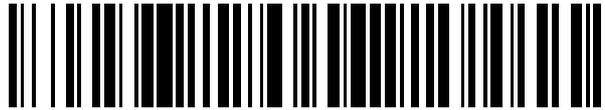


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 796 498**

51 Int. Cl.:

H04L 12/841 (2013.01)

H04L 12/811 (2013.01)

H04L 12/853 (2013.01)

H04L 12/927 (2013.01)

H04J 3/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.09.2016 PCT/KR2016/009934**

87 Fecha y número de publicación internacional: **09.03.2017 WO17039421**

96 Fecha de presentación y número de la solicitud europea: **05.09.2016 E 16842387 (9)**

97 Fecha y número de publicación de la concesión europea: **20.05.2020 EP 3343851**

54 Título: **Procedimiento y dispositivo para regular un retardo de reproducción**

30 Prioridad:

04.09.2015 KR 20150125875

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.11.2020

73 Titular/es:

**SAMSUNG ELECTRONICS CO., LTD. (50.0%)
129, Samsung-ro, Yeongtong-gu
Suwon-si, Gyeonggi-do 16677, KR y
INDUSTRY-UNIVERSITY COOPERATION
FOUNDATION HANYANG UNIVERSITY ERICA
CAMPUS (50.0%)**

72 Inventor/es:

**SUNG, HO-SANG;
KANG, SANG-WON;
KIM, JONG-HYUN;
OH, EUN-MI;
LI, YAXING y
JUNG, KYUNG-HUN**

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 796 498 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

Procedimiento y dispositivo para regular un retardo de reproducción

Campo técnico

5 La presente divulgación se refiere a la gestión de memoria intermedia de fluctuación y, más específicamente, a un procedimiento y a un aparato para ajustar un retardo de reproducción sin un aumento excesivo de complejidad y a un procedimiento y aparato de modificación de escala de tiempo.

Antecedentes de la técnica

10 Un sistema de comunicación basado en red de paquetes es apenas fiable y es inestable, ya que no se produce un caso en el que un paquete se transmite con un retardo aleatorio o se pierde. Para compensar este problema, se proporciona una función de ajuste de errores activa de solicitar una retransmisión cuando un paquete transmitido se retrasa por un valor permitido o más o se pierde, pero se produce un retardo adicional debido a esta función, y por lo tanto es difícil que esta función se aplique a un servicio de llamadas que soporta una conversación en tiempo real.

15 Por lo tanto, muchos esquemas de permitir que un codificador de voz emita de forma continua una señal de voz mediante el mantenimiento constantemente de un flujo de un paquete recibido, mientras se ha desarrollado una reducción de un retardo de paquetes y/o una pérdida de paquetes. El esquema más eficiente en un entorno en el que una característica de red varía a lo largo del tiempo entre estos esquemas puede ser un esquema de gestión de memoria intermedia de fluctuación (en lo sucesivo, citada como JBM) capaz de ajustar un retardo de reproducción no solo en un intervalo de silencio, sino también en un intervalo de voz utilizando el procesamiento de modificación de escala de tiempo (en adelante, denominado como TSM). La necesidad de reducir un retardo de reproducción o corregir un error de paquetes para cumplir con varias circunstancias de la red es el aumento significativo en el esquema de JBM.

20 "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Codec for Enhanced Voice Services EVS Codec Jitter Buffer Management (Release 12)", 3GPP BORRADOR; 1 de agosto de 2014, XP050806026, divulga la solución de gestión de la memoria intermedia de fluctuaciones, para el códec para servicios de voz mejorados (EVS). Se requieren memorias intermedias de fluctuación en las comunicaciones basadas en paquetes, tal como MTSI 3GPP, para suavizar la fluctuación entre llegadas de paquetes de medios entrantes para una reproducción ininterrumpida.

Otro ejemplo de la técnica anterior se puede encontrar en el documento WO2005/045830.

Descripción de realizaciones

30 Problema técnico

Se proporcionan un procedimiento y un aparato para ajustar un retardo de reproducción mediante el uso de una memoria intermedia de fluctuación o un módulo de modificación de escala de tiempo (TSM) sin un aumento excesivo de complejidad y un procedimiento y un aparato de modificación de escala de tiempo.

Solución al problema

35 De acuerdo con un aspecto de la presente divulgación, un procedimiento de ajuste de retardo de reproducción incluye: ajustar un excedente de retardo de reproducción en base a un valor de diferencia entre un primer retardo de reproducción obtenido en un primer esquema y un segundo retardo de reproducción obtenido en un segundo esquema, y determinar un tipo de adaptación de una trama actual según si una trama anterior es una trama activa; y cuando el tipo de adaptación determinado es una adaptación basada en señales, realizar una modificación de escala de tiempo (TSM) de acuerdo con un esquema de adaptación determinado de acuerdo con un resultado de comparación entre el primer retardo de reproducción y el segundo retardo de reproducción y un resultado de comparación entre un retardo objetivo y el primer retardo de reproducción.

El primer esquema se puede basar en un códec de servicio de voz mejorada (EVS), y el segundo esquema puede basarse en un modelo E.

45 La realización de la TSM puede incluir: cuando una diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es mayor que un umbral, determinar una reducción de escala de tiempo como el esquema de adaptación; y cuando la diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es menor que el umbral, determinar la ampliación o la reducción de la escala de tiempo como el esquema de adaptación basado en una diferencia entre el retardo objetivo y el primer retardo de reproducción.

50 De acuerdo con un aspecto de la presente divulgación, un procedimiento de modificación de escala de tiempo (TSM) incluye: determinar un rango de búsqueda reducido para una trama actual basado en una ubicación óptima de una trama pasada; determinar si buscar un rango restante excluyendo el rango de búsqueda reducido; estimar una similitud con respecto al rango de búsqueda reducido y al rango restante y buscar ubicaciones óptimas que tengan una similitud

máxima; determinar, como una ubicación óptima final, una de las ubicaciones óptimas recuperadas en el rango de búsqueda reducido y la ubicación óptima recuperada en el rango restante; y realizar una reducción o ampliación de la escala de tiempo en función de la ubicación óptima final.

5 De acuerdo con un aspecto de la presente divulgación, un aparato de ajuste de retardo de reproducción incluye: una unidad de control de adaptación configurada para ajustar un excedente de retardo de reproducción en base a un valor de diferencia entre un primer retardo de reproducción obtenido en un primer esquema y un segundo retardo de reproducción obtenido en un segundo esquema y determinar un tipo de adaptación de una trama actual según si una trama anterior es una trama activa; una unidad de memoria intermedia de fluctuación configurada para almacenar tramas para decodificar, y para realizar la adaptación en las tramas almacenadas cuando el tipo de adaptación
10 determinado es la adaptación basada en trama; una unidad de decodificación configurada para decodificar una trama proporcionada desde la memoria intermedia de fluctuación; y una unidad de aplicación de modificación de escala de tiempo (TSM) configurada para realizar la adaptación en la trama decodificada cuando el tipo de adaptación determinado es una adaptación basada en el tiempo.

15 La unidad de aplicación de TSM puede estar configurada además para determinar la reducción de escala de tiempo como un esquema de adaptación cuando una diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es mayor que un umbral y para determinar la escala de tiempo de ampliación o reducción como el esquema de adaptación basado en una diferencia entre un retardo objetivo y el primer retardo de reproducción cuando la diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es menor que el umbral.

20 De acuerdo con un aspecto de la presente divulgación, un aparato de modificación de escala de tiempo (TSM) incluye un aparato para determinar un rango de búsqueda reducido para una trama actual basado en una ubicación óptima de una trama pasada, determinar si buscar un rango restante excluyendo el rango de búsqueda reducido, estimar una similitud con respecto al rango de búsqueda reducido y al rango restante, buscar ubicaciones óptimas que tengan una similitud máxima, determinar, como una ubicación óptima final, una de las ubicaciones óptimas recuperadas en el
25 rango de búsqueda reducido y la ubicación óptima recuperada en el rango restante, y realizar una reducción o ampliación de la escala de tiempo en función de la ubicación óptima final.

La invención se define mediante las reivindicaciones adjuntas.

Efectos ventajosos de la divulgación

La reducción de retardo de reproducción o la corrección de errores de paquetes se puede realizar para cumplir diversas circunstancias de red sin un aumento excesivo de complejidad.

Breve descripción de los dibujos

La figura 1 es un diagrama de bloques de un aparato de ajuste de retardo de reproducción de acuerdo con una realización.
La figura 2 es un diagrama de bloques de un aparato de ajuste de retardo de reproducción de acuerdo con otra realización.
35 La figura 3 es un esquema de flujo para describir una operación de una unidad de control de adaptación que se muestra en la figura 1.
La figura 4 muestra un ejemplo en el que se puede seleccionar un factor de ajuste de tiempo según se seleccione el ajuste.
La figura 5 ilustra un procesamiento de reducción de la escala de tiempo.
40 La figura 6 ilustra un procesamiento de aumento de la escala de tiempo.
La figura 7 es un diagrama de bloques que ilustra un esquema adicional de reducción de retardo de reproducción cuando se determina la reducción de retardo de reproducción mediante modelo E.
La figura 8 es un diagrama de bloques de un aparato de modificación de escala de tiempo (TSM) de acuerdo con una realización.
45 La figura 9 es un diagrama de bloques de una unidad de estimación de similitud mostrada en la figura 8.
La figura 10 es un diagrama de bloques de un dispositivo multimedia de acuerdo con una realización.
La figura 11 es un diagrama de bloques de un dispositivo multimedia de acuerdo con otra realización.

Modo de divulgación

50 La presente divulgación puede permitir que diversos tipos de cambio o modificación y diversos cambios de forma, y realizaciones específicas se ilustrarán en los dibujos y se describen en detalle en la descripción detallada. Sin embargo, debe entenderse que las realizaciones específicas no limitan la presente divulgación a una forma de divulgación específica, sino que incluyen todas las modificadas, equivalentes o reemplazadas dentro del espíritu técnico y del alcance técnico de la presente divulgación. En la descripción de las realizaciones, cuando se determina que una descripción específica de características bien conocidas relevantes puede ocultar los elementos esenciales
55 de la presente divulgación, se omite una descripción detallada de la misma.

Aunque los términos, tales como 'primero' y 'segundo', se pueden utilizar para describir diversos elementos, los elementos pueden no estar limitados por los términos. Los términos se pueden usar para clasificar un cierto elemento

de otro elemento.

Los términos utilizados en esta memoria descriptiva son aquellos términos generales ampliamente utilizados en la técnica en la actualidad, pero los términos pueden variar de acuerdo con la intención de los expertos ordinarios en la técnica, los precedentes, o una nueva tecnología en la técnica. Además, el solicitante puede seleccionar términos específicos y, en este caso, el significado detallado de los mismos se describirá en la descripción detallada. Por lo tanto, los términos utilizados en la memoria descriptiva deben entenderse no como nombres simples, sino basados en el significado de los términos y de la descripción general.

Una expresión en singular incluye una expresión en plural a menos que sean claramente diferentes entre sí en su contexto. En la presente divulgación, debe entenderse que los términos, tales como 'incluir' y 'tener', se usan para indicar la existencia de una característica implementada, número, etapa, operación, elemento, parte o una combinación de los mismos sin excluir por adelantado la posibilidad de la existencia o adición de una o más características, números, etapas, operaciones, elementos, partes o combinaciones de los mismos.

A continuación, se describirán las realizaciones en detalle con referencia a los dibujos adjuntos.

La figura 1 es un diagrama de bloques de un aparato de ajuste de retardo de reproducción de acuerdo con una realización.

El aparato mostrado en la figura 1 puede incluir una unidad 110 de desempaquetado, una unidad 120 de estimación de retardo objetivo, una memoria 130 intermedia de eliminación de fluctuación, una unidad 140 de decodificación de audio, una unidad 150 de aplicación de modificación de escala de tiempo (TSM), una unidad 160 de estimación de retardo de reproducción, una unidad 170 de control de adaptación, y una memoria 180 intermedia de receptor.

Con referencia a la figura 1, la unidad 110 de desempaquetado puede analizar un flujo de paquetes recibido a través de un canal de transmisión o un módem y extraer tramas incrustadas en un paquete y metadatos para estimar la fluctuación de red. Las tramas extraídas pueden proporcionarse a la memoria intermedia 130 de eliminación de fluctuación. Por ejemplo, el paquete puede ser un paquete de protocolo de transporte en tiempo real (RTP) que tiene un formato de carga útil RTP. Por ejemplo, las tramas pueden ser tramas codificadas por un códec de servicio de voz mejorado (EVS). Cuando un aparato de decodificación recibe un flujo de bits, la unidad 110 de desempaquetado puede incluirse opcionalmente. La unidad 110 de desempaquetado puede calcular y asignar una marca de tiempo de medios para cada trama existente en un paquete RTP recibido. El paquete RTP puede incluir un encabezado RTP y una carga útil RTP. En el encabezado RTP, un campo de datos relacionado con la memoria 130 intermedia de eliminación de fluctuación puede incluir una marca de tiempo RTP y un número de secuencia RTP. De acuerdo con una realización, una estructura de datos de cada trama extraída por la unidad 110 de desempaquetado puede incluir datos de carga útil de trama que incluyen un tamaño de carga útil, una marca de tiempo de llegada de un paquete RTP que incluye una trama correspondiente, una marca de tiempo de medios de una unidad de escala de tiempo RTP derivada de una marca de tiempo RTP del paquete, una escala de tiempo RTP definida en una especificación de formato de carga útil RTP, un indicador de descriptor de inserción de silencio (SID) y un indicador de copia parcial. Para operar la memoria 130 intermedia de eliminación de fluctuación independientemente de un códec de audio, el indicador SID debe proporcionarse a la memoria 130 intermedia de eliminación de fluctuación. Por ejemplo, para EVS, códecs adaptativos de velocidad múltiple (AMR) y de banda ancha AMR (AMR-WB), el indicador SID puede determinarse dependiendo del tamaño de los datos de carga útil de la trama.

La unidad 120 de estimación de retardo objetivo puede estimar la fluctuación de entrada y un retardo objetivo de un paquete. La unidad 120 de estimación de retardo objetivo puede estimar la fluctuación de la red para controlar un retardo de extracción de la memoria 130 intermedia de eliminación de fluctuación. El retardo objetivo puede obtenerse combinando un valor de estimación de fluctuación a largo plazo y un valor de estimación de fluctuación a corto plazo.

En lo sucesivo, se describirá un procedimiento de obtención del valor de estimación de fluctuación a largo plazo, de acuerdo con una realización.

Para cada trama recibida como un paquete RTP a través de una red, se añade una entrada a una matriz tal como una cola de primero en entrar y de primero en salir (FIFO), en el que la entrada incluye un retardo d_i , un desplazamiento o_i , y una marca de tiempo RTP t_i . En una circunstancia específica, un período de tiempo almacenado en la cola FIFO puede diferir de varias entradas almacenadas. Debido a esto, el tamaño de la ventana de la cola FIFO puede ser limitado. Por ejemplo, cuando se reciben 50 paquetes durante un segundo, la cola FIFO puede incluir un total de 500 entradas correspondientes a 10 segundos, y un intervalo de tiempo correspondiente a una diferencia de marca de tiempo RTP entre la última trama y la trama más antigua puede ser de 10 segundos. Cuando es necesario almacenar más entradas, se elimina la entrada más antigua. Aquí, la matriz puede incluir entradas válidas. En estas condiciones, el valor de estimación de fluctuación a largo plazo j_i puede obtenerse usando la Ecuación 1 a continuación.

$$j_i = \max(d_{i-500}, \dots, d_i) - \min(d_{i-500}, \dots, d_i) \quad (1)$$

A continuación, se describirá un procedimiento de obtención del valor de estimación de fluctuación a corto plazo, de acuerdo con una realización. Básicamente, este procedimiento es similar al procedimiento del valor de estimación de

la fluctuación a largo plazo.

Un tamaño de ventana de una primera matriz está limitado a un máximo de 50 entradas y un lapso de tiempo máximo de un segundo. Como resultado, el primer valor de fluctuación temporal resultante k_i se calcula como la diferencia entre un valor de retardo de percentil del 94 % d_{p94} actualmente almacenado en la primera matriz y el valor de retardo mínimo de la primera matriz usando la Ecuación 2 a continuación.

$$k_i = d_{p94} - \min(d_{i-50}, \dots, d_i) \quad (2)$$

A continuación, un segundo valor de la fluctuación temporal l_i puede generarse mediante la compensación de diferentes desplazamientos a largo plazo FIFO y de corto plazo FIFO. El segundo valor de fluctuación temporal l_i se puede obtener agregando una diferencia entre un valor de desplazamiento mínimo almacenado en la primera matriz y un valor de desplazamiento mínimo almacenado en la FIFO a largo plazo al primer valor de fluctuación temporal k_i por la Ecuación 3 a continuación.

$$l_i = k_i + \min(o_{i-50}, \dots, o_i) - \min(o_{i-500}, \dots, o_i) \quad (3)$$

El segundo valor de la fluctuación temporal l_i se añade a una segunda matriz que tiene un tamaño de ventana máximo de 200 entradas y un período de tiempo máximo de cuatro segundos.

Finalmente, un valor máximo de fluctuación de la segunda matriz puede ser redondeado como un múltiplo entero de un tamaño de trama, y como resultado, el valor de estimación de fluctuación a corto plazo m_i puede obtenerse utilizando la Ecuación 4 a continuación.

$$m_i = \lceil \max(l_{i-200}, \dots, l_i) / 20ms \rceil * 20ms \quad (4)$$

Mientras tanto, la reproducción puede estar adaptada de tal manera que una diferencia entre un retardo de reproducción y un retardo objetivo se reduce al mínimo. Un retardo de transmisión de paquetes en una red puede incluir una suma de un componente fijo y un componente variable. El componente fijo puede incluir factores inevitables, tal como el tiempo de propagación a través de un medio físico y un tiempo de procesamiento mínimo, y el componente variable puede ocurrir, por ejemplo, por fluctuación de red debido a la programación. Dado que la memoria intermedia de eliminación de fluctuación no espera relojes sincronizados entre un transmisor y un receptor, el componente de retardo fijo no puede estimarse solo con información obtenible de un paquete RTP recibido, y por lo tanto el componente de retardo fijo puede ignorarse.

Mientras tanto, para estimar el componente variable de retardo, dos valores básicos, es decir, el retardo d_i y el desplazamiento o_i se pueden obtener usando una llegada marca de tiempo r_i y un medio de sello de tiempo T_i . La estimación del componente de retardo puede realizarse para cada trama. Aquí i indica una trama actual, e $i-1$ indica una trama anterior. Un retardo d_0 de una trama recibida por primera vez en una sesión puede ser 0. El retardo d_i de cada trama puede obtenerse usando la Ecuación 5 a continuación.

$$d_i = (r_i - r_{i-1}) - (t_i - t_{i-1}) + d_{i-1} \quad (5)$$

El retardo d_i puede calcularse en una unidad de milisegundo.

Una ventana deslizante puede almacenar el desplazamiento o_i y el retardo d_i en la Ecuación 6 a continuación para una trama recibida.

$$o_i = r_i - t_i \quad (6)$$

Una diferencia almacenada entre un retardo mínimo y un retardo máximo puede utilizarse como un componente de fluctuación variable, es decir, un valor de estimación de la fluctuación de la red.

El retardo objetivo y el retardo de reproducción pueden calcularse en base a un desplazamiento mínimo almacenado, es decir, un desplazamiento calculado con respecto a una trama recibida con el menor retardo de transmisión en comparación con todas las tramas actualmente incluidas en la ventana deslizante.

La unidad 120 de estimación de retardo objetivo puede estimar diferentes retardos objetivo de acuerdo con los tipos de trama y ubicaciones. Por ejemplo, se pueden estimar diferentes retardos objetivo para una trama activa, una trama inactiva y una primera trama después de inactiva.

Para estimar un retardo objetivo para una trama activa, un retardo objetivo mínimo, es decir, un umbral más bajo u_i , y un retardo objetivo máximo, es decir, un umbral superior v_i , mediante la combinación de un valor de estimación de

fluctuación a corto plazo y un valor de estimación de fluctuación a largo plazo. El umbral inferior u_i y el umbral superior v_i pueden obtenerse utilizando la Ecuación 7 a continuación.

$$\begin{aligned} v_i &= m_i + 60ms + g \\ u_i &= \min(j_i + 20ms + g + h, v_i) \end{aligned} \quad (7)$$

5 En el presente documento, g indica un retardo de trama redundante excedente y puede ser ajustado a 0, y h indica un retardo reservado se puede establecer en 15. Estas variables g y h pueden modificarse de acuerdo con las realizaciones.

Para una trama activa, cuando un retardo de reproducción supera cada umbral, la adaptación basada en la señal puede realizarse.

10 Mientras tanto, para una trama inactiva, para aumentar o disminuir un retardo de reproducción de manera que el retardo de reproducción sigue un retardo objetivo w_i para la trama inactiva, la adaptación basada en tramas se puede realizar mediante la inserción o la eliminación de una trama SIN_DATOS. El retardo objetivo w_i puede obtenerse a partir de un valor mínimo de un valor de estimación de fluctuación a largo plazo y un valor de estimación de fluctuación a corto plazo mediante la Ecuación 8 a continuación.

$$w_i = \min(j_i + h, m_i) \quad (8)$$

15 Mientras tanto, para una primera trama activa después de una trama inactiva, para aumentar un retardo de reproducción de manera que el retardo de reproducción coincide con un retardo objetivo z_i para la primera trama activa, la adaptación basada en tramas se puede realizar mediante la inserción de una trama SIN_DATOS. El retardo objetivo z_i puede obtenerse a partir de una media del umbral inferior u_i y del umbral superior v_i de una trama activa mediante la Ecuación 9 a continuación.

$$z_i = (u_i + v_i + h/4)/2 \quad (9)$$

20 La memoria 130 intermedia de eliminación de fluctuación puede almacenar las tramas extraídas por la unidad 110 de desempaquetado, para la decodificación y la reproducción. Cada valor estadístico puede actualizarse en correspondencia con las tramas almacenadas. Las tramas almacenadas en la memoria 130 intermedia de eliminación de fluctuación pueden no proporcionarse directamente a la unidad 140 de decodificación de audio, y la adaptación basada en tramas puede realizarse para suavizar la fluctuación de la red. Ejemplos de adaptación basada en tramas pueden incluir insertar una trama oculta, quitar una trama almacenada en la memoria 130 intermedia de eliminación de fluctuación y agregar o quitar una trama de ruido de confort. En detalle, cuando la existencia de una trama perdida se señala a la unidad 140 de decodificación de audio, se genera una trama oculta para la trama perdida y se proporciona a la memoria 130 intermedia de eliminación de fluctuación. Cuando la trama oculta se proporciona a la memoria 130 intermedia de eliminación de fluctuación, puede aumentar un retardo de reproducción y, por lo tanto, cuando el retardo de reproducción después de la reproducción de la última trama es mayor que un retardo objetivo, se puede eliminar una primera trama a decodificar después de insertar la trama oculta. En una circunstancia específica, por ejemplo, en un estado de transmisión discontinua (DTX), un retardo de reproducción puede aumentar insertando una trama que incluye ruido de confort. Para la adaptación basada en tramas, se puede decodificar una primera trama activa después de DTX después de decodificar una trama de ruido de confort. Cuando la adaptación se realiza insertando una trama de ruido de confort, es decir, una trama SIN_DATOS, entre dos tramas SID o entre una trama SID y una primera trama activa, la distorsión de la señal puede minimizarse. Mientras tanto, se puede eliminar una trama de ruido de confort omitiendo la decodificación de una trama SIN_DATOS. Además, se puede eliminar una trama de ruido de confort antes de una primera trama activa después de DTX para controlar un retardo de reproducción entre tramas SID.

40 Un paquete RTP puede transmitirse en una red con la fluctuación de la red, es decir, un retardo variable en el tiempo, y puede reordenarse, perderse, o copiarse. La memoria 130 intermedia de eliminación de fluctuación puede almacenar tramas incluidas en un paquete RTP recibido desde una red y proporcionar las tramas a la unidad 150 de decodificación de audio en un orden preciso. La memoria 130 intermedia de eliminación de fluctuación puede tener una estructura de memoria intermedia de anillo que tiene una capacidad fija. Para evitar un retardo excesivo y el uso de la memoria en un entorno específico, la memoria 130 intermedia de eliminación de fluctuación puede tener una capacidad asignada para almacenar datos de audio activos de tres segundos, es decir, hasta 150 entradas, cuando una longitud de trama después de la inicialización es de 20 ms. En un caso de desbordamiento, la trama más antigua puede eliminarse de la memoria 130 intermedia de eliminación de fluctuación. La profundidad de la memoria 130 intermedia de eliminación de fluctuación es dinámica y puede ser controlada por la unidad 170 de control de adaptación.

La unidad 140 decodificación de audio puede decodificar tramas proporcionadas desde la memoria 130 intermedia de

eliminación de fluctuación a datos de modulación de código de pulso (PCM). Por ejemplo, se puede usar un códec EVS para la decodificación.

La unidad 150 de aplicación de TSM puede llevar a cabo la adaptación basada en la señal para cambiar un retardo de reproducción para una señal decodificada por la unidad 140 de audio de decodificación, es decir, los datos PCM.

5 La unidad 150 de aplicación de TSM puede realizar TSM para reducir el tiempo o ampliar el tiempo de la señal decodificada por la unidad 140 de decodificación de audio, en respuesta a un modo de adaptación determinado por la unidad 170 de control de adaptación. La unidad 150 de aplicación de TSM puede generar muestras adicionales para aumentar un retardo de reproducción o puede eliminar muestras de la señal decodificada por la unidad 140 de decodificación de audio para disminuir un retardo de reproducción.

10 La unidad 160 de estimación de retardo de reproducción puede generar información con respecto a un retardo de reproducción actual debido a la memoria 130 intermedia de eliminación de fluctuación. Se puede calcular un retardo de reproducción cada vez que se activa la unidad 140 de decodificación de audio independientemente de si la unidad 140 de decodificación de audio decodifica una trama recibida, oculta una trama perdida o genera ruido de confort para una trama inactiva.

15 La unidad 170 de control de adaptación puede seleccionar un tipo basado en tramas adaptación o un tipo de adaptación basada en el tiempo basado en un retardo objetivo y un retardo de reproducción obtenido de acuerdo con un estado y la ubicación de una trama y determinar un modo de adaptación en cada tipo de adaptación.

20 La memoria 180 intermedia de receptor puede almacenar temporalmente datos PCM proporcionados desde la unidad 150 de aplicación TSM y la salida de los datos PCM en un tamaño de trama fija. La memoria 180 intermedia de receptor puede incluir una cola FIFO para datos PCM. Cuando se realiza la adaptación basada en la señal, la unidad 150 de aplicación de TSM no genera una trama de una longitud fija, por ejemplo, 20 ms, y, por lo tanto, la memoria 180 intermedia de receptor se utiliza para emitir datos PCM de una longitud fija.

25 La figura 2 es un diagrama de bloques de un aparato de ajuste de retardo de reproducción de acuerdo con una realización. De acuerdo con una realización, el aparato de ajuste de retardo de reproducción puede incluir un módulo 210 de estimación de retardo de reproducción y un módulo 230 de control de adaptación. El módulo 210 de estimación de retardo de reproducción y el módulo 230 de control de adaptación pueden corresponder a la unidad 160 de estimación de retardo de reproducción y la unidad 170 de control de adaptación en la figura 1.

30 En la figura 2, el módulo 210 de estimación de retardo de reproducción puede incluir una primera unidad 211 de estimación de retardo de reproducción y una segunda unidad 213 de estimación de retardo de reproducción, y el módulo 230 de control de adaptación puede incluir una unidad 231 de control de adaptación, una unidad 233 de determinación de reducción de retardo de reproducción y una unidad 235 de reducción de retardo de reproducción.

35 Con referencia a la figura 2, la primera unidad 211 de estimación de retardo de reproducción puede generar información con respecto a un retardo de reproducción actual debido a la memoria 130 intermedia de eliminación de fluctuación. Un retardo de reproducción anterior p_{k-1} puede indicar un retardo de memoria intermedia entre un tiempo de reproducción anterior $k-1$ en el que se reprodujo una trama y un momento en que se pudo recibir la trama. El retardo de reproducción anterior p_{k-1} se puede obtener utilizando la Ecuación 10 a continuación.

$$p_{k-1} = q_{k-1} - \min(o_{i-500}, \dots, o_i) + b_k \quad (10)$$

40 En este documento, b_k indica un intervalo de muestras en memoria intermedia en la memoria 180 intermedia de receptor en un tiempo de reproducción k , y $\min(o_{i-500}, \dots, o_i)$ indica un desplazamiento mínimo almacenado en un FIFO a largo plazo independientemente de qué trama se esté reproduciendo actualmente. Un desplazamiento de reproducción anterior q_{k-1} puede recalcularse usando una hora actual del sistema s_{k-1} y una marca de tiempo RTP t_{k-1} de una salida de trama desde la memoria 130 intermedia de eliminación de fluctuación cada vez que se emite una trama recibida desde la memoria 130 intermedia de eliminación de fluctuación. El desplazamiento de reproducción anterior q_{k-1} se puede obtener utilizando la Ecuación 11 a continuación.

$$q_{k-1} = s_{k-1} - t_{k-1} \quad (11)$$

45 Puesto que el desplazamiento de reproducción anterior q_{k-1} utiliza una marca de tiempo RTP, la reproducción anterior q_{k-1} no se puede calcular para una trama oculta o ruido de confort, y en este caso, un valor de estimación puede utilizarse. Se puede estimar que un desplazamiento de reproducción actual q_k puede ser el mismo que el desplazamiento de reproducción anterior q_{k-1} . El desplazamiento de reproducción actual q_k puede actualizarse cada vez que se realiza una adaptación basada en tramas, en la que se agrega un intervalo de tramas a insertar al desplazamiento de reproducción actual q_k , y un intervalo de tramas a eliminar se deduce del desplazamiento de reproducción actual q_k .

50 La segunda unidad 213 de estimación de retardo de reproducción puede estimar un retardo de reproducción basado en modelo E definido por la Unión Internacional de Telecomunicaciones - Sector de Normalización de

Telecomunicación (UIT-T). Es decir, se puede estimar un retardo de reproducción de tal manera que se maximice un valor de evaluación de calidad de sonido, es decir, un factor de clasificación de transmisión R, utilizado en el modelo E, que puede estar representado por la Ecuación 12.

$$R = R_0 - I_s - I_d - I_{e,eff} + A \quad (12)$$

5 En el presente documento, R_0 indica una señal de base a la proporción de ruido, I_s indica un factor de degradación simultánea, I_D designa un factor de degradación de retardo, $I_{e,eff}$ indica un factor de degradación del equipo, y A indica un factor de ventaja. Es decir, I_d e $I_{e,eff}$ son factores de degradación de la calidad del sonido que ocurren de acuerdo con un retardo y pérdida de transmisión, respectivamente. Por lo tanto, se puede estimar un retardo de reproducción para minimizar I_d y $I_{e,eff}$.

10 $I_{e-eff,WB}$, que indica un factor de degradación del equipo para una señal de banda ancha, puede definirse en la Ecuación 13 a continuación.

$$I_{e-eff,WB} = I_{e,WB} + (95 - I_{e,WB}) \cdot \frac{P_{pl}}{\frac{P_{pl}}{Ráfaga} + B_{pl}} \quad (13)$$

15 En este documento, $I_{e,WB}$ indica un factor de degradación de códec considerado cuando no hay pérdida de paquetes, P_{pl} indica una tasa de pérdida de paquetes, Ráfaga indica una velocidad de ráfaga, y B_{pl} indica un factor de robustez de pérdida de paquetes de códec considerado. Es decir, $I_{e,WB}$ y B_{pl} son factores definidos de acuerdo con un códec.

La tasa de pérdida de paquetes P_{pl} puede definirse por la Ecuación 14 a continuación.

$$P_{pl} = (1 - F(d)) \quad (14)$$

En el presente documento, $F(d)$ indica una función de distribución acumulativa del retardo de paquetes.

20 Mientras tanto, I_d^{WB} , que indica un factor de degradación de retardo para una señal de banda ancha, puede definirse en la Ecuación 15 a continuación.

$$I_d^{WB} = 0,0103d + 0,1006(d - 168)H(d - 168) \quad (15)$$

En el presente documento, d indica un retardo absoluto de extremo a extremo en una conexión de paquetes y puede obtenerse por una suma de un retardo de codificación, un retardo de empaquetado, un retardo de decodificación, y un retardo de red. Mientras tanto, $H(x)$ tiene un valor de 0 cuando $x < 0$ y 1 cuando $x \geq 0$.

25 De acuerdo con una realización, basado en la Ecuación 16 a continuación, puede encontrarse un valor de retardo de transmisión que provoca una suma I_m de I_d e $I_{e,eff}$ que son factores de degradación de la calidad de sonido para ser minimizados para maximizar el valor de evaluación de la calidad de sonido R.

$$d_{\text{modelo E}}^i : I_m(d_{\text{modelo E}}^i) = \min_{d_{\min} \leq d \leq d_{\max}} I_m(d) \quad (16)$$

30 En el presente documento, d_{\min} se puede establecer en 0, y d_{\max} puede establecerse en 450, pero la presente realización no esté limitada.

35 La unidad 231 de control de adaptación puede comparar un retardo de reproducción obtenido de acuerdo con un estado y la ubicación de una trama y un retardo objetivo, y cuando una diferencia entre los mismos excede un umbral, la unidad 231 de control de adaptación puede desencadenar un tipo de adaptación y corregir el retardo de reproducción y una profundidad de la memoria 130 intermedia de retardo. Para el ruido de confort para el que está activo un indicador SID, se activa un tipo de adaptación basado en trama de modo que la adaptación 130 basada en trama puede ser realizada por la memoria 130 intermedia de eliminación de fluctuación. Mientras tanto, para un intervalo activo, se activa un tipo de adaptación basada en señal de tal manera que la unidad 150 de aplicación de TSM pueda realizar una adaptación basada en el tiempo.

40 La unidad 233 de determinación de reducción de retardo de reproducción puede comparar el retardo de reproducción estimado por la primera unidad 211 de estimación de retardo de reproducción y el retardo de reproducción estimado

por la segunda unidad 213 de estimación de retardo de reproducción y determinar un esquema de reducción de retardo de reproducción en correspondencia con el resultado de la comparación.

5 La unidad 235 de reducción de retardo de reproducción puede reducir un retardo de reproducción mediante el uso de tiempo adicional de reducción o reducción de un excedente de retardo de reproducción basado en el sistema de reducción retardo de reproducción determinada por la unidad 233 de determinación de reducción de retardo de reproducción.

La figura 3 es un esquema de flujo para describir una operación de la unidad de control de adaptación que se muestra en la figura 1.

10 Con referencia a la figura 3, un tipo de adaptación basado en trama y un tipo de adaptación basado en señal se identifican de acuerdo con una señal, y un tipo de adaptación puede determinarse comparando un retardo objetivo estimado por la unidad 115 de estimación de retardo objetivo y un retardo de reproducción estimado por la unidad 119 de estimación de retardo de reproducción. Una primera unidad de reducción (operación 321) y una segunda unidad de reducción (operación 311) son partes recién agregadas.

15 En la figura 3, en la operación 311, se puede mantener o reducir un excedente de retardo de reproducción basado en una segunda unidad de reducción (530 de la figura 5) se describirá a continuación. En detalle, cuando se cumple una condición de que se puede reducir un retardo de reproducción estimado, es decir, la ecuación 17 a continuación, o un valor mínimo en una memoria intermedia en la que las diferencias entre los retardos objetivo estimados por la estimación 120 de retardo objetivo y los retardos de reproducción estimados por el la unidad 160 de estimación de retardo de reproducción almacenados es mayor que un valor predeterminado, el excedente de retardo de reproducción puede reducirse. De lo contrario, se puede mantener un retardo de reproducción inicialmente establecido.

En la operación 313, se puede determinar si una trama anterior es una trama activa. Si la trama anterior es una trama activa en la operación 313, se puede activar un tipo de adaptación basada en señal, y si la trama anterior no es una trama activa, se puede activar un tipo de adaptación basada en trama.

25 En la operación 315, se puede determinar si una trama posterior es una trama activa, si la trama previa no es una trama activa como el resultado de la determinación en la operación 313.

En las operaciones 317 y 319, un modo de adaptación tal como supresión o adición de ruido de confort (CN) se puede determinar y se proporciona a una memoria intermedia de fluctuación (130 de la figura 1) de acuerdo con un resultado de comparación entre p_{k-1} y w_k , si la trama posterior no es una trama activa como resultado de la determinación en la operación 315.

30 En la operación 318, un modo de adaptación tal como la adición de CN puede determinarse y se proporciona a la memoria intermedia de eliminación de fluctuación (130 de la figura 1) según un resultado de comparación entre p_{k-1} y z_k , si la trama posterior es una trama activa como resultado de la determinación en la operación 315.

35 En la operación 331, la memoria intermedia de eliminación de fluctuación (130 de la figura 1) elimina o agrega ruido de confort en función del modo de adaptación determinado de acuerdo con el resultado de comparación en la operación 317, 318 o 319.

40 Mientras tanto, en la operación 321, se puede determinar si una condición de que una primera unidad de reducción (510 de la figura 5) que se describirá a continuación se realiza si se cumple la trama anterior como una trama activa como resultado de la determinación en la operación 313. En detalle, cuando se satisface la Ecuación 17 a continuación, y no se ha producido una pérdida de paquetes en un número predeterminado de paquetes recibidos previamente, por ejemplo, 1000 paquetes, se determina que una red es estable, y un modo de adaptación como la reducción de tiempo adicional puede determinarse y proporcionarse a una unidad de aplicación TSM (170 de la figura 1).

45 En la operación 323, un modo de adaptación tal como reducción o ampliación de escala de tiempo puede determinarse y se proporciona a la unidad de aplicación TSM 170 de la figura 1) de acuerdo con un resultado de comparación entre p_{k-1} y v_k o u_k , si un modo de adaptación como la reducción de la escala de tiempo adicional no se determina como el resultado de la determinación en la operación 321.

En la operación 333, la unidad de aplicación TSM (170 de la figura 1) puede reducir o ampliar una escala de tiempo basada en el modo de adaptación determinado de acuerdo con el resultado de la determinación en la operación 321 y los resultados de comparación en las operaciones 323 y 325.

50 La figura 4 muestra un ejemplo en el que se puede seleccionar un factor de ajuste de tiempo según se seleccione el ajuste.

Con referencia a la figura 4, una unidad de reducción de escala de tiempo 410 puede funcionar de acuerdo con el resultado de determinación en la operación 321 o el resultado de comparación en la operación 323 en la figura 3. Es decir, cuando se satisface la Ecuación 17 siguiente, y no se ha producido una pérdida de paquetes en un número predeterminado de paquetes recibidos previamente, por ejemplo, 1000 paquetes, se puede reducir una escala de

tiempo. En otras palabras, cuando p_{k-1} es mayor que v_k , una escala de tiempo puede reducirse.

En lo sucesivo, el procesamiento de reducción de la escala de tiempo de acuerdo con una realización se describirá en detalle con referencia a la figura 5.

5 El procesamiento de reducción de la escala de tiempo es para reducir un tamaño de trama L a L_{salida} . L_{salida} puede diferir de L en, por ejemplo, 2,5 a 10 ms, y como resultado, un intervalo de trama puede ser de 10 a 17,5 ms. Una cantidad de escala de tiempo puede variar dependiendo de la ubicación de un segmento candidato de coincidencia óptimo (en adelante, denominado ubicación óptima) que tiene la mayor similitud con un primer segmento de una señal de entrada. Una ubicación de inicio de búsqueda y una ubicación de finalización de búsqueda que se utilizarán para buscar un segmento candidato desde una trama de entrada para la reducción de la escala de tiempo pueden variar según la frecuencia de muestreo. Se puede obtener una trama de salida de acuerdo con la reducción de la escala de tiempo mediante el desvanecimiento cruzado de un primer segmento de una trama de entrada y un segmento candidato de correspondencia óptimo. En este caso, la trama de salida se desplaza por una ubicación P_{sinc} óptima, y su tamaño es $L - P_{sinc}$. Las muestras después del segmento candidato de coincidencia óptima se pueden agregar a una señal fusionada de modo que se mantenga la continuidad con las tramas posteriores. De acuerdo con una realización, dado que no se requiere una muestra anticipada durante el procedimiento de reducción de la escala de tiempo, no se produce un retardo excedente, pero se puede requerir una memoria intermedia de reserva para generar una trama de salida de longitud fija a partir de una trama de salida de longitud variable.

Una unidad 430 de ampliación de escala de tiempo puede operar en función del resultado de la comparación en la operación 325 en la figura 3. Es decir, cuando p_{k-1} es menor que u_k , se puede ampliar una escala de tiempo.

20 En lo sucesivo, el procesamiento de ampliación de la escala de tiempo de acuerdo con una realización se describirá en detalle con referencia a la figura 6.

25 El procesamiento de ampliación de la escala de tiempo es para ampliar un tamaño de trama L a L_{salida} . L_{salida} puede diferir de L en, por ejemplo, 2,5 a 15 ms, y como resultado, un intervalo de trama puede ser de 22,5 a 35 ms. Una cantidad de escala de tiempo puede variar dependiendo de la ubicación de un segmento candidato de coincidencia óptimo (en adelante, denominado ubicación óptima) que tiene la mayor similitud con un primer segmento de una señal de entrada. Una ubicación de inicio de búsqueda y una ubicación de finalización de búsqueda que se utilizarán para buscar un segmento candidato desde una trama de entrada para la ampliación de la escala de tiempo pueden variar según la frecuencia de muestreo. Se puede obtener una trama de salida de acuerdo con la ampliación de la escala de tiempo mediante el desvanecimiento cruzado de un primer segmento de una trama de entrada y un segmento candidato de correspondencia óptimo. En este caso, la trama de salida se desplaza por una ubicación P_{sinc} óptima, y su tamaño es $L - P_{sinc}$. Aquí, P_{sinc} tiene un valor negativo basado en la ubicación de inicio y la ubicación final. Las muestras después del segmento candidato de coincidencia óptima se pueden agregar a una señal fusionada de modo que se mantenga la continuidad con las tramas posteriores. De acuerdo con una realización, no se requiere una muestra anticipada durante el procesamiento de ampliación de escala de tiempo, pero se requiere una trama previa. Por lo tanto, no se produce un retardo excedente, pero se puede requerir una memoria intermedia de reserva para generar una trama de salida de longitud fija a partir de una trama de salida de longitud variable.

Una unidad 450 de mantenimiento de escala de tiempo puede operar en función del resultado de la comparación en la operación 325 en la figura 3. Es decir, cuando p_{k-1} es mayor o igual que u_k , se puede mantener una escala de tiempo tal como está.

40 La figura 7 ilustra un esquema de reducción de retardo de reproducción cuando se determina la reducción de retardo de reproducción por modelo E.

En la figura 7, una primera unidad de reducción 710 puede reducir un retardo de reproducción usando una reducción de escala de tiempo adicional. Primero, la primera unidad 710 de reducción puede verificar la condición de que se puede reducir un retardo de reproducción estimado, comparando el retardo de reproducción estimado por la primera unidad 211 de estimación de retardo de reproducción y el retardo de reproducción estimado por la segunda unidad 213 de estimación de retardo de reproducción por la Ecuación 17 a continuación.

$$P_{k-1}^{EVS} > P_k^{\text{modelo E}} + 25 \quad (17)$$

Mientras tanto, un valor de diferencia entre los dos retardos de emisión puede almacenarse en una memoria intermedia que tiene una capacidad de 200 entradas.

50 Cuando la Ecuación 17 anterior se satisface, y no se ha producido la pérdida de paquetes en un número predeterminado de paquetes previamente llegados, por ejemplo, 1000 paquetes, se determina que una red es estable, y un retardo de reproducción se puede reducir mediante el uso de reducción adicional de tiempo.

Una segunda unidad 730 de reducción puede reducir un retardo de reproducción mediante la reducción de un excedente de retardo de reproducción. La segunda unidad 730 de reducción puede reducir el excedente de retardo

de reproducción establecido inicialmente en un primer valor predeterminado, por ejemplo, 60, a un segundo valor predeterminado, por ejemplo, 40, cuando se cumple la condición de la primera unidad 710 de reducción. Además,

cuando un valor mínimo en una memoria intermedia en la que los valores de diferencia de P^{EVS}_{k-1} y $P^{\text{modelo E}}_k$ si se almacenan es mayor que un tercer valor predeterminado, por ejemplo, 20, el excedente de retardo de reproducción puede reducirse a un cuarto valor predeterminado, por ejemplo, 20.

La figura 8 es un diagrama de bloques de un aparato TSM acuerdo con una realización.

El aparato mostrado en la figura 8 puede incluir una unidad 810 de cálculo de energía, una unidad 820 de estimación de similitud, una unidad 830 de control de calidad y una unidad 840 de adición de solapamiento.

Con referencia a la figura 8, la unidad 810 de cálculo de energía puede calcular la energía de una señal decodificada. En detalle, la unidad 810 de cálculo de energía puede analizar una señal de entrada que incluye una trama previa, en unidades de subsegmentos, por ejemplo, en unidades de 1 ms. Cuando la energía de todos los subsegmentos a fusionar es un umbral o menos, una trama puede escalarse como máximo. El umbral puede ser, por ejemplo, - 65 dB, pero no está limitado al mismo. Para la reducción de la escala de tiempo, un valor de desplazamiento s puede establecerse en una ubicación final de búsqueda en correspondencia con cada frecuencia de muestreo de modo que una trama de 20 ms se reduzca a una trama de 10 ms. Para la ampliación de la escala de tiempo, el valor de desplazamiento s puede establecerse en una ubicación de inicio de búsqueda en correspondencia con cada frecuencia de muestreo de manera que una trama de 20 ms se amplía a una trama de 35 ms. Se puede obtener una trama de salida mediante el desvanecimiento cruzado de un primer segmento y un segmento desplazado en s de una trama de entrada.

La unidad 820 de estimación de similitud puede determinar una ubicación óptima para TSM mediante el uso de una correlación obtenida en base a la energía calculada por la unidad 810 de cálculo de energía.

La unidad 830 de control de calidad puede prevenir la aparición de la distorsión de una señal de TSM con respecto a la ubicación óptima. En detalle, después de que la unidad 820 de estimación de similitud estima la similitud, se puede calcular una medición de calidad q . La medición de calidad q se puede obtener utilizando una correlación cruzada normalizada. La medición de calidad q puede usarse para determinar si se realiza TSM.

La unidad 840 de adición de solapamiento 840 puede realizar TSM con respecto a la ubicación óptima. La unidad 840 de adición de solapamiento puede generar una trama de salida escalada por el procesamiento de reducción de escala de tiempo o el procesamiento de ampliación de escala de tiempo.

La figura 9 es un diagrama de bloques de la unidad de estimación de similitud mostrada en la figura 8.

La unidad de estimación de similitud mostrada en la figura 9 puede incluir una unidad 910 de determinación de rango de búsqueda, una primera unidad 920 de estimación de similitud, una unidad 930 de determinación de búsqueda de rango restante, una segunda unidad 940 de estimación de similitud y una unidad 950 de determinación de ubicación óptima.

Con referencia a la figura 9, la unidad 910 de determinación del rango de búsqueda puede determinar un rango de búsqueda reducido en un porcentaje predeterminado, por ejemplo, 50 %, para buscar primero una ubicación óptima determinada en una trama anterior.

La primera unidad 920 de estimación de similitud puede buscar una ubicación óptima que tenga la máxima similitud estimando la similitud dentro del rango de búsqueda determinado.

La unidad 930 de determinación de búsqueda del rango restante obtiene una similitud normalizada C_τ aplicando la Ecuación 18 siguiente para la ubicación óptima determinada por la primera unidad 920 de estimación de similitud.

$$C_\tau = \frac{\sum_{n=0}^{L_{seg}} x(n)x(n+\tau)}{\sqrt{\sum_{n=0}^{L_{seg}} x^2(n) \cdot \sum_{n=0}^{L_{seg}} x^2(n+\tau)}} \quad (18)$$

Aquí, τ indica la ubicación óptima determinada por la primera unidad 920 de estimación de similitud. Cuando C_τ es mayor que un umbral, finaliza la búsqueda de ubicación óptima, de lo contrario se determina que la búsqueda de ubicación óptima se realiza con respecto al rango restante.

La segunda unidad 940 de estimación de similitud puede buscar una ubicación óptima que tenga la máxima similitud dentro del rango restante nuevamente.

La unidad 950 de determinación de ubicación óptima puede seleccionar, como ubicación óptima final, una ubicación

óptima que tenga una mayor similitud entre la ubicación óptima determinada por la primera unidad 920 de estimación de similitud y la ubicación óptima determinada por la segunda unidad 940 de estimación de similitud. La reducción o ampliación de la escala de tiempo se puede realizar en función de la ubicación óptima final.

La figura 10 es un diagrama de bloques de un dispositivo multimedia de acuerdo con una realización.

5 Un dispositivo 1000 multimedia mostrado en la figura 10 puede incluir una unidad 1010 de comunicación y un módulo 1030 de decodificación. Además, una unidad 1050 de almacenamiento para almacenar una señal de audio reconstruida obtenida como resultado de decodificación puede incluirse adicionalmente dependiendo del uso de la señal de audio reconstruida. Además, el dispositivo 1000 multimedia puede incluir además un altavoz 1070. Es decir, la unidad 1050 de almacenamiento y el altavoz 1070 pueden incluirse opcionalmente. En el presente documento, el
10 módulo 1030 de decodificación puede integrarse con otro componente (no mostrado) incluido en el dispositivo 1000 multimedia para implementar uno o más procesadores (no mostrados).

Con referencia a la figura 10, la unidad 1010 de comunicación puede recibir al menos una de una secuencia de bits codificada y una señal de audio proporcionada desde el exterior, o transmitir al menos una de una señal de audio reconstruida obtenida como resultado de decodificación del módulo 1030 de decodificación y una secuencia de bits
15 de audio obtenida por un resultado de codificación.

La unidad 1010 de comunicación está configurada para poder transmitir y recibir datos hacia y desde un dispositivo multimedia externo o un servidor a través de una red inalámbrica tal como Internet inalámbrico, intranet inalámbrica, red telefónica inalámbrica, red inalámbrica de área local (LAN), Wi-Fi, Wi-Fi Direct (WFD), tercera generación (3G),
20 4G, Bluetooth, asociación de datos por infrarrojos (IrDA), identificación por radiofrecuencia (RFID), banda ultra ancha (UWB), Zigbee o comunicación de campo cercano (NFC) o una red cableada, como una red telefónica cableada o Internet por cable.

El módulo 1030 de decodificación puede recibir un paquete o flujo de bits proporcionado a través de la unidad 1010 de comunicación y realizar procesado de ajuste de retardo de reproducción o procesamiento de TSM de acuerdo con las realizaciones descritas anteriormente cuando se realiza la decodificación. En el presente documento, de acuerdo
25 con una realización, el procesamiento de ajuste de retardo de reproducción puede incluir: ajustar un excedente de retardo de reproducción en base a un valor de diferencia entre un primer retardo de reproducción obtenido en un primer esquema y un segundo retardo de reproducción obtenido en un segundo esquema, determinar un tipo de adaptación de una trama actual según si una trama anterior es una trama activa; y cuando el tipo de adaptación determinado es una adaptación basada en señales, realizar una TSM de acuerdo con un esquema de adaptación determinado de acuerdo con un resultado de comparación entre el primer retardo de reproducción y el segundo retardo
30 de reproducción y un resultado de comparación entre un retardo objetivo y el primer retardo de reproducción. Mientras tanto, de acuerdo con una realización, el procesamiento de TSM puede incluir determinar un rango de búsqueda reducido para la trama actual en función de una ubicación óptima de una trama pasada, determinar si se debe buscar un rango restante al excluir el rango de búsqueda reducido, estimar la similitud con respecto al rango de búsqueda reducido y al rango restante, buscando ubicaciones óptimas que tengan la máxima similitud, determinando, como
35 ubicación óptima final, una de las ubicaciones óptimas recuperadas en el rango de búsqueda reducido y la ubicación óptima recuperada en el rango restante, y realizar una reducción de la escala de tiempo o ampliación en función de la ubicación óptima final.

La unidad 1050 de almacenamiento puede almacenar la señal de audio reconstruida generada por el módulo 1030 de decodificación. Mientras tanto, la unidad 1050 de almacenamiento puede almacenar varios programas requeridos para
40 operar el dispositivo 1000 multimedia.

El altavoz 1070 puede emitir la señal de audio reconstruida generada por el módulo 4230 de decodificación al exterior.

La figura 11 es un diagrama de bloques de un dispositivo multimedia de acuerdo con otra realización.

Un dispositivo 1100 multimedia mostrado en la figura 11 puede incluir una unidad 1110 de comunicación, un módulo 1120 de codificación y un módulo 1130 de decodificación. Además, una unidad 1140 de almacenamiento para almacenar un flujo de bits de audio obtenido como resultado de codificación o una señal de audio reconstruida obtenida
45 como resultado de decodificación puede incluirse adicionalmente dependiendo del uso del flujo de bits de audio o la señal de audio reconstruida. Además, el dispositivo 1100 multimedia puede incluir además un micrófono 1150 o un altavoz 1060. Aquí, el módulo 1120 de codificación y el módulo 1130 de decodificación pueden integrarse con otro
50 componente (no mostrado) incluido en el dispositivo 1100 multimedia para implementar uno o más procesadores (no mostrados). Algunos de los componentes mostrados en la figura 11 corresponden a los componentes en el dispositivo 1000 multimedia mostrado en la figura 10, y por lo tanto se omite una descripción detallada de los mismos.

La unidad 1110 de comunicación puede recibir al menos uno de audio y un flujo de bits codificado proporcionado desde el exterior o transmitir al menos uno de audio reconstruido y un flujo de bits de audio obtenido por un resultado
55 de codificación del módulo 1120 de codificación.

El módulo 1120 de codificación puede generar un flujo de bits o paquete incrustando varios códigos en el mismo para codificar una señal de audio o de voz.

El módulo 1130 de decodificación puede implementarse en correspondencia con o independientemente del módulo 1120 de codificación. El módulo 1130 de decodificación puede recibir un paquete o flujo de bits proporcionado a través de la unidad 1110 de comunicación y aplicar procesado de ajuste de retardo de reproducción o procesamiento de TSM de acuerdo con las realizaciones descritas anteriormente cuando se realiza la decodificación.

- 5 La unidad 1140 de almacenamiento puede almacenar varios programas requeridos para operar el dispositivo 1100 multimedia.

El micrófono 1150 puede proporcionar una señal de audio de un usuario o del exterior al módulo 1120 de codificación.

- 10 Los dispositivos 1000 y 1100 multimedia mostrados en las figuras 10 y 11 pueden incluir un terminal exclusivo de comunicación de voz que incluye un teléfono, un teléfono móvil o similar, un terminal exclusivo de transmisión o música que incluye un televisor (TV), un reproductor de MP3 o similar, un dispositivo terminal de fusión de la voz terminal exclusivo de comunicación y el terminal exclusivo de difusión o música, o un terminal de usuario de un sistema de teleconferencia o interacción, pero no están limitados a los mismos. Además, el dispositivo 1000 o 1100 multimedia puede usarse como un cliente, un servidor o un transductor dispuesto entre un cliente y un servidor.

- 15 Mientras tanto, cuando el dispositivo 1000 u 1100 multimedia es, por ejemplo, un teléfono móvil, aunque no se muestra, una unidad de entrada de usuario tal como un teclado, una unidad de visualización para mostrar una interfaz de usuario o información procesada por el teléfono móvil, y se puede incluir además un procesador para controlar la función general del teléfono móvil. Además, el teléfono móvil puede incluir además una unidad de cámara que tiene una función de captura de imágenes y al menos un componente para realizar las funciones requeridas en el teléfono móvil.

- 20 Mientras tanto, cuando el dispositivo 1000 u 1100 multimedia es, por ejemplo, un televisor, aunque no se muestra, una unidad de entrada de usuario tal como un teclado, una unidad de visualización para mostrar la información de transmisión recibida y un procesador para controlar la función general de la TV también se puede incluir. Además, el televisor también puede incluir al menos un componente para realizar las funciones requeridas en el televisor.

- 25 Las realizaciones descritas anteriormente pueden ser escritas como programas ejecutables por ordenador y se pueden implementar en ordenadores digitales de uso general que ejecutan los programas utilizando un medio de grabación legible por ordenador. Además, una estructura de datos, instrucciones de programa o un archivo de datos que se pueda utilizar en las realizaciones descritas anteriormente de la presente divulgación se pueden grabar en un medio de grabación legible por ordenador a través de diversos medios. El medio de grabación legible por ordenador puede incluir todos los tipos de dispositivos de almacenamiento en los que se almacenan datos legibles por ordenador. Ejemplos del medio de grabación legible por ordenador pueden incluir medios magnéticos tales como discos duros, disquetes y cintas magnéticas, medios ópticos de grabación como memorias compactas de solo lectura (CD-ROM) y discos versátiles digitales (DVD), medios magnetoópticos tales como discos flopticales y dispositivos de hardware que están especialmente configurados para almacenar y ejecutar comandos de programas, tales como ROM, RAM y memorias flash. Alternativamente, el medio de grabación legible por ordenador puede ser un medio de transmisión a través del cual se transmiten comandos de programa de designación de señal, una estructura de datos o similares.
- 30 Ejemplos de los comandos del programa pueden incluir un código de lenguaje de alto nivel que puede ser ejecutado por un ordenador usando un intérprete, así como un código de lenguaje de máquina creado por un compilador.
- 35

REIVINDICACIONES

1. Un procedimiento de ajuste de retardo de reproducción que comprende:

ajustar un excedente de retardo de reproducción en base a un valor de diferencia entre un primer retardo de reproducción debido a una memoria intermedia de eliminación de fluctuación y un segundo retardo de reproducción estimado en base a un modelo E;
 determinar un tipo de adaptación de una trama actual según si una trama anterior es una trama activa; y cuando el tipo de adaptación determinado es adaptación basada en señal, realizar una modificación de escala de tiempo, TSM, de acuerdo con un esquema de adaptación determinado de acuerdo con un resultado de comparación entre el primer retardo de reproducción y el segundo retardo de reproducción y un resultado de comparación entre un retardo objetivo y el primero retardo de reproducción, en el que la realización de la TSM comprende:

determinar un rango de búsqueda reducido, reducido desde un rango de búsqueda original, para una trama actual, en base a una ubicación de un segmento candidato coincidente determinada en una trama anterior;
 buscar un primer segmento candidato coincidente dentro del rango de búsqueda reducido mediante la estimación de similitud;
 determinar si buscar dentro del rango restante, excepto el rango de búsqueda reducido del rango de búsqueda original;
 buscar un segundo segmento candidato coincidente dentro del rango restante estimando la similitud;
 determinar, como ubicación final de un segmento candidato coincidente, una ubicación que tiene una mayor similitud entre la ubicación del primer segmento candidato coincidente y la ubicación del segundo segmento candidato coincidente; y
 realizar una reducción de la escala de tiempo para reducir el tamaño de una trama o una ampliación de escala de tiempo para ampliar el tamaño de la trama en función de la ubicación final.

2. El procedimiento de ajuste de retardo de reproducción de la reivindicación 1, en el que la realización de la TSM comprende:

cuando una diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es mayor que un umbral, determinar la reducción de la escala de tiempo como el esquema de adaptación; y
 cuando la diferencia entre el primera retardo de reproducción y el segundo retardo de reproducción es menor que el umbral, determinar la ampliación o reducción de la escala de tiempo como el esquema de adaptación en base a una diferencia entre el retardo objetivo y el primer retardo de reproducción.

3. Un aparato de ajuste de retardo de reproducción que comprende:

una unidad (170) de control de adaptación configurada para ajustar un excedente de retardo de reproducción en base a un valor de diferencia entre un primer retardo de reproducción debido a una unidad (130) de memoria intermedia de eliminación de fluctuación y un segundo retardo de reproducción estimado en base a un modelo E y determinar un tipo de adaptación de una trama actual según si una trama anterior es una trama activa;
 la unidad (130) de memoria intermedia de fluctuación configurada para almacenar tramas a decodificar, y para realizar la adaptación en las tramas almacenadas cuando el tipo de adaptación determinado es una adaptación basada en tramas;
 una unidad (140) de decodificación configurada para decodificar una trama proporcionada desde la unidad (130) de memoria intermedia de fluctuación; y
 una unidad (150) de aplicación de modificación de escala de tiempo, TSM, configurada para realizar la adaptación en la trama decodificada cuando el tipo de adaptación determinado es una adaptación basada en el tiempo, en el que la unidad de aplicación de TSM está configurada además para:

determinar un rango de búsqueda reducido, reducido desde un rango de búsqueda original, para una trama actual, en base a una ubicación de un segmento candidato coincidente determinada en una trama anterior,
 buscar un primer segmento candidato coincidente dentro del rango de búsqueda reducido mediante la estimación de similitud,
 determinar si buscar dentro del rango restante, excepto el rango de búsqueda reducido del rango de búsqueda original,
 buscar un segundo segmento candidato coincidente dentro del rango restante estimando la similitud,
 determinar, como ubicación final, una ubicación que tiene una mayor similitud entre la ubicación del primer segmento candidato coincidente y la ubicación del segundo segmento candidato coincidente; y
 realizar una reducción de la escala de tiempo para reducir el tamaño de una trama o una ampliación de escala de tiempo para ampliar el tamaño de la trama en función de la ubicación final.

4. El aparato de ajuste de retardo de reproducción de la reivindicación 3, en el que la unidad de aplicación de TSM puede estar configurada además para determinar la reducción de escala de tiempo como un esquema de adaptación cuando una diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es mayor que un umbral y para determinar la ampliación o reducción de escala de tiempo de como el esquema de adaptación en base

a una diferencia entre un retardo objetivo y el primer retardo de reproducción cuando la diferencia entre el primer retardo de reproducción y el segundo retardo de reproducción es menor que el umbral.

5. Un medio de grabación legible por ordenador que tiene grabado en el mismo un programa que cuando es ejecutado por un ordenador realiza el procedimiento de ajuste de retardo de reproducción de la reivindicación 1.

5

FIG. 1

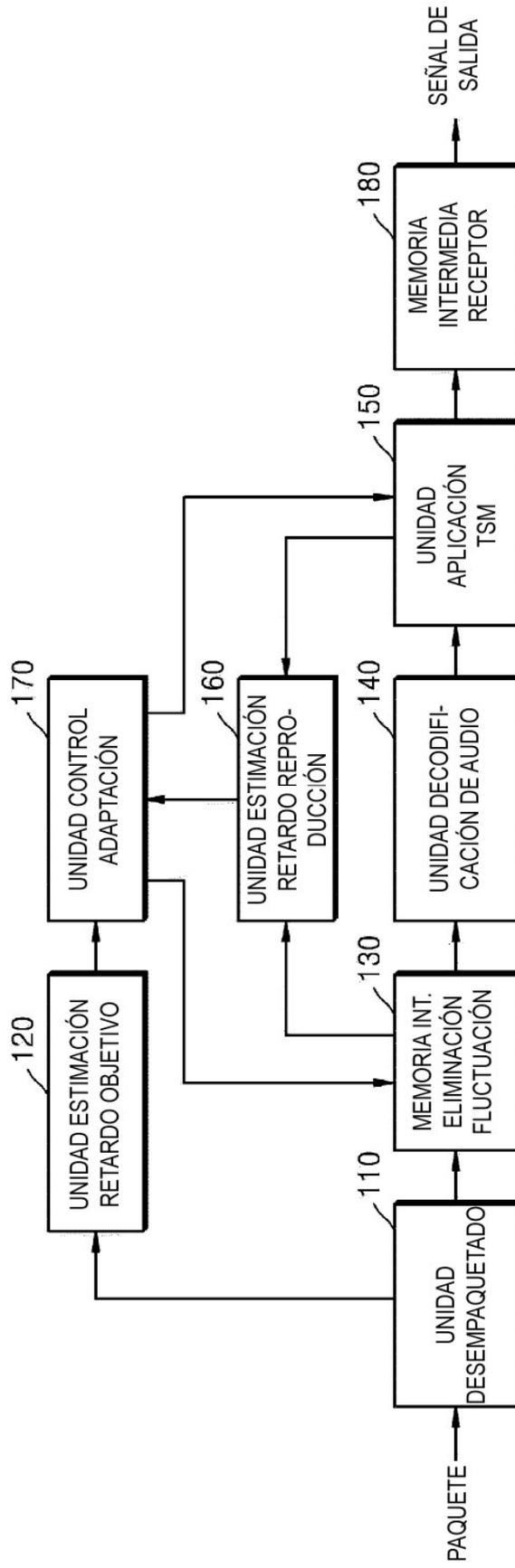


FIG. 2

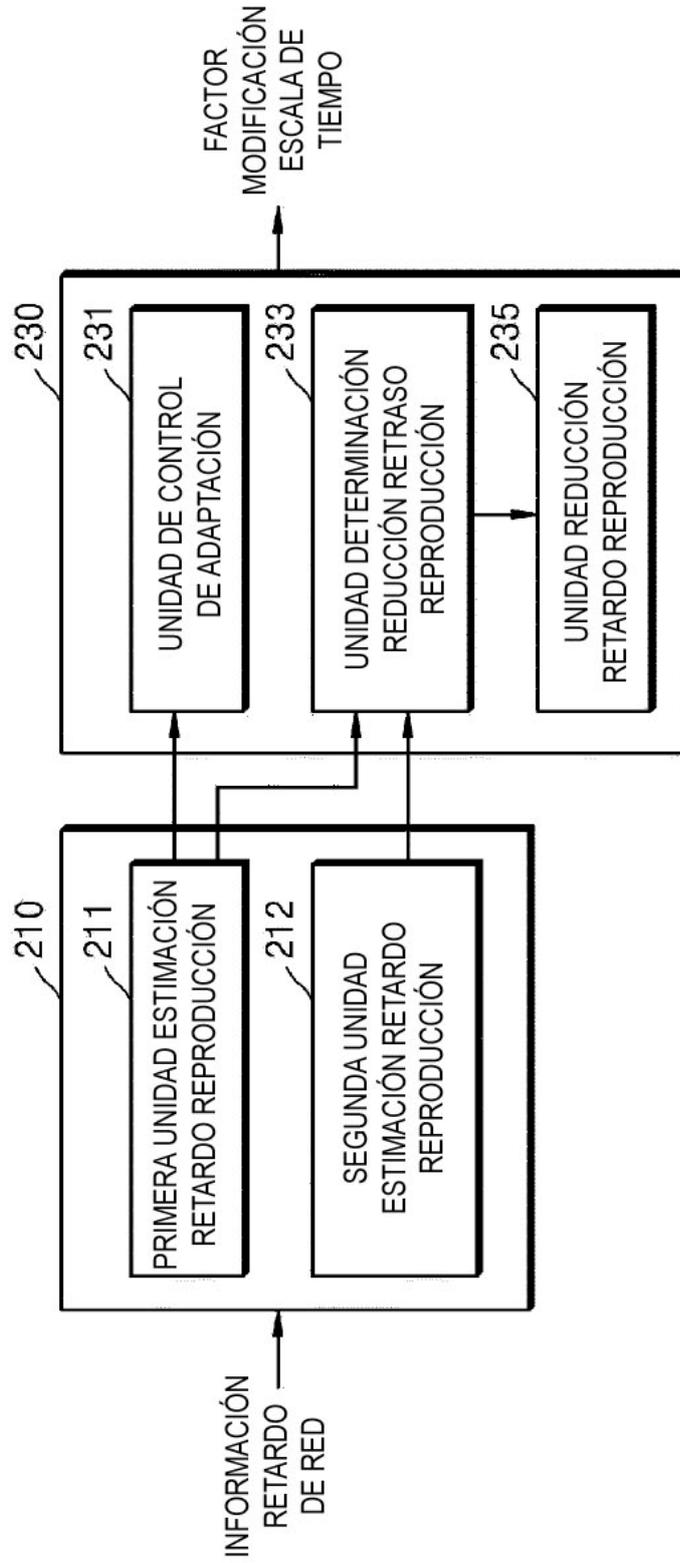


FIG. 3

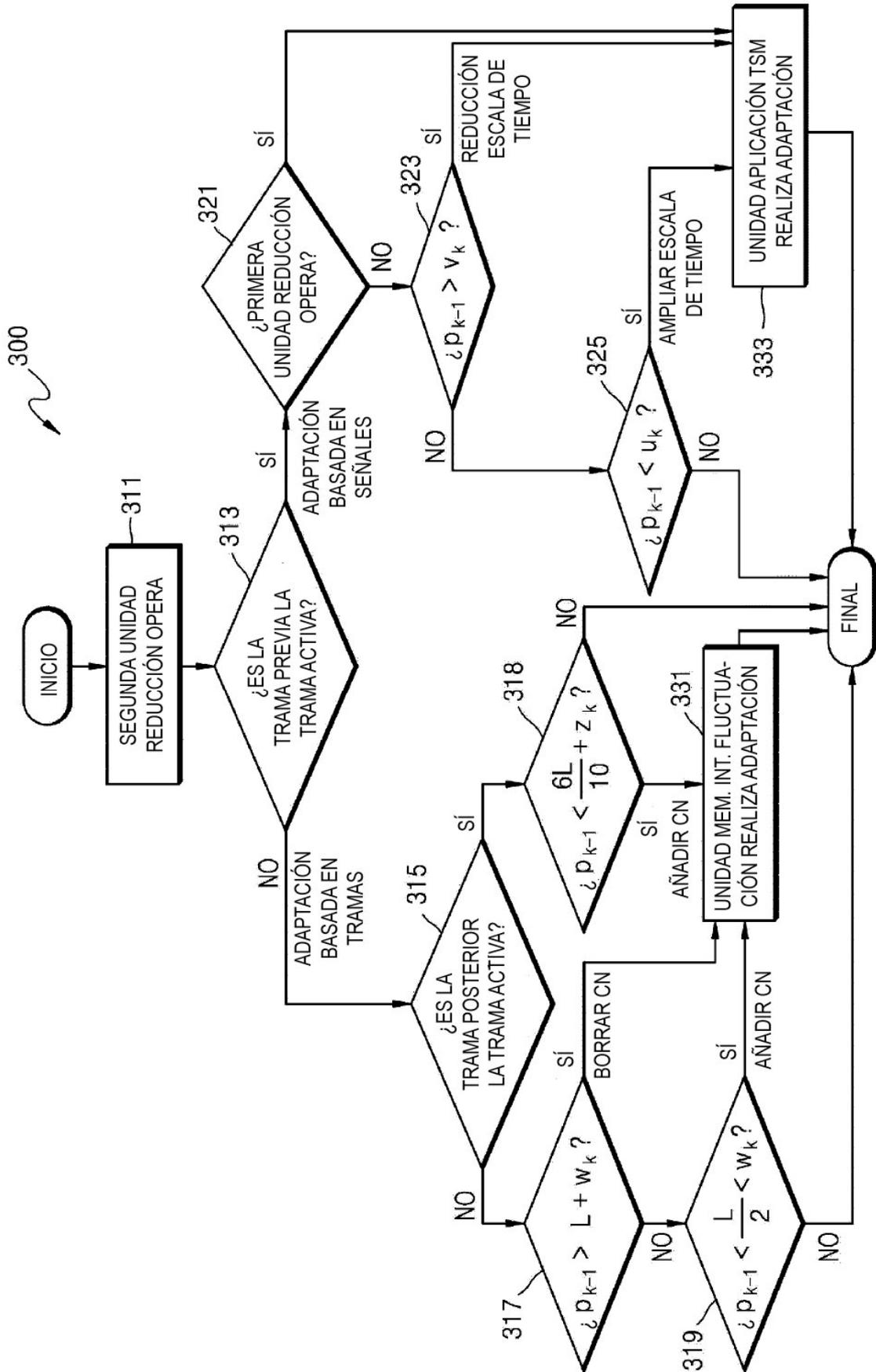


FIG. 4

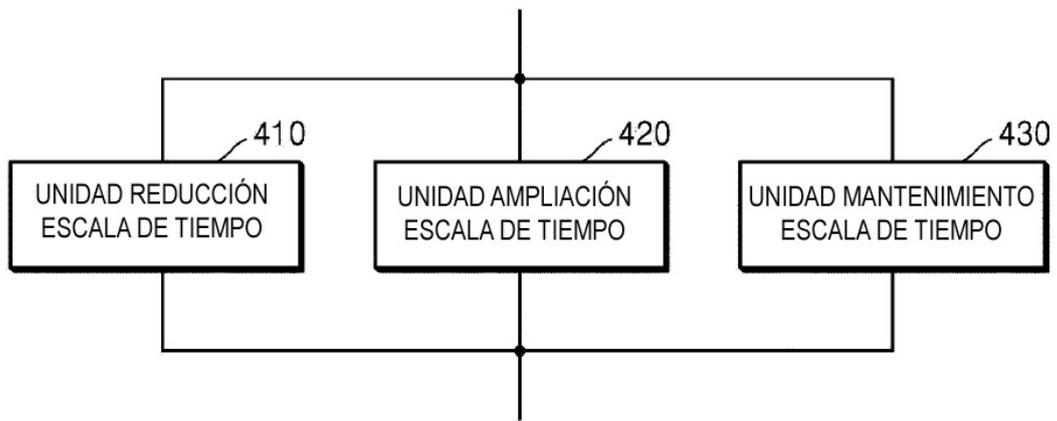


FIG. 5

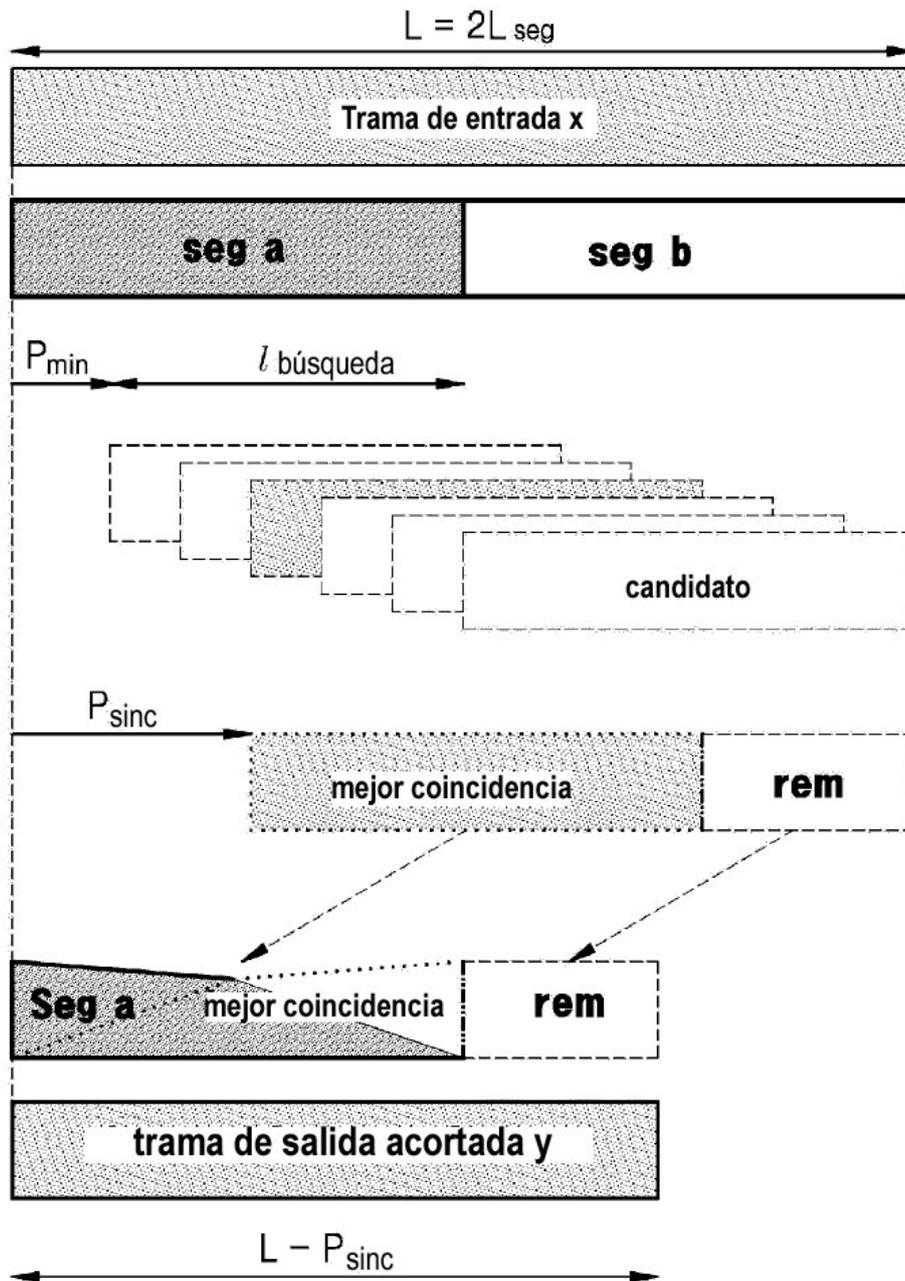


FIG. 6

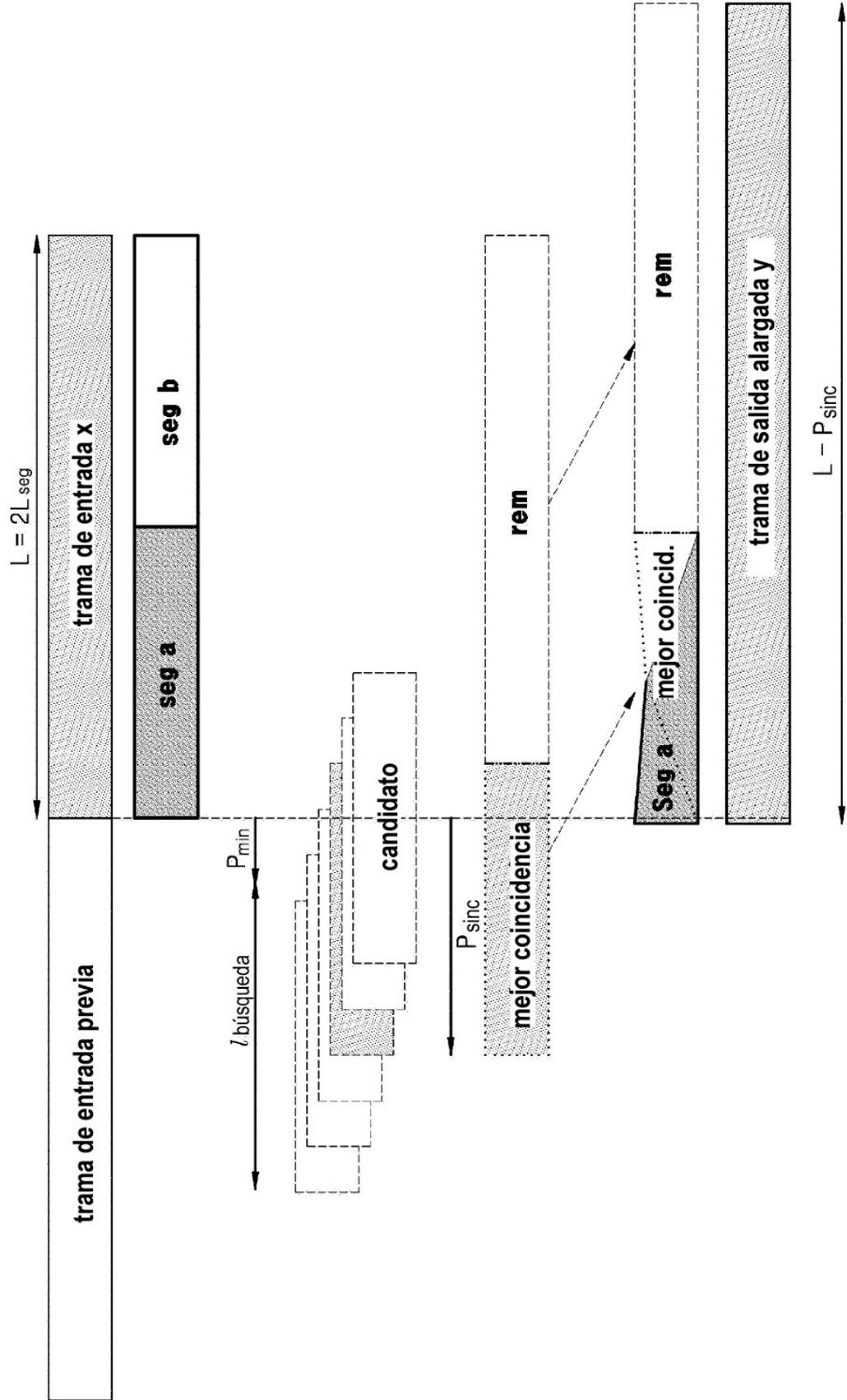


FIG. 7

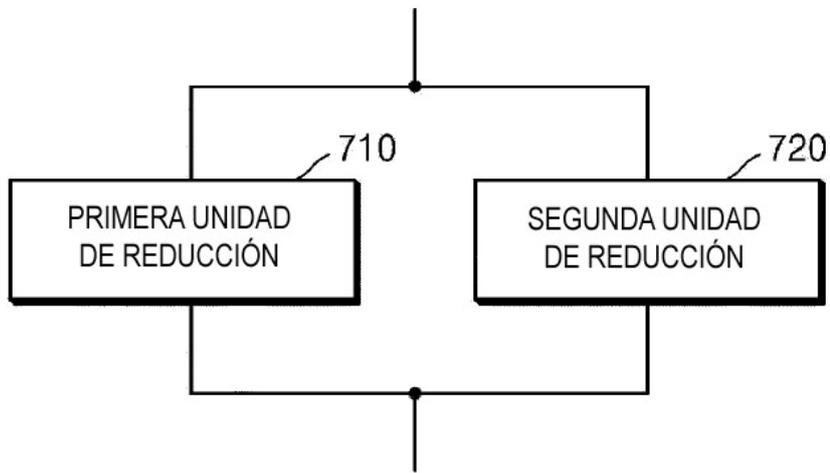


FIG. 8

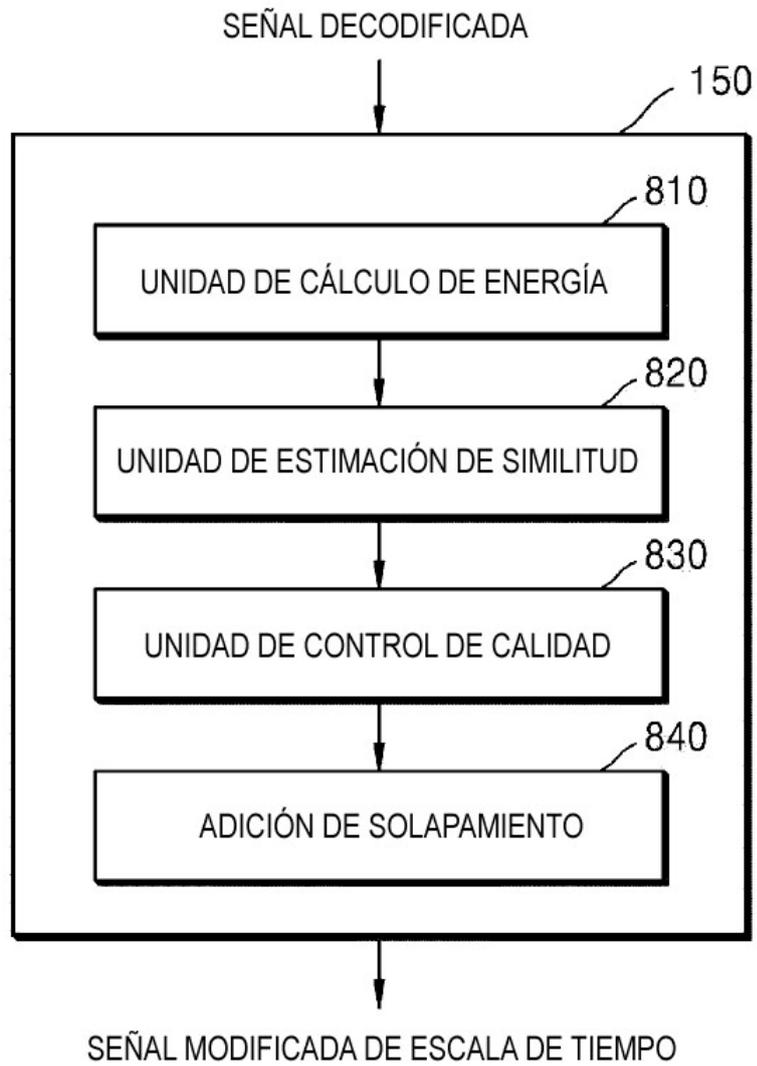


FIG. 9

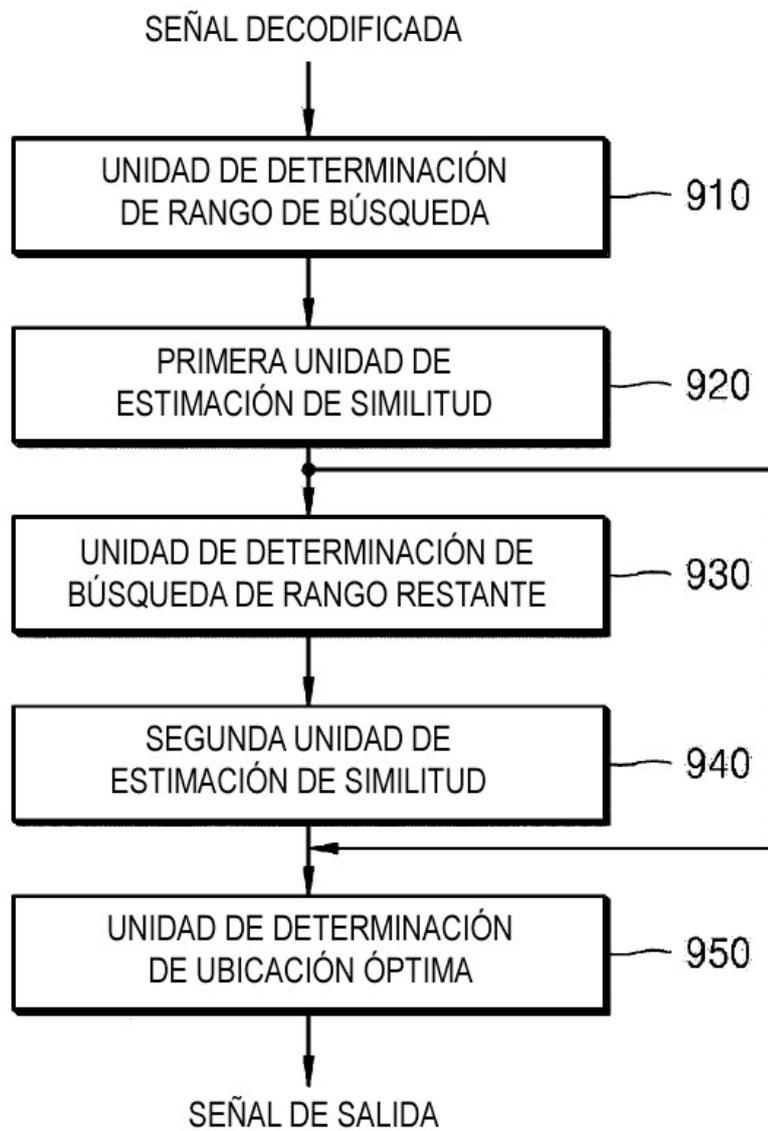


FIG. 10

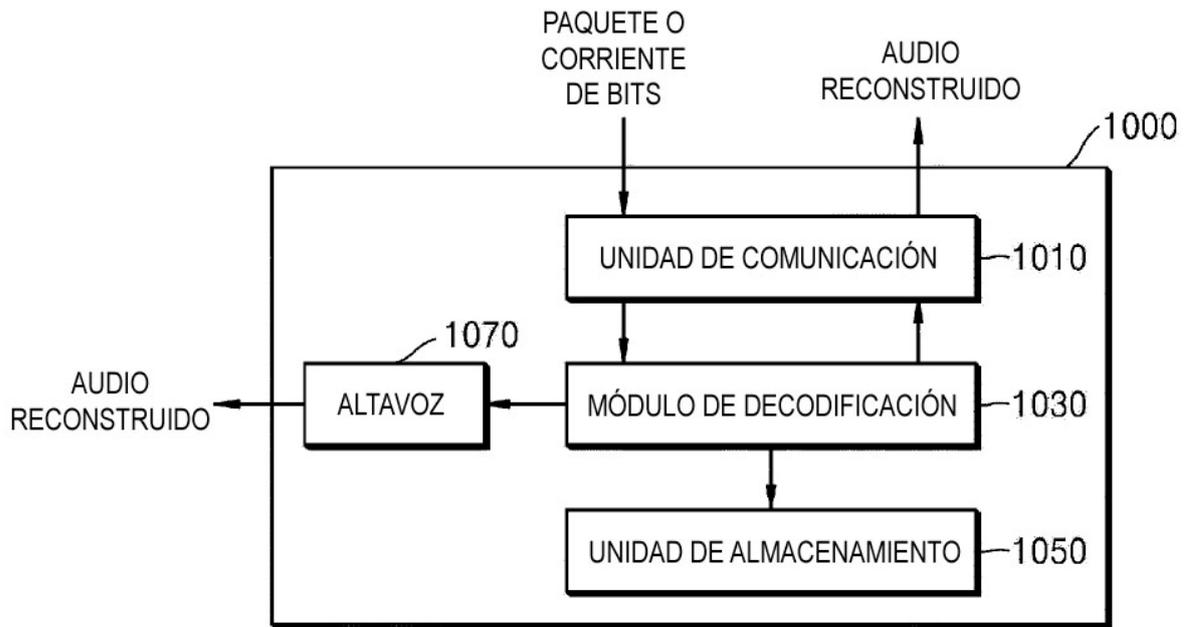


FIG. 11

