(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 687 617**

(51) Int. Cl.:

G10L 19/012 (2013.01)
G10L 19/008 (2013.01)
G10L 25/03 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **14.02.2014** **E 17176159 (6)**

(97) Fecha y número de publicación de la concesión europea: **20.06.2018** **EP 3244404**

(54) Título: **Generación de ruido de confort**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
26.10.2018

(73) Titular/es:

TELEFONAKTIEBOLAGET LM ERICSSON (PUBL) (100.0%)
164 83 Stockholm, SE

(72) Inventor/es:

ERIKSSON, ANDERS

(74) Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 687 617 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

Generación de ruido de confort

Campo técnico

La solución descrita en el presente documento se refiere en general al procesamiento de señales de audio, y en particular a la generación de ruido de confort.

Antecedentes

Los productos de procesamiento de voz usan el ruido de confort, CN, para replicar el ruido de fondo con una señal generada artificialmente. Esto puede usarse, por ejemplo, en el control de eco residual en canceladores de eco usando un procesador no lineal, NLP, donde el NLP bloquea la señal contaminada por eco e inserta CN para no introducir un espectro perceptualmente molesto y una falta de coincidencia de nivel de la señal transmitida. Otra aplicación de CN es en codificación de voz en el contexto de supresión de silencio o transmisión discontinua, DTX, donde, para ahorrar ancho de banda, el transmisor solo envía una representación altamente comprimida de las características espectrales del ruido de fondo y se reproduce el ruido de fondo como un CN en el receptor.

Como el verdadero ruido de fondo está presente en períodos en los que la supresión de silencio/NLP o DTX no está activa, el CN debe hacer coincidir este ruido de fondo lo más fielmente posible. La coincidencia espectral se logra, por ejemplo, produciendo el CN como una señal de pseudorruido de forma espectral. El CN se genera con mayor frecuencia usando un filtro de ponderación espectral y una señal de pseudorruido de accionamiento. Esto puede realizarse en el dominio del tiempo, $n(t) = H(z)\, w(t)$, o en el dominio de la frecuencia, $n(t) = \mathrm{IFFT}\,(H(f)*W(f))$, donde $H(z)$ y $H(f)$ son la representación de la conformación espectral en el dominio del tiempo y de la frecuencia, respectivamente, y $w(t)$ y $W(f)$ son secuencias de ruido de accionamiento adecuadas, por ejemplo, una señal de pseudorruido. El documento de patente US 2013/0006622 propone una técnica para generar un ruido de confort en el contexto de una llamada de conferencia.

Sin embargo, cuando se aplica generación de ruido de confort a señales estéreo u otras señales de audio multicanal, el resultado a menudo no es satisfactorio. De hecho, los oyentes pueden experimentar efectos desagradables.

Sumario

Sería deseable lograr un ruido de confort de alta calidad para múltiples canales de audio. La solución divulgada en el presente documento se refiere a un procedimiento para generar ruido de confort, que replica las características espaciales del ruido de fondo además de las características espectrales comúnmente usadas.

De acuerdo con un primer aspecto, se proporciona un método, que debe ser realizado por un nodo de transmisión. El método comprende determinar las características espectrales de las señales de audio en al menos dos canales de audio de entrada. El método comprende además determinar una coherencia espacial entre las señales de audio en los canales de audio de entrada respectivos; y señalizar la información de señalización sobre las características espectrales de las señales de audio en al menos dos canales de audio de entrada e información sobre la coherencia espacial entre las señales de audio en los canales de audio de entrada, a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

De acuerdo con un segundo aspecto, se proporciona un método, que debe ser realizado por un nodo de recepción. El método comprende obtener información sobre características espectrales de señales de audio de entrada en al menos dos canales de audio. El método comprende además obtener información sobre una coherencia espacial entre las señales de audio de entrada en al menos dos canales de audio. El método comprende además generar ruido de confort para al menos dos canales de audio de salida, basándose en la información obtenida sobre características espectrales y coherencia espacial.

De acuerdo con un tercer aspecto, se proporciona un nodo de transmisión. El nodo de transmisión comprende medios de procesamiento, por ejemplo, en forma de un procesador y una memoria, en el que la memoria contiene instrucciones ejecutables por el procesador, por lo que el nodo de transmisión puede funcionar para realizar el método de acuerdo con el segundo aspecto. Es decir, el nodo de transmisión está operativo para determinar las características espectrales de las señales de audio en al menos dos canales de audio de entrada y para señalizar la información sobre las características espectrales de las señales de audio en al menos dichos dos canales de audio de entrada. La memoria contiene además instrucciones ejecutables por dicho procesador por lo que el nodo de transmisión es además operativo para determinar la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y para señalizar la información sobre la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

De acuerdo con un cuarto aspecto, se proporciona un nodo de recepción. El nodo de recepción comprende medios de procesamiento, por ejemplo, en forma de un procesador y una memoria, en el que la memoria contiene instrucciones ejecutables por el procesador, por lo que el nodo de transmisión puede funcionar para realizar el método de acuerdo con el tercer aspecto anterior. Es decir, el nodo de recepción está operativo para obtener características espectrales de señales de audio en al menos dos canales de audio de entrada. El nodo de recepción es además operativo para obtener una coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y para generar ruido de confort, para al menos dos canales de audio de salida, basándose en la información obtenida sobre las características espectrales y la coherencia espacial.

La solución de acuerdo con los aspectos descritos anteriormente permite la generación de ruido de confort de alta calidad para múltiples canales.

Breve descripción de los dibujos

Los anteriores y otros objetos, características y ventajas de la solución divulgada en el presente documento serán evidentes a partir de la siguiente descripción más particular de las realizaciones tal como se ilustra en los dibujos adjuntos. Los dibujos no son necesariamente a escala, sino que se hace hincapié en ilustrar los principios de la solución divulgada en el presente documento.

La figura 1 es un diagrama de flujo de un método realizado por una disposición, de acuerdo con una realización de ejemplo.

La figura 2 es un diagrama de flujo de un método realizado por una disposición y/o un nodo de transmisión, de acuerdo con una realización de ejemplo.

La figura 3 es un diagrama de flujo de un método realizado por una disposición y/o un nodo de recepción, de acuerdo con una realización de ejemplo.

La figura 4 es un diagrama de flujo de un método realizado por un nodo de transmisión, de acuerdo con una realización de ejemplo.

La figura 5 es un diagrama de flujo de un método realizado por una disposición y/o un nodo de recepción, de acuerdo con una realización de ejemplo.

Las figuras 6 y 7 ilustran las disposiciones de acuerdo con las realizaciones de ejemplo.

Las figuras 8 y 9 ilustran los nodos de transmisión de acuerdo con realizaciones de ejemplo.

Las figuras 10 y 11 ilustran los nodos receptores de acuerdo con las realizaciones de ejemplo.

Descripción detallada

Una forma sencilla de generar ruido de confort, CN, para múltiples canales, por ejemplo, estéreo, es generar CN basándose en uno de los canales de audio. Es decir, deriva las características espectrales de la señal de audio en dicho canal y controla un filtro espectral para formar el CN a partir de una señal de pseudorruido que se emite en múltiples canales, es decir, aplica el CN desde un canal a todos los canales de audio. Sin embargo, si busca un ruido estéreo más realista, otra forma directa es derivar las características espectrales de las señales de audio en todos los canales y usar múltiples filtros espectrales y múltiples señales de pseudorruido, una para cada canal, generando así tantos CN como canales de salida. Sin embargo, aunque podría esperarse que el último método replicara el ruido de fondo en estéreo con un buen resultado, este no es siempre el caso. Los oyentes que están sujetos a este tipo de CN a menudo experimentan que hay algo extraño o molesto con el sonido. Por ejemplo, los oyentes pueden tener la experiencia de que la fuente de ruido está ubicada dentro de su cabeza, lo que puede ser muy desagradable.

El inventor se ha dado cuenta de este problema y ha encontrado una solución, que se describe en detalle a continuación. El inventor se ha dado cuenta de que, para mejorar el CN multicanal, también deben tenerse en cuenta las características espaciales de las señales de audio en los múltiples canales de audio al generar el CN. Sin embargo, no es obvio cómo lograr esto. El inventor ha resuelto el problema encontrando una manera de determinar o estimar la coherencia espacial de las señales de audio de entrada, y después configurar la generación de señales CN de manera que estas señales CN tengan una coherencia espacial que coincida con la de las señales de audio de entrada. Cabe señalar que, incluso después de haber identificado que se podría usar la coherencia espacial, no es una tarea sencilla lograr esto. Para simplificar, la solución que se describe a continuación se describe para dos canales de audio, también denominados "izquierdo" y "derecho", o "x" e "y", es decir estéreo. Sin embargo, el concepto podría generalizarse a más de dos canales.

La coherencia espacial del ruido de fondo puede obtenerse usando la función de coherencia C(f) = |S_xy(f)|^2/(S_x(f)*S_y(f)) donde S_x(f) es el espectro medio de la señal del canal izquierdo, S_y(f) es el espectro medio de la señal del canal derecho, y S_xy(f) es el espectro cruzado de las señales del canal izquierdo y derecho. Estos espectros pueden, por ejemplo, ser estimados por medio del periodograma usando la transformada rápida de Fourier (FFT).

De manera similar, los filtros de conformación espectral CN se pueden obtener como una función de la raíz cuadrada de los espectros de señal S_x(f) y S_y(f). Otras tecnologías, por ejemplo, el modelado AR, también se pueden emplear para estimar los filtros de conformación espectral CN.

Se puede obtener un CN correlacionado espacial y espectralmente como

$$n_l(t) = ifft(H_1(f)*(W_1(f) + G(f)*W_2(f)))$$

$$n_r(t) = ifft(H_2(f)*(W_2(f) + G(f)*W_1(f)))$$

donde H_1(f) y H_2(f) son funciones de ponderación espectral obtenidas como función de los espectros de señal S_x(f) y S_y(f), G(f) es una función de la función de coherencia C(f), y W_1(f) y W_2(f) son componentes pseudoaleatorios de fase/ruido.

La estimación de las características de ruido de fondo espacial y espectral,

Cm(f): coherencia espacial

H_l(f): características espectrales del canal izquierdo (sqrt (S_l(f))

H_r(f): características espectrales del canal derecho (sqrt (S_r(f))

pueden obtenerse usando la transformada de Fourier de la señal del canal izquierdo, x, y derecho, y, durante los períodos de solo ruido, como se ejemplifica en el siguiente pseudocódigo:

```
X = fft(x, N_FFT);
M = abs(X(1:(N_FFT/2))).^2/2/L;
Sx = RHO*Sx + (1-RHO)*M;
M_1 = sqrt(min(Sx, 2*M));
H_1 = [M_1; M_1(end); flipud(M_1(2:end))];
Y = fft(y, N_FFT);
M = abs(Y(1:(N_FFT/2))).^2/2/L;
Sy = RHO*Sy + (1-RHO)*M;
M_r = sqrt(min(Sy, 2*M));
H_r = [M_r; M_r(end); flipud(M_r(2:end))];
crossCorr = RHO*crossCorr + (1-RHO)*x'*y)^2/
                    (x'*x) / (y'*y) ;
Sxy = RHO*Sxy + (1-RHO)*
(X(1:(N_FFT/2))).*conj(Y(1:(N_FFT/2)))/2/L;
C = (abs(Sxy).^2)./(eps+Sx.*Sy);
Cm = (31/32)*Cm + (1/32)*C;
```

El ruido de confort correlacionado espacial y espectralmente puede luego reproducirse usando la transformada de Fourier inversa de una suma de secuencias de ruido ponderado en frecuencia como se describe a continuación.

La representación espectral del ruido de confort puede formularse como, para los canales izquierdo y derecho, respectivamente:

$$n_l(f) = H_1(f)*(W_1(f) + G(f)*W_2(f))$$

$$N_r(f) = H_2(f)*(W_2(f) + G(f)*W_1(f))$$

donde W_1(f) y W_2(f) son preferiblemente secuencias de ruido aleatorio con una magnitud unida representada en el dominio de la frecuencia. Bajo la suposición de que W_1(f) y W_2(f) son secuencias pseudoblancas

independientes con magnitud unitaria, la función de coherencia de N_l(f) y N_r(f) es igual a (omitiendo el parámetro f)

$$C_N(f) = (|H_1|^2*|H_2|^2*|2*G|^2) / (|H_1|^2*|H-2|^2*(1 + G^2)^2 = 4 G^2 / (1 + G^2)^2$$

Por lo tanto, para obtener una coherencia espacial similar del ruido de confort a partir de la señal estéreo original, es decir, que C_N(f) = C(f); G(f) puede derivarse de la identidad C(f) = 4 G(f)^2 / (1 + G(f)^2)^2 como

$$G(f) = sqrt(2 – C(f) – sqrt((2 – C(f)^2 – C(f)))$$

La coincidencia espectral se obtiene al observar que el espectro de N_l(f) y N_r(f) debe ser igual a S_N_l(f) = |H_1(f)|^2*(1 + G(f)^2) y S_N_r(f) = |H_2(f)|^2*(1 + G(f)^2). A partir de esto, H_1(f) y H_2(f) se pueden elegir para que S_N_l f) y S_N_r(f) coincidan con el espectro del ruido de fondo original en el canal izquierdo y derecho, |H_l f)|^2 y |H_r(f)|^2, respectivamente, como

$$H_1(f) = H_l(f) / sqrt(1 + G(f)^2)$$

$$H_2(f) = H_r(f) / sqrt(1 + G(f)^2)$$

Con el fin de reducir la complejidad, se puede observar que la coherencia de las señales de ruido normalmente solo es significativa para las bajas frecuencias, por lo tanto, puede reducirse el intervalo de frecuencia para el que se realizarán los cálculos. Es decir, los cálculos pueden realizarse solo para un intervalo de frecuencia, por ejemplo, donde la coherencia espacial C(f) excede un umbral, por ejemplo, 0,2.

Un procedimiento simplificado puede usar solo la correlación del ruido de fondo en los canales izquierdo y derecho, g, en lugar de la función de coherencia C(f) anterior. La versión simplificada de solo usar la correlación del ruido de fondo del canal izquierdo y derecho puede implementarse reemplazando G(f) en la expresión para H_1(f) y H_2(f) con un escalar calculado similar a G(f) pero con el factor de correlación escalar en lugar de la función de coherencia C(f).

El procedimiento puede implementarse como se describe en el siguiente pseudocódigo:

```
 seed = exp(i*2*pi*rand(N_FFT/2-1, 1));
W_1 = [rand(1); seed; rand(1); conj(flipud(seed))];
 seed = exp(i*2*pi*rand(N_FFT/2-1, 1));
W_2 = [rand(1); seed; rand(1); conj(flipud(seed))];
 if (useCoherence)
   Gamma = (1 - 2./Cm);
   Gamma = -Gamma - sqrt(Gamma.^2 - Cm);
   Gamma = sqrt(Gamma);
   G = [Gamma; Gamma(end); flipud(Gamma(2:end))];
   CrossCorr(frame) = mean(Cm);
   H_1 = H_1./sqrt(1+G.^2);
   H_2 = H_r./sqrt(1+G.^2);
   N_1 = H_1.*(W_1 + G.*W_2);
   N_r = H_2.*(W_2 + G.*W_1);
 else
   if (useCorrelation)
      gamma = (1 - 2/crossCorr);
      gamma = -gamma - sqrt(gamma^2 - crossCorr);
      gamma = sqrt(gamma);
   else
      gamma = 0;
   end
   H_1 = H_l/sqrt(1+gamma^2);
   H_2 = H_r/sqrt(1+gamma^2);
   N_1 = H_1.*(W_1 + gamma*W_2);
   N_r = H_2.*(W_2 + gamma*W_1);
 end
 n_l = sqrt(N_FFT)*ifft(N_l);
 n_r = sqrt(N_FFT)*ifft(N_r);
 n_l = n_l(1:(L+N_overlap));
 n_r = n_r(1:(L+N_overlap));
 noise(ind, 1) =
[overlapWindow.*n_l(1:N_overlap)+overlap_l;
                        n_l((N_overlap+1):L)];
 overlap_l = flipud(overlapWindow).*n_l((L+1):end);
 noise(ind, 2) =
[overlapWindow.*n_r(1:N_overlap)+overlap_r;
                        n_r((N_overlap+1):L)];
 overlap_r = flipud(overlapWindow).*n-r((L+1):end);
```

En la descripción anterior, el ruido de confort se genera en el dominio de la frecuencia, pero el método puede implementarse usando representaciones de filtro de dominio del tiempo de los filtros de conformación espacial y espectral.

Para el control de eco residual, el ruido de confort resultante puede utilizarse en un NLP selectivo de dominio de la frecuencia que solo bloquea ciertas frecuencias, mediante una ponderación espectral posterior.

Para la aplicación de codificación de voz, pueden usarse varias tecnologías para que el generador de CN obtenga la ponderación espectral y espacial, y la invención se puede usar independientemente de estas tecnologías. Las posibles tecnologías incluyen, entre otras, por ejemplo, la transmisión de parámetros de AR que representan el ruido de fondo a intervalos de tiempo regulares o que estima continuamente el ruido de fondo durante la transmisión de voz regular. De manera similar, la coherencia espacial puede ser modelada usando, por ejemplo, una función sinc y transmitida a intervalos regulares, o continuamente estimada durante el habla.

En los párrafos siguientes, se describirán diferentes aspectos de la solución divulgada en el presente documento en más detalle con referencias a ciertas realizaciones y a los dibujos adjuntos. Con fines de explicación y no de limitación, se exponen detalles específicos, tales como escenarios y técnicas particulares, a fin de proporcionar una comprensión completa de las diferentes realizaciones. Sin embargo, otras realizaciones pueden apartarse de estos detalles específicos.

Método de ejemplo realizado por una disposición, figura 1

A continuación se describirá un método de ejemplo para la generación de CN realizado por una disposición en un dispositivo o sistema con referencia a la figura 1. Se debe asumir que la disposición tiene carácter técnico. El método es adecuado para la generación de ruido de confort para una pluralidad de canales de audio, es decir, al menos dos canales de audio. La disposición puede ser de diferentes tipos. Puede comprender un cancelador de eco ubicado en un nodo de red o un dispositivo, o puede comprender un nodo de transmisión y un nodo de recepción que funcionan para codificar y decodificar señales de audio, y aplicar supresión de silencio o un esquema DTX durante períodos de silencio relativo, por ejemplo, habla no activa.

La figura 1 ilustra el método que comprende determinar 101 las características espectrales de las señales de audio en al menos dos canales de audio de entrada. El método comprende además determinar 102 la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y generar 103 ruido de confort, para al menos dos canales de audio de salida, basándose en las características espectrales determinadas y la coherencia espacial.

Se supone que la disposición ha recibido la pluralidad de señales de audio de entrada en la pluralidad de canales de audio, por ejemplo, a través de uno o más micrófonos o de alguna fuente de audio multicanal, como el almacenamiento de archivos de audio. La señal de audio en cada canal de audio se analiza con respecto a sus contenidos de frecuencia, y las características espectrales, indicadas, por ejemplo, H_l(f) y H_r(f), se determinan de acuerdo con un método adecuado. Esto es lo que se ha hecho en los métodos de la técnica anterior para la generación de ruido de confort. Estas características espectrales también podrían denominarse características espectrales del canal, en el sentido de que un canal que tenga las características espectrales H_l(f) generaría la señal de audio l(t) a partir de, por ejemplo, ruido blanco. Es decir, las características espectrales se consideran como un filtro de conformación espectral. Debe observarse que estas características espectrales no comprenden ninguna información relacionada con ninguna correlación cruzada entre las señales o canales de audio de entrada.

Sin embargo, aquí, se determina otra característica más de las señales de audio, concretamente, una relación entre las señales de audio de entrada en forma de la coherencia espacial C entre las señales de audio de entrada. En general, el concepto de coherencia está relacionado con la estabilidad o predictibilidad de la fase. La coherencia espacial describe la correlación entre las señales en diferentes puntos del espacio, y a menudo se presenta como una función de la correlación frente a la distancia absoluta entre los puntos de observación.

En un ejemplo con dos señales de audio de entrada, l(t) y r(t), donde "l" representa "izquierda" y "r" significa "derecha", estas señales de audio se introducen en la disposición, por ejemplo, a través de un micrófono estéreo. Estas señales podrían alternativamente denominarse x(t) e y(t), que se usa en una parte previa de la descripción. La figura 2 es una ilustración esquemática de un proceso, que muestra tanto acciones como señales, donde las dos señales de entrada pueden verse como la señal 201 del canal izquierdo y la señal 202 del canal derecho. Las características espectrales del canal izquierdo, expresadas como H_l(f), se estiman en 203, y las características espectrales del canal derecho, H_r(f), se estiman 204. Esto podría, como se describió previamente, realizarse usando el análisis de Fourier de las señales de audio de entrada. Entonces, la coherencia espacial C_lr se estima 205 basándose en las señales de audio de entrada y posiblemente reutilizando los resultados de la estimación 203 y 204 de las características espectrales de las respectivas señales de audio de entrada.

La generación de ruido de confort se ilustra de forma de ejemplo en la figura 3, que muestra tanto acciones como señales. Una primera, W_1, y una segunda, W_2, secuencia de pseudorruido se generan en 301 y 302, respectivamente. Entonces, se genera 303 una señal de ruido del canal izquierdo basándose en las estimaciones de las características del espectro del canal izquierdo H_l y la coherencia espacial C_lr; y basándose en las secuencias W_1 y W_2 de pseudorruido generadas. Además, se genera 304 una señal de ruido del canal derecho basándose en las características espectrales del canal derecho estimadas H_l y coherencia espacial C_lr, y las secuencias W_1 y W_2de pseudorruido. Se han descrito anteriormente más detalles sobre cómo se hace esto y se describirán con más detalle a continuación.

Cuando la disposición es del tipo cancelador de eco, la determinación de información espectral y espacial y la generación de ruido de confort se realizan en la misma entidad, que podría ser un NLP. En ese caso, la información espectral y espacial no es señalizada necesariamente a otra entidad o nodo, sino que solo se procesa dentro del cancelador de eco. El cancelador de eco podría ser parte/estar ubicado en, por ejemplo, dispositivos, tales como teléfonos inteligentes; mezcladores y diferentes tipos de nodos de red.

Método de ejemplo realizado por un nodo de transmisión, figura 4

Un método de ejemplo, realizado por un nodo de transmisión, para soportar la generación de ruido de confort, se describirá a continuación con referencia a la figura 4. El nodo de transmisión, que alternativamente podría ser denominado, por ejemplo, nodo de codificación, se debe suponer que tiene carácter técnico. El método es adecuado para soportar la generación de ruido de confort para una pluralidad de canales de audio, es decir, al menos dos canales de audio. El nodo de transmisión puede funcionar para codificar señales de audio, y para aplicar supresión de silencio o un esquema DTX durante períodos de silencio relativo, por ejemplo, periodos de habla no activa. El nodo de transmisión puede ser un dispositivo inalámbrico y/o por cable, tal como un equipo de usuario, UE, una tableta, un ordenador o cualquier nodo de red que reciba u obtenga señales de audio para ser codificadas. El nodo de transmisión puede ser parte de la disposición descrita anteriormente.

La figura 4 ilustra el método que comprende determinar 401 las características espectrales de las señales de audio en al menos dos canales de audio de entrada. El método comprende además determinar 402 la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y señalizar 403 de información sobre las características espectrales de las señales de audio en al menos dos canales de audio de entrada e información

sobre la coherencia espacial entre las señales de audio en los canales de audio de entrada, a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

En un caso de ejemplo con dos señales de audio de entrada, es decir estéreo, el procedimiento para determinar las características espectrales y la coherencia espacial puede corresponder al ilustrado en la figura 2, que también se describe anteriormente.

La señalización de información sobre las características espectrales y la coherencia espacial puede comprender una transmisión explícita de estas características, por ejemplo, H_l, H_r y C_lr, o, puede comprender la transmisión o el transporte de alguna otra representación o indicación, implícita o explícita, a partir de la cual pueden derivarse las características espectrales de las señales de audio de entrada y la coherencia espacial entre las señales de audio de entrada.

La coherencia espacial puede determinarse aplicando una función de coherencia en una representación de las señales de audio en al menos dichos dos canales de audio de entrada. Por ejemplo, la coherencia espacial C_{xy} entre dos señales, x e y de las al menos dos señales de audio de entrada, podría determinarse como: $C_{xy} = |S_{xy}|^2/(S_{xx}^2 * S_{yy}^2)$; donde S_{xy} es la densidad espectral cruzada entre x e y, y S_{xx} y S_{yy} es la densidad autoespectral de x e y respectivamente.

En un ejemplo estéreo, cuando se denominan las señales de entrada "l" y "r", esto se denominaría como $C_lr = |S_{lr}|^2/(S_{ll}^2 * S_{rr}^2)$ o $C_lr = |S_{lr}|^2/(Sl^2 * Sr^2)$. Cabe señalar que $Sx \approx |H_x|^2$. Por lo tanto, al determinar las características espectrales H para cada señal de audio, o canal, y la coherencia espacial C entre los canales, estos parámetros deben señalizarse a un nodo de recepción. En el caso de aplicar la solución en un cancelador de eco, como se describió anteriormente, los parámetros determinados se usan para generar ruido de confort dentro de la misma entidad.

En una implementación simplificada, la coherencia C(f) podría estimarse, es decir, aproximarse, con la correlación cruzada de/entre las señales de audio en los respectivos canales de audio de entrada. Este sería un factor de correlación escalar, es decir, un valor constante, que podría derivarse integrando la función de coherencia C(f) a través de un intervalo de frecuencia. Esto aún daría un resultado mejor que cuando no se usa información de coherencia espacial.

Las señales de audio de entrada son señales de audio "reales", a partir de las cuales las características espectrales y la coherencia espacial podrían derivarse o determinarse de la manera descrita en el presente documento. Esta información se debe usar entonces para generar ruido de confort, es decir, una señal de ruido sintetizado que ha de imitar o reproducir el ruido de fondo en los canales de audio de entrada.

<u>Método de ejemplo realizado por un nodo de recepción, figura 5</u>

Un método de ejemplo, para generar ruido de confort, realizado por un nodo de recepción, por ejemplo dispositivo u otra entidad técnica, se describirá a continuación con referencia a la figura 5. Se supone que el nodo de recepción tiene carácter técnico. El método es adecuado para la generación de ruido de confort para una pluralidad de canales de audio, es decir, al menos dos canales de audio.

La figura 7 ilustra el método que comprende obtener información 501 sobre características espectrales de señales de audio de entrada en al menos dos canales de audio. El método comprende además obtener información 502 sobre la coherencia espacial entre las señales de audio de entrada en al menos dos canales de audio. El método comprende además generar ruido de confort para al menos dos canales de audio de salida, basándose en la información obtenida sobre características espectrales y coherencia espacial.

La obtención de información podría comprender recibir la información de un nodo de transmisión o determinar la información basándose en señales de audio, dependiendo del tipo de entidad al que se hace referencia, en términos de cancelador de eco o nodo de decodificación, que se describirá con más detalle a continuación. La información obtenida corresponde a la información determinada o estimada como se describe anteriormente junto con los métodos realizados por una disposición o por un nodo de transmisión. La información obtenida sobre las características espectrales y la coherencia espacial puede comprender los parámetros explícitos, por ejemplo, para estéreo: H_l, H_r y C_lr, o, puede comprender alguna otra representación o indicación, implícita o explícita, a partir de la cual podrían derivarse las características espectrales de las señales de audio de entrada y la coherencia espacial entre las señales de audio de entrada.

La generación de ruido de confort comprende generar señales de ruido de confort para cada uno de al menos dichos dos canales de audio de salida, donde el ruido de confort tiene características espectrales correspondientes a las señales de audio de entrada, y una coherencia espacial que corresponde a la de las señales de audio de entrada. Cómo se puede hacer esto en detalle se ha descrito anteriormente y se describirá más adelante.

La generación de una señal de ruido de confort N_1 para un canal de audio de salida puede comprender determinar una función de conformación espectral H_1, basándose en la información sobre características espectrales de una de las señales de audio de entrada y la coherencia espacial entre la señal de audio de entrada y al menos otra señal de audio de entrada. La generación puede comprender además aplicar la función de conformación espectral H_1 a una primera señal de ruido aleatorio W_1 y a una segunda señal de ruido aleatorio W_2(f), donde W_2(f) está ponderada G(f) basándose en la coherencia entre la señal de audio de entrada y al menos otra señal de audio de entrada.

En el ejemplo estéreo, la señal de ruido de confort N_l(f) para el canal de audio de salida izquierdo puede derivarse como N_l(f) = H_1(f)*(W_1(f) + G(f)*W_2(f)), donde G(f) se deriva como G(f) = sqrt(2 - C_lr(f) - sqrt((2 - C_lr(f))^2 - C_lr(f))), y H_1(f) se deriva como H_1(f) = H_l(f)/sqrt(1 + G(f)^2). Esto también se describe más arriba en esta descripción. Como se mencionó anteriormente y se ilustra, por ejemplo, en la figura 3, W_1(f) y W_2(f) indican señales de ruido aleatorio, que se generan como base para el ruido de confort. Las señales de ruido aleatorio se conforman en las respectivas señales de ruido de confort mediante el uso de funciones de conformación espectral o filtros y componentes que representan una contribución de coherencia espacial. Es decir, mirando el ejemplo para estéreo, N_l(f) = H_1(f)*(W_1(f) + G(f)*W_2(f)), por ejemplo G(f) W_2(f) está relacionado con la coherencia espacial.

Dado que el ruido de confort se genera para reproducir el ruido de fondo de las señales de audio de entrada, se desea que la coherencia espacial entre las señales de ruido de confort de salida sea lo más cercana posible a la coherencia espacial entre las señales de audio de entrada. Con las señales de entrada l y r, y las señales de salida n_l y n_r, esto corresponde al ajuste C_nlnr = C_lr.

Cuando el nodo de recepción se refiere al lado del decodificador de un códec, y podría ser denominado, por ejemplo, nodo de decodificación, la obtención de información comprende recibir la información de un nodo de transmisión como el descrito anteriormente. Este sería el caso, por ejemplo, cuando el audio codificado se transfiere entre dos dispositivos en un sistema de comunicación inalámbrica, por ejemplo, comunicación D2D (dispositivo a dispositivo) o comunicación celular a través de una estación base u otro punto de acceso. Durante los períodos de DTX, se puede generar ruido de confort en el nodo de recepción, en lugar de que el ruido de fondo en el nodo de transmisión sea codificado y transferido en su totalidad. Es decir, en este caso, la información se deriva o determina a partir de señales de audio de entrada en otro nodo, y luego es señalizada al nodo de recepción.

Por otro lado, si el nodo de recepción se refiere a un nodo que comprende un cancelador de eco, que obtiene la información y genera ruido de confort, la obtención de información comprende determinar la información basándose en señales de audio de entrada en al menos dos canales de audio. Es decir, la información no se deriva ni se determina en otro nodo y luego se transfiere desde el otro nodo, sino que se determina a partir de una representación de las señales de audio de entrada "reales". Las señales de audio de entrada pueden, en este caso, obtenerse a través de, por ejemplo, uno o más micrófonos, o de un almacenamiento de archivos o datos de audio multicanal.

Al menos cuando "nodo de recepción" se refiere a un nodo de lado del decodificador, el nodo de recepción puede funcionar para decodificar audio, tal como voz, y para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El nodo de recepción además puede funcionar para aplicar supresión de silencio o un esquema de DTX que comprende, por ejemplo, transmisión de tramas SID (descriptor de inserción de silencio) durante la inactividad del habla. El nodo de recepción puede ser, por ejemplo, un teléfono celular, un UE, una tableta, un ordenador o cualquier otro dispositivo capaz de comunicación por cable y/o inalámbrica y de decodificación de audio.

Disposiciones de ejemplo, figuras 6 y 7

Las realizaciones descritas en el presente documento también se refieren a una disposición. La disposición podría comprender una entidad, como se ilustra en la figura 6; o dos entidades, como se ilustra en la figura 7. La disposición 600 de una sola entidad se ilustra para representar una solución relacionada con, por ejemplo, un cancelador de eco, que determina las características espectrales y espaciales de las señales de audio de entrada, y genera un ruido de confort basándose en estas características determinadas para una pluralidad de canales de salida. La disposición 600 podría ser o comprender un nodo de recepción como se describe a continuación que tiene una función de cancelación de eco.

La disposición 700 de dos entidades se ilustra para representar una solución de unidad de codificación/decodificación; donde la determinación de las características espectrales y espaciales es realizada en una entidad o nodo 710, y luego señalizada a otra entidad o nodo 720, donde se genera el ruido de confort. La entidad 710 podría ser un nodo de transmisión, como se describe a continuación; y la entidad 720 podría ser un nodo de recepción como se describe a continuación que tiene una función de lado del decodificador.

La disposición comprende al menos un procesador 603, 711, 712 y al menos una memoria 604, 712, 722, donde al menos dicha memoria contiene instrucciones 605, 713, 714 ejecutables por al menos dicho procesador. Mediante la ejecución de las instrucciones, la disposición es operativa para determinar las características espectrales de las

señales de audio en al menos dos canales de audio de entrada; determinar la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y además generar ruido de confort para al menos dos canales de audio de salida, basándose en las características espectrales determinadas y la coherencia espacial.

Nodo de transmisión de ejemplo, figura 8

Las realizaciones descritas en el presente documento también se refieren a un nodo 800 de transmisión. El nodo de transmisión está asociado con las mismas características técnicas, objetos y ventajas que el método descrito anteriormente e ilustrado, por ejemplo, en las figuras 2 y 4. El nodo de transmisión se describirá brevemente para evitar repeticiones innecesarias. El nodo 800 de transmisión podría ser, por ejemplo, un equipo de usuario UE, tal como un UE de LTE, un dispositivo de comunicación, una tableta, un ordenador o cualquier otro dispositivo capaz de comunicación inalámbrica y/o por cable. El nodo de transmisión puede funcionar para comunicarse en uno o más sistemas de comunicación inalámbrica, tales como UMTS, E-UTRAN o CDMA 2000 y/o en uno o más tipos de redes de comunicación de corto alcance.

A continuación, se describirá un nodo 800 de transmisión de ejemplo, adaptado para permitir la realización de un método descrito anteriormente realizado por un nodo de transmisión, con referencia a la figura 8.

El nodo de transmisión puede funcionar para aplicar supresión de silencio o un esquema DTX, y puede funcionar para comunicarse con otros nodos o entidades en una red de comunicación.

La parte del nodo de transmisión que está relacionada principalmente con la solución sugerida en el presente documento se ilustra como un grupo 801 rodeado por una línea interrumpida/discontinua. El grupo 801, y posiblemente otras partes del nodo de transmisión, se adaptada para permitir la realización de uno o más de los métodos o procedimientos descritos anteriormente e ilustrados, por ejemplo, en la figura 4. El nodo de transmisión puede comprender una unidad 802 de comunicación para comunicarse con otros nodos y entidades, y puede comprender una funcionalidad adicional 807 útil para que el nodo 110 de transmisión sirva para su propósito como nodo de comunicación. Estas unidades se ilustran con una línea discontinua.

El nodo de transmisión ilustrado en la figura 8 comprende medios de procesamiento, en este ejemplo en forma de un procesador 803 y una memoria 804, donde dicha memoria contiene instrucciones 805 ejecutables por dicho procesador, por lo que el nodo de transmisión puede funcionar para realizar el método descrito anteriormente. Es decir, el nodo de transmisión está operativo para determinar las características espectrales de las señales de audio en al menos dos canales de audio de entrada y para señalizar información sobre las características espectrales de las señales de audio en al menos dichos dos canales de audio de entrada. La memoria 804 contiene además instrucciones ejecutables por dicho procesador por lo que el nodo de transmisión es además operativo para determinar la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y para señalizar información sobre la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

Como se mencionó anteriormente, la coherencia espacial se puede determinar aplicando una función de coherencia en una representación de las señales de audio en al menos dichos dos canales de audio de entrada. Además, la coherencia espacial C_{xy} entre dos señales, x e y, de al menos dos señales, se puede determinar como: $C_{xy} = |S_{xy}|^2/(S_{xx}^2 * S_{yy}^2)$; donde S_{xy} es la densidad espectral cruzada entre x e y, y S_{xx} y S_{yy} es la densidad autoespectral de x e y respectivamente. La coherencia puede aproximarse como una correlación cruzada entre las señales de audio en los respectivos canales de audio de entrada.

El programa informático 805 puede transportarse mediante un medio de almacenamiento legible por ordenador conectable al procesador. El producto del programa informático puede ser la memoria 804. El medio de almacenamiento legible por ordenador, por ejemplo, la memoria 804, puede realizarse como, por ejemplo, una RAM (memoria de acceso aleatorio), ROM (memoria de solo lectura) o una EEPROM (ROM programable borrable eléctrica). Además, el programa informático puede transportarse mediante un medio legible por ordenador por separado, tal como un CD, DVD, USB o memoria flash, desde el cual el programa podría descargarse en la memoria 804. Alternativamente, el programa informático puede almacenarse en un servidor u otra entidad conectada a una red de comunicación a la que tiene acceso el nodo de transmisión, por ejemplo, a través de la unidad 802 de comunicación. El programa informático puede luego descargarse del servidor a la memoria 804.

El programa informático podría ser transportado además por una portadora no tangible, tal como una señal electrónica, una señal óptica o una señal de radio.

El grupo 801 y otras partes del nodo de transmisión podrían implementarse, por ejemplo, por uno o más de: un procesador o microprocesador y un software y almacenamiento adecuados, por lo tanto, un dispositivo lógico programable, PLD, u otro u otros componentes electrónicos/circuito o circuitos de procesamiento configurados para realizar las acciones mencionadas anteriormente. Aunque las instrucciones descritas en las realizaciones divulgadas anteriormente se implementan como un programa informático 805 para ser ejecutado por el procesador 803, al

menos una de las instrucciones puede implementarse en realizaciones alternativas al menos parcialmente como circuitos de hardware.

El grupo 801 puede implementarse alternativamente y/o describirse esquemáticamente como se ilustra en la figura 9. El grupo 901 comprende una unidad 903 de determinación, para determinar las características espectrales de las señales de audio en al menos dos canales de audio de entrada, y para determinar la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada. El grupo comprende además una unidad 904 de señalización para señalizar información sobre las características espectrales de las señales de audio en al menos dos canales de audio de entrada, y para señalizar información sobre la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción

El nodo 900 de transmisión podría ser, por ejemplo, un equipo de usuario UE, tal como un UE de LTE, un dispositivo de comunicación, una tableta, un ordenador o cualquier otro dispositivo capaz de comunicación inalámbrica. El nodo de transmisión puede funcionar para comunicarse en uno o más sistemas de comunicación inalámbrica, tales como UMTS, E-UTRAN o CDMA 2000 y/o en uno o más tipos de redes de comunicación de corto alcance.

La coherencia espacial puede determinarse, mediante el nodo 900 de transmisión, aplicando una función de coherencia en una representación de las señales de audio en al menos dichos dos canales de audio de entrada. Además, la coherencia espacial C_{xy} entre dos señales, x e y, de al menos dos señales, se puede determinar como: $C_{xy} = |S_{xy}|^2/(S_{xx}^2 * S_{yy}^2)$; donde S_{xy} es la densidad espectral cruzada entre x e y, y S_{xx} y S_{yy} es la densidad autoespectral de x e y respectivamente. La coherencia puede aproximarse como una correlación cruzada entre las señales de audio en los respectivos canales de audio de entrada.

El grupo 901 y otras partes del nodo de transmisión podrían implementarse, por ejemplo, por uno o más de: un procesador o microprocesador y un software y almacenamiento adecuados, por lo tanto, un dispositivo lógico programable, PLD, u otro componente o componentes electrónicos/circuito o circuitos de procesamiento configurados para realizar las acciones mencionadas anteriormente.

El nodo 900 de transmisión, ilustrado en la figura 9, puede comprender además una unidad 902 de comunicación para comunicarse con otras entidades, una o más memorias 907, por ejemplo, para almacenar información y funcionalidades adicionales 908, tales como procesamiento de señales y/o interacción del usuario.

Nodo de recepción de ejemplo, figura 10

Las realizaciones descritas en el presente documento también se refieren a un nodo 1000 de recepción. El nodo de recepción está asociado con las mismas características técnicas, objetos y ventajas que el método descrito anteriormente e ilustrado, por ejemplo, en las figuras 3 y 5. El nodo de recepción se describirá brevemente para evitar repeticiones innecesarias. El nodo 1000 de recepción podría ser, por ejemplo, un equipo de usuario UE, tal como un UE de LTE, un dispositivo de comunicación, una tableta, un ordenador o cualquier otro dispositivo capaz de comunicación inalámbrica. El nodo de recepción puede funcionar para comunicarse en uno o más sistemas de comunicación inalámbrica, tales como UMTS, E-UTRAN o CDMA 2000 y/o en uno o más tipos de redes de comunicación de corto alcance.

El nodo de recepción puede funcionar para aplicar supresión de silencio o un esquema de DTX, y puede funcionar para comunicarse con otros nodos o entidades en una red de comunicación; al menos cuando el nodo de recepción se describe en una función como una unidad decodificadora que recibe información espectral y espacial de un nodo de transmisión. A continuación, se describirá un nodo 1000 de recepción de ejemplo, adaptado para permitir la realización de un método descrito anteriormente realizado por un nodo de recepción, con referencia a la figura 10.

La parte del nodo de recepción que está relacionada principalmente con la solución sugerida en el presente documento se ilustra como un grupo 1001 rodeado por una línea interrumpida/discontinua. El grupo 1001 y posiblemente otras partes del nodo de recepción están adaptados para permitir la realización de uno o más de los métodos o procedimientos descritos anteriormente e ilustrados, por ejemplo, en las figuras 1, 3 o 5. El nodo de recepción puede comprender una unidad 1002 de comunicación para comunicarse con otros nodos y entidades, y puede comprender una funcionalidad adicional 1007, tal como un procesamiento y/o comunicación adicional de la señal e interacción del usuario. Estas unidades se ilustran con una línea discontinua.

El nodo de recepción ilustrado en la figura 10 comprende medios de procesamiento, en este ejemplo en forma de un procesador 1003 y una memoria 1004, en el que dicha memoria contiene instrucciones 1005 ejecutables por dicho procesador, por lo que el nodo de transmisión puede funcionar para realizar el método descrito anteriormente. Es decir, el nodo de recepción está operativo para obtener, es decir, recibir o determinar, las características espectrales de las señales de audio en al menos dos canales de audio de entrada. La memoria 1004 contiene además instrucciones ejecutables por dicho procesador por lo que el nodo de recepción es además operativo para obtener, es decir, recibir o determinar, la coherencia espacial entre las señales de audio en los respectivos canales de audio

de entrada; y para generar ruido de confort, para al menos dos canales de audio de salida, basándose en la información obtenida sobre las características espectrales y la coherencia espacial.

La generación de una señal de ruido de confort N_1 para un canal de audio de salida puede comprender determinar una función de conformación espectral H_1, basándose en la información sobre características espectrales de una de las señales de audio de entrada y la coherencia espacial entre la señal de audio de entrada y al menos otra señal de audio de entrada La generación puede comprender además aplicar la función de conformación espectral H_1 a una primera señal de ruido aleatorio W_1 y a una segunda señal de ruido aleatorio W_2(f), donde W_2(f) se pondera basándose en la coherencia entre la señal de audio de entrada y al menos otra señal de audio de entrada.

La obtención de información puede comprender recibir la información de un nodo de transmisión. Alternativamente, el nodo de recepción puede comprender un cancelador de eco, y la obtención de información puede comprender la determinación de la información basándose en señales de audio de entrada en al menos dos canales de audio. Es decir, como se describió anteriormente, en el caso de la función de cancelación de eco, la determinación de las características espectrales y espaciales está determinada por la misma entidad, por ejemplo, un NLP. En este último caso, el "receptor" en el nodo de recepción puede estar asociado, por ejemplo, con la recepción de al menos dos señales del canal de audio, por ejemplo, a través de un micrófono.

El grupo 1001 puede implementarse alternativamente y/o describirse esquemáticamente como se ilustra en la figura 11. El grupo 1101 comprende una unidad 1103 de obtención, para obtener información sobre las características espectrales de las señales de audio de entrada en al menos dos canales de audio; y para obtener información sobre la coherencia espacial entre las señales de audio de entrada en al menos dos canales de audio. El grupo 1101 comprende además una unidad 1104 de generación de ruido para generar ruido de confort para al menos dos canales de audio de salida, basándose en la información obtenida sobre las características espectrales y la coherencia espacial.

El nodo 1100 de recepción podría ser, por ejemplo, un equipo de usuario UE, tal como un UE de LTE, un dispositivo de comunicación, una tableta, un ordenador o cualquier otro dispositivo capaz de comunicación inalámbrica y/o por cable. El nodo de recepción puede funcionar para comunicarse en uno o más sistemas de comunicación inalámbrica, tales como UMTS, E-UTRAN o CDMA 2000 y/o en uno o más tipos de redes de comunicación de corto alcance.

Como para el nodo 1000 de recepción, la generación de una señal de ruido de confort N_1 para un canal de audio de salida puede comprender la determinación de una función de conformación espectral H_1, basándose en la información sobre características espectrales de una de las señales de audio de entrada y la coherencia espacial entre la señal de audio de entrada y al menos otra señal de audio de entrada. La generación puede comprender además aplicar la función de conformación espectral H_1 a una primera señal de ruido aleatorio W_1 y a una segunda señal de ruido aleatorio W_2(f), donde W_2(f) se pondera basándose en la coherencia entre la señal de audio de entrada y al menos otra señal de audio de entrada.

La obtención de información puede comprender recibir la información de un nodo de transmisión. Alternativamente, el nodo de recepción puede comprender un cancelador de eco, y la obtención de información puede comprender la determinación de la información basándose en señales de audio de entrada en al menos dos canales de audio.

El grupo 1101 y otras partes del nodo de recepción podrían implementarse, por ejemplo, por uno o más de: un procesador o microprocesador y un software y almacenamiento adecuados, por lo tanto, un dispositivo lógico programable, PLD, u otro componente o componentes electrónicos/circuito o circuitos de procesamiento configurados para realizar las acciones mencionadas anteriormente.

El nodo 1100 de recepción, ilustrado en la figura 11, puede comprender además una unidad 1102 de comunicación para comunicarse con otras entidades, una o más memorias 1107, por ejemplo, para almacenar información y funcionalidades adicionales 1107, tales como procesamiento de señales y/o interacción del usuario.

Debe entenderse que la elección de unidades o módulos que interactúan, así como la denominación de las unidades, son solo para fines ilustrativos, y las disposiciones, los nodos de transmisión y recepción adecuados para ejecutar cualquiera de los métodos descritos anteriormente pueden configurarse en una pluralidad de formas alternativas para poder ejecutar las acciones de proceso sugeridas.

También debe observarse que las unidades o módulos descritos en esta divulgación se deben considerar como entidades lógicas y no necesariamente como entidades físicas separadas.

REIVINDICACIONES

1.- Método, realizado por un nodo de transmisión para soportar la generación de ruido de confort para al menos dos canales de audio, comprendiendo el método:

- determinar (401) las características espectrales de las señales de audio en al menos dos canales de audio de entrada;

- señalizar (403) información sobre las características espectrales de las señales de audio en al menos dichos dos canales de audio de entrada;

- determinar (402) una coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; y

- señalizar (403) información sobre la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

2.- Método de acuerdo con la reivindicación 1, en el que la coherencia espacial se determina aplicando una función de coherencia en las señales de audio en al menos dichos dos canales de audio de entrada.

3.- Método de acuerdo con la reivindicación 1 o 2, en el que la coherencia espacial C_{xy} entre dos señales, x e y de las al menos dos señales, se determina como: $C_{xy} = |S_{xy}|^2/(S_{xx}2 * S_{yy}2)$; donde S_{xy} es la densidad espectral cruzada entre x e y, y S_{xx} y S_{yy} es la densidad autoespectral de x e y respectivamente.

4.- Método de acuerdo con cualquiera de las reivindicaciones 1-3, en el que la coherencia se aproxima como una correlación cruzada entre las señales de audio en los respectivos canales de audio de entrada.

5.- Método realizado por un nodo de recepción para generar ruido de confort para al menos dos canales de audio, comprendiendo el método:

- obtener información sobre las características espectrales de las señales de audio de entrada en al menos dos canales de audio;

- obtener información sobre la coherencia espacial entre las señales de audio de entrada en al menos dos canales de audio; y

- generar ruido de confort para al menos dos canales de audio de salida, basándose en la información obtenida sobre las características espectrales y la coherencia espacial.

6.- Método de acuerdo con la reivindicación 5, en el que la generación de una señal de ruido de confort para un canal de audio de salida comprende:

- determinar una función de conformación espectral, basándose en la información sobre características espectrales de una de las señales de audio de entrada y la coherencia espacial entre la señal de audio de entrada y al menos otra señal de audio de entrada; y

- aplicar la función de conformación espectral a una primera señal de ruido aleatorio y a una segunda señal de ruido aleatorio, donde la segunda señal de ruido aleatorio se pondera basándose en la coherencia entre la señal de audio de entrada y al menos dicha otra señal de audio de entrada.

7.- Método de acuerdo con la reivindicación 5 o 6, en el que la obtención de información comprende recibir la información de un nodo de transmisión.

8.- Método de acuerdo con la reivindicación 5 o 6, en el que el nodo de recepción comprende un cancelador de eco, y la obtención de información comprende determinar la información basándose en señales de audio de entrada en al menos dos canales de audio.

9.- Nodo (800) de transmisión para soportar la generación de ruido de confort para al menos dos canales de audio, estando configurado el nodo de transmisión para:

- determinar las características espectrales de las señales de audio en al menos dos canales de audio de entrada;

- señalizar información sobre las características espectrales de las señales de audio en al menos dichos dos canales de audio de entrada a un nodo de recepción;

- determinar una coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada; e

- señalizar información sobre la coherencia espacial entre las señales de audio en los respectivos canales de audio de entrada a un nodo de recepción, para la generación de ruido de confort para al menos dos canales de audio en el nodo de recepción.

10.- Nodo de transmisión de acuerdo con la reivindicación 9, en el que la coherencia espacial se determina aplicando una función de coherencia en una representación de las señales de audio en al menos dichos dos canales de audio de entrada.

11.- Nodo de transmisión de acuerdo con la reivindicación 9 o 10, en el que la coherencia espacial C_{xy} entre dos señales, x e y, de al menos dos señales, se determina como:

$$C_{xy} = |Sxy|^2/(S_{xx}{}^2 * S_{yy}{}^2);$$

donde S_{xy} es la densidad espectral cruzada entre x e y, y S_{xx} y S_{yy} es la densidad autoespectral de x e y respectivamente.

12.- Nodo de transmisión de acuerdo con cualquiera de las reivindicaciones 9-11, en el que la coherencia se aproxima como una correlación cruzada entre las señales de audio en los respectivos canales de audio de entrada.

13.- Nodo (1000) de recepción para generar ruido de confort para al menos dos canales de audio, estando configurado el nodo de recepción para:

- obtener información sobre las características espectrales de las señales de audio en al menos dos canales de audio;

- obtener información sobre la coherencia espacial entre las señales de audio en al menos dichos dos canales de audio; y

- generar ruido de confort para al menos dos canales de audio de salida, basándose en la información obtenida sobre las características espectrales y la coherencia espacial.

14.- Nodo de recepción de acuerdo con la reivindicación 13, en el que la generación de una señal de ruido de confort para un canal de audio de salida comprende:

- determinar una función de conformación espectral, basándose en la información sobre características espectrales de una de las señales de audio y la coherencia espacial entre la señal de audio y al menos otra señal de audio; y

- aplicar la función de conformación espectral a una primera señal de ruido aleatorio y a una segunda señal de ruido aleatorio, donde la segunda señal de ruido aleatorio se pondera basándose en la coherencia entre la señal de audio y al menos dicha otra señal de audio.

15.- Un nodo de recepción de acuerdo con la reivindicación 13 o 14, en el que la obtención de información comprende recibir la información de un nodo de transmisión.

16.- Un nodo de recepción de acuerdo con la reivindicación 13 o 14, en el que el nodo de recepción comprende un cancelador de eco, y la obtención de información comprende determinar la información basándose en señales de audio de entrada en al menos dos canales de audio.

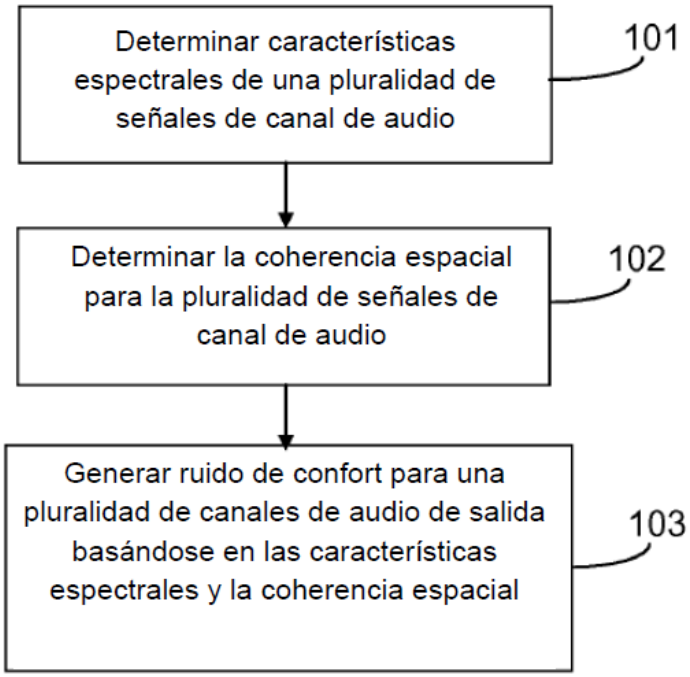

Figura 1

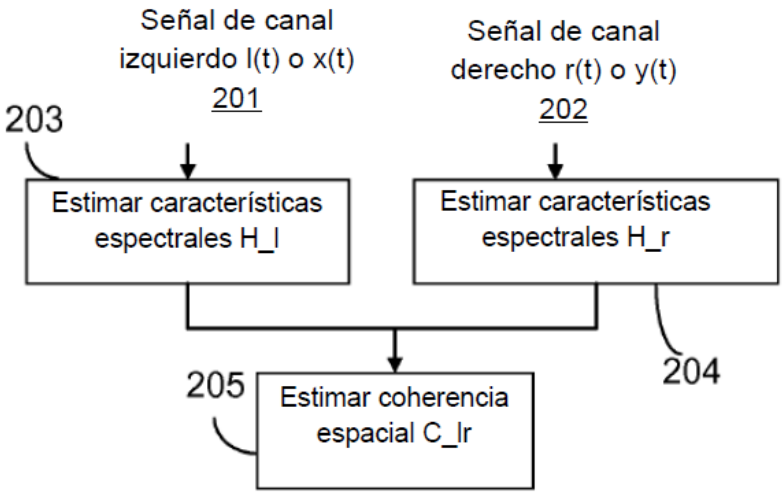

Figura 2

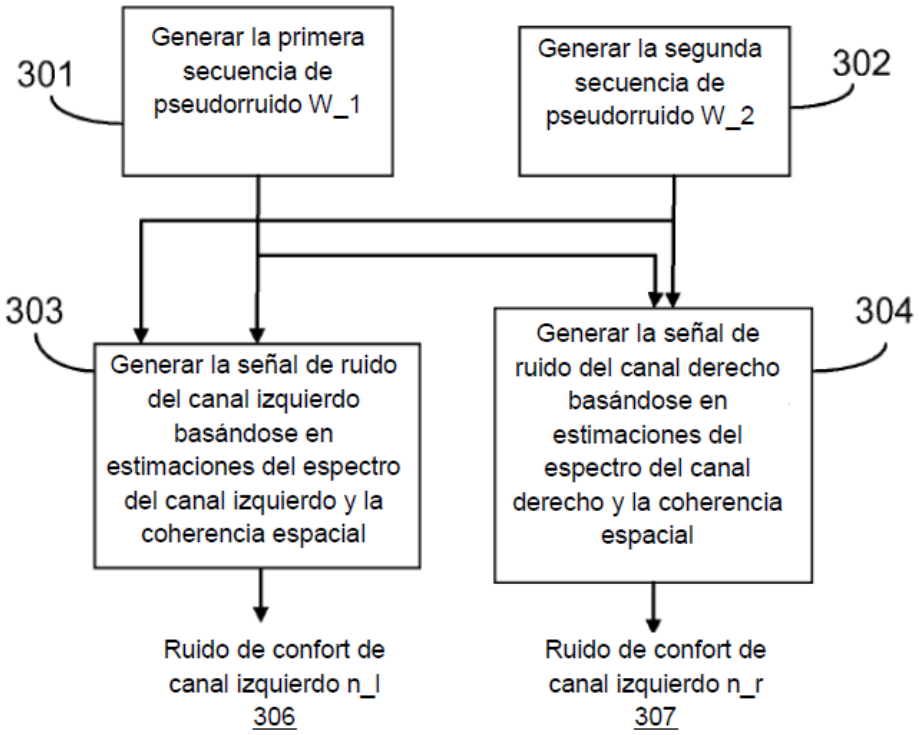

Figura 3

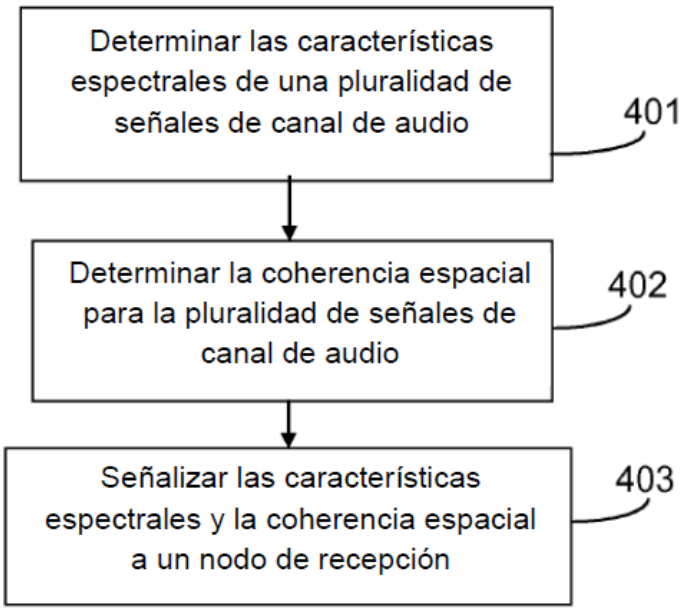

Figura 4

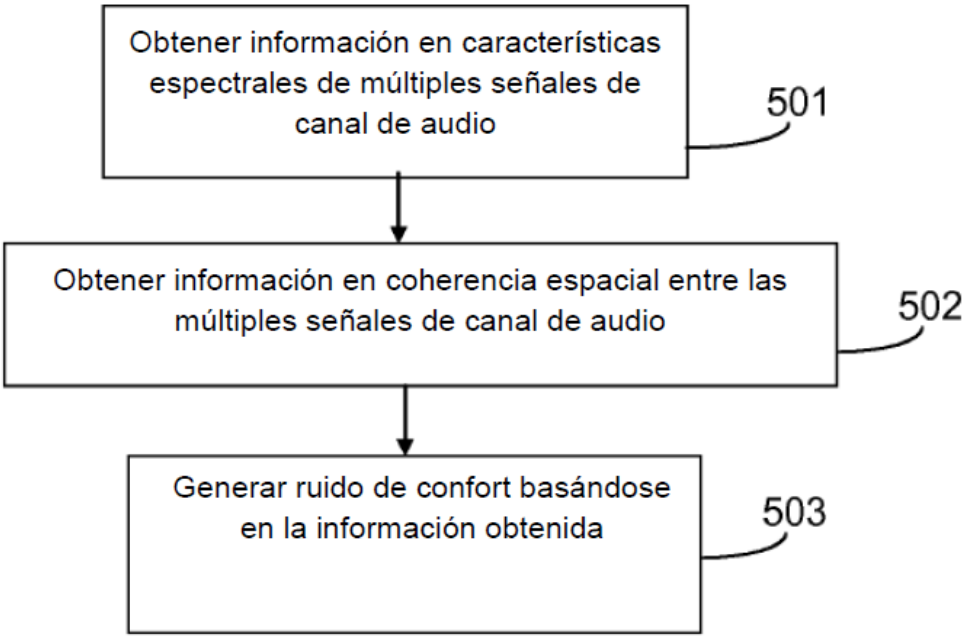
Obtener información en características espectrales de múltiples señales de canal de audio
501
Obtener información en coherencia espacial entre las múltiples señales de canal de audio
502
Generar ruido de confort basándose en la información obtenida
503

Figura 5

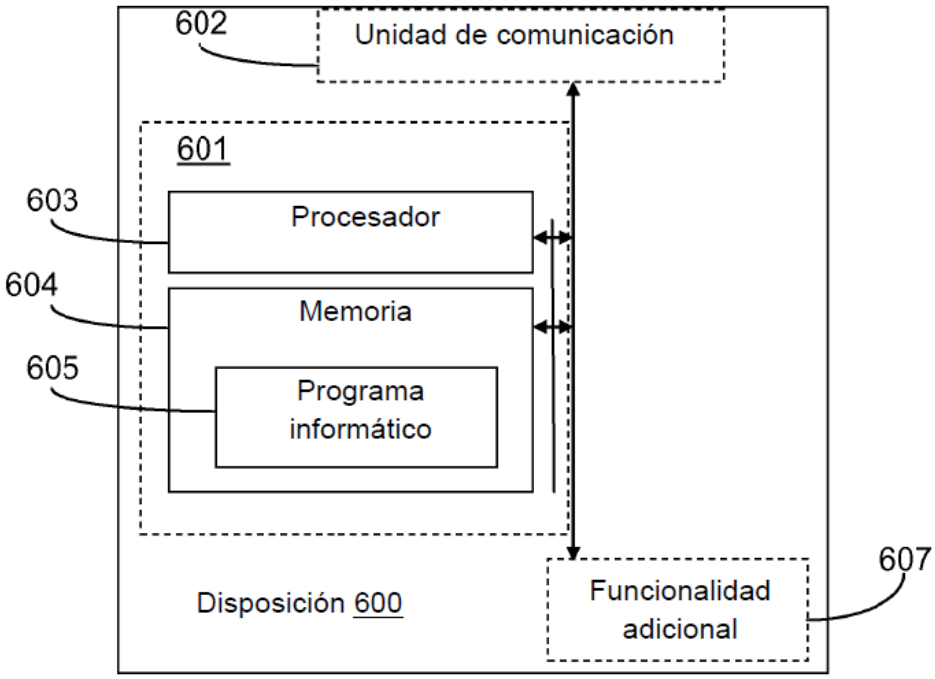

Figura 6

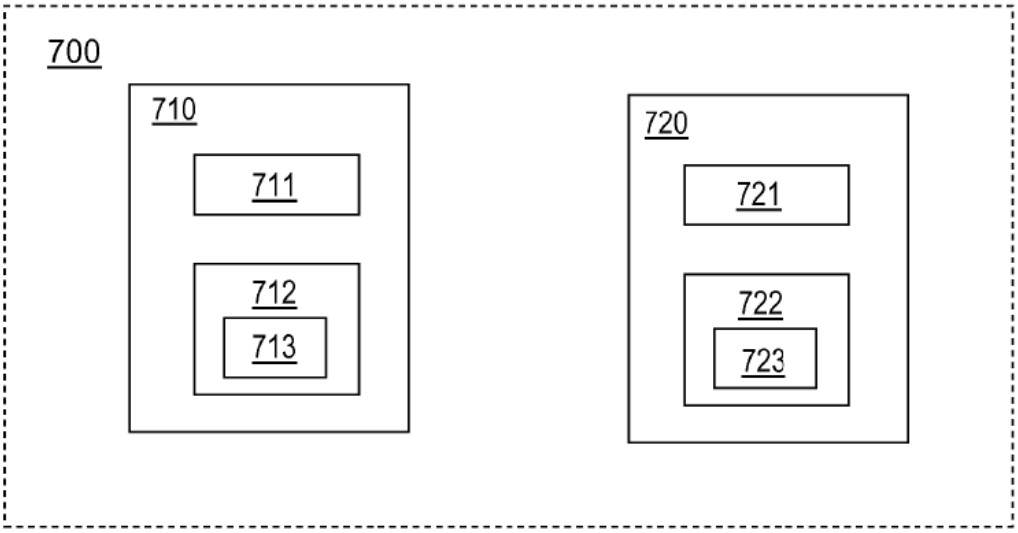

Figura 7

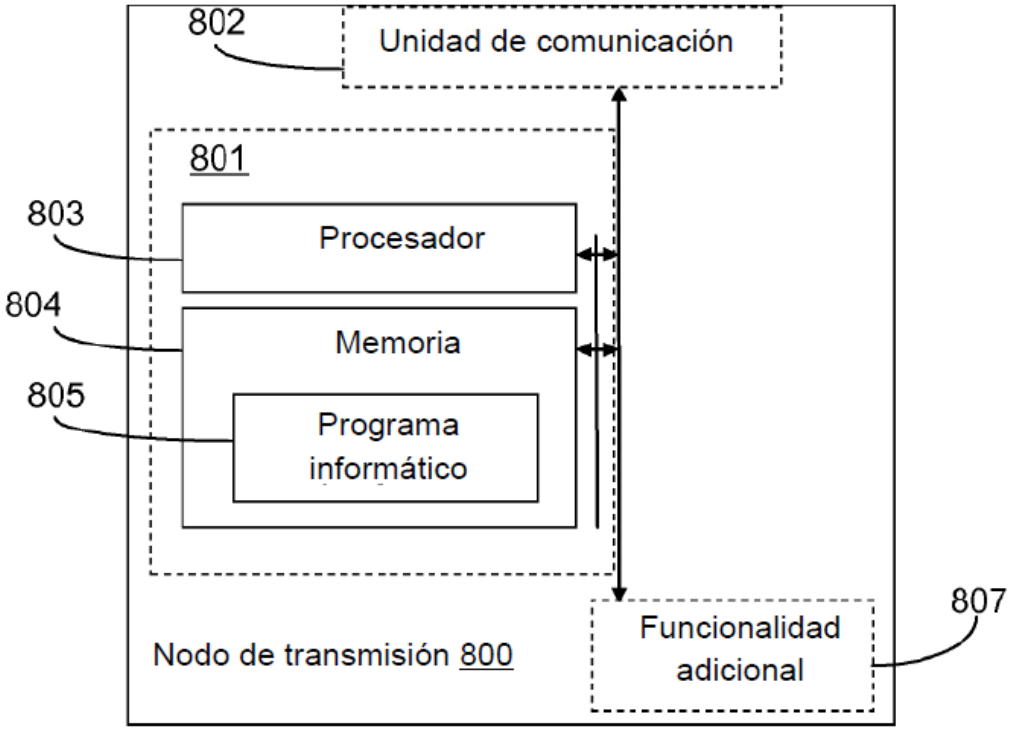

Figura 8

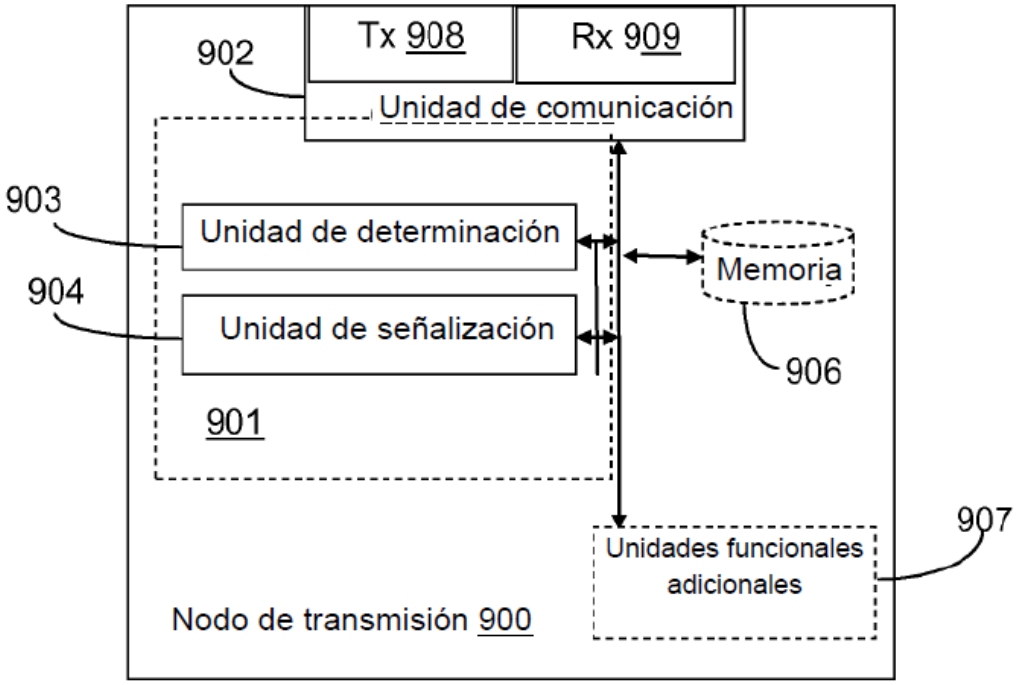

Figura 9

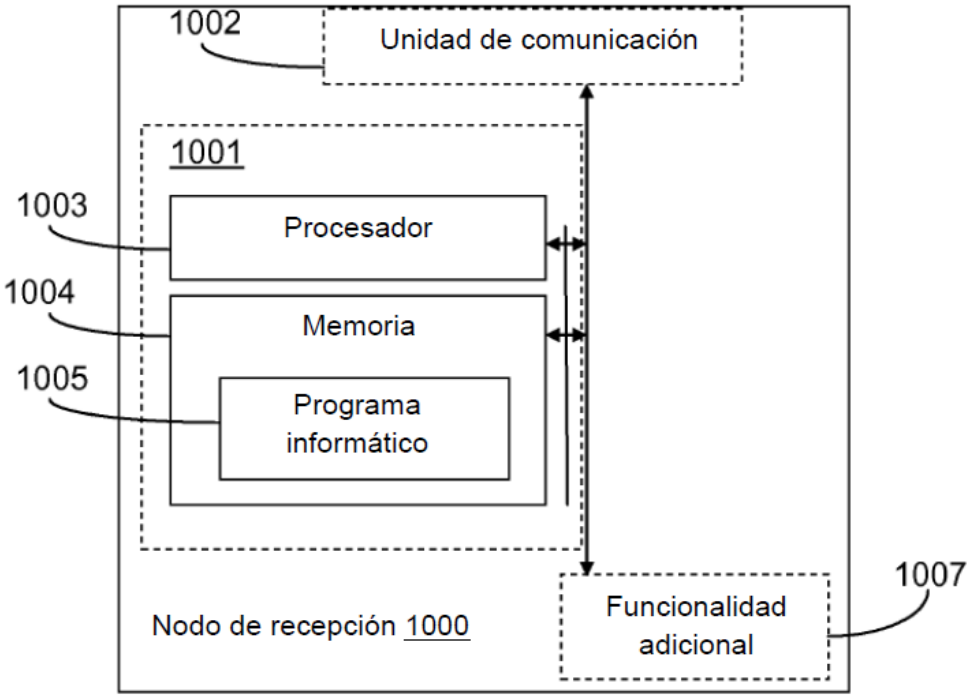

Figura 10

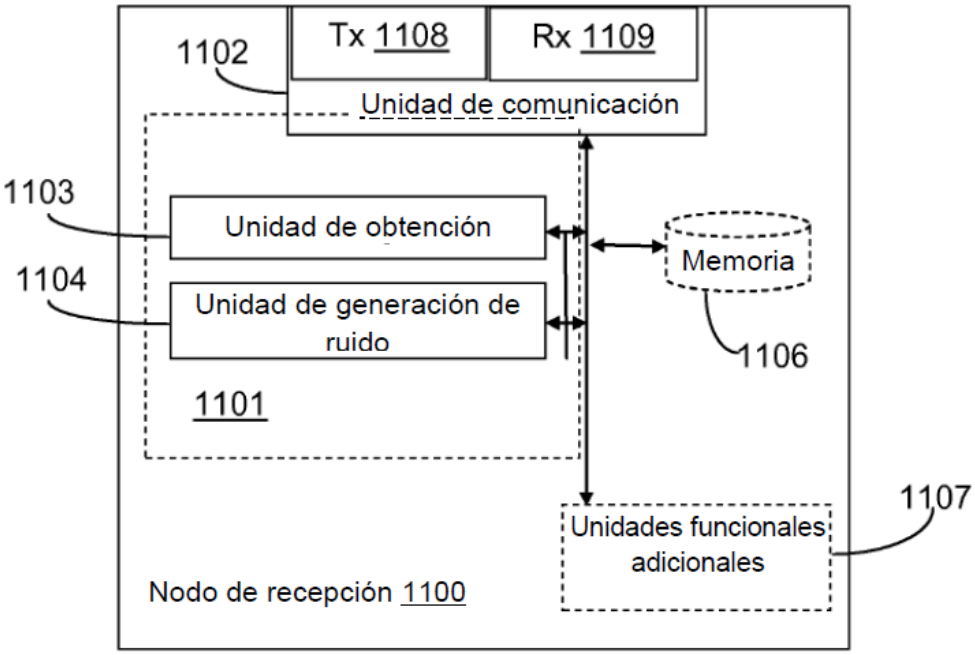

Figura 11