerirse menos operaciones matemáticas de adición y de multiplicación para conseguir un resultado igualmente útil en términos de la calidad de audición percibida.

10 De acuerdo con realizaciones proporcionadas a modo de ejemplo adicionales, la operación de tratamiento compuesta en el instante de tiempo adyacente puede computarse basándose en un valor recibido de los parámetros del tipo seleccionado y en un valor predicho de los parámetros del tipo no seleccionado. Es también posible la situación inversa en la que la operación de tratamiento compuesta en el instante de tiempo adyacente puede computarse basándose en un valor predicho de los parámetros del tipo seleccionado y en un valor recibido de los parámetros del tipo no seleccionado. Situaciones en las que un valor del mismo tipo de parámetro es un valor recibido en el instante de predicción y un valor predicho en el instante de tiempo adyacente, pueden producirse si, por ejemplo, los mismos instantes del conjunto con el que está asociado el tipo de parámetro seleccionado se encuentran situados estrictamente entre los instantes de tiempo del conjunto con el que está asociado el tipo de parámetro no seleccionado.

15 De acuerdo con realizaciones proporcionadas a modo de ejemplo, la operación de tratamiento compuesta en el instante de tiempo adyacente puede ser computada basándose en un valor recibido de los parámetros del tipo de parámetro seleccionado, y en un valor recibido de los parámetros del tipo de parámetro no seleccionado. Tales situaciones pueden producirse, por ejemplo, si se reciben valores exactos de parámetros de los dos tipos para los límites de trama, pero también –para el tipo seleccionado– para un instante de tiempo a medio camino entre los límites. Entonces el instante de tiempo adyacente es un instante de tiempo asociado con un límite de trama, y el instante de tiempo de predicción está situado a medio camino entre los límites de trama.

20 De acuerdo con realizaciones proporcionadas a modo de ejemplo adicionales, el método puede incluir, de manera adicional, seleccionar, basándose en las primera y segunda configuraciones de interpolación, una configuración de interpolación compuesta de acuerdo con una regla de selección predefinida, de tal manera que la interpolación de las respectivas operaciones de tratamiento compuestas computadas es de conformidad con la configuración de interpolación compuesta. La regla de selección predefinida puede haberse definido para el caso de que las primera y segunda configuraciones de interpolación sean iguales, y puede también haberse definido para el caso de que las primera y segunda configuraciones de interpolación sean diferentes. Como ejemplo de ello, si la primera configuración de interpolación es lineal (y, preferiblemente, si existe una relación lineal entre parámetros y propiedades cuantitativas de la operación de mejora de diálogo) y la segunda configuración de interpolación es constante por tramos, puede seleccionarse la configuración de interpolación compuesta de manera que sea lineal.

25 De acuerdo con realizaciones proporcionadas a modo de ejemplo, la predicción del valor de los parámetros del tipo no seleccionado en el instante de predicción se realiza de acuerdo con la configuración de interpolación para los parámetros del tipo no seleccionado. Esto puede implicar el uso de un valor exacto del parámetro del tipo no seleccionado, en un instante de tiempo del conjunto asociado con el tipo no seleccionado, que es adyacente al instante de predicción.

30 De acuerdo con realizaciones proporcionadas a modo de ejemplo, la operación de tratamiento compuesta se computa como una única operación matricial y, a continuación, se aplica al subconjunto de la pluralidad de señales de mezcla descendente. Preferiblemente, las etapas de mezclar en sentido descendente y aplicar la mejora del diálogo se llevan a cabo como operaciones matriciales definidas por los parámetros de reconstrucción y los parámetros para la mejora del diálogo. Como configuración de interpolación compuesta, puede seleccionarse una configuración de interpolación lineal, y el valor interpolado de las respectivas operaciones de tratamiento compuestas computadas puede computarse por interpolación matricial lineal. La interpolación puede restringirse a los elementos de matriz que cambian entre el instante de predicción y el instante de tiempo adyacente, a fin de reducir la complejidad computacional.

35 De acuerdo con realizaciones proporcionadas a modo de ejemplo, las señales de mezcla descendente recibidas pueden ser segmentadas en tramas temporales, y el método puede incluir, en una operación de estado estacionario, una etapa consistente en recibir al menos un valor de los tipos de parámetro respectivos que se aplica exactamente en un instante de tiempo de cada trama temporal. Tal y como se utiliza en esta memoria, la expresión 'estado estacionario' se refiere a un funcionamiento que no implica la presencia de porciones inicial y final, por ejemplo, una canción, y a un funcionamiento que no implica transitorios internos que necesiten subdivisión de tramas.

40 De acuerdo con un segundo aspecto, se proporciona un producto de programa informático que comprende un medio legible por computadora con instrucciones para llevar a cabo el método del primer aspecto. El medio legible por computadora puede ser un medio o dispositivo legible por computadora no transitorio.

De acuerdo con un tercer aspecto, se proporciona un descodificador para mejorar el diálogo en un sistema de audio, de tal manera que el descodificador comprende:

un componente de recepción, configurado para recibir:

5 una pluralidad de señales de mezcla descendente, que son una mezcla en sentido descendente d una pluralidad de canales más grande,

parámetros para la mejora del diálogo, de tal manera que los parámetros se definen con respecto a un subconjunto de la pluralidad de canales que incluye canales que comprenden diálogo, de modo que el subconjunto de la pluralidad de canales se mezcla en sentido descendente hasta formar un subconjunto de la pluralidad de señales de mezcla descendente, y

10 parámetros de reconstrucción, que permiten la reconstrucción paramétrica de canales que son mezclados en sentido descendente hasta formar el subconjunto de la pluralidad de señales de mezcla descendente;

15 un componente de mezcla en sentido ascendente, configurado para mezclar en sentido ascendente únicamente el subconjunto de la pluralidad de señales de mezcla descendente paraméricamente basándose en los parámetros de reconstrucción, a fin de reconstruir únicamente un subconjunto de la pluralidad de canales que incluye el subconjunto de la pluralidad de canales con respecto al cual se definen los parámetros para la mejora del diálogo; y

20 un componente de mejora del diálogo, configurado para aplicar una mejora del diálogo al subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, utilizando los parámetros para la mejora del diálogo con el fin de proporcionar al menos una señal mejorada en diálogo; y

un componente de mezcla, configurado para proporcionar versiones mejoradas en diálogo del subconjunto de la pluralidad de señales de mezcla descendente, mezclando la al menos una señal mejorada en diálogo con al menos otro de dichos subconjuntos de la pluralidad de señales de mezcla descendente.

25 Generalmente, el segundo y el tercer aspectos pueden comprender las mismas características y ventajas del primer aspecto.

II. Ejemplos de realización

30 La Figura 1a y la Figura 1b ilustran esquemáticamente una configuración de canales de 7.1+4 (correspondiente a una configuración de altavoz de 7.1+4) con tres canales delanteros L, C, R, dos canales de entorno LS, RS, dos canales traseros LB, RB, cuatro canales elevados TFL, TFR, TBL, TBR, y un canal de efectos de baja frecuencia, LFE. En el procedimiento de codificar la configuración de canales de 7.1+4, los canales son, por lo común, mezclados en sentido descendente (esto es, combinados) en un número más pequeño de señales, a las que se hace referencia como señales de mezcla descendente. En el procedimiento de mezcla en sentido descendente, los canales pueden ser combinados de diferentes maneras para formar diferentes configuraciones de mezcla descendente. La Figura 1a ilustra una primera configuración de mezcla descendente de 5.1 100a, con señales de mezcla descendente l, c, r, ls, rs, lfe. Los círculos de la figura indican qué canales son mezclados en sentido descendente para formar qué señales de mezcla descendente. La Figura 1b ilustra una segunda configuración de mezcla descendente de 5.1 100b, con señales de mezcla descendente l, c, r, tl, tr, lfe. La segunda configuración de mezcla descendente de 5.1 100b es diferente de la configuración de mezcla descendente 5.1 100a en que los canales se combinan de un modo diferente. Por ejemplo, en la primera configuración de mezcla descendente 100a, los canales L y TFL son mezclados en sentido descendente para formar la señal de mezcla descendente l, en tanto que, en la segunda configuración de mezcla descendente 100b, los canales L, LS, LB son mezclados en sentido descendente para formar la señal de mezcla descendente l. Se hace referencia, en ocasiones, a la configuración de mezcla descendente, en esta memoria, como un esquema de mezcla en sentido descendente que describe qué canales son mezclados en sentido descendente para formar qué señales de mezcla descendente. La configuración de mezcla descendente, o esquema de mezcla en sentido descendente, puede ser dinámica por cuanto puede variar ente tramas temporales de un sistema de codificación de audio. Por ejemplo, puede utilizarse el primer esquema de mezcla en sentido descendente 100a en algunas tramas temporales, en tanto que puede utilizarse el segundo esquema de mezcla en sentido descendente 100b en otras tramas temporales. En el caso de que el esquema de mezcla en sentido descendente varíe dinámicamente, el codificador puede enviar datos al descodificador que indican qué esquema de mezcla en sentido descendente se ha utilizado a la hora de codificar los canales.

45 La Figura 2 ilustra un descodificador 200 de la técnica anterior para la mejora del diálogo. El descodificador comprende tres componentes principales, un componente de recepción 202, un componente de mezcla en sentido ascendente, o reconstrucción, 204, y un componente de mejora de diálogo (DE –“dialog enhancement”–) 206. El descodificador 200 es del tipo que recibe una pluralidad de señales de mezcla descendente 212, reconstruye la configuración completa 218 de los canales basándose en las señales de mezcla descendente 212 recibidas, lleva a cabo una mejora del diálogo con respecto a la configuración de canales completa 218, o al menos un subconjunto de esta, y suministra

como salida una configuración completa de canales mejorados en diálogo 220.

Con más detalle, el componente de recepción 202 se ha configurado para recibir una corriente de datos 210 (a la que se hace referencia, en ocasiones, como corriente de bits) de un codificador. La corriente de datos 210 puede comprender diferentes tipos de datos, y el componente de recepción 202 puede descodificar la corriente de datos recibida 210 en diferentes tipos de datos. En este caso, la corriente de datos comprende una pluralidad de señales de mezcla descendente 212, parámetros de reconstrucción 214 y parámetros de mejora de diálogo 216.

El componente de mezcla en sentido ascendente 204 reconstruye, entonces, la configuración completa de canales basándose en la pluralidad de señales de mezcla descendente 212 y en los parámetros de reconstrucción 214. En otras palabras, el componente de mezcla en sentido ascendente 204 reconstruye todos los canales 218 que fueron mezclados en sentido descendente para obtener las señales de mezcla descendente 212. Por ejemplo, el componente de mezcla en sentido ascendente 204 puede reconstruir toda la configuración de canales paramétricamente, basándose en los parámetros de reconstrucción 214.

En el ejemplo ilustrado, las señales de mezcla descendente 212 corresponden a las señales de mezcla descendente de una de las configuraciones de mezcla descendente de 5.1 de las Figuras 1a y 1b, y los canales 218 corresponden a los canales de la configuración de canales de 7.1+4 de las Figuras 1a y 1b. Sin embargo, los principios del descodificador 200 se aplicarán, por supuesto, a otras configuraciones de canales / configuraciones de mezcla descendente.

Los canales 218 reconstruidos, o al menos un subconjunto de los canales 218 reconstruidos, son entonces sometidos a una mejora del diálogo por parte del componente de mejora de diálogo 206. Por ejemplo, el componente de mejora de diálogo 206 puede llevar a cabo una operación matricial sobre los canales 218 reconstruidos, o al menos sobre un subconjunto de los canales 218 reconstruidos, al objeto de suministrar como salida canales mejorados en diálogo. Tal operación matricial se define, por lo común, por los parámetros de mejora de diálogo 216.

A modo de ejemplo, el componente de mejora de diálogo 206 puede someter los canales C, L, R a una mejora del diálogo, a fin de proporcionar canales mejorados en diálogo C_{DE} , L_{DE} , R_{DE} , en tanto que los demás canales únicamente se hacen pasar a su través como se indica por las líneas discontinuas de la Figura 2. En tal situación, los parámetros de mejora de diálogo se han definido únicamente con respecto a los canales C, L, R, esto es, con respecto a un subconjunto de la pluralidad de canales 218. Por ejemplo, los parámetros de mejora de diálogo 216 pueden definir una matriz 3 x 3 que puede ser aplicada a los canales C, L, R.

$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ R_{DE} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} C \\ L \\ R \end{bmatrix}$$

Alternativamente, los canales no implicados en la mejora del diálogo pueden hacerse pasar a su través por medio de la matriz de mejora de diálogo con 1 en las posiciones diagonales correspondientes y 0 en todos los demás elementos de las filas y columnas correspondientes.

$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ TFL \\ R_{DE} \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & 0 & m_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{21} & m_{22} & 0 & m_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{31} & m_{32} & 0 & m_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} C \\ L \\ TFL \\ R \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix}$$

El componente de mejora de diálogo 206 puede llevar a cabo una mejora del diálogo con arreglo a diferentes modos. Un primer modo, al que se hace referencia en la presente memoria como mejora paramétrica independiente de canales, se ilustra en la Figura 3. La mejora del diálogo se lleva a cabo con respecto a al menos un subconjunto de los canales 218 reconstruidos, por lo común, los canales que comprenden diálogo, aquí, los canales L, R, C. Los

parámetros de mejora de diálogo 216 comprenden un conjunto de parámetros para cada uno de los canales que se han de mejorar. En el ejemplo ilustrado, los conjuntos de parámetros vienen dados por parámetros p_1 , p_2 , p_3 correspondientes a los canales L, R, C, respectivamente. En principio, los parámetros transmitidos en este modo representan la contribución relativa del diálogo a la energía de mezcla, para un mosaico o losa de tiempo-frecuencia de un canal. El factor de ganancia g puede ser expresado como:

$$g = 10^{\frac{G}{20}} - 1$$

donde G es una ganancia de mejora de diálogo, expresada en dB. La ganancia de mejora de diálogo G puede, por ejemplo, ser suministrada como entrada por un usuario, y no está, en consecuencia, incluida, por lo común, en la corriente de datos 210 de la Figura 2.

10 Cuando se está en el modo de mejora paramétrica independiente de canales, el componente de mejora de diálogo 206 multiplica cada canal por su parámetro p_i correspondiente y el factor de ganancia g , y añade entonces el resultado al canal con el fin de producir canales de mejora de diálogo 220, aquí, L_{DE} , R_{DE} , C_{DE} . Utilizando notación matricial, esto puede escribirse como:

$$X_e = (I + \text{diag}(p) \cdot g) \cdot X$$

15 donde X es una matriz que tiene por filas los canales 218 (L, R, C), X_e es una matriz que tiene por filas los canales mejorados en diálogo 220, p es un vector fila con entradas correspondientes a los parámetros de mejora de diálogo p_1 , p_2 , p_3 para cada canal, y $\text{diag}(p)$ es una matriz diagonal que tiene las entradas de p en la diagonal.

En la Figura 4 se ilustra un segundo modo de realización, al que se hace referencia en esta memoria como predicción de diálogo de múltiples canales. En este modo, el componente de mejora de diálogo 206 combina múltiples canales 218 en una combinación lineal con el fin de predecir una señal de diálogo 419. Aparte de la adición coherente de la presencia del diálogo en múltiples canales, esta solución puede aprovecharse de la sustracción del ruido de fondo en un canal que comprende diálogo utilizando otro canal sin diálogo. Para este propósito, los parámetros de mejora de diálogo 216 comprenden un parámetro para cada canal 218, que define el coeficiente del canal correspondiente a la hora de formar la combinación lineal. En el ejemplo ilustrado, los parámetros de mejora de diálogo 216 comprenden parámetros p_1 , p_2 , p_3 correspondientes a los canales L, R, C, respectivamente. Por lo común, pueden utilizarse algoritmos de optimización de error cuadrático medio mínimo (MMSE –“minimum mean square error”–) para generar los parámetros de predicción en el lado del codificador.

El componente de mejora de diálogo 206 puede entonces mejorar, es decir, producir ganancia en, la señal de diálogo predicha 419 por aplicación de un factor de ganancia g , y añadir la señal de diálogo mejorada a los canales 218, a fin de producir los canales mejorados en diálogo 220. Para añadir la señal de diálogo mejorada a los canales correctos en la posición espacial adecuada (de otro modo, ello no mejorará el diálogo con la ganancia esperada), el reparto entre los tres canales es transmitido por coeficientes de representación, aquí, r_1 , r_2 , r_3 . Con la restricción de que los coeficientes de representación son conservativos de la energía, esto es:

$$r_1^2 + r_2^2 + r_3^2 = 1$$

35 El tercer coeficiente de representación r_3 puede ser determinado a partir de los dos primeros coeficientes, tal como:

$$r_3 = \sqrt{1 - r_1^2 - r_2^2}.$$

Utilizando notación matricial, la mejora del diálogo llevada a cabo por el componente de mejora de diálogo 206 cuando se encuentra en el modo de predicción de diálogo de múltiples canales, puede escribirse como:

$$X_e = (I + g \cdot H \cdot P) \cdot X$$

40 o:

$$X_e = \begin{bmatrix} 1 + g \cdot r_1 \cdot p_1 & g \cdot r_1 \cdot p_2 & g \cdot r_1 \cdot p_3 \\ g \cdot r_2 \cdot p_1 & 1 + g \cdot r_2 \cdot p_2 & g \cdot r_2 \cdot p_3 \\ g \cdot r_3 \cdot p_1 & g \cdot r_3 \cdot p_2 & 1 + g \cdot r_3 \cdot p_3 \end{bmatrix} \cdot X$$

45 donde I es la matriz identidad, X es una matriz que tiene por filas los canales 218 (L, R, C), X_e es una matriz que tiene por filas los canales mejorados en diálogo 220, P es un vector fila con entradas correspondientes a los parámetros de mejora de diálogo p_1 , p_2 , p_3 para cada canal, H es un vector columna que tiene como entradas los coeficientes de representación r_1 , r_2 , r_3 , y g es el factor de ganancia, de manera que:

$$g = 10^{\frac{G}{20}} - 1.$$

De acuerdo con un tercer modo, al que se hace referencia en esta memoria como híbrido de forma de onda-paramétrico, el componente de mejora de diálogo 206 puede combinar cualquiera de entre el primer y el segundo

modos con la transmisión de una señal de audio adicional (una señal de forma de onda) que representa diálogo. Esta última es, por lo común, codificada a una baja velocidad de transferencia de bits, lo que provoca señales espurias bien audibles cuando se escuchan por separado. Dependiendo de las propiedades de señal de los canales 218 y del diálogo, así como de la velocidad de transferencia de bits asignada a la codificación de la señal de forma de onda de diálogo, el codificador también determina un parámetro de fusión, α_c , que indica el modo como han de dividirse las contribuciones de ganancia ente la contribución paramétrica (originaria del primer o del segundo modo) y la señal de audio adicional que representa diálogo.

En combinación con el segundo modo, la mejora del diálogo del tercer modo puede escribirse como:

$$X_e = H \cdot g_1 \cdot d_c + (I + H \cdot g_2 \cdot P) \cdot X$$

o:

$$X_e = \begin{bmatrix} 1 + g_2 \cdot r_1 \cdot p_1 & g_2 \cdot r_1 \cdot p_2 & g_2 \cdot r_1 \cdot p_3 & g_1 \cdot r_1 \\ g_2 \cdot r_2 \cdot p_1 & 1 + g_2 \cdot r_2 \cdot p_2 & g_2 \cdot r_2 \cdot p_3 & g_1 \cdot r_2 \\ g_2 \cdot r_3 \cdot p_1 & g_2 \cdot r_3 \cdot p_2 & 1 + g_2 \cdot r_3 \cdot p_3 & g_1 \cdot r_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ d_c \end{bmatrix}$$

donde d_c es la señal de audio adicional que representa diálogo, con:

$$g_1 = \alpha_c \cdot \left(10^{\frac{G}{20}} - 1 \right),$$

$$g_2 = (1 - \alpha_c) \cdot \left(10^{\frac{G}{20}} - 1 \right).$$

Para la combinación con la mejora independiente de canal (el primer modo), se recibe para cada canal 218 una señal de audio $d_{c,i}$ que representa diálogo. Escribiendo $D_c = \begin{pmatrix} d_{c,1} \\ d_{c,2} \\ d_{c,3} \end{pmatrix}$, la mejora del diálogo puede escribirse como:

$$X_e = g_1 \cdot D_c + (I + \text{diag}(p) \cdot g_2) \cdot X.$$

La Figura 5 ilustra un descodificador 500 de acuerdo con realizaciones proporcionadas a modo de ejemplo. El descodificador 500 es del tipo que descodifica una pluralidad de señales de mezcla descendente, que son una mezcla descendente de una pluralidad más grande de canales, para su reproducción subsiguiente. En otras palabras, el descodificador 500 es diferente del descodificador de la Figura 2 en que no está configurado para reconstruir la configuración de canales completa.

El descodificador 500 comprende un componente de recepción 502 y un bloque de mejora de diálogo 503 que comprende un componente de mezcla en sentido ascendente 504, un componente de mejora de diálogo 506 y un componente de mezcla 508.

Como se explica con referencia a la Figura 2, el componente de recepción 502 recibe una corriente de datos 510 y la descodifica en sus componentes, en este caso, una pluralidad de señales de mezcla descendente 512 que constituyen una mezcla descendente de una pluralidad más grande de canales (véanse las Figuras 1a y 1b), parámetros de reconstrucción 514 y parámetros para mejora de diálogo 516. En algunos casos, la corriente de datos 510 comprende, de manera adicional, datos indicativos de parámetros de mezcla 522. Por ejemplo, los parámetros de mezcla pueden formar parte de los parámetros para la mejora del diálogo. En otros casos, los parámetros de mezcla 522 se encuentran ya disponibles en el descodificador 500, por ejemplo, pueden estar codificados en hardware en el descodificador 500. En otros casos, se dispone de parámetros de mezcla 522 para múltiples conjuntos de parámetros de mezcla, y los datos de la corriente de datos 510 proporcionan una indicación de qué conjunto de entre estos múltiples conjuntos de parámetros de mezcla se utiliza.

Los parámetros para mejora de diálogo 516 se definen, por lo común, con respecto a un subconjunto de la pluralidad de canales. Los datos que identifican el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para la mejora del diálogo, pueden estar incluidos en la corriente de datos recibida 510, por ejemplo, como parte de los parámetros para mejora de diálogo 516. Alternativamente, el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo puede estar codificado en hardware en el descodificador 500. Por ejemplo, haciendo referencia a la Figura 1a, los parámetros para mejora de diálogo 516 pueden haberse definido con respecto a canales L, TFL que se han mezclado en sentido descendente para formar la señal de mezcla descendente l, el canal C que está comprimido en la señal de mezcla descendente c, y los canales R, TFR que se han mezclado en sentido descendente para formar la señal de mezcla descendente r. Para propósitos de ilustración, se supone que el diálogo está únicamente presente los canales L, C y R. Ha de apreciarse que los parámetros para mejora de diálogo 516 pueden haberse definido con respecto a canales que comprenden diálogo, tales como los canales L, C, R, pero también pueden haberse definido con respecto a canales que no comprenden diálogo, tales como los canales TFL, TFR, en este ejemplo. De esta manera, el ruido de fondo de un canal que

comprende diálogo puede, por ejemplo, sustraerse utilizando otro canal sin diálogo.

El subconjunto de canales con respecto a los cuales se definen los parámetros para mejora de diálogo 516, se mezcla en sentido descendente hasta formar un subconjunto 512a de la pluralidad de señales de mezcla descendente 512. En el ejemplo ilustrado, el subconjunto 512a de señales de mezcla descendente comprende las señales de mezcla descendente c, l y r. Este subconjunto de señales de mezcla descendente 512a se suministra como entrada al bloque de mejora de diálogo 503. El subconjunto relevante 512a de señales de mezcla descendente puede, por ejemplo, encontrarse basándose en el conocimiento del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo y el esquema de mezcla en sentido descendente.

El componente de mezcla en sentido ascendente 514 se sirve de técnicas paramétricas como es sabido en la técnica, para la reconstrucción de canales que son mezclados en sentido descendente hasta formar el subconjunto de señales de mezcla descendente 512a. La reconstrucción está basada en los parámetros de reconstrucción 514. En particular, el componente de mezcla en sentido ascendente 504 reconstruye el subconjunto de la pluralidad de canales con respecto a los cuales se han definido los parámetros para mejora de diálogo 516. En algunas realizaciones, el componente de mezcla en sentido ascendente 504 reconstruye únicamente el subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo 516. Tales realizaciones proporcionadas a modo de ejemplo se describirán con referencia a la Figura 7. En otras realizaciones, el componente de mezcla en sentido ascendente 504 reconstruye al menos un canal, además del subconjunto de la pluralidad con respecto al cual se han definido los parámetros para mejora de diálogo 516. Tales realizaciones proporcionadas a modo de ejemplo se describirán con referencia a la Figura 6.

Los parámetros de reconstrucción no solo pueden ser variables en el tiempo, sino que también pueden ser dependientes de la frecuencia. Por ejemplo, los parámetros de reconstrucción pueden adoptar diferentes valores para diferentes bandas de frecuencias. Esto generalmente favorecerá la calidad de los canales reconstruidos.

Como es sabido en la técnica, la mezcla paramétrica en sentido ascendente puede, generalmente, incluir la formación de señales descorrelacionadas a partir de las señales de entrada que son sometidas a la mezcla en sentido ascendente, y la reconstrucción de señales paramétricamente basándose en las señales de entrada y en las señales descorrelacionadas. Véase, por ejemplo, el libro "Spatial Audio Processing: MPEG Surround and Other Applications" (Tratamiento de audio espacial: entorno de MPEG y otras aplicaciones), por Jeroen Breebaart y Christof Faller, ISBN: 978-9-470-03350-0. Sin embargo, el componente de mezcla en sentido ascendente 504 lleva, preferiblemente, a cabo mezcla paramétrica en sentido ascendente sin utilizar ninguna de tales señales descorrelacionadas. Las ventajas obtenidas por el uso de señales descorrelacionadas se ven, en este caso, reducidas por la subsiguiente mezcla en sentido descendente llevada a cabo en el componente de mezcla 508. Por lo tanto, el uso de señales descorrelacionadas puede, ventajosamente, ser omitido por el componente de mezcla en sentido ascendente 504, con lo que se ahorra complejidad de computación. De hecho, el uso de señales descorrelacionadas en la mezcla en sentido ascendente, en combinación con la mejora del diálogo, tendrá como resultado una peor calidad, puesto que podría dar lugar a una reverberación descorrelacionadora en el diálogo.

El componente de mejora de diálogo 506 aplica, entonces, la mejora del diálogo al subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo 516, a fin de producir al menos una señal mejorada en diálogo. En algunas realizaciones, la señal mejorada en diálogo corresponde a versiones mejoradas en diálogo del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo 516. Esto se explicará con mayor detalle más adelante, con referencia a la Figura 6. En otras realizaciones, la señal mejorada en diálogo corresponde a un componente de diálogo predicho y mejorado del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo 516. Esto se explicará con mayor detalle más adelante, con referencia a la Figura 7.

Similarmente a los parámetros de reconstrucción, los parámetros para mejora de diálogo pueden variar en el tiempo, así como con la frecuencia. Con mayor detalle, los parámetros para mejora de diálogo pueden adoptar diferentes valores para diferentes bandas de frecuencias. El conjunto de bandas de frecuencias con respecto al cual se han definido los parámetros de reconstrucción puede diferir del conjunto de bandas de frecuencias con respecto al cual se han definido los parámetros de mejora de diálogo.

El componente de mezcla 508 lleva a cabo, entonces, una mezcla basándose en la al menos una señal de mejora de diálogo, a fin de proporcionar versiones mejoradas en diálogo 520 del subconjunto 512a de señales de mezcla descendente. En el ejemplo ilustrado, las versiones mejoradas en diálogo 520 del subconjunto 512a de señales de mezcla descendente vienen dadas por c_{DE} , l_{DE} , r_{DE} , que corresponden a señales de mezcla descendente c, l, r, respectivamente.

La mezcla en sentido descendente puede realizarse de acuerdo con parámetros de mezcla 522 que describen una contribución de la al menos una señal mejorada en diálogo a las versiones mejoradas en diálogo 520 del subconjunto de señales de mezcla descendente 512a. En algunas realizaciones, véase la Figura 6, la al menos una señal mejorada en diálogo se mezcla conjuntamente con canales que han sido reconstruidos por el componente de mezcla en sentido ascendente 504. En tales casos, los parámetros de mezcla 522 pueden corresponderse a un esquema de mezcla en sentido descendente, véanse las Figuras 1a y 1b, que describe en cuáles de las señales de mezcla descendente

mejoradas en diálogo 520 ha de mezclarse cada canal. En otras realizaciones, véase la Figura 7, la al menos una señal mejorada en diálogo se mezcla conjuntamente con el subconjunto 512a de señales de mezcla descendente. En tal caso, los parámetros de mezcla 522 pueden corresponder a factores de ponderación que describen el modo como la al menos una señal mejorada en diálogo debe ser ponderada dentro del subconjunto 512a de señales de mezcla descendente.

La operación de mezcla en sentido ascendente llevada a cabo por el componente de mezcla en sentido ascendente 504, la operación de mejora de diálogo llevada a cabo con el componente de mejora de diálogo 506, y la operación de mezcla llevada a cabo por el componente de mezcla 508 son, por lo común, operaciones lineales cada una de las cuales puede definirse por una operación matricial, esto es, por un producto de matriz-vector. Esto es, al menos, verdadero si se omiten las señales descorrelacionadoras en la operación de mezcla en sentido ascendente. En particular, la matriz asociada con la operación de mezcla en sentido ascendente (U) se define por / puede ser deducida de los parámetros de reconstrucción 514. A este respecto, ha de apreciarse que sigue siendo posible el uso de señales descorrelacionadoras en la operación de mezcla en sentido ascendente, pero que la creación de las señales descorrelacionadas no forma entonces parte de la operación matricial para la mezcla en sentido ascendente. La operación de mezcla en sentido ascendente con descorrelacionadores puede verse como una aproximación en dos estadios. En un primer estadio, las señales de mezcla descendente suministradas como entrada son aportadas a una matriz de descorrelación preliminar, y las señales suministradas como salida tras la aplicación de la matriz de descorrelación preliminar son, cada una de ellas, aportadas a un descorrelacionador. En un segundo estadio, las señales de mezcla descendente suministradas como entrada y las señales suministradas como salida desde los descorrelacionadores son aportadas a la matriz de mezcla en sentido ascendente, en la que los coeficientes de la matriz de mezcla en sentido ascendente correspondientes a las señales de mezcla descendente suministradas como entrada forman lo que se denomina la «matriz de mezcla en sentido ascendente seca», y los coeficientes correspondientes a las señales suministradas como salida desde los descorrelacionadores forman lo que se denomina la «matriz de mezcla en sentido ascendente mojada». Cada matriz subordinada, o submatriz, guarda una relación de correspondencia con la configuración de canales de mezcla en sentido ascendente. Cuando no se utilizan las señales descorrelacionadoras, la matriz asociada con la operación de mezcla en sentido ascendente está configurada para operar únicamente sobre las señales de entrada 512a, y las columnas relacionadas con las señales descorrelacionadas (la matriz de mezcla en sentido ascendente mojada) no están incluidas en la matriz. En otras palabras, la matriz de mezcla en sentido ascendente corresponde, en este caso, a la matriz de mezcla en sentido ascendente seca. Sin embargo, como se ha destacado anteriormente, el uso de señales descorrelacionadoras tendrá, por lo común, como resultado, en este caso, una peor calidad.

La matriz asociada con la operación de mejora de diálogo (M) se define por / puede deducirse de los parámetros para mejora de diálogo 516, y la matriz asociada con la operación de mezcla (C) se define por / puede deducirse de, los parámetros de mezcla 522.

Puesto que la operación de mezcla en sentido ascendente, la operación de mejora del diálogo y la operación de mezcla son, todas ellas, operaciones lineales, las matrices correspondientes pueden ser combinadas, por multiplicación de matrices, en una única matriz E (entonces, $X_{DE} = E \cdot X$, con $E = C \cdot M \cdot U$). Aquí, X es un vector columna de las señales de mezcla descendente 512a, y X_{DE} es un vector columna de las señales de mezcla descendente mejoradas en diálogo 520. De esta forma, el bloque completo de mejora de diálogo 503 puede corresponder a una única operación matricial que se aplica al subconjunto 512a de señales de mezcla descendente con el fin de producir las versiones mejoradas en diálogo 520 del subconjunto 512a de señales de mezcla descendente. De acuerdo con esto, los métodos que se describen en esta memoria pueden implementarse de una manera muy eficiente.

La Figura 6 ilustra un descodificador 600 que corresponde a una realización proporcionada a modo de ejemplo del descodificador 500 de la Figura 5. El descodificador 600 comprende un componente de recepción 602, un componente de mezcla en sentido ascendente 604, un componente de mejora de diálogo 606 y un componente de mezcla 608.

Similarmente al descodificador 500 de la Figura 5, el componente de recepción 602 recibe una corriente de datos 610 y la descodifica en una pluralidad de señales de mezcla descendente 612, parámetros de reconstrucción 614 y parámetros para la mejora del diálogo 616.

El componente de mezcla en sentido ascendente 604 recibe un subconjunto 612a (correspondiente al subconjunto 512a) de la pluralidad de señales de mezcla descendente 612. Para cada una de las señales de mezcla descendente del subconjunto 612a, el componente de mezcla en sentido ascendente 604 reconstruye todos los canales que han sido mezclados en sentido descendente en la señal de mezcla descendente ($X_u = U \cdot X$). Esto incluye canales 618a con respecto a los cuales se han definido los parámetros para mejora de diálogo, así como canales 618b que no han de estar implicados en la mejora del diálogo. Haciendo referencia a la Figura 1b, los canales 618a con respecto a los cuales se han definido los parámetros para la mejora del diálogo, pueden, por ejemplo, corresponder a los canales L, LS, C, R, RS, y los canales 618b que no han de implicarse en la mejora del diálogo pueden corresponder a los canales LR, RB.

Los canales 618a con respecto a los cuales se han definido los parámetros para la mejora del diálogo (X'_u) son entonces sometidos a una mejora del diálogo por el componente de mejora de diálogo 606 ($X_e = M \cdot X'_u$), en tanto que los canales 618b que no se han de implicar en la mejora del diálogo (X''_u) se encuentran puenteando el componente

de mejora de diálogo 606.

El componente de mejora de diálogo 606 puede aplicar cualquiera del primer, segundo y tercer modos de mejora de diálogo descritos en lo anterior. En caso de que se aplique el tercer modo, la corriente de datos 610 puede comprender, como se ha explicado anteriormente, una señal de audio que representa diálogo (esto es, una forma de onda que representa diálogo), destinada a ser aplicada en la mejora del diálogo conjuntamente con el subconjunto 618a de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo $(X_e = M \cdot \begin{bmatrix} X_u' \\ D_c \end{bmatrix})$.

Como resultado de ello, el componente de mejora de diálogo 606 suministra como salida señales mejoradas en diálogo 619 que, en este caso, corresponden a versiones mejoradas en diálogo del subconjunto 618a de canales con respecto al cual se han definido los parámetros para la mejora del diálogo. A modo de ejemplo, las señales de mejora de diálogo 619 pueden corresponder a versiones mejoradas en diálogo de los canales L, LS, C, R, RS de la Figura 1b.

El componente de mezcla 608 mezcla entonces las señales mejoradas en diálogo 619 conjuntamente con los canales 618b que no están implicados en la mejora del diálogo $(X_{DE} = C \cdot \begin{bmatrix} X_e \\ X_u'' \end{bmatrix})$, con el fin de producir versiones mejoradas en diálogo 620 del subconjunto 612a de señales de mezcla descendente. El componente de mezcla 608 efectúa la mezcla de acuerdo con el esquema de mezcla en sentido descendente vigente, tal como el esquema de mezcla en sentido descendente ilustrado en la Figura 1b. En este caso, los parámetros de mezcla 622 corresponden, de este modo, a un esquema de mezcla en sentido descendente que describe en qué señal de mezcla descendente 620 ha de ser mezclado cada canal 619, 618b. El esquema de mezcla en sentido descendente puede ser estático y, por tanto, conocido por el descodificador 600, lo que significa que siempre se aplica el mismo esquema de mezcla en sentido descendente, o bien el esquema de mezcla en sentido descendente puede ser dinámico, lo que significa que puede variar de una trama a otra, o puede ser uno de diversos esquemas conocidos en el descodificador. En este último caso, se incluye en la corriente de datos 610 una indicación referente al esquema de mezcla en sentido descendente.

En la Figura 6, el descodificador está equipado con un componente de remodelación opcional 630. El componente de remodelación 630 puede ser utilizado para la conversión entre diferentes esquemas de mezcla en sentido descendente, por ejemplo, para la conversión del esquema 100b al esquema 100a. Se aprecia que el componente de remodelación 630 deja, por lo común, sin cambios las señales c y lfe, es decir, actúa como un componente de paso a su través, o paso libre, con respecto a estas señales. El componente de remodelación 630 puede recibir y operar (no se muestra) basándose en varios parámetros tales como, por ejemplo, los parámetros de reconstrucción 614 y los parámetros para mejora de diálogo 616.

La Figura 7 ilustra un descodificador 700 que corresponde a una realización proporcionada a modo de ejemplo del descodificador 500 de la Figura 5. El descodificador 700 comprende un componente de recepción 702, un componente de mezcla en sentido ascendente 704, un componente de mejora de diálogo 706, y un componente de mezcla 708.

Similarmente al descodificador 500 de la Figura 5, el componente de recepción 702 recibe una corriente de datos 710 y la descodifica en una pluralidad de señales de mezcla descendente 712, parámetros de reconstrucción 714 y parámetros para mejora de diálogo 716.

El componente de mezcla en sentido ascendente 704 recibe un subconjunto 712a (correspondiente al subconjunto 512a) de la pluralidad de señales de mezcla descendente 712. En contraste con la realización descrita con respecto a la Figura 6, el componente de mezcla en sentido ascendente 704 únicamente reconstruye el subconjunto 718a de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo 716 $(X_u' = U' \cdot X)$. Haciendo referencia a la Figura 1b, los canales 718a con respecto a los cuales se han definido los parámetros para la mejora del diálogo pueden corresponder, por ejemplo, a los canales C, L, LS, R, RS.

El componente de mejora de diálogo 706 lleva entonces a cabo una mejora del diálogo en los canales 718a con respecto a los cuales se han definido los parámetros para la mejora del diálogo $(X_d = M_d \cdot X_u')$. En este caso, el componente de mejora de diálogo 706 prosigue prediciendo un componente de diálogo basándose en los canales 718a, mediante la formación de una combinación lineal de los canales 718a, de acuerdo con un segundo modo de mejora de diálogo. Los coeficientes utilizados a la hora de formar la combinación lineal, denotados por p_1 a p_5 en la Figura 7, están incluidos en los parámetros para mejora de diálogo 716. El componente de diálogo predicho es entonces mejorado por la multiplicación de un factor de ganancia g, con el fin de producir una señal mejorada en diálogo 719. El factor de ganancia g puede expresarse como:

$$g = 10^{\frac{G}{20}} - 1$$

donde G es una ganancia de mejora de diálogo expresada en dB. La ganancia de mejora de diálogo G puede, por ejemplo, ser suministrada como entrada por un usuario y no está, por tanto, por lo común, incluida en la corriente de datos 710. Ga de apreciarse que, en caso de que haya varias componentes de diálogo, el procedimiento de predicción y mejora anterior puede ser aplicado una sola vez por cada componente de diálogo.

La señal mejorada en diálogo predicha 719 (es decir, las componentes de diálogo predichas y mejoradas) es entonces

mezclada para formar el subconjunto 712a de señales de mezcla descendente con el fin de producir versiones mejoradas en diálogo 720 del subconjunto 712a de señales de mezcla descendente ($X_{DE} = C \cdot \begin{bmatrix} X_d \\ X \end{bmatrix}$). La mezcla se realiza de acuerdo con parámetros de mezcla 722 que describen una contribución de la señal mejorada en diálogo 719 a las versiones mejoradas en diálogo 720 del subconjunto de señales de mezcla descendente. Los parámetros de mezcla son, por lo común, incluidos en la corriente de datos 710. En este caso, los parámetros de mezcla 722 corresponden a factores de ponderación r_1, r_2, r_3 que describen el modo como la al menos una señal mejorada en diálogo 719 ha de ser ponderada dentro del subconjunto 712a de señales de mezcla descendente:

$$X_{DE} = x + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \cdot X_d = \begin{bmatrix} r_1 & 1 & 0 & 0 \\ r_2 & 0 & 1 & 0 \\ r_3 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_d \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Con mayor detalle, los factores de ponderación pueden corresponder a coeficientes de representación que describen el reparto de la al menos una señal mejorada en diálogo 719 con respecto al subconjunto 712a de señales de mezcla descendente, de tal manera que la señal mejorada en diálogo 719 es añadida a las señales de mezcla descendente 712a en las posiciones espaciales correctas.

Los coeficientes de representación (los parámetros de mezcla 722) de la corriente de datos 710 pueden corresponder a los canales mezclados en sentido ascendente 718a. En el ejemplo ilustrado, existen cinco canales mezclados en sentido ascendente 718a y puede haber, por tanto, póngase por caso, cinco coeficientes de representación correspondientes rc_1, rc_2, \dots, rc_5 . Los valores de r_1, r_2, r_3 (que corresponden a las señales de mezcla descendente 712a) pueden ser entonces calculados a partir de rc_1, rc_2, \dots, rc_5 , en combinación con el esquema de mezcla en sentido descendente. Cuando una multiplicidad de los canales 718a corresponden a la misma señal de mezcla descendente 712a, los coeficientes de representación de diálogo pueden ser sumados. Por ejemplo, en el ejemplo ilustrado, se tiene que $r_1 = rc_1, r_2 = rc_2 + rc_3$, y $r_3 = rc_4 + rc_5$. Esta puede ser también una suma ponderada en caso de que la mezcla en sentido descendente de los canales se realizase utilizando coeficientes de mezcla en sentido descendente.

Ha de apreciarse que, también en este caso, el componente de mejora de diálogo 706 puede hacer uso de una señal de audio adicionalmente recibida que representa diálogo. En tal caso, la señal mejorada en diálogo predicha 719 puede ser ponderada conjuntamente con la señal de audio que representa diálogo, antes de ser suministrada como entrada al componente de mezcla 708 ($X_d = (1 - \alpha_c) \cdot M_d \cdot X'_u + \alpha_c \cdot g \cdot D_c$). La ponderación apropiada viene dada por un parámetro de fusión α_c incluido en los parámetros para mejora de diálogo 716. El parámetro de fusión α_c indica el modo como deben dividirse las contribuciones a la ganancia entre la componente de diálogo predicha 719 (según se ha descrito en lo anterior) y la señal de audio adicional que representa diálogo D_c . Esto es análogo a lo que se ha descrito con respecto al tercer modo de mejora del diálogo, cuando se combina con el segundo modo de mejora del diálogo.

En la Figura 7, el decodificador está equipado con un componente de remodelación 730. El componente de remodelación 730 puede ser utilizado para la conversión entre diferentes esquemas de mezcla en sentido descendente, por ejemplo, para la conversión del esquema 100b al esquema 100a. Se destaca que el componente de remodelación 730 deja, por lo común, las señales c y lfe sin cambios, es decir, actúa como un componente de paso a su través, o paso libre, con respecto a estas señales. El componente de remodelación 730 puede recibir y operar (no se muestra) basándose en diversos parámetros tales como, por ejemplo, los parámetros de reconstrucción 714 y los parámetros para mejora de diálogo 716.

Lo anterior se ha explicado principalmente con respecto a una configuración de canales de 7.1+4 y una mezcla en sentido descendente de 5.1. Debe entenderse, sin embargo, que los principios de los decodificadores y de los métodos de decodificación que se describen en esta memoria se aplican igualmente bien a otras configuraciones de canales y de mezcla en sentido descendente.

La Figura 8 es una ilustración de un codificador 800 que puede ser utilizado para codificar una pluralidad de canales 818, algunos de los cuales incluyen diálogo, al objeto de producir una corriente de datos 810 para su transmisión al decodificador. El codificador 800 puede ser utilizado con cualquiera de los decodificadores 200, 500, 600, 700. El codificador 800 comprende un componente de mezcla en sentido descendente 805, un componente de codificación para mejora de diálogo 806, un componente de codificación paramétrica 804 y un componente de transmisión 802.

El codificador 800 recibe una pluralidad de canales 818, por ejemplo, los de las configuraciones de canales 100a, 100b representadas en las Figuras 1a y 1b.

El componente de mezcla en sentido descendente 805 mezcla en sentido descendente la pluralidad de canales 818 para obtener una pluralidad de señales de mezcla descendente 812, las cuales son entonces aportadas al componente de transmisión 802 para su inclusión en la corriente de datos 810. La pluralidad de canales 818 puede, por ejemplo, ser mezclada en sentido descendente de acuerdo con un esquema de mezcla en sentido descendente, tal como el ilustrado en la Figura 1a o en la Figura 1b.

La pluralidad de canales 818 y las señales de mezcla descendente 812 son suministradas como entrada al componente de codificación paramétrica 804. Basándose en sus señales suministradas como entrada, el componente de codificación paramétrica 804 calcula los parámetros de reconstrucción 814 que permiten la reconstrucción de los canales 818 a partir de las señales de mezcla descendente 812. Los parámetros de reconstrucción 814 pueden ser calculados, por ejemplo, utilizando algoritmos de optimización de error cuadrático medio mínimo (MMSE –“minimum mean square error”–), como es conocido en la técnica. Los parámetros de reconstrucción 814 son entonces aportados al componente de transmisión 802 para su inclusión en la corriente de datos 810.

El componente de codificación para mejora de diálogo 806 calcula parámetros para mejora de diálogo 816 basándose en uno o más de la pluralidad de canales 818 y en una o más señales de diálogo 813. Las señales de diálogo 813 representan diálogo puro. Es de destacar que el diálogo ya está mezclado dentro de uno o más de los canales 818. En los canales 818 puede haber, por tanto, uno o más componentes de diálogo que se corresponden con las señales de diálogo 813. Por lo común, el componente de codificación para mejora de diálogo 806 calcula los parámetros para mejora de diálogo 816 utilizando algoritmos de optimización de error cuadrático medio mínimo (MMSE). Tales algoritmos pueden proporcionar parámetros que hacen posible la predicción de las señales de diálogo 813 a partir de algunos de la pluralidad de canales 818. Los parámetros para mejora de diálogo 816 pueden, de esta forma, haberse definido con respecto a un subconjunto de la pluralidad de canales 818, a saber, aquellos a partir de los cuales pueden predecirse las señales de diálogo 813. Los parámetros para predicción de diálogo 816 son aportados al componente de transmisión 802 para su inclusión en la corriente de datos 810.

En conclusión, la corriente de datos 810, en consecuencia, comprende al menos la pluralidad de señales de mezcla descendente 812, los parámetros de reconstrucción 814 y los parámetros para mejora de diálogo 816.

Durante el funcionamiento normal del descodificador, valores de los parámetros de diferentes tipos (tales como los parámetros para la mejora del diálogo, o los parámetros de reconstrucción) son recibidos repetidamente por el descodificador con ciertas velocidades de transmisión. Si las velocidades de transmisión con las que se reciben los diferentes valores de parámetros son más bajas que la velocidad de transmisión con la que se ha de calcular la salida del descodificador, puede ser necesario interpolar los valores de los parámetros. Si el valor de un parámetro genérico p es conocido en los instantes t_1 y t_2 de tiempo, de manera que son $p(t_1)$ y $p(t_2)$, respectivamente, el valor $p(t)$ del parámetro en un tiempo intermedio $t_1 \leq t < t_2$ puede ser calculado utilizando diferentes esquemas de interpolación. Un ejemplo de tal esquema, al que se hace referencia en esta memoria como configuración de interpolación lineal, puede calcular el valor intermedio utilizando interpolación lineal, por ejemplo, $p(t) = p(t_1) + [p(t_2) - p(t_1)] (t - t_1) / (t_2 - t_1)$. Otra configuración, a la que se hace aquí referencia como configuración de interpolación constante por tramos, puede, en lugar de ello, incluir la conservación de un valor de parámetro fijo en uno de los valores conocidos durante todo el intervalo temporal, por ejemplo, $p(t) = p(t_1)$ o $p(t) = p(t_2)$, o una combinación de los valores conocidos, tal como, por ejemplo, el valor medio $p(t) = [p(t_1) + p(t_2)] / 2$. Puede haberse incorporado en el descodificador información acerca de qué esquema de interpolación debe utilizarse para un cierto tipo de parámetro durante un cierto intervalo de tiempo, o bien esta puede proporcionarse al descodificador de diferentes maneras, tal como de manera conjunta con los propios parámetros o como información adicional contenida en la señal recibida.

En un ejemplo ilustrativo, un descodificador recibe valores de parámetro para un primer y un segundo tipos de parámetros. Los valores recibidos de cada tipo de parámetro son exactamente aplicables en un primer ($T1 = \{t11, t12, t13, \dots\}$) y en un segundo ($T2 = \{t21, t22, t23, \dots\}$) conjuntos de instantes de tiempo, respectivamente, y el descodificador también tiene acceso a información acerca de cómo se han de interpolar los valores de cada tipo de parámetro en el caso de que sea necesario estimar un valor en un instante de tiempo que no está presente en el correspondiente conjunto. Los valores de parámetro controlan propiedades cuantitativas de las operaciones matemáticas sobre las señales, operaciones que, por ejemplo, pueden ser representadas como matrices. En el ejemplo que sigue, se supone que la operación controlada por el primer tipo de parámetro viene representada por una primera matriz A, que la operación controlada por el segundo tipo de parámetro viene representada por una segunda matriz B, y que los términos «operación» y «matriz» pueden utilizarse de forma intercambiable en el ejemplo. En un instante de tiempo en que es necesario calcular un valor de salida desde el descodificador, ha de computarse una operación de tratamiento compuesta correspondiente a la composición de ambas operaciones. Si se supone, adicionalmente, que la matriz A es la operación de mezcla en sentido ascendente (controlada por los parámetros de reconstrucción) y que la matriz B es la operación de aplicar una mejora del diálogo (controlada por los parámetros para la mejora del diálogo), entonces, en consecuencia, la operación de tratamiento compuesta de mezclar en sentido ascendente seguido de una mejora del diálogo se representa por el producto matricial BA.

Métodos para computar la operación de tratamiento compuesta se ilustran en las Figuras 9a-9e, en las que el tiempo discurre a lo largo del eje horizontal y unas marcas de verificación temporal indican los instantes de tiempo en que se ha de computar una operación de tratamiento compuesta (instantes de tiempo de salida). En las figuras, los triángulos corresponden a la matriz A (representando la operación de mezcla en sentido ascendente), los círculos, a la matriz B (representando la operación de aplicar la mejora del diálogo), y los cuadrados, a la matriz de operación compuesta BA (representando la operación compuesta de mezclar en sentido ascendente seguido de la mejora del diálogo). Los triángulos y los círculos llenos indican que la matriz respectiva se conoce exactamente (esto es, que los parámetros que controlan la operación que representa la matriz se conocen exactamente) en el instante de tiempo correspondiente, mientras que los triángulos y los círculos vacíos indican que el valor de la matriz respectiva es

predicho o interpolado (utilizando, por ejemplo, cualquiera de las configuraciones de interpolación esbozadas anteriormente). Un cuadrado lleno indica que la matriz de operación compuesta BA ha sido computada, en el instante de tiempo correspondiente, por ejemplo, por una matriz producto de las matrices A y B, y un cuadrado vacío indica que el valor de BA ha sido interpolado desde un instante de tiempo anterior. Por otra parte, las flechas de trazo de puntos indican entre qué instantes de tiempo se ha llevado a cabo una interpolación. Por último, una línea continua horizontal que une instantes de tiempo indica que el valor de una matriz se supone constante por tramos en ese intervalo.

En la Figura 9a se ha ilustrado un método para computar una operación de tratamiento compuesta BA que no hace uso de la presente invención. Los valores recibidos para las operaciones A y B se aplican exactamente en los instantes de tiempo t_{11} , t_{21} y t_{12} , t_{22} , respectivamente, y, para computar la matriz de operación de tratamiento compuesta en cada instante de tiempo de salida, el método interpola cada matriz individualmente. A fin de completar cada etapa de avance en el tiempo, la matriz que representa la operación de tratamiento compuesta es computada como el producto de los valores predichos A y B. Aquí, se supone que cada matriz se ha de interpolar utilizando una configuración de interpolación lineal. Si la matriz A tiene N' filas y N columnas, y la matriz B tiene M filas y N' columnas, cada etapa de avance en el tiempo requerirá $O(MN'N)$ operaciones de multiplicación por cada banda de parámetros (a fin de llevar a cabo la multiplicación requerida para computar la matriz de tratamiento compuesta BA). Una alta densidad de instantes de tiempo de salida y/o un elevado número de bandas de parámetros conllevará, por tanto, el riesgo (debido a la relativamente elevada complejidad computacional de una operación de multiplicación, en comparación con una operación de adición) de requerir una demanda elevada de los recursos computacionales. A fin de reducir la complejidad computacional, puede utilizarse el método alternativo ilustrado en la Figura 9b. Al computar la operación de tratamiento compuesta (por ejemplo, llevando a cabo una multiplicación de matrices) únicamente en los instantes de tiempo en que los valores de parámetro cambian (es decir, cuando son exactamente aplicables valores recibidos, en t_{11} , t_{21} y t_{12} , t_{22}), la matriz de operación de tratamiento compuesta BA puede ser interpolada directamente, en lugar de interpolar las matrices A y B por separado. Al hacerlo así, si las operaciones se representan por matrices, cada etapa de avance en el tiempo (entre los instantes de tiempo en que los valores de parámetro exactos cambian) requerirá, entonces, únicamente $O(NM)$ operaciones (para la suma de matrices) por cada banda de parámetros, y la reducida complejidad computacional requerirá una menor demanda de los recursos computacionales. También, si las matrices A y B son tales, que $N' > N \times M / (N + M)$, la matriz que representa la operación de tratamiento compuesta BA tendrá un menor número de elementos que los que se encuentran en las matrices individuales A y B, combinadas. El método de interpolar la matriz BA directamente requerirá, sin embargo, que tanto A como B sean conocidas en los mismos instantes de tiempo. Cuando los instantes de tiempo para los que A está definida (al menos parcialmente) son distintos de los instantes de tiempo para los que B está definida, se requiere un método de interpolación mejorado. Tal método mejorado, de acuerdo con realizaciones proporcionadas a modo de ejemplo de la presente invención, se ilustra en las Figuras 9c-9e. En asociación con la exposición de las Figuras 9a-9e, se supone, por simplicidad, que la matriz de operación de tratamiento compuesta BA es computada como producto de las matrices individuales. A y B, cada una de las cuales se ha generado basándose en valores de parámetro (recibidos o predichos / interpolados). En otras situaciones, puede resultar igualmente o más ventajoso computar la operación representada por la matriz BA directamente a partir de los valores de parámetro, sin pasar por una representación en forma de dos factores matriciales. En combinación con cualquiera de las técnicas ilustradas con referencia a las Figuras 9c-9e, cada una de estas soluciones cae dentro del alcance de la presente invención.

En la Figura 9c se ilustra una situación en la que el conjunto T1 de instantes de tiempo para el parámetro correspondiente a la matriz A incluye un valor temporal t_{12} que no está presente en el conjunto T2 (instantes de tiempo para el parámetro correspondiente a la matriz B). Ambas matrices se han de interpolar utilizando una configuración de interpolación lineal, y el método identifica el instante de predicción $t_p = t_{12}$ en que se ha de predecir el valor de la matriz B (utilizando, por ejemplo, interpolación). Una vez que se ha encontrado el valor, el valor de la matriz de operación de tratamiento compuesta BA en t_p puede ser computado multiplicando A y B. Para continuar, el método computa el valor de BA en un instante de tiempo adyacente $t_a = t_{11}$ y, seguidamente, interpola BA entre t_a y t_p . El método puede también computar, si se desea, el valor de BA en otro instante de tiempo adyacente $t_a = t_{13}$, e interpolar BA de t_p a t_a . Incluso aunque se requiera una multiplicación matricial adicional (en $t_p = t_{12}$), el método permite interpolar la matriz de operación de tratamiento compuesta BA directamente, con lo que sigue reduciendo la complejidad computacional en comparación con el método de, por ejemplo, la Figura 9a. Como se ha dicho anteriormente, la operación de tratamiento compuesta puede, alternativamente, ser computada directamente a partir de los valores de parámetro (recibidos o predichos / interpolados), en lugar de como un producto explícito de dos matrices que, a su vez, dependen de los valores de parámetro respectivos.

En el caso previo, únicamente el tipo de parámetro correspondiente a tenía instantes de tiempo que no estaban incluidos entre los instantes del tipo de parámetro correspondiente a B. En la Figura 9d se ilustra una situación diferente en la que el instante de tiempo t_{12} falta en el conjunto T2, y en la que el instante de tiempo t_{22} falta en el conjunto T1. Si se ha de computar un valor de BA en un instante de tiempo intermedio t' entre t_{12} y t_{22} , el método puede predecir tanto el valor de B en $t_p = t_{12}$ como el valor de A en $t_a = t_{22}$. Una vez computada la matriz de operación de tratamiento compuesta BA en ambos tiempos, BA puede ser interpolada para encontrar su valor en t' . En general, el método únicamente lleva a cabo multiplicaciones de matrices en los instantes de tiempo en que cambian valores de parámetro (es decir, en los instantes de tiempo de los conjuntos T1 y T2 en los que los valores recibidos son aplicables de forma exacta). Entre medias, la interpolación de la operación de tratamiento compuesta tan solo requiere adiciones de

matrices que tienen menos complejidad computacional que sus contrapartidas de multiplicación.

En los ejemplos anteriores, se ha supuesto que todos los parámetros de interpolación son lineales. En la Figura 9e se ilustra un método para la interpolación también cuando los parámetros han de ser interpolados inicialmente utilizando diferentes esquemas. En la figura, los valores del parámetro correspondiente a la matriz A se mantienen de manera que sean constantes por tramos hasta que se llega al instante de tiempo t_{12} , en el que los valores cambian bruscamente. Si los valores de parámetro son recibidos según una pauta por tramas, cada trama puede portar señalización que indica un instante de tiempo en el que un valor recibido se aplica exactamente. En el ejemplo, el parámetro correspondiente a B ha recibido valores aplicables de forma exacta en t_{21} y t_{22} , y el método puede, primeramente, predecir el valor de B en el instante de tiempo t_p inmediatamente precedente t_{12} . Una vez computada la matriz de operación de tratamiento compuesta BA en t_p , y $t_a = t_{11}$, la matriz BA puede ser interpolada entre t_a y t_p . El método puede entonces predecir el valor de B en un nuevo instante de predicción $t_p = t_{12}$, computar los valores de BA en t_p y $t_a = t_{22}$, e interpolar BA directamente entre t_p y t_a . Una vez más, la operación de tratamiento compuesta BA ha sido interpolada a través del intervalo, y se ha encontrado su valor en todos los instantes de tiempo de salida. En comparación con la situación anterior, según se ilustra en la Figura 9a, en el caso de que A y B hubieran sido interpoladas individualmente, y BA computada por multiplicación de A y B en cada instante de tiempo de salida, se necesita un número reducido de multiplicaciones matriciales y la complejidad computacional se rebaja.

Equivalentes, extensiones, alternativas y miscelánea

Realizaciones adicionales de la presente invención resultarán evidentes para una persona experta en la técnica tras el estudio de la anterior descripción. Incluso aunque la presente descripción y los dibujos divulguen realizaciones y ejemplos, la invención no está limitada por estos ejemplos específicos. Pueden realizarse numerosas modificaciones y variaciones sin apartarse del alcance de la presente invención, que se define por los dibujos que se acompañan. Cualesquiera signos de referencia que aparecen en las reivindicaciones no deben ser entendidos como limitativos de su alcance.

Adicionalmente, pueden concebirse y llevarse a efecto variaciones en las realizaciones divulgadas por parte de la persona experta a la hora de poner en práctica la invención, a partir del estudio de los dibujos, la descripción y las reivindicaciones que se acompañan. En las reivindicaciones, la expresión “que comprende” no excluye otros elementos o etapas, y el artículo indefinido “un” o “uno” no excluye una pluralidad. El mero hecho de que ciertas medidas se refieran en diferentes reivindicaciones dependientes entre sí no indica que no pueda usarse de forma ventajosa una combinación de estas medidas.

Los sistemas y métodos divulgados en lo anterior de esta memoria pueden ser implementados como software, *firmware*, o software instalado de forma permanente en hardware, hardware, o una combinación de los mismos. En la implementación en hardware, la división de tareas entre las unidades funcionales referidas en la anterior descripción no se corresponde, necesariamente, a la división en unidades físicas; al contrario, un único componente físico puede tener múltiples capacidades funcionales, y una única tarea puede llevarse a cabo por varios componentes físicos en cooperación. Ciertos componentes, o todos los componentes, pueden ser implementados como software ejecutado por un procesador de señal digital o microprocesador, o bien ser implementados como hardware o como un circuito integrado específico de la aplicación. Tal software puede ser distribuido en medios legibles por computadora, los cuales pueden comprender medios de almacenamiento informáticos (o medios no transitorios) y medios de comunicación (o medios transitorios). Como es bien conocido por una persona experta en la técnica, la expresión “medios de almacenamiento informáticos” incluye medios tanto volátiles como no volátiles, y tanto extraíbles como no extraíbles, implementados en cualquier método o tecnología para el almacenamiento de información, tal como en instrucciones legibles por computadora, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento informáticos incluyen RAM, ROM, EEPROM, memoria de tipo flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otros dispositivos de almacenamiento de disco óptico, cintas magnéticas, cinta magnética, dispositivos de almacenamiento de disco magnético u otros dispositivos de almacenamiento magnéticos, o cualquier otro medio que pueda ser utilizado para almacenar la información deseada y al que pueda accederse por parte de una computadora, si bien no están limitados por estos. Por otra parte, es bien conocido para la persona experta que los medios de comunicación incorporan, por lo común, instrucciones legibles por computadora, estructuras de datos, módulos de programa u otros datos dispuestos en una señal de datos modulada, tal como una onda portadora u otro mecanismo de transporte, e incluyen cualesquiera medios de suministro de información.

REIVINDICACIONES

1.- Un método para la mejora del diálogo en un descodificador de un sistema de audio, de tal manera que el método comprende las etapas de:

5 recibir una pluralidad de señales de mezcla descendente (512, 612, 712), que constituyen una mezcla en sentido descendente de una pluralidad más grande de canales;

recibir parámetros para mejora de diálogo (516, 616, 716), de tal manera que los parámetros se han definido con respecto a un subconjunto de la pluralidad de canales que incluye canales que comprenden diálogo, de tal modo que el subconjunto de la pluralidad de canales se mezcla en sentido descendente para formar un subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente;

10 recibir parámetros de reconstrucción (514, 614, 714) que permiten una reconstrucción paramétrica de los canales que están mezclados en sentido descendente en el subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente;

15 mezclar en sentido ascendente únicamente el subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente paraméricamente, basándose en los parámetros de reconstrucción (514, 614, 714), a fin de reconstruir únicamente un subconjunto (618a, 618b, 718a) de la pluralidad de canales que incluye el subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716);

20 aplicar una mejora del diálogo al subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716), utilizando los parámetros para mejora de diálogo (516, 616, 716), a fin de proporcionar al menos una señal mejora en diálogo (619, 719); y

proporcionar versiones mejoradas en diálogo (520, 620, 720) del subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente, al mezclar la al menos una señal mejorada en diálogo (619, 719) con al menos otra de dicho subconjunto de la pluralidad de señales de mezcla descendente (618b, 712a).

25 2.- El método de acuerdo con la reivindicación 1, en el cual, en la etapa de mezclar en sentido ascendente únicamente el subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente paraméricamente, se utilizan señales no descorrelacionadas para reconstruir tan solo un subconjunto (618a, 618b, 718a) de la pluralidad de canales que incluye el subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716).

30 3.- El método de acuerdo con la reivindicación 1, en el cual la mezcla se realiza de acuerdo con parámetros de mezcla (522, 622, 722) que describen una contribución de la al menos una señal mejorada en diálogo (619, 719) a las versiones mejoradas en diálogo (620, 720) del subconjunto de la pluralidad de señales de mezcla descendente.

35 4.- El método de acuerdo con una cualquiera de las reivindicaciones precedentes, en el cual la etapa de mezclar en sentido ascendente únicamente el subconjunto (512a, 712a) de la pluralidad de señales de mezcla descendente paraméricamente comprende reconstruir tan solo el subconjunto (718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 716),

de tal manera que la etapa de aplicar la mejora del diálogo comprende predecir y mejorar una componente de diálogo a partir del subconjunto de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo, utilizando los parámetros para mejora de diálogo (516, 716), a fin de proporcionar la al menos una señal mejorada en diálogo (719), y

40 en el cual la mezcla comprende mezclar la al menos una señal mejorada en diálogo (719) con el subconjunto (712a) de la pluralidad de señales de mezcla descendente; y/o

45 que comprende adicionalmente recibir una señal de audio que representa diálogo, de tal modo que la etapa de aplicar la mejora del diálogo comprende aplicar la mejora del diálogo al subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716), utilizando adicionalmente la señal de audio que representa diálogo.

5.- El método de acuerdo con una cualquiera de las reivindicaciones precedentes, que comprende adicionalmente recibir parámetros de mezcla (522, 622, 722) para mezclar la al menos una señal mejorada en diálogo (619, 719) con al menos otra de dicho subconjunto de la pluralidad de señales de mezcla descendente (618b, 712a).

50 6.- El método de acuerdo con la reivindicación 3 o la reivindicación 5, en el cual las etapas de mezclar en sentido ascendente únicamente el subconjunto de la pluralidad de señales de mezcla descendente, aplicar la mejora del diálogo y mezclar se llevan a cabo como operaciones matriciales definidas por los parámetros de reconstrucción, los parámetros para la mejora del diálogo y los parámetros de mezcla, respectivamente, y que, opcionalmente, comprende, de manera adicional, combinar, por multiplicación de matrices, las operaciones matriciales

correspondientes a las etapas de mezclar en sentido ascendente únicamente el subconjunto de la pluralidad de señales de mezcla descendente, aplicar la mejora del diálogo y mezclar, en una única operación matricial, antes de su aplicación al subconjunto de la pluralidad de señales de mezcla descendente.

5 7.- El método de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que los parámetros de mejora de diálogo (516, 616, 716) y los parámetros de reconstrucción (514, 614, 714) son dependientes de la frecuencia, y, opcionalmente, en el cual los parámetros para mejora de diálogo (516, 616, 716) se definen con respecto a un primer conjunto de bandas de frecuencias, y los parámetros de reconstrucción se definen con respecto a un segundo conjunto de bandas de frecuencias, de tal modo que el segundo conjunto de bandas de frecuencias es diferente del primer conjunto de bandas de frecuencias.

10 8.- El método de acuerdo con cualquiera de las reivindicaciones precedentes, en el cual:

se reciben repetidamente valores de los parámetros para mejora de diálogo (516, 616, 716) y son asociados con un primer conjunto de instantes de tiempo ($T1 = \{t11, t12, t13, \dots\}$) en los que los valores respectivos se aplican de forma exacta, de tal manera que se ha de llevar a cabo una primera configuración de interpolación predefinida (I1) entre instantes de tiempo consecutivos; y

15 se reciben repetidamente valores de los parámetros de reconstrucción (514, 614, 714) y son asociados con un segundo conjunto de instantes de tiempo ($T2 = \{t21, t22, t23, \dots\}$) en los que los valores respectivos se aplican de forma exacta, de tal manera que se ha de llevar a cabo una segunda configuración de interpolación predefinida (I2) entre instantes de tiempo consecutivos,

de tal modo que el método comprende, adicionalmente:

20 seleccionar un tipo de parámetro que consiste, bien en parámetros para la mejora del diálogo o bien en parámetros de reconstrucción, y de tal manera que el conjunto de instantes de tiempo asociado con el tipo seleccionado comprende al menos un instante de predicción que es un instante de tiempo (t_p) que está ausente del conjunto asociado con el tipo no seleccionado;

predecir un valor de los parámetros del tipo no seleccionado en el instante de predicción (t_p); y

25 computar, basándose en al menos el valor predicho de los parámetros del tipo no seleccionado y en un valor recibido de los parámetros del tipo seleccionado, una operación de tratamiento compuesta que representa al menos la mezcla en sentido ascendente de únicamente el subconjunto de señales de mezcla descendente (512a, 612a, 712a), seguida por la mejora del diálogo en el instante de predicción (t_p); y

30 computar, basándose en al menos un valor de los parámetros del tipo seleccionado y en un valor de los parámetros del tipo no seleccionado, de los cuales al menos uno es un valor recibido, dicha operación de tratamiento compuesta en un instante de tiempo adyacente (t_a) del conjunto asociado con el tipo seleccionado o con el no seleccionado,

35 de tal manera que dichas etapas de mezclar en sentido ascendente tan solo el subconjunto de la pluralidad de señales de mezcla descendente (512a, 612a, 712a) y aplicar la mejora del diálogo se llevan a cabo entre el instante de predicción (t_p) y el instante de tiempo adyacente (t_a) por medio de un valor interpolado de la operación de tratamiento compuesta computada, y, opcionalmente, de modo que el tipo seleccionado de parámetros son los parámetros de reconstrucción (514, 614, 714).

40 9.- El método de acuerdo con la reivindicación 8, en el cual dicha operación de tratamiento compuesta, en el instante de tiempo adyacente (t_a), es computada basándose en un valor recibido de los parámetros del tipo seleccionado y en un valor recibido de los parámetros del tipo no seleccionado.

10.- El método de acuerdo con la reivindicación 8 o la reivindicación 9, que comprende adicionalmente:

seleccionar, basándose en las primera y segunda configuraciones de interpolación, una configuración de interpolación compuesta (I3) de acuerdo con una regla de selección predefinida,

45 en el cual dicha interpolación de las respectivas operaciones de tratamiento compuestas computadas es de acuerdo con la configuración de interpolación compuesta, y, opcionalmente, de manera que la regla de selección predefinida se ha definido para el caso de que las primera y segunda configuraciones de interpolación sean diferentes, y de tal modo que, en respuesta a que la primera configuración de interpolación (I1) es lineal y la segunda configuración de interpolación (I2) es constante por tramos, se selecciona la interpolación lineal como configuración de interpolación compuesta.

50 11.- El método de acuerdo con cualquiera de las reivindicaciones 8 a 10, en el cual la predicción del valor de los parámetros del tipo no seleccionado en el instante de predicción (t_p) se realiza de acuerdo con la configuración de interpolación para los parámetros del tipo no seleccionado.

12.- El método de acuerdo con cualquiera de las reivindicaciones 8 a 11, en el que la operación de tratamiento compuesta se computa como una única operación matricial antes de ser aplicada al subconjunto de la pluralidad de señales de mezcla descendente (512a, 612a, 712a), y, opcionalmente, en el cual:

se selecciona la interpolación lineal como configuración de interpolación compuesta; y

5 el valor interpolado de las respectivas operaciones de tratamiento compuestas computadas se computa por interpolación matricial lineal.

13.- El método de acuerdo con cualquiera de las reivindicaciones precedentes, en el cual la mezcla de la al menos una señal mejorada en diálogo (619, 719) con al menos otra señal (618b, 712a) está restringida a una selección incompleta de la pluralidad de señales de mezcla descendente (512, 612, 712).

10 14.- Un producto de programa informático que comprende un medio legible por computadora, con instrucciones para llevar a cabo el método de acuerdo con una cualquiera de las reivindicaciones 1-13, cuando dicho producto de programa se hace marchar en una computadora.

15.- Un decodificador (500, 600, 700) para mejorar el diálogo en un sistema de audio, de tal manera que el decodificador comprende:

15 un componente de recepción (502, 602, 702), configurado para recibir:

una pluralidad de señales de mezcla descendente (512, 612, 712), que son una mezcla en sentido descendente de una pluralidad más grande de canales,

20 parámetros para mejora de diálogo (516, 616, 716), de tal modo que los parámetros se han definido con respecto a un subconjunto de la pluralidad de canales que incluye canales que comprenden diálogo, de manera que el subconjunto de la pluralidad de canales se mezcla en sentido descendente hasta formar un subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente, y

parámetros de reconstrucción (514, 614, 714), que permiten la reconstrucción paramétrica de canales que están mezclados en sentido descendente hasta formar el subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente;

25 un componente de mezcla en sentido ascendente (504, 604, 704), configurado para mezclar en sentido ascendente únicamente el subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente paraméricamente, basándose en los parámetros de reconstrucción (514, 614, 714), a fin de reconstruir únicamente un subconjunto (618a, 618b, 718a) de la pluralidad de canales que incluye el subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716); y

30 un componente de mejora de diálogo (506, 606, 706), configurado para aplicar la mejora del diálogo al subconjunto (618a, 718a) de la pluralidad de canales con respecto al cual se han definido los parámetros para mejora de diálogo (516, 616, 716), utilizando los parámetros para mejora de diálogo (516, 616, 716), a fin de proporcionar al menos una señal mejorada en diálogo (619, 719); y

35 un componente de mezcla (508, 608, 708), configurado para proporcionar versiones mejoradas en diálogo (520, 620, 720) del subconjunto (512a, 612a, 712a) de la pluralidad de señales de mezcla descendente al mezclar la al menos una señal mejorada en diálogo (619, 719) con al menos otra señal de dicho subconjunto de la pluralidad de señales de mezcla descendente (618b, 712a).

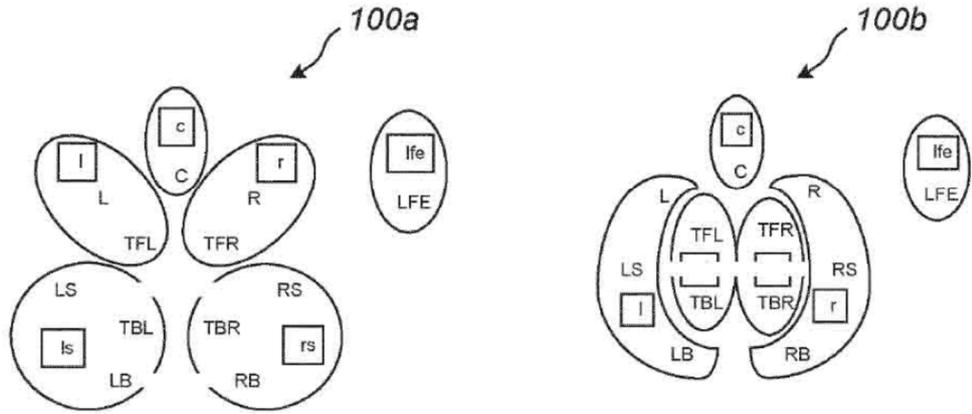


Fig. 1a

Fig. 1b

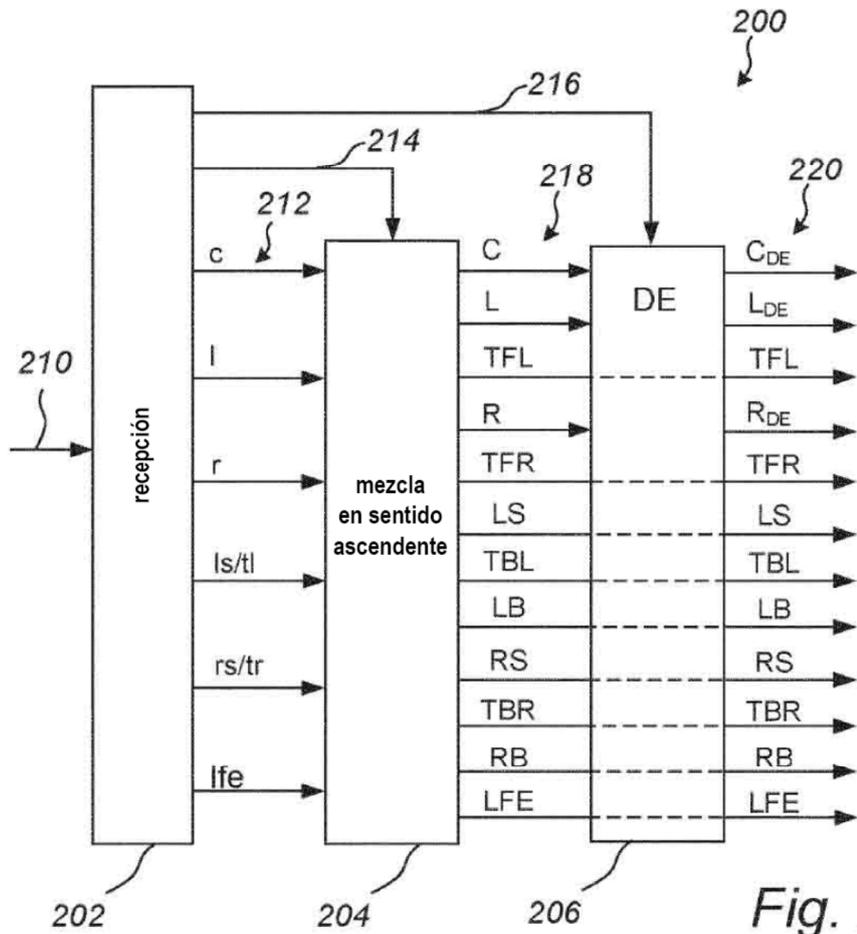


Fig. 2

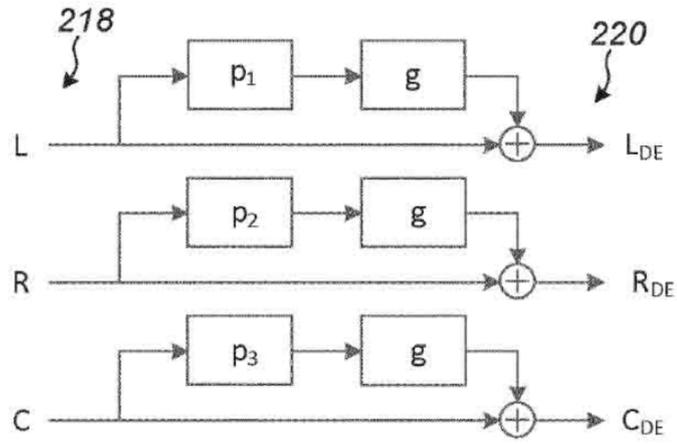


Fig. 3

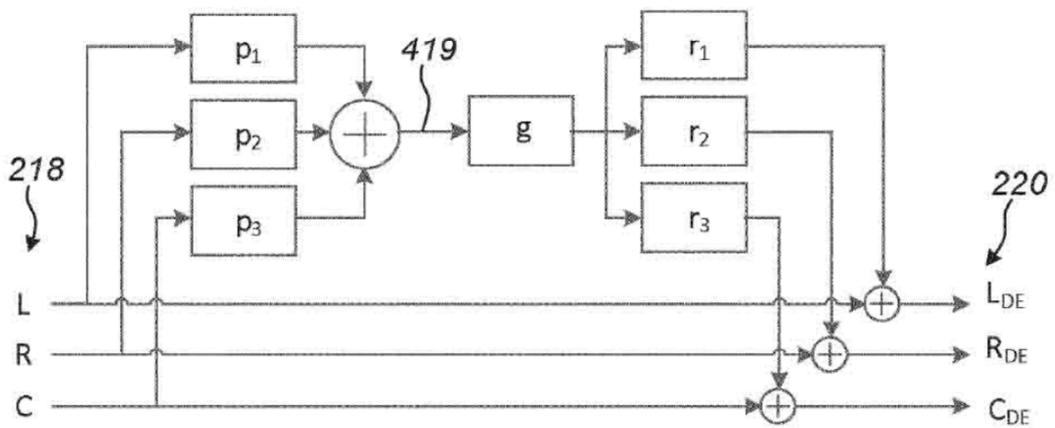


Fig. 4

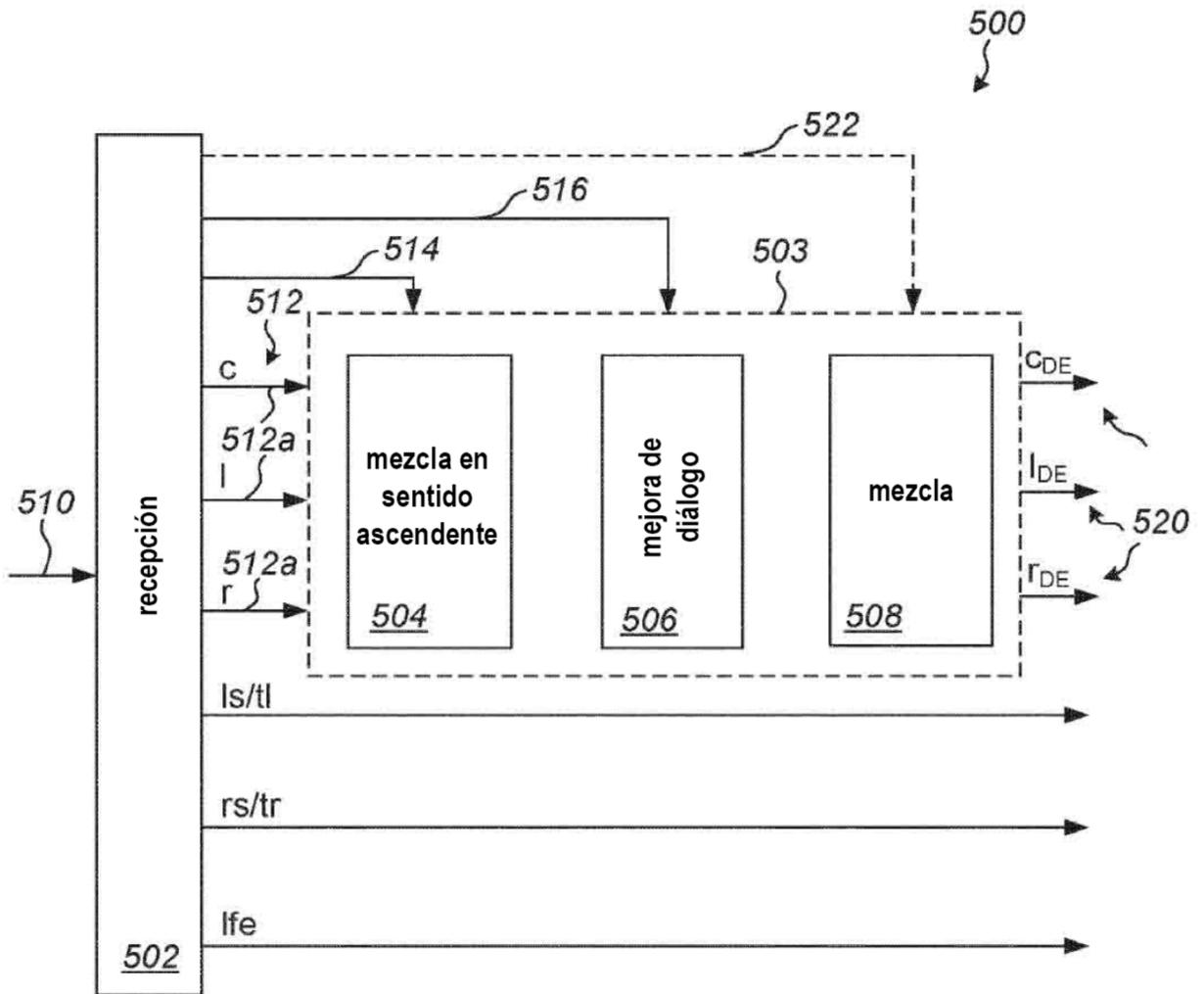


Fig. 5

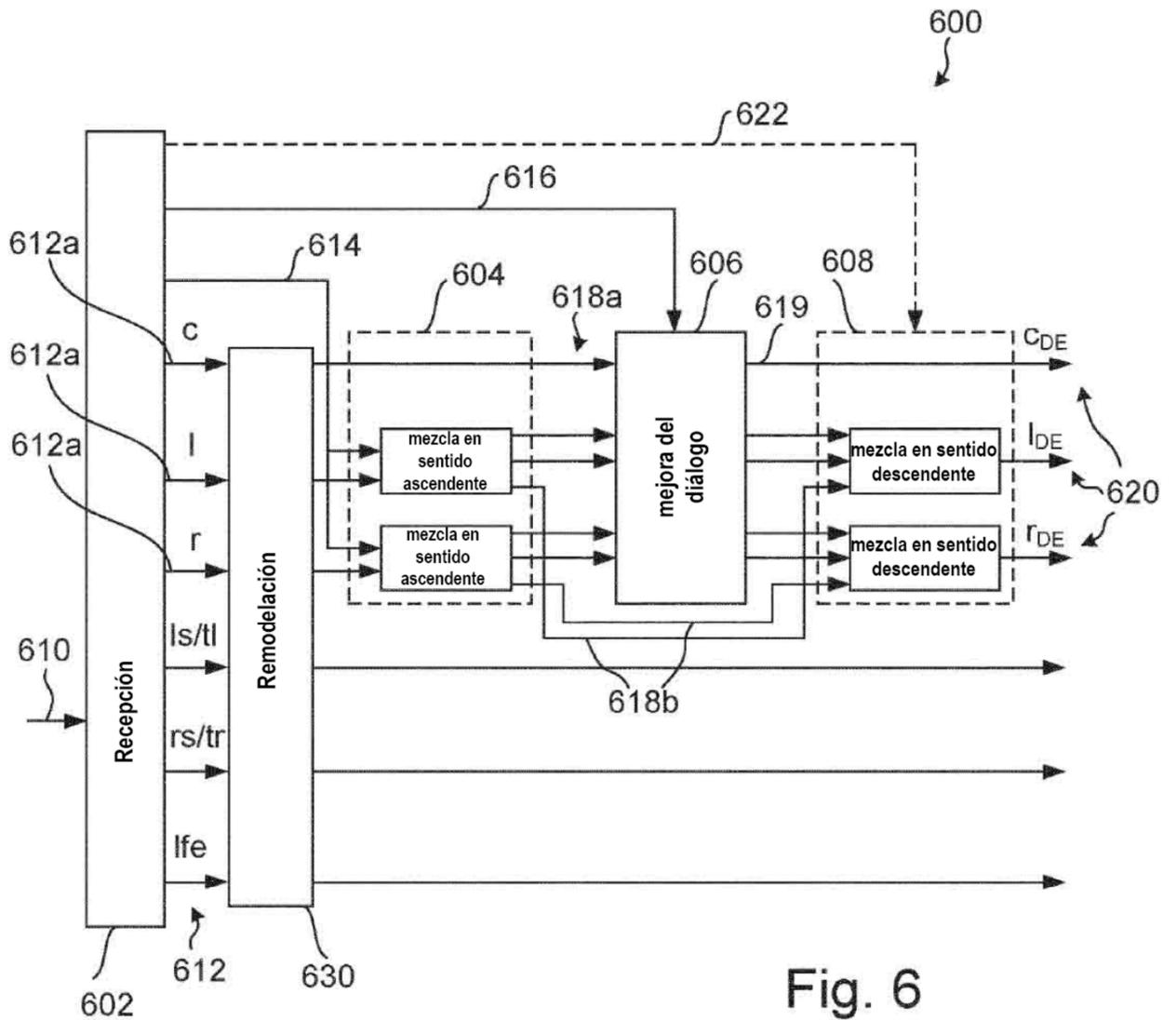


Fig. 6

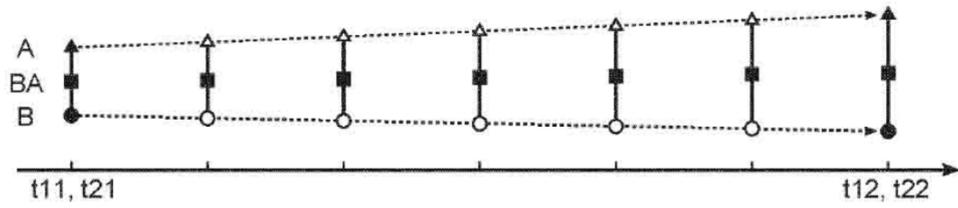


Fig. 9a

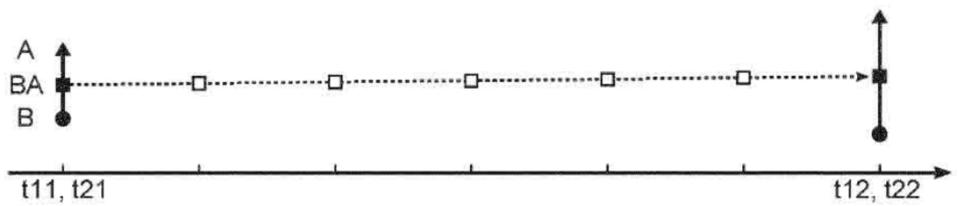


Fig. 9b

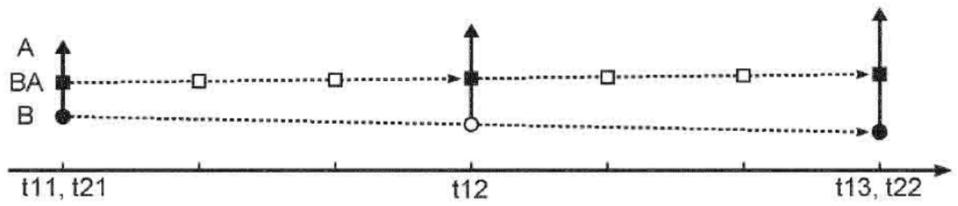


Fig. 9c

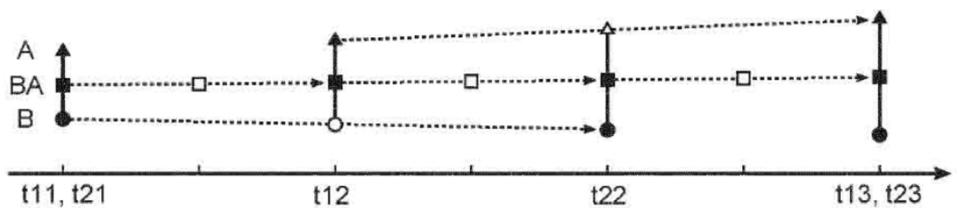


Fig. 9d

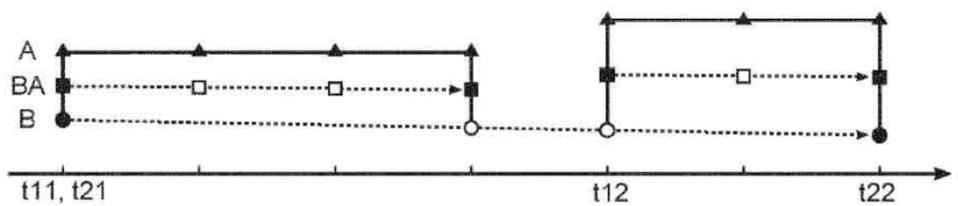


Fig. 9e