

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 377 017**

51 Int. Cl.:  
**G10L 21/04** (2006.01)  
**G11B 20/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07760954 .3**  
96 Fecha de presentación: **19.04.2007**  
97 Número de publicación de la solicitud: **2011118**  
97 Fecha de publicación de la solicitud: **07.01.2009**

54 Título: **Procedimiento y aparato para el ajuste automático de la velocidad de reproducción de datos de audio**

30 Prioridad:  
**25.04.2006 US 411074**

45 Fecha de publicación de la mención BOPI:  
**21.03.2012**

45 Fecha de la publicación del folleto de la patente:  
**21.03.2012**

73 Titular/es:  
**INTEL CORPORATION  
2200 MISSION COLLEGE BOULEVARD  
SANTA CLARA, CA 95052, US**

72 Inventor/es:  
**SHIRES, Glen**

74 Agente/Representante:  
**Carpintero López, Mario**

ES 2 377 017 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 aparato para el ajuste automático de la velocidad de reproducción de datos de audio

**Campo técnico**

5 Las realizaciones de la presente invención son pertinentes a los reproductores multimedia que reproducen datos de audio. Más específicamente, las realizaciones de la presente invención versan acerca de un procedimiento y un aparato para el ajuste automático de la velocidad de reproducción de datos de audio.

**Antecedentes**

10 Existen reproductores multimedia que permiten la reproducción de grabaciones de audio y de sesiones de audio-vídeo a una velocidad que es superior a la velocidad normal. Esto permite que los usuarios escuchen y vean estas sesiones en un periodo de tiempo menor. El uso de estas características puede ser común, por ejemplo, en aplicaciones empresariales, en las que los empleados ven y/o escuchan sesiones de formación, reuniones, conferencias y presentaciones. El uso de estas características también puede ser común en aplicaciones de ocio, por ejemplo, cuando los usuarios escuchan la radio o audio bajo demanda o ven la televisión. Estas características permiten que la reproducción a mayor velocidad esté libre de fallos de audio y vídeo.

15 Típicamente, los usuarios descubren que la reproducción de datos de audio resulta inteligible y comprensible a velocidades de reproducción aproximadamente entre 1,2 y 1,9 veces la velocidad normal de reproducción. Sin embargo, la velocidad óptima puede variar durante la reproducción debido a la velocidad de la voz de la persona que habla, al ruido de fondo, a la presencia de pausas de silencio o rellenas, y a otros criterios que pueden cambiar en el curso de la reproducción de los datos de audio.

20 Los reproductores multimedia actuales permiten que los usuarios ajusten manualmente la velocidad de reproducción de los datos de audio. Cuando la velocidad óptima de reproducción cambia frecuentemente en el curso de la reproducción de datos de audio, realizar ajustes manualmente puede resultar inconveniente. Además, cuando se realiza un ajuste manual, un oyente puede reaccionar solamente a cambios en los datos de audio. La demora experimentada en la detección y la reacción al cambio en los datos de audio puede dar como resultado la reproducción de porciones de datos de audio a una velocidad que resulte incomprensible para el oyente. Esto puede hacer que el oyente vuelva a reproducir los datos de audio, negando así algunos de los beneficios de la reproducción a mayor velocidad.

25 El ajuste automático de la velocidad de reproducción de los datos de audio ha sido dado a conocer previamente. En lo que sigue se identifican dos ejemplos de publicaciones anteriores a modo de antecedente de la presente invención.

La publicación de solicitud de patente estadounidense US 2002/0010916 A1 da a conocer un procedimiento y un aparato que controla la velocidad de reproducción de datos de audio correspondientes a un flujo de voz en continuo. Usando un reconocimiento de voz, se determina la tasa de velocidad de los datos de audio y se la compara con una tasa diana. En base a esta comparación, se aumenta o se disminuye la tasa para que coincida con la tasa diana.

35 La publicación de solicitud de patente estadounidense US 2005/0149329 A1 describe un aparato para cambiar la velocidad de reproducción de voz grabada que incluye una memoria que almacena una pluralidad de mensajes de voz grabados y una pluralidad de tablas de características. Cada tabla de características está asociada con un mensaje individual de los mensajes vocales e incluye parámetros intermedios basados en los estados de fluctuación de las tramas de voz del mensaje grabado de voz asociado. Un módulo de reproducción recibe una entrada que especifica un mensaje de voz grabado en la memoria que debe ser reproducido y la velocidad con la que el mensaje de voz grabado debe ser reproducido. En respuesta a esta entrada, el módulo de reproducción usa un conjunto de reglas de decisión para modificar el mensaje de voz especificado en base a los parámetros de las tramas de voz en la tabla de características asociada con el mensaje de voz especificado y la velocidad de reproducción especificada, antes de la reproducción del mensaje de voz especificado.

45 La presente invención proporciona un procedimiento para la gestión de datos de audio, según se define en la reivindicación 1, un medio accesible por máquina, según se define en la reivindicación 5, para la implementación de tal procedimiento y un aparato de ajuste de la velocidad de reproducción, según se define en la reivindicación 6. Notablemente, los datos de audio se convierten de un dominio temporal a un dominio frecuencial, se extraen características de los datos de audio en un dominio frecuencial y las tasas de cambio de las características extraídas en el dominio frecuencial se usan y se miden para generar uno o más parámetros de control de la velocidad de reproducción que son usados después para ajustar automáticamente la velocidad de reproducción de los datos de audio.

**Breve descripción de los dibujos**

Las características y las ventajas de las realizaciones de la presente invención están ilustradas a título de ejemplo y no se pretende que limiten el alcance de las realizaciones de la presente invención a las realizaciones particulares mostradas.

5 La Figura 1 es un diagrama de bloques de un sistema ejemplar en el cual puede implementarse una realización ejemplar de la presente invención.

La Figura 2 es un diagrama de bloques de una unidad de ajuste de la velocidad de reproducción según una realización ejemplar de la presente invención.

10 La Figura 3 es un diagrama de bloques de una unidad integradora de la tasa de cambio según una realización ejemplar de la presente invención.

La Figura 4 es un diagrama de flujo que ilustra un procedimiento para la gestión de datos de audio según una primera realización de la presente invención.

La Figura 5 es un diagrama de flujo que ilustra un procedimiento para la gestión de datos de audio según una segunda realización de la presente invención.

15 La Figura 6 es un diagrama de flujo que ilustra un procedimiento para la generación de un valor de control de la velocidad de reproducción según una realización de la presente invención.

**Descripción detallada**

20 En la descripción siguiente, se presenta nomenclatura específica con fines de explicación para proporcionar una comprensión cabal de realizaciones de la presente invención. Sin embargo, será evidente para un experto en la técnica que estos detalles específicos pueden no ser requeridos para poner en práctica las realizaciones de la presente invención. En otros casos, se muestran en forma de diagrama de bloques circuitos, dispositivos y procedimientos bien conocidos para evitar oscurecer innecesariamente realizaciones de la presente invención.

25 La Figura 1 es un diagrama de bloques de una primera realización de un sistema en el que puede implementarse una realización de la presente invención. El sistema es un sistema 100 de ordenador. El sistema 100 de ordenador incluye uno o más procesadores que procesan señales de datos. Según se muestra, el sistema 100 de ordenador incluye un primer procesador 101 y un procesador enésimo 105, pudiendo n ser cualquier número. Los procesadores 101 y 105 pueden ser microprocesadores de ordenador de conjunto complejo de instrucciones, microprocesadores informáticos de conjunto reducido de instrucciones, microprocesadores con palabras de instrucciones muy largas, procesadores que implementen una combinación de conjuntos de instrucciones u otros dispositivos procesadores. Los procesadores 101 y 105 pueden ser procesadores de núcleos múltiples, con múltiples núcleos procesadores en cada chip. Los procesadores 101 y 105 están acoplados a un bus 110 de CPU que transmite señales de datos entre los procesadores 101 y 105 y otros componentes en el sistema 100 de ordenador.

35 El sistema 100 de ordenador incluye una memoria 113. La memoria 113 incluye una memoria principal que puede ser un dispositivo de memoria dinámica de acceso aleatorio (DRAM). La memoria 113 puede almacenar instrucciones y código representados por señales de datos que pueden ser ejecutados por los procesadores 101 y 105. Una memoria de almacenamiento temporal (almacenamiento temporal del procesador) puede residir en cada uno de los procesadores 101 y 105 para almacenar señales de datos procedentes de la memoria 113. El almacenamiento temporal puede acelerar los accesos a memoria de los procesadores 101 y 105 aprovechando la localidad del acceso. En una realización alternativa del sistema 100 de ordenador, el almacenamiento temporal puede residir de forma externa a los procesadores 101 y 105.

40 Un controlador puente 111 de memoria está acoplado al bus 110 de CPU y a la memoria 113. El controlador puente 111 de memoria dirige señales de datos entre los procesadores 101 y 105, la memoria 113 y otros componentes en el sistema 100 de ordenador y comunica las señales de datos entre el bus 110 de CPU, la memoria 113 y un primer bus 120 de entrada/salida (E/S).

45 El primer bus 120 de E/S puede ser un solo bus o una combinación de múltiples buses. El primer bus 120 de E/S proporciona enlaces de comunicaciones entre componentes en el sistema 100 de ordenador. Un controlador 121 de red está acoplado al primer bus 120 de E/S. El controlador 121 de red puede enlazar el sistema 100 de ordenador con una red de ordenadores (no mostrada) y soporta la comunicación entre las máquinas. Un controlador 122 del dispositivo de visualización está acoplado al primer bus 120 de E/S. El controlador 122 del dispositivo de visualización permite el acoplamiento de un dispositivo de visualización (no mostrado) al sistema 100 de ordenador y actúa como interfaz entre el dispositivo de visualización y el sistema 100 de ordenador.

Un segundo bus 130 de E/S puede ser un solo bus o una combinación de múltiples buses. El segundo bus 130 de E/S proporciona enlaces de comunicaciones entre componentes en el sistema 100 de ordenador. El dispositivo 131

de almacenamiento de datos está acoplado al segundo bus 130 de E/S. El dispositivo 131 de almacenamiento de datos puede ser una unidad de disco duro, una unidad de disquete flexible, un dispositivo de CD-ROM, un dispositivo de memoria flash u otro dispositivo de almacenamiento masivo. Una interfaz 132 de entrada está acoplada al segundo bus 130 de E/S. La interfaz 132 de entrada puede ser, por ejemplo, un teclado y/o un controlador de ratón u otra interfaz de entrada. La interfaz 132 de entrada puede ser un dispositivo dedicado o puede residir en otro dispositivo, como un controlador de bus u otro controlador. La interfaz 132 de entrada permite el acoplamiento de un dispositivo de entrada al sistema 100 de ordenador y transmite señales de datos desde un dispositivo de entrada al sistema 100 de ordenador. Un controlador 133 de audio está acoplado al segundo bus 130 de E/S. El controlador 133 de audio opera coordinando la grabación y la reproducción de sonidos. Un puente 123 de bus acopla el primer bus 120 de E/S al segundo bus 130 de E/S. El puente 123 de bus opera regulando y comunicando señales de datos entre el primer bus 120 de E/S y el segundo bus 130 de E/S.

Según una realización de la presente invención, una unidad 140 de ajuste de la velocidad de reproducción puede ser implementada en el sistema 100 de ordenador. Según una realización, el sistema 100 de ordenador lleva a cabo la gestión de los datos de audio en respuesta a la ejecución de secuencias de instrucciones por parte del procesador 101 en la memoria 113 representada por la unidad 140 de ajuste de la velocidad de reproducción. Tales instrucciones pueden ser leídas a la memoria 113 desde otros medios legibles por ordenador, tales como el almacenamiento 131 de datos o desde un ordenador conectado a la red por medio del controlador 112 de red. La ejecución de las secuencias de instrucciones en la memoria 113 hace que el procesador soporte la gestión de los datos de audio. Según una realización de la presente invención, la unidad 140 de ajuste de la velocidad de reproducción identifica una condición en los datos de audio. La unidad 140 de ajuste de la velocidad de reproducción ajusta automáticamente una velocidad de reproducción de los datos de audio en respuesta a la identificación de la condición. La condición puede ser, por ejemplo, una velocidad del habla, el ruido de fondo, una pausa rellena u otra condición.

La Figura 2 es un diagrama de bloques de una unidad 200 de ajuste de la velocidad de reproducción según una realización ejemplar de la presente invención. La unidad 200 de ajuste de la velocidad de reproducción puede ser usada para implementar la unidad 140 de ajuste de la velocidad de reproducción mostrada en la Figura 1. Debería apreciarse que la unidad 200 de ajuste de la velocidad de reproducción puede residir en otros tipos de sistemas. La unidad 200 de ajuste de la velocidad de reproducción incluye una pluralidad de módulos que puede ser implementada en un soporte lógico. En realizaciones alternativas puede usarse circuitería de soporte físico en lugar en un soporte lógico, o en combinación con el mismo, para llevar a cabo la gestión de datos de audio. Así, las realizaciones de la presente invención no están limitadas a ninguna combinación específica de circuitería de soporte físico y de soporte lógico.

La unidad 200 de ajuste de la velocidad de reproducción incluye una unidad extractora 210 de características. La unidad extractora 210 de características extrae características de los datos de audio que recibe. Según una realización de la presente invención, la unidad extractora 210 de características transforma los datos de audio de un dominio temporal a un dominio frecuencial e identifica características en el dominio frecuencial. En una realización, las características pueden basarse en energías de subbanda. En esta realización, las características pueden ser identificarse usando coeficientes cepstrales de frecuencias Mel o usando otras técnicas o procedimientos. Según una realización alternativa, las características pueden basarse en características de fonema. En esta realización, las características de fonema pueden ser identificadas por coincidencia de patrones o clasificación de patrones contra señales de voz de referencia, usando un modelo oculto de Márkov, un alineamiento de Viterbi o saltos temporales dinámicos o usando otras técnicas o procedimientos. Debería apreciarse que las características pueden basarse en otras propiedades y ser identificadas usando otras técnicas.

La unidad 200 de ajuste de la velocidad de reproducción incluye una unidad integradora 220 de la tasa de cambio. La unidad integradora 220 de la tasa de cambio reconoce una condición en la que los datos de audio incluyen un habla que es producida con una velocidad que ha cambiado. Según una realización, la unidad integradora 220 de la tasa de cambio produce una salida que corresponde a la tasa de cambio, promediada en el tiempo, de las características desde la unidad 210. El integrador 220 de la tasa de cambio puede generar un valor de control de la velocidad de reproducción que puede ser usado para ajustar la velocidad de reproducción de los datos de audio. Según una realización en la que las características se basan en energías de subbanda, la unidad integradora 220 de la tasa de cambio puede medir una diferencia entre muestras consecutivas de una característica. Tomando una media de las mediciones de una pluralidad de características, se identifica una tasa global de cambio de las características. La tasa de cambio puede usarse para determinar una tasa de cambio del habla y un valor apropiado de control de la velocidad de reproducción que ha de ser generado. Según una realización en la que las características se basan en fonemas, la tasa de cambio de las clasificaciones de fonemas pueden ser promediadas en el tiempo para generar un valor apropiado de control de la velocidad de reproducción.

La unidad 200 de ajuste de la velocidad de reproducción puede incluir una unidad comparadora 230. La unidad comparadora 230 reconoce cuándo hay presentes otras condiciones en los datos de audio. La unidad comparadora 230 puede generar uno o más valores de control de la velocidad de reproducción que pueden ser usados para ajustar la velocidad de reproducción de los datos de audio en base a las condiciones. Según una realización de la unidad 200 de ajuste de la velocidad de reproducción, la unidad comparadora 230 puede comparar las

características de los datos de audio con características en modelos de voz que pueden reflejar condiciones diferentes. Las características de los datos de audio pueden compararse con modelos de voz que reflejan cantidades altas y bajas de ruido de fondo para determinar un grado de ruido de fondo presente en los datos de audio y la calidad de la grabación. Según una realización de la presente invención, si hay presente un grado elevado de ruido de fondo en los datos de audio, la unidad comparadora 230 genera un valor de control de la velocidad de reproducción que disminuye una velocidad de reproducción. Las características de los datos de audio pueden ser comparadas con modelos de voz que reflejan pausas en el habla o pausas rellenas con expresiones que no contribuyen al contenido de los datos de audio para determinar si puede acelerarse o editarse una porción de los datos de audio durante la reproducción. Debería apreciarse que también pueden detectarse de forma similar otras condiciones. Por ejemplo, la unidad comparadora 230 puede generar valores de control de la velocidad de reproducción para ajustar la velocidad de reproducción de los datos de audio en base a cambios en imágenes de vídeo.

La unidad 200 de ajuste de la velocidad de reproducción incluye una unidad 240 de procesamiento de datos de audio. La unidad 240 de procesamiento de datos de audio recibe uno o más valores de control de la velocidad de reproducción. Cuando la unidad 240 de procesamiento de datos de audio recibe más de un valor de control de la velocidad de reproducción, puede tomar una media de los valores, calcular una media ponderada de los valores o tomar un valor mínimo o máximo. La unidad 240 de procesamiento de datos de audio también recibe los datos de audio que deben reproducirse y ajusta una velocidad de reproducción de los datos de audio en respuesta a los uno o más valores de control de la velocidad de reproducción. Según una realización de la presente invención, la unidad 240 de procesamiento de datos de audio puede ajustar la velocidad de reproducción llevando a cabo un muestreo selectivo, solapamiento y suma sincronizados, escalado armónico o llevando a cabo otros procedimientos u otras técnicas.

La unidad 200 de ajuste de la velocidad de reproducción puede incluir una unidad 250 de retardo temporal. La unidad 250 de retardo temporal retarda el momento en el que la unidad 240 de procesamiento de datos de audio recibe los datos de audio. Insertando un retardo, la unidad 250 de retardo temporal permite que la unidad integradora 220 de la tasa de cambio y la unidad comparadora 230 analicen las características de los datos de audio y generen valores apropiados de control de la velocidad de reproducción antes de que los datos de audio sean reproducidos por la unidad 240 de procesamiento de datos de audio.

Según una realización de la unidad 200 de ajuste de la velocidad de reproducción, la unidad extractora 210 de características, la unidad integradora 220 de la tasa de cambio, la unidad comparadora 230, la unidad 240 de procesamiento de datos de audio y la unidad 250 de retardo temporal pueden ser implementadas usando cualquier procedimiento, técnica o circuitería apropiados. Debería apreciarse que algunos de los componentes mostrados pueden ser opcionales, tales como la unidad comparadora 230 y la unidad 250 de retardo temporal.

La Figura 3 es un diagrama de bloques de una unidad integradora 300 de la tasa de cambio según una realización ejemplar de la presente invención. La unidad integradora 300 de la tasa de cambio puede ser implementada como una realización de la unidad integradora 220 de la tasa de cambio mostrada en la Figura 2. La unidad integradora 300 de la tasa de cambio incluye una pluralidad de unidades de diferencia. Según una realización de la unidad integradora 300 de la tasa de cambio, se proporciona una unidad de diferencia para cada tipo de característica procesada por la unidad integradora 300 de la tasa de cambio. El bloque 310 representa una primera unidad de diferencia. El bloque 311 representa una  $n$ -ésima unidad de diferencia, pudiendo ser  $n$  cualquier número. Las unidades 310 y 311 de diferencia comparan las propiedades de las características recibidas desde una unidad extractora de características de diferentes periodos de tiempo y calculan un valor absoluto de la diferencia (valor absoluto de la diferencia). Por ejemplo, la unidad 310 de diferencia puede calcular el valor absoluto de la diferencia de una característica de un primer tipo identificada en el instante  $t$  y de una característica del primer tipo identificada en  $t-1$ . La unidad 311 de diferencia puede calcular el valor absoluto de la diferencia de una característica de un segundo tipo identificada en el instante  $t$  y una característica del segundo tipo identificada en  $t-1$ .

La unidad integradora 300 de la tasa de cambio puede incluir una pluralidad de unidades de ponderación opcionales. Según una realización de la unidad integradora 300 de la tasa de cambio, se proporciona una unidad de ponderación para cada tipo de característica procesado por la unidad integradora 300 de la tasa de cambio. El bloque 320 representa una primera unidad de ponderación. El bloque 321 representa una  $n$ -ésima unidad de ponderación. Cada unidad de ponderación pondera el valor absoluto de la diferencia de un tipo de característica. Las unidades 320 y 321 de ponderación pueden aplicar un coeficiente de ponderación a los valores absolutos de la diferencia en base a las propiedades de las características.

La unidad integradora 300 de la tasa de cambio incluye una unidad sumadora 330. La unidad sumadora 330 suma los valores absolutos ponderados de diferencias recibidos por las unidades 320 y 321 de ponderación.

La unidad integradora 300 de la tasa de cambio incluye una unidad 340 de control de la velocidad de reproducción. La unidad 340 de control de la velocidad de reproducción genera un valor de control de la velocidad de reproducción a partir de la suma de los valores absolutos ponderados de diferencias. Según una realización de la unidad integradora 300 de la tasa de cambio, la unidad 340 de control de la velocidad de reproducción toma una media de

la suma de los valores absolutos ponderados de diferencias. Según una realización alternativa, la unidad 340 de control de la velocidad de reproducción integra la suma de los valores absolutos ponderados de diferencias en un periodo de tiempo.

5 La Figura 4 es un diagrama de flujo que ilustra un procedimiento para la gestión de datos de audio según una primera realización de la presente invención. En 401, los datos de audio son transformados de un dominio temporal a un dominio frecuencial. Según una realización de la presente invención, puede aplicarse una transformada rápida de Fourier a los datos de audio para transformarlos de un dominio temporal a un dominio frecuencial.

10 En 402, se identifican características de los datos de audio transformados al dominio frecuencial. Según una realización de la presente invención, las características pueden basarse en energías de subbanda. En esta realización, las características se identifican usando coeficientes cepstrales de frecuencias Mel. Según una realización alternativa de la presente invención, las características puede basarse en características de fonemas.

15 En 403, se genera una medida de la tasa de cambio de las características. Según una realización de la presente invención, la medida de la tasa de cambio de las características puede ser generada analizando las características de los datos de audio. La medida de la tasa de cambio de las características puede ser usada para identificar una condición en la que ha cambiado una velocidad vocal de una persona que habla. Según una realización de la presente invención, se genera un valor de control de la velocidad de reproducción.

20 En 404, se ajusta una velocidad de reproducción de los datos de audio. El ajuste se basa en la tasa de cambio de las características determinada en 403, según se refleja por medio del valor de control de la velocidad de reproducción. Según una realización de la presente invención, la velocidad de reproducción del audio puede ajustarse llevando a cabo un muestreo selectivo, solapamiento y suma sincronizados, escalado armónico o llevando a cabo otros procedimientos.

25 La Figura 5 es un diagrama de flujo que ilustra un procedimiento para la gestión de datos de audio según una segunda realización de la presente invención. En 501, los datos de audio son transformados de un dominio temporal a un dominio frecuencial. Según una realización de la presente invención, puede aplicarse una transformada rápida de Fourier a los datos de audio para transformarlos de un dominio temporal a un dominio frecuencial.

En 502, se identifican características de los datos de audio transformados al dominio frecuencial. Según una realización de la presente invención, las características pueden basarse en energías de subbanda. En esta realización, las características se identifican usando coeficientes cepstrales de frecuencias Mel. Según una realización alternativa de la presente invención, las características puede basarse en características de fonemas.

30 En 503, se genera una medida de la tasa de cambio de las características. Según una realización de la presente invención, la medida de la tasa de cambio de las características puede ser generada analizando las características de los datos de audio. La medida de la tasa de cambio de las características puede ser usada para identificar una condición en la que ha cambiado una velocidad vocal de una persona que habla. Según una realización de la presente invención, se genera un valor de control de la velocidad de reproducción.

35 En 504, se comparan las características de los datos de audio identificados en 502 con características en modelos de voz que reflejan condiciones diferentes para determinar la presencia de las condiciones. Por ejemplo, pueden compararse características de los datos de audio con modelos de voz que reflejan cantidades altas y bajas de ruido de fondo para determinar un grado de ruido de fondo presente en los datos de audio. Las características de los datos de audio pueden ser comparadas también con modelos de voz que reflejan pausas en el habla o pausas rellenas con expresiones que no contribuyen al contenido de los datos de audio para determinar si puede acelerarse una porción de los datos de audio durante la reproducción o si puede eliminarse u omitirse. Debería apreciarse que también pueden detectarse otras condiciones. Según una realización de la presente invención, se generan uno o más valores de control de la velocidad de reproducción.

45 En 505, se determina un ajuste de la velocidad de reproducción a partir de los valores de control de la velocidad de reproducción generados. Según una realización de la presente invención, los valores de control de la velocidad de reproducción son promediados para determinar el grado de ajuste a realizar en la velocidad de reproducción de los datos de audio. Según una realización alternativa de la presente invención, se toma una media ponderada de los valores de control de la velocidad de reproducción para determinar el grado de ajuste a realizar en la velocidad de reproducción de los datos de audio.

50 En 506, se ajusta una velocidad de reproducción de los datos de audio. El ajuste se basa en la media o la media ponderada de los valores de control de la velocidad de reproducción generados. Según una realización de la presente invención, la velocidad de reproducción del audio puede ajustarse llevando a cabo un muestreo selectivo, solapamiento y suma sincronizados, escalado armónico o llevando a cabo otros procedimientos.

55 La Figura 6 es un diagrama de flujo que ilustra un procedimiento para la generación de un valor de control de la velocidad de reproducción según una realización de la presente invención. El procedimiento mostrado en la Figura 6 puede ser usado para implementar los bloques 403 y 503 mostrados en las Figuras 4 y 5. En 601, se determinan

valores absolutos de diferencias para una pluralidad de tipos de características. Según una realización de la presente invención, el valor absoluto se toma de la diferencia de cada tipo de característica medida en un primer instante y un segundo instante.

5 En 602, se ponderan los valores absolutos de diferencias de los tipos de características. Según una realización de la presente invención, los valores absolutos de diferencias de los tipos de características se ponderan en base a propiedades de las características.

En 603, se suman conjuntamente los valores absolutos ponderados de diferencias.

10 En 604, se genera un valor de control de la velocidad de reproducción a partir de la suma de los valores absolutos ponderados de diferencias. Según una realización de la presente invención, se toma una media de la suma de los valores absolutos ponderados de diferencias. Según una realización alternativa, la suma de los valores absolutos ponderados de diferencias se integra en un periodo de tiempo.

15 Según una realización de la presente invención, un procedimiento de gestión de los datos de audio incluye la identificación de una condición en los datos de audio y el ajuste automático de una velocidad de reproducción de los datos de audio en respuesta a la identificación de la condición. La condición puede incluir un cambio producido en la velocidad del habla, la presencia de ruido de fondo, la presencia de una pausa o de una pausa rellena en el habla. Ajustando automáticamente la velocidad de reproducción, las realizaciones de la presente invención permiten que los oyentes se concentren en los datos de audio que se están reproduciendo sin tener que distraerse por tener que ajustar manualmente la velocidad de reproducción.

20 Las Figuras 4-6 son diagramas de flujo que ilustran procedimientos según realizaciones de la presente invención. algunas de las técnicas ilustrados en estas figuras pueden llevarse a cabo secuencialmente, en paralelo o en un orden distinto del descrito. Debería apreciarse que no se requiere que se lleven a cabo todas las técnicas descritas, que pueden añadirse técnicas adicionales o que algunas de las técnicas ilustradas pueden ser sustituidas con otras técnicas.

25 Las realizaciones de la presente invención pueden ser proporcionadas como un producto de programa de ordenador, o soporte lógico, que puede incluir un artículo fabricado en un medio accesible por máquina o legible por máquina que tiene instrucciones. Las instrucciones en el medio accesible por máquina o legible por máquina pueden ser usadas para programar un sistema de ordenador u otro dispositivo electrónico. El medio legible por máquina puede incluir, sin limitación, disquetes flexibles, discos ópticos, CD-ROM y discos magneto-ópticos u otro tipo de medios/medio legible por máquina adecuados para almacenar o transmitir instrucciones electrónicas. Las técnicas descritas en el presente documento no están limitadas a ninguna configuración particular de soporte lógico. Pueden encontrar aplicabilidad en cualquier entorno informático o de procesamiento. Las expresiones "medio accesible por máquina" o "medio legible por máquina" usadas en el presente documento incluirán cualquier medio que sea capaz de almacenar, codificar o transmitir una secuencia de instrucciones para su ejecución por una máquina y que hagan que la máquina lleve a cabo uno cualquiera de los procedimientos descritos en el presente documento. Además, es común en la técnica hablar de que un soporte lógico, en una u otra forma (por ejemplo, programa, procedimiento, proceso, aplicación, módulo, unidad, lógica, etcétera) emprenda una acción o provoque un resultado. Tales expresiones son meramente una forma apropiada de enunciar que la ejecución del soporte lógico por un sistema de procesamiento hace que el procesador lleve a cabo una acción para producir un resultado.

40 En la anterior memoria, las realizaciones de la presente invención han sido descritas con referencia a realizaciones ejemplares específicas de la misma. Sin embargo, será evidente que pueden realizarse a la misma diversas modificaciones y diversos cambios sin apartarse del espíritu y el alcance más amplios de las realizaciones de la presente invención. En consecuencia, la memoria y los dibujos han de ser considerados en un sentido ilustrativo, no restrictivo.

**REIVINDICACIONES**

1. Un procedimiento para el ajuste automático de la velocidad de reproducción de datos de audio que comprende:
  - 5 identificar (502) una primera condición en los datos de audio relativa a una velocidad del habla y una segunda condición en los datos de audio relativa a ruido de fondo convirtiendo (501) los datos de audio desde un dominio temporal a un dominio frecuencial, extrayendo características de los datos de audio en el dominio frecuencial y midiendo (503) una tasa de cambio de las características extraídas en el dominio frecuencial generando uno o más valores (401-403; 501-503) de control de la velocidad de reproducción en respuesta a la primera condición, y comparar (504) las características con un modelo de voz para generar uno o más valores adicionales de control de la velocidad de reproducción en respuesta a la segunda condición; y
  - 10 ajustar automáticamente (506) una velocidad de reproducción de los datos de audio en respuesta a todos los valores (404; 506) de control de la velocidad de reproducción.
2. El procedimiento de la reivindicación 1 en el que el ajuste automático de una velocidad de reproducción de los datos de audio en respuesta a todos los valores de control de la velocidad de reproducción comprende:
  - 15 tomar una media de todos los valores de control de la velocidad de reproducción generados; y
  - aplicar la media de todos los valores (506) de control de la velocidad de reproducción.
3. El procedimiento de la reivindicación 1 en el que las características comprenden, al menos, una de:
  - (a) energías de subbanda; o
  - (b) características (502) de fonema.
- 20 4. El procedimiento de la reivindicación 1 en el que el ajuste de la velocidad de reproducción de los datos de audio comprende llevar a cabo, al menos, uno de:
  - (a) muestreo selectivo;
  - (b) solapamiento y suma sincronizados; o
  - (c) escalado armónico.
- 25 5. Un medio de almacenamiento de instrucciones accesible por máquina que, cuando se ejecuta, hace que la máquina lleve a cabo el procedimiento de una cualquiera de las reivindicaciones 1-4.
6. Un aparato (200) de ajuste de la velocidad de reproducción que comprende:
  - una unidad extractora (210) de características para convertir datos de audio de un dominio temporal a un dominio frecuencial e identificar características de los datos de audio en el dominio frecuencial;
  - 30 una unidad integradora (220) de la tasa de cambio para identificar una condición relativa a una velocidad del habla a partir del cambio de la tasa de las características identificadas en el dominio frecuencial y generar uno o más valores de control de la velocidad de reproducción;
  - una unidad comparadora (230) para comparar las características de los datos de audio identificados en el dominio frecuencial con características en modelos de voz para identificar una condición relativa al ruido de fondo y para generar otro u otros valores de control de la velocidad de reproducción; y
  - 35 una unidad procesadora (240) de datos de audio para ajustar una velocidad de reproducción de los datos de audio en respuesta a todos los valores de control de la velocidad de reproducción.
7. El ajuste de la velocidad de reproducción de la reivindicación 6 en el que la unidad procesadora (240) de datos de audio toma una media de los uno o más valores de control de la velocidad de reproducción generados a partir del integrador de la tasa de cambio y de la unidad comparadora.
- 40 8. El aparato de ajuste de la velocidad de reproducción de la reivindicación 6 en el que la unidad procesadora (240) de datos de audio toma una media ponderada de los uno o más valores de control de la velocidad de reproducción generados a partir del integrador de la tasa de cambio y de la unidad comparadora.
9. El aparato de ajuste de la velocidad de reproducción de la reivindicación 6 en el que la unidad procesadora (240) de datos de audio toma un mínimo o un máximo de los uno o más valores de control de la velocidad de reproducción generados a partir del integrador de la tasa de cambio y de la unidad comparadora.
- 45

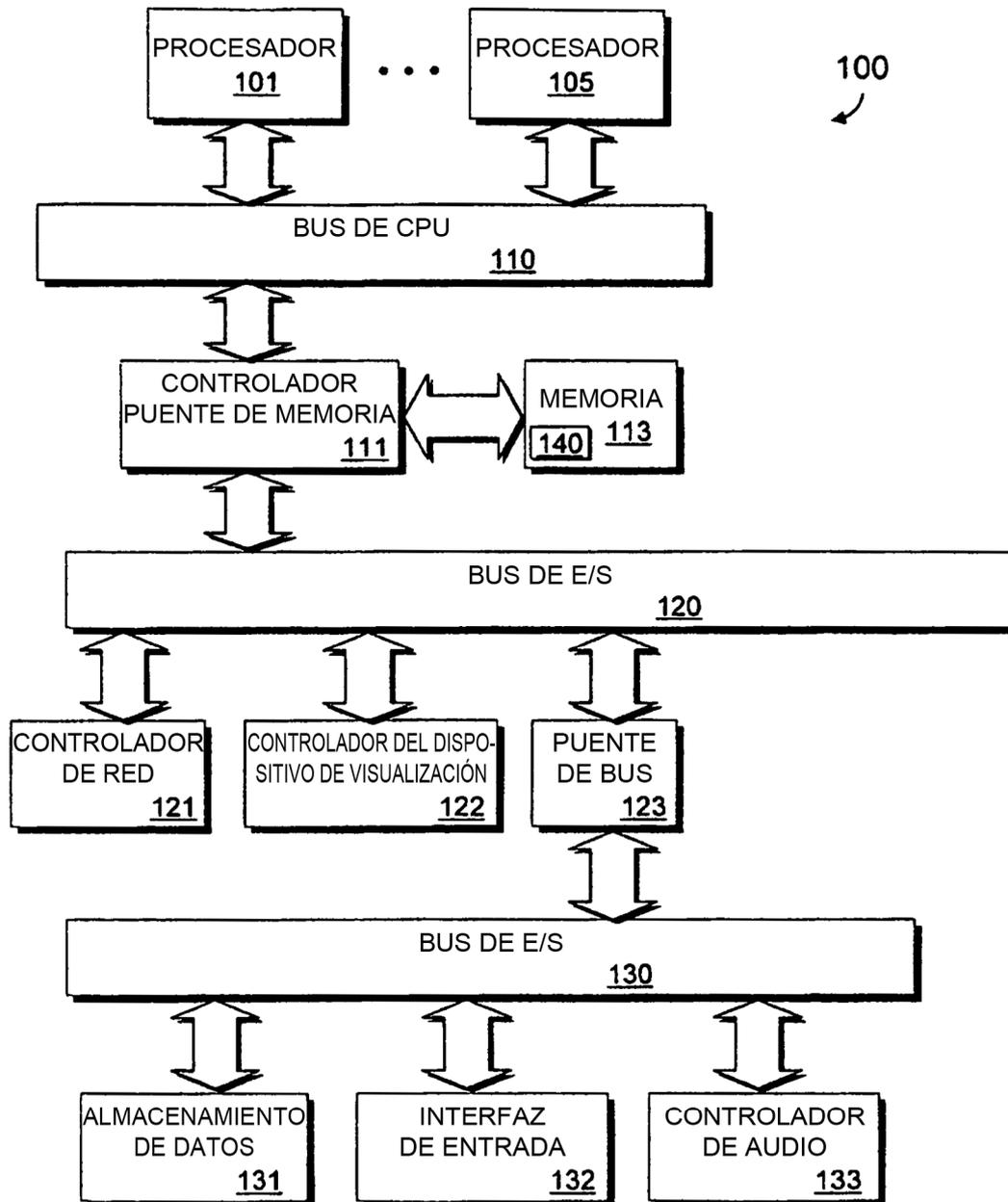


FIG. 1

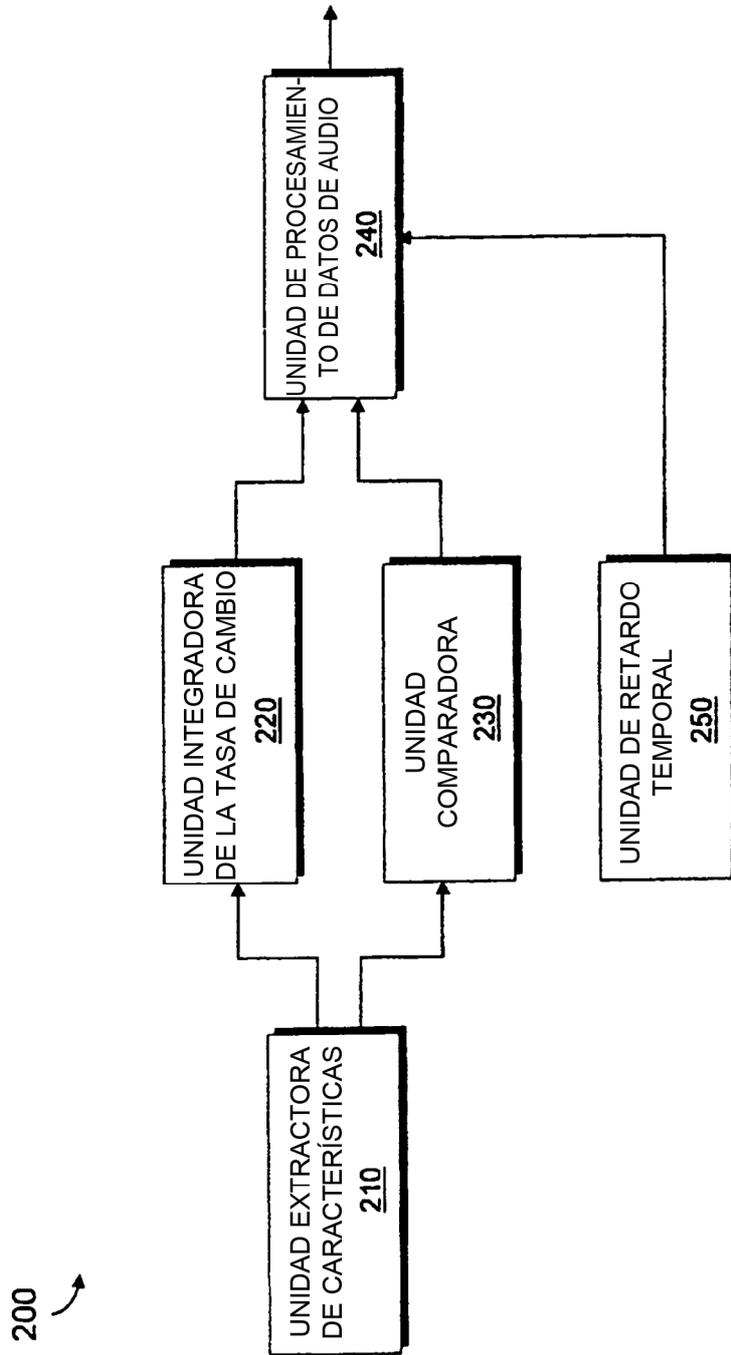
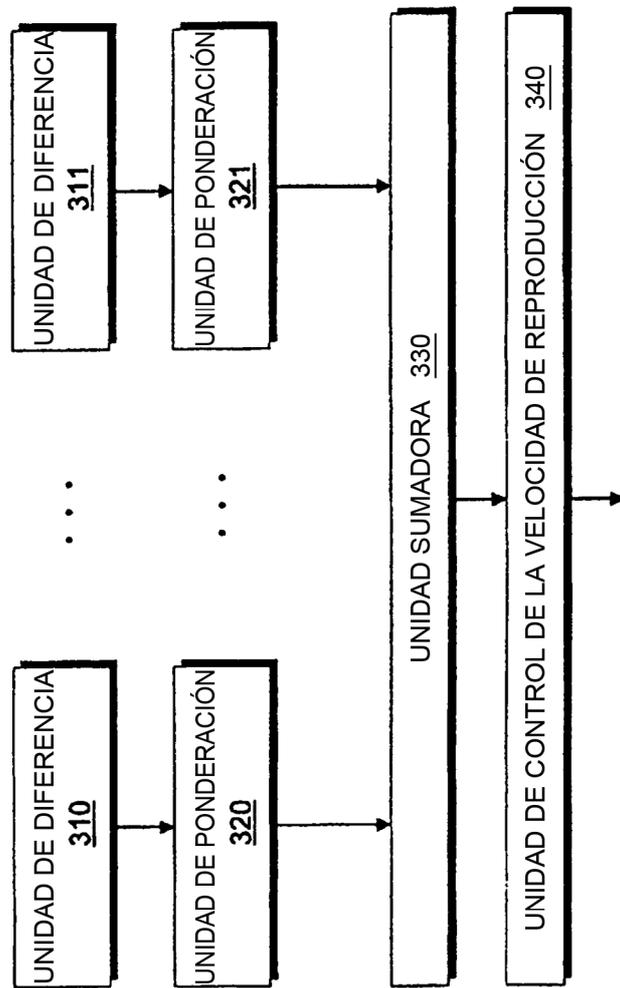


FIG. 2

300 ↷



**FIG. 3**

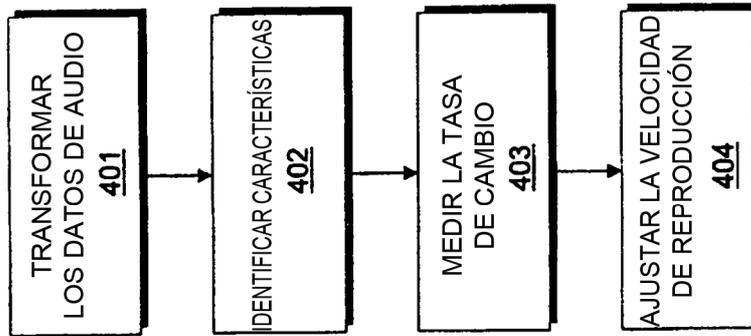


FIG. 4

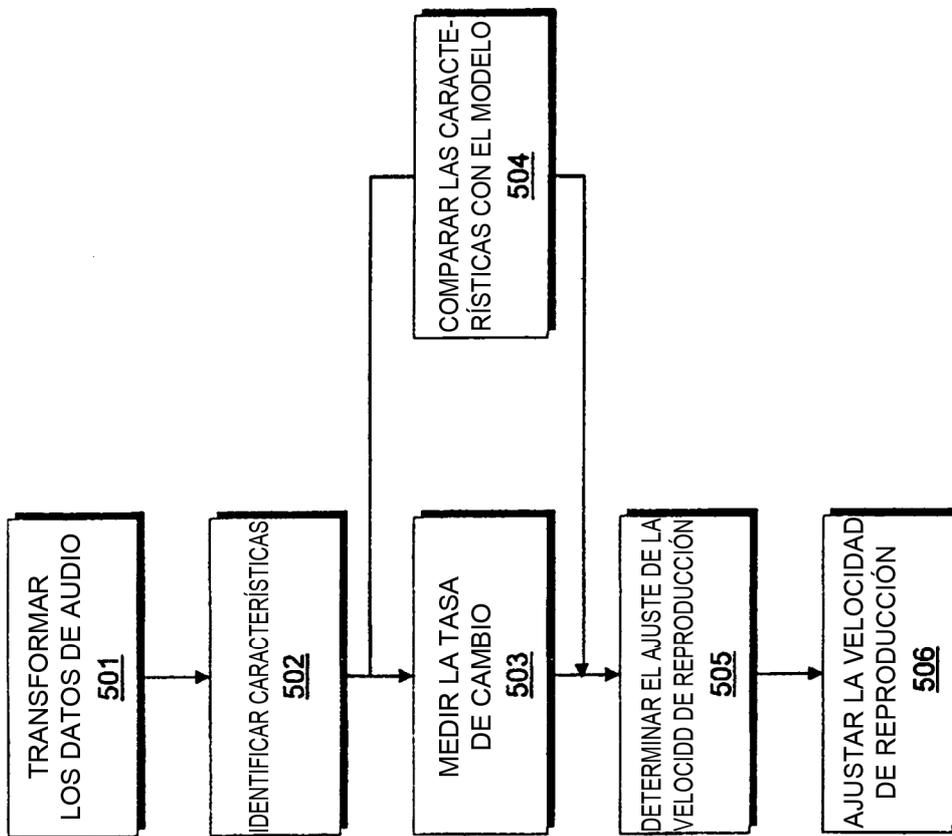
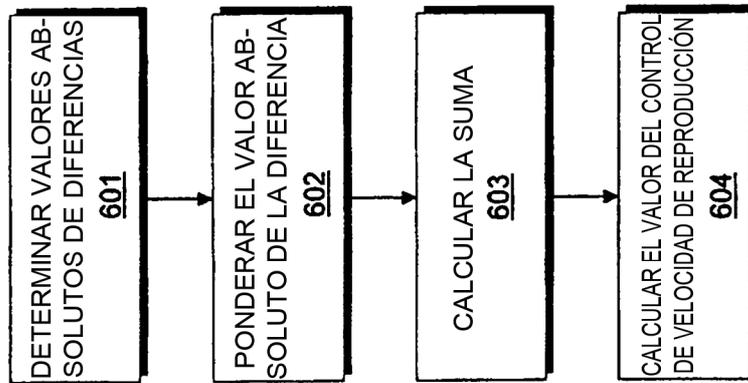


FIG. 5



**FIG. 6**