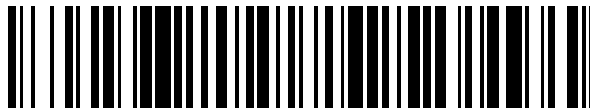


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 626 054**

51 Int. Cl.:

**H04M 3/56**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.12.2012 PCT/EP2012/076972**

87 Fecha y número de publicación internacional: **03.07.2014 WO14101944**

96 Fecha de presentación y número de la solicitud europea: **27.12.2012 E 12816078 (5)**

97 Fecha y número de publicación de la concesión europea: **22.02.2017 EP 2939407**

54 Título: **Sistema de conferencia y proceso para activación por voz en el sistema de conferencia**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**21.07.2017**

73 Titular/es:

**ROBERT BOSCH GMBH (100.0%)  
Postfach 30 02 20  
70442 Stuttgart, DE**

72 Inventor/es:

**VAN DER SCHAAR, HANS;  
DERKX, RENE;  
SMAAK, MARC;  
BONARIUS, JOCHEM y  
KEES, JANSE**

74 Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

ES 2 626 054 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 de conferencia y proceso para activación por voz en el sistema de conferencia

Estado de la técnica

5 La invención se refiere a un sistema de conferencia. Más específicamente, la invención se refiere a un sistema de conferencia que comprende una pluralidad de unidades delegadas, cada unidad delegada tiene un micrófono para recibir una señal de audio de un módulo circundante y un módulo de servicio central que maneja una pluralidad de canales de contribución, con lo cual la salida de audio de los canales de constitución contribuye a una salida de audio amplificada del sistema de conferencia, con lo cual cada unidad delegada se adapta para transmitir una solicitud de confirmación de canal de contribución al módulo de servicio central, con lo cual el módulo de servicio se adapta para otorgar la solicitud y asignar un canal de contribución a la unidad delegada de solicitud, fijando de esta manera la unidad delegada de solicitud en un estado activo. Adicionalmente la invención se refiere a un proceso para activación por voz en el sistema de conferencia.

10 Se utilizan sistemas de conferencia por ejemplo en debates y usualmente comprende una pluralidad de unidades delegadas con micrófonos, con lo cual en un debate cada participante en el debate utiliza su propia unidad delegada. Usualmente la unidad delegada tiene un interruptor o similar, que permite al participante enfrente de la unidad delegada solicitar, que su micrófono se active, de modo que la voz del participante se ingresa en el sistema de conferencia y se amplifica mediante el sistema de conferencia.

15 Por ejemplo, el documento EP 1 686 835 A1, probablemente representa la técnica anterior más cercana, divulga un sistema de conferencia con una pluralidad de dichas unidades delegadas, con lo cual se colocan diversos indicadores sobre la unidad delegada para indicar diferentes estados de la unidad delegada.

20 De acuerdo con la invención se propone un sistema de conferencia con las características de la reivindicación 1. Adicionalmente se presenta un proceso con las características de la reivindicación 11. Se divulgan realizaciones preferidas o ventajosas de la invención mediante las reivindicaciones dependientes, la descripción y las figuras adjuntas.

25 Resumen de la invención

De acuerdo con la invención se proporciona un sistema de conferencia, que se puede operar para que sea instalado en una sala de conferencias o en una sala plenaria y se utiliza por ejemplo por parte de políticos u otros participantes que debaten entre sí. El sistema de conferencia también puede ser denominado acuerdo de conferencia

30 El sistema de conferencia comprende una pluralidad de unidades delegadas. Preferiblemente más de 50, especialmente más de 100 unidades delegadas se utilizan en el sistema de conferencia. Cada unidad delegada tiene un micrófono para recibir una señal de audio de un entorno. La señal de audio es especialmente una señal de voz de un participante de pie o sentado en frente de la unidad delegada y de esta manera en frente del micrófono. Preferiblemente la unidad delegada se incorpora como una unidad superior de mesa que tiene un cuerpo base opcionalmente con o sin un orador, con lo cual el micrófono se dispone sobre el cuerpo base. Para procesamiento de datos se prefiere, que la unidad delegada comprenda una unidad de procesamiento local, integrada preferiblemente en el cuerpo base.

35 El sistema de conferencia comprende adicionalmente un módulo de servicio central, que puede funcionar para manejar una pluralidad de canales de contribución, especialmente canales de amplificación, con lo cual la salida de audio de los canales de contribución está contribuyendo a una salida de audio amplificada del sistema de conferencia. La salida de audio amplificada del sistema de conferencia es la atmósfera de audio resultante proporcionada por el sistema de conferencia en el entorno. Preferiblemente el sistema de conferencia puede funcionar para amplificar señales de audio de las unidades delegadas a señales de audio amplificadas al utilizar los canales de contribución, con lo cual las señales de audio amplificadas hacen parte de la atmósfera audio, por ejemplo, en el salón de plenarias o sala de reuniones.

40 Cada unidad delegada se adapta para enviar o transmitir una solicitud de confirmación de canal de contribución al módulo de servicio central. Con la solicitud la unidad delegada solicita que se conecte con uno de los canales de contribución con el fin de transmitir la propia señal de audio al canal de contribución, de tal manera que la señal de audio se amplifica y forma una parte de la salida de audio amplificada y de esta manera una parte de la atmósfera de audio. El módulo de servicio central se adapta a para otorgar la solicitud y asignar uno de los canales de contribución a la unidad delegada solicitante. Con este procedimiento la unidad delegada solicitante se fija en un estado activo y cambia su estado desde una unidad delegada pasiva hasta una unidad delegada activa.

45 El módulo de servicio central se puede incorporar como un ordenador, especialmente como un servidor. La comunicación entre el módulo de servicio central y las unidades delegadas es preferiblemente una comunicación digital, especialmente utilizando un protocolo de red, para enviar la solicitud y/o para transmitir las señales de audio.

De acuerdo con la invención, la unidad delegada se adapta para activar la solicitud y, por consiguiente, la transmisión de la solicitud mediante activación por voz. El sistema de conferencia permite de esta manera, que las unidades delegadas se cambian desde el estado pasivo hasta el estado activo mediante activación de voz. La activación por voz significa, que un orador enfrente de la unidad delegada tiene sólo que empezar a hablar con el fin de iniciar el activador de la solicitud. La ventaja del modo de activación por voz es, que los participantes de las discusiones no tienen que solicitar contribuir a una discusión al presionar un botón o similar, pero pueden unir la discusión simplemente al empezar a hablar. El método de activación por voz no sólo es útil para un debate interactivo, sino que también se puede utilizar en otros métodos de debate para limitar el número de canales que contribuyen a la baja de audio.

Una de las unidades delegadas, que puede activar la solicitud, se denomina posible unidad delegada de solicitud, con lo cual cada una de las unidades delegada puede ser dicha unidad de solicitud posible. La unidad delegada de solicitud posible se adapta para activar la solicitud en caso de que se cumpla por lo menos una primera condición de activación, que reclama que su nivel de señal de audio de la unidad delegada de solicitud posible es mayor que un valor de prueba individual para cada unidad delegada que está en estado activo. El propio nivel de señal de audio de la unidad delegada de solicitud posible es el nivel de señal de audio recibida por el micrófono de unidad delegada de solicitud posible. Para cada unidad delegada activa se estima o calcula un valor de una prueba individual. El valor de prueba individual es un nivel de señal de audio calculado o estimado de la unidad de solicitud posible que resulta de una señal de voz o audio proporcionada a las unidades delegadas activas y que se acopla en el micrófono de la unidad de solicitud posible.

La idea subyacente de la invención es, que en caso que las unidades delegadas se dispongan cerca entre sí, por ejemplo a una distancia menor de 1 metro, especialmente menor de 60 centímetros, y con respecto al hecho de que las unidades delegadas frecuentemente tienen un rango dinámico grande de entrada de audio, la unidad delegada no puede determinar, si es la unidad delegada de solicitud posible o una unidad delegada vecina sin información adicional. Adicionalmente las condiciones de la sala (reflexión y reverberación) pueden provocar puntos en la sala en donde se acumula el audio al nivel que se reconoce falsamente como un orador.

Al realiza la prueba de la primera condición de activador, la unidad de legada de solitud posible prueba si la propia señal de audio es recibida de un orador o un participante que habla en el propio micrófono o se recibe una señal de voz desde un orador o un participante que habla en la una unidad delegada vecina. En ese caso, el valor de prueba individual de la unidad delegada vecina sería mayor que el propio nivel de señal de audio y la unidad delegada de solicitud posible concluye que no se activaría por voz. Por lo tanto, la unidad delegada de solicitud prueba si la primera condición de activador es verdadera o falsa.

El valor de prueba individual para una unidad delegada de prueba, se estima preferiblemente, especialmente calculada al multiplicar un factor de acoplamiento acústico individual, que se define para cada par de unidades delegadas de solicitud posible y unidad delegada de prueba, con el nivel de señal de audio de la unidad delegada de prueba, durante el período de prueba y, opcionalmente, con un factor umbral. Por lo tanto, para cada par (unidad delegada de solicitud posible <-> unidad delegada de prueba) se define un factor de acoplamiento acústico individual. El nivel de señal de audio de la unidad delegada de prueba se define como el nivel de señal de audio recibida del micrófono de la unidad delegada de prueba respectiva. El período de prueba preferiblemente es más corto de 1 segundo, especialmente menor de 0.1 segundo. Preferiblemente solo las unidades delegadas activa se evalúan como delegados de prueba.

En una realización alterna de la invención, el máximo de los niveles de señal de audio del delegado de prueba para los pocos últimos periodos de tiempo se utiliza en cambio del nivel de señal de audio del periodo de prueba actual. Por ejemplo, por lo menos los últimos 3, 5 o 10 periodos de tiempo se utilizan para determinar el máximo. Esta alternativa mejora la robustez del valor y, de esta manera, la robustez de la activación de voz.

En un desarrollo adicional de la invención, se actualiza el valor de prueba individual para un bloque de muestras, por ejemplo, un bloque de muestras 1024 mediante una frecuencia de índice de muestreo de 48 kHz, de tal manera que se actualiza el valor de prueba individual cada 21 ms. Se pueden utilizar otros valores para el índice muestreo y longitud de bloque. Se prefiere adicionalmente que la prueba de la primera condición de activador se realice para cada muestra. Para lo tanto para el ejemplo mencionado anteriormente, la prueba se realiza cada 1/48000 s.

El factor de acoplamiento acústico individual describe o se representa por la relación entre el nivel de señal de audio de la unidad delegada de solicitud posible y el nivel de señal de audio de la unidad delegada de prueba en caso de que se genera una señal de voz desde un orador que utiliza la unidad delegada de prueba. En otras palabras, el factor de acoplamiento acústico individual se puede derivar en la situación de que un orador habla en el micrófono de la unidad delegada de prueba, con lo cual la señal de audio de la unidad delegada de solicitud posible se divide en el nivel de señal de audio de la unidad delegada de solicitud para recibir el factor de acoplamiento acústico. En otras palabras, el factor de acoplamiento acústico describe la relación de que el micrófono de la unidad delegada de solicitud posible y el micrófono de la unidad delegada de prueba reciben de la señal de voz.

El sistema de conferencia también es tolerante en el caso en el que la unidad delegada de solicitud posible recibe una señal de audio de su propio u otro orador. En este caso, el factor de acoplamiento acústico individual también convergerá actualmente hacia el audio desde el orador, evitando por lo tanto que se active una solicitud por la señal de orador. Esto es posible, debido a que se sabe qué unidades delegadas (activas) están contribuyendo a la señal del

orador. Para reducir el acoplamiento del altavoz local al micrófono es posible implementar un cancelador de eco corto acústico, que filtra la señal del orador de las unidades delegadas de su señal de micrófono.

5 En una realización preferida cada unidad delegada comprende una unidad de factores que contienen ID individuales de las otras unidades delegadas (que comprenden unidades delegadas pasivas y activas) y los factores de acoplamiento individuales. Se subrayará, que las tablas de factor de las diversas unidades delegadas son diferentes entre sí. Adicionalmente la unidad delegada comprende una tabla de nivel de señal de audio que contiene ID individuales de las otras unidades delegadas, que están en estado activo y el nivel de señal de audio durante el período de prueba. Con estas dos tablas la unidad delegada de solicitud posible es capaz de realizar la prueba que se relaciona con la primera condición activadora.

10 En una realización preferida de la invención las tablas de factor se gestionan por las unidades delegadas y se presentan preferiblemente en la unidad delegada. Por ejemplo, las unidades delegadas tienen una memoria para presentar la tabla de factores. La tabla de nivel de señal de audio se proporciona mediante el módulo de servicio central. Como la tabla de nivel de señal de audio se debe actualizar para cada período de prueba, la tabla de nivel de señal de audio puede por ejemplo se distribuida a las unidades delegadas mediante métodos de distribución  
15 multidifusión o transmisión.

Una posible mejora de la invención la unidad delegada se adapta para estimar los factores de acoplamiento acústicos individuales para cada una de las unidades delegadas en una forma iterativa, con lo cual en cada etapa de iteración se mejora un valor de inicio del factor de acoplamiento acústico individual. En una realización posible todos los factores de acoplamiento individuales se fijan en  $1.0 = 0$  dB durante instalación o inicialización del sistema de conferencia. Tan pronto como se crea una situación para estimar los factores de acoplamiento individuales, se realiza una siguiente etapa de iteración. Dicha situación es dada solo a un parlante que utiliza una unidad delegada del sistema de conferencia. En esta situación los factores de acoplamiento individuales de todas las otras unidades delegadas que se relacionan con la unidad delegada utilizada se pueden mejorar en una etapa iterativa. Por lo tanto, el sistema de conferencia es de autoaprendizaje y de esta manera se auto optimiza con el tiempo.

25 En una posible mejora adicional de la invención una de las unidades delegadas, que solicita un canal de contribución sobre la base de los datos de un primer período de prueba y que se dedica a un canal de contribución al asignar el canal de contribución a la unidad delegada, se adapta para revisar la solicitud y de esta manera la asignación al volver a probar por lo menos la primera condición activadora sobre la base de los datos de un segundo período de prueba. La idea subyacente de esta mejora es la siguiente situación: por ejemplo, tres unidades delegadas dispuestas adyacentes entre sí están en un estado pasivo durante el primer periodo de prueba y todos los tres reciben la misma señal de voz. En esta situación puede suceder, que todas las tres unidades delegadas activen una solicitud como se describió anteriormente, ya que las tres unidades delegadas están en un estado pasivo y no se prueban ellas mismas entre sí. Después que cada una de las tres unidades delegadas se asigna a un canal de contribución, la primera condición se vuelva a ensayar sobre la base de los datos de un segundo período de prueba, preferiblemente siguiendo  
30 después del primer periodo de prueba, especialmente es el siguiente período de prueba después del primer periodo de prueba. Durante el segundo periodo de prueba, las tres unidades delegadas están en estado activo. Al probar la primera condición de activador sobre la base de los datos del segundo periodo de prueba, dos de las tres unidades probaran a primera condición de activador como falsa, ya que las tres unidades delegadas prueban la primera condición activadora entre sí.

40 En una realización de mejora adicional de la invención, se reivindica que la unidad delegada se adapta para activar la solicitud en caso de que por lo menos la primera condición de actividad y una segunda condición de activador se cumple, con lo cual la segunda condición activadora define, que el nivel de señal de audio de la unidad delegada de solicitud posible es mayor que un nivel de ruido de referencia durante el período de prueba. Los datos del nivel de ruido se proporcionan preferiblemente por el módulo de servicio central y se proporcionan especialmente junto con la  
45 tabla de nivel de señal de audio.

Una mejora posible adicional para evitar el otorgamiento de solicitudes de dos unidades delegada, que resultan de la misma señal de voz es que el módulo de servicio central se adapta para otorgar solo una solicitud durante un tiempo muerto preseleccionado. Regresando al ejemplo anterior, todas las unidades delegadas enviarán su solicitud en una corta ventana de tiempo. El módulo de servicio central otorgará sólo la primera solicitud y denegará las otras solicitudes durante el tiempo muerto.

55 En una mejora adicional de la invención, la unidad delegada comprende un dispositivo de indicación de parlante para indicar un estado de parlante de la unidad delegada, con lo cual se activa la indicación de parlante en caso de que la unidad delegada este en el estado activo como una primera condición de indicación y bajo una segunda condición de indicación de que se detecta un tono de voz. Con esta mejora se asegura que el dispositivo de indicación sólo se activará si la señal de audio de activación es una señal de voz.

Características adicionales, ventajas y detalles de la invención serán evidentes mediante la descripción de una realización de la invención. La figura muestra:

La figura 1 es un diagrama de bloques de un sistema de conferencia como una realización de la invención.

La figura 1 muestra una vista general esquemática de un sistema 1 de conferencia como una realización de la invención. El sistema 1 de conferencia comprende una pluralidad de unidades 2 delegadas etiquetadas adicionalmente con las letras i, p1, p2, p3, p(n-1), pn y un módulo 3 de servicio central, que se conectan con las unidades 2 delegadas mediante una red 4. El módulo 3 tal de servicio central se puede incorporar como un servidor de ordenador u otro servidor o puede formar un módulo separado a un servidor.

Cada una de las unidades 2 delegadas comprende un micrófono 5 para recibir una señal de voz de un orador o un participante de una discusión. El módulo 3 de servicio central organiza una pluralidad de canales de contribución, con lo cual los canales de contribución son canales amplificadores, de tal manera que una señal de audio, que se envía desde la unidad 2 delegada a uno de los canales de contribución se amplificará y emitirá como una señal de audio amplificada a un entorno.

El sistema de conferencia puede por ejemplo ser instalado en una sala de plenarias, con lo cual cada asiento de la plenaria se equipa con una de las unidades 2 delegadas. Durante una discusión en la sala de plenarias un participante del debate que utilizar una de las unidades delegadas puede hablar en el micrófono 5 de la unidad 2 delegada de tal manera que se recibe una señal de audio por la unidad 2 delegada. La señal de audio se transmite al canal de contribución, se amplifica y emite en la sala de plenaria, de tal manera que los otros participantes pueden escuchar la señal de audio.

Con el fin de tener un debate bien regulado, algunas de las unidades 2 delegadas están en un estado A activo y permiten al participante hablar en la discusión y algunas de las unidades 2 delegadas están en estado P pasivo, con lo cual la señal de audio no se amplifica ni emite en el piso. En caso de que las unidades 2 delegadas estén en el estado A activo se asigna uno de los canales de la contribución del módulo de servicio central a la unidad 2 delegada.

Durante operación las unidades 2 delegadas se conmutan del estado P pasivo a un estado A activo mediante un método de activación por voz como se ejemplifica adelante:

Cada unidad 2 delegada en un estado P pasivo solicita una confirmación de canal de contribución cuando se cumple una primera condición opcional y una segunda condición de activación:

La primera condición de activación es una condición de ruido direccional: el nivel de señal de audio de entrada de la unidad 2 delegada de solicitud posible está bien por encima del nivel de audio acoplado estimado, es decir, está bien por encima del nivel de señal de audio estimada que resulta de recibir una señal de voz de un orador que utiliza otra unidad 2 delegada.

La segunda condición de activador es una condición de ruido difuso: El nivel de señal de audio de entrada de unidad 2 delegada de solicitud posible está bien por encima de un nivel de referencia (por ejemplo el nivel de ruido de fondo de piso).

Los datos, que también se denominan como metadatos de audio, requeridos como información externa para que cada unidad 2 delegada determine las dos condiciones anteriores es (1) una tabla ALT de todas las unidades 2 delegadas en los estados activos que sean ID de identificación únicos y su nivel de señal de audio durante un período de prueba T y (2) el nivel N de referencia (ruido de fondo). La tabla ALT puede por ejemplo tener la estructura:

nivel p2  $X_{p2}(T)$

nivel p(n-1)  $X_{p(n-1)}(T)$

nivel PN  $X_{pn}(T)$

Nivel de ruido N

por lo que p2, p(n-1), pn representan el ID y nivel  $X_{p\#}$  el nivel de señal de audio durante el período de prueba T. El nivel N de ruido se explicará posteriormente.

En una implementación posible, no limitante el nivel es un valor dentro de un rango [0, 1] descrito por un entero sin signo de 16 bits. Para calcular el nivel de se determinan los niveles de audio durante un bloque de muestras, por ejemplo, durante las muestras 1024. Para cada sub-bloque de 32 muestras se calcula la raíz cuadrada y el resultado se coloca en un filtro promedio exponencial. Para el nivel de ruido se calcula el nivel utilizando un algoritmo (por ejemplo, densidad de ruido espectral) sobre el audio de piso (que es una mezcla de todos los canales de contribución).

Los metadatos audio se recolectan y distribuyen mediante el módulo 3 de servicio central. Desde un punto de vista práctico es suficiente distribuir sólo periódicamente los metadatos de audio, por ejemplo, todas las muestras 1024 para ahorrar ancho de banda de comunicación. Los metadatos de audio se pueden distribuir eficientemente al utilizar métodos de distribución de multidifusión o trasmisión.

Cuando se reciba una solicitud de canal de contribución, el módulo 3 de servicio central otorga la asignación del canal de contribución si uno disponible. Si se ocupan todos los canales de contribución responden con una respuesta de denegación. Si la unidad 2 delegada en el estado A activo no cumple más ninguna de las condiciones, se solicita

liberación del canal de contribución después de un período de tiempo de espera. El periodo de espera evita que se solicite una liberación debido a una pausa pequeña en la voz. Una confirmación de canal de contribución o liberación siempre se tiene que solicitar desde el módulo 3 de servicio central, debido a que pueden existir requerimientos que resultarían en una denegación. Por ejemplo, el requisito de que por lo menos una unidad 2 delegada siempre debe permanecer activa.

5

La primera condición de activador puede ser vista en la siguiente ecuación:

$$X_i(k) > \gamma_{dir,p} \max_{P \in P(\kappa B)} \{W_{p,i}(\kappa B) X_{p,max}(\kappa B)\}$$

con:

k: Tiempo discreto

10 B: Bloque con una longitud de bloque de una pluralidad de muestras, por ejemplo 1024 muestras, que definen el tiempo o período de prueba;

K: Índice de franja de tiempo discreta para períodos de bloque B.

$X_{p,max}(\kappa B)$ : El nivel de señal de audio máximo de la unidad delegada p para los últimos pocos, por ejemplo por lo menos 3 a 10 periodos de tiempo antes del periodo k

15  $X_i(k)$ : El nivel de audio de la unidad i delegada de solicitud posible durante el periodo k;

$\gamma_{dir}$ : Factor de umbral para esta condición;

P: La recolección de unidades 2 delegadas en el estado A activo durante el periodo k;

$W_{p,i}(\kappa B)$ : El factor de acoplamiento acústico entre la unidad p delegada y la unidad i delegada durante el periodo k.

20 Por lo tanto la primera condición de activador prueba si el nivel de señal de audio de la unidad delegada i como la unidad delegada de solicitud posible es mayor que un valor de prueba de referencia de cada una de las unidades p delegadas activas multiplicada con el factor umbral. La función max sirve como una preselección, ya que extrae el mayor valor de prueba de referencia. El valor de prueba de referencia es por lo tanto el producto del nivel de señal de audio máxima de la unidad p delegada durante el periodo k y el factor de acoplamiento entre la unidad p delegada de prueba y la unidad i delegada de solicitud posible durante el periodo k.

25 El factor de acoplamiento acústico individual  $W_{p,i}(\kappa B)$  describe una relación entre el nivel  $X_i$  de señal de audio de la unidad i delegada de solicitud posible y el nivel  $X_p$  de señal de audio de la unidad p delegada de prueba en caso en que se genere una señal de voz desde un orador utilizando la unidad p delegada de prueba. De esta manera el factor de acoplamiento acústico individual puede ser diferente de cada uno para cada unidad p delegada de prueba. La primera condición de activador se cumplirá si se proporciona la señal de voz por un parlante que está enfrente del micrófono 4 de la unidad i delegada de solicitud posible y no enfrente de la unidad p delegada de prueba.

30

En una implementación posible, la longitud "B" de bloque es el intervalo de muestra 1024. "k" es el tiempo discreto, dependiente de la frecuencia de muestreo. Por lo menos la primera, preferiblemente ambas condiciones de activador son evaluadas actualmente todos los períodos de muestra: primero se actualiza el nivel X de audio utilizando la última muestra preferiblemente usando promedio exponencial, luego se hace la comparación. De lo contrario, por ejemplo, un retardo en la muestra 1024, en el peor de los casos provocaría que el sistema posiblemente pierda las primeras letras de la oración de los oradores. Los valores de prueba  $W_{p,i}(\kappa B)$   $X_{p,max}(\kappa B)$  son el valor (N) de ruido y sólo se actualizan cuando se reciben los nuevos metadatos de audio (lo que ocurre cada bloque B).

35

Los factores  $W_{p,i}(\kappa B)$  de acoplamiento acústico individuales se estiman y se alcanzan utilizando un algoritmo de mínimos cuadrados normalizado estándar. Su objetivo es filtrar rápidamente la convergencia de coeficientes para minimizar el error (= nivel de error). De nuevo se utiliza el periodo KB para descripción.

40

En una primer etapa, un nivel  $R_{p,i}$  residual, se determina a partir de la unidad p delegada a la unidad i delegada, con lo cual la unidad p delegada es solo la unidad 2 delegada en el estado A activo. En la situación en donde sólo una única unidad p delegada se activa, todas las otras unidades 2 delegadas ajustan dinámicamente su estimación de factor de acoplamiento acústico a la unidad delegada activa p, utilizando el nivel de señal de audio de su entrada 4 de micrófono y el nivel de señal de audio de la unidad p delegada activa simple, con lo cual el nivel de señal de audio de la unidad p delegada activa simple se distribuye a todas las unidades 2 delegadas por el módulo 3 de servicio central.

45

$$R_{p,i}(\kappa B) = X_i(\kappa B) - W_{p,i}(\kappa B) X_{p,max}(\kappa B)$$

En la siguiente etapa, se actualiza el factor de acoplamiento acústico:

$$W_{p,i}([\kappa + 1]B) = W_{p,i}(\kappa B) + \mu \frac{R_{p,i}(\kappa B) X_{max,p}(\kappa B)}{\max(\varepsilon\{[X_{p,max}(\kappa B)]^2\}, \varepsilon\{[R_{p,i}(\kappa B)]^2\}, thr)}$$

con:

$W_{p,i}(\kappa B)$ : el factor de acoplamiento acústico actualizado de la unidad p delegada a la unidad i delegada.

$\mu$ : la constante de tiempo de índice de convergencia

5  $\varepsilon\{\}$ : una función de promedio exponencial;

thr: un umbral "inferior" para evitar picos durante inicialización.

Como un valor de partida, todos los factores W de acoplamiento acústicos se pueden fijar al valor 1.0=0 dB.

Una posible implementación para la función promedio exponencial se define como:

$$X_i(k) = \beta X_i(k-1) + (1-\beta)|x_i(k)|$$

10

En donde el factor de suavizamiento beta se determina utilizando:

$$\beta = \exp\left[\frac{-1}{T_{exp} F_s}\right]$$

con:

$T_{exp}$ : la constante de tiempo exponencial

15  $F_s$ : la frecuencia de muestreo.

Se pueden utilizar otras implementaciones conocidas.

La función promedio (móvil) exponencial se describió anteriormente, pero para las funciones  $\varepsilon\{\}$ - (que determinan el nivel de potencia promedio) se realiza en el cuadrado de la entrada y el índice de actualización para cada período de bloque es KB. Refiérase a  $\varepsilon\{X_{p,max}(kB)^2\}$  como  $P_{xx,p}(KB)$ . Entonces:

20

$$P_{xx,p}(KB) = \beta P_{xx,p}([\kappa - 1]B) + (1-\beta)X_{p,max}^2(\kappa B)$$

Como el ingreso para esta función es el nivel máximo del pasado para los bloques del ejemplo 5, y este nivel se determina por la función de promedio exponencial del audio, esto parece un trabajo doble, pero este suavizamiento se prefiere para el algoritmo NLMS con el fin de que converja rápidamente. El otro valor  $P_{rr,p,i}(KB) = \varepsilon\{R_{p,i}(kB)^2\}$  es ventajoso para que reaccione con las interrupciones externas del algoritmo, por ejemplo: en un sistema en donde sólo una unidad delegada es activa, el factor de acoplamiento a esa unidad delegada se actualiza. Si el orador detrás de una unidad delegada no activa empezaría a hablar, esa unidad delegada enviaría una solicitud. Sin embargo, tomaría decenas de milisegundos al sistema otorgar esta solicitud: mientras que la voz del orador puede provocar que el factor de acoplamiento se actualice utilizando una entrada incorrecta. Sin embargo, debido al error grande/señal residual, el  $P_{rr,p,i}(KB)$  se incrementará rápidamente, evitando la actualización rápida de los factores de acoplamiento.

25

30

Como resultado, cada unidad 2 delegada conserva una tabla que contiene las estimaciones de factores de acoplamiento acústico para cada unidad 2 delegada. Las tablas de factor de acoplamiento acústico se indican en la figura 1 con ACFi, ACF1, ACF2...ACFn.

La segunda condición de activador, es decir, la condición de ruido difuso puede verse en la siguiente ecuación:

35

$$X_i(k) > \gamma_{dif} N(\kappa B)$$

con:

$\gamma_{dif}$ : el factor umbral para esta condición

N: el nivel de referencia (ruido de fondo) (del ALT).

Debido a que el sistema 1 de conferencia es un sistema distribuido, se pueden presentar retardos y latencias en la comunicación, lo que se puede manejar como sigue.

5 Antes que se otorgue a la unidad 2 delegada del orador un canal de contribución, otras unidades 2 delegadas también solicitan una confirmación del canal debido al acoplamiento acústico. Por lo tanto, sólo se otorga la primera solicitud de confirmación, en la que después se deniegan todas las solicitudes por un determinado tiempo (denominado "tiempo muerto"). Este tiempo muerto debe ser suficientemente largo de tal manera que los metadatos distribuidos contienen información sobre la unidad 2 delegada del orador.

Para evitar una avalancha, de repetición de solicitudes, una unidad 2 delegada debe esperar un período de tiempo antes de enviar una nueva solicitud, después que se ha negado la solicitud anterior.

10 Debido a que los metadatos se envían sólo una vez cada x muestras, la última información conocida puede quedarse atrás. En ese caso, no se puede evitar que un inicio de voz del orador active una confirmación en una o más unidades 2 delegadas (esto ocurre más frecuentemente cuando las estimaciones de acoplamiento han convergido a su valor final). Para resolver esto, una unidad 2 delegada espera por una actualización de metadatos directamente después que se ha otorgado su solicitud de confirmación de canal: si es claro de los nuevos metadatos que una solicitud de confirmación se activa mediante acoplamiento acústico, la unidad 2 delegada solicita inmediatamente una liberación de canal (es decir, sin tiempo de espera).

Detección/identificación de voz:

20 En el sistema 1 de conferencia una unidad 2 delegada puede aún solicitar y recibir un canal de contribución debido a una perturbación (clic de esfera, tos, etc.). Esto es aceptable, debido a que el canal se libera rápidamente. La unidad 2 delegada comprende un dispositivo 6 de indicación, que indica con una luz o LED el estado activo o pasivo de la unidad 2 delegada. Sin embargo: para una discusión se preferiría indicar solo un orador real para el público al activar el dispositivo de indicación en la unidad 2 delegada. En una posible realización se propone separar la indicación de la asignación de canal.

25 El dispositivo 6 de indicación en la unidad 2 delegada con el canal de contribución asignada, Es decir, en el estado A activa, se activa tan pronto como se detecta la voz (tono) en su señal de audio. Para limitar los requerimientos de procesamiento, será suficiente realizar la detección de voz (tono) solo en el canal de contribución más alto. También se prefiere realizar la detección de voz (tono) en el módulo 3 de servicio central, para reducir los requerimientos de hardware en las unidades delegada.

30 Posibles ventajas del sistema 1 de debate son que el acoplamiento acústico entre las unidades 2 delegadas se determina dinámicamente. Requiere que se distribuya información limitada, con la cual las unidades 2 delegadas pueden determinar si tienen un orador real como entrada de audio. Este puede manejar retardos en la comunicación en intercambio de información. Estas mejoras permiten al sistema 1 de conferencia ser flexible y escalable. El sistema 1 de conferencia, especialmente la activación por voz, es muy robusto, en razón a que las unidades 2 delegadas vecinas no activaran o solo se activaran muy poco debido al acoplamiento acústico cuando el orador empieza a hablar.

35 El sistema 1 de conferencia, especialmente las unidades 2 delegadas, son de autoaprendizaje, con lo cual después de un corto período, es fácil para otros oradores participar en el debate, incluso en dispositivos vecinos. El sistema 1 de conferencia es escalable, en razón a que el sistema 1 de conferencia funciona en configuraciones pequeñas y muy grandes, sin exigir configuración manual. La sobrecarga de comunicación es baja, debido al uso de metadatos periódicos, por ejemplo, distribuido utilizando multidifusión o transmisión. En resumen, el sistema 1 de conferencia mejora la robustez y flexibilidad al determinar el acoplamiento acústico y, opcionalmente las condiciones de habla en

40 las unidades 2 delegadas.



## REIVINDICACIONES

1. Sistema (1) de conferencia que comprende una pluralidad unidades (2) delegadas, cada unidad (2) delegada tiene un micrófono (5) para recibir una señal de audio de un entorno,
- 5 un módulo (3) de servicio central que maneja una pluralidad de canales de contribución, con lo cual la salida de audio de los canales de contribución contribuye a una salida de audio amplificada del sistema (1) de conferencia, con lo cual cada unidad (2) delegada se adapta para enviar una solicitud para una confirmación de canal de contribución al módulo (3) de servicio central, con lo cual el módulo (2) de servicio central se adapta para otorgar la solicitud y asignar un canal de contribución a la unidad (2) delegada de solicitud, configurando de esta manera la
  - 10 unidad (i) delegada de solicitud en un estado (A) activo, caracterizado porque la unidad (2) delegada se adapta para activar la solicitud mediante activación por voz, con lo cual la solicitud se activa en caso de que se cumpla por lo menos una primera condición de activador que define que el nivel de señal de audio de una de las unidades (2) delegadas, como una unidad (i) delegada de solicitud posible, es mayor que un valor de prueba individual para cada unidad (2) delegada en el estado (A) activo, con lo cual el nivel de señal de audio es el
    - 15 nivel de señal de audio recibida por el micrófono de la unidad (i) delegada de solicitud posible, con lo cual para cada unidad delegada activa el valor de prueba individual es un nivel de señal de audio estimado para la unidad (i) delegada de solicitud posible que resulta de una señal de voz o audio proporcionada a las otras unidades (2) delegadas activas.
  - 20 2. Sistema (1) de conferencia de acuerdo con la reivindicación 1, caracterizado porque el valor de prueba individual para una unidad (p) delegada de prueba se deriva al multiplicar un factor de acoplamiento acústico individual entre la unidad (i) delegada de solicitud posible y la unidad (p) delegada de prueba con el nivel de señal de audio de la unidad (p) delegada de prueba durante el período de prueba y, opcionalmente, con un factor umbral.
  - 25 3. Sistema (1) de conferencia de acuerdo con la reivindicación 1, caracterizado porque el valor de prueba individual para una unidad (p) delegada de prueba se deriva al multiplicar un factor de acoplamiento acústico individual entre la unidad (i) delegada de solicitud posible y la unidad (p) delegada de prueba con el máximo del nivel de señal de audio de la unidad (p) delegada de prueba durante los últimos pocos períodos de prueba y opcionalmente con un factor umbral.
  - 30 4. Sistema (1) de conferencia de acuerdo con una de las reivindicaciones 2 o 3 precedentes, caracterizado porque cada unidad (2) delegada comprende una tabla (ACF) de factores que contiene ID individuales de las otras unidades delegadas y los factores de acoplamiento individuales y una tabla (ALT) de nivel de señal de audio que contiene ID individuales de las otras unidades (2) delegadas activas y el nivel de señal de audio durante el período de prueba.
  - 35 5. Sistema (1) de conferencia de acuerdo con la reivindicación 4, caracterizado porque la tabla (ACF) de factores se gestiona por y/o presenta en las unidades (2) delegadas y que se proporciona la tabla (ALT) de señal de audio por el módulo (3) de servicio central.
  - 40 6. Sistema (1) de conferencia de acuerdo con la reivindicación 2 a 5, caracterizado porque la unidad (2) delegada se adapta para estimar los factores de acoplamiento acústicos individuales para cada una de las unidades (2) delegadas en una forma iterativa, con lo cual en cada etapa de iteración se mejora un valor de partida del factor de acoplamiento acústico individual.
  - 45 7. Sistema (1) de conferencia de acuerdo con una de las reivindicaciones 2 a 6 precedentes, caracterizado porque la unidad (2) delegada, que solicita un canal de contribución sobre la base de los datos de un primer período (KB) de prueba y que se dedica a un canal de contribución al asignar el canal de contribución a la unidad (2) delegada se adapta para revisar la solicitud y, de esta manera, la asignación al volver a probar por lo menos la primera condición sobre la base de los datos de un segundo período de prueba ((k+1) B).
  - 50 8. Sistema (1) de conferencia de acuerdo con una de las reivindicaciones 2 a 7 precedentes, caracterizada porque la unidad (i) delegada se adapta para activar la solicitud en caso de que por lo menos una primera condición de activador y una segunda condición de activador se cumpla, con lo cual la segunda condición de activador reclama que el nivel de señal de audio de la unidad (p) delegada de solicitud posible sea mayor que un nivel (N) de ruido de referencia durante el período de prueba.
  - 55 9. Sistema (1) de conferencia de acuerdo con una de las reivindicaciones 2 a 8 precedentes, caracterizado porque el módulo (3) de servicio central se adapta para otorgar sólo una solicitud durante el tiempo muerto preseleccionado.
  - 60 10. Sistema (1) de conferencia de acuerdo con una cualquiera de las reivindicaciones precedentes, caracterizado porque la unidad (2) delegada comprende un dispositivo (6) de indicación de orador para indicar el estado un orador

de la unidad (2) delegada, con lo cual el dispositivo (6) de indicación de orador se activa en caso de que se cumple una primera condición de indicación, que reclama que la unidad (2) delegada está en un estado activo y bajo una segunda condición de indicación que reclama que se detecta una voz, especialmente un tono de voz.

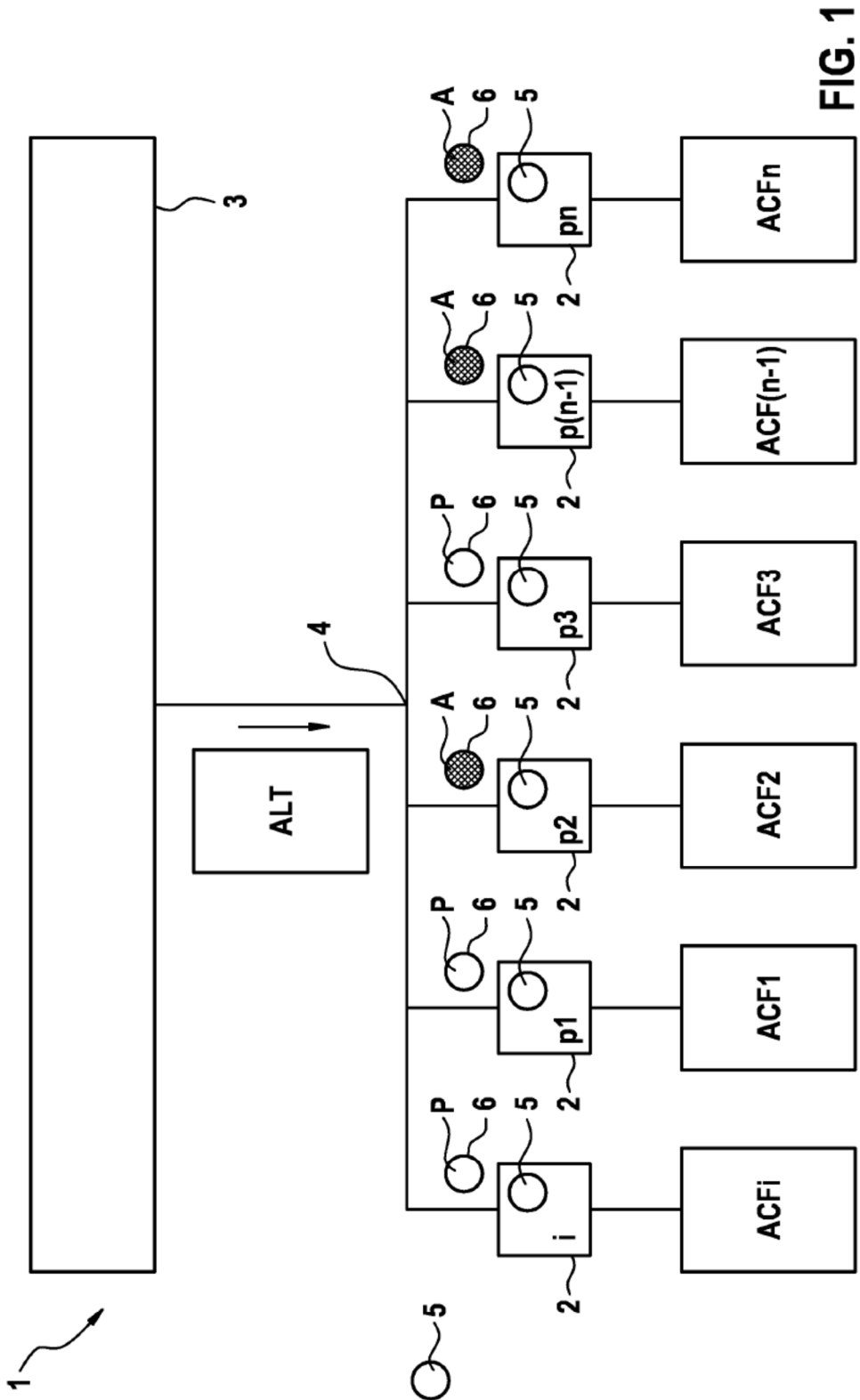


FIG. 1