

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 570 967**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/008 (2013.01)

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.10.2008** **E 13176079 (5)**

97 Fecha y número de publicación de la concesión europea: **06.04.2016** **EP 2650877**

54 Título: **Método y aparato para proporcionar audio alineado de múltiples canales**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.05.2016

73 Titular/es:

TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)
(100.0%)
164 83 Stockholm, SE

72 Inventor/es:

JONES, ANTHONY RICHARD

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 570 967 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 y aparato para proporcionar audio alineado de múltiples canales

Sector técnico

5 La invención hace referencia a la codificación de audio en general y, en particular, a un método y a un aparato para proporcionar audio alineado de múltiples canales.

Antecedentes de la invención

10 Los estándares de codificación audiovisual modernos, tales como MPEG-1 y MPEG-2, proporcionan medios para el transporte de múltiples componentes de audio y video dentro de un único flujo de transporte. Los componentes de audio individuales y separados pueden ser alineados con componentes de video seleccionados. El audio sincronizado de múltiples canales, tal como el sonido envolvente, solo es proporcionado en términos de un único componente de audio de sonido envolvente premezclado, por ejemplo, un único componente de audio Dolby 5.1. No obstante, actualmente no existe ningún medio para el transporte de manera sincronizada de componentes de audio de múltiples canales individualizados.

15 En particular, las especificaciones de audio MPEG-1 y MPEG-2 (ISO/IEC 11172-3 e ISO/IEC 13818-3 respectivamente) describen medios de codificación y de empaquetamiento de señales de audio digitales. Estos incluyen esquemas que están especificados para soportar varias formas de sonido de múltiples canales que utilizan un único componente de flujo de transporte MPEG-2. Estas provisiones son compatibles con lo anterior, con el sistema de audio MPEG-1 previo. En la técnica anterior, solo es posible asegurar la sincronización requerida de los canales mediante el ensamblaje de los diferentes canales de audio en un solo componente de transporte. Estos esquemas requieren:

[a] la utilización de métodos de compresión de sonido envolvente (por ejemplo, Dolby 5.1) o bien

[b] la utilización de técnicas de compresión patentadas, o bien

[c] la utilización de audio no comprimido.

25 La utilización de métodos de compresión de sonido envolvente reduce la tasa de bits requerida por los múltiples canales aprovechando las redundancias existentes entre los diferentes canales, y también las características del sistema auditivo humano, que hacen que ciertas características espaciales del sonido sean indetectables y de este modo puedan ser ocultas durante el procesamiento. Estos complejos esquemas proporcionan medios adecuados para manejar una única etapa de codificación en la cual solo se espera una operación de codificación y de descodificación, pero no son ideales para señales que, por razones prácticas y operativas (por ejemplo, 30 alimentaciones de fuentes desde una ubicación remota a las instalaciones centrales de edición), necesitan ser recodificadas quizás varias veces en las redes de transmisión. Esto se debe a problemas de concatenación resultantes de múltiples operaciones de codificación secuenciales que degradan la calidad del audio. Este es particularmente el caso en el que la capacidad está limitada, haciendo que la tasa de bits se reduzca substancialmente, dejando poco espacio de cabecera para tratar tales degradaciones en la codificación y la 35 transmisión concatenadas.

La utilización de técnicas de compresión patentadas típicamente requiere la utilización de equipos patentados externos adicionales que conllevan un mayor gasto y mayores complicaciones operativas. Este método puede asimismo sufrir la misma degradación de la calidad que la que produce la concatenación de más de una codificación/descodificación.

40 Mientras que si el audio es enviado en formato no comprimido (por ejemplo, muestras de PCM Lineal no comprimidas), la tasa de datos requerida es una tasa de datos muy elevada (por ejemplo, aproximadamente 3 Mbit/s por cada par de dos canales).

45 Aunque lo anterior no supone generalmente un problema cuando se proporcionan medios audiovisuales finalizados a consumidores, presenta un problema para la industria de la producción de medios audiovisuales, dado que la industria aprovecha cada vez más la ubicuidad de las redes de datos de alta velocidad modernas para enviar medios audiovisuales "no procesados" (es decir, el material de fuente utilizado para producir televisión, películas y otros medios) de manera instantánea en forma comprimida entre instalaciones de producción o, de hecho desde las instalaciones de producción hacia los puntos de distribución de red de televisión o de audio, por ejemplo, transmisores terrestres, enlaces ascendentes vía satélite o cabeceras de transmisión por cable.

50 Por ejemplo, los equipos de cámaras de localización típicamente proporcionan material audiovisual a los estudios centrales de televisión, para su edición y distribución a estaciones de televisión filiales para una eventual emisión a espectadores. Los estándares de codificación audiovisual mencionados anteriormente no permiten el envío de audio sincronizado de múltiples canales sin pre-mezcla, sumándose con ello a la complejidad de sus equipos de campo, o impidiéndoles proporcionar audio de múltiples canales.

Existe una particular necesidad de poder transmitir audio de múltiples canales que tiene el requisito de un alineamiento preciso de canal a canal, de tal manera que las señales de audio pueden ser codificadas a continuación como audio de sonido envolvente en el que el alineamiento temporal de múltiples canales es importante, utilizando los estándares MPEG anteriores dado que una mayoría de los equipos de producción ya están configurados para su utilización con estos estándares. De acuerdo con ello la presente invención propone métodos y aparatos que proporcionan un mecanismo económico y conveniente para proporcionar audio de múltiples canales manteniendo la calidad del sonido y un alineamiento temporal preciso entre los canales.

El documento Z US 2008/001 3614 A1 describe un dispositivo y un método para generar un flujo de datos para generar una representación de múltiples canales. La sincronización en el tiempo se proporciona mediante el cálculo de la información de huella que se lleva a cabo en el lado del codificador para al menos un canal de base. La información de huella se inserta en un flujo de datos conectado en el tiempo con los datos adicionales de múltiples canales. En el descodificador, la información de la huella se calcula desde el al menos un canal de base y se utiliza junto con la información de huella extraída del flujo de datos para calcular y compensar un desfase de tiempo entre el flujo de datos con información adicional de múltiples canales y el flujo de datos con el al menos un canal de base, por ejemplo, por medio de una correlación, para obtener una representación sincronizada de múltiples canales.

El documento US 2004/0049379 A1 describe la codificación y la descodificación de múltiples canales. Un codificador de audio lleva a cabo una transformada de múltiples canales de pre-procesamiento sobre la fecha del audio de múltiples canales, variando la transformada con el fin de controlar la calidad. El codificador agrupa múltiples ventanas de diferentes canales en uno o más cuadros y proporciona información de configuración del mosaico, lo que permite al codificador aislar los transitorios que aparecen en un canal particular con pequeñas ventanas, pero utiliza ventanas grandes en otros canales. Utilizando una variedad de técnicas, el codificador efectúa transformadas de múltiples canales flexibles que aprovechan de manera efectiva la correlación intercanales. Un descodificador de audio lleva a cabo el procesamiento y la descodificación correspondientes. Además, el descodificador lleva a cabo una transformada de múltiples canales de pre-procesamiento para alguno de varios propósitos diferentes.

El documento XP030014396, ISSN: 0000-0341, "Text of ISO/IEC 13818-1:200X (3rd edition)", 75. MPEG meeting; 16-01-2006 – 20-01-2006; Bangkok; no. N7904, es una recomendación de TUT para el estándar internacional 13818-1. Este estándar es la especificación ISO para los sistemas de flujo de transporte MPEG-2. Define cómo se proporciona la hora actual (PCR) para un programa, y cómo se señalan las horas de presentación (PTS) para cada componente individual.

Compendio

Las realizaciones de la presente invención proporcionan un método de codificación de audio de acuerdo con la reivindicación 1 y de inclusión del citado audio codificado en un flujo de transporte digital, que comprende recibir en una entrada del codificador una pluralidad de señales de audio situadas en el mismo momento del tiempo, asignando marcas de tiempo idénticas por unidad de tiempo a todas de la pluralidad de señales de audio situadas en el mismo momento del tiempo, e incorporando las señales de audio con marcas de tiempo idénticas en el flujo de transporte digital.

Opcionalmente, la etapa de recepción comprende además muestrear las señales de audio situadas en el mismo momento del tiempo para formar tramas de datos de audio de un tamaño predeterminado, y alinear las citadas tramas de datos de audio para mantener el que las señales de audio estén situadas en el mismo momento del tiempo, y en la que la etapa de asignar marcas de tiempo idénticas se lleva a cabo en las tramas alineadas de datos de audio.

Opcionalmente, el método comprende además comprimir las tramas alineadas de datos de audio con ajustes de configuración de codificador de audio antes de asignar las marcas de tiempo, y asignar los datos de audio comprimidos y con marcas de tiempo idénticas a una pluralidad de canales mono de un flujo de transporte.

Opcionalmente, la pluralidad de canales mono comprende uno o más componentes de audio dual mono convencionales.

Opcionalmente, el tamaño predeterminado es el tamaño de una unidad de acceso en el estándar MPEG, y el flujo de transporte de video es un flujo de transporte MPEG-1 o MPEG-2.

Opcionalmente, las marcas de tiempo son marcas de tiempo de presentación.

Opcionalmente, el método de cualquier reivindicación precedente, en el que la etapa de incorporación del audio en un flujo de video digital comprende la multiplexación de los datos de audio comprimidos y con marcas de tiempo idénticas en un flujo de transporte.

Las realizaciones de la presente invención proporcionan asimismo un método de descodificación de un flujo de transporte digital que incluye audio codificado de acuerdo con alguno de los métodos de codificación anteriores, que comprenden recibir una pluralidad de señales de audio con marcas de tiempo idénticas, representativas de una pluralidad de canales de audio individuales situados en el mismo momento del tiempo, detectar las marcas de

tiempo para determinar marcas de tiempo compartidas, y proporcionar la pluralidad de canales de audio individuales situados en el mismo momento del tiempo de acuerdo con las marcas de tiempo detectadas como canales múltiples.

Opcionalmente, la pluralidad de señales de audio con marcas de tiempo idénticas han sido muestreadas y alineadas para formar tramas alineadas de datos de audio y en las que las marcas de tiempo idénticas han sido aplicadas a las tramas alineadas de datos de audio.

Opcionalmente, las tramas alineadas de datos de audio han sido comprimidas antes de la asignación de las marcas de tiempo, y el método comprende además descomprimir las tramas de datos de audio para producir las señales de audio individuales para ser producidas.

Opcionalmente, la etapa de producir la pluralidad de canales de audio individuales situados en el mismo momento del tiempo comprende presentar el audio utilizando la marca de tiempo de solo una de las señales de audio situadas en el mismo momento del tiempo.

Opcionalmente, el flujo de transporte digital es un flujo de transporte de video digital, y las tramas alineadas de datos de audio comprenden paquetes PES.

Las realizaciones de la presente invención proporcionan asimismo un aparato de codificación adaptado para llevar a cabo cualquiera de los métodos de codificación anteriores.

Las realizaciones de la presente invención proporcionan asimismo un aparato de decodificación adaptado para llevar a cabo cualquiera de los métodos de decodificación anteriores.

Breve descripción de los dibujos

A continuación, se describirán un método y un aparato para proporcionar audio alineado de múltiples canales, únicamente a modo de ejemplo, y con referencia a los dibujos que se acompañan, en los cuales:

la figura 1 muestra un diagrama de bloques esquemático de una porción de un aparato de codificación mono analógico o digital de acuerdo con la técnica anterior;

la figura 2 muestra un diagrama de bloques esquemático de una porción de un aparato de decodificación mono analógico o digital de acuerdo con la técnica anterior;

la figura 3 muestra un diagrama de bloques esquemático de una porción de un aparato de codificación mono estéreo o dual, analógico o digital de acuerdo con la técnica anterior;

la figura 4 muestra un diagrama de bloques esquemático de una porción de un aparato de decodificación mono estéreo o dual, analógico o digital de acuerdo con la técnica anterior;

la figura 5 muestra un diagrama de flujo de una porción de codificación del método para proporcionar audio alineado de múltiples canales de acuerdo con una realización de la invención;

la figura 6 muestra un diagrama de flujo de una porción de decodificación del método para proporcionar audio alineado de múltiples canales de acuerdo con una realización de la invención;

la figura 7 muestra un diagrama de bloques esquemático de una porción de un aparato de codificación analógico o digital de múltiples canales de acuerdo con una realización de la invención;

la figura 8 muestra un diagrama de bloques esquemático de una porción de un aparato de decodificación analógico o digital de múltiples canales de acuerdo con una realización de la invención.

Descripción detallada

A continuación, se describirá una realización de la invención con referencia a los dibujos que se acompañan, en los cuales a partes o etapas iguales o similares se les asignan números de referencia iguales o similares.

Lo que sigue se basa en el estándar MPEG-2. No obstante, resultará evidente que la invención subyacente es igualmente aplicable a otros estándares de audio comprimido que soportan codificación dual-mono, tales como codificación de audio avanzada (AAC – Advanced Audio Coding, en inglés) o Dolby digital.

Las especificaciones de audio MPEG-1 y MPEG-2 describen medios de codificación y empaquetamiento de señales de audio digital. Los datos de audio procesados se pasan a la capa de sistemas MPEG (ISO/IEC 13818-1) para su posterior empaquetamiento en un flujo de transporte (TS – Transport Stream, en inglés) antes de que sea transmitido a través de redes de comunicación tal como los sistemas de telecomunicaciones o de emisión. Estas reglas de empaquetamiento MPEG definen una estructura que proporciona sintaxis a los flujos de bits. En particular, los flujos de bits contienen marcas de tiempo que son utilizadas por el decodificador para controlar la temporización del audio producido decodificado y restaurado. Estas marcas de tiempo se utilizan para una temporización precisa de los componentes tanto de audio como de video.

Los estándares MPEG definen dos tipos de marca de tiempo - una marca de tiempo de decodificador (DTS – Decoder Time Stamp, en inglés), que define cuándo deben ser presentados al decodificador los datos codificados recibidos, y marcas de tiempo de presentación (PTS), que definen cuándo debe el sistema proporcionar el audio o el video para que, respectivamente, se oiga o se vea. Es el último tipo de marca de tiempo el que se utiliza con mayor frecuencia.

Gestionando estas marcas de tiempo tal como se describe con más detalle a continuación, un sistema de transmisión audiovisual de acuerdo con una realización de la invención es capaz de presentar de manera apropiada las diferentes señales de audio separadas de un conjunto de múltiples canales para su codificación o decodificación al mismo tiempo, consiguiendo con ello la sincronización requerida entre el conjunto de múltiples canales.

La figura 1 muestra un diagrama de bloques esquemático de una porción de un aparato de codificación mono analógico o digital de acuerdo con la técnica anterior, que ilustra el flujo sistemático de datos de audio mediante un proceso de codificación, tal como por ejemplo MPEG-2. El proceso de decodificación es el proceso inverso de este, y se muestra en la figura 2.

Todos los ejemplos de las figuras muestran entradas duales analógica 110 y digital 105, pasando las entradas analógicas a través de un convertidor 120 de analógico a digital (A/D – Analogue to Digital, en inglés) para su digitalización antes de ser introducidas en el codificador 130. El audio digital 105 es introducido directamente en el codificador 130. Los canales separados se denotan mediante etiquetas a–d. No obstante, resultará evidente que la presente invención no está limitada a ningún conjunto de canales establecido, y es completamente escalable, y la entrada de audio puede ser solo analógica, solo digital o en formato dual, como se muestra.

Cuando la entrada es en forma analógica, el sonido analógico es muestreado digitalmente, por ejemplo, en forma de modulación de código por impulsos lineal (PCM – Linear Pulse Code Modulation, en inglés), antes de entrar en el codificador 130, donde se convierte en una forma reducida de bits.

El codificador 130 produce múltiples flujos de bits digitales codificados, uno para cada canal de audio separado, en una función de empaquetamiento 140, que empaqueta el audio en muestras de audio. Grupos definidos de muestras de audio son ensamblados y asociados en el dominio codificado mediante bloques de bits denominados unidades de acceso. Cada unidad de acceso es una porción empaquetada de audio, por ejemplo, una trama de 1152 muestras de audio.

Los canales empaquetados separados son a continuación multiplexados entre sí mediante el multiplexador 150, para formar un flujo de transporte 160.

El aparato de decodificación se muestra en la figura 2, y es esencialmente el proceso inverso. El flujo de transporte 160 es desmultiplexado por el desmultiplexador 250, que proporciona los canales de audio separados empaquetados, para ser desempaquetados mediante la función de desempaquetado 240, antes de ser decodificados en la etapa de decodificación 235 y es producido bien como un flujo digital 105 directo o a través del convertidor 220 de digital a analógico en forma analógica 110.

Las figuras 3 y 4 muestran el aparato de codificación y de decodificación para casos estéreo dual mono o sincronizados. Es posible añadir múltiples pares estéreo o dual - mono a un sistema, pero estos pares no serán bloqueados juntos dado que la especificación MPEG no realiza ninguna provisión explícita para ello (distinta de las opciones de sonido envolvente que sufren los problemas descritos en la sección de antecedentes) y, de este modo, permanecen como entidades separadas con marcas de tiempo separadas, siendo cada una reconstruida de manera independiente a la salida del decodificador.

Pueden existir varios canales de audio independientes, por ejemplo, pistas de sonido de diferentes idiomas, para su inclusión en algún flujo de transporte dado, estando cada uno codificado de manera separada.

Existen varias asociaciones diferentes entre los grupos de audio de entrada y sus contrapartidas codificados, dependiendo del número de canales requerido y de los criterios de calidad y las asignaciones de tasa de bits para cada canal elegido por el operador del sistema. El modo normal de operación es que estos canales de audio sean codificados de manera independiente, y no existe ningún requisito especial para bloquearlos entre sí.

Algunos de estos canales pueden ser asociados con una señal de video que los acompaña (es decir, en la que el audio es un sonido de video o de televisión) y el sistema alineará estas señales con su respectivo video de manera apropiada utilizando marcas de tiempo que son comunes para los flujos de video y de audio. El alineamiento de audio en este caso no es muy preciso – solo es necesario asegurar que se cumplen los requisitos de la sincronización de labios (Lip-Sync, en inglés). Este nivel de alineamiento no tiene el nivel de precisión necesario para el sonido envolvente de múltiples canales.

Por lo tanto, es normal que cada señal de audio monoaural independiente, monoaural dual o par estéreo (véase la figura 3) tenga una identidad separada (es decir, un flujo elemental) dentro del flujo de salida multiplexado y, de este

modo, cada uno tiene su propia marca de tiempo generada de manera independiente por el aparato de codificación durante la etapa de empaquetado, y se utiliza de manera independiente en el descodificador.

Brevemente, la solución propuesta a los inconvenientes de la técnica anterior descritos anteriormente es adaptar los formatos de transmisión MPEG-2 normales utilizados para los canales monoaural estándar o estéreo de dos canales aprovechando los controles de temporización proporcionados para estos casos y extendiéndolos a los de la situación de múltiples canales. De este modo, los descodificadores de acuerdo con las realizaciones de la invención son capaces de presentar múltiples canales de audio alineados con exactitud, y esto resuelve entonces el problema de sincronización y evita la concatenación de los sistemas de codificación y la degradación de la calidad recibida.

La solución es completamente compatible con la sintaxis de MPEG-2 existente y, de este modo, los descodificadores normales podrán presentar el audio de múltiples canales en la relación temporal convencional, y el método permite su repetición en sistemas concatenados sin temor a una degradación de la calidad, aunque sin el mismo grado de precisión de alineamiento que un descodificador de acuerdo con una realización de la invención.

Con más detalle, en el método de sincronización de múltiples canales propuesto, las diferentes señales de audio para las cuales se requiere un tratamiento en modo separado y sincronizado son procesadas con los mismos controles de temporización, de tal manera que se asignan las mismas marcas de tiempo en la sintaxis de la transmisión, de manera que un descodificador mantendrá también el alineamiento.

La figura 5 muestra una porción de un método de codificación 500 de acuerdo con una realización de la presente invención.

En la etapa 510, un número (N) predefinido de canales de audio independientes, que deben ser sincronizados y transportados sobre un único flujo de transporte sin ser convertidos en un solo componente, son introducidos en el aparato de codificación. El aparato de codificación forma K muestras de audio alineadas por unidad de tiempo, tomando una muestra de cada canal de audio de entrada, en el que las muestras corresponden al mismo instante de tiempo.

El aparato de codificación forma N/2 tramas de K muestras de audio alineadas por unidad de tiempo (etapa 520), en el que cada trama corresponde a la misma hora original, pero para canales de audio individuales, listos para su compresión utilizando el método de compresión elegido en la etapa 530 para formar unidades de acceso, utilizando típicamente compresión de audio dual-mono para cada par de canales de audio.

A las tramas comprimidas (es decir, las unidades de acceso) de las muestras de audio se les asignan a continuación marcas de tiempo idénticas, típicamente en forma de un campo de cabecera, en la etapa 540.

Las tramas de audio comprimidas con marca de tiempo son encapsuladas (es decir, empaquetadas) en paquetes PES que contienen pares dual mono del respectivo estándar en uso, por ejemplo, el estándar MPEG-2, en la etapa 550. La parte restante del proceso de codificación es igual que para el caso normal, es decir el audio empaquetado es empaquetado para su transporte y multiplexado con cualquier video relacionado (si aplica), y los otros canales, en un flujo de transporte 160 de salida.

La figura 6 muestra el proceso de descodificación inverso, de acuerdo con una realización de la invención.

En particular, el método de descodificación comprende recibir N/2 pares de canales de audio mono 610, detectar las marcas de tiempo 620, determinar qué pares comparten marcas de tiempo 630, descomprimir aquellos en N unidades de acceso de muestras de audio mono correspondientes a la misma hora de presentación 640, y a continuación proporcionar el audio descomprimido para presentar las N muestras exactamente a la misma hora, de acuerdo con la marca de tiempo 650 común única.

Resultará evidente que la provisión de alineamiento, compresión y marca de tiempo puede ser llevada a cabo mediante un solo componente de hardware del aparato de codificación, y los procesos inversos mediante un solo componente de hardware del aparato de descodificación.

El aparato de codificación para llevar a cabo el método de codificación descrito anteriormente de acuerdo con una realización de la invención se muestra en la figura 7, en la que puede apreciarse que existe una etapa adicional (es decir, la etapa de organización en tramas de múltiples canales 770) de procesamiento provista para alinear las diferentes señales de audio y para disponer y proporcionar para la utilización de una marca de tiempo común entre canales de audio separados pero sincronizados en la etapa de empaquetamiento 140.

El método y aparato preferiblemente opera utilizando canales dual mono para albergar los canales de audio separados pero sincronizados. Por ello, el aparato de codificación de la figura 7, 700 (y su aparato de descodificación correspondiente de la figura 8, 800) se muestra con codificador/descodificador separados y empaquetado/desempaquetado por par de canales de audio.

La figura 7 muestra un ejemplo que tiene cuatro canales de audio separados para ser sincronizados entre sí, con capacidad de entrada dual (analógica/digital). Los canales analógicos se pasan a través de un A/D 120 (a-d) para su

digitalización antes de ser proporcionados a una etapa de organización en tramas 770. Las entradas digitales son proporcionadas directamente en la etapa de organización en tramas 770.

La etapa de organización en tramas 770 crea bloques de muestras de audio situados en el mismo momento del tiempo a partir de todos los canales de audio, y los marca para su procesamiento junto con marcas de tiempo idénticas para todas las demás muestras de audio situadas en el mismo momento del tiempo. Esto, típicamente, toma la forma de una señal de sincronización 780 de marca de tiempo, que se pasa a la etapa de empaquetamiento 140, avanzando en el desarrollo del procesamiento.

Mientras tanto, las muestras de audio son proporcionadas en una etapa de codificación 730 estándar, como tramas situadas en el mismo momento del tiempo de pares muestreados dual mono formados en la etapa de organización en tramas 770, que a su vez proporciona las muestras de audio codificadas a la etapa de empaquetamiento 140, en la que son empaquetadas de acuerdo con la señal de sincronización 780 de marca de tiempo proporcionada por la etapa de organización en tramas 770.

Una realización preferida utilizaría bloques de muestras del tamaño de una unidad de acceso, y las marcas de tiempo de presentación (PTS) asociadas, perteneciendo las unidades de acceso a pares de múltiples canales que son comprimidos utilizando un único procesador de señal digital, lo que resulta en un conjunto de paquetes PES con valores de PTS idénticos, que contienen audio comprimido correspondiente exactamente a muestras de datos de audio originales situadas en el mismo momento del tiempo.

Donde existe un número impar de canales de entrada y se utilizan canales dual mono como mecanismo de transporte, entonces uno de los canales dual mono puede estar relleno simplemente con silencio.

Las salidas de cada una de las cadenas dual mono (par de funciones de codificador y de empaquetamiento) son a continuación multiplexadas entre sí de la manera normal por el multiplexador 150, para proporcionar un flujo de transporte de salida 160.

El aparato de decodificación 800 de acuerdo con una realización de la invención se muestra en la figura 8.

La operación de decodificación descomprime unidades de acceso independientes de audio correspondientes a múltiples componentes de audio dual-mono, manteniendo sus marcas de tiempo de presentación 835. Las tramas de las muestras decodificadas son a continuación presentadas por la etapa de presentación de tramas 870 en tiempos idénticos, de acuerdo con la marca de tiempo común que es compartida entre ellas. De este modo, múltiples pares de muestras correspondientes a la hora exacta de la muestra situada en el mismo momento del tiempo de presentan juntos, consiguiendo con ello el objetivo de mantener un alineamiento de audio de canal a canal exacto a través de múltiples pares de canales en toda la cadena de procesamiento de codificación/decodificación.

De este modo, el esquema completo para la sincronización de varios canales de audio utiliza las siguientes características en el aparato de codificación:

- Las muestras que están situadas en el mismo momento del tiempo a la entrada a través de múltiples canales de audio son conformadas en tramas alineadas de muestras de audio para adaptarse a los tamaños de la unidad de acceso comprimida.

- Las tramas de audio alineadas son comprimidas con configuraciones de codificador de audio idénticas, preferiblemente asignando dos canales monoaurales (como par) a cada componente de audio comprimido. No obstante, es asimismo posible utilizar canales estéreo o canales mono individuales como, o en lugar de, el par dual mono.

- A las unidades de acceso comprimidas se les asignan preferiblemente valores de marcas de tiempo de presentación idénticos, o marcas de tiempo de decodificador (DTS) – Decoder Time Stamps, en inglés) con un retardo de tiempo predeterminado.

- Los componentes de audio comprimidos son transmitidos como componentes de audio comprimidos mono de dos canales convencionales en el flujo de transporte MPEG-2.

En el aparato de decodificación (es decir, ubicación de recepción):

- se decodifican múltiples componentes de audio comprimidos, con el resultado de ser múltiples conjuntos (es decir, canales decodificados) de tramas descomprimidos de muestras de audio que tienen marcas de tiempo idénticas a través de los canales para cualquier punto dado de los flujos respectivos.

- Las tramas de audio descomprimidas para múltiples canales son presentadas a la salida utilizando la marca de tiempo de presentación de un solo componente, de tal manera que las muestras de audio de salida están situadas en el mismo momento del tiempo (o en un periodo de tiempo predeterminado después de un DTS).

El método y aparato descritos anteriormente proporcionan medios mediante los cuales es posible transmitir varios canales de audio a través de un sistema de comunicaciones de tal manera que permanecen sincronizados con otro con una precisión de muestra. Los medios anteriores para permitir esto estaban limitados a pares estéreo y a codificación de sonido envolvente, lo que provoca degradaciones en la calidad cuando se concatenan múltiples etapas de codificación. El presente método y aparato evitan las degradaciones de la calidad de los sistemas de la técnica anterior, y niegan la necesidad de soluciones de sonido envolvente más complejas y en ocasiones patentadas.

Por lo tanto, las realizaciones de la presente invención proporcionan medios para enviar un audio de múltiples canales "sin procesar" (es decir, no mezclado todavía en forma de sonido envolvente) a través del mismo flujo de transporte que el video con el que está relacionado, reduciendo con ello la degradación en la calidad del sonido debida a la concatenación, y otros problemas con otros métodos de transporte de audio conocidos previamente. Esto evita asimismo la necesidad de utilizar un procesamiento de sonido envolvente con pérdidas antes de la transmisión o PCM lineal no comprimido de ancho de banda grande.

La presente invención es particularmente adecuada para emitir transmisión de video de calidad que utiliza audio de múltiples canales sin convertirlo en un único componente (por ejemplo, sonido envolvente 5.1). No obstante, resultará evidente que las realizaciones de la presente invención pueden ser aplicadas igualmente a flujos de transporte solo de audio, tal como los utilizados para proporcionar sonido de radio de múltiples canales u otros similares.

La presente invención resulta particularmente beneficiosa en sistemas en los que se envía audio comprimido para su procesamiento en sonido envolvente en otra ubicación. Esto sucede dado que cuando se utilizan tales fuentes de sonido comprimido en la mezcla envolvente, el desalineamiento de las muestras de audio comprimido puede provocar aberraciones de compresión, lo que a su vez puede provocar desajustes de audio no deseados en la mezcla final de audio envolvente.

Una implementación típica comprenderá un aparato de codificación de acuerdo con una realización de la invención en un extremo de un enlace de comunicaciones, y un aparato de decodificación de acuerdo con una realización de la invención en el otro extremo. Tales pares de sistema pueden ser repetidos a través de múltiples enlaces de comunicaciones, si se requiere.

El método descrito anteriormente puede ser llevado a cabo mediante cualquier hardware adaptado o diseñado adecuadamente. Es posible asimismo realizar porciones del método en un conjunto de instrucciones, almacenadas en un medio legible por ordenador, que cuando se cargan en un ordenador, un procesador de señal digital (DSP) o similar, hacen que el ordenador lleve a cabo el método descrito anteriormente en esta memoria.

Igualmente, el método puede ser realizado como un hardware programado especialmente o diseñado, un circuito integrado que opera para llevar a cabo el método sobre los datos de audio cargados en el citado circuito integrado. El circuito integrado puede estar formado como parte de un dispositivo informático de propósito general, tal como un PC y otros, o puede estar formado como parte de un dispositivo más especializado, tal como una consola de juegos, un teléfono móvil, un dispositivo informático portátil o un codificador/decodificador de audio/video de hardware.

Una realización hardware de ejemplo es la de una matriz de puertas programable en campo (FPGA – Field Programmable Gate Array, en inglés) programada para llevar a cabo el método descrito y/o proporcionar el aparato descrito, estando la FPGA situada en una placa hija de un servidor de video montado en un bastidor que se encuentra en un centro de datos, para su utilización, por ejemplo, en un sistema de televisión IPTV y/o, un estudio de televisión, o una furgoneta de enlace ascendente de video de localización que da soporte a un equipo de informativos sobre el terreno.

Otra realización hardware de ejemplo de la presente invención es la de un emisor de audio y video, que comprende un par de transmisor y receptor, en la que el transmisor comprende el aparato de codificación y el receptor comprende el aparato de decodificación, en la que cada aparato de codificación está realizado como un Circuito integrado específico para una aplicación (ASIC – Application Specific Integrated Circuit, en inglés).

Resultará evidente para un experto en la materia que el orden exacto y el contenido de las etapas llevadas a cabo en el método descrito en esta memoria pueden ser alterados de acuerdo con los requisitos de un conjunto particular de parámetros de ejecución, tal como la velocidad de codificación y otros. Además, resultará evidente que diferentes realizaciones del aparato descrito pueden implementar selectivamente ciertas características de la presente invención en diferentes combinaciones, de acuerdo con los requisitos de una implementación particular de la invención como conjunto. De acuerdo con ello, la numeración de las reivindicaciones no debe ser considerada como una limitación estricta de la capacidad de mover características entre reivindicaciones, y como tales porciones de reivindicaciones dependientes pueden ser utilizadas libremente.

Apéndice:

Existe otro método de codificación de audio y que incluye el citado audio codificado en un flujo de transporte digital, que comprende: recibir en la entrada de un codificador una pluralidad de señales situadas en el mismo momento del

tiempo; asignar marcas de tiempo idénticas por unidad de tiempo a todas de la pluralidad de señales de audio situadas en el mismo momento del tiempo; e incorporar las señales de audio con marcas de tiempo idénticas en el flujo de transporte digital.

- 5 La etapa de recibir puede comprender, además: muestrear las señales situadas en el mismo momento del tiempo para formar tramas de datos de audio de un tamaño predeterminado; y alinear las citadas tramas de datos de audio para mantener el que las señales de audio se encuentren en el mismo momento del tiempo; y donde la etapa de asignar marcas de tiempo idénticas se lleva a cabo en las tramas alineadas de los datos de audio.

- 10 El método puede comprender, además: comprimir las tramas alineadas de datos de audio con un codificador de audio idéntico; ajustes de codificación antes de asignar las marcas de tiempo; y asignar los datos de audio comprimidos y con marcas de tiempo idénticas a una pluralidad de canales mono de un flujo de transporte.

La pluralidad de canales mono puede comprender uno o más componentes de audio dual mono convencionales.

El tamaño predeterminado puede ser el tamaño de una unidad de acceso en el estándar MPEG, y el flujo de transporte de video puede ser un flujo de transporte MPEG-1 o MPEG-2. Las marcas de tiempo pueden ser marcas de tiempo de presentación.

- 15 La etapa de incorporar el audio en un flujo de video digital puede comprender: multiplexar los datos de audio comprimidos y con marcas de tiempo idénticas en un flujo de transporte.

- 20 Existe además un método de descodificación de un flujo de transporte digital que incluye audio codificado de acuerdo con cualquiera de lo anterior, comprendiendo además el método: recibir una pluralidad de señales de audio con marcas de tiempo idénticas, representativas de una pluralidad de canales de audio individuales situados en el mismo momento del tiempo; detectar las marcas de tiempo para determinar marcas de tiempo compartidas; y producir la pluralidad de canales de audio individuales situados en el mismo momento del tiempo de acuerdo con las marcas de tiempo detectadas como múltiples canales.

- 25 La pluralidad de señales de audio con marcas de tiempo idénticas pueden ser muestreadas y alineadas para formar tramas alineadas de datos de audio, y donde las marcas de tiempo idénticas pueden ser aplicadas a las tramas alineadas de datos de audio.

Las tramas alineadas de datos de audio pueden ser comprimidas antes de la asignación de las marcas de tiempo y el método puede comprender, además: descomprimir las tramas de datos de audio para producir las señales de audio individuales para su emisión.

- 30 La etapa de producir la pluralidad de canales de audio individuales situados en el mismo momento del tiempo puede comprender presentar el audio utilizando la marca de tiempo de una sola de las señales de audio situadas en el mismo momento del tiempo.

El flujo de transporte digital puede ser un flujo de transporte de video digital, y las tramas alineadas de datos de audio pueden comprender paquetes PES.

- 35 Existe además un aparato de codificación adaptado para llevar a cabo cualquiera de los métodos de codificación descritos anteriormente.

Existe además un aparato de descodificación adaptado para llevar a cabo cualquiera de los métodos de descodificación descritos anteriormente.

- 40 Existe además un sistema de transporte digital que comprende: al menos un aparato de codificación tal como el descrito anteriormente; al menos un aparato de descodificación tal como el descrito anteriormente; y un enlace de comunicaciones entre ellos.

Existe además un medio legible por ordenador, que contiene instrucciones, las cuales, cuando son ejecutadas, hacen que la lógica del ordenador lleve a cabo cualquiera de los métodos descritos anteriormente.

REIVINDICACIONES

1. Procedimiento de codificación de audio y de la inclusión del citado audio codificado en un flujo de transporte digital, que comprende:
5 recibir, en una entrada de codificador, de una pluralidad de señales de audio situadas en el mismo momento del tiempo;
muestrear de la pluralidad de señales de audio situadas en el mismo momento del tiempo para formar una pluralidad de tramas alineadas de datos de audio de un tamaño predeterminado, en el cual tramas alineadas de datos de audio corresponden a un mismo periodo de tiempo;
comprimir la pluralidad de tramas alineadas de datos de audio para crear tramas comprimidas;
10 asignar marcas de tiempo idénticas por unidad de tiempo en las tramas comprimidas; e
incorporar las tramas comprimidas en una pluralidad de flujos elementales del flujo de transporte digital.
2. Procedimiento de acuerdo con la reivindicación 1, en el cual la compresión comprende, además:
comprimir la pluralidad de tramas alineadas de datos de audio con ajustes de configuración idénticos del codificador de audio antes de la asignación de las marcas de tiempo idénticas; y
15 asociar la pluralidad de tramas alineadas a una pluralidad de canales mono del flujo de transporte digital.
3. Procedimiento de acuerdo con la reivindicación 2, en el cual la pluralidad de canales mono comprende uno o varios componentes de audio dual mono convencionales.
4. Procedimiento de acuerdo con la reivindicación 1, en el cual el tamaño predeterminado es el tamaño de una unidad de acceso en el estándar MPEG, y el flujo de transporte de video es un flujo de transporte MPEG-1 o MPEG-2.
20
5. Procedimiento de acuerdo con la reivindicación 1, en el cual las marcas de tiempo son marcas de tiempo de presentación.
6. Procedimiento de acuerdo con la reivindicación 1, en el cual la etapa de incorporación comprende, además:
multiplexar los datos de audio de marcas de tiempo idénticas en el flujo de transporte digital.
- 25 7. Procedimiento de decodificación de un flujo de transporte digital, que comprende:
recibir un flujo de transporte digital que incluye audio codificado;
obtener, a partir de una pluralidad de flujos elementales del flujo de transporte digital, tramas comprimidas;
descomprimir las tramas comprimidas para crear una pluralidad de tramas alineadas de datos de audio de un tamaño predeterminado representativas de una pluralidad de canales de audio individuales situados en el mismo
30 momento del tiempo, en el cual las tramas alineadas de datos de audio corresponden a un mismo periodo de tiempo;
detectar las marcas de tiempo de cada trama de datos de audio entre la pluralidad de tramas de datos de audio para determinar tramas con marcas de tiempo idénticas de datos de audio; y
presentar las tramas con marcas de tiempo idénticas de datos de audio en tiempos idénticos mediante la utilización de las marcas de tiempo de las tramas de datos de audio entre la pluralidad de tramas de datos de audio que son representativas de una trama de audio individual entre la pluralidad de canales de audio individuales situados en el mismo momento del tiempo.
35
8. Procedimiento de acuerdo con la reivindicación 7, en el cual el audio codificado ha sido muestreado y alineado para formar la pluralidad de tramas alineadas de datos de audio y en el cual las marcas de tiempo idénticas han sido aplicadas a la pluralidad de tramas alineadas de datos de audio.
- 40 9. Procedimiento de acuerdo con la reivindicación 7, en el cual el flujo de transporte digital es un flujo de transporte de video digital, y la pluralidad de tramas alineadas de datos de audio comprende paquetes de PES.
10. Codificador para la codificación de un audio y la inclusión del citado audio en un flujo de transporte digital, comprendiendo del codificador:
45 un procesador;

un soporte de memoria legible por ordenador no transitorio que comprende además instrucciones legibles por ordenador, las cuales, cuando son ejecutadas por el procesador, están configuradas para efectuar:

la recepción, en una entrada, de una pluralidad de señales de audio situadas en el mismo momento del tiempo;

5 el muestreo de la pluralidad de señales de audio situadas en el mismo momento del tiempo para formar una pluralidad de tramas alineadas de datos de audio de un tamaño predeterminado, en el cual tramas alineadas de datos de audio corresponden a un mismo periodo de tiempo;

la compresión de la pluralidad de tramas alineadas de datos de audio para crear tramas comprimidas;

la asignación de marcas de tiempo idénticas por unidad de tiempo a las tramas comprimidas; y

la incorporación de las tramas comprimidas en una pluralidad de flujos elementales del flujo de transporte digital.

10 11. Codificador de acuerdo con la reivindicación 10, en el que las instrucciones legibles por ordenador configuradas para efectuar la compresión, cuando son ejecutadas por el procesador, están además configuradas para efectuar:

la compresión de la pluralidad de tramas alineadas de datos de audio con ajustes de configuración idénticos del codificador de audio antes de la asignación de las marcas de tiempo idénticas; y

15 la asignación de la pluralidad de tramas alineadas de datos de audio a una pluralidad de canales mono del flujo de transporte digital.

12. Codificador de acuerdo con la reivindicación 11, en el cual la pluralidad de canales mono comprenden uno o varios componentes de audio mono dual convencionales.

20 13. Codificador de acuerdo con la reivindicación 10, en el cual el tamaño predeterminado es el tamaño de una unidad de acceso en el estándar MPEG, y el flujo de transporte de video es un flujo de transporte MPEG-1 o MPEG-2.

14. Codificador de acuerdo con la reivindicación 10, en el cual las marcas de tiempo son marcas de tiempo de presentación.

25 15. Codificador de acuerdo con la reivindicación 10, en el cual las instrucciones legibles por ordenador, cuando son ejecutadas por el procesador, están además configuradas para efectuar la incorporación del audio en un flujo de video digital mediante:

multiplexación de la pluralidad de tramas alineadas de datos de audio en el flujo de transporte digital.

16. Descodificador para la descodificación de un flujo de transporte digital, que comprende:

un procesador;

30 un soporte de memoria legible por ordenador transitorio que comprende además instrucciones legibles por ordenador, las cuales, cuando son ejecutadas por el procesador, están configuradas para efectuar:

la recepción del flujo de transporte digital que comprende un audio codificado; la obtención, a partir de una pluralidad de flujos elementales del flujo de transporte digital, de tramas comprimidas;

35 la descompresión de las tramas comprimidas para crear una pluralidad de tramas alineadas de datos de audio de un tamaño predeterminado representativas de una pluralidad de canales de audio individuales situados en el mismo momento del tiempo, en el cual las tramas alineadas de datos de audio corresponden a un mismo periodo de tiempo;

la detección de las marcas de tiempo de cada trama de datos de audio entre la pluralidad de tramas alineadas de datos de audio para determinar tramas con marcas de tiempo idénticas de datos de audio; y

40 la presentación de tramas de marca de tiempo idénticas de datos de audio en momentos idénticos mediante la utilización de las marcas de tiempo de tramas de datos de audio entre la pluralidad de tramas de datos de audio que son representativas de un canal de audio individual entre la pluralidad de canales de audio individuales situados en el mismo momento del tiempo.

17. Descodificador de acuerdo con la reivindicación 16, en el cual el flujo de transporte digital es un flujo de transporte de video digital, y la pluralidad de tramas alineadas de datos de audio comprende paquetes de PES.

45

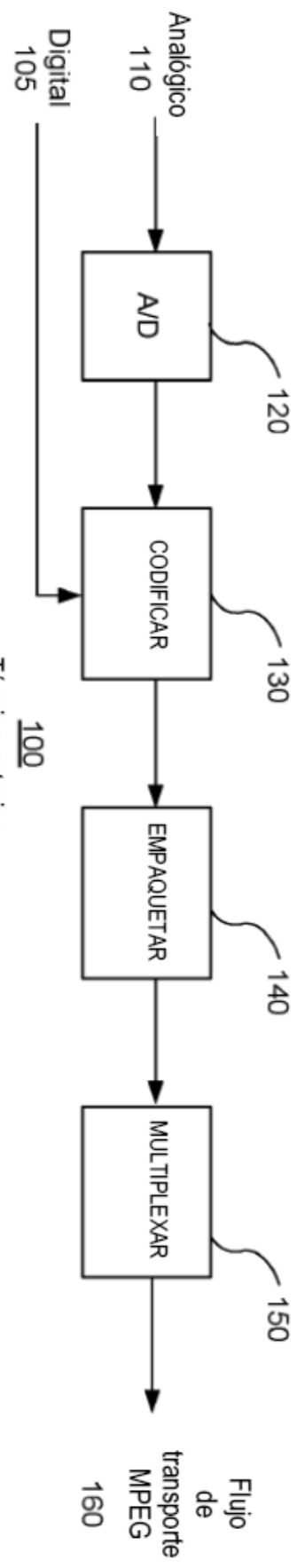


FIG. 1

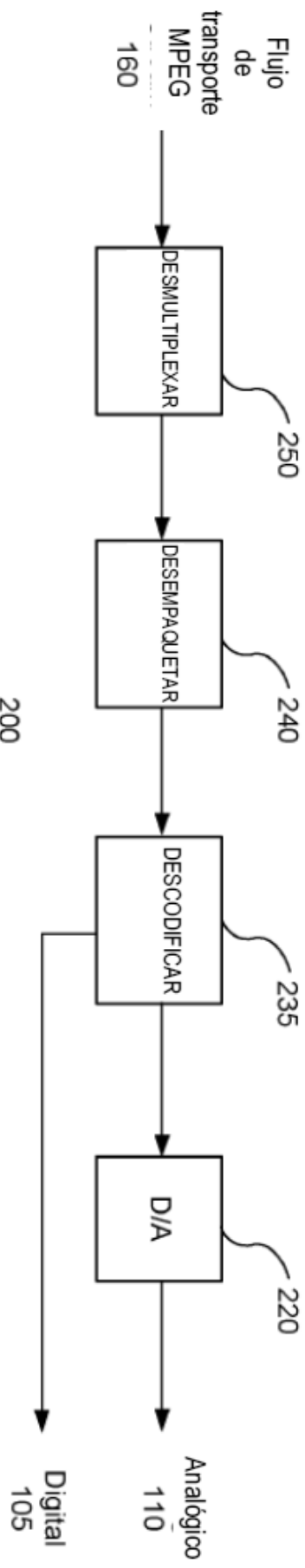


FIG. 2

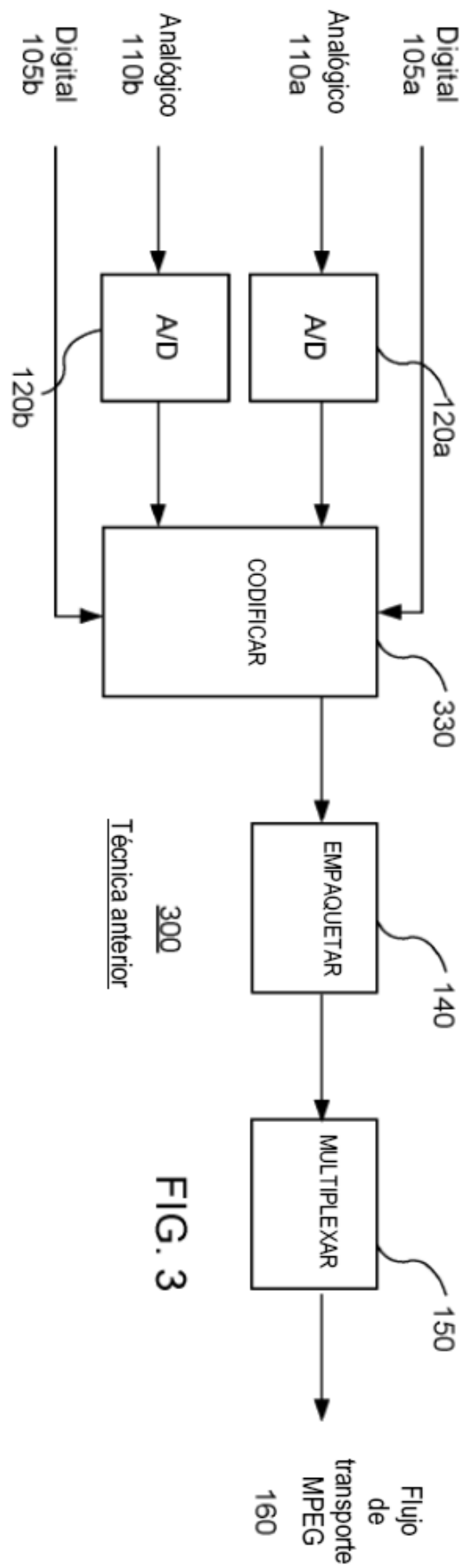


FIG. 3

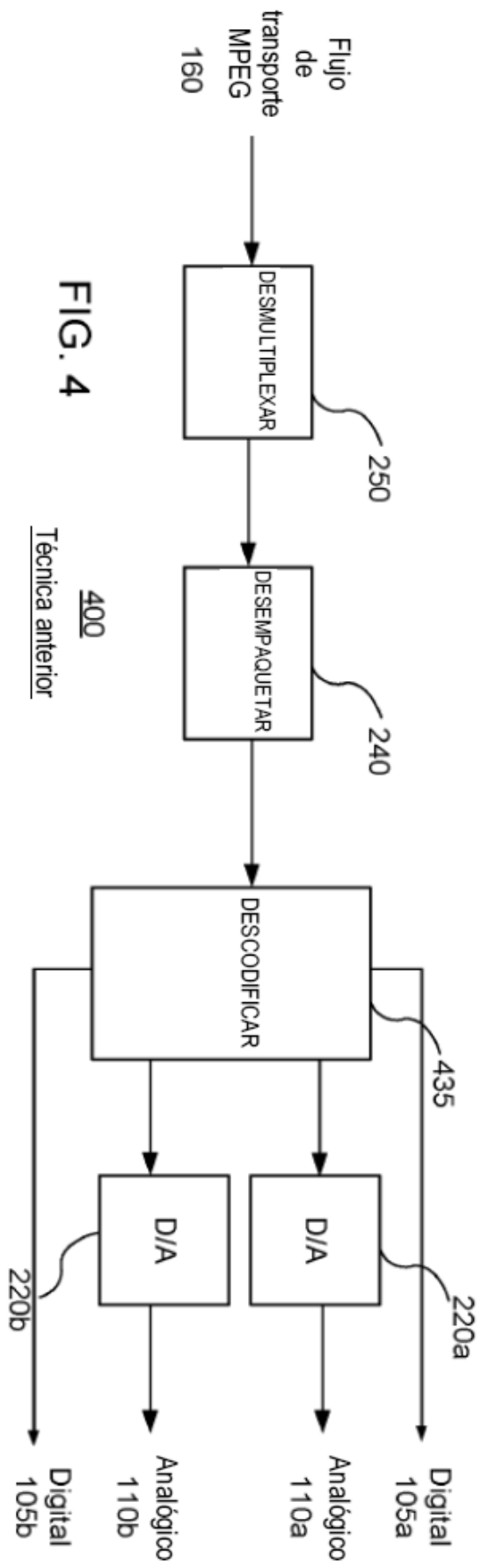


FIG. 4



FIG. 5

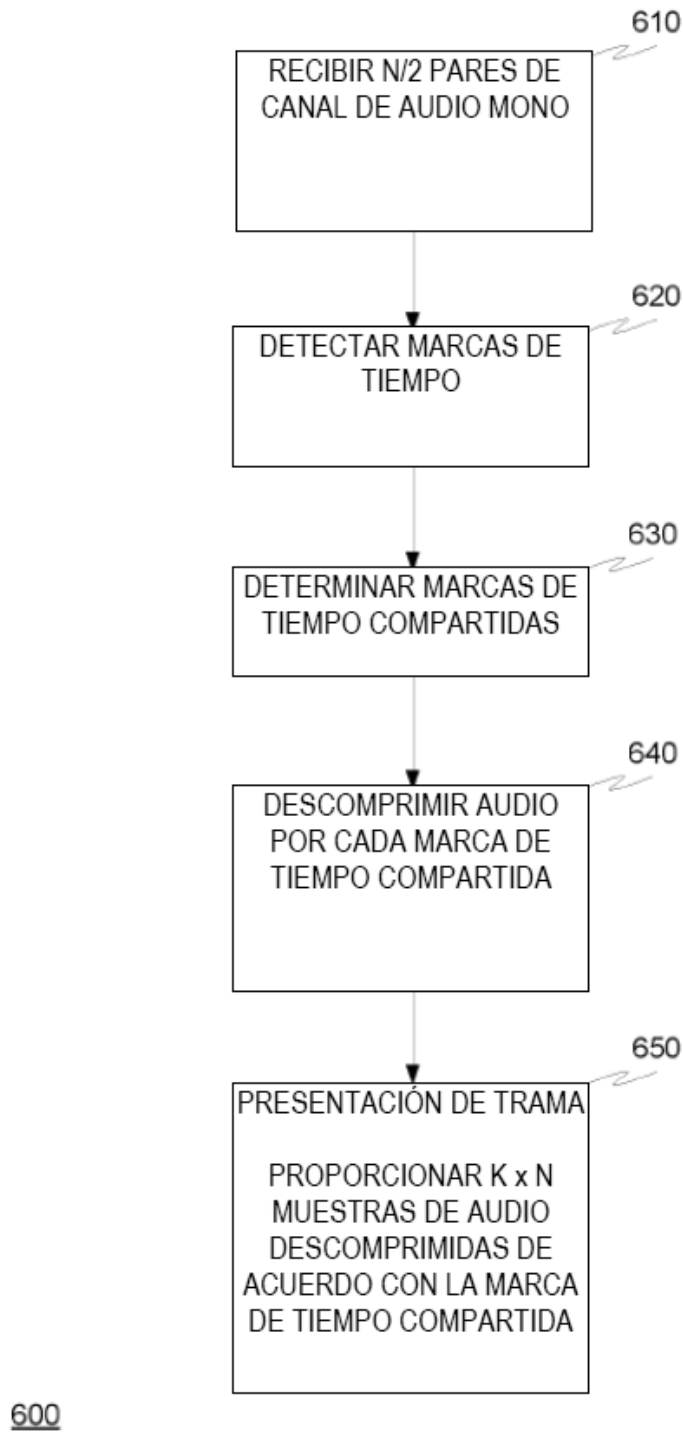


FIG. 6

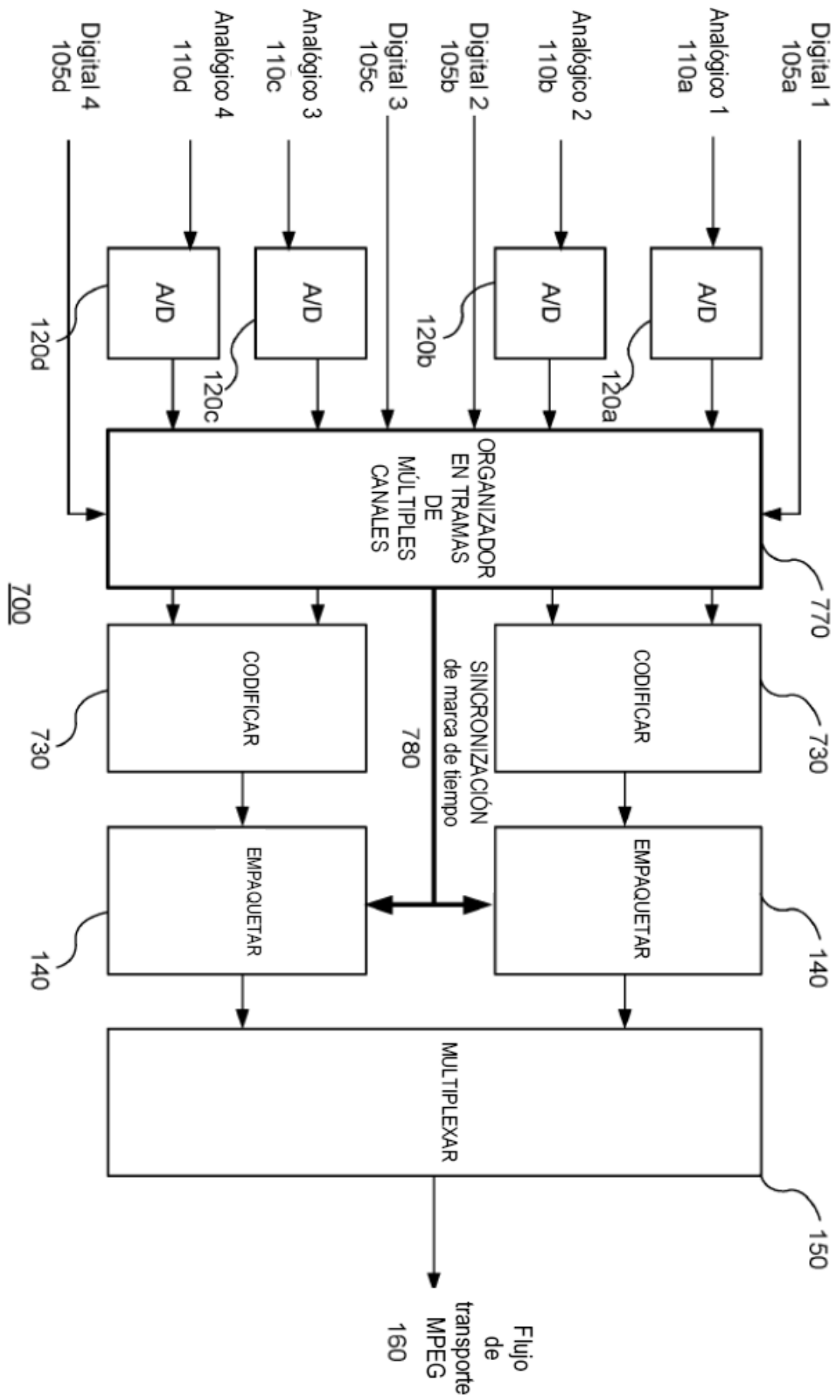


FIG. 7

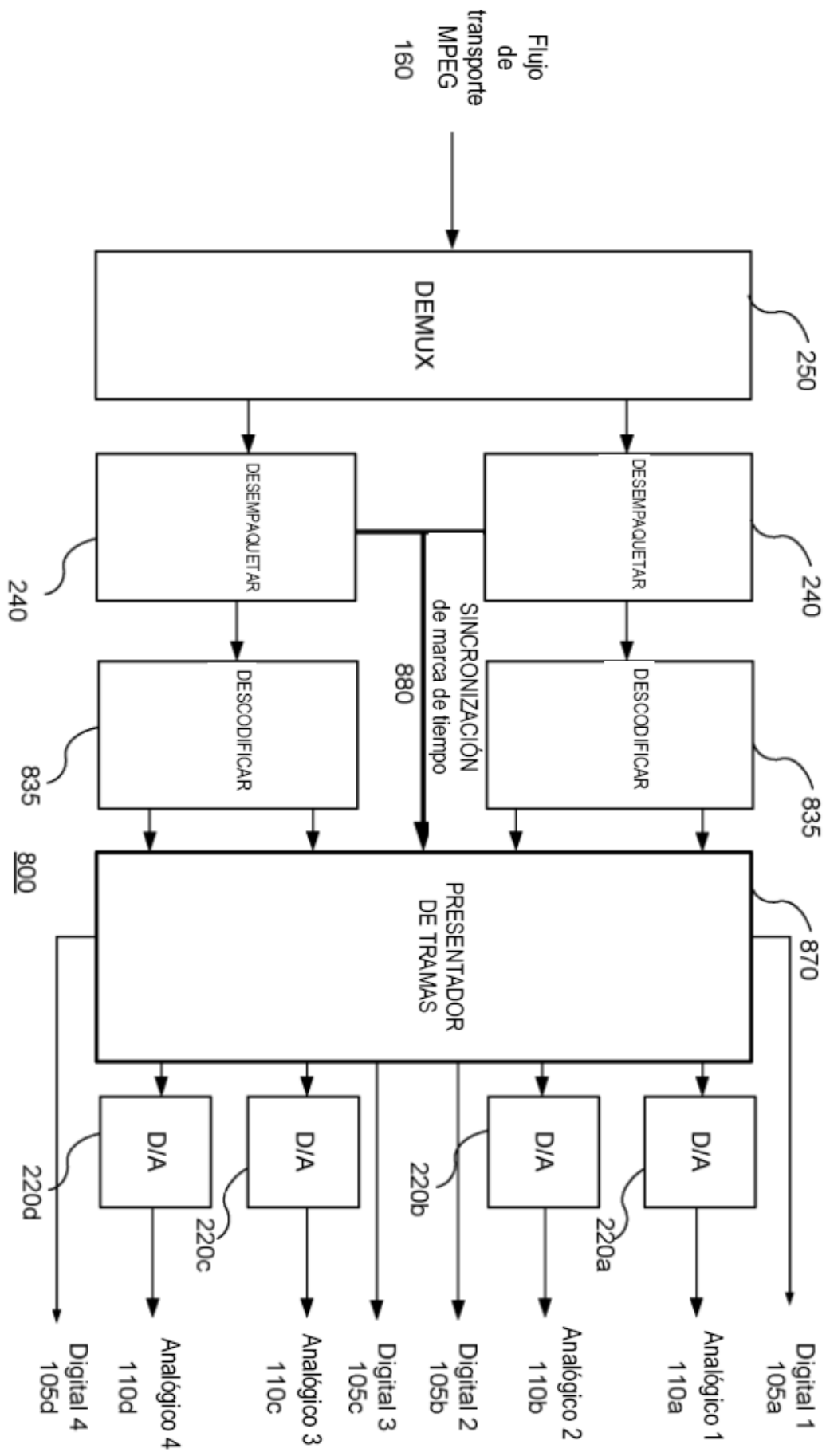


FIG. 8