

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 430 250**

51 Int. Cl.:

H04R 5/027 (2006.01)

H04R 5/033 (2006.01)

H04M 1/60 (2006.01)

H04M 3/56 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.10.2007 E 07817647 (6)**

97 Fecha y número de publicación de la concesión europea: **13.03.2013 EP 2084937**

54 Título: **Procedimiento y dispositivo para la grabación, transmisión y reproducción de eventos de sonido para aplicaciones de comunicación**

30 Prioridad:

12.10.2006 DE 102006048295

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

19.11.2013

73 Titular/es:

**PAVEL, ANDREAS MAX (100.0%)
VIA BESSO, 7
6900 LUGANO, CH**

72 Inventor/es:

PAVEL, ANDREAS MAX

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 430 250 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Procedimiento y dispositivo para la grabación, transmisión y reproducción de eventos de sonido para aplicaciones de comunicación.

5

Campo de la invención

La presente invención se refiere a un procedimiento según el preámbulo de la reivindicación 1, así como también a un dispositivo según el preámbulo de la reivindicación 7.

10

En el campo de la transmisión remota estereofónica de eventos de sonido, las rutas de transmisión completas dúplex en tiempo real, estereofónicamente designadas de alta calidad, son comunes en la zona de tecnología de radio y estudio las cuales, sin embargo, están vinculadas a puntos de transferencia de red estacionaria. Además, son conocidas las conexiones punto a punto, inalámbricas de corto rango, estereofónicas de alta calidad similares, las cuales son principalmente utilizadas para entrevistas de programas en el exterior.

15

En el campo de las conferencias telefónicas, por otro lado, se conocen muchas propuestas para la grabación, transmisión y reproducción estereofónica de señales telefónicas, ya sea para la mejor identificación del(los) participante(s) en una conversación individual, o para mejorar la inteligibilidad de voz o, en cualquier caso, imitar una mezcla de panorama y colocar en una sola oreja fuentes individuales recibidas (altavoces) en un cierto lugar con el panorama estéreo.

20

Ninguna de las técnicas actuales para la transmisión estereofónica en el campo de estudio de grabaciones y radio, ni las propuestas existentes para una configuración estereofónica de conferencias telefónicas, se refieren a la zona central de la presente invención, es decir, preferentemente, la transmisión móvil de imágenes acústicas personales en tiempo real, que se refieren al respectivo hablante, el campo mismo, que es también nuevo en esta formulación conceptual, que dirige una nueva tarea y función. Para una mejor comprensión, parece justificado puntualizar ya en este punto que el término imagen acústica personal se refiere al entorno acústico, tal como el hablante individual lo nota en sí mismo en el momento de llamar binauralmente, por ejemplo, con los dos oídos y con la inclusión del comportamiento de reflexión, inflexión y resonancia en relación con la cabeza y respectivamente causada por él.

25

30

De este modo, como estado de la técnica comparable más próximo, es posible hacer referencia a las siguientes publicaciones a modo de ejemplo: JP 06268722 A, DE 33 03 418 A1, GB 2 416 955 A, DE 39 30 278 C1, DE 100 20 857 A1, WO 99/53673 A, WO 98/42161 A, DE 37 37 873 A1, JP 2004 023535 A, US 4 088 849 A, EP 0 724 352 A2, DE 40 41 319 A1, EP 0 358 028 A2, JP 02217100 A, DE 100 20 857 A1.

35

Un ejemplo de una grabación, transmisión y reproducción estereofónica de alta calidad de eventos de sonido para aplicaciones de comunicación en telefonía se muestra en el documento JP 06268722, en el que, sin embargo, teléfonos que funcionan estereofónicamente están conectados a un dispositivo de transmisión central a través de un único conducto de conexión. Cada teléfono del participante dispone de un par de auriculares y de dos micrófonos, es decir un micrófono derecho e izquierdo. La transmisión de las señales estereofónicamente recibidas a cada participante se lleva a cabo mediante un conducto de conexión telefónica en el procedimiento de multiplexión-tiempo, que hace necesaria una división y preparación posteriores de una señal transmitida, que consiste normalmente en paquetes de impulsos no completamente sincronizados en el tiempo, en una señal de auriculares derechos e izquierdos.

40

45

Esta disposición básicamente compleja sirve para realizar una conversación social especialmente popular en Japón, es decir, el denominado fenómeno del karaoke. En este contexto, es importante que cada uno de los participantes individuales reciba la pieza musical en el auricular a través de un circuito adicional, que está también conectado a un dispositivo de transmisión central (red) y que es apto para entretenimientos musicales de alta calidad, como el karaoke. Posteriormente, el participante canta. Para este fin, y también para la grabación y transmisión de la voz que está cantando están previstos ambos micrófonos, estando, de este modo, los micrófonos colocados juntos y directamente frente a la boca del hablante/cantante y lejos de los auriculares. La figura 3 en el documento JP 06 268722 A muestra claramente lo que se quiere decir, de manera que la voz que canta también pueda ser grabada de manera precisa. Los auriculares están obviamente conectados entre sí mediante una banda para los auriculares/banda de apriete elástica y se colocan en las orejas. Los micrófonos se encuentran alejados de los auriculares y están situados directamente frente a la boca del cantante, tal como muestra claramente la figura 3, si la cabeza del cantante se tiene en cuenta en esta disposición. Los micrófonos no están situados en la zona auditiva, sino lejos de ella, lo cual se es comprensible, puesto que en el documento se indica que los dos micrófonos pequeños están previstos para grabar de manera estereofónica el canto del hablante, sea lo que sea lo que signifique esto; una grabación estereofónica de una voz que canta con dos micrófonos situados directamente a la derecha y a la izquierda, y frente a la boca de la persona que está cantando no tiene demasiado sentido, puesto que ambos micrófonos graban prácticamente los mismo. Dicha disposición no permite la transmisión de una imagen acústica personal, sino que está dirigida a alcanzar otro objetivo.

50

55

60

65

El documento DE 33 03 418 A1 se refiere a un dispositivo auxiliar para tartamudos. Una telefonía estéreo o un

equipo correspondiente para este uso no se desprenden de este documento. En el caso de personas tartamudas, evidentemente es habitual proporcionarles un denominado ruido de enmascaramiento en forma de un ruido blanco fijado a 300 Hz en este caso mediante unos auriculares ajustados con sonido muy alto, con el fin de mejorar el habla, de manera que sea posible utilizar auriculares abiertos o cerrados. Sin embargo, en el caso de los auriculares cerrados preferentemente utilizados, la voz de la persona tartamuda solo le llega al oído desde fuera de una manera muy amortiguada, lo cual significa que se escucha a sí mismo muy mal debido al efecto pantalla de los auriculares cerrados con respecto a ruidos ambientales exteriores, lo cual, por otro lado, resulta esencial para un tratamiento curativo, este documento propone fijar uno o varios micrófonos en la zona de los auriculares y permitir de este modo una transmisión de sonidos monofónica o estereofónica desde estos micrófonos directamente a los auriculares, en los que están fijados. Este dispositivo se corresponde esencialmente a la disposición de audífonos. No se divulgan los dos canales de transmisión que son necesarios para la transmisión de una imagen acústica personal. Un adaptador telefónico está previsto; sin embargo, un tráfico telefónico correspondiente funciona, no obstante, sin una configuración estereofónica, puesto que el adaptador (véase, el dibujo) solo recibe una señal monoaural directamente de un micrófono para la boca a través de un preamplificador y por su parte, suministra la monoseñal a los respectivos amplificadores de ambos auriculares. El conducto simplemente se bifurca detrás del adaptador, de manera que ambos auriculares puedan ser cargados. Esto no está relacionado con la presente invención, puesto que el problema relevante para las personas tartamudas se resuelve con la ayuda de un dispositivo para oír.

El documento GB 3 416 955 A correspondiente no se refiere a telefonía estereofónica, sino que propone, en este caso en comparación con el documento DE 37 37 973 A1 al que se hará referencia posteriormente, en el caso de conferencias telefónicas, a la posibilidad de una asignación espacial de las personas que participan en la conversación con la ayuda de un "procesador espacial". De este modo, los componentes lingüísticos entrantes de los participantes en una conferencia son codificados por un detector, de tal manera que incluyan adicionalmente una determinación de la posición o además permitan una asignación espacial mediante el/los oyente(s). En ninguno de los ejemplos de formas de realización mostrados en este documento se graba de manera binaural; solo están previstos monocanales salientes; la impresión estéreo, mejor dicho, impresión de posicionamiento, obtenida por el oyente es generada por el procesador espacial y además pasa a través de los altavoces.

El documento DE 39 30 278 C1 se refiere a una forma de realización estéreo-telefónica con la reproducción exclusiva a través de los altavoces, de manera que una movilidad posiblemente deseada de los abonados telefónicos no es posible debido a esta circunstancia. No se indica dónde están situados los micrófonos, la asignación a los auriculares no es posible de ningún modo, dado que no hay auriculares. Con mayor detalle, el problema descrito es que normalmente son necesarios dos canales de transmisión en el caso estereofónico, pudiendo ser uno de ellos opcionalmente más largo que el otro o pudiendo ser conectado posteriormente, de manera que se debe contar con retrasos temporales y también con diferencias en el tiempo de funcionamiento entre las dos señales que pertenecen a una conexión estéreo. A modo de ayuda, están previstos registros, que almacenan una señal anterior durante siempre que, como determinen los impulsos de sincronización, ambos canales de datos de uso puedan evaluarse al mismo tiempo.

El documento DE 100 20 857 A1 se refiere a la aplicación de una simulación estereofónica para una unidad de telecomunicaciones móvil en conexión con un microrreproductor-grabador (en inglés, "microrecordplayer"), por ejemplo, un teléfono móvil con reproductor de MP3. En este dispositivo, los auriculares o audífonos correspondientes son, por supuesto, están previstos para disfrutar de música estereofónica de alta calidad. Además, está previsto por lo menos un micrófono, preferentemente varios micrófonos, que están dispuestos en un denominado dispositivo de fijación y exposición/de sonido de pabellón auricular/cabeza, es decir, unos denominados cascos con auriculares (en inglés, "headset"). Los cascos con auriculares están separados de la unidad de teléfono móvil y están conectados con la misma de manera inalámbrica. Esta conexión inalámbrica es la que incluye la necesaria conversión digital-analógica así como analógica-digital de dos canales/estereofónica. Estas disposiciones se refieren a los denominados medios de conexión "Bluetooth" inalámbricos entre el dispositivo real y los auriculares y el micrófono en los cascos de auriculares. El documento DE 100 20 857 A1 ve de manera evidente un punto significativo en dichas unidades de teléfonos móviles con reproductores de MP3 en combinación con un teléfono móvil con MP3 con protecciones electromagnéticas para controlar los momentos de estrés/ efecto biológico para intensidades de campo que se consideran demasiado elevadas. Con este fin, se propone disponer arena de sílice natural o cuarzo rosa, que se introduce de nuevo en el interior de los tubos de plástico/cobre alargados, en sistemas de tubos realizados a partir de chapa de cobre o chapa de acero, reduciendo de este modo los efectos del estrés o reacciones de estrés "electrosmog" con respecto a las reacciones corporales. No siempre es posible extraer instrucciones inequívocas para acciones técnicas a partir del documento DE 100 20 857 A1 con la claridad necesaria y deseable. Sin embargo, es esencial que las medidas de transmisión estereofónicas reales en la transmisión de telecomunicaciones no se prevean, sino que únicamente aparezcan en la zona de transmisión entre el dispositivo "Bluetooth" y el conjunto de audición y habla. Esto se pone de manifiesto, por ejemplo, en la referencia de la columna 2, líneas 54-59 según la cual las señales alternas de voz/audio para la reproducción de sonido están mezcladas individualmente y filtradas en dirección binaural en distintas direcciones de recepción seleccionables. Este filtrado de dirección utiliza filtros de paso alto/paso bajo adecuados y filtros similares para ajustar la dirección de recepción según el documento DE 37 37 873 C2 que se detallará a continuación, también en el sentido de la capacidad de selección; mientras que en la grabación binaural real, naturalmente es imposible una selección de dirección de la respectiva fuente acústica y ni siquiera está prevista.

El documento WO 99/53673 muestra un sistema de teleconferencia comparativamente complicado, en el que de nuevo los procesadores de determinación espacial están previstos para la identificación de los interlocutores, así como los denominados concentradores, que detectan los canales activos respectivos y los transmiten mediante conexiones de multicanal. En el documento WO 98/42161 A2, la transmisión telefónica de un evento de sonido tridimensional ocurre a través de micrófonos dispuestos para ser estacionarios en frente del(los) participante(s) a una distancia uno del otro, en conjunto con un ordenador personal, con la distancia correspondiente aproximadamente a la amplitud de una cabeza humana. Preferentemente, los micrófonos están dispuestos dentro de formas de orejas artificiales, ya que el arreglo completo está supuesto para parecerse a una cabeza artificial, o al menos para seguir el principio de la denominada estereofónica de dispositivo de separación (o *Trennkörper-Stereophonie*, un término alemán que se refiere a técnicas de captura de sonido estereofónico que hace uso de dos micrófonos separados por un objeto de tamaño de una cabeza acústicamente opaco). Los altavoces, se proporcionan y se disponen a distancia entre sí o en ambos lados para la reproducción de señales estereofónicas recibidas de esta manera, a partir de los lados opuestos respectivos, de este modo completando la estructura. Una pluralidad de circuitos especiales para filtración, compresión, reducción de datos, y posiblemente, compensación cruzada, también se usa, en particular, para compensar las distorsiones especiales que resultan cuando una señal es primero grabada por una cabeza simulada o disposición de micrófono de *Trennkörper-Stereophonie*, y después viene al oyente respectivo vía altavoces.

El dispositivo como se describe por el documento WO 98/42161 A2, es básicamente considerado neutral al usuario. Por lo tanto, dicho dispositivo no está dispuesto en una persona subjetiva como la invención que va a ser explicada en detalle en los siguientes esfuerzos por hacer, los cuales transmiten la imagen subjetiva del participante y de este modo al oyente personal de conformidad con un ambiente cambiante acústico, como se refiere al participante de conversación respectivo. Por el contrario, en el documento WO 98/42162 A2, el ambiente acústico es siempre transmitido a partir de la misma perspectiva, la cual es capturada por la cabeza simulada rígidamente montada. A tal grado, este dispositivo conocido se comporta de una manera neutral para todas las personas que participan en el evento acústico. Esta situación puede ser deseable para una conferencia telefónica, puesto que permite a cada participante individual, ser localizado en una posición diferente y de este modo, permite la identificación fácil de cada hablante, cuando el ambiente es grabado por la cabeza simulada, con la condición de que las personas no se muevan durante la conversación. El documento WO 98/42161 A2, también menciona la posibilidad de usar auriculares para la reproducción de llamadas de conferencia entrantes, en una forma meramente accesoria, lo cual puede hacer aún más fácil la localización de los participantes individuales de una conferencia entrante. Pero esto podría hacer la comunicación dentro de un grupo de participantes oyentes usando auriculares en la misma ubicación, extremadamente difícil.

A partir del punto de vista de esta concepción básica y el mero propósito de ser completamente cubierta, la invención actualmente reivindicada conduce lejos de la disposición descrita en el documento WO 98/42161 A2, en muchas formas y actualmente se mueve en sentido contrario:

- 1) la perspectiva personal la cual es la base de la presente invención, con su cabeza y movimientos corporales típicos, no permite un posicionamiento fiable y estable de cualquier participante en la conversación con el ambiente que está siendo grabado:
- 2) en el lado de reproducción, la voz del usuario actual de un dispositivo basado en la invención, como lo adquiere por su equipo y reproduce a un participante en conversación a una diferente ubicación, podría ser percibido fuera de un grupo de conversación potencial dispuesto alrededor del mismo, es decir, podría ser percibido cerca o dentro de la cabeza del participante remoto (en la localización de cabeza).
- 3) Ambas condiciones contradicen el propósito del documento WO 98/42161 A2, el cual es permitido para una distribución espacial predecible y estable de participantes de conferencia individual dispuestos alrededor de una mesa, a partir de la perspectiva de una tercera parte remota, la cual no está físicamente presente en la conversación. Además,
- 3) la disposición que funciona con batería del equipo técnico usado en el cuerpo de conformidad con las características de la invención reivindicada, no podría solamente ser innecesario para el objetivo del documento WO 98/42161 A2, sino podría ser contrario al mismo, el cual es capturar la disposición espacial sin cambio de la conversación estacionaria a través de la operación de un sistema de telecomunicación fijo instalado en una sala de conferencia.

Para no tener que llevar consigo constantemente cabezas artificiales ordinarias para realizar una caracterización binaural, especialmente en entornos externos, el documento US 4 088 849 A utiliza la cabeza de la persona de grabación misma, utilizando formas de simulación de oreja artificial que contienen micrófonos en el exterior de los auriculares de monitorización usados por él, mientras los auriculares de izquierda y derecha están conectados entre sí por la venda flexible usual. Las señales registradas son alimentadas en una grabadora de cinta y reproducidas a través de los auriculares inmediatamente después para permitir la monitorización inmediata de las grabaciones del evento de sonido. De este modo, el usuario es su propia "cabeza artificial" con orejas de simulación externas. El

documento no hace alusión a una transmisión remota de señales.

Otra posibilidad para la identificación de participantes en la conferencia telefónica, en donde una transmisión de señal estereofónica no se toma en consideración, se muestra en el documento EP 0 724 352 A2. Un dispositivo de conmutación de telecomunicación digital incluye un diagrama con los datos de identificación de todos los participantes. Quienquiera que hable más alto, se desconecta automáticamente y la identificación correspondiente se cambia en los dispositivos de los otros participantes para indicar a la persona del discurso.

A partir de otro contexto, es decir en un sistema para comunicación de audio y vídeo, usada, por ejemplo, en muestras a larga distancia vía satélite, un interruptor de micrófono que funciona deliberadamente ya se conoce como tal también, véase el documento DE 40 41 319 A1.

Para mejorar la calidad de la información de las aplicación, se conoce (JP 02217100 AA), el hecho de prever un micrófono de apoyo frontal adicional para mezcla de voz, que se conecta, cuando la voz del altavoz exceda un valor de umbral predeterminado.

Para mejorar la calidad de las aplicaciones de comunicación, se hace referencia a la posibilidad de estimular una transmisión binaural (DE 37 37 873 C2), las señales binaurales proporcionadas a un oyente a través de auriculares o audífonos con filtros especiales (por ejemplo, pasos altos, pasos bajos, líneas de retardo, filtros de todos los pasos y tales), que consisten en señales mono que están provistas de una información de distancia y direccional (la cual se conoce como mezclado direcciones binaural). A través de esto, y ajustando los filtros según las llamadas entrantes a partir de los diversos participantes de conversación, las voces pueden ser asignadas a diferentes direcciones de oyentes, los cuales pueden significativamente, mejorar la inteligibilidad de las voces entrantes simultáneamente de varios participantes en la conversación, particularmente en un ambiente ruidoso.

Finalmente, un dispositivo interruptor de comunicación de multiplexión-tiempo digital, puede ser derivado a partir del documento EP 0 358 028 A2, con una memoria de voz que puede ser utilizada como memoria de conferencia y expandida por celdas de memoria adicionales. Dentro de dicha disposición, un bucle de retroalimentación conecta la salida de la memoria de voz con su entrada. Los aspectos estereofónicos no son tomados en consideración.

La presente invención es concebida para resolver el problema de permitir la transmisión, y en particular para la transmisión móvil de imágenes oyentes tridimensionales personales, en tiempo real, a través del medio de telefonía estereofónica, el medio es adaptado a esta tarea o propósito como se requiera.

La invención resuelve este problema por los factores caracterizadores de la reivindicación principal o de la primera reivindicación de dispositivo y con ello, establece un nuevo campo: la transmisión de imágenes auditivas personales en tiempo real.

A través de la recepción de sonidos binaural, esto es, la grabación de sonidos en la zona auditiva de cada participante de conversación, de este modo, se obtienen imágenes acústicas relacionadas con la cabeza natural, que corresponden a la realidad de la vida en la aproximación más grande. Es decir, cada participante, a través de sus auriculares o audífonos respectivos, percibe el ambiente relacionado con la cabeza a través de los auriculares o casos, en el que su compañero en la conversación está actualmente localizado, incluyendo la voz en dicho ambiente y solamente en dicho ambiente del compañero, y de este modo, con todas las reflexiones, difracciones, y resonancias logradas dentro de dicho ambiente o influenciado por este. Este es también un factor principal para proporcionar buena inteligibilidad de voz, puesto que las circunstancias precisas son replicadas en las cuales las regiones del cerebro que procesan la voz de cada persona, se han acostumbrado y se han adaptado desde el comienzo de la evolución del lenguaje, es decir, a percibir el sonido completo de una voz dentro del espectro específico de resonancias, difracciones y reflexiones generadas dentro de un ambiente particular con relación a al propio cuerpo y no el espectro de bajo corte del sonido estrecho de las transmisiones de voz telefónicas practicadas hasta ahora.

A este respecto, otro fenómeno es que el significado de la invención efectivamente sucede en la supresión de la percepción de ruidos de interferencia debido a que pueden ser bien localizados por el participante oyente y puede por lo tanto, ser seleccionado *a priori*, por no ser parte de la conversación. Esto también es una capacidad especial del oído y cerebro humanos del oído y se muestra particularmente bien en el denominado "efecto de fiesta cóctel" que es a menudo mencionado en este contexto: a pesar del ruido actual que se origina por la habitual superposición de voces, los presentes no tienen casi ningún problema en distinguir exactamente a los hablantes individuales en la distancia, o en concentrarse en aquellos, en los que estén interesados.

La percepción de todos los otros eventos de sonido no seleccionados de este modo de la misma intensidad y aún de eventos de sonido que son aún más ruidosos, son desconocidamente sorprendidos o debilitados a un nivel que no obstaculiza el entendimiento. Utilizando este fenómeno natural, la invención permite una conversación natural que es inmediatamente orientada a cualquier participante en la conversación particular, también en situaciones de conferencia telefónica de cualquier tipo, lo cual se logra realizando una grabación binaural del ambiente del sitio como se libera a la cabeza del participante.

Para un mejor entendimiento exactamente de este aspecto de la invención, se debe indicar que una transmisión binaural de alta calidad permite a un participante en la conversación al otro extremo de una conexión telefónica para experimentar el mundo acústico local propio a partir de la perspectiva relacionada con la propia persona, con todas sus calidades de sonido percibidas, secuencias de tono y otras características espaciales como si fuera una pieza de "cine acústico", sea una que está en un club de Jazz en New York, en el carnaval en Río, o en una playa con oleaje y ruido de gaviotas.

Dentro de esta perspectiva, existe también la posibilidad de agregar o mezclar otras secuencias de tono o sonido en la señal de estéreo binaural transmitida que contiene el sonido local y ambiente de voz: por ejemplo, música o sonidos o todo aquello que sea almacenado en el teléfono móvil que se está usando o en un reproductor de música digital propio, apropiadamente atenuado en su dinámica, para no interferir con la conversación. Si la "localización en la cabeza" que resulta normalmente de las señales de audio convencionales agregadas se evita, la codificación direccional binaural se puede proporcionar para este propósito. La integración de estas diversas funciones adicionales, tales como un teléfono, un reproductor de MP3, consola de videojuegos, ordenador, y similares, en un dispositivo pequeño único representa el estado general hoy día de la técnica, y puede también ser parte de una forma de realización preferida de la presente invención.

A pesar de las demandas relativamente altas hechas en la transferencia de datos alámbricas o inalámbricas por una conexión dúplex completa binaural de banda ancha en un ambiente real -sea aquella vía interruptor de circuito o estructuras de red de interruptor de paquetes-, la transmisión de calidad adecuada para el propósito de la invención se puede lograr con las amplitudes de banda de red y calidad de servicios hoy día disponibles, aunque la selección apropiada a partir de los procedimientos de codificación o decodificación de canal y señal y sus implementaciones potenciales disponibles hoy día. Las conexiones de comunicación de alta calidad en la zona de tecnología de estudio y emisión mencionadas anteriormente, las cuales son realizadas a través de estructuras de red alámbricas de banda ancha, así como también inalámbricamente a través de una comunicación de emisión de punto a punto o con la ayuda de procedimientos de acumulación de canal en redes telefónicas celulares, son ejemplos altamente desarrollados que muestran que existen prerequisites técnicos para la realización de la comunicación binaural en el sentido de la presente invención. La denominada telefonía de Internet, la cual es usualmente conocida como VoIP ("voice-over-internet-protocol", Protocolo de Internet sobre Voz), es una aplicación especial de estructuras de red conmutada de paquete mencionada previamente que pueden ser usadas en conjunto con una interfaz de comunicación inalámbricas existentes, tales como WiMAX o sus sucesores potenciales, tales como Hiper-LAN/2, como parte de las estructuras y procedimientos mencionados anteriormente, que son adecuados para la implementación de la comunicación en tiempo real binaural contemplada con calidad de servicio adecuada.

Ventajas especiales que incrementan o explotan considerablemente las posibilidades de la presente invención, resultan en particular a partir de la movilidad personal garantizada por las medidas citadas en la reivindicación 2, las cuales extienden la transmisión de imágenes auditivas del personal vivo a las secuencias móviles y estructuras de vida personales, abarcando de este modo la completa variedad de situaciones de la vida real, en lugar de limitarse al ambiente local de un conector de línea fijo o, en el caso de una conexión inalámbrica local, a la zona de recepción muy estrecha del mismo.

Dicha telefonía móvil en particular, puede ser vista como la principal aplicación o ciertamente como la aplicación potencial más amplia de estereofonía binaural. Aunque esto nunca ha sido públicamente reconocido antes, estas dos tecnologías están hechas literalmente una para la otra, por decirlo de algún modo, tanto por sus configuraciones técnicas respectivas, como por sus aplicaciones prácticas respectivas. Mediante la fusión de una conexión dúplex móvil en tiempo real con la tecnología de transmisión binaural, el concepto poderosamente emergente de la denominada "telepresencia", puede ser materializado con mayor eficiencia en el campo acústico.

Con referencia al campo específico de tecnología de conferencia telefónica, la cual, sin embargo, no representa la zona central de la aplicación primaria de la invención reivindicada, la invención reivindicada ofrece la ventaja de que para el primer tiempo ambos o en el caso de las denominadas conferencias telefónicas, indiferentemente de si viene de varios participantes que se encuentran en la misma habitación o de uno o varios participantes localizados en otra parte, o si se originan a partir de varias ubicaciones, todos los participantes son habilitados para tener conversaciones con cada uno de los otros participantes, de forma que, en el caso de situaciones móviles, ya sea que la persona hablante cambie la ubicación o que su ambiente acústico cambie debido a que otras personas se unen en, las secuencias de acontecimiento continuamente cambiante, en otras palabras, tal perspectiva oyente actual de la persona, son siempre transmitidas en su vivacidad total. El resultado es la impresión de que el participante en la conversación oyente está, por así decirlo, en la misma habitación con el hablante en el otro extremo, sometido a las mismas funciones de reflexión y difracción cambiantes que normalmente ocurren en una conversación viva, activa, con otra persona en un cierto ambiente, el cual se caracteriza no al menos, por una movilidad personal deseable, y a la cual uno está acostumbrado absolutamente de manera natural.

Debido a este hecho, pero también debido a que la distancia de los micrófonos binauralmente receptores de la boca de la persona hablante respectiva no cambia, las relaciones dinámicas permanecen sin cambio, lo cual significa que el volumen no tiene que ser constantemente ajustado, lo cual ayuda a mantener el entendimiento de voz apropiado

que es de alta calidad comparado con la banda estrecha, prácticamente transmisión de frecuencia de “mera voz” que es todavía practicada exclusivamente hoy día y carece completamente de la calidad de vida de espacios naturales y las estructuras de reflexión y resonancia, difracción multifacéticas, que pueden ser atribuidas al ambiente listo. Esto también carece de la compleja superposición generada por el cuerpo humano, es decir, por el tronco, los
5 hombros, la cabeza, etc., los cuales se disponen finalmente en la función de transmisión relacionada con la cabeza de dos canales, de conformidad con la invención.

Las medidas mencionadas en las reivindicaciones subordinadas describen mejoras y desarrollos ventajosos de la conexión de teléfono estereofónica caracterizadas en la reivindicación principal y en la primera reivindicación de
10 dispositivo.

Dibujos

Algunas formas de realización de la invención se describirán en los dibujos y serán además, ejemplificadas en la
15 siguiente descripción. En los dibujos:

La Figura 1 muestra, en una representación esquemática, una primera forma de realización de la presente invención en la forma de una conexión estereofónica telefónica con dos participantes en diferentes ubicaciones; y

La Figura 2 muestra una segunda forma de realización de la invención, en la cual un primer participante es conectado a través de una conexión estereofónica telefónica con otros tres participantes quienes están juntos en otra ubicación, a modo de conferencia telefónica.

Descripción detallada de ejemplos de formas de realización

La idea fundamental de la presente invención es transportar el ambiente real de cada participante de conversación, a la otra respectiva parte en forma de imágenes acústicas relacionadas- de cabeza tridimensional personal, por medio de una conexión de telecomunicaciones, independientemente de si la conexión se realiza completamente por cable o completa o parcialmente de forma inalámbrica, particularmente también durante la transmisión móvil,
20 disponiendo cada participante al menos de un micrófono doble fijo a señales binaurales grabadas y de audífonos o auriculares estereofónicos.

La Figura 1 muestra lo que significa. El participante A_0 , cuya cabeza está indicada por el número de referencia 10, está conectado a través de un enlace de teléfono estereofónico con el participante B_0 con la cabeza 11.
35

Cada participante A_0 y B_0 usa una combinación 12 contiguamente a cada oreja o dentro de cada uno de sus oídos, pero en cualquier caso dentro de cada zona auditiva, consistiendo dicha combinación en un transductor que genera sonido, usualmente un auricular o audífono 13, y un micrófono 14, y en conjunto las dos combinaciones proporcionan una grabación estereofónica y también la reproducción de eventos de sonido. Los micrófonos 14 se encuentran, por lo tanto, contiguamente o dentro de las zonas auditivas, de manera que, al funcionar estereofónicamente en conjunto, sean capaces de grabar exactamente las imágenes acústicas, denominadas imágenes relacionadas con la cabeza, las cuales, en efecto, representan el ambiente acústico real del participante. Se entiende que se debe tener cuidado en la integración de los micrófonos con los transductores que generan sonido vecino, de este modo, los auriculares o audífonos, son elaborados en una forma y/o son provistos de los
40 medios que son adecuados para evitar el eco y la retroalimentación, de manera que el respectivo participante en la conversación, no tiene su propia retroalimentación de voz a él. Dicho aislamiento mutuo entre los auriculares o audífonos y micrófonos, que asegura libertad de la retroalimentación puede ser rutinariamente realizado por el experto en la materia.

Como se mencionó, los transductores que generan sonidos, pueden ser de varios tipos, por ejemplo, auriculares supraaurales o, preferentemente, audífonos, de manera que las disposiciones de soporte que rodean la cabeza pueden ser evitadas. En cualquier caso, para la amplificación y ecualización de las señales grabadas o reproducidas, los dos micrófonos (los cuales forman en conjunto un micrófono estéreo), así como también los dos transductores que generan sonido 13, están cada uno seguido por unos circuitos amplificadores/ecualizadores 15a para los transductores que generan sonido y 15b para los micrófonos, a los cuales están conectados a través de interfaces de dos canales bilaterales 16a. Si se consideran las combinaciones 12 asignadas a cada participante como un primer montaje, entonces los circuitos amplificadores/ecualizadores con la interfaz asignada 16a, forman un segundo montaje 17, el cual para las transmisiones de dos canales está conectado inalámbrica o alámbricamente al terminal de comunicación 18 conectado aguas abajo, el cual entonces, nuevamente inalámbrica o alámbricamente,
50 asegura la transmisión de la señal de dos canales a la red.

Independientemente de si se utilizan auriculares externos, es decir, auriculares cerrados o abiertos supra-aurales o un sistema de reproducción en la oreja, las señales telefónicas estereofónicas relacionadas con la cabeza siempre aparecen, las cuales en el caso de audífonos, a los cuales están fijados los micrófonos o están asociados de otro modo, incluso una parte del pabellón auricular se percibiría como un cuerpo de reflexión, difracción y resonancia,
65 mejorando de manera complementaria la naturalidad de las señales emitidas.

- Debido a las considerables posibilidades ofrecidas hoy día, y predeciblemente en el futuro también, por el desarrollo técnico continuo de procedimientos con respecto a la integración de componentes y la miniaturización creciente, se produce una especial ventaja por la utilización de audífonos también porque las combinaciones miniaturizadas 12 en este caso, pueden ser realizadas, respectivamente, con un audífono y un micrófono aún sin conductos alámbricos y por lo tanto, de manera diferente a las ilustraciones en los dibujos; con una batería de suministro común para el audífono y el micrófono de cada combinación 12 y un transmisor de distancia ultra-corta común al siguiente montaje 17, se logra una mayor comodidad para el usuario. Ninguna de las conexiones alámbricas está colgando alrededor de la cabeza del participante y con la excepción de la combinación y micrófono y el generador transductor de sonido ligeramente obturado en cada oído, no se produce ningún tipo de incomodidad. Como cualquier usuario de un reproductor de MP3 portátil como los "iPods" conocidos, los audífonos son usualmente particularmente ventajosos en tanto que se asemejan a auriculares abiertos, esto es, no aíslan al usuario de su ambiente acústico, facilitando de este modo cualquier clase deseada de comunicación.
- Actualmente, los bloques de circuito respectivos en los montajes 12, 17, y 18, se explican por sí mismos para el experto en la materia, en la leyenda indicada en la figura. Los circuitos ecualizantes son usados para la estandarización de señal, los cuales pueden ser necesarios cuando los respectivos participantes en la conversación trabajan con diferentes receptores de cabeza realizados en dos series inalámbricas para orejas, consistiendo cada una de ellas en un micrófono y un audífono, para lograr una comparabilidad con las otras señales. Esto también puede ser importante debido a la libertad deseada de retroalimentación, la cual también puede depender del posicionamiento del micrófono. En esta forma, los ecualizadores proporcionan compensaciones que finalmente suministrarán una señal estandarizada a la interfaz, que conecta con el terminal de comunicación.
- Para la separación deseada o significativa de los montajes individuales, las interfaces 16a, 16b y 16c son requeridas por tener un diseño de dos canales de alta calidad correspondientemente. Están conectadas por alambre o inalámbricamente, a través de ondas electromagnéticas, a las interfaces correspondientes del siguiente montaje.
- Fundamentalmente, se indica que la separación y asignación de los diferentes montajes constructivos y/o bloques de circuitos realizados principalmente en el dibujo sirven para proporcionar un mejor entendimiento y representación visual de las funciones básicas según la invención. Se entenderá que, no menos atribuible al progreso técnico deseado o para propósitos diferentes o utilidad en la asignación de las partes variadas o su diseño, se puede realizar o utilizar otro agrupamiento de los bloques de circuito, así como unos circuitos de procesamiento de señal conformados de otro modo y conectados entre sí.
- La figura 2 describe una forma de realización ventajosa de la invención, en la que al menos existen varios participantes B, B', B'' en un lado que, en este caso, se localizan en el mismo lugar, con cada participante B, B', B'' lleva un receptor de cabeza que comprende una combinación 12, la cual consiste en un micrófono y en un transductor que genera sonido para cada oído, como en el participante en la conversación A, con quien cada uno de los participantes B, B', B'' tiene una conexión de dos canales a través de la red. Para este propósito, cada una de las terminales de comunicación participantes 18' en la Figura 2, y además los participantes A', A'' en la conferencia pueden ser localizados en la zona del participante A, así como también es modificado en lo referente a que se proporciona un circuito de selección de función de grabación adicional 19 en la interfaz 16b' de entrada múltiples de dos canales. Esto sirva para el propósito, en una primera variante, de decidir automáticamente cuál de dicho par de micrófonos del participante de conversación es para ser cambiado a la salida de interfaz de red 16c' y así de este modo, ser liberado para transmisión vía la red. Esto puede ocurrir, por ejemplo, comparando las dinámicas de las señales de voz generadas por los participantes B, B', B'' a un tiempo dado o determinando en cuales de los participantes están hablando todos. El circuito de selección de función de grabación entonces bloquea la transmisión de la señal de micrófono a partir de los otros participantes, pero por supuesto, no la transmisión de señal de sonido a los transductores que generan sonido de cada participante adicional.
- También existe la posibilidad de que las señales ambientales y de habla que proceden de cada hablante sean suministradas de retorno complementariamente, es decir, a través del dispositivo, o mejor a través del transductor a los otros participantes que se encuentran en la misma habitación.
- Si en algún momento, el participante correspondiente conectado que habla en primer lugar del grupo de tres, es decir, por ejemplo el participante B, y otro participante, quizás B'', comienza a hablar, entonces el circuito de selección de función de grabación del dispositivo de comunicación 18', en este caso, cambia automáticamente el participante B'' en la interfaz de red 16c'. Sin embargo, esto no significa que el participante A en el otro extremo de la red pueda necesariamente escuchar solamente al participante de la conferencia que se ha desconectado en último lugar; sin embargo, él por supuesto, continuará escuchando, pero aún de una forma más reservada a partir de las condiciones ambientales respectivas, todos los otros participantes a través del micrófono estéreo del participante B'', de manera que en este caso también, el panorama estéreo completo en el sonido tridimensional resultará para el participante A, prácticamente como si compartiera el mismo ambiente con los participantes B, B', B''.
- Es posible, además de o en lugar de una conmutación automática, diseñarlo para una operación arbitraria, de manera que, por ejemplo, el participante que desee hablar, pueda operar deliberadamente un cambio en su

disposición por medio del cual será cambiado por el circuito de selección de función de grabación en la red. También se pueden disponer teclas de silencio si, por ejemplo, se mantiene una conversación corta discreta.

5 Es también ventajoso disponer una pantalla de control, diodos que emiten luz, o medios similares, en la zona del
circuito de función de grabación o en otra ubicación apropiada, para mostrar cuál de los participantes de la
conversación está siendo conmutado por el circuito de selección de función de grabación 19 a la salida de la interfaz
de red 16c'. Puesto que los otros componentes de cambio de la variante de conferencia telefónica de la Figura 2,
corresponden a los bloques del circuito de la Figura 1, no es necesario tratarlo en este punto. La Figura 2 también
omite la repetición de los signos de referencia numerados de los bloques del circuito que ya han sido discutidos y
representados con sus funciones en la Figura.
10

Con respecto a los decodificadores o codificadores de señal y canal en el terminal de comunicación, se debe
agregar que los decodificadores y codificadores realizan una conversión digital/análoga y viceversa, así como
también determinaciones de amplitud de banda (la amplitud de las señales puede ser al menos 3.4 kHz u 8 kHz de
15 16 kHz). También aseguran las diferencias de retardo de grupo más bajas posibles, mientras se tiene cuidado de no
cambiar la coherencia entre los canales durante la codificación y decodificación que ajusta la señal total a la
respectiva red, las señales estéreo actuales son ya multiplexadas en conjunto con cualquiera de los datos auxiliares
en una señal única en este punto. Además, estos codificadores y decodificadores proporcionan la redundancia
requerida, así como también la detección y corrección por error. Idealmente, los tiempos de funcionamiento
20 unidireccionales que incluyen la codificación/decodificación de señal y transmisión, se mantienen por debajo de 120
milisegundos, de manera que una transmisión de señal sincronizada es asegurada sin retardos de interferencia.

Otra forma de realización ventajosa de la invención, debe ser mencionada también, la cual consiste en disponer un
micrófono individual adicional, preferentemente próximo a la boca de cada participante en la conversación, el cual es
25 mezclado en la señal estéreo como un tipo de micrófono de apoyo para además, mejorar la inteligibilidad, por
ejemplo, o alternativamente, puede reemplazar completamente la señal estéreo del par de micrófono binaural. Sin
embargo, haciendo esto, se regresa a la zona de telefonía monofónica convencional, aunque los receptores de
cabeza binaurales estén siendo usados, todavía se podría implementar esta posibilidad cuando, en ciertas
circunstancias y posiblemente aún a partir del inicio de una conversación, la transmisión estereofónica binaural de
30 información ambiental, no sea relevante o no deseada. Esto puede ocurrir, por ejemplo, cuando en el curso de una
conversación, tal como una operación monofónica, es conmutada para el transporte de información vocal, con ello,
la amplitud de banda de la transmisión de señal, así como también los costes relacionados pueden ser reducidos.
Las medidas correspondientes pueden ser integradas en la configuración existente sin ningún problema, con el
complemento de un fácil accionamiento del conmutador por el lado del participante.
35

Se entiende que todas las características mencionadas en la descripción, en las siguientes reivindicaciones, y en
particular también en los dibujos adjuntos, pueden ser esenciales para la presente invención por sí mismas, así
como también en cualquier número de combinaciones entre ellas.

REIVINDICACIONES

- 5 1. Procedimiento para la grabación, transmisión y reproducción estereofónica de eventos de sonido para aplicaciones de comunicación en telefonía, estando previstos unos auriculares para cada participante y estando disponibles unos micrófonos, caracterizado porque a cada una de las zonas auditivas de cada participante, está asociada una combinación espacialmente estrecha, pero conectada sin retroalimentación, formada respectivamente por un audífono o auricular y un micrófono, a través de la cual el ambiente real grabado binauralmente, relacionado con la cabeza de cada participante grabado en su comportamiento de reflexión, difracción y resonancia es transmitido a cada uno de los otros participantes en forma de imágenes sonoras y acústicas estereofónicas correspondientes a través de una conexión de dos canales.
- 10
- 15 2. Procedimiento según la reivindicación 1, caracterizado porque la doble combinación, que consiste, en cada caso, en una combinación asociada con cada oreja es parte de un receptor y un emisor telefónico móvil, que funciona con batería, el cual está montado en el cuerpo de una persona que participa en el tráfico telefónico o que es llevado por dicha persona para transmitir las respectivas imágenes acústicas personales relacionadas con la cabeza.
- 20 3. Procedimiento según la reivindicación 1 o 2, caracterizado porque en el caso de conferencias telefónicas, cada participante en la conversación que se encuentra en la misma habitación, es cambiado selectivamente a una red que conecta a todos los participantes en la conversación, ya sea mediante una conmutación automática, producida por el propio proceso de conversación, o mediante una conmutación que puede funcionar arbitrariamente a través de un circuito de selección de función de grabación.
- 25 4. Procedimiento según la reivindicación 1, 2 o 3, caracterizado porque en el caso de varios participantes en una conferencia telefónica, que se encuentran en una misma habitación, además de la transmisión espacial acústica natural, por lo menos la señal de conversación es suministrada a los auriculares o audífonos binaurales de cada uno de los participantes que no esté hablando en un respectivo momento, a través de un bucle de retroalimentación eléctrico, por ejemplo, en un aparato de comunicación.
- 30 5. Procedimiento según una de las reivindicaciones 1 a 4, caracterizado porque complementariamente a la transmisión ambiental y de habla estereofónica, las secuencias de sonido y tono, que son almacenadas en el respectivo teléfono estereofónico y también móvil utilizado son transmitidas según los deseos de cada uno de los participantes en la conversación.
- 35 6. Procedimiento según la reivindicación 5, caracterizado porque para evitar una localización en la cabeza de dichas señales de audio mezcladas se lleva a cabo una codificación direccional binaural.
- 40 7. Dispositivo para realizar el procedimiento según una o varias de las reivindicaciones 1 a 6, caracterizado porque para la grabación, transmisión y reproducción para cada zona auditiva de un participante en una conversación (A, B, B', B''), está prevista una combinación (12), respectivamente de un transductor (13) y un micrófono (4), estando el transductor y el micrófono en cada combinación (12) dispuestos espacialmente próximos entre sí para evitar fenómenos de eco y/o de retroalimentación, para la grabación, transmisión y reproducción binaural simultánea de imágenes sonoras y acústicas estereofónicas reales, incluyendo el comportamiento de reflexión, difracción y resonancia relacionado con el entorno y la persona.
- 45 8. Dispositivo según la reivindicación 7, caracterizado porque la doble combinación que permite la grabación y reproducción binaural, formada respectivamente por una combinación que consiste en un transductor y un micrófono en cada oreja de cada uno de los participantes en una conversación telefónica, forma parte de un dispositivo telefónico móvil que funciona con batería, que es llevado por cada participante en la conversación.
- 50 9. Dispositivo según la reivindicación 7 u 8, caracterizado porque para el postprocesado de dos canales de las señales obtenidas o recibidas binauralmente, están previstos, respectivamente, unos circuitos de amplificación/ecualización (15a, 15b) separadamente para cada uno de los transductores y los micrófonos, además de unos circuitos codificadores y decodificadores de señal y canal.
- 55 10. Dispositivo según la reivindicación 7, 8 o 9, caracterizado porque entre los circuitos de procesamiento de señales individuales, respectivamente, en función de su asignación entre sí, las interfaces de dos canales (16a, 16b), así como una interfaz terminal de red (16c), están conectadas con una transmisión alámbrica o inalámbrica entre las interfaces o entre la interfaz terminal de red y la red.
- 60 11. Dispositivo según una o varias de las reivindicaciones 7 a 10, caracterizado porque en el caso de varios participantes en una conversación que se encuentran en una misma habitación, está prevista una interfaz múltiple de dos canales (16b), con la cual están conectados los participantes en una conversación individuales, tras una preparación de señal propia correspondiente, y porque está previsto un circuito de selección de función de grabación (19) para una conmutación que puede ser accionada automática o arbitrariamente de señales de micrófono grabadas para transmitir las posteriormente a la red de uno de los participantes en la conversación (B) al otro (B', B'').
- 65

5 12. Dispositivo según una de las reivindicaciones 7 a 11, caracterizado porque aguas abajo de un primer grupo de circuitos que comprende las dos combinaciones de zonas auditivas (12), formadas, respectivamente, por un micrófono (14) y un auricular o audífono (13), está montado un segundo grupo de circuitos (17), que consiste en unos circuitos amplificadores/ecualizadores para auriculares o micrófonos, que está conectado con un terminal de comunicación (18) conectado aguas abajo a través de unas interfaces de dos canales (16a, 16b) asignadas a ambos lados, y conectadas de manera alámbrica o inalámbrica, comprendiendo el terminal de comunicación (18) unos codificadores y decodificadores de señal, así como unos codificadores y decodificadores de canal y, en el caso de conferencias telefónicas hacia una interfaz múltiple (16b) en el lado de salida, un circuito de selección de función de grabación (19) con unas pantallas de control y monitorización para la conmutación, que se puede accionar automáticamente o arbitrariamente del respectivo participante en la conversación que esté hablando en cada momento en la red telefónica.

10 13. Dispositivo según una o varias de las reivindicaciones 7 a 12, caracterizado porque presenta una disposición de conmutación, a través de la cual la comunicación estereofónica puede ser conmutada bilateralmente a una operación monofónica en cualquier momento deseado.

15 14. Dispositivo según una o varias de las reivindicaciones 7 a 13, caracterizado porque comprende un tercer micrófono de apoyo situado cerca de la boca para la adición de una señal de habla, en caso de necesidad, a la señal estéreo transmitida o para permitir el funcionamiento monofónico según la reivindicación 13.

20

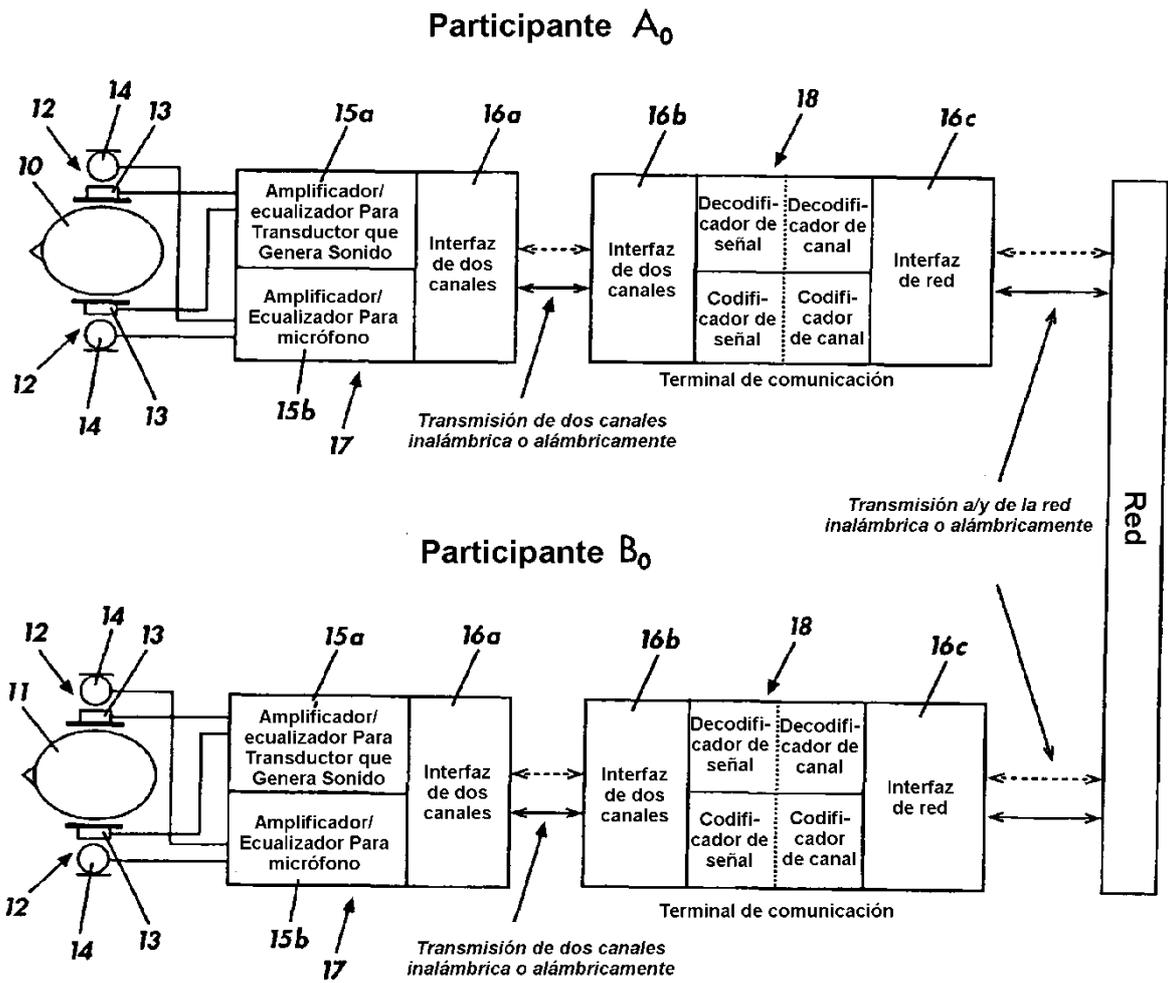


Fig. 1

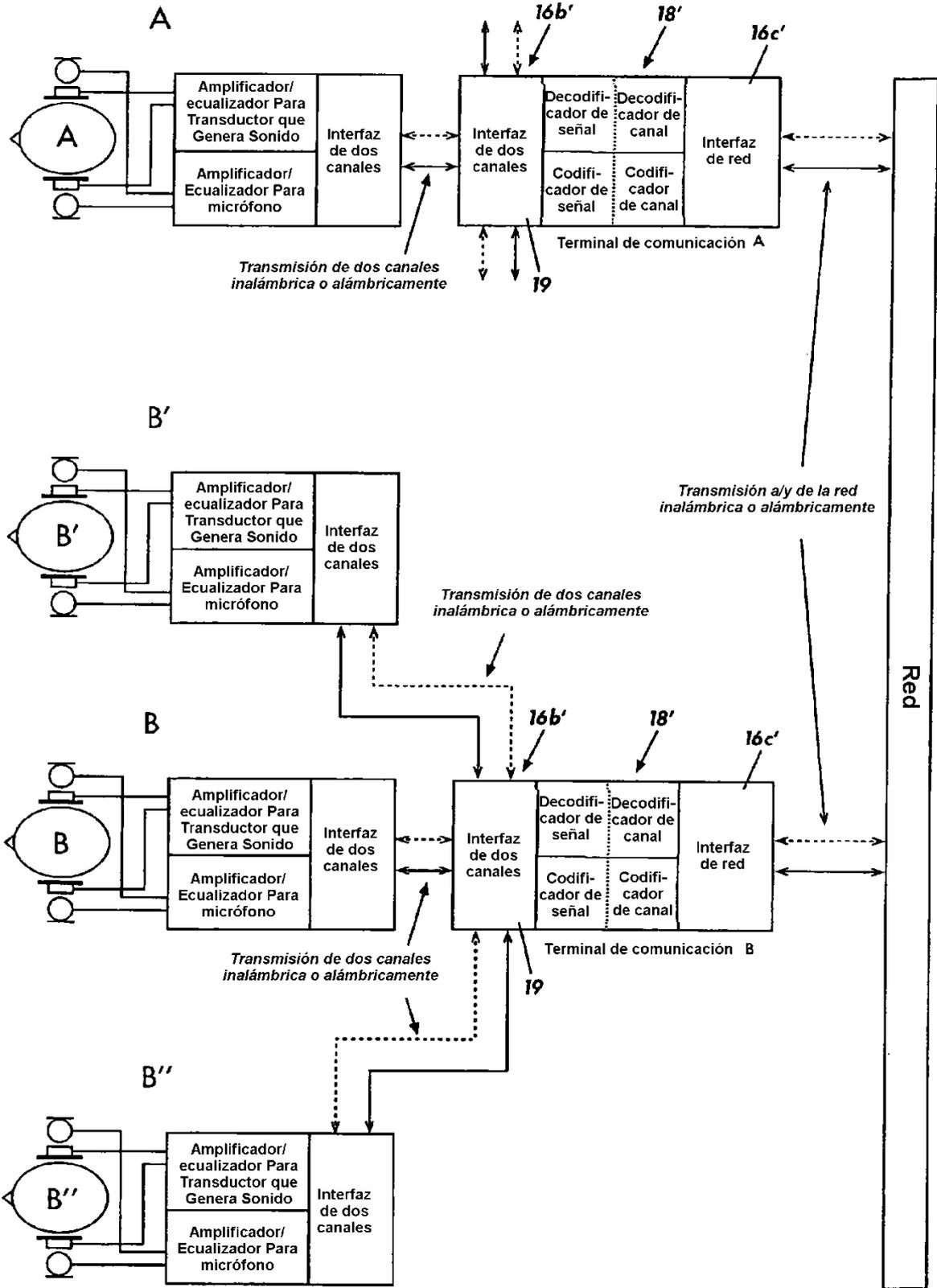


Fig. 2