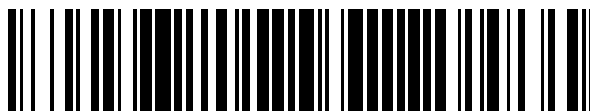


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 548 977**

51 Int. Cl.:

H03G 3/32 (2006.01)

G10L 21/0208 (2013.01)

H03G 9/00 (2006.01)

H03G 9/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.05.2009 E 09739994 (3)**

97 Fecha y número de publicación de la concesión europea: **29.07.2015 EP 2283484**

54 Título: **Sistema y método para generación de sonido dinámico**

30 Prioridad:

02.05.2008 US 150970

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

22.10.2015

73 Titular/es:

**DOLBY LABORATORIES LICENSING
CORPORATION (100.0%)**

**100 Potrero Avenue
San Francisco, CA 94103-4813, US**

72 Inventor/es:

TAENZER, JON C.

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 548 977 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

Sistema y método para generación de sonido dinámico

5 CAMPO TÉCNICO

La presente descripción está relacionada con reproducción de audio, por ejemplo en sistemas de comunicación bidireccionales como por ejemplo teléfonos móviles y walkie-talkies.

10 ANTECEDENTES

El ruido ambiental puede interferir a veces con el suministro de información de audio. En un sistema de comunicación bidireccional por ejemplo, en el cual el hablante de extremo lejano está en una posición alejada del oyente de extremo cercano, el hablante de extremo lejano, ignorante de las condiciones de ruido en la posición del oyente, puede no adoptar medidas para compensar el hecho de que se produzcan episodios de ruido disruptivos (instantáneos o sostenidos). Por ejemplo, el hablante, no consciente de un coche que pasa por la posición del oyente, puede no elevar su voz para mantener la audibilidad para el oyente, y las palabras del hablante pueden no ser oídas o comprendidas por el oyente, incluso si el sistema fuera eléctricamente y mecánicamente capaz de manejar dicha compensación. La incapacidad del oyente de discernir el habla del hablante en dichas circunstancias es debida al fenómeno psicológico bien conocido llamado "enmascaramiento", que es cuando el ruido local, al ser suficientemente alto, cubre o enmascara, a la señal sonora de extremo lejano reproducida. Por supuesto, este problema no está limitado a los sistemas de comunicación bidireccionales, y el ruido ambiental puede interferir de manera similar con voces pregrabadas, o con cualquier información de audio prealmacenada que se esté reproduciendo. La Patente US 2003/0223597 A1 describe un método para control de la ganancia dinámica.

La Patente US 5.434.922 describe un sistema para compensar el nivel de ruido en el interior de un vehículo midiendo el nivel de la música y el nivel de ruido dentro del vehículo.

La Patente US 6.011.853 está relacionada con ecualización en un teléfono digital.

La Patente US 6.507.623 B1 describe un método para reducción de ruido.

La Patente US 2006/0034447 A1 describe un método para captura clara de señal.

Las Patentes US 5.907.823, US 4.887.299, JP 2000-196388 A, EP 2066025 A2 y JP H05-110362 A describen métodos para control del rango dinámico.

35 VISIÓN GLOBAL

La invención está definida por las reivindicaciones independientes 1 y 9. En las reivindicaciones dependientes se describen realizaciones preferentes.

40 BREVE DESCRIPCIÓN DE LOS DIBUJOS

Los dibujos adjuntos, los cuales están incorporados en esta especificación y constituyen una parte de la misma, ilustran uno o más ejemplos de realizaciones y, junto con la descripción de realizaciones de ejemplo, sirven para explicar los principios e implementaciones de las realizaciones.

45 En los dibujos:

La Figura 1A es un diagrama de un sistema 100 de comunicación de audio bidireccional que permite que dos usuarios se comuniquen a distancia el uno con el otro.

La Figura 1B es un diagrama de bloques de un dispositivo 102 de comunicación del sistema 100 de comunicación de la Figura 1A.

50 La Figura 2 es un diagrama de bloques de un sistema de generación de sonido dinámico.

La Figura 3 es un diagrama de bloques de un sistema 300 de manipulación del rango dinámico que puede constituir parte del procesador 202 de la Figura 2 (y del controlador 110 de la Figura 1B).

La Figura 4 es un diagrama de bloques relacionado con un proceso de transformada de Fourier de tiempo corto (STFT) estándar.

55 La Figura 5 es un diagrama de flujo de un proceso de manipulación del rango dinámico.

La Figura 6 es un diagrama de bloques de los módulos o circuitos de procesamiento utilizados en la implementación del proceso de la Figura 5.

La Figura 7A es un gráfico de diferentes curvas de compresión/expansión.

La Figura 7B es un gráfico de curvas de nivel de ganancia para una sub-banda.

60 La Figura 7C es una curva de nivel del nivel de la señal de reproducción utilizada para excitar el altavoz 206.

DESCRIPCIÓN DE REALIZACIONES DE EJEMPLOS

En este documento se describen realizaciones de ejemplo en el contexto de un sistema y método para generación de sonido dinámico. Las personas con experiencia ordinaria en la técnica se darán cuenta de que la descripción que sigue es sólo ilustrativa y que no tiene intención de ser limitativa de ninguna manera. Otras realizaciones se insinuarán rápidamente a dichas personas con experiencia en la técnica que se benefician de esta descripción. Se

hará ahora referencia en detalle a implementaciones de las realizaciones de ejemplo como las ilustradas en los dibujos adjuntos. En la medida de lo posible, para hacer referencia a elementos iguales o similares se utilizarán los mismos indicadores de referencia en todos los dibujos y a lo largo de toda la descripción que sigue.

Para mayor claridad, no se muestran y se describen todos los rasgos rutinarios de las implementaciones descritas en este documento. Por supuesto, se observará que en el desarrollo de cualquier implementación real de este tipo, se deben tomar numerosas decisiones específicas de implementación para alcanzar los objetivos específicos del desarrollador, tales como el cumplimiento de restricciones relacionadas con la aplicación y con el negocio, y que estos objetivos específicos variarán de una implementación a otra y de un desarrollador a otro. Además, se observará que un esfuerzo de desarrollo como este podría ser complejo y de larga duración, pero que sin embargo sería una tarea rutinaria de ingeniería para las personas con experiencia ordinaria en la técnica que se benefician de esta descripción.

De acuerdo con esta descripción, los componentes, etapas del proceso, y/o estructuras de datos descritos en este documento se pueden implementar utilizando diferentes tipos de sistemas operativos, plataformas informáticas, programas informáticos, y/o máquinas de propósito general. Además, las personas con experiencia ordinaria en la técnica reconocerán que también se pueden utilizar dispositivos que tengan una naturaleza de propósito menos general, tales como dispositivos conectados por cable, matrices de puertas programables in situ (FPGAs), circuitos integrados específicos para aplicaciones (ASICs), o similares, sin apartarse del alcance de los conceptos de la invención definidos por las reivindicaciones adjuntas.

En los casos en que un método que comprende una serie de etapas de proceso es implementado mediante un ordenador o una máquina y esas etapas de proceso se pueden almacenar como una serie de instrucciones que pueden ser leídas por la máquina, se pueden almacenar en un medio físico tal como un dispositivo informático de memoria (por ejemplo, una ROM (Memoria de Sólo Lectura), una PROM (Memoria de Sólo Lectura Programable), una EEPROM (Memoria de Sólo Lectura Programable con Borrado Eléctrico), una Memoria FLASH, una memoria USB, y similares), en un medio de almacenamiento magnético (por ejemplo, una cinta, un disco magnético, y similares), en un medio de almacenamiento óptico (por ejemplo, un CD-ROM, un DVD-ROM, una tarjeta de papel, una cinta de papel y similares) y otros tipos de memoria de programa.

Las realizaciones de ejemplo descritas en este documento se presentan en el contexto de un proceso implementado utilizando un proceso de señal digital. Se reconocerá que cada etapa de proceso se puede conseguir con implementaciones alternativas, por ejemplo, utilizando circuitos analógicos. Aunque el hardware que soporta una implementación analógica parecería diferente a la implementación hardware en el dominio digital, la naturaleza fundamental de cada una de las correspondientes etapas de proceso es equivalente. De esta forma, los procesos descritos en este documento tienen el objetivo de ser aplicables a cualquier implementación hardware en el dominio analógico o digital.

La Figura 1A es un diagrama de un sistema 100 de comunicación de audio bidireccional que permite que dos usuarios se comuniquen a distancia el uno con el otro. Cada usuario está provisto de un dispositivo 102 de comunicación, mostrado con mayor detalle en el diagrama de bloques de la Figura 1B. Cada dispositivo 102 de comunicación incluye micrófono 104, altavoz 106, transmisor-receptor 108 y procesador o controlador 110. En un primer "circuito" de comunicación, la voz del usuario que se encuentra en una posición remota o de extremo lejano es recogida por un micrófono 104 del dispositivo 102 de comunicación en esa posición del usuario, y es transmitida, de forma inalámbrica o de otra forma, para que sea reproducida por un altavoz 106 del dispositivo 102 de comunicación en la posición del usuario local o de extremo cercano. De manera similar, en un segundo "circuito" de comunicación, la voz del usuario que se encuentra en la posición local o de extremo cercano es recogida por un micrófono 104 de un dispositivo 102 de comunicación de extremo cercano y es reproducida por un altavoz 106 en la posición remota o de extremo lejano.

Se considera que el sistema 100 de comunicación es un sistema bidireccional, ya que contiene dos "circuitos" de comunicación como se ha descrito. Sin embargo, se debería comprender que las implementaciones descritas en este documento están relacionadas con los "circuitos" de comunicación individualmente y, por lo tanto, no están limitadas a sistemas bidireccionales. Al contrario, también son aplicables a sistemas unidireccionales, en los cuales un usuario local o de extremo cercano es sólo capaz de oír a un usuario remoto, y no está equipado para hablar al usuario remoto, o viceversa. De forma incluso más general, las implementaciones descritas en este documento son aplicables a sistemas que pueden ser exclusivamente para reproducción o presentación de información de audio, como por ejemplo música, señales sonoras y voces pregrabadas, con independencia del estado o posición de la fuente de la información de audio, y no tiene por qué estar implicado ningún usuario remoto o ninguna fuente de audio remota. Estos sistemas incluyen por ejemplo sistemas de audio portátiles y no-portátiles tales como walkmans, reproductores de discos compactos, equipos de sonido domésticos, televisores, agendas electrónicas (PDAs), etc. En dichos sistemas, a diferencia de en el sistema 100 de comunicación bidireccional, la reproducción no se efectúa necesariamente en tiempo real, es decir, la información de audio no se presenta necesariamente al mismo tiempo que es creada.

Volviendo a la Figura 1B, la información que se espera que transmita el transmisor-receptor 108 en este ejemplo son señales sonoras tales como la voz del usuario, que es recogida por el micrófono 104 y convertida en señales eléctricas que son reenviadas al transmisor-receptor directamente, o por medio de un controlador 110 representado. Cuando la información recogida se hace pasar a través del controlador 110, ésta se puede empaquetar de forma apropiada para su transmisión de acuerdo con la aplicación y/o protocolo concretos que deban ser observados entre los dispositivos 102 del sistema 100 de comunicación. Después de este empaquetado, el cual puede ser uno de entre numerosos tipos de modulación, por ejemplo, la información es reenviada al transmisor-receptor 108 para su transmisión. A la inversa, el transmisor-receptor 108 sirve para reenviar la información que recibe, de forma inalámbrica o de otra forma, al controlador 110 para su “desempaquetado” y, como se detalla más adelante, para su procesamiento y manipulación de tal manera que cuando la información se convierta a forma acústica durante su reproducción por el altavoz 106, permanezca inteligible –o conserve en la medida de lo posible su mensaje o carácter original– con independencia del entorno de ruido en el cual está inmerso el usuario que escucha.

El transmisor-receptor 108 está configurado para realizar transmisión y/o recepción de información, y puede tener la forma de un único componente. De forma alternativa, se pueden utilizar componentes independientes específicos para cada una de estas dos funciones. La transmisión se puede realizar de forma inalámbrica, por medio de señales de radio modulada por ejemplo, o de una forma con conexión por cable utilizando cableado eléctrico convencional, o incluso de forma óptica utilizando fibras o a través de línea visual directa.

Dado que, en el ejemplo de las Figuras 1A y 1B, el hablante de extremo lejano se encuentra en una posición alejada del oyente de extremo cercano, el hablante puede ignorar las condiciones de ruido en la posición del oyente, y el hablante puede no adoptar medidas para compensar el hecho de que se produzcan episodios de ruido disruptivo (instantáneo o sostenido) en la posición distante del oyente porque el hablante puede no ser consciente de que se producen. Para proporcionar un ejemplo sencillo, el hablante, no consciente de un coche que pasa por la posición del oyente, puede no elevar su voz para mantener su audibilidad para el oyente, y sus palabras pueden no ser oídas o comprendidas por el oyente, incluso si el sistema fuera eléctricamente o mecánicamente capaz de manejar dicha compensación. La incapacidad del oyente para discernir el habla del hablante en tales circunstancias es debida al fenómeno psicológico bien conocido llamado “enmascaramiento”, que es cuando, al ser suficientemente alto, el ruido local cubre, o enmascara, a la señal sonora de extremo lejano reproducida. No obstante, como se establece en este documento, se puede hacer que el controlador 110 sea consciente de los episodios de audio disruptivos, y dicho controlador, y, si es necesario, los otros componentes, pueden estar equipados para adoptar las medidas necesarias para compensarlos o para minimizar su impacto.

Como se explicó anteriormente, aunque la descripción incluida en este documento puede estar dentro del contexto de un sistema de comunicación bidireccional, esto no pretende ser limitativo, y también se contemplan aplicaciones en sistemas que son exclusivamente para reproducción de información de audio, como por ejemplo música, señales sonoras y voces pregrabadas, con independencia del estado o naturaleza de la fuente de información de audio, y en tiempo real o de otra manera. La Figura 2 ilustra una aplicación generalizada de este tipo, en la cual, en un sistema 200 de generación de sonido dinámico, un procesador 202 opera sobre información de audio proporcionada por una fuente 204 de información de audio, manipulando la información y adoptando las medidas necesarias para compensar condiciones ambientales de audición complicadas, antes de enviarla en forma de una excitación de salida o señal de reproducción a un altavoz 206 para su presentación o reproducción a un usuario. En el sistema 200, un indicador 208 de ruido de audio genera una representación o peso del ruido de audio ambiental en la posición de reproducción. En estos casos, los sistemas de reproducción pueden estar provistos de un micrófono, si no hay uno ya disponible. La manipulación y compensación se realiza en tiempo real y puede ser continua o en forma de muestreos instantáneos discretos. La representación o peso, a la cual en lo que sigue se puede hacer referencia como los indicios de ruido ambiental, o indicios de ruido, se proporciona al procesador 202, que la utiliza, en conjunto con la señal de información procedente de la fuente 204 de información, para efectuar la necesaria compensación en la reproducción.

El indicador 208 a partir del cual se pueden derivar los indicios puede ser un simple micrófono, o una matriz de micrófonos (por ejemplo el micrófono o los micrófonos 104 de la Figura 1B), que se utiliza(n) para detectar ruido ambiental en la posición de reproducción. De forma alternativa (o adicional), los indicios de ruido se pueden derivar a partir de operaciones de procesamiento auxiliares que se realizan en otra parte del sistema, o en un sistema conectado, con el mismo propósito o con un propósito relacionado, o con un propósito totalmente diferente. Por ejemplo, en un sistema bidireccional, los indicios de ruido se pueden derivar a partir de un algoritmo de reducción de ruido utilizado en el extremo cercano para mejorar una señal de audio saliente en presencia del ruido ambiental. Mediante dicho algoritmo de reducción de ruido se puede obtener una determinación del ruido ambiental de varias maneras, y esta determinación se puede utilizar para proporcionar los indicios de ruido que necesita el sistema 200 de generación de sonido dinámico para mejorar la reproducción. El algoritmo de reducción de ruido para la señal de audio saliente utiliza a menudo métodos multibanda para crear un conjunto de valores de atenuación que se aplican, mediante multiplicación, a la señal con ruido saliente. Los valores de atenuación pueden ser un número entre “0” y “1”. Cuando se aplican a la señal con ruido saliente, estos valores actúan reduciendo el ruido en dicha señal atenuando partes de la señal con ruido que se considera que son mayoritariamente o únicamente ruido, al tiempo que no atenuando, o atenuando hasta un grado menor, partes que se considera que son la señal deseada. El sistema 200 de generación de sonido dinámico puede obtener los indicios de ruido restando a “1” cada valor de

atenuación. A continuación, el sistema 200 de generación de sonido dinámico puede aplicar los valores de “anti-atenuación” derivados de esta forma a la señal con ruido original, para derivar de ese modo los indicios de ruido a partir del indicador 208 de ruido. Además, en una variación, explicada en detalle más adelante, puede ser deseable utilizar los propios valores de atenuación de la siguiente manera: 1) elevándolos al cuadrado de manera que representen un porcentaje de potencia, 2) sumando los valores resultantes dentro de cada banda de frecuencia para obtener una medida porcentual total de potencia de no-ruido por banda, 3) calculando la potencia total de la señal con ruido original en cada banda, y 4) multiplicando el porcentaje de ruido, el cual es el 100% menos el porcentaje de potencia de no-ruido, multiplicado por la potencia total para obtener una medida de potencia de sólo-ruido en cada banda.

Como se explica más adelante con mayor detalle, el sistema 200 de generación de sonido dinámico de la Figura 2 se puede configurar para que opere como un procesador de rango dinámico multibanda que utilice compresión para ajustar el rango dinámico de la señal que debe ser reproducida en el entorno de extremo cercano con ruido. El sistema 200 aplica control selectivo de la ganancia a la señal de reproducción de salida que se utiliza para excitar el altavoz 206, siendo dicho control de la ganancia una función de la información de audio procedente de la fuente 204 de información, y de las condiciones de ruido ambientales (los indicios de ruido) procedentes del indicador 208 de ruido. Dependiendo de las circunstancias concretas, partes de la señal impulsora de salida pueden no verse sometidas a ninguna compresión, o incluso pueden verse sometidas a expansión hacia abajo.

La Figura 3 es un diagrama de bloques de un sistema 300 de manipulación del rango dinámico que puede constituir parte del procesador 202 de la Figura 2 (y del controlador 110 de la Figura 1B). El procesador 202 puede incluir otros módulos o circuitos de procesamiento que tengan funciones diferentes a las que se describen más adelante en este documento. Aunque, como se ha explicado anteriormente, el procesamiento llevado a cabo por el sistema 300 de manipulación del rango dinámico se puede implementar como un proceso de banda única o como un proceso multibanda, el ejemplo de la Figura 3 está relacionado con la técnica multibanda. En ella, a cada banda del proceso multibanda se le denominará sub-banda y las realizaciones de ejemplo se analizarán en el contexto de un proceso de dos sub-bandas, aunque esto no pretende ser limitativo. En una realización de ejemplo, se selecciona el corte entre las dos sub-bandas de manera que esté en aproximadamente 1 kHz. Por lo general, las señales de audio por debajo de 1 kHz contienen sonidos vocales del habla que pueden denotar sonoridad del habla, mientras que las que están por encima de 1 kHz por lo general contienen los sonidos consonantes, con mayor contenido de información. Más concretamente, y sólo a modo de ejemplo, la sub-banda inferior se extiende desde aproximadamente 50 Hz hasta aproximadamente 1 kHz, mientras que la sub-banda superior se extiende desde aproximadamente 1 kHz hasta aproximadamente 3,5 kHz. En el ejemplo de dos sub-bandas, el módulo 306 de procesamiento de señal proporciona el proceso de manipulación del rango dinámico de la sub-banda inferior, mientras que el módulo 308 de procesamiento de señal proporciona el proceso de manipulación del rango dinámico de la sub-banda superior. Una implementación de una única banda requiere sólo un módulo 306 de procesamiento de señal, el cual estaría configurado entonces para incluir todo el espectro deseado, mientras que una implementación de muchas sub-bandas requiere módulos de procesamiento de señal adicionales, como se indica mediante los módulos de procesamiento de señal mostrados en líneas discontinuas en la Figura 3.

El sistema 300 de manipulación del rango dinámico recibe la señal indicadora de ruido de audio en una primera entrada 301 del mismo, y proporciona esa señal a un primer módulo 302 de descomposición en frecuencias, donde se divide en intervalos (bins) o bandas de frecuencia. Como se ha explicado anteriormente, la señal indicadora de ruido de audio es indicativa del ruido ambiental en la posición de reproducción. El sistema 300 de manipulación del rango dinámico recibe la señal de información de audio en una segunda entrada 303, y se la proporciona a un segundo módulo 304 de descomposición en frecuencias. El módulo 304 de descomposición en frecuencias divide también esta señal de información de audio en intervalos o bandas de frecuencia.

En los sistemas de comunicación bidireccional de las Figuras 1A y 1B, esta señal de información de audio es representativa de la voz del hablante situado en la posición remota, de extremo lejano, tal como es recogida por el micrófono 104 del dispositivo 102 de comunicación en la posición de extremo lejano y transmitida al dispositivo 102 de comunicación de extremo cercano donde debe ser reproducida. Sin embargo, de forma más general, la información de audio procede de una fuente 204 de información como se muestra en la Figura 2 y puede ser, por ejemplo, música almacenada, voces pregrabadas y similares, y no está limitada necesariamente a ser reproducida en tiempo real.

En el ejemplo de realización de dos sub-bandas, cada uno de los módulos 302 y 304 de descomposición en frecuencias aplica un proceso de transformada de Fourier de tiempo corto (STFT) estándar a las respectivas señales de entrada procedentes de las entradas 301 y 303. Este proceso se explica haciendo referencia al diagrama de flujo 400 de la Figura 4. Se toman (402) muestras en el tiempo de las señales y se dividen estas muestras en tramas (404), a las cuales se les aplican ventanas (406) (las técnicas de ventana que implican multiplicación son conocidas). Se puede realizar un completado con ceros (408), y las tramas de datos temporales completadas con ceros se transforman (410) en el dominio de la frecuencia utilizando, en este ejemplo, transformada rápida de Fourier (FFT). Otros métodos de transformación disponibles incluyen transformadas de coseno, transformadas de Hartley, transformadas de ondículas (transformadas wavelet), etc. A modo de ejemplo, se utilizan palabras de 16 bits y una frecuencia de muestreo de 8 ksp/s (miles de muestras por segundo), con una trama de 384 muestras a la que

se aplican ventanas con una ventana de audio de reconstrucción perfecta y que se completa con ceros hasta 512 muestras para incrementar la resolución en frecuencia del resultado. Se obtiene de esta forma un número total de 257 "Intervalos" de frecuencia (incluyendo la frecuencia de Nyquist y semi-intervalos DC). Se utiliza un avance de trama de 56 muestras para proporcionar un funcionamiento a 143 tramas por segundo. De esta manera, para

En la realización de ejemplo en el dominio de la frecuencia, los intervalos de frecuencia se ensamblan en grupos, o sub-bandas. En este ejemplo, el grupo de intervalos que va desde 1 kHz hasta 3,5 kHz forman una sub-banda (la "sub-banda superior") y el grupo de intervalos que va desde 50 Hz hasta 1 kHz forman una segunda sub-banda, independiente (la "sub-banda inferior"). De esta manera, cada uno de los dos conjuntos de intervalos que constituyen representaciones en el dominio de la frecuencia de la señal indicadora del ruido de audio y de la señal de información de audio, respectivamente, se dividen en dos grupos de intervalos, o sub-bandas, proporcionando una sub-banda alta y una sub-banda baja para cada una de la señal indicadora del ruido y de la señal de información de audio.

La realización anteriormente descrita se lleva a cabo en el dominio de la frecuencia. Sin embargo, en una realización alternativa en el dominio del tiempo, la descomposición en frecuencias se puede realizar con filtros paso banda. Para el ejemplo de dos sub-bandas, existen dos conjuntos de dos filtros paso banda, comprendiendo cada uno de dichos conjuntos un filtro paso banda que deja pasar el rango de frecuencias asociado para la sub-banda inferior y un filtro paso banda que deja pasar el rango de frecuencias asociado para la sub-banda superior. Uno de los conjuntos de dos filtros paso banda está asociado con la señal indicadora de ruido y uno de los conjuntos de dos filtros paso banda está asociado con la señal de información de audio. De esta manera, en esta realización de ejemplo, existen cuatro filtros paso banda que realizan el proceso de descomposición en frecuencias.

los subprocesadores 306, 308 desarrollan una señal de control de manipulación del rango dinámico, como se describe en la Figura 5, la cual muestra un diagrama de flujo del proceso, y en la Figura 6, la cual es un diagrama de bloques de los módulos o circuitos de procesamiento utilizados en su implementación. Para cada trama de datos, el proceso mostrado en la Figura 5 se realiza para cada sub-banda. Para cada una de las sub-bandas inferior y superior, el sub-procesador 306, 308 asociado realiza una determinación de potencia de la trama, la cual es la base del control de la ganancia aplicado a la señal de excitación de la reproducción que es enviada al altavoz 206 (Figura 2). Para derivar las determinaciones de potencia de la trama para las señales de ruido y de información respectivamente, se elevan al cuadrado los valores de magnitud para cada intervalo de frecuencia de cada señal y se suman los resultados de todos los intervalos para una sub-banda en todo el rango de frecuencia de la sub-banda. Esto se muestra en 501, 501' en la Figura 5 y en 601, 601' en la Figura 6. Cada determinación de potencia de la trama es realizada por circuitos o módulos de elevación al cuadrado situados dentro de 601, 601', por ejemplo el 601a (Figura 6), el cual realiza la elevación al cuadrado de los valores de magnitud para conformar valores de potencia del intervalo, y un circuito o módulo 601, 601' sumador, por ejemplo el 601b (Figura 6), que realiza la suma de los valores de potencia del intervalo dentro de la respectiva sub-banda de cada señal. Los valores resultantes de suma de potencias del intervalo son convertidos a continuación (503 y 503' de la Figura 5) en valores logarítmicos de suma de potencias del intervalo por convertidores 603, 603' logarítmicos (Figura 6), proporcionando un total de cuatro valores de potencia de la trama para cada trama: uno asociado con la señal de ruido en la sub-banda inferior, uno asociado con la señal de información en la sub-banda inferior, uno asociado con la señal de ruido en la sub-banda superior, y uno asociado con la señal de información en la sub-banda superior. Por supuesto, la elevación al cuadrado es sólo una de muchas técnicas de determinación de potencia posibles, y se contemplan otras. El proceso de la sub-banda inferior suma los valores individuales de potencia de los intervalos desde el número de intervalo 3 hasta el número de intervalo 64 ambos inclusive para el valor de potencia de trama de la sub-banda inferior de ruido en el ejemplo aquí tratado, y también hace lo mismo para el valor de potencia de trama de la sub-banda inferior de información. A continuación, se puede normalizar cada resultado multiplicando, si se desea, por un valor de normalización, que puede ser por ejemplo 10, como se muestra en 604, 604' en la Figura 6, y que puede ser diferente para cada sub-banda y para cada señal.

Los valores de potencia de la trama así obtenidos varían de forma natural de una trama a otra, especialmente para los valores de potencia de la trama de ruido, y puede ser deseable reducir esta variabilidad. Por consiguiente, los cuatro flujos resultantes de valores de potencia de la trama se hacen pasar a través de filtros 605 y 605' de suavizado, los cuales rastrean los respectivos valores de potencia de la trama. Es preferible que este suavizado se realice en el dominio logarítmico, de modo que las cuatro corrientes se conviertan en primer lugar en corrientes de valores de potencia logarítmicos como se ha descrito anteriormente. El suavizado se muestra en 505 y 505' en la Figura 5 y, en general, es un proceso de filtrado paso bajo realizado por los filtros 605 y 605' de suavizado paso bajo de la Figura 6. Cualquier filtro paso bajo de este tipo es aplicable, pero en este ejemplo concreto se utiliza un filtro descrito por la siguiente ecuación:

$$P_i = P_{i-1} + \max[-\beta, \alpha \cdot (\{10 + \log S_i\} - P_{i-1})] \quad (1)$$

donde P_i es la potencia rastreada para la trama, señal y sub-banda actuales, P_{i-1} es la potencia rastreada para la trama anterior, α y β son coeficientes del filtro, S_i es la potencia sumada para la trama, señal y sub-banda actuales, e i es el índice de la trama. En el ejemplo aquí tratado, se utilizan los valores $\alpha = 0,83$ (los valores típicos varían desde aproximadamente 0,5 hasta aproximadamente 0,95) y $\beta = 0,134$ (los valores típicos varían desde aproximadamente 0,01 hasta aproximadamente 0,5).

Debido a que los valores de potencia de la trama de ruido tienen mayor variabilidad que los valores de potencia de la trama de información, y además pueden necesitar una respuesta más lenta para impedir que se introduzca un carácter de sonido "de bombeo" en la salida final del sistema de compresión, los dos valores de potencia de la trama de ruido de la sub-banda se suavizan aún más aplicando (507) un filtro 607 de rastreo (de nuevo, estos filtros se encuentran dentro de la clase de los filtros paso bajo y cualquier filtro de este tipo es generalmente aplicable para esta función). En el sistema de dos sub-bandas de ejemplo, este suavizado viene descrito por la ecuación

$$N_i = N_{i-1} + \delta \cdot |P_i - N_{i-1}| \cdot (P_i - N_{i-1}) \quad (2)$$

donde N_i es la potencia de ruido suavizada y rastreada para la trama y la sub-banda actuales, N_{i-1} es la potencia de ruido suavizada y rastreada para la trama anterior y δ es el coeficiente de suavizado del filtro. Por ejemplo, se utiliza el valor $\delta = 0,4$ (los valores típicos varían desde aproximadamente 0,1 hasta aproximadamente 0,8).

En la realización aquí descrita, los filtros 605 y 605' de suavizado pueden tener las mismas propiedades y, si se implementan como código DSP (procesador de señales digitales), se puede usar el mismo código para ambos filtros, usándose el filtrado de rastreo para el componente de ruido conseguido mediante la adición del filtro 607 de rastreo. De forma alternativa, el filtro 605 de suavizado y el filtro 607 de rastreo se pueden combinar en un único filtro, el cual puede tener una constante de tiempo mayor que la del filtro 605'.

A partir de la potencia de ruido suavizada and rastreada, se calcula (509, 609) un valor de codo de compresión para cada sub-banda de acuerdo con la ecuación

$$K_i = N_i + \mu \quad (3)$$

donde K_i es el valor de codo de compresión para la trama actual y μ es un factor de desfase. Se utiliza el valor $\mu = -1,2$.

La importancia del valor de codo de compresión se explica haciendo referencia a las curvas de la Figura 7A, la cual ilustra una curva de ganancia unitaria/compresión cero representada como una línea continua, una combinación de expansión hacia abajo y compresión suave, separadas por un umbral mostrado en aproximadamente 52 dB, representada como una línea discontinua, y una expansión hacia abajo y una compresión fuerte separadas por un umbral y representada como una línea de rayas y puntos. Cada una de las tres curvas de "compresión" de la familia representa la respuesta a un nivel diferente de ruido local, representando la línea continua la respuesta a un nivel de ruido de extremo cercano bajo, de 60 dB, representando la línea discontinua la respuesta a un nivel de ruido de extremo cercano mayor, de 75 dB, y representando la línea de rayas y puntos la respuesta a un nivel de ruido de extremo cercano grande, de 90 dB. Los propios niveles de ruido asociados se muestran con líneas horizontales, correspondiendo la línea de ruido continua a la curva de "compresión" continua, y así sucesivamente. La parte de las curvas de compresión por encima de las respectivas líneas de nivel de ruido correspondientes o, con mayor precisión, la parte por encima de una línea aproximadamente 5 dB por debajo de las respectivas líneas de nivel de ruido correspondientes, representa la parte que puede ser oída eficazmente, con una inteligibilidad razonable, por el oyente de extremo cercano. Es decir, se sabe que los humanos pueden comprender habla por encima de aproximadamente -5 dB de relación señal-ruido (SNR).

En el gráfico de la Figura 7A, se puede ver que para las tres curvas de "compresión", aunque la cantidad de ruido correspondiente es diferente de unas a otras, y la cantidad de compresión es diferente de unas a otras, la parte de cada curva de "compresión" que está dentro del rango de inteligibilidad basado en el ruido de extremo cercano respectivo es similar para las tres curvas. En otras palabras, mediante el uso de un proceso de control que ajuste los parámetros de compresión de forma automática en respuesta a ruido ambiental de extremo cercano, se puede

ajustar el rango dinámico de la señal de excitación de reproducción, comprimiendo la señal en un rango dinámico más pequeño a medida que aumenta el ruido local, de tal manera que la señal, que se origina con un rango dinámico en silencio típico de una persona, se comprimirá a un rango dinámico que representa la zona de audibilidad por encima del nivel de enmascaramiento de ruido de extremo cercano. Los niveles absolutos en los gráficos son ejemplos, pero constituyen una representación razonable de niveles que podrían ser aplicables a un sistema de comunicación. En los gráficos de ejemplo, los ejes están escalados a un SPL en dB acústico equivalente. De esta manera, el eje de la señal de entrada representa el nivel acústico equivalente de la señal de extremo lejano que habría estado en el oído del usuario si éste pudiera estar escuchando directamente al hablante de extremo lejano en el ejemplo de sistema de comunicación bidireccional, el eje de la señal de salida representa el nivel acústico real del sonido proporcionado por el hablante, y el eje de la señal de ruido representa el nivel acústico real del ruido local.

Este valor de codo de compresión se utiliza, junto con la potencia de la señal suavizada, para crear la ganancia de compresión (o atenuación), en 511 y 611, que se aplicará posteriormente a cada sub-banda de la señal de información de entrada para crear la señal de salida que se utilice para excitar el altavoz 206 (Figura 2). La ganancia de compresión se define como

$$G_i = \text{if} \left\{ K_i < \eta, 0, \text{else if} \left[P_i \geq \eta, \left(\frac{K_i - \eta}{\mu - \eta} \right) \cdot (\mu - P_i), \text{else} \left(\frac{K_i - \eta}{\eta} \right) \cdot P_i \right] \right\} \quad (4)$$

donde G_i es la ganancia de compresión, expresada en forma logarítmica, para la trama y sub-banda actuales, y η y μ son parámetros de ganancia de compresión. En este ejemplo, se utiliza $\eta = 0,4$ (los valores típicos están entre aproximadamente 0,1 y aproximadamente 2) y $\mu = 10$ (los valores típicos oscilan desde aproximadamente 1 hasta aproximadamente 100). El resultado de esta ecuación es la curva de nivel de ganancia mostrada en la Figura 7B, donde, en la citada Figura 7B, con fines de representación gráfica, la ganancia de la ecuación (4) se pasa a dB multiplicando por 10.

Para un instante de tiempo dado, y para una sub-banda concreta, la Ecuación 4 define la curva de compresión, donde el valor de codo de compresión es la única variable. El funcionamiento de esta ecuación es tal que los parámetros definen un nivel de señal de entrada mínimo que representa al SPL audible equivalente mínimo en silencio, un nivel de señal máximo que representa al SPL audible equivalente máximo, y un nivel umbral de codo de compresión que representa al menor SPL acústico que está pensado para retener audibilidad cuando el oyente está en un entorno con ruido. Para la realización de ejemplo ilustrada por las curvas mostradas en la Figura 7A, el nivel de presión sonora de entrada equivalente mínimo se establece mediante los valores de los parámetros proporcionados anteriormente para que sea aproximadamente 20 dB. En este punto, el cual está a la izquierda de la parte representada de la gráfica, un nivel de señal de entrada equivalente de 20 dB de SPL producirá un nivel de señal de salida equivalente de 20 dB de SPL, el cual gráficamente es el punto en el que las tres curvas de compresión se cortan más allá del lado izquierdo de la gráfica. De nuevo, para la realización de ejemplo ilustrada por las curvas mostradas en la Figura 7A, el nivel de presión sonora de entrada equivalente máximo se establece mediante los valores de los parámetros proporcionados anteriormente para que sea aproximadamente 110 dB. En este punto máximo, el cual está a la derecha de la parte representada de la gráfica, un nivel de señal de entrada equivalente de 110 dB de SPL producirá un nivel de señal de salida equivalente de 110 dB de SPL. Esto se ve gráficamente como el punto en el que las tres curvas de compresión se cortan más allá del lado derecho de la gráfica. En tercer lugar, de nuevo para la realización de ejemplo ilustrada por las curvas mostradas en la Figura 7A, el nivel umbral de codo de compresión (el nivel de la señal de entrada por encima de la cual se produce compresión) se establece mediante los valores de los parámetros proporcionados anteriormente para que sea aproximadamente 52 dB. En este nivel de señal de entrada, el valor de codo (ganancia) de compresión se aplica a la señal. De esta manera, la altura de la curva de compresión en ese punto de nivel de señal de entrada es igual al nivel de señal de entrada más el valor de codo de compresión, cuando ambos se expresan en términos logarítmicos (dB).

Una forma de visualizar el control de compresión de valor único es reconocer que los puntos mínimo y máximo son substancialmente fijos y que el nivel de salida es igual al nivel de entrada en esos dos puntos fijos. Para ayudar a la visualización, si se estirase una banda de goma entre esos dos puntos, y se aplicara el valor de codo de compresión elevando la banda de goma en una cantidad igual al valor de codo de compresión en el punto del nivel umbral de codo de compresión, entonces la banda de goma adoptaría la forma de la curva de compresión, es decir, una forma como la de las curvas mostradas en la Figura 7A. De esta manera, mediante la aplicación de la Ecuación 4, un único valor de codo de compresión determina toda la curva de compresión.

Una característica de este método de control de compresión es la expansión hacia abajo que se crea a la izquierda del umbral de codo de compresión, como es evidenciado por la pendiente mayor que uno de las curvas de compresión en esta zona (Figura 7A). Seleccionando este umbral para que esté justo por debajo del menor nivel

esperado de la señal de entrada de interés, se impide la sobre amplificación de componentes de ruido de bajo nivel incluidos en la señal de entrada y, en lugar de esto, se suprimen estos componentes no deseados.

Después de que la ganancia de compresión G_i haya sido definida por la Ecuación 4, la ganancia se puede convertir a forma lineal como se explica a continuación:

$$Gf_i = 10^{\frac{G_i}{2}} \quad (5)$$

donde Gf_i es el factor de ganancia lineal para la trama actual y sub-banda de señal. De esta manera, como ilustra la Ecuación (5) para esta realización de ejemplo, las ganancias de compresión se vuelven a convertir del dominio logarítmico al dominio lineal dividiendo por "2" el valor proporcionado por la ecuación (4) y elevando 10 a esa potencia para crear factores de ganancia lineal. Esto se muestra en 513 en la Figura 5, como realizado por una parte de convertidor antilogarítmico y expansor 613 en la Figura 6. Esto produce un factor de ganancia lineal por trama para cada sub-banda.

En este punto, existen dos factores de ganancia lineal de la sub-banda correspondientes a cada trama de datos. Estos factores se vuelven a expandir a 257 factores de ganancia del intervalo independientes asignando el factor de ganancia lineal apropiado a cada número de intervalo asociado con todos los intervalos de cada sub-banda. (513 de la Figura 5, 613 de la Figura 6). Se asigna un valor de cero a esos intervalos que no están incluidos en el interior de ninguna sub-banda. Dado que sub-bandas contiguas normalmente tendrán diferentes factores de ganancia lineal, si no se aplica ningún suavizado en frecuencia, habrá un escalón en la ganancia a través de las fronteras de sub-banda contiguas, y esto podría provocar una distorsión de señal audible. Se puede aplicar un suavizado a través de los factores de ganancia del intervalo de frecuencia (515 de la Figura 5) después de que las ganancias de compresión se hayan vuelto a convertir al dominio lineal y se hayan expandido en factores (513) de ganancia del intervalo independientes. Opcionalmente, se puede aplicar el suavizado, como se muestra con mayor detalle como un convolucionador / interpolador de splines 615 en la Figura 6. Utilizando un conjunto de valores de ajuste, conocido como kernel, el cual se puede leer, por ejemplo, de la tabla 614, o se puede calcular a partir de una ecuación, u otro método similar, el convolucionador / interpolador de splines 615 suaviza los valores de ganancia de intervalo a través de la transición entre sub-bandas contiguas realizando una convolución del kernel de suavizado con los factores de ganancia de intervalo. De forma alternativa, un kernel corto, conocido como una spline, puede ser multiplicado por los factores de ganancia de intervalo para suavizar los factores a través de las fronteras. Otros métodos para suavizar a través de la frecuencia son bien conocidos en la técnica, y cualquier método de este tipo se contempla. El proceso de suavizado produce un conjunto de ganancias de compresión de intervalo, una asociada con cada intervalo de frecuencias. Estas ganancias de compresión resultantes se aplican a la señal de información multiplicando la señal en cada intervalo por la respectiva ganancia de compresión de intervalo, en 517, 617. La señal de salida procesada resultante para cada sub-banda se muestra en 519 en la Figura 5 y en 619 en la Figura 6.

Después de aplicar las ganancias de compresión de intervalo, las señales 519, 619 multibanda de frecuencia atenuadas resultantes se combinan en el proceso de solape-y-suma de STFT inversa bien conocido. Esto se muestra como el bloque 310 de recomposición de señal en la Figura 3. El resultado de esta recomposición es una señal de salida de banda completa comprimida, en 312, que es la señal de salida que se utiliza para excitar el altavoz 206 (Figura 2). De forma alternativa, en una implementación en el dominio del tiempo de ejemplo, se multiplica cada señal de sub-banda limitada por banda por su factor de ganancia lineal de sub-banda respectivo en 517, 617 (Figuras 5 y 6), y las múltiples señales de salida procesadas para cada sub-banda se suman unas a otras en la etapa o circuito 310 de recombinación para proporcionar una señal de salida de banda completa comprimida, en 312, que es la señal que es la señal de salida utilizada para excitar el altavoz 206 (Figura 2).

El gráfico de curvas de nivel de ganancia mostrado en la Figura 7B muestra la curva de nivel de ganancias para una sub-banda. En la gráfica de curvas de nivel de ganancias de la Figura 7B, el eje vertical es la ganancia aplicada a la sub-banda asociada de la señal de información de reproducción utilizada para excitar el altavoz 206. Se puede ver que es función del nivel del ruido ambiental local (de extremo cercano) de entrada procedente del indicador 208 y del nivel instantáneo de la señal de información de entrada procedente de la fuente 204 de información. La Figura 7C es una curva de nivel del ruido ambiental (de extremo cercano) local de entrada procedente del indicador 208 y del nivel instantáneo de la señal de información de entrada procedente de la fuente 204 de información.

Se debería resaltar que en el procesamiento de las señales de ruido y de información, se usan las potencias de las señales en lugar de sus magnitudes. Utilizando las potencias de las señales se minimizan las necesidades de potencia de cálculo para conseguir la señal de salida procesada. Sin embargo, de forma alternativa, también se pueden usar las magnitudes, y esta operación también se contempla. Asimismo, el suavizado se realiza con filtros en el dominio logarítmico en lugar de en el dominio lineal, proporcionando de esa forma simetría logarítmica alrededor de una media logarítmica, lo cual es sensiblemente diferente al filtrado en el dominio lineal. Dado que se sabe que la percepción de sonoridad del oído humano sigue una característica logarítmica, este proceso

proporciona resultados de reproducción de sonido muy naturales, a diferencia del procesamiento en el dominio lineal. No obstante, estos rasgos no pretenden ser limitativos, y también se contemplan otras alternativas tanto lineales como no lineales.

5 Como se ha explicado en este documento, una ventaja conseguida por el sistema anterior es el control de la compresión a través del control del valor de codo de compresión (el valor de la ganancia aplicada en el umbral de compresión). En otras palabras, el nivel de señal de salida se puede mantener en el nivel de la señal de entrada (ganancia = 1) en un SPL máximo (por ejemplo 110 dB de SPL, rango de 90 a 120) y en un SPL mínimo (por ejemplo, 20 dB de SPL, rango de 10 a 40). Entonces se puede calcular una ganancia instantánea a aplicar en el
10 umbral de codo de compresión (por ejemplo, 52 dB de SPL, rango de 40 a 75). El umbral de codo de compresión es el nivel de señal de entrada en el cual se produce la transición de expansión hacia abajo a compresión a medida que la señal de entrada aumenta a través del umbral de codo de compresión. Se determina la ganancia (el valor de codo de compresión) a aplicar en el umbral de compresión en función del ruido de extremo cercano, y se deriva la curva de compresión/expansión basada en ese valor de codo de compresión. Todos los demás puntos de la curva de compresión se determinan simplemente por interpolación lineal entre esos tres puntos (véase la Figura 7A). De esta manera, la ganancia aplicada para cualquier trama y para cualquier banda es función de la amplitud de la señal de entrada y de la curva de compresión/expansión, y la propia curva de compresión/expansión es una función del nivel de ruido de extremo cercano. De esta forma, la ganancia aplicada se controla mediante la amplitud de la señal de extremo lejano y la amplitud del ruido de extremo cercano. Esta técnica significa que sólo es necesario calcular un
20 número, y ese número puede definir toda la curva de compresión/expansión (por trama, por sub-banda).

De esta forma, se observará que las mediciones de potencia por sí mismas no son primordiales. De forma más importante, lo que se consigue son cálculos simplificados que no requieren "perfiles" o almacenamiento, y que proporcionan cantidades de compresión/expansión que responden tanto al nivel de la señal de extremo lejano como al nivel de la señal de extremo cercano.

Aunque se han mostrado y descrito realizaciones y aplicaciones, resultaría evidente para las personas con experiencia en la técnica que se benefician de esta descripción que son posibles muchas más modificaciones que las anteriormente mencionadas sin alejarse de los conceptos de la invención definidos por las reivindicaciones adjuntas.

Por ejemplo, el rastreo de la potencia de una señal, ya sea ruido o información, es esencialmente una forma de detección de envolvente, de modo que cualquier tipo de detección de envolvente (o de seguimiento/seguir de envolvente) puede estar dentro del alcance de la invención. Por lo tanto, la invención no debe estar restringida salvo por las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un sistema de manipulación del rango dinámico que comprende:

una primera entrada configurada para proporcionar una señal de audio;
una segunda entrada configurada para proporcionar una señal de indicación de ruido para ruido ambiental en una posición de reproducción; y
un procesador configurado para:

recibir la señal de audio y la señal de indicación de ruido,
tomar muestras en el tiempo de cada una de la señal de audio y de la señal de indicación de ruido,
dividir las muestras en el tiempo de cada una de la señal de audio y de la señal de indicación de ruido en tramas, y transformar cada una de la señal de audio y de la señal de indicación de ruido en una pluralidad de sub-bandas de frecuencia,
para cada trama y para cada sub-banda de frecuencia, determinar una potencia de ruido suavizada a partir de la señal de indicación de ruido, y determinar además una potencia de ruido suavizada y rastreada a partir de la potencia de ruido suavizada aplicando un filtro paso bajo a la potencia de ruido suavizada,
para cada trama y para cada sub-banda de frecuencia, determinar una potencia de ruido suavizada a partir de la señal de audio,
para cada trama y para cada sub-banda de frecuencia, calcular un valor de codo de compresión a partir de la potencia de ruido suavizada y rastreada de acuerdo con la ecuación

$$K_i = N_i + \text{factor de desfase},$$

donde K_i es el valor de codo de compresión para la trama y la sub-banda de frecuencia actuales y N_i es la potencia de ruido suavizada y rastreada para la trama y la sub-banda de frecuencia actuales,
para cada trama y para cada sub-banda de frecuencia, determinar una ganancia de compresión de acuerdo con la potencia de audio suavizada y curva de ganancia definida por

$$G_i = \text{if} \left\{ K_i < \eta, 0, \text{else if} \left[P_i \geq \eta, \left(\frac{K_i - \eta}{\mu - \eta} \right) \cdot (\mu - P_i), \text{else} \left(\frac{K_i - \eta}{\eta} \right) \cdot P_i \right] \right\},$$

en la cual G_i es la ganancia de compresión en forma logarítmica para la trama y la sub-banda de frecuencia actuales, P_i es la potencia de audio suavizada para la trama y la sub-banda de frecuencia actuales, y η y μ son parámetros de ganancia de compresión, y
para cada trama y para cada sub-banda de frecuencia, aplicar la ganancia de compresión a la señal de audio para generar una señal de salida controlada en ganancia.

2. Sistema de acuerdo con la reivindicación 1, en el cual la determinación de potencia se realiza sobre dos sub-bandas para cada una de la señal de audio y de la señal de indicación de ruido.

3. Sistema de acuerdo con la reivindicación 2, en el cual las dos sub-bandas van respectivamente desde aproximadamente 50 Hz hasta aproximadamente 1 kHz, y desde aproximadamente 1 kHz hasta aproximadamente 3,5 kHz.

4. Sistema de acuerdo con la reivindicación 1, en el cual la determinación de potencia se realiza en el dominio logarítmico para al menos una de la señal de audio y de la señal de indicación de ruido.

5. Sistema de acuerdo con la reivindicación 1, en el cual la señal de salida controlada en ganancia utiliza un valor de ganancia que es sometido a una función de suavizado.

6. Sistema de acuerdo con la reivindicación 1, en el cual la señal de salida controlada en ganancia es presentada por un altavoz de acuerdo con una señal de excitación de reproducción, y la señal de indicación de ruido es proporcionada por una fuente de señal de indicación de ruido basada en ruido ambiental en la posición del altavoz.

7. Sistema de acuerdo con la reivindicación 1, en el cual la señal de audio es información de audio pre-grabada.

8. Sistema de acuerdo con la reivindicación 1, en el cual la señal de audio es información de audio generada en tiempo real.

9. Un método para compensación de ruido ambiental en una posición de reproducción, que comprende:

recibir una señal de ruido ambiental indicativa del ruido ambiental,

recibir una señal de audio,
 tomar muestras en el tiempo de cada una de la señal de audio y de la señal de ruido ambiental, dividir las
 muestras en el tiempo de cada una de la señal de audio y de la señal de ruido ambiental en tramas, y
 transformar cada una de la señal de audio y de la señal de ruido ambiental en una pluralidad de sub-bandas
 de frecuencia,
 para cada trama y para cada sub-banda de frecuencia, determinar una potencia de ruido suavizada a partir de
 la señal de ruido ambiental y determinar, además, una potencia de ruido suavizada y rastreada a partir de la
 potencia de ruido suavizada aplicando un filtro paso bajo a la potencia de ruido suavizada,
 para cada trama y para cada sub-banda de frecuencia, determinar una potencia de ruido suavizada a partir de
 la señal de audio,
 para cada trama y para cada sub-banda de frecuencia, calcular un valor de codo de compresión a partir de la
 potencia de ruido suavizada y rastreada de acuerdo con la ecuación

$$K_i = N_i + \text{factor de desfase},$$

donde K_i es el valor de codo de compresión para la trama y la sub-banda de frecuencia actuales y N_i es la
 potencia de ruido suavizada y rastreada para la trama y la sub-banda de frecuencia actuales,
 para cada trama y para cada sub-banda de frecuencia, determinar una ganancia de compresión de acuerdo
 con la potencia de audio suavizada y con una curva de ganancia definida por

$$G_i = \text{if} \left\{ K_i < \eta, 0, \text{else if} \left[P_i \geq \eta, \left(\frac{K_i - \eta}{\mu - \eta} \right) \cdot (\mu - P_i), \text{else} \left(\frac{K_i - \eta}{\eta} \right) \cdot P_i \right] \right\},$$

donde G_i es la ganancia de compresión en forma logarítmica para la trama y la sub-banda de frecuencia
 actuales, P_i es la potencia de audio suavizada para la trama y la sub-banda de frecuencia actuales, y η y μ
 son parámetros de ganancia de compresión, y
 para cada trama y para cada sub-banda de frecuencia, aplicar la ganancia de compresión a la señal de audio
 para generar una señal de salida controlada en ganancia.

10. Método de acuerdo con la reivindicación 9, en el cual la determinación de potencia se realiza sobre dos sub-
 bandas para cada una de la señal de audio y de la señal de ruido ambiental.

11. Método de acuerdo con la reivindicación 10, en el cual las dos sub-bandas van respectivamente desde
 aproximadamente 50 Hz hasta aproximadamente 1 kHz, y desde aproximadamente 1 kHz hasta aproximadamente
 3,5 kHz.

12. Método de acuerdo con la reivindicación 9, en el cual la determinación de potencia se realiza en el dominio
 logarítmico para al menos una de la señal de audio y la señal de ruido ambiental.

13. Método de acuerdo con la reivindicación 9, que comprende además el suavizado de un valor de control de
 ganancia aplicado a la señal de salida.

14. Método de acuerdo con la reivindicación 9, que comprende además la generación de la señal de audio a partir
 de una fuente pre-grabada.

15. Método de acuerdo con la reivindicación 9, que comprende además la generación de la señal de audio en tiempo
 real.

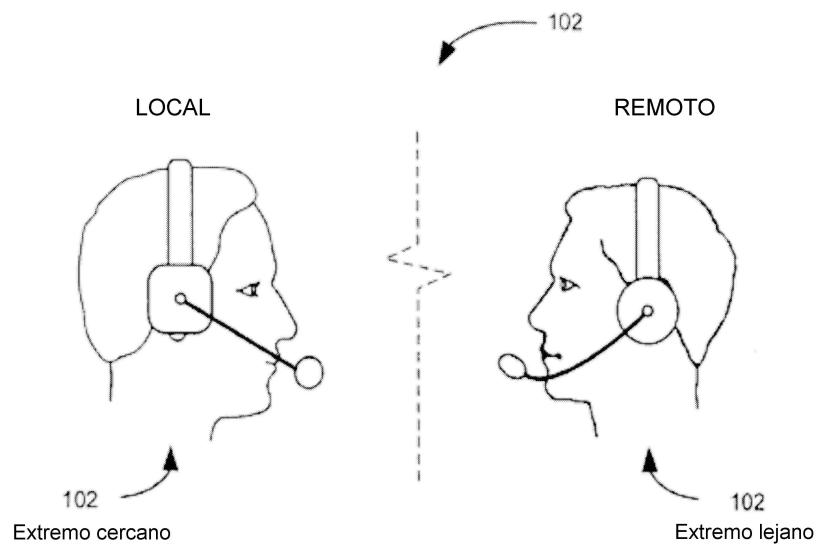


FIG. 1A

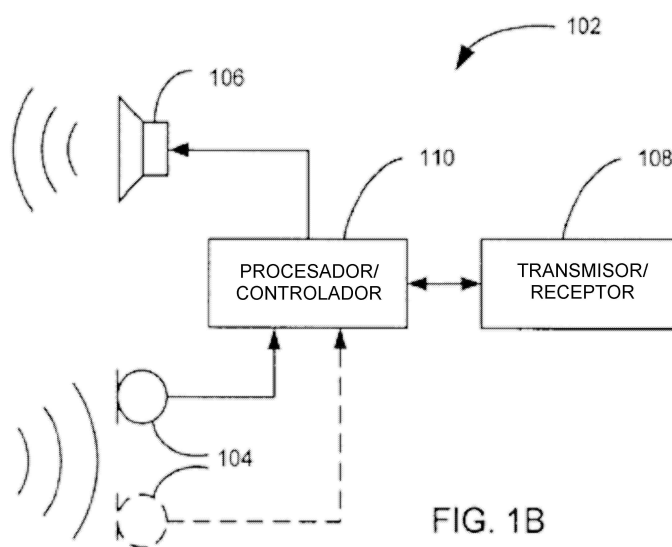


FIG. 1B



FIG. 2

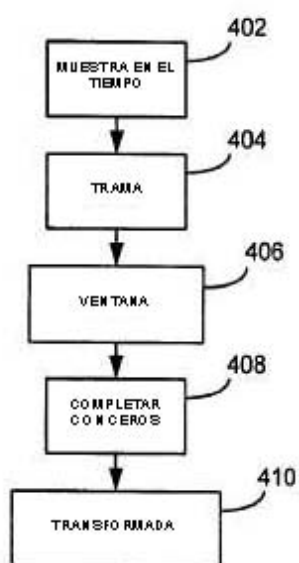


FIG. 4

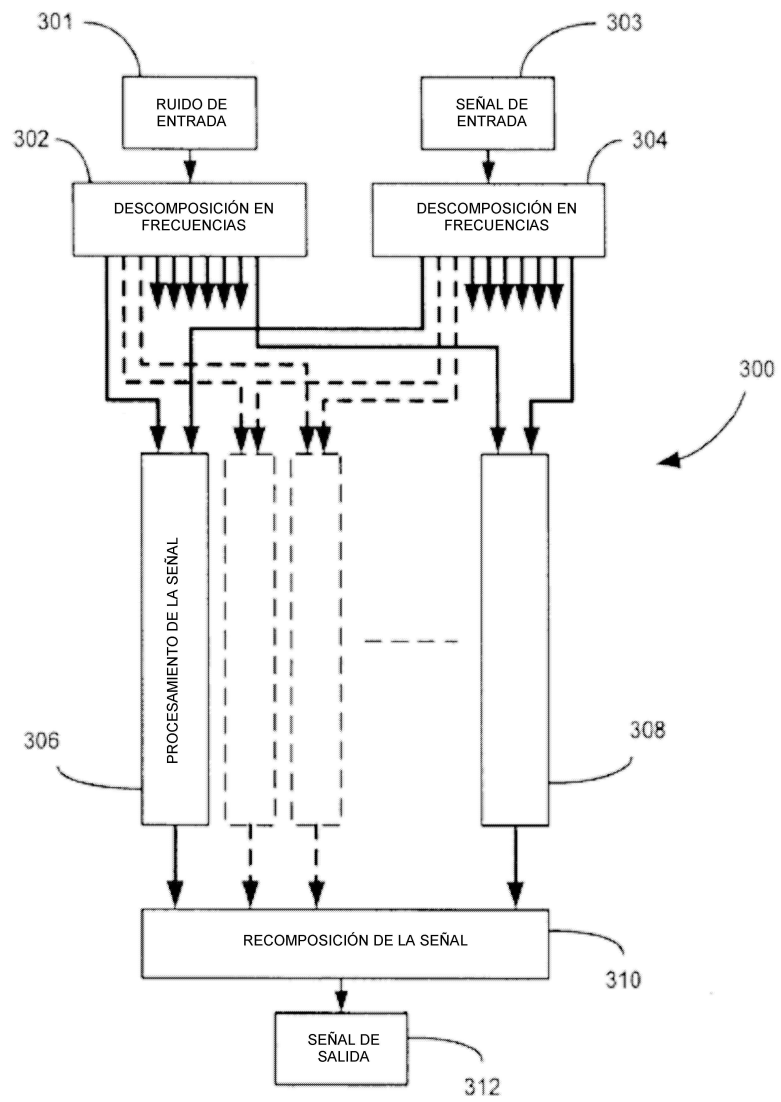


FIG. 3

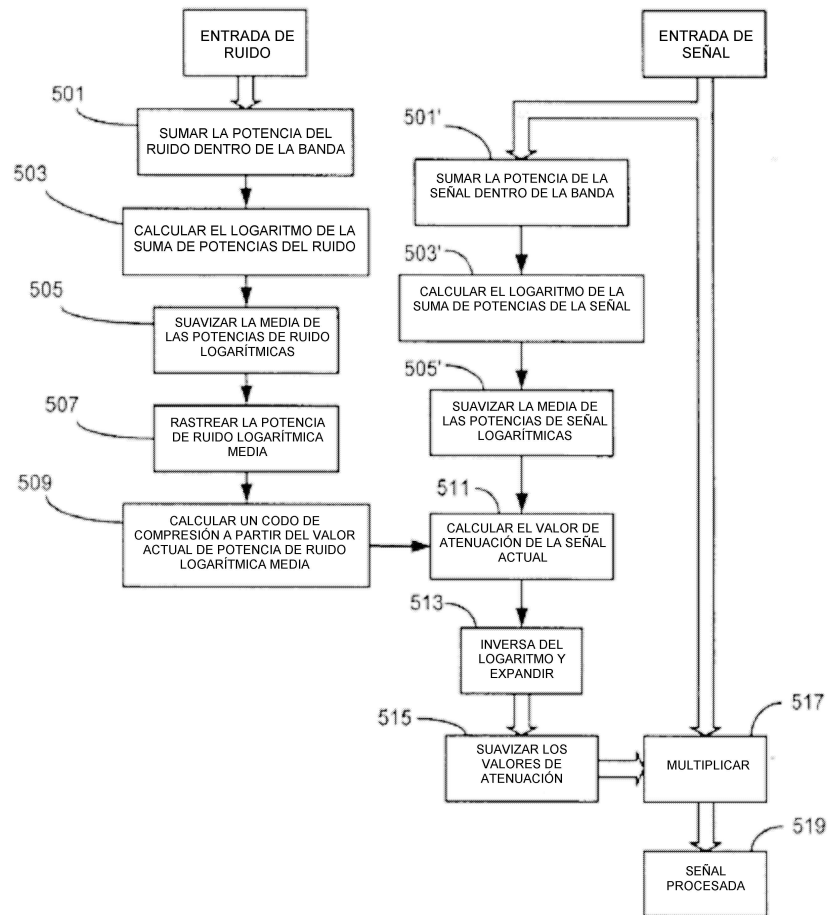


FIG. 5

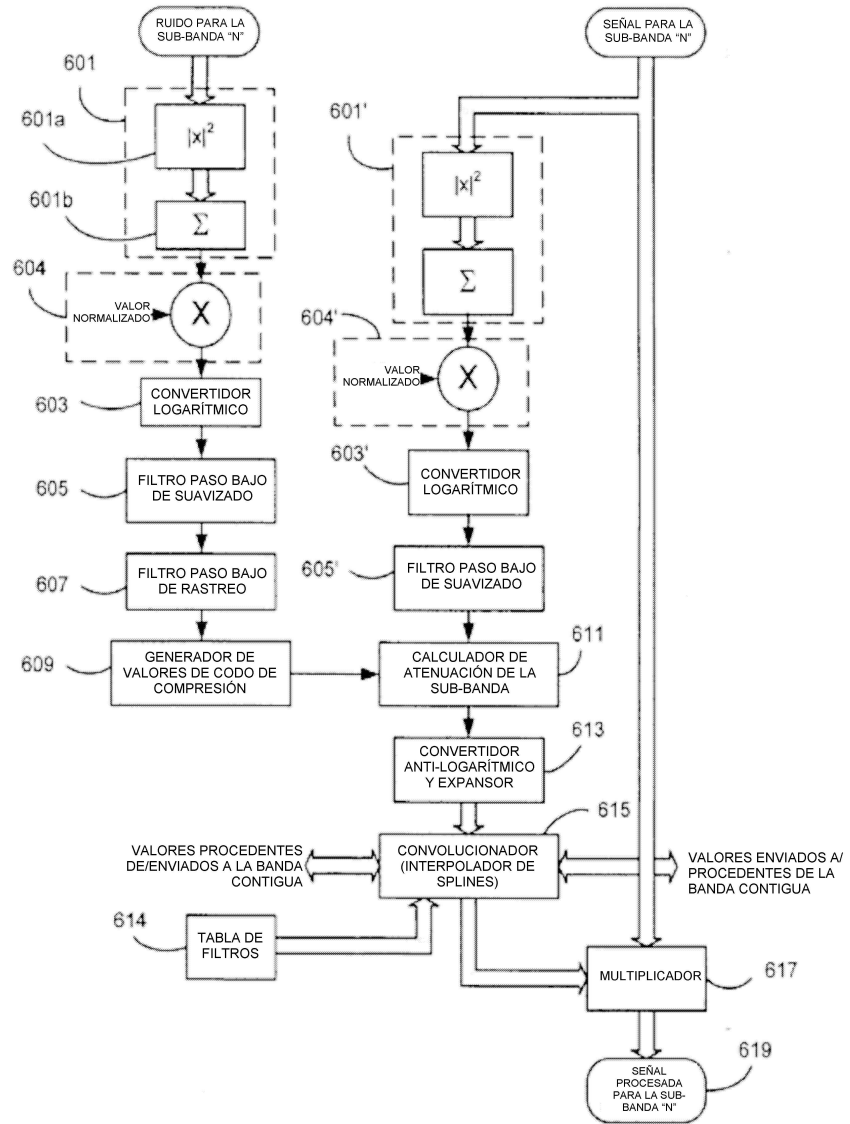


FIG. 6

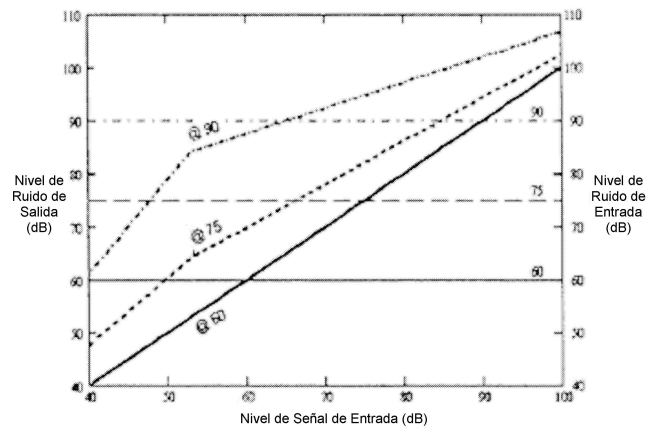


FIG. 7A

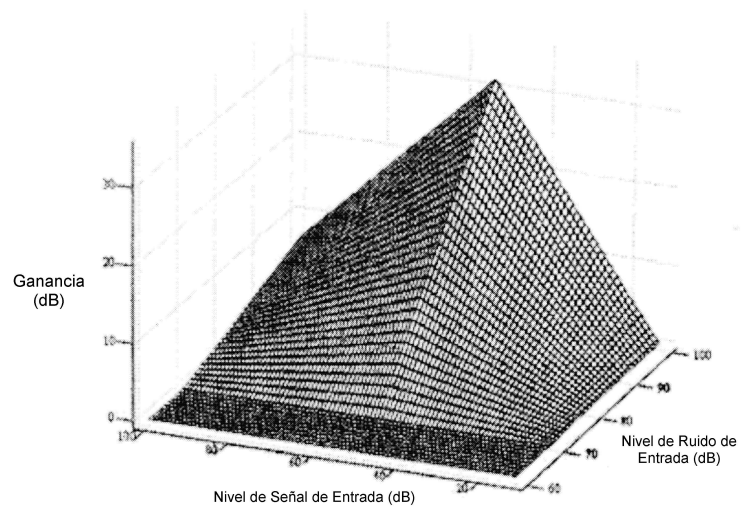


FIG. 7B

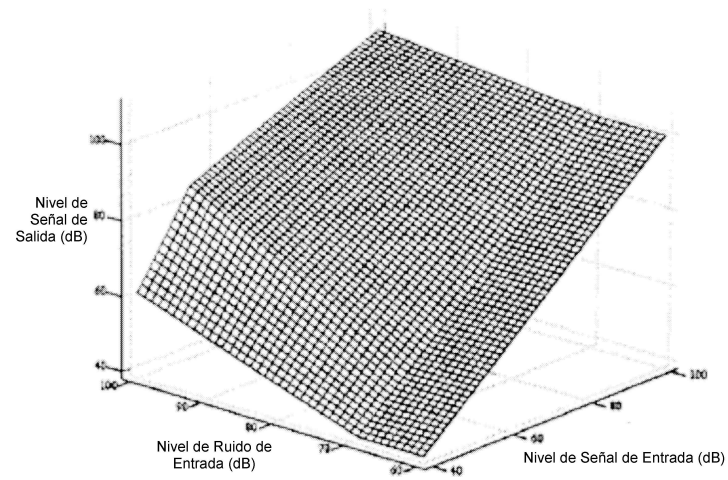


FIG. 7C