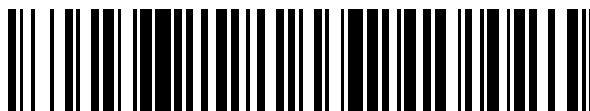


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 732 373**

51 Int. Cl.:

H04R 29/00 (2006.01)

G10L 21/02 (2013.01)

G08B 29/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.05.2011 PCT/EP2011/057622**

87 Fecha y número de publicación internacional: **15.11.2012 WO12152323**

96 Fecha de presentación y número de la solicitud europea: **11.05.2011 E 11721020 (3)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019 EP 2708040**

54 Título: **Sistema y método para emitir y controlar especialmente una señal de audio en un entorno usando una medida de inteligibilidad objetivo**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
22.11.2019

73 Titular/es:

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

72 Inventor/es:

**VAN DER SCHAAR, HANS;
HAN, OOSTEROM;
HEUSDENS, RICHARD y
HENDRIKS, RICHARD**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 732 373 T3

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

DESCRIPCIÓN

Sistema y método para emitir y controlar especialmente una señal de audio en un entorno usando una medida de inteligibilidad objetivo.

5 La invención se refiere a un sistema y a un método para emitir una señal de audio en un entorno. Más específicamente la invención se refiere a un sistema para emitir una señal de audio en un entorno, comprendiendo el sistema: una fuente de audio para proporcionar la señal de audio, al menos un altavoz para emitir la señal de audio, y al menos un micrófono para recibir una señal acústica desde el entorno, mediante el cual la señal acústica está basada en la señal de audio y puede comprender componentes molestos. La invención también se refiere a un método que usa el sistema.

10 Estado de la técnica

15 Los sistemas de megafonía u otros sistemas para emitir señales de audio, como música, voz o anuncios, en diferentes localizaciones como supermercados, escuelas, universidades, auditorios son ampliamente conocidos. Estos sistemas normalmente comprenden una fuente de audio, por ejemplo un micrófono o un grabador, y una pluralidad de altavoces, que están distribuidos de manera local en las localizaciones, para emitir la señal de audio desde la fuente de audio.

20 En realizaciones sencillas, estos sistemas tienen una amplificación ajustable, de modo que el volumen de la señal de audio emitida por los altavoces puede ajustarse a un valor deseado. En sistemas más sofisticados, la amplificación se hace dependiente del ruido y otros componentes molestos en las localizaciones. En algunos de estos sistemas se calcula una relación de señal a ruido (SNR), que a menudo se determina como el cociente: (salida amplificada)/(señal ambiente detectada - salida amplificada), mediante el cual la señal ambiente detectada puede detectarse por un micrófono en las localizaciones. Un enfoque de este tipo se desvela, por ejemplo, en el documento US 5.434.922 A en la conexión de una radio para un automóvil.

El documento EP 1 808 853 A1, desvela un sistema de megafonía que compara una señal de audio deseada con una señal de audio molesta y calcula un factor de amplificación para amplificar la señal de audio.

25 El documento DE 10 2007 031 064 A1 desvela un dispositivo de emergencia para un sistema de advertencia de emergencia electro-acústico, que tiene un micrófono montado en altavoces para medir presión acústica de altavoces individuales, mediante el cual el altavoz operando la señal analógica se compara con la señal radiada para evaluar el nivel de ruido en el entorno. Adicionalmente el documento desvela que las señales de audio pueden repetirse hasta que el nivel de ruido sea lo suficientemente bajo.

30 El documento US 2009 012 794A1 desvela un sistema para la mejora de la inteligibilidad de altavoces que se dirigen a un público objetivo. El sistema está adaptado para proporcionar realimentación de inteligibilidad a un altavoz, hablando para un público, y comprende un primer micrófono en el lado del altavoz y un segundo micrófono en el lado del público. Ambos micrófonos están conectados a medios de procesamiento que están dispuestos para calcular un valor de inteligibilidad basándose en ambas señales de los micrófonos. Los medios de señalización, preferentemente en el lado del público, están dispuestos para generar una señal de realimentación de inteligibilidad dependiendo del valor de inteligibilidad calculado. Los medios de señalización están dispuestos para generar dicha señal de realimentación de inteligibilidad en una forma óptica, visible para el altavoz referido.

40 El documento US 2007 147 625 A1 desvela un sistema y método para detectar y remediar niveles inaceptables de inteligibilidad de voz. El sistema evalúa audio de prueba recibido transmitido a través de, y recibido en, un espacio o región de interés. La inteligibilidad se mejora modificando la tasa, tono, amplitud y energía de bandas de frecuencia durante la presentación de la señal de voz.

45 El documento US 2005 135 637 A1, que representa la técnica anterior más cercana, desvela un sistema de medición y método que combinan un sistema de anuncio de audio con una pluralidad de sensores espaciados para evaluar la inteligibilidad de audio emitido desde altavoces del sistema de anuncio de audio. El procesamiento puede tener lugar en alguno o todos los sensores así como en un elemento de control común. Las evaluaciones pueden basarse en el uso de un método de índice de inteligibilidad de voz apropiado. El método de índice de inteligibilidad de voz usa una señal de audio artificial para evaluar la inteligibilidad.

Descripción de la Invención

50 La invención propone un sistema con las características de la reivindicación 1 y un método con las características de la reivindicación 12. Se desvelan realizaciones ventajosas o preferidas de la invención mediante las reivindicaciones dependientes, la descripción y/o las figuras según se adjuntan.

De acuerdo con la invención se desvela un sistema para emitir una señal de audio en un entorno, especialmente en un entorno acústico. El sistema puede realizarse como una pequeña escala, por ejemplo un sistema de mano como un teléfono móvil, un asistente digital personal (pda) un ordenador de tableta, etc. Puede realizarse como un sistema a media escala o privado como un coche o un aparato estéreo doméstico, conjunto de televisión, etc.
5 preferentemente, el sistema es un sistema a gran escala o público como un sistema de megafonía, etc.

Por consiguiente, el entorno puede ser - por ejemplo - el adyacente o cercano rodeando para el sistema a pequeña escala, una sala o el espacio interior de un vehículo para el sistema a media escala. En caso del sistema a gran escala también es posible que el sistema proporcione la señal de audio para una sala de conferencias o salón de conferencias como el entorno o para una pluralidad de salas como una pluralidad de entornos.

10 La señal de audio se realiza como una señal que lleva información dirigida a personas que permanecen en el entorno o que usan el entorno. La información llevada por la señal de audio es específicamente información hablada y se realiza, por ejemplo, como un anuncio, un mensaje o como un discurso. En otra realización de la invención la información llevada por la señal de audio es música o una combinación de música e información hablada.

15 La fuente de audio puede realizarse como una unidad de generación de señal de audio, por ejemplo un micrófono, especialmente un transductor, o como una unidad de reproducción de señal de audio, por ejemplo un grabador o un ordenador, que emite señales de audio habladas informáticas. Opcionalmente la fuente de audio está acoplada a un amplificador y/o una unidad de atenuación para amplificar o atenuar la señal de audio.

20 El sistema comprende adicionalmente al menos un altavoz, que emite la señal de audio en el entorno. En caso de los sistemas a pequeña escala, únicamente el altavoz o la disposición de altavoces pueden estar presentes, en caso de los sistemas a media escala, una pluralidad de altavoces pueden estar distribuidos en la sala o espacio interior. En caso de los sistemas a gran escala, al menos un altavoz está dispuesto en cada sala, que se proporciona por el sistema con la señal de audio, de modo que el sistema puede comprender una pluralidad de altavoces, que están distribuidos localmente.

25 Al menos se proporciona un micrófono para recibir una señal acústica del entorno. El micrófono puede realizarse como cualquier clase de un transductor, que convierte la señal acústica en una señal eléctrica. La señal acústica está basada en la señal de audio, especialmente comprende la señal de audio o al menos partes o fragmentos de la señal de audio. Los componentes molestos de la señal acústica están basados en ecos, errores de transmisión, reverberaciones y/o ruido en el entorno o son resultantes del mismo.

30 De acuerdo con la invención, el sistema comprende un módulo de análisis, que está adaptado o es operable para analizar la señal acústica. Durante la etapa de análisis, se realiza una medida de inteligibilidad objetivo, como resultado de la etapa de análisis o del método de medida de inteligibilidad objetivo se deriva o calcula o estima una medida de inteligibilidad. La medida de inteligibilidad se define como una característica de cómo de comprensible es la información, especialmente la voz o el anuncio, insertado por la señal de audio en la señal acústica.

35 La medida de inteligibilidad es preferentemente un valor, especialmente un valor dependiente del tiempo o una pluralidad de valores, por ejemplo un vector o matriz de valores, especialmente una pluralidad de valores dependientes del tiempo. Una pluralidad de valores, por ejemplo, es ventajoso en caso de que una pluralidad de entornos diferentes, por ejemplo salas, debieran controlarse de manera independiente o separada unos de los otros, de modo que se proporciona un valor para cada entorno. También es posible que la medida de inteligibilidad sea dependiente de la frecuencia, de modo que se proporciona una pluralidad de valores para una señal acústica de una
40 localización, mediante la cual la pluralidad de valores de inteligibilidad hace referencia a frecuencias diferentes o bandas de frecuencia diferentes de la señal acústica. La medida de inteligibilidad puede derivarse, por ejemplo, por uno de los siguientes métodos de medida de inteligibilidad objetivo:

- AI Índice artificial,
- SII Índice de inteligibilidad de voz (ANSI S3.5-1997)
- STI Índice de transmisión de voz
- SSNR SNR segmentaria
- LLR Relación de probabilidad logarítmica
- IS Itakura-Saito
- CEP Medida de distancia cepstral
- WSS Métrica de pendiente espectral ponderada
- FWS SSNR de frecuencia normalizada y ponderada
- PESO PESO
- DAU Modelo auditivo dau
- CSII SII de coherencia
- CSTI STI basada en covarianza
- STOI Medida de inteligibilidad objetivo de tiempo breve

Las referencias para los métodos de medida de inteligibilidad objetivo anteriormente mencionados pueden hallarse en el artículo científico de Cees Taal, Richard Hendriks, Richard Heusdens, Jesper Jensen: *Intelligibility Prediction of Single-Channel Noise-Reduced Speech*; in ITG-Fachtagung Sprachkommunikation. 06. - 08.10.2010 en Bochum, Alemania (ISBN 978-3-8007-3300-2).

- 5 La medida de inteligibilidad se usa como una señal de realimentación en el sistema. Como se explica continuación, la señal de realimentación puede acoplarse por ejemplo de vuelta al sistema para mejorar o controlar la inteligibilidad de la señal acústica o para establecer un protocolo de la medida de inteligibilidad, por ejemplo, como una prueba o una tabla de correspondencia o para iniciar otras reacciones de los sistemas como repetir la señal de audio para mejorar la inteligibilidad. Adicionalmente o como alternativa la señal de realimentación puede acoplarse de vuelta en una unidad de indicación del sistema, que indica a un operador de llamada o un altavoz que la señal de audio se emitió, por ejemplo, con una mala inteligibilidad.

- 15 El sistema de acuerdo con la invención muestra diversas ventajas: la configuración del sistema es fácil, puesto que un ajuste o intervalo de la medida de inteligibilidad deseada son casi suficientes. La medida de inteligibilidad como una señal de realimentación es un valor expresivo y una medida directa para el rendimiento del sistema, puesto que es en general el objetivo principal de un sistema para emitir una señal de audio en un entorno que la señal de audio sea inteligible y no por ejemplo si la relación de señal a ruido se mantiene o no a un cierto nivel.

- 20 En una realización preferida de la invención, el módulo de análisis o el mismo sistema funcionan en tiempo real, de modo que la señal de realimentación también está acoplada de vuelta en tiempo real. El tiempo real en la conexión del sistema significa que la medida de inteligibilidad se proporciona con un pequeño retardo, por ejemplo, menor que 2 s, preferentemente menor que 1 s y especialmente menor que 0,5 s. Esta realización tiene la ventaja, de que una reacción del sistema o del operador de llamada o del altavoz puede proporcionarse también de manera oportuna o también en tiempo real. Esta realización es la base, por ejemplo, para un sistema, que adapta la señal de audio en tiempo real en dependencia de la medida de inteligibilidad.

- 25 La aplicación principal del sistema puede hallarse en la transmisión de información hablada, como un anuncio, un mensaje o una voz, etc. Por lo tanto se prefiere que la medida de inteligibilidad sea una medida para la inteligibilidad de voz de la señal acústica. Se han enumerado anteriormente diversas posibilidades para derivar la medida de inteligibilidad, especialmente la medida de inteligibilidad de voz. En realizaciones alternativas, el sistema puede proporcionar una medida de inteligibilidad para música, de modo que el sistema se preocupa acerca de la inteligibilidad de música, por ejemplo en una sala de conciertos o en un coche.

- 30 De acuerdo con la invención, el módulo de análisis es operable para comparar la señal de audio como una señal limpia con la señal acústica como una señal con ruido para derivar la medida de inteligibilidad de la señal acústica. Para mejorar el resultado, se prefiere que las dos señales estén alineadas en el tiempo antes de la comparación.

- 35 En una realización práctica, la medida de inteligibilidad objetivo está basada en la STOI - Medida de Inteligibilidad Objetivo de Breve Tiempo como se desvela por ejemplo en el artículo científico Cees H. Taal, Richard C. Hendriks, Richard Heusdens, Jesper Jensen: *a short-time objective intelligibility measure for time-frequency weighted noisy speech*; in International Conference on Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE, ISBN: 978-1-4244-4295-9. Especialmente, la medida de inteligibilidad objetivo está basada en la comparación de la distribución de frecuencia de la señal de audio de tiempo alineado y la señal acústica durante un breve periodo de tiempo, por ejemplo menor que 1 s, especialmente menor que 0,5 s.

- 40 En una realización preferida, el sistema comprende un control automático de volumen con un bucle de control, que está adaptado para controlar el volumen (o energía) de la señal de audio emitida por el al menos un altavoz, mediante la cual la medida de inteligibilidad se usa como la señal de realimentación en el bucle de control. En esta realización se proporciona un control automático de volumen basado en medida de inteligibilidad. El volumen puede controlarse usando una ganancia o un factor de amplificación como una variable de accionamiento. El bucle de control puede realizarse, por ejemplo, como un control de bucle cerrado, pero también son posibles otras estrategias de control como lógica difusa, etc. La ventaja de esta realización es que, el sistema mantendrá la inteligibilidad, especialmente la inteligibilidad de voz de la señal acústica de acuerdo con un punto de ajuste o intervalo predefinidos, y por lo tanto asegura que todas las señales acústicas son inteligibles. Especialmente en caso de usar el módulo de análisis en un modo en tiempo real, el sistema puede reaccionar instantáneamente por ejemplo, en elevaciones del ruido de fondo, sin desestabilizar el sistema.

- 55 En un desarrollo de la invención, el módulo de análisis es operable para proporcionar la medida de inteligibilidad para al menos dos o una pluralidad de bandas de frecuencia de la señal acústica, mediante la cual se calcula un valor de inteligibilidad para cada una de las bandas de frecuencia. Adicionalmente el control automático de volumen usa los al menos dos valores de inteligibilidad para controlar los volúmenes de las bandas de frecuencia de la señal de audio de manera separada y/o independientemente entre sí. Este desarrollo permite que el sistema adapte el volumen en diferentes bandas de frecuencia de manera separada para compensar las fuentes de ruido en ciertos

intervalos de frecuencia.

5 En una posible realización de este desarrollo, el control automático de volumen está adaptado para mantener constante la energía o volumen global en el entorno de la señal de audio emitida o dentro de un intervalo predefinido. En esta realización, el sistema permite mantener la energía global o volumen constante mientras mantiene una inteligibilidad predefinida. Por ejemplo en caso de que la inteligibilidad de una primera banda de frecuencia sea alta y la inteligibilidad de una segunda banda de frecuencia sea baja, se reduce el volumen de la primera banda de frecuencia y se aumenta el volumen de la segunda banda de frecuencia, de modo que la inteligibilidad de todas las bandas de frecuencia es suficiente o por encima de un nivel predefinido y el volumen global se mantiene constante o al menos se mantiene dentro de intervalos deseados o predefinidos.

10 De acuerdo con la invención el sistema comprende un módulo de repetición, que está adaptado para repetir la misma señal de audio u otra, sustituyendo la señal de audio en caso de que la medida de inteligibilidad sea peor que un valor o umbral predefinido. En este caso la señal de realimentación se usa como una base para una decisión de si debe emitirse o no la señal de audio un tiempo adicional.

15 En una posible realización adicional, el sistema puede comprender un módulo de protocolo, que es operable para establecer un protocolo de la medida de inteligibilidad de la señal acústica. En esta realización la señal de realimentación se usa para establecer un protocolo de si la señal de audio/acústica era inteligible para las personas en el entorno. El protocolo derivado del módulo de protocolo puede mantener metadatos acerca de la señal de audio, tiempo de difusión o emisión de la señal de audio, la localización de la difusión o emisión de la señal de audio en el entorno y la medida de inteligibilidad. Este protocolo puede usarse beneficiosamente como una prueba o una evidencia de que se emitió una cierta señal de audio de manera inteligible en una cierta área.

20 En una realización adicional más de la invención, se proporciona un módulo de información, que está adaptado para informar a un usuario del sistema de la medida de inteligibilidad o una representativa o una equivalente de la misma. El módulo de información puede comprender, por ejemplo, indicadores visuales como semáforos, que indican si una señal de audio recién emitida era o no inteligible. En caso de que la señal de audio no se emitiera de manera inteligible, el usuario tiene la posibilidad de reaccionar y - por ejemplo - puede repetir la señal de audio. En caso de que el módulo de información indique que la señal de audio se emitía de manera inteligible, el usuario recibirá una confirmación positiva.

En una realización práctica el sistema se realiza como un sistema de megafonía o como un sistema de refuerzo de sonido que comprende una pluralidad de altavoces como se ha descrito anteriormente.

30 En una posible realización, el sistema, especialmente el sistema de megafonía, comprende una unidad de altavoz con un transductor o un micrófono e indicadores visuales que indican si una señal de audio recién emitida era o no inteligible.

35 Una materia objeto adicional de la invención es un método para controlar, corregir y/o indicar la medida de inteligibilidad de una señal de audio generada por el sistema como se ha descrito anteriormente y/o de acuerdo con una de las reivindicaciones anteriores, mediante el cual la medida de inteligibilidad se usa como una señal de realimentación en el sistema.

Efectos, características y ventajas adicionales se harán evidentes mediante la descripción de realizaciones preferidas de la invención y las figuras según se adjunta. Las figuras muestran:

40 Figura 1 un diagrama de bloques de un sistema para emitir una señal de audio en un entorno como una realización de la invención;

Figura 2 un diagrama de bloques del módulo de control del sistema en la Figura 1;

Figura 3 un diagrama de bloques del módulo de control de la Figura 2 en otra realización.

45 La Figura 1 es un diagrama de bloques que ilustra un sistema 1 para emitir una señal 2 de audio amplificada en un entorno 3. El sistema 1 comprende al menos un altavoz 4 para emitir la señal 2 de audio amplificada en el entorno 3 acústico y al menos un micrófono 5 para recibir una señal 6 acústica de dicho entorno 3 acústico. La señal 6 acústica comprende partes de la señal 2 de audio emitida y adicionalmente componentes molestos del entorno 3 como reverberaciones de eco y ruido 7 adicional, que pueden resultar del entorno 3 o del sistema 1 mismo como ruido de amplificador, etc. El sistema 1 comprende adicionalmente o está acoplado a un medio de generación de señal de audio (no mostrado) por ejemplo un grabador o un micrófono para un altavoz, que genera la señal 8 de audio no amplificada u original. La señal 8 de audio se amplifica por un amplificador 9.

En esta realización, el sistema 1 se realiza como un sistema de megafonía o un sistema de refuerzo de sonido, que podría comprender una pluralidad de altavoces 4 y también una pluralidad de micrófonos 5. Un sistema de megafonía puede usarse en escuelas, supermercados u otros lugares, mediante el cual se forman una pluralidad de entornos 3 acústicos en los que está dispuesto al menos un altavoz 4 y un micrófono 5. Un entorno 3 acústico puede realizarse como una sala, por ejemplo una sala de clase.

Como se indica en la Figura 1, la señal 6 acústica (convertida en una señal eléctrica) se guía en un módulo 10 de control, que se explicará en relación con la figura 2. Adicionalmente la señal 8 de audio original se guía en el módulo 10 de control. Como una salida, el módulo 10 de control comprende una ruta de señal 11 de ganancia al amplificador 9, de modo que el módulo 10 de control es operable para controlar la ganancia del amplificador 9 y por lo tanto el volumen de la señal 2 de audio amplificada.

La Figura 2 ilustra los componentes del módulo 10 de control, que muestra dos entradas para recibir la señal 8 de audio y la señal 6 acústica y una salida para enviar la señal de ganancia 11 al amplificador 9. En una primera etapa, la señal 8 de audio se retarda por una unidad 12 de retardo para estar alineada en el tiempo con la señal 6 acústica. El retardo de tiempo entre la señal 8 de audio y la señal 6 acústica resulta de longitudes diferentes de las rutas de señal y puede eliminarse o compensarse según se describe o por otra manera. Las dos señales 6 y 8 se transforman a un módulo 13 de análisis, que está adaptado para analizar las dos señales 6 y 8 y para proporcionar una medida de inteligibilidad de una medida de inteligibilidad objetivo.

El método de medida de inteligibilidad objetivo usado en el módulo 13 de análisis preferentemente muestra una baja complejidad con alta correlación a la inteligibilidad de voz subjetiva de la señal 6 acústica.

Ejemplo:

El método propuesto como un ejemplo es una función de la voz limpia y procesada, indicada por x e y, respectivamente, que corresponde a la señal 8 de audio y a la señal 6 acústica. El modelo está diseñado para una tasa de muestreo de 10000 Hz, para cubrir el intervalo de frecuencia relevante para inteligibilidad de voz. Cualesquiera señales a otras tasas de muestra deberían re-muestrearse. Adicionalmente, se supone que la señal limpia y la procesada están ambas alineadas en el tiempo, por ejemplo por la unidad 12 de retardo. En primer lugar, se obtiene una representación TF (Frecuencia de Tiempo) segmentando ambas señales al 50 % de solapamiento, tramas en ventanas de Hanning con una longitud de 256 muestras, donde cada trama se rellena a cero hasta 512 muestras y se transforma por Fourier. A continuación, se realiza un análisis de la banda de un tercio de octava agrupando binarios de DFT. En total se usan 15 bandas de un tercio de octava, donde la frecuencia central más inferior se establece igual a 150 Hz. Sea $\hat{x}(k,m)$ que indica el binario de DFT de orden k de la trama de orden m de la voz limpia. La norma de la banda de un tercio de octava de orden j, denominada como una unidad de TF, se define a continuación como,

$$X_j(m) = \sqrt{\sum_{k=k_1(j)}^{k_2(j)-1} |\hat{x}(k,m)|^2}$$

donde k_1 y k_2 indican los bordes de banda de un tercio de octava, que se redondean al binario de DFT más cercano. La representación de TF de la voz procesada se obtiene de manera similar, y se indicará por $Y_j(m)$. La medida de inteligibilidad intermedia para una unidad de TF, es decir $d_j(m)$, depende de una región de N unidades de TF consecutivas desde ambas de

$X_j(n)$ y $Y_j(n)$, donde $n \in M$ y $M = \{(m-N+1), (m-N+2) \dots m-1, m\}$. En primer lugar, se aplica un procedimiento de normalización local, escalando todas las unidades de TF de $Y_j(n)$ con un factor

$$\alpha = \left(\sum_n X_j(n)^2 / \sum_n Y_j(n)^2 \right)^{1/2}$$

de manera que su energía equivale a la energía de voz limpia, dentro de esa región de TF. A continuación, $\alpha Y_j(n)$ se recorta al límite inferior de la relación de señal a distorsión (SDR), que definimos como,

$$SDR_j(n) = 10 \log_{10} \left(\frac{X_j(n)^2}{(\alpha Y_j(n) - X_j(n))^2} \right)$$

Por lo tanto

$$Y' = \max(\min(\alpha Y, X + 10^{-\beta/20} X), X - 10^{-\beta/20} X),$$

- 5 donde Y' representa la unidad de TF normalizada y recortada y β indica el límite de SDR inferior. Lo índices de trama y banda de un tercio de octava se omiten por conveniencia de notación. La medida de inteligibilidad intermedia se define como una estimación del coeficiente de correlación lineal entre las unidades de TF procesadas limpia y modificada,

$$d_j(m) = \frac{\sum_n \left(X_j(n) - \frac{1}{N} \sum_l X_j(l) \right) \left(Y'_j(n) - \frac{1}{N} \sum_l Y'_j(l) \right)}{\sqrt{\sum_n \left(X_j(n) - \frac{1}{N} \sum_l X_j(l) \right)^2 \sum_n \left(Y'_j(n) - \frac{1}{N} \sum_l Y'_j(l) \right)^2}}$$

- 10 donde $l \in M$. Finalmente, la OIM eventual simplemente se proporciona por el promedio de la medida de inteligibilidad intermedia sobre todas las bandas y tramas,

$$d = \frac{1}{JM} \sum_{j,m} d_j(m),$$

- 15 donde M representa el número total de tramas y J el número de bandas de un tercio de octava. La correlación máxima se obtiene con $\beta = 15$ y $N=30$, que significa que la medida intermedia depende de información de voz desde los últimos 384 ms. El retardo para proporcionar la medida de inteligibilidad es aproximadamente 400 ms y por lo tanto se proporciona en tiempo real.

- 20 La OIM como un ejemplo de una medida de inteligibilidad o un valor similar de otro método de medida de inteligibilidad objetivo se transfiere a un control 14 automático de volumen como una señal de realimentación, que compara la medida de inteligibilidad a ciertos umbrales para determinar si la ganancia del amplificador 9 ha de aumentarse, reducirse o mantenerse constante para mantener una medida de inteligibilidad predefinida. La ganancia está delimitada de manera superior e inferior a ciertos niveles predeterminados. El módulo 10 de control o el control 14 automático de volumen pueden detectar silencios en la voz de la señal 8 de audio. Durante breves pausas la ganancia se paraliza y durante pausas largas, después de que el eco se ha extinguido, el nivel de ruido se detecta directamente y este se traduce en una ganancia adecuada, para cuando el sistema 1 reinicia la transmisión de un mensaje.

- 25 Las ventajas principales, que pueden alcanzarse con la invención son como sigue: en primer lugar su simplicidad, no ha de completarse configuración extensiva en la instalación, un ajuste sencillo de la inteligibilidad deseada o intervalo o medida de inteligibilidad y el retardo acústico inicial para el micrófono 5 lo harán. Debido a que la acústica de la sala no tiene que modelarse en este sistema 1 es adecuado para cualquier espacio. La complejidad computacional también se reduce drásticamente si se elige el método de medida de inteligibilidad objetivo correcto. Este sistema 1 puede reaccionar instantáneamente en elevaciones en el ruido de fondo, sin desestabilizar el sistema. Pero la ventaja principal es que hay una realimentación directa al sistema 1 y opcionalmente al operador de llamada sobre la inteligibilidad del mensaje transportado. Si la inteligibilidad (medida) es baja, la ganancia tiene que aumentarse. Los sistemas conocidos generalmente se adaptan a la relación de señal a ruido medida, sin embargo, esto no es siempre una buena medida de la inteligibilidad de un mensaje. Asegurándose que el mensaje es inteligible es en general el objetivo global de un sistema de megafonía y no de si la relación de señal a ruido se mantiene a un cierto nivel.

La Figura 3 ilustra detalles adicionales del módulo 10 de control en la Figura 2. En la Figura 3, la medida de inteligibilidad se acopla de vuelta en un módulo 15 de procesamiento. Se proporciona el módulo 15 de procesamiento y opcionalmente el control 14 automático de volumen.

5 De acuerdo con la invención, el módulo 15 de procesamiento se realiza como un módulo de repetición, que está adaptado para repetir la señal 2 de audio en caso de que la medida de inteligibilidad como una señal de realimentación sea peor que un valor o umbral predefinido. Esta realización puede usarse en caso de que el sistema 1 proporcione anuncios o mensajes en el entorno 3 acústico. En caso de que el anuncio no fuera inteligible, el anuncio se repite automáticamente o se proporciona otro anuncio sustitutivo.

10 Por ejemplo, la inteligibilidad medida se analiza en un número de tramas durante un mensaje o anuncio. Si se clasifican demasiadas tramas consecutivas, o demasiadas tramas de media como que son ininteligibles o tienen inteligibilidad baja, el módulo de repetición podría proporcionar una advertencia al sistema 1 y opcionalmente al operador de llamada de que el mensaje o anuncio puede no ser inteligible para todos los oyentes y que el mensaje debería repetirse.

15 En una segunda realización, no cubierta por las reivindicaciones, el módulo 15 de procesamiento se realiza como un módulo de protocolo, que usa la medida de inteligibilidad como una señal de realimentación para establecer el protocolo de la inteligibilidad de las señales 8 de audio emitidas. En algunas aplicaciones es importante conocer si un anuncio era o no inteligible. Para tener una prueba para la inteligibilidad, el módulo de protocolo proporciona un registro diario como es conocido, por ejemplo a partir de máquinas de facsímil.

20 En una tercera realización, no cubierta mediante las reivindicaciones, el módulo 15 de procesamiento se realiza como un módulo de información, que está adaptado para informar a un usuario del sistema acerca de la inteligibilidad o no inteligibilidad de la señal acústica. Es posible, por ejemplo, que los medios de generación de señal de audio sean un micrófono y la información para el usuario se alimente en una lámpara de indicación, como un semáforo, que está mecánicamente acoplada o adyacente al micrófono, permitiendo una realimentación en tiempo real al usuario, de si el anuncio o voz eran o no inteligibles.

25 Deberá indicarse que la primera realización puede realizarse junto con cualquiera de la segunda y la tercera realización, o incluso ambas, en un sistema 1 como una realización adicional de la invención.

En una realización sencilla de la invención, la medida de inteligibilidad es un valor o un escalar. En realizaciones más sofisticadas, la medida de inteligibilidad puede realizarse como un vector o una matriz multidimensional.

30 Por ejemplo es posible, que una pluralidad de entornos 3 acústicos se controlen u observen, de modo que la medida de inteligibilidad es un vector, mediante el cual cada entrada del vector se asigna a un único entorno 3 acústico. Los entornos 3 acústicos pueden hacer referencia a áreas separadas, por ejemplo salas. Como alternativa, los entornos 3 acústicos pueden hacer referencia a un área común, por ejemplo una sala o salón de conferencias, mediante el cual el sistema 1 asegura que en cualquier lugar del área común se asegura la inteligibilidad.

35 También es posible, que el sistema 1 adapte el volumen en diferentes bandas de frecuencia de manera separada para compensar fuentes de ruido en ciertos intervalos de frecuencia de manera separada. En este caso la medida de inteligibilidad es un vector, mediante el cual cada entrada del vector se asigna a una banda de frecuencia de la señal 6 acústica o de la señal 8 de audio. Opcionalmente, el volumen general o global o nivel de energía del entorno acústico se mantiene inferior mientras mantiene la inteligibilidad. Esta alternativa podría también servir para aumentar adicionalmente la inteligibilidad si se ha alcanzado un nivel de ganancia máximo en otras bandas. Sin embargo, esto podría reducir la naturaleza del mensaje reproducido.

Adicionalmente es posible usar el sistema 1 para una pluralidad de entornos 3 acústicos, mediante los cuales se controlan de manera separada bandas de frecuencia separadas, de modo que la medida de inteligibilidad es una matriz.

45 Aunque la invención se ilustró por medio de ejemplo por un sistema de megafonía, la invención puede usarse también en otros sistemas de emisión de señal de audio como teléfonos móviles, aparatos estéreo de coche, conjuntos de televisión, etc.

REIVINDICACIONES

1. Sistema (1) para emitir una señal (2, 8) de audio en un entorno (3), comprendiendo el sistema (1):

una fuente de audio para proporcionar la señal (2, 8) de audio,

al menos un altavoz (4) para emitir la señal (2) de audio,

5 al menos un micrófono (5) para recibir una señal (6) acústica del entorno (3), mediante la cual la señal (6) acústica está basada en la señal (2) de audio y puede comprender componentes (7) molestos,

un módulo (13) de análisis para analizar la señal (6) acústica y para proporcionar una medida de inteligibilidad de un método de medida de inteligibilidad objetivo, la medida de inteligibilidad se usa como una señal de realimentación y mediante la cual el módulo (13) de análisis está adaptado para comparar la señal (2, 8) de audio

10 con la correspondiente señal (6) acústica para derivar la medida de inteligibilidad,

caracterizado por

que la señal de audio se realiza como una señal que lleva información, mediante la cual la información llevada por la señal de audio es una información hablada y/o música, y por que el sistema (1) comprende adicionalmente un módulo (15) de repetición, que está adaptado para repetir la señal (2, 8) de audio u otra señal de audio sustitutiva en caso de que la medida de inteligibilidad sea peor que un valor o umbral predefinido.

15

2. Sistema (1) de acuerdo con la reivindicación 1, **caracterizado por que** el módulo (13) de análisis está adaptado para analizar la señal (6) acústica con un retardo menor que 2 s, preferentemente menor que 1 s y especialmente menor que 0,5 s.

3. Sistema (1) de acuerdo con la reivindicación 1 o 2, **caracterizado por que** la medida de inteligibilidad es una característica para la inteligibilidad de voz de la señal (6) acústica o que la medida de inteligibilidad es una característica para la inteligibilidad de música de la señal (6) acústica.

20

4. Sistema (1) de acuerdo con la reivindicación 3, **caracterizado por que** el método de medida de inteligibilidad objetivo está basado en la comparación de la distribución de frecuencia de la señal (2, 8) de audio, especialmente de una señal (2, 8) de audio alineada en el tiempo, y la señal (6) acústica durante un breve periodo de tiempo, menor que 2 s, preferentemente menor que 1 s y especialmente menor que 0,5 s.

25

5. Sistema (1) de acuerdo con una de las reivindicaciones anteriores, **caracterizado por** un control (14) automático de volumen que comprende un bucle de control, que está adaptado para controlar el volumen o la energía de la señal (2) de audio emitida por el al menos un altavoz (4), mediante la cual la medida de inteligibilidad se usa como la señal de realimentación en el bucle de control.

6. Sistema (1) de acuerdo con la reivindicación 5, **caracterizado por que** el módulo (13) de análisis está adaptado para proporcionar la medida de inteligibilidad para al menos dos bandas de frecuencia diferentes de la señal (6) acústica y por que el control (14) automático de volumen está adaptado para controlar los volúmenes o energías de las bandas de frecuencia de la señal (2) de audio de manera separada.

30

7. Sistema (1) de acuerdo con la reivindicación 6, **caracterizado por que** el control (14) automático de volumen está adaptado para mantener la energía global de la señal (2) de audio en el entorno (3) constante o dentro de un intervalo dado.

35

8. Sistema (1) de acuerdo con una de las reivindicaciones anteriores, **caracterizado por** un módulo (15) de protocolo, que está adaptado para establecer el protocolo de la medida de inteligibilidad de la señal (6) acústica.

9. Sistema (1) de acuerdo con una de las reivindicaciones anteriores, **caracterizado por** un módulo (15) de información, que está adaptado para informar a un usuario del sistema (1) acerca de la medida de inteligibilidad o una representativa o una equivalente de la misma.

40

10. Sistema (1) de acuerdo con una de las reivindicaciones anteriores, caracterizado como un sistema de megafonía o como un sistema de refuerzo de sonido.

11. Sistema (1) de acuerdo con la reivindicación 10, **caracterizado por que** la fuente de audio comprende una unidad de altavoz con un transductor, especialmente un micrófono, y un indicador visual que indica la medida de inteligibilidad o una representativa o una equivalente de la misma.

45

12. Método para controlar, corregir y/o indicar la medida de inteligibilidad de una señal (2, 8) de audio generada por un sistema (1) de acuerdo con una de las reivindicaciones anteriores, **caracterizado por que** la medida de inteligibilidad se usa como una señal de realimentación en el sistema (1).

50

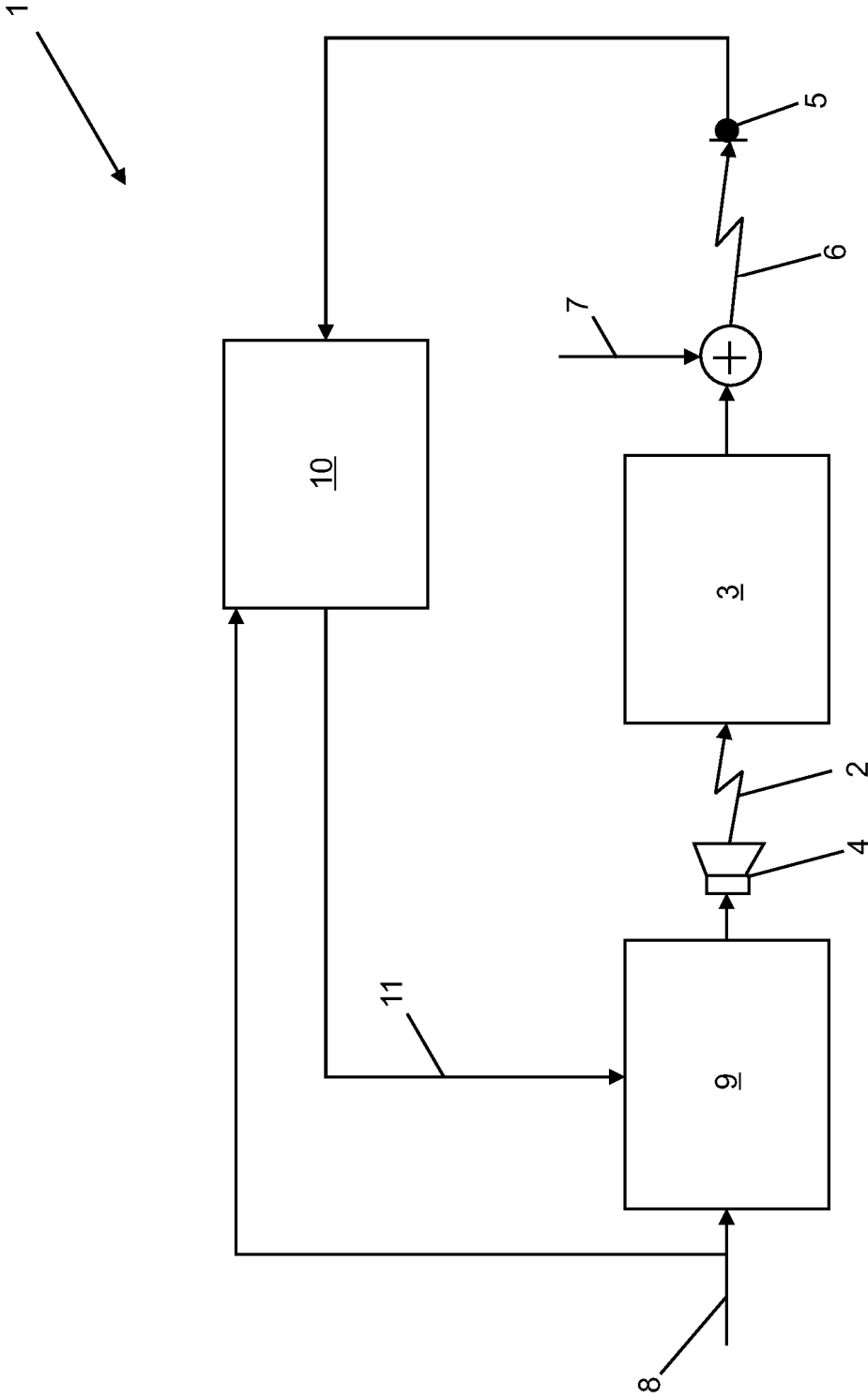


Fig. 1

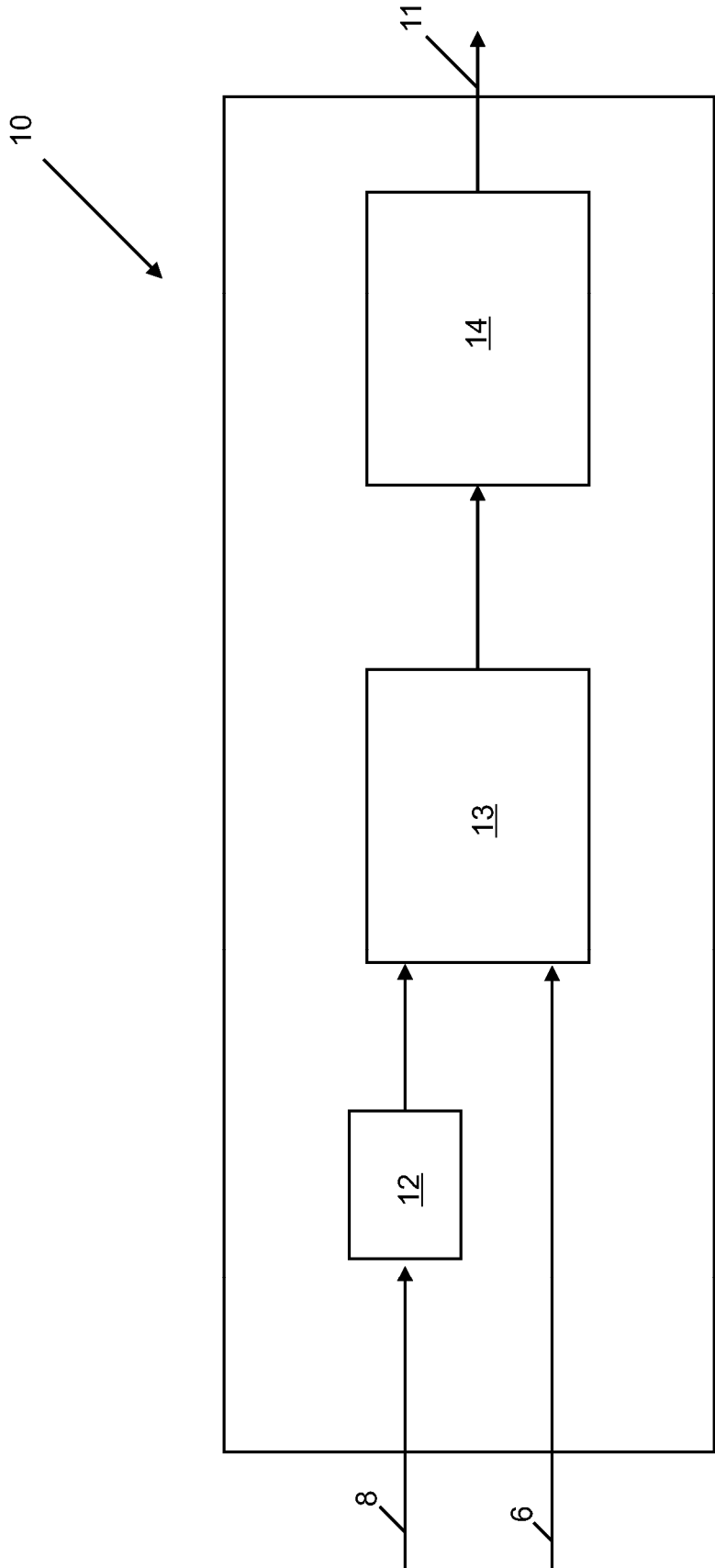


Fig. 2

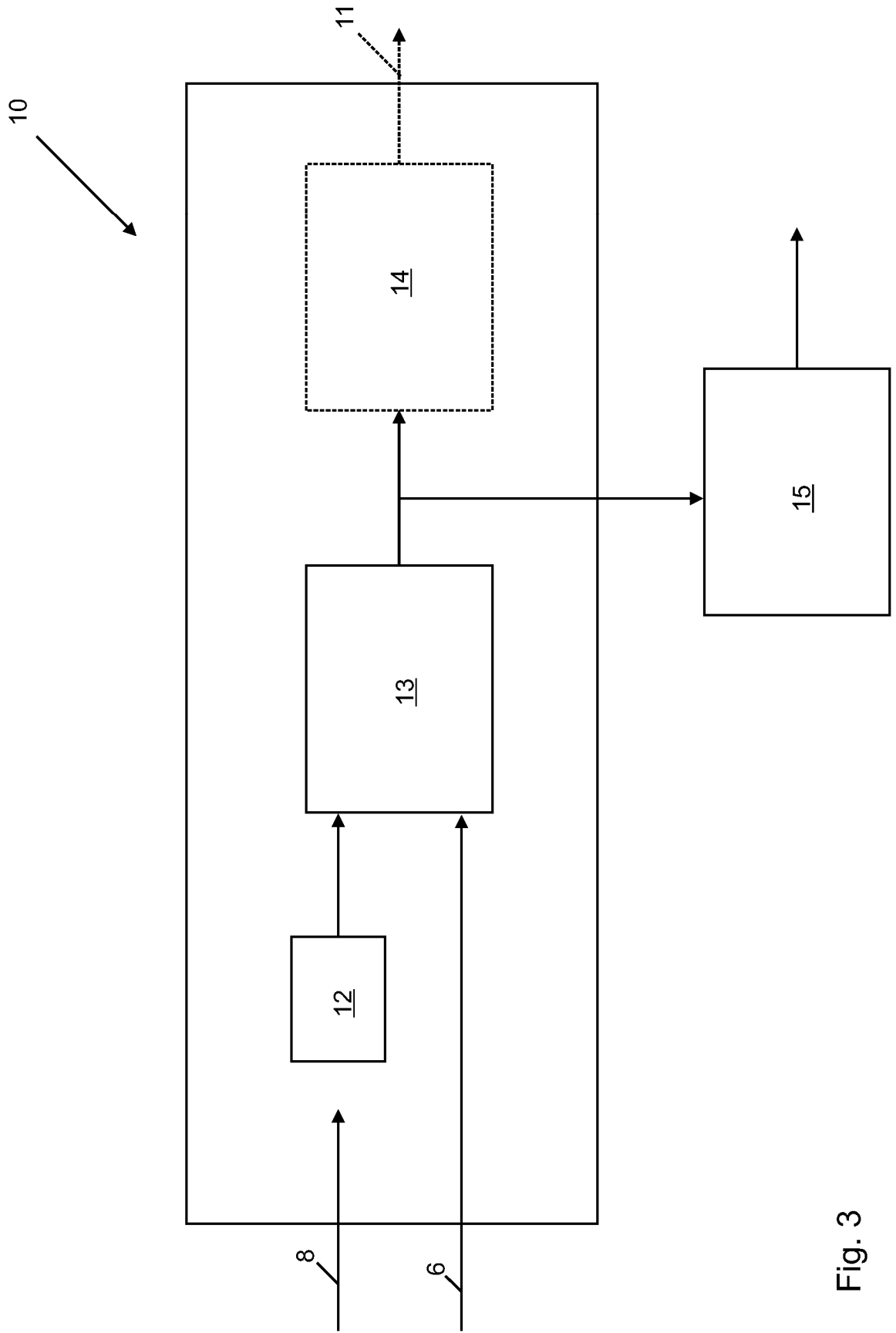


Fig. 3