



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 347 760**

51 Int. Cl.:
G10L 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **01273554 .4**

96 Fecha de presentación : **19.11.2001**

97 Número de publicación de la solicitud: **1356461**

97 Fecha de publicación de la solicitud: **29.10.2003**

54 Título: **Procedimiento y dispositivo de reducción de ruido.**

30 Prioridad: **30.01.2001 FR 01 01220**

45 Fecha de publicación de la mención BOPI:
04.11.2010

45 Fecha de la publicación del folleto de la patente:
04.11.2010

73 Titular/es: **FRANCE TELECOM**
6 place d'Alleray
75015 Paris, FR

72 Inventor/es: **Scalart, Pascal;**
Marro, Claude y
Mauuary, Laurent

74 Agente: **Justo Bailey, Mario de**

ES 2 347 760 T3

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

DESCRIPCIÓN

Procedimiento y dispositivo de reducción de ruido.

5 La presente invención se refiere a las técnicas de procesamiento de señal que sirve para reducir el nivel de ruido presente en una señal de entrada.

Un importante ámbito de aplicación es el del procesamiento de la señal de audio (voz o música), incluyendo de manera no limitativa:

- 10 • teleconferencia y videoconferencia en medio ruidoso (en sala dedicada o bien a partir de ordenadores multimedia, etc.);
- telefonía: procesamiento a nivel de los terminales, fijos o móviles y/o en las redes de transporte;
- 15 • terminales manos libres, especialmente de oficina, de vehículo o móviles;
- toma de sonido en lugares públicos (estación, aeropuerto, etc.);
- 20 • toma de sonido manos libres en vehículos;
- reconocimiento de voz robusta en el entorno acústico;
- toma de sonido para cine y en los medios de comunicación (radio y televisión, por ejemplo para periodismo deportivo o conciertos, etc.).

30 La invención se puede aplicar asimismo a cualquier ámbito en el que se desee extraer una información útil a partir de una observación con ruido. Se pueden mencionar especialmente los siguientes ámbitos: imagen submarina, teledetección submarina, procesamiento de señales biomédicas (EEG, ECG, imagen biomédica, etc.).

Un problema característico de la toma de sonido se refiere al entorno acústico en el que se coloca el micrófono de toma de sonido y, más concretamente, el hecho de que, debido a la imposibilidad de controlar totalmente dicho entorno, una señal perturbadora (calificada de ruido) está asimismo presente dentro de la señal de observación.

35 Con objeto de mejorar la calidad de la señal, se desarrollan sistemas de reducción de ruido con objeto de extraer la información útil efectuando un procesamiento de la señal de observación con ruido. Cuando la señal de audio es una señal de voz transmitida a distancia, estos sistemas permiten incrementar su inteligibilidad y reducir la fatiga del interlocutor. Además de dichas aplicaciones de comunicación hablada, la mejora de la calidad de la señal de voz se revela asimismo útil para el reconocimiento de voz, cuyas prestaciones quedan fuertemente alteradas cuando el 40 utilizador se encuentra en un entorno ruidoso.

La elección de una técnica de procesamiento de la señal para realizar la operación de reducción de ruido depende, en primer lugar, del número de observaciones disponibles en la entrada del procesamiento. En la presente descripción, 45 se considerará el caso en que está disponible una única señal de observación. Los métodos de reducción de ruido adaptados a esta problemática monosensor se basan principalmente en técnicas de procesamiento de la señal tales como el filtrado adaptativo mediante adelanto/retraso, el filtrado paramétrico de Kalman o también el filtrado mediante modificación espectral a corto plazo.

50 Esta última familia (filtrado mediante modificación espectral a corto plazo) agrupa prácticamente el conjunto de soluciones utilizadas en los equipos industriales debido a la sencillez de los conceptos en juego y a la gran disponibilidad de herramientas básicas (por ejemplo, la transformada de Fourier discreta) necesarias para su programación. Sin embargo, el desarrollo de dichas técnicas de reducción de ruido se basa en gran medida en la posibilidad de efectuar con facilidad dichos procesamiento en tiempo real en un dispositivo de procesamiento de la señal, sin introducir 55 distorsiones relevantes en la señal disponible a la salida del procesamiento. En los métodos de dicha familia, el procesamiento consiste la mayoría de las veces en evaluar una función de transferencia de un filtro reductor de ruido, y en realizar el filtrado a partir de una multiplicación en el ámbito espectral, lo que permite operar la reducción de ruido mediante atenuación espectral a corto plazo, mediante un procesamiento por bloques.

60 Se escribe $x(n)$ la señal de observación con ruido procedente de la mezcla de la señal deseada $s(n)$ y el ruido de perturbación $b(n)$, designando n el índice temporal en tiempo discreto. La elección de una representación en tiempo discreto está unida a una aplicación orientada al procesamiento digital de la señal, pero se observará que los métodos descritos más adelante se aplican asimismo a las señales en tiempo continuo. Se analiza la señal por segmentos sucesivos o tramas de índice k de longitud constante. Algunas formas de anotación habitualmente utilizadas para 65 representaciones en los ámbitos discretos tiempo y frecuencia son:

- $X(k, f)$: transformada de Fourier (f es el índice de frecuencia de la k -ésima trama (k es el índice de trama) de la señal analizada $x(n)$;

- $S(k,f)$: transformada de Fourier de la k -ésima trama de la señal deseada $s(n)$;
- $\hat{v}$: estimación de una magnitud (temporal o frecuencial) v ; por ejemplo, $\hat{S}(k,f)$ es la estimación de la transformada de Fourier de la señal deseada;
- $\gamma_{uu}(f)$: densidad espectral de potencia (DSP) de una señal $u(n)$.

En la mayoría de las técnicas de reducción de ruido, la señal con ruido $x(n)$ soporta un filtrado en el ámbito frecuencial para producir una señal útil estimada $\hat{s}(n)$ lo más próxima posible a la señal original $s(n)$ desprovista de cualquier perturbación. Como se ha indicado anteriormente, esta operación de filtrado consiste en reducir cada componente frecuencial f de la señal con ruido habida cuenta de la relación señal/ruido (RSB) estimada en dicha componente. Dicha RSB función de la frecuencia f se escribe en este caso $\eta(k,f)$ para la trama k .

Para cada una de las tramas, se multiplica primero la señal por una ventana de ponderación que permite mejorar la estimación posterior de las magnitudes espectrales necesarias para el cálculo del filtro reductor de ruido. Cada trama así ventaneada se analiza a continuación en el ámbito espectral (generalmente con la ayuda de la transformada de Fourier discreta en su versión rápida). Esta operación se denomina transformación de Fourier a corto plazo (TFCT). Esta representación frecuencial $X(k,f)$ de la señal observada permite al mismo tiempo estimar la función de transferencia $H(k,f)$ del filtro reductor de ruido, y aplicar dicho filtro en el ámbito espectral mediante simple multiplicación entre dicha función de transferencia y el espectro a corto plazo de la señal con ruido, es decir:

$$\hat{S}(k,f) = H(k,f).X(k,f) \quad (1)$$

El regreso en el ámbito temporal de la señal así obtenida se efectúa a continuación mediante simple transformada espectral inversa. La síntesis de la señal a la que se ha quitado el ruido se efectúa generalmente mediante una técnica de solapamiento e incorporación de bloques (OLA, "overlap add") o una técnica de salvaguarda de bloques (OLS, "overlap save"). Esta operación de reconstrucción de la señal en el ámbito temporal se denomina transformación de Fourier a corto plazo inverso (TFCTI).

Se encontrará una descripción detallada de los métodos de atenuación espectral a corto plazo en las referencias: J.S. Lim, A.V. Oppenheim, "Enhancement and bandwidth compression of noisy speech", Proceedings of the IEEE, vol. 67, pp. 1586-1604, 1979; y R.E. Crochiere, L.R. Rabiner, "Multirate digital signal processing", Prentice Hall, 1983.

Las principales tareas realizadas por dicho sistema de reducción de ruido son:

- la detección de actividad vocal (DAV);
- la estimación de la densidad espectral de potencia (DSP) de ruido durante los instantes de actividad vocal;
- la aplicación de una atenuación espectral a corto plazo evaluada a partir de una regla de supresión de las componentes espectrales de ruido;
- la síntesis de la señal procesada a partir de una técnica del tipo OLS u OLA.

La elección de la regla de supresión de las componentes de ruido es importante, ya que determina la calidad de la señal transmitida. Generalmente, dichas reglas de supresión sólo modifican la amplitud $|X(k,f)|$ de las componentes espectrales de la señal con ruido, no su fase. Generalmente, se realizan las siguientes hipótesis:

- estadísticamente, no existe correlación entre el ruido y la señal útil;
- la señal útil es intermitente (presencia de períodos de silencio en los que se puede estimar el ruido);
- el oído humano no es sensible a la fase de la señal (véase D.L. Wang, J.S. Lim, "The unimportance of phase in speech enhancement", IEEE Trans. On ASSP, vol. 30, nº 4, pp. 679-681, 1982).

La atenuación espectral a corto plazo $H(k,f)$, aplicada a la señal de observación $X(k,f)$ en la trama de índice k a la componente frecuencial f , se determina generalmente a partir de la estimación de la relación señal/ruido local $\eta(k,f)$. Una característica común al conjunto de reglas de supresión es su comportamiento asintótico, dado por:

$$\begin{aligned}
 H(k,f) &\approx 1 \quad \text{para} \quad \eta(k,f) \gg 1 \\
 H(k,f) &\approx 0 \quad \text{para} \quad \eta(k,f) \ll 1
 \end{aligned}
 \tag{2}$$

Las reglas de supresión habitualmente empleadas son:

- la sustracción espectral en potencia (véase el artículo ya citado de J.S. Lim y A.V. Oppenheim), para la que la función de transferencia $H(k,f)$ del filtro reductor de ruido se escribe:

$$H(k,f) = \sqrt{\frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)}} \tag{3}$$

- la sustracción espectral en amplitud (véase S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", IEEE Trans. On Audio, Speech and Signal Processing, vol. 27, nº 2, pp. 113-120, Abril 1979), para la que la función de transferencia se escribe:

$$H(k,f) = 1 - \sqrt{\frac{\gamma_{bb}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)}} \tag{4}$$

- el empleo directo del filtro de Wiener (véase el artículo ya citado de J.S. Lim y A.V. Oppenheim), para el que la función de transferencia $H(k,f)$ se escribe:

$$H(k,f) = \frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f) + \gamma_{ss}(k,f)} \tag{5}$$

En dichas expresiones, $\gamma_{ss}(k,f)$ y $\gamma_{bb}(k,f)$ representan respectivamente las densidades espectrales de potencia de la señal útil y del ruido presente dentro de la componente frecuencial f de la señal de observación $X(k,f)$ en la trama de índice k .

A partir de las expresiones (3)-(5), es posible estudiar, en función de la relación señal/ruido local medida en una componente frecuencial dada f , el comportamiento de la atenuación espectral aplicada a la señal con ruido. Se comprueba que todas las reglas dan lugar a una atenuación idéntica cuando la relación señal/ruido local es importante. La regla de sustracción en potencia es óptima en el sentido del máximo de verosimilitud para modelos gaussianos (véase O. Cappé, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor", IEEE Trans. On Speech and Audio Processing, vol. 2, nº 2, pp. 345-349, Abril 1994). Pero es aquella para la que la potencia del ruido sigue siendo más importante a la salida del procesamiento. Para el conjunto de las reglas de supresión, se observa que basta una escasa variación de la relación señal/ruido local alrededor del valor de corte para pasar del caso de la atenuación total ($H(k,f) = 0$) al caso de una modificación espectral despreciable ($H(k,f) = 1$).

Esta última propiedad constituye una de las causas del fenómeno calificado de "ruido musical". En efecto, el ruido ambiente, caracterizado al mismo tiempo por componentes deterministas y aleatorias, sólo se puede caracterizar durante los períodos de no actividad vocal. Debido a la presencia de dichas componentes aleatorias, existen muy fuertes variaciones entre la contribución real de una componente frecuencial f del ruido durante los períodos de actividad vocal y su estimación de media realizada en varias tramas durante los instantes de no actividad vocal. Debido a esta diferencia, la estimación de la relación señal/ruido local puede fluctuar alrededor del nivel de corte y, por lo tanto, generar a la salida del procesamiento componentes espectrales que aparecen y desaparecen, y cuya duración de vida media no supera estadísticamente el orden de magnitud de la ventana de análisis considerada. La generalización de este comportamiento en el conjunto del ancho de banda introduce un ruido residual audible y molesto, calificado de "ruido musical".

Varios estudios están orientados a reducir la influencia de dicho ruido. Las soluciones preconizadas se desglosan según varios ejes:

- media de las estimaciones a corto plazo (véase el artículo ya citado de S.F. Boll);

• sobreestimación del espectro de potencia del ruido (véase M. Berouti *et al*, “Enhancement of speech corrupted by acoustic noise”, Int. Conf. On Speech, Signal Processing, pp. 208-211, 1979; y P. Lockwood, J. Boudy, “Experiments with a non-linear spectral subtractor, hidden Markov models and the projection for robust speech recognition in cars”, Proc. of EUSIPCO’91, pp. 79-82, 1991);

• seguimiento de los mínimos de la densidad espectral de ruido (véase R. Martin, “Spectral subtraction based on minimum statistics”, in Signal Processing VII: Theories and Applications, EUSIPCO’94, pp. 1182-1185, Sept. 1994).

Varios estudios han tratado asimismo del establecimiento de nuevas reglas de supresión basadas en modelos estadísticos de las señales de voz y de ruido aditivo. Estos estudios han permitido introducir nuevos algoritmos de “decisiones flexibles”, ya que poseen un grado de libertad adicional con relación a los métodos clásicos (véase R.J. Mac Aulay, M.L. Malpass, “Speech enhancement using a soft-decision noise suppression filter”, IEEE trans. On Audio, Speech and Signal Processing, vol. 28, nº 2, pp. 137-145, Abril 1980; Y. Ephraim, D. Malah, “Speech enhancement using optimal non-linear spectral amplitude estimation”, Int. Conf. on Speech, Signal Processing, pp. 1118-1121, 1983; Y. Ephraim, D. Malah, “Speech enhancement using a minimum mean square error short-time spectral amplitude estimator”, IEEE Trans. on ASSP, vol. 32, nº 6, pp. 1109-1121, 1984).

Las reglas de modificación espectral a corto plazo anteriormente mencionadas presentan las siguientes características:

• El cálculo de la atenuación espectral a corto plazo se basa en la estimación de la relación señal/ruido en cada una de las componentes espectrales, haciendo intervenir cada una de las ecuaciones (3)-(5) la cantidad:

$$\eta(k,f) = \frac{\gamma_{ss}(k,f)}{\gamma_{bb}(k,f)} \quad (6)$$

De este modo, las prestaciones de la técnica de reducción de ruido (distorsiones, reducción efectiva del nivel de ruido) se rigen por la pertinencia de este estimador de la relación señal/ruido.

• Estas técnicas se basan en un procesamiento por bloques (con posibilidad de solapamiento entre los bloques sucesivos) que consiste en filtrar el conjunto de muestras de una trama dada, presente en la entrada del dispositivo de reducción de ruido, mediante una única atenuación espectral. Esta propiedad reside en el hecho de que la aplicación del filtro se efectúa mediante una multiplicación en el ámbito espectral. Esto es especialmente apremiante cuando la señal presente en la trama actual no verifica las hipótesis de estacionalidad en el segundo orden, por ejemplo en el caso del inicio o el final de una palabra, o en el caso de una trama mixta sonora/sorda.

• La multiplicación efectuada en el ámbito espectral corresponde en realidad a una operación de convolución cíclica. En la práctica, para evitar las distorsiones, la operación que se intenta efectuar es una convolución lineal, lo que requiere al mismo tiempo añadir cierto número de muestras nulas a cada trama de entrada (técnica denominada del “zero padding”) y efectuar un procesamiento adicional orientado a limitar el soporte temporal de la respuesta impulsional del filtro reductor de ruido. Satisfacer el requisito de convolución temporal incrementa por lo tanto necesariamente el orden de la transformada espectral y, como consecuencia, la complejidad aritmética del procesamiento de reducción de ruido. La técnica más utilizada para limitar el soporte temporal de la respuesta impulsional del filtro reductor de ruido consiste en introducir una restricción en el ámbito temporal, lo que requiere (i) una primera transformación espectral “inversa” que permite obtener la respuesta impulsional $h(k,n)$ a partir del conocimiento de la función de transferencia del filtro $H(k,f)$, (ii) una limitación del número de puntos de dicha respuesta impulsional, que conduce a un filtro temporal truncado $h'(k,n)$, y (iii) una segunda transformación espectral “directa” que permite obtener la función de transferencia modificada $H'(k,f)$ a partir de la respuesta impulsional truncada $h'(k,n)$.

• En la práctica, cada trama de análisis se multiplica por una ventana de análisis $w(n)$ antes de efectuar la operación de transformada espectral. Cuando el filtro reductor de ruido es del tipo de todo paso (es decir $H(k,f) \approx 1, \forall f$), la ventana de análisis debe cumplir la siguiente condición

$$\sum_k w(n-k.D) = 1 \quad (7)$$

si se desea que se verifique la condición de reconstitución perfecta. En esta ecuación, el parámetro D representa la diferencia (en número de muestras) entre dos tramas de análisis sucesivas. Por otra parte, la elección de la ventana de ponderación $w(n)$ (típicamente del tipo Hanning, Hamming, Blackman, etc.) condiciona el ancho del lóbulo principal

de $W(f)$ así como la amplitud de los lóbulos secundarios (con relación a la del lóbulo principal). Si el lóbulo principal es ancho, la aproximación de las transiciones rápidas de la transformada de la señal original es muy mala. Si la amplitud relativa de los lóbulos secundarios es grande, la aproximación obtenida posee oscilaciones molestas, especialmente alrededor de las discontinuidades. Por lo tanto, es difícil cumplir al mismo tiempo el requisito de análisis espectral pertinente (elección del ancho del lóbulo principal y de la amplitud de los lóbulos laterales) y el requisito de escasa demora introducida por el proceso de filtrado de reducción de ruido (desfase temporal entre la señal a la entrada y a la salida del procesamiento). Respetar el segundo requisito conduce a utilizar tramas sucesivas sin solapamiento alguno y, por lo tanto, una ventana de análisis del tipo rectangular, lo que no conduce a efectuar un análisis espectral pertinente. El único medio de respetar estos dos requisitos al mismo tiempo consiste en efectuar un análisis espectral a partir de una primera transformación espectral realizada en una trama ponderada mediante una ventana de análisis adecuada (para efectuar una correcta estimación espectral), y en realizar paralelamente una segunda transformada espectral sobre los datos no ventaneados (para efectuar la operación de convolución mediante multiplicación espectral). Dicha técnica se revela en la práctica demasiado costosa en términos de complejidad aritmética.

Los documentos EP-A-0710947 y EP 0918317 describen un dispositivo de reducción de ruido acoplado a un anulador de eco. La reducción de ruido se opera mediante un filtrado por bloques en el ámbito temporal, por medio de una respuesta impulsional obtenida mediante transformación de Fourier inversa de la función de transferencia $H(k,f)$ estimada en función de la relación señal/ruido en el transcurso del análisis espectral.

Un objetivo principal de la presente invención es mejorar las prestaciones de los métodos de reducción de ruido.

La invención propone de este modo un procedimiento según la reivindicación 1 de reducción de ruido en sucesivas tramas de una señal de entrada, que incluye las siguientes etapas para por lo menos algunas de las tramas:

- calcular un espectro de la señal de entrada mediante transformación en el ámbito frecuencial;
- obtener un estimador de nivel de ruido en función de la frecuencia;
- calcular un primer estimador de nivel de señal útil en la trama en función de la frecuencia;
- calcular la función de transferencia de un primer filtro reductor de ruido sobre la base del primer estimador de nivel de señal útil y del estimador de nivel de ruido;
- calcular un segundo estimador de nivel de señal útil en la trama en función de la frecuencia, combinando el espectro de la señal de entrada y la función de transferencia del primer filtro reductor de ruido;
- calcular la función de transferencia de un segundo filtro reductor de ruido sobre la base del segundo estimador de nivel de señal útil y del estimador de nivel de ruido; y
- utilizar la función de transferencia del segundo filtro reductor de ruido en una operación de filtrado de la trama para producir una señal con ruido reducido.

Los niveles de ruido y de señal útil que se estiman son típicamente DSP, o más generalmente magnitudes en correlación con dichas DSP.

El cálculo en dos pasos, cuya particularidad reside en una actualización más rápida de la DSP de la señal útil $\gamma_{ss}(k,f)$, proporciona al segundo filtro reductor de ruido dos ventajas apreciables con relación a los métodos anteriores. Por una parte, se consigue un seguimiento más rápido de las no estacionalidades de la señal útil, especialmente durante las variaciones rápidas de su envuelta temporal (por ejemplo de los ataques o de las extinciones para señal de voz durante una transición silencio/voz). Por otra parte, se estima mejor el filtro reductor de ruido, lo que se traduce en un incremento de las prestaciones del método (mayor reducción de ruido y menor degradación de la señal útil).

El procedimiento se puede generalizar a los casos en que se efectúan más de dos pasos. A partir de la p -ésima función de transferencia obtenida ($p \geq 2$), se vuelve a calcular entonces el estimador de nivel de señal útil, y se vuelve a evaluar una $(p+1)$ -ésima función de transferencia para la reducción de ruido. La definición anterior del procedimiento vale asimismo para aquellos casos en que se realizan $P > 2$ pasos: basta considerar que el "primer estimador de nivel de señal útil" según dicha definición es el obtenido en el transcurso del $(P-1)$ -ésimo paso. En la práctica, se observan buenas prestaciones del procedimiento con $P = 2$.

En una realización ventajosa del procedimiento, el cálculo del espectro incluye una ponderación de la trama de señal de entrada mediante una función de ventaneo y una transformación de la trama ponderada hacia el ámbito frecuencial, siendo asimétrica la función de ventaneo de manera a aplicar una ponderación más fuerte a la mitad más reciente de la trama que a la mitad menos reciente de la trama.

La elección de dicha función de ventaneo permite concentrar el peso de la estimación espectral hacia las muestras más recientes, permitiendo al mismo tiempo una ventana con buenas propiedades espectrales (subida controlada de los lóbulos secundarios). Esto permite seguir con rapidez las variaciones de la señal. Cabe subrayar que este modo de

ES 2 347 760 T3

cálculo del espectro para el análisis en frecuencia se puede asimismo aplicar cuando la estimación de la función de transferencia del filtro reductor de ruido se realiza en un único paso.

5 El procedimiento se puede utilizar cuando el filtrado de la señal de entrada se efectúa por bloques en el ámbito frecuencial, mediante los métodos ya citados de atenuación espectral a corto plazo. La señal a la que se ha retirado el ruido se produce entonces en la forma de sus componentes espectrales $\hat{S}(k,f)$, que pueden ser explotadas directamente (por ejemplo en una aplicación de codificación o de reconocimiento de voz) o transformadas hacia el ámbito temporal para obtener explícitamente la señal $\hat{s}(n)$.

10 Sin embargo, en un modo de realización preferido del procedimiento, se determina una respuesta impulsional de filtro reductor de ruido para la trama actual a partir de una transformación hacia el ámbito temporal de la función de transferencia del segundo filtro reductor de ruido, y se realiza la operación de filtrado de la trama en el ámbito temporal por medio de la respuesta impulsional determinada para dicha trama.

15 Ventajosamente, la determinación de la respuesta impulsional de filtro reductor de ruido para la trama actual incluye entonces las siguientes etapas:

- transformar hacia el ámbito temporal la función de transferencia del segundo filtro reductor de ruido para obtener una primera respuesta impulsional; y

20 - truncar la primera respuesta impulsional a una longitud de truncado correspondiente a un número de muestras sensiblemente menor (típicamente por lo menos cinco veces menor) que el número de puntos de la transformación hacia el ámbito temporal.

25 Esta limitación del soporte temporal del filtro reductor de ruido presenta una doble ventaja. Por una parte, permite evitar los problemas de repliegue temporal (respeto de la convolución lineal). Por otra parte, asegura un alisado que permite evitar los efectos de un filtro demasiado agresivo, que podrían degradar la señal útil. Puede acompañarse de una ponderación de la respuesta impulsional truncada mediante una función de ventaneo en un número de muestras correspondiente a la longitud de truncado. Cabe subrayar que esta limitación del soporte temporal del filtro es asimismo aplicable cuando la estimación de la función de transferencia se realiza en un único paso.

30 Cuando se realiza el filtrado en el ámbito temporal, es ventajoso subdividir la trama actual en varias subtramas y calcular para cada subtrama una respuesta impulsional interpolada a partir de la respuesta impulsional de filtro reductor de ruido determinada para la trama actual y de la respuesta impulsional de filtro reductor de ruido determinada para por lo menos una trama anterior. La operación de filtrado de la trama incluye entonces un filtrado de la señal de cada subtrama en el ámbito temporal, de conformidad con la respuesta impulsional interpolada calculada para dicha subtrama.

40 Este procesamiento en subtramas permite aplicar un filtro reductor de ruido que varía dentro de una misma trama y, por lo tanto, bien adaptado a las no estacionalidades de la señal procesada. En el caso del procesamiento de una señal de voz, esto se observa especialmente en las tramas mixtas, es decir que presentan sonidos sonoros y sordos. Cabe subrayar que este procesamiento en subtramas se puede asimismo aplicar cuando la estimación de la función de transferencia del filtro se realiza en un único paso.

45 Otro aspecto de la presente invención se refiere a un dispositivo, según la reivindicación 10, de reducción de ruido adecuado para la aplicación del procedimiento anterior.

Otras particularidades y ventajas de la presente invención aparecerán en la siguiente descripción de ejemplos de realización no limitativos, en referencia a los dibujos adjuntos, en los cuales:

50 - la figura 1 muestra un esquema sinóptico de un dispositivo de reducción de ruido apto para la aplicación del procedimiento de la invención;

55 - la figura 2 muestra un esquema sinóptico de una unidad de estimación de la función de transferencia de un filtro reductor de ruido utilizable en un dispositivo según la figura 1;

- la figura 3 muestra un esquema sinóptico de una unidad de filtrado temporal utilizable en un dispositivo según la figura 1; y

60 - la figura 4 muestra un gráfico de una función de ventaneo utilizable en una realización particular del procedimiento.

65 Las figuras 1 a 3 muestran una representación de un dispositivo según la invención en forma de unidades distintas. En una implementación típica del procedimiento, las operaciones de procesamiento de señal se efectúan, como es usual, mediante un procesador de señal digital que ejecuta programas cuyos distintos módulos funcionales corresponden a dichas unidades.

En referencia a la figura 1, un dispositivo de reducción de ruido según la invención incluye una unidad 1 que distribuye la señal de entrada $x(n)$, como una señal de audio digital, en sucesivas tramas de longitud L muestras (indexadas por un entero k). Cada trama de índice k es ponderada (multiplicador 2) multiplicándola por una función de ventaneo $w(n)$, lo que produce la señal $x_w(k,n) = w(n).x(k,n)$ para $0 \leq n < L$.

El paso en el ámbito frecuencial se realiza aplicando la transformada de Fourier discreta (TFD) a las tramas ponderadas $x_w(k,n)$ por medio de una unidad 3 que proporciona la transformada de Fourier $X(k,f)$ de la trama actual.

Para los pasos en los ámbitos tiempo-frecuencia y viceversa empleados en la invención, la TFD de la transformada inversa hacia el ámbito temporal (TFDI) utilizada aguas abajo en su caso (unidad 7), es ventajosamente una transformada de Fourier rápida (o FFT, "Fast Fourier Transform") y una transformada de Fourier inversa rápida (o IFFT, "Inverse Fast Fourier Transform") respectivamente. También se pueden utilizar otras transformaciones tiempo-frecuencia, como la transformada de ondillas.

Una unidad 4 de detección de actividad vocal (DAV) permite discriminar las tramas sólo de ruido y las tramas de voz, y proporciona una indicación binaria de actividad de voz δ para la trama actual. Se puede utilizar cualquier método conocido de DAV, tanto si opera en el ámbito temporal sobre la base de la señal $x(k,n)$ o, como se indica en trazo discontinuo, en el ámbito frecuencial sobre la base de la señal $X(k,f)$.

La DAV controla la estimación de la DSP del ruido por la unidad 5. De este modo, para cada trama k_b de "sólo ruido" detectada por la unidad 4 ($\delta = 0$), se estima la densidad espectral de potencia de ruido $\hat{\gamma}_{bb}(k_b, f)$ mediante la siguiente expresión recursiva:

$$\begin{cases} \hat{\gamma}_{bb}(k_b, f) = \alpha(k_b) \cdot \hat{\gamma}_{bb}(k_b - 1, f) + (1 - \alpha(k_b)) |X(k_b, f)|^2 \\ \hat{\gamma}_{bb}(k, f) = \hat{\gamma}_{bb}(k_b, f) \end{cases} \quad (10)$$

donde k_b es bien la trama de ruido actual si $\delta = 0$, bien la última trama de ruido si $\delta = 1$ (k es detectada como trama de señal útil), y $\alpha(k_b)$ es un parámetro de alisado que puede variar con el transcurso del tiempo.

Obsérvese que el modo de cálculo de $\hat{\gamma}_{bb}(k_b, f)$ no se limita a este estimador con alisado exponencial, pudiendo utilizarse por parte de la unidad 5 cualquier otro estimador de DSP.

A partir del espectro $X(k, f)$ de la trama actual y de la estimación de nivel de ruido $\hat{\gamma}_{bb}(k_b, f)$, otra unidad 6 estima la función de transferencia (FT) del filtro reductor de ruido $\hat{H}(k, f)$. La unidad 7 aplica la TFDI a dicha FT para obtener la respuesta impulsional correspondiente $\hat{h}(k, n)$.

Una función de ventaneo $w_{filt}(n)$ se aplica a dicha respuesta impulsional $\hat{h}(k, n)$ mediante un multiplicador 8 con objeto de obtener la respuesta impulsional $\hat{h}_w(k, n)$ del filtro temporal del dispositivo de reducción de ruido. La operación efectuada por la unidad de filtrado 9 para producir la señal temporal a la que se ha quitado el ruido $\hat{s}(n)$ es, en su principio, una convolución de la señal de entrada por la respuesta impulsional $\hat{h}_w(k, n)$ determinada para la trama actual.

La función de ventaneo $w_{filt}(n)$ tiene un soporte netamente más corto que la longitud de una trama. Dicho de otro modo, la respuesta impulsional $\hat{h}(k, n)$ resultante de la TFDI se trunca antes de aplicarle la ponderación mediante la función $w_{filt}(n)$. Preferiblemente, la longitud de truncado L_{filt} , expresada en número de muestras, es por lo menos cinco veces menor que la longitud de la trama. Es típicamente del orden de magnitud de la décima parte de dicha longitud de trama.

Los L_{filt} coeficientes más significativos de la respuesta impulsional son objeto de la ponderación mediante la ventana $w_{filt}(n)$, que es por ejemplo una ventana de Hamming o de Hanning de longitud L_{filt} para

$$\hat{h}_w(k, n) = w_{filt}(n) \cdot \hat{h}(k, n) \quad \text{para} \quad 0 \leq n < L_{filt} \quad (11)$$

La limitación del soporte temporal del filtro reductor de ruido permite evitar los problemas de repliegue temporal, con objeto de respetar la convolución lineal. Además, asegura un alisado que permite evitar los efectos de un filtro demasiado agresivo que podrían degradar la señal útil.

La figura 2 ilustra una organización preferida de la unidad 6 de estimación de la función de transferencia $\hat{H}(k, f)$ del filtro reductor de ruido, que depende de la DSP del ruido $b(n)$ y de la de la señal útil $s(n)$.

Se ha descrito cómo la unidad 5 puede estimar la DSP del ruido $\hat{\gamma}_{bb}(k,f)$. Pero la DSP $\hat{\gamma}_{ss}(k,f)$ de la señal útil no puede obtenerse directamente debido a la mezcla de la señal y del ruido durante los períodos de actividad vocal. Para preestimarla, el módulo 11 de la unidad 6 de la figura 2 utiliza por ejemplo un estimador de decisión dirigida (véase Y. Ephraim, D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator", IEEE Trans. on ASSP, vol. 32, n° 6, pp. 1109-1121, 1984), de conformidad con la siguiente expresión:

$$\hat{\gamma}_{ss1}(k,f) = \beta(k) |\hat{S}(k-1,f)|^2 + (1 - \beta(k)) P[|X(k,f)|^2 - \hat{\gamma}_{bb}(k,f)] \quad (12)$$

donde $\beta(k)$ es un parámetro baricéntrico que puede variar en el transcurso del tiempo y $\hat{S}(k-1,f)$ es el espectro de la señal útil estimada con relación a la trama anterior de índice $k-1$ (por ejemplo $\hat{S}(k-1,f) = \hat{H}(k-1,f).X(k-1,f)$, obtenido mediante el multiplicador 12 de la figura 2). La función P asegura el umbralado de la magnitud $|X(k,f)|^2 - \hat{\gamma}_{bb}(k,f)$ que corre el riesgo de ser negativa en caso de error de estimación. Se da mediante:

$$P[z(k,f)] = \begin{cases} z(k,f) & \text{si } z(k,f) > 0 \\ \hat{\gamma}_{bb}(k,f) & \text{si no} \end{cases} \quad (13)$$

Cabe subrayar que el cálculo de $\hat{\gamma}_{ss1}(k,f)$ no se limita a este estimador de decisión dirigida. En efecto, se puede utilizar un estimador con alisado exponencial o cualquier otro estimador de densidad espectral de potencia.

Mediante el módulo 13, se calcula una preestimación de la FT del filtro reductor de ruido para la trama actual, como una función de las DSP estimadas $\hat{\gamma}_{ss1}(k,f)$ y $\hat{\gamma}_{bb}(k,f)$:

$$\hat{H}_1(k,f) = F(\hat{\gamma}_{ss1}(k,f), \hat{\gamma}_{bb}(k,f)) \quad (14)$$

Este módulo 13 puede emplear especialmente la regla de sustracción espectral en potencia ($F(y,z) = \sqrt{\frac{y}{y+z}}$,

según (3)), en amplitud ($F(y,z) = 1 - \sqrt{\frac{z}{y+z}}$, según (4)), o también la del filtro de Wiener en bucle abierto

($F(y,z) = \frac{y}{y+z}$, según (5)).

Habitualmente, la función de transferencia final del filtro reductor de ruido se obtiene mediante la ecuación (14). Para mejorar las prestaciones del filtro, se propone estimarlo mediante un procedimiento iterativo en dos pasos. El primer paso consiste en las operaciones efectuadas por los módulos 11 a 13.

La función de transferencia $\hat{H}_1(k,f)$ así obtenida se reutiliza para afinar la estimación de la DSP de la señal útil. La unidad 6 (multiplicador 14 y módulo 15) calcula para ello la magnitud $\hat{\gamma}_{ss}(k,f)$ dada por:

$$\hat{\gamma}_{ss}(k,f) = |\hat{H}_1(k,f).X(k,f)|^2 \quad (15)$$

ES 2 347 760 T3

El segundo paso consiste entonces, para el módulo 16, en calcular el estimador final $\hat{H}(k, f)$ de la función de transferencia del filtro reductor de ruido a partir de la estimación afinada de la DSP de la señal útil:

$$\hat{H}(k, f) = F(\hat{\gamma}_{ss}(k, f), \hat{\gamma}_{bb}(k, f)) \quad (16)$$

pudiendo ser la función F la misma que la empleada por el módulo 13.

Este cálculo en dos pasos permite una actualización más rápida de la DSP de la señal útil $\hat{\gamma}_{ss}(k, f)$ y una mejor estimación del filtro.

La figura 3 ilustra una organización preferida de la unidad de filtrado temporal 9, basada en una subdivisión de la trama actual en N subtramas, permitiendo así aplicar una función de reducción de ruido capaz de evolucionar dentro de una misma trama de señal.

Un módulo 21 efectúa una interpolación de la respuesta impulsional truncada y ponderada $\hat{h}_w(k, n)$ con objeto de obtener un juego de $N \geq 2$ respuestas impulsionales de filtros de subtramas $\hat{h}_w^{(i)}(k, n)$ para i entre 1 y N .

El filtrado en subtrama se puede implementar por medio de un filtro transversal 23 de longitud L_{filt} cuyos coeficientes $\hat{h}_w^{(i)}(k, n)$ ($0 \leq n < L_{\text{filt}}$, $1 \leq i \leq N$) son presentados en cascada por el selector 22 sobre la base del índice i de la subtrama actual. Las subtramas de la señal a filtrar se obtienen mediante una subdivisión de la trama de entrada $x(k, n)$. El filtro transversal 23 calcula así la señal con ruido reducido $\hat{s}(n)$ mediante convolución de la señal de entrada $x(n)$ con los coeficientes $\hat{h}_w^{(i)}(k, n)$ relativos a la subtrama actual.

Las respuestas $\hat{h}_w^{(i)}(k, n)$ de los filtros de subtramas se pueden calcular mediante el módulo 21 como sumas ponderadas por la respuesta impulsional $\hat{h}_w(k, n)$ determinada para la trama actual y por la $\hat{h}_w(k-1, n)$ determinada para la trama anterior. Cuando las subtramas corresponden a un recorte regular de la trama, la función de mezcla ponderada puede ser especialmente:

$$\hat{h}_w^{(i)}(k, n) = \left(\frac{N-i}{N}\right)\hat{h}_w(k-1, n) + \left(\frac{i}{N}\right)\hat{h}_w(k, n) \quad (17)$$

Se observará que el caso al que se aplica directamente el filtro $\hat{h}_w(k, n)$ corresponde a $N = 1$ (sin subtramas).

Ejemplo 1

Este ejemplo de dispositivo se adapta a una aplicación para la comunicación vocal, especialmente en pre-procesamiento de un codificador de voz de baja velocidad de transferencia.

Se utilizan ventanas no solapables para reducir al máximo teórico el retraso introducido por el procesamiento, ofreciendo al mismo tiempo al utilizador la posibilidad de elegir un ventaneo adaptado a la aplicación. Esto es posible ya que el ventaneo de la señal de entrada del dispositivo no está sometido a una limitación de reconstrucción perfecta.

En dicha aplicación, la función de ventaneo $w(n)$ aplicada por el multiplicador 2 es ventajosamente asimétrica, con objeto de realizar una ponderación más fuerte en la mitad más reciente de la trama que en la mitad menos reciente.

Como se muestra en la figura 4, la ventana de análisis asimétrica $w(n)$ se puede construir por medio de dos medias ventanas de Hanning de distintos tamaños L_1 y L_2 :

$$w(n) = \begin{cases} 0,5 - 0,5 \cdot \cos\left(\frac{\pi n}{L_1}\right) & \text{para } 0 \leq n < L_1 \\ 0,5 + 0,5 \cdot \cos\left(\frac{\pi(n-L_1+1)}{L_2}\right) & \text{para } L_1 \leq n < L_1 + L_2 = L \end{cases} \quad (18)$$

Muchos codificadores de voz para móviles utilizan tramas de longitud 20 ms y funcionan a la frecuencia de muestreo $F_e = 8$ kHz (es decir 160 muestras por trama). En el ejemplo representado en la figura 4, se ha elegido $L = 160$, $L_1 = 120$ y $L_2 = 40$.

La elección de dicha ventana permite concentrar el peso de la estimación espectral hacia las muestras más recientes, garantizando al mismo tiempo una buena ventana espectral. El método propuesto permite dicha elección debido a que no se tiene limitación alguna de reconstrucción perfecta de la señal en la síntesis (señal reconstruida a la salida mediante filtrado temporal).

Para tener una mejor resolución frecuencial, las unidades 3 y 7 utilizan una FFT de longitud $L_{FFT} = 256$. Esta elección está motivada asimismo dado que la FFT es digitalmente óptima cuando se aplica a tramas de longitud en potencia 2. Por lo tanto, es necesario prolongar previamente el bloque ventaneado $x_w(k,n)$ por $L_{FFT} - L = 96$ muestras nulas ("zero-padding"):

$$x_w(k,n) = 0 \quad \text{para} \quad L \leq n < L_{FFT} \quad (19)$$

La detección de actividad vocal utilizada en este ejemplo es un método clásico basado en comparaciones de energías a corto plazo y a largo plazo de la señal. La actualización de la estimación de la densidad espectral de potencia de ruido $\hat{\gamma}_{bb}(k,f)$ se realiza mediante estimación con alisado exponencial, de conformidad con la expresión (10) con $\alpha(k_b) = 0,8553$, correspondiente a una constante de tiempo de 128 ms, considerada suficiente para asegurar un compromiso entre una estimación fiable y un seguimiento de las variaciones temporales de la estadística del ruido.

La FT del filtro de reducción de ruido $\hat{H}_1(k,f)$ se preestima de conformidad con la fórmula (5) (filtro de Wiener en bucle abierto), tras haber preestimado la DSP de la señal útil según el estimador con decisión dirigida definido en (12) con $\beta(k) = 0,98$. Se reutiliza la misma función F por parte del módulo 16 para producir la estimación final $\hat{H}(k,f)$ de la FT.

Debido a que la FT $\hat{H}(k,f)$ es de valores reales, el filtro temporal se vuelve causal mediante:

$$\begin{cases} \hat{h}_{\text{caus}}(k,n) = \hat{h}(k, n+L/2) & \text{para} \quad 0 \leq n < L/2 \\ \hat{h}_{\text{caus}}(k,n) = \hat{h}(k, n-L/2) & \text{para} \quad L/2 \leq n < L \end{cases} \quad (20)$$

A continuación, se seleccionan los $L_{\text{filt}} = 21$ coeficientes de dicho filtro, que se ponderan mediante una ventana de Hanning $w_{\text{filt}}(n)$ de longitud L_{filt} , valor correspondiente a las muestras significativas para esta aplicación:

$$\hat{h}_w(k,n) = w_{\text{filt}}(n) \cdot \hat{h}_{\text{caus}}(k, n + \frac{L}{2} - \frac{L_{\text{filt}}-1}{2}) \quad \text{para} \quad 0 \leq n < L_{\text{filt}} \quad (21)$$

con:

$$w_{\text{filt}}(n) = 0,5 - 0,5 \cdot \cos\left(\frac{2\pi n}{L_{\text{filt}}-1}\right) \quad \text{para} \quad 0 \leq n < L_{\text{filt}} \quad (22)$$

El filtrado temporal se realiza mediante $N = 4$ filtros de subtramas $\hat{h}_w^{(i)}(k,n)$ obtenidos por las funciones de mezcla ponderada dadas por (17). A continuación, estos cuatro filtros se aplican por medio de un filtrado transversal de longitud $L_{\text{filt}} = 21$ a las cuatro subtramas de la señal de entrada $x^{(i)}(k,n)$, obteniéndose éstas mediante una extracción contigua de cuatro subtramas de tamaño $L/4 = 40$ muestras de señal de observación $x(k,n)$:

$$x^{(i)}(k,n) = x(k,n) \quad \text{para} \quad (i-1) \cdot L/N \leq n < i \cdot L/N \quad (22)$$

Ejemplo 2

Este ejemplo de dispositivo se adapta a una aplicación para el reconocimiento de voz robusto (en un entorno ruidoso).

En este ejemplo, se utilizan tramas de análisis de longitud L que presentan solapados mutuos de $L/2$ muestras entre dos tramas sucesivas, y la ventana utilizada es del tipo Hanning:

$$w(n) = 0,5 - 0,5 \cdot \cos\left(\frac{2\pi n}{L-1}\right) \quad \text{para } 0 \leq n < L \quad (23)$$

Se fija la longitud de las tramas en 20 ms, es decir $L = 160$ a la frecuencia de muestreo $F_e = 8$ kHz, y se completan las tramas mediante 96 muestras nulas ("zero-padding") para la FFT.

En este ejemplo, el cálculo de la FT del filtro reductor de ruido se basa en una relación de raíces cuadradas de las densidades espectrales de potencia del ruido $\hat{\gamma}_{bb}(k, f)$ y de la señal útil $\hat{\gamma}_{ss}(k, f)$ y, por consiguiente, en los módulos de

la estimada del ruido $|\hat{B}(k, f)| = \sqrt{\hat{\gamma}_{bb}(k, f)}$ y de la señal útil $|\hat{S}(k, f)| = \sqrt{\hat{\gamma}_{ss}(k, f)}$.

La detección de actividad vocal utilizada en este ejemplo es un método clásico existente basado en comparaciones de energías a corto plazo y a largo plazo de la señal. La actualización de la estimación del módulo de la señal de ruido

$|\hat{B}(k, f)| = \sqrt{\hat{\gamma}_{bb}(k, f)}$ se realiza mediante estimación con alisado exponencial:

$$\begin{cases} |\hat{B}(k_b, f)| = \alpha |\hat{B}(k_b - 1, f)| + (1 - \alpha) |X(k_b, f)| \\ |\hat{B}(k, f)| = |\hat{B}(k_b, f)| \end{cases} \quad (24)$$

donde k_b es la trama de ruido actual o la última trama de ruido (si se detecta k como trama de señal útil). La magnitud de alisado α se elige constante e igual a 0,99, es decir una constante de tiempo de 1,6 s.

La FT del filtro de reducción de ruido $\hat{H}_1(k, f)$ es preestimada por el módulo 13, según:

$$\hat{H}_1(k, f) = F(|\hat{S}(k, f)|, |\hat{B}(k, f)|)$$

con:

$$F(y, z) = \frac{y}{y + z}$$

El hecho de calcular la raíz cuadrada permite efectuar estimaciones sobre los módulos, que están ligados al RSB $\eta(k, f)$ mediante:

$$\eta(k, f) = \frac{|\hat{S}(k, f)|^2}{|\hat{B}(k, f)|^2}$$

El estimador de la señal útil en el módulo $|\hat{S}(k,f)|$ se obtiene mediante: (28)

$$|\hat{S}(k,f)| = \beta |\hat{S}(k-1,f)|^2 + (1 - \beta) P[|X(k,f)| - |\hat{B}(k,f)|] \quad (28)$$

con $\beta(k) = 0,98$.

El multiplicador 14 realiza el producto de la FT preestimada $\hat{H}_1(k,f)$ por el espectro $X(k,f)$, y el módulo del resultado (y no su cuadrado) se obtiene en 15 para proporcionar la estimación afinada de $|\hat{S}(k,f)|$, a partir de la cual el módulo 16 produce la estimación final $\hat{H}(k,f)$ de la FT, utilizando la misma función F que en (25).

A continuación, se obtiene la respuesta temporal $\hat{h}_w(k,n)$ exactamente de la misma manera que en el ejemplo 1 (paso a ámbito temporal, restitución de la causalidad, selección de las muestras significativas y ventaneo). La única diferencia reside en la elección del número de coeficientes seleccionado L_{filt} , que se fija en $L_{\text{filt}} = 17$ en este ejemplo.

El filtrado de la trama de entrada $x(k,n)$ se realiza aplicándole directamente la respuesta temporal del filtro de reducción de ruido obtenida $\hat{h}_w(k,n)$. El hecho de no realizar el filtrado en subtramas equivale a tomar $N = 1$ en la expresión (17).

REIVINDICACIONES

1. Procedimiento de reducción de ruido en sucesivas tramas de una señal de entrada $x(n)$, que incluye las siguientes etapas para por lo menos algunas de las tramas:

- calcular un espectro $X(k,f)$ de la señal de entrada mediante transformación en el ámbito frecuencial;
- obtener un estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$ del espectro $X(k,f)$ en función de la frecuencia;
- calcular un primer estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ en la trama en función de la frecuencia;
- calcular la función de transferencia $\hat{H}_1(k,f)$ de un primer filtro reductor de ruido sobre la base del primer estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ y del estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$;
- calcular un segundo estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ en la trama en función de la frecuencia, combinando el espectro $X(k,f)$ de la señal de entrada y la función de transferencia $\hat{H}_1(k,f)$ del primer filtro reductor de ruido;
- calcular la función de transferencia $\hat{H}(k,f)$ de un segundo filtro reductor de ruido sobre la base del segundo estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ y del estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$; y
- utilizar la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido en una operación de filtrado de la trama para producir una señal con ruido reducido.

2. Procedimiento según la reivindicación 1, en el que el cálculo del espectro $X(k,f)$ incluye una ponderación de la trama de señal de entrada mediante una función de ventaneo $w(n)$ y una transformación de la trama ponderada hacia el ámbito frecuencial, siendo la función de ventaneo asimétrica de manera a aplicar una ponderación más fuerte a la mitad más reciente de la trama que a la mitad menos reciente de la trama.

3. Procedimiento según la reivindicación 1 o 2, en el que se determina una respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ para la trama actual a partir de una transformación hacia el ámbito temporal de la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido, y se realiza la operación de filtrado de la trama en el ámbito temporal por medio de la respuesta impulsional determinada para dicha trama.

4. Procedimiento según la reivindicación 3, en el que la determinación de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ para la trama actual incluye las siguientes etapas:

- transformar hacia el ámbito temporal la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido para obtener una primera respuesta impulsional; y
- truncar la primera respuesta impulsional a una longitud de truncado correspondiente a un número de muestras sensiblemente menor que el número de puntos de la transformación hacia el ámbito temporal.

5. Procedimiento según la reivindicación 4, en el que la determinación de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ para la trama actual incluye además la siguiente etapa:

- ponderar la respuesta impulsional truncada mediante una función de ventaneo $w_{filt}(n)$ en un número de muestras correspondiente a dicha longitud de truncado.

6. Procedimiento según una cualquiera de las reivindicaciones 3 a 5, en el que se subdivide la trama actual en varias subtramas y se calcula para cada subtrama una respuesta impulsional interpolada $\hat{h}_w^{(i)}(k,n)$ a partir de una respuesta impulsional de filtro reductor de ruido determinada para la trama actual y de la respuesta impulsional de filtro reductor de ruido determinada para por lo menos una trama anterior, y en el que la operación de filtrado de la trama incluye un filtrado de la señal de cada subtrama en el ámbito temporal, de conformidad con la respuesta impulsional interpolada calculada para dicha subtrama.

7. Procedimiento según la reivindicación 6, en el que las respuestas impulsionales interpoladas $\hat{h}_w^{(i)}(k,n)$ se calculan para las distintas subtramas de la trama actual como sumas ponderadas de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ determinada para la trama actual y de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k-1,n)$ determinada para la trama anterior.

8. Procedimiento según la reivindicación 7, en el que la respuesta impulsional interpolada $\hat{h}_w^{(i)}(k,n)$ calculada para la i-ésima subtrama de la trama actual ($1 \leq i < N$) es igual a $(N-i)/N$ veces la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k-1,n)$ determinada para la trama anterior más i/N veces la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ determinada para la trama actual, siendo N el número de subtramas de la trama actual.

9. Procedimiento según una cualquiera de las reivindicaciones anteriores, en el que la señal de entrada $x(n)$ es una señal de audio.

10. Dispositivo de reducción de ruido en una señal de entrada $x(n)$ que incluye:

- medios (1-3) de cálculo de un espectro $X(k,f)$ de una trama de la señal de entrada mediante transformación en el ámbito frecuencial;

- medios (5) de obtención de un estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$ del espectro $X(k,f)$ en función de la frecuencia;

- medios (11) de cálculo de un primer estimador de nivel de señal útil $\hat{\gamma}_{ss1}(k,f)$ en la trama en función de la frecuencia;

- medios (13) de cálculo de la función de transferencia $\hat{H}_1(k,f)$ de un primer filtro reductor de ruido sobre la base del primer estimador de nivel de señal útil $\hat{\gamma}_{ss1}(k,f)$ y del estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$;

- medios (14-15) de cálculo de un segundo estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ en la trama en función de la frecuencia, combinando el espectro $X(k,f)$ de la señal de entrada y la función de transferencia $\hat{H}_1(k,f)$ del primer filtro reductor de ruido;

- medios (16) de cálculo de la función de transferencia $\hat{H}(k,f)$ de un segundo filtro reductor de ruido sobre la base del segundo estimador de nivel de señal útil $\hat{\gamma}_{ss}(k,f)$ y del estimador de nivel de ruido $\hat{\gamma}_{bb}(k,f)$; y

- medios (7-9) de filtrado de la trama por medio de la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido para producir una señal con ruido reducido.

11. Dispositivo según la reivindicación 10, en el que los medios de cálculo de espectro incluyen medios (2) de ponderación de la trama de señal de entrada $x(n)$ mediante una función de ventaneo $w(n)$, y medios (3) de transformación de la trama ponderada hacia el ámbito frecuencial, siendo la función de ventaneo asimétrica de manera a aplicar una ponderación más fuerte a la mitad más reciente de la trama que a la mitad menos reciente de la trama.

12. Dispositivo según la reivindicación 10 u 11, que incluye medios (7-8) de determinación de una respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ para la trama actual a partir de una transformación hacia el ámbito temporal de la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido, en el que los medios de filtrado (9) operan en el ámbito temporal por medio de la respuesta impulsional determinada para la trama actual.

13. Dispositivo según la reivindicación 12, en el que los medios de determinación de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ incluyen medios (7) de transformación hacia el ámbito temporal de la función de transferencia $\hat{H}(k,f)$ del segundo filtro reductor de ruido, para obtener una primera respuesta impulsional, y medios (8) de truncado de la primera respuesta impulsional a una longitud de truncado correspondiente a un número de muestras sensiblemente menor que el número de puntos de la transformación hacia el ámbito temporal.

14. Dispositivo según la reivindicación 13, en el que los medios de determinación de la respuesta impulsional de filtro reductor de ruido incluyen medios (8) de ponderación de la respuesta impulsional truncada mediante una función de ventaneo $w_{fil}(n)$ en un número de muestras correspondiente a dicha longitud de truncado.

15. Dispositivo según una cualquiera de las reivindicaciones 12 a 14, que incluye medios para subdividir la trama actual en varias subtramas, y medios (21) de cálculo de una respuesta impulsional interpolada $\hat{h}_w^{(i)}(k,n)$ para cada subtrama a partir de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ determinada para la trama actual y de la respuesta impulsional de filtro reductor de ruido determinada para por lo menos una trama anterior, en el que los medios de filtrado (9) incluyen un filtro (23) para filtrar la señal de cada subtrama en el ámbito temporal, de conformidad con la respuesta impulsional interpolada calculada para dicha subtrama.

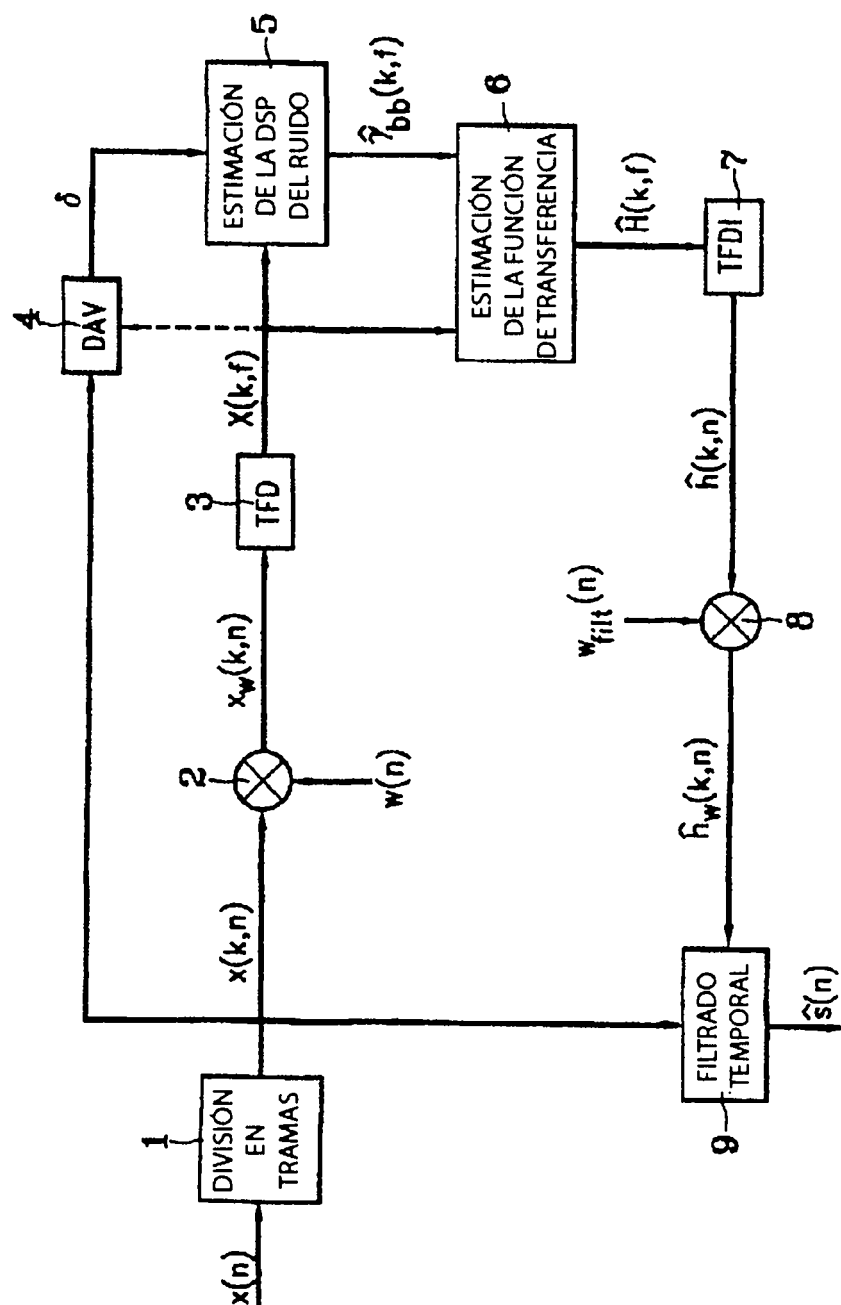
16. Dispositivo según la reivindicación 15, en el que los medios de cálculo de respuesta impulsional interpolada están dispuestos para calcular las respuestas impulsionales interpoladas $\hat{h}_w^{(i)}(k,n)$ para las distintas subtramas de la trama actual como sumas ponderadas de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ determinada para la trama actual y de la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k-1,n)$ determinada para la trama anterior.

ES 2 347 760 T3

17. Dispositivo según la reivindicación 16, en el que la respuesta impulsional interpolada $\hat{h}_w^{(i)}(k,n)$ calculada para la i-ésima subtrama de la trama actual ($1 \leq i \leq N$) es igual a $(N-i)/N$ veces la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k-1,n)$ determinada para la trama anterior más i/N veces la respuesta impulsional de filtro reductor de ruido $\hat{h}_w(k,n)$ determinada para la trama actual, siendo N el número de subtramas de la trama actual.

18. Dispositivo según una cualquiera de las reivindicaciones 10 a 17, en el que la señal de entrada $x(n)$ es una señal de audio.

FIG. 1



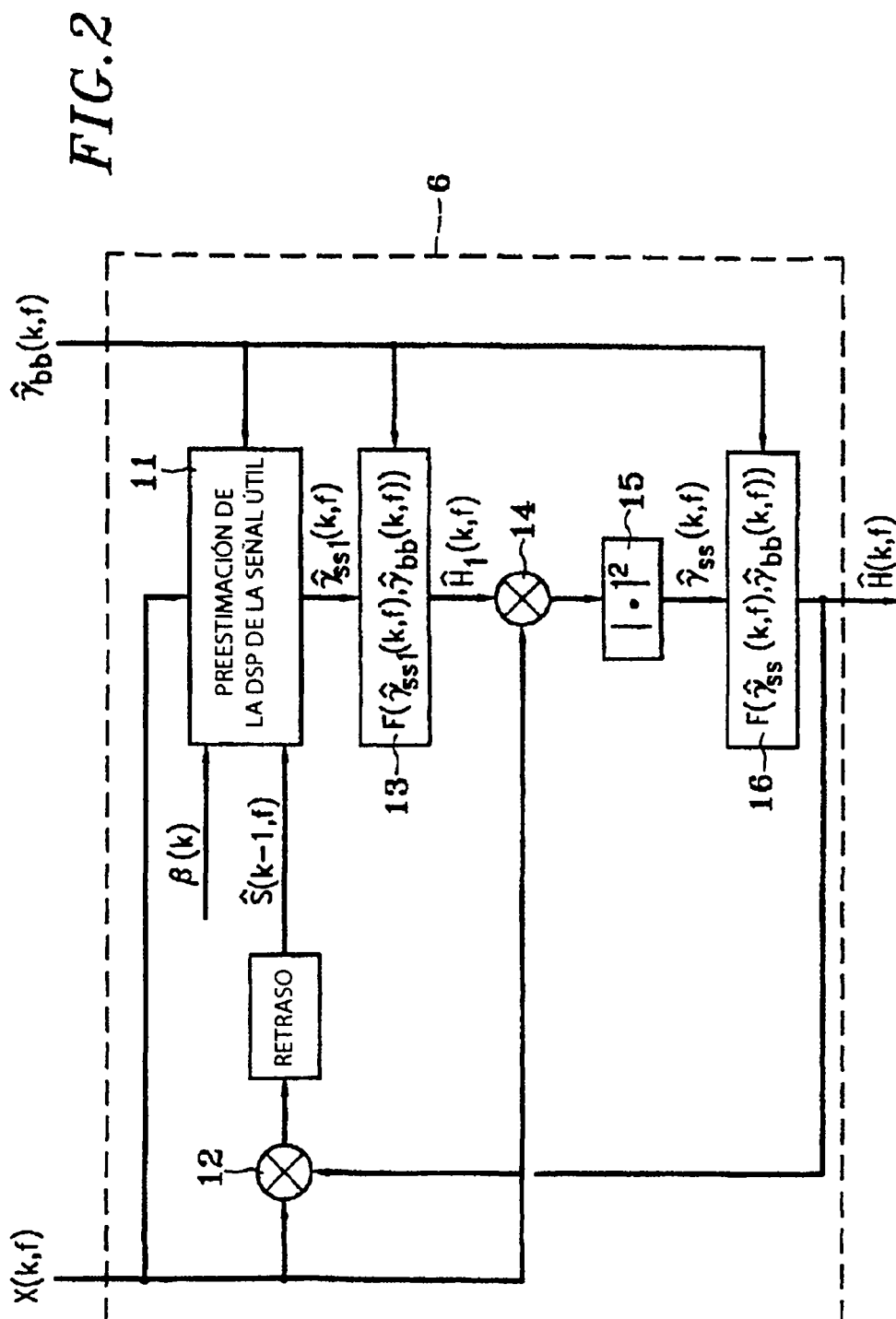


FIG.3

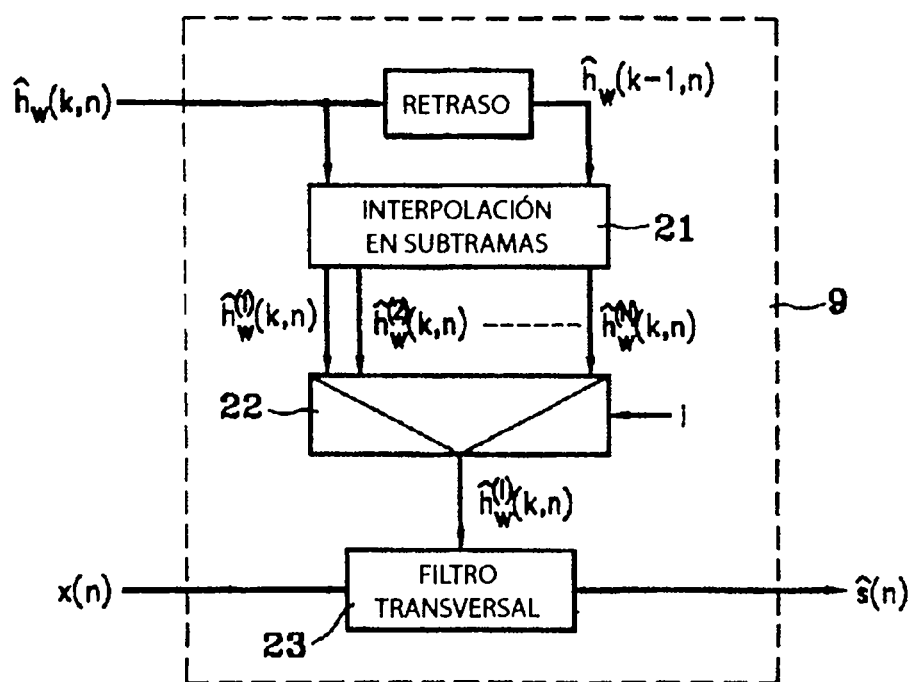


FIG.4

