

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 598 113**

51 Int. Cl.:

**G10L 19/26** (2013.01)

**G10L 21/0364** (2013.01)

**G10L 19/008** (2013.01)

**G10L 25/90** (2013.01)

**G10L 19/107** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.12.2007 PCT/SE2007/051077**

87 Fecha y número de publicación internacional: **31.12.2008 WO09002245**

96 Fecha de presentación y número de la solicitud europea: **21.12.2007 E 07861172 (0)**

97 Fecha y número de publicación de la concesión europea: **10.08.2016 EP 2171712**

54 Título: **Método y disposición para mejorar señales de audio espaciales**

30 Prioridad:

**27.06.2007 US 929440 P**

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

**25.01.2017**

73 Titular/es:

**TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)  
(100.0%)**

**164 83 Stockholm, SE**

72 Inventor/es:

**KARLSSON, ERLENDUR y  
DE BACHTIN, SEBASTIAN**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

ES 2 598 113 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

Método y disposición para mejorar señales de audio espaciales

**Campo técnico**

5 La presente invención se relaciona con señales de audio grabadas en estéreo y espaciales en general, y específicamente con métodos y disposiciones para mejorar esas señales en una aplicación de teleconferencia.

**Antecedentes**

10 Unas pocas horas de reunión cara a cara entre partes ubicadas en lugares geográficos diferentes ha demostrado ser un modo muy eficaz de construir relaciones de negocios duraderas, poner a un grupo de proyecto al día, intercambiar ideas e información y mucho más. La desventaja de esas reuniones es el gran gasto que resulta de viajar y posiblemente incluso alojarse durante la noche, lo que a menudo hace que estas reuniones sean demasiado caras y engorrosas de organizar. Mucho se ganaría si una reunión se pudiera organizar de forma que cada parte pudiera participar en la reunión desde su propia localización geográfica y las diferentes partes pudieran comunicarse tan fácilmente entre ellas como si estuvieran todas reunidas juntas en una reunión cara a cara. Esta visión de la telepresencia ha insuflado nueva vida a la investigación y desarrollo de sistemas de video-teleconferencia, donde se están poniendo grandes esfuerzos en el desarrollo de métodos para crear una conciencia espacial percibida que se parezca a la de una reunión cara a cara real.

15 Un factor importante de una conversación en la vida real es la capacidad de la especie humana de ubicar a los participantes usando solo la información del sonido. El audio espacial, que es explicado con mayor detalle más abajo, es el sonido el que contiene pistas binaurales, y esas pistas son usadas para ubicar la fuente del sonido. En una teleconferencia que use audio espacial, es posible disponer a los participantes en una sala de reuniones virtual, donde la voz de cada participante es percibida como si se originara desde una dirección específica. Cuando un participante puede ubicar a otros participantes en la imagen en estéreo, es más fácil centrarse en cierta voz y determinar quién está diciendo qué.

20 En una aplicación de teleconferencia que soporta audio espacial, un puente de conferencia en la red es capaz de dar un renderizado de audio espacializado (3D) de una sala de reuniones virtual a cada uno de los participantes. La espacialización mejora la percepción de una reunión cara a cara y permite a cada participante ubicar a los otros participantes en puestos diferentes en el espacio de audio virtual renderizado alrededor de él/ella, lo que de nuevo hace más fácil al participante mantener el seguimiento de quién está diciendo qué.

25 Se puede crear una teleconferencia de muchas maneras diferentes. Uno puede escuchar la conversación a través de auriculares o altavoces usando señales estéreo o mono. El sonido puede ser obtenido mediante un micrófono usando o señales estéreo o mono. El micrófono estéreo se puede usar cuando varios participantes están en la misma sala física y la imagen estéreo en la sala debe ser transferida a los otros participantes situados en otro lugar. Las personas sentadas a la izquierda son percibidas como ubicadas a la izquierda en la imagen estéreo. Si la señal del micrófono es en mono entonces la señal puede ser transformada a una señal en estéreo, donde el sonido mono se sitúa en una imagen en estéreo. El sonido será percibido como teniendo una situación en la imagen estéreo, usando un renderizado del audio espacializado de una sala de reuniones virtual.

30 Para participantes de un terminal multimedia avanzado el renderizado espacial se puede hacer en el terminal, mientras que para participantes con terminales más simples el renderizado debe ser realizado por la aplicación de conferencia en la red y entregado al usuario final como una señal en estéreo codificada binaural. Para ese caso particular, sería beneficioso si descodificadores de voz estándar que ya están disponibles en los terminales estándar pudieran ser usados para descodificar la señal binaural codificada.

35 Un códec de interés particular es el así llamado Predicción Lineal Excitada por Código Algebraico (ACELP) basado en el codificador de Banda Ancha de Multi Tasa Adaptativa (AMR-WB) [1-2]. Es un descodificador-mono, pero podría ser usado potencialmente para codificar los canales izquierdo y derecho de la señal en estéreo independientemente cada una de la otra.

40 Audiometrías de teleconferencia codificadas con AMR-WB relacionada con grabaciones en estéreo y señales binaurales renderizadas sintéticamente han mostrado que el códec a menudo introduce aberraciones de codificación que son bastante molestas y distorsionan la imagen espacial de la señal de sonido. El problema es más grave para los modos de operación con una tasa baja de bits, como 12.65 kbs, pero incluso ha sido encontrado en modos de operación a mayores tasas de bits. La señal de voz de estéreo es codificada con un codificador de voz mono donde los canales izquierdo y derecho son codificados por separado. Es importante que el codificador preserve las pistas binaurales necesarias para ubicar el sonido. Cuando sonidos estéreo son codificados de esta forma, extrañas aberraciones pueden ser oídas durante la escucha simultánea de ambos canales. Cuando los canales izquierdo y derecho son reproducidos separadamente, las aberraciones no son tan molestas. Las aberraciones pueden ser explicadas como ruido espacial, porque el ruido no es percibido dentro de la cabeza. Es más difícil decidir donde se origina el ruido espacial desde una imagen de estéreo, que es molesta de escuchar para el usuario.

Una escucha más minuciosa del material codificado con AMR-WB ha revelado que los problemas surgen principalmente cuando hay una vocal fuerte aguda en la señal o cuando hay dos o más vocales simultáneas en la señal y el codificador tiene problemas estimando la frecuencia principal del tono. Más análisis de señal han revelado también que la mayor parte de la distorsión de señal anteriormente mencionada está situada en la zona de baja frecuencia desde 0 Hz hasta justo debajo del tono de frecuencia más bajo en la señal.

Si el códec de AMR -WB ha de usarse como se ha descrito anteriormente, es necesario mejorar la señal codificada en el intervalo de frecuencias bajas descrito anteriormente.

Voiceage Corporation ha desarrollado una mejora de tono selectivo en frecuencia de voz sintetizada [3-4]. Sin embargo, las audiometrías han revelado que el método no consigue mejorar las señales codificadas satisfactoriamente, pues la mayor parte de la distorsión todavía podía ser oída. Recientes análisis de señal del método han mostrado que solo mejora el intervalo de frecuencias inmediatamente alrededor de la frecuencia de tono más baja y deja intacta la mayor parte de la distorsión, la cual está situada en el intervalo de frecuencias desde 0 Hz hasta justo debajo de la frecuencia de tono más baja.

El documento WO03/102923 describe un método y una disposición para la mejora de tono selectivo en frecuencia de voz sintetizada por el post procesamiento de una señal de sonido descodificada en vista de mejorar la calidad percibida de esta señal de sonido descodificada.

Debido a lo anterior, hay una necesidad de métodos y disposiciones que posibiliten la mejora de señales codificadas con ACELP para reducir el ruido espacial.

### Resumen

Un objetivo general de la presente invención es posibilitar teleconferencias mejoradas.

Otro objetivo de la presente invención es posibilitar mayores mejoras de señales de audio espaciales.

Un objetivo específico de la presente invención es posibilitar mayores mejoras de señales espaciales codificadas con ACELP en un sistema de teleconferencia.

Básicamente, la presente invención describe un método de mejora de señales de audio espaciales recibidas, por ejemplo señales de audio codificadas con ACELP en un sistema de teleconferencia. Inicialmente es recibida (S10) una señal de audio codificada con ACELP que comprende una pluralidad de bloques. Para cada bloque un tipo de señal es estimado (S20) basándose en la señal recibida. Además, para cada bloque una frecuencia de tono es estimada (S30) basándose en la señal recibida. Subsecuentemente, los parámetros de filtrado son determinados (S40) basándose en al menos uno de los tipos de señal estimados y dicha frecuencia de tono estimada. Finalmente, la señal recibida es filtrada en paso alto (S50) basándose en los parámetros de filtro determinados para proporcionar una señal de salida filtrada en paso alto.

Además, todos los canales de una señal de audio multicanal están sujetos a los pasos de estimación y subsecuentemente determinación S41 de parámetros de filtro conjuntos para los canales. Finalmente, todos los canales son filtrados en paso alto usando los mismos parámetros de filtro conjuntos. La invención se expone mediante las reivindicaciones anexas.

Las ventajas de la presente invención comprenden:

Señales de audio espaciales mejoradas.

Señales de audio espaciales con ruido espacial reducido.

Sesiones de teleconferencia mejoradas.

### 40 Breve descripción de los dibujos

La invención, junto con otros objetos y ventajas de la misma, puede ser entendida mejor refiriéndose a la siguiente descripción tomada junto a los dibujos que se acompañan, en los cuales:

Fig. 1 es un diagrama de flujo esquemático de una realización de la presente invención;

Fig. 2 es un diagrama de flujo esquemático de otra realización de la presente invención;

45 Fig. 3a es un diagrama de bloques esquemático de una disposición de acuerdo a la presente invención;

Fig. 3b es un diagrama de bloques esquemático de una disposición de acuerdo a la presente invención;

Fig. 4 es un diagrama de una comparación entre la mejora de acuerdo a la presente invención y el conocido test de MUSHRA para una señal con distorsiones;

Fig. 5 es un diagrama de una comparación entre la mejora de acuerdo a la presente invención y el conocido test de MUSHRA para una señal sin distorsiones.

### Abreviaturas

	ACELP	Predicción Lineal Excitada por Código Algebraico
5	AMR-WB	Banda Ancha de Multi Tasa Adaptativa
	AMR-WB+	Banda Ancha Extendida de Multi Tasa Adaptativa
	FIR	Respuesta Finita al Impulso
	Hz	Hercio
	IIR	Respuesta Infinita al Impulso
10	MUSHRA	Estímulos Múltiples con Referencia y Anclaje Ocultos
	WB	Banda Ancha
	VMR-WB	Banda Ancha de Multi Modo de Tasa Variable

### Descripción detallada

15 La presente invención será descrita en el contexto de señales codificadas con Predicción Lineal Excitada por Código Algebraico (ACELP) en Banda Ancha de Multi-Tasa Adaptativa (AWR\_WB). Sin embargo, es apreciado que puede ser igualmente aplicado a otros sistemas similares que utilicen ACELP.

20 Cuando los inventores han probado el método Voiceage de la técnica anterior en material relacionado con teleconferencias, el conocido método no ha conseguido mejorar las señales codificadas satisfactoriamente. El análisis de señal del método ha mostrado que solo mejora el intervalo de frecuencias inmediatamente alrededor de frecuencia de tono más baja y deja intacta la mayor parte de la distorsión, la cual está situada en el intervalo de frecuencias desde 0 Hz hasta justo debajo de la frecuencia de tono más baja.

25 Para posibilitar el aumento de la mejora de señales de audio espacial, los inventores han descubierto que es necesario reducir o incluso eliminar la distorsión descrita anteriormente filtrando en paso alto la señal codificada con un filtro de paso alto variable en el tiempo, donde para cada bloque de señal la frecuencia de corte del filtro de paso alto se actualiza en función del tipo de señal estimado y las frecuencias de tono del bloque de señal. En otras palabras, la presente descripción está relacionada generalmente con un método de filtrado de paso alto de una señal espacial con un filtro de paso alto variable en el tiempo de forma que sigue el tono de la señal.

30 Con referencia a la Fig. 1, es recibida S10 una señal de audio, por ejemplo, una señal codificada con ACELP, que se compone de una pluralidad de bloques. Cada bloque de la señal recibida es sometido a un proceso de estimación en el que un tipo de señal S20 es estimado basándose en la señal recibida y/o un conjunto de parámetros del decodificador. Subsecuentemente, o en paralelo, una frecuencia de tono S30 para el bloque es estimada, también basándose en una o ambas de las señales recibidas y los parámetros del decodificador. Basándose en el tono estimado y/o el tipo de señal, un conjunto de parámetros de filtrado S40 son determinados para el bloque. Finalmente, la señal recibida es filtrada en paso alto S50 basándose en los parámetros del filtro determinados para proporcionar una señal de audio de salida filtrada en paso alto.

40 De acuerdo con otra realización, el filtrado en paso alto es posibilitado por medio de uno u opcionalmente una secuencia de filtros (o filtros paralelos). Los filtros potenciales a usar se componen de: filtros de Respuesta Finita al Impulso (FIR), filtros de IIR (Respuesta Infinita al Impulso). Preferiblemente, se usan una pluralidad de filtros de IIR paralelos de tipo elíptico. En una realización preferida, tres filtros de IIR paralelos son usados para posibilitar el proceso de filtrado de paso alto.

45 Específicamente, y con referencia a la Fig. 2, de acuerdo con otra realización de la presente invención, una señal de audio espacial multicanal es aportada o recibida S10. Para cada bloque y canal, el tipo de señal y la frecuencia de tono son determinados o estimados S20, S30. Subsecuentemente, los parámetros del filtro son determinados para cada canal S40 y adicionalmente, los parámetros del filtro conjuntos son determinados S41 para los bloques y los canales. Finalmente, todos los canales de la señal de audio espacial multicanal son filtrados en paso alto (S50) basándose en los parámetros del filtro conjuntos determinados. Un caso especial de la señal multicanal es una señal estéreo con dos canales.

50 El paso de determinar los parámetros del filtro conjuntos S41 es, de acuerdo con una realización específica, posibilitado por la determinación de una frecuencia de corte para cada canal basándose en el tipo de señal y la frecuencia de tono estimados, y formando los parámetros del filtro conjuntos basándose en la frecuencia de corte más baja. También se pueden utilizar otros criterios de frecuencia en el proceso.

De acuerdo con otra posible realización (no mostrada) de la presente invención, los parámetros del filtro son determinados únicamente basándose en el tipo de señal estimado. El paso de estimación del tono S30, en ese caso, comprende el paso adicional de determinar si es necesario añadir la estimación de tono para determinar parámetros del filtro más precisos. Si el paso de determinar revela que ese es el caso, el tono es estimado y los parámetros del filtro son determinados basándose tanto en tipo de señal como en el tono. Si el paso de estimación de tono se considera superfluo, entonces los parámetros del filtro son determinados basándose solo en el tipo de señal.

Con referencia a la Fig. 3a, una realización de una disposición 1 para mejorar señales de audio espaciales de acuerdo con la presente invención se describirá a continuación.

Además de las unidades ilustradas, la disposición 1 puede contener cuantas unidades (no mostradas) sean necesarias para recibir y transmitir señales de audio espaciales. Estas están indicadas por la casilla general de entrada/salida I/O en el dibujo. La disposición 1 comprende una unidad 10 para proporcionar o recibir una señal de audio espacial, estando la señal dispuesta como una pluralidad de bloques. Otra unidad 20 proporciona estimaciones del tipo de señal para cada bloque recibido, basándose en los parámetros del descodificador proporcionados y el bloque de señal recibido. Subsecuentemente, o en paralelo, una unidad de estimación de tono 30 estima la frecuencia de tono del bloque de señal recibido, también basándose en parámetros proporcionados del descodificador y el bloque de señal recibido. Se proporciona una unidad 40 de determinación de parámetros del filtro. La unidad 40 usa el tipo de señal estimada y/o la frecuencia de tono estimada para determinar los parámetros del filtro apropiados para una unidad de filtro de paso alto 50.

De acuerdo a otra realización, la disposición 1 está más adaptada para utilizar las unidades descritas anteriormente para mejorar señales de audio espaciales en estéreo o incluso de multicanal. Para ese caso, las unidades 20, 30 para estimar el tipo de señal y la frecuencia de tono están adaptadas para realizar las estimaciones para cada canal de la señal multicanal. Además, la unidad de filtro 40 (o una unidad de filtro alternativa 41) está adaptada para utilizar los parámetros de filtro respectivos determinados (o directamente el tono y tipo de señal estimados) para determinar los parámetros del filtro conjuntos. Finalmente, el filtro de paso alto 50 es adaptado para filtrar en paso alto todos los múltiples canales de la señal recibida con los mismos parámetros de filtro conjuntos.

Las casillas representadas en la realización de la Fig. 3a pueden ser implementadas en software o igualmente bien en hardware, o una mezcla de ambos.

De acuerdo a otra realización, una disposición de la presente invención consta de un primer bloque en la Fig. 3b que es el clasificador de Señal y estimador de Tono, bloques 20, 30, que, para cada bloque de señal de la señal recibida representada por la señal sintética  $x(n)$ , estiman el tipo de señal y las frecuencias de tono del bloque de señal a partir de un conjunto de parámetros del descodificador además de la propia señal sintética. El bloque 40 de evaluación de parámetros de filtro toma entonces el tipo de señal estimada y las frecuencias de tono y evalúa los parámetros de filtro apropiados para el filtro de paso alto. Finalmente, el bloque 50 del filtro de paso alto variable en el tiempo toma los parámetros del filtro actualizados y realiza el filtrado en paso alto de la señal sintética  $x(n)$ .

En general, el método usará ambos parámetros del descodificador además de la señal sintética cuando se estima el tipo de señal y las frecuencias de tono, pero puede también optar a usar solo uno u otro.

Como la señal de interés es una señal estéreo y el descodificador es un descodificador mono, la clasificación de señal y la estimación del tono se realiza para ambos canales izquierdo y derecho. Sin embargo, como es importante no distorsionar la imagen espacial de la señal estéreo, ambos canales necesitan ser filtrados con el mismo filtro de paso alto variante en el tiempo. El método decide por tanto qué canal requiere la frecuencia de corte más baja (basándose en los parámetros de filtro determinados para cada canal, respectivamente) y usa esa frecuencia de corte cuando evalúa los coeficientes del filtro del filtro de paso alto conjunto que es usado para filtrar ambos canales.

En una realización de la invención, la clasificación del tipo de señal es muy simple. Simplemente determina si el bloque de señal contiene un componente de banda de paso fuerte y estrecha de frecuencia central baja en el típico intervalo de frecuencias de la voz humana, de aproximadamente 100-500 Hz. Si se encuentra un tal componente de banda de paso estrecha la frecuencia central del componente, es estimada como la frecuencia de tono más baja del bloque de señal. La frecuencia de corte del filtro es evaluada justo por debajo de esa frecuencia de tono más baja y los parámetros del filtro para esa frecuencia de corte son evaluados y enviados al filtro de paso alto variable en el tiempo. Cuando no se encuentra componente de banda de paso estrecha la frecuencia de corte, es reducida a 50 Hz.

Para conseguir que este tipo de filtrado de paso alto variante en el tiempo trabaje adecuadamente y para obtener una implementación eficiente de él, hay varias cuestiones de diseño que deben ser cuidadosamente consideradas. Aquí hay una lista de las cuestiones más importantes.

1. El filtro de paso alto debe estar adaptado para suprimir el ruido no deseado bajo la frecuencia de tono más baja sin distorsionar el componente de tono. Esto requiere de una clara transición entre la banda de corte y la banda de paso.

2. El filtrado necesita también ser computado eficazmente, lo que requiere tan pocos parámetros del filtro como sea posible.

3. Para cumplir eficientemente los requisitos 1 y 2, la estructura del así llamado filtro de IIR puede ser elegida de acuerdo a una realización. Probando el método de la invención, se ha establecido que se obtienen resultados razonablemente buenos usando filtros elípticos de sexto orden.

4. La estabilidad del filtrado de IIR variante en el tiempo es un asunto no trivial. Para garantizar la estabilidad los filtros de IIR de sexto orden se pueden descomponer en tres filtros de segundo orden, lo que proporciona control completo sobre los polos de cada filtro de segundo orden y así se garantiza la estabilidad de la operación de filtrado completa.

10 Incluso aunque estas soluciones de diseño de filtros han sido usadas en una realización de la invención, no son de ningún modo restrictivas a la invención. Alguien experto en la técnica reconoce fácilmente que otras estructuras de filtros y mecanismos de control de estabilidad se pueden usar en su lugar.

**Ventajas de la invención**

15 El desempeño de la invención en comparación a señales codificadas no mejoradas y otros métodos de mejora ha sido evaluado a través de una audiometría MUSHRA [5] en dos conjuntos de señales de test. El primer conjunto de señales contenía señales que tenían graves distorsiones de codificación, mientras el segundo conjunto contenía señales sin ninguna distorsión grave. Con el primer conjunto, el objetivo era evaluar cuán grande era el perfeccionamiento que el método de mejora descrito en esta invención estaba proporcionando, mientras que el segundo conjunto de señales fue usado para mostrar si el método de mejora causaba alguna degradación audible a señales que no tenían ninguna distorsión grave de codificación.

Los codificadores y métodos de mejora evaluados en el test están resumidos en la Tabla 1 a continuación.

Tabla 1: Comparación de los métodos de mejora

Señal de Salida	Codificación y mejora
ref	Señal original sin codificar
modo7filt	AMR-WB, 23.05 kbit/s y filtrado de acuerdo con la invención.
modo7	AMR-WB, 23.05 kbit/s.
modo2filt	AMR-WB, 12.65 kbit/s y filtrado de acuerdo con la invención.
modo2	AMR-WB, 12.65 kbit/s.
bpf2	AMR-WB, 12.65 kbit/s y filtrado con el mejorador de tono de Voiceage.
wb+	AMR-WB+, 13.6 kbit/s, con un marco fijo de 20 ms. El AMR-WB+ fue forzado a solo codificar en modo ACELP [6].
vmr	VMR-WB, 12.65 kbit/s [7]
ancla	Señal sin codificar original que es filtrada en paso bajo a 3.5 kHz.

25 Los resultados del test MUSHRA se dan en las Fig. 4 y Fig. 5. La Fig. 4 muestra los resultados para un conjunto de señales con graves distorsiones de codificación, mientras la Fig. 5 muestra los resultados para un conjunto de señales sin ninguna aberración grave de codificación.

30 De la Fig. 4 se puede ver que el método de mejora de esta invención mejora la calidad de las señales codificadas en aproximadamente 15 puntos de MUSHRA tanto para modo 2 como para modo 7 del material codificado con AMR-WB, lo que es una mejora significativa. La Fig. 4 también muestra que el modo 2 mejorado obtiene aproximadamente la misma puntuación de MUSHRA que obtiene el modo 7, el cual requiere el doble de tasa de bit que el modo 2. Esto muestra que el método de mejora está trabajando muy bien y que la baja tasa de bit de 12.65 kbps de tasa de bit por canal podría ser satisfactoriamente usada para codificar señales estéreo y binaurales para aplicaciones de teleconferencia que admitan audio espacial.

35 Los resultados en la Fig. 5 claramente muestran que el método de mejora de acuerdo con la presente invención no está añadiendo distorsiones audibles al material de test que no tenía ninguna distorsión de codificación grave, lo que es una cuestión importante para el método de mejora.

Con estos resultados, está claro que el método de mejora está proporcionando una mejora significativa de las señales codificadas distorsionadas y que con esas mejoras de por ejemplo el códec de AMR-WB combinado con el método de mejora de esta invención puede ser exitosamente usado en aplicaciones de teleconferencia para proporcionar señales binaurales grabadas en estéreo o generadas sintéticamente. Sin el método de mejora, por otro

lado, la calidad de las señales estéreo o binaurales proporcionadas por el decodificador de AMR-WB podría ser demasiado baja para la aplicación prevista.

Será entendido por aquellos expertos en la técnica que varias modificaciones y cambios pueden ser realizados a la presente invención sin desviarse del alcance de la misma, que está definido por las reivindicaciones anexas.

## 5 Referencias

- [1] 3GPP, July 2005. TS 26.190 v6.1.1 (2005-07), Speech codec speech processing function, Adaptive Multi-Rate - Wideband (AMR-WB) speech codec, Release 6.
- [2] BRUNO BESSETTE, REDWAN SALAMI, ROCH LEFEBVRE, MILAN JELÍNEK, JANI ROTOLA-PUKKILA, JANNE VAINIO, HANNU MIKKOLA, KARI JÄRVINEN. November 2002. The Adaptive Multirate Wideband Speech Codec (AMR-WB), IEEE Transaction on speech and audio processing, vol 10, no 8.
- [3] 3GPP, 2007-03, TS 26.290 V7.0.0, Page 57.
- [4] Patent Application WO 03/102923 A2.
- [5] ITU-R RECOMMENDATION BS. 1535-1, 2001, Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA), International Telecommunications Union, Geneva, Switzerland.
- [6] 3GPP, 2007-03. TP 26.290 v7.0.0, Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec, Release 6.
- [7] 3GPP2, 2005-04.C.S0052-A v1.0, Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB).

**REIVINDICACIONES**

1. Un método para mejorar señales de audio espaciales, caracterizado por:  
 recibir (S10) una señal de audio multicanal codificado con ACELP que comprende una pluralidad de bloques;  
 para cada bloque recibido de cada canal de dicha señal de entrada multicanal
  - 5        estimar (S20) un tipo de señal basándose en la señal recibida, determinando si dicho bloque de señal consta de un componente de banda de paso fuerte y estrecho del tono humano con una frecuencia central en el intervalo de 100-500 Hz;  
           estimar (S30) una frecuencia de tono basándose en la señal recibida;
  - 10        determinar (S40) parámetros de filtrado basándose en al menos uno de dicho tipo de señal estimado y dicha frecuencia de tono estimada, y  
           formar (S41) parámetros de filtro conjuntos basándose en los respectivos parámetros del filtro determinados para dichos canales múltiples
  - filtrar en paso alto (S50) dicha señal recibida basándose en dichos parámetros de filtro conjuntos determinados para proporcionar una señal de salida filtrada en paso alto.
- 15    2. El método de acuerdo a la reivindicación 1, en el que dicho paso de formar (S41) parámetros de filtro conjuntos comprende determinar una frecuencia de corte para cada canal basándose en el tipo de señal estimado y la frecuencia de tono, y formar dichos parámetros de filtro conjuntos basándose en la frecuencia de corte más baja.
3. El método de acuerdo a la reivindicación 1, en el que dicha señal de entrada multicanal es una señal estéreo.
- 20    4. El método de acuerdo a la reivindicación 1, en el que dicho paso de estimación del tono S30 comprende el paso adicional S31 de determinar si la estimación de tono es necesaria, y realizar dichas estimación de tono basándose en dicho paso de determinación.
5. El método de acuerdo a la reivindicación 4, en el que si dicho paso de determinación S31 necesita estimar el tono, la estimación del tono de dicha señal recibida y la determinación (S40) de dichos parámetros de filtrado se basa en ambos de dicho tipo de señal estimado y dicha frecuencia de tono estimada.
- 25    6. El método de acuerdo a cualquiera de los pasos anteriores, en el que dicha señal espacial es una señal de ACELP de AMR-WB.
7. Una disposición (1) para mejorar señales de audio espaciales recibidas, caracterizada por:  
 medios para recibir (10) una señal de audio multicanal codificada con ACELP que consta de una pluralidad de bloques;
- 30    medios para estimar (20) un tipo de señal para cada bloque de señal de cada canal basándose en la señal recibida, determinando si dicho bloque de señal comprende un componente de banda de paso fuerte y estrecha de tono humano con una frecuencia central en el intervalo de 100-500 Hz;  
 medios para estimar (30) una frecuencia de tono para cada bloque de señal de cada canal basándose en la señal recibida;
- 35    medios para determinar (40) parámetros de filtrado basándose en dicho tipo de señal estimado y dicha frecuencia de tono estimada, comprendiendo dichos medios (40) medios para formar (41) parámetros de filtro conjuntos basándose en los respectivos parámetros de filtro determinados para dichos canales múltiples; y  
 medios para filtrar en paso alto (50) dicha señal recibida basándose en dichos parámetros de filtro conjuntos determinados para proporcionar una señal de salida filtrada en paso alto.
- 40    8. La disposición de acuerdo a la reivindicación 7, en la que dichos medios de filtrado en paso alto constan de una pluralidad de filtros.
9. La disposición de acuerdo a la reivindicación 8, en la que dichos filtros constan de filtros de Respuesta Finita al Impulso o filtros de Respuesta Infinita al Impulso.
- 45    10. La disposición de acuerdo a la reivindicación 8, dichos filtros constan de filtros de Respuesta al Impulso Infinita elípticos.



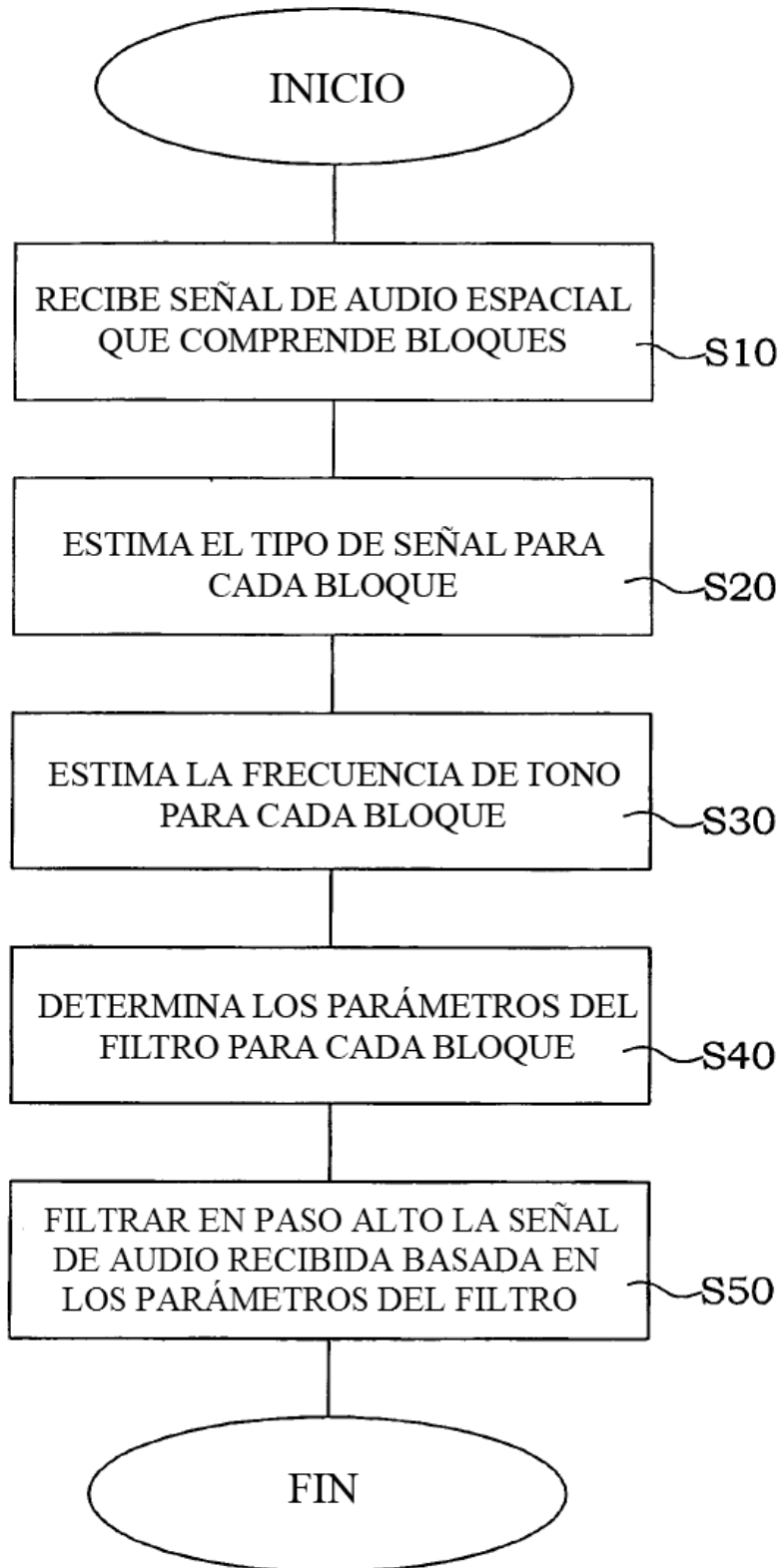


Fig. 1

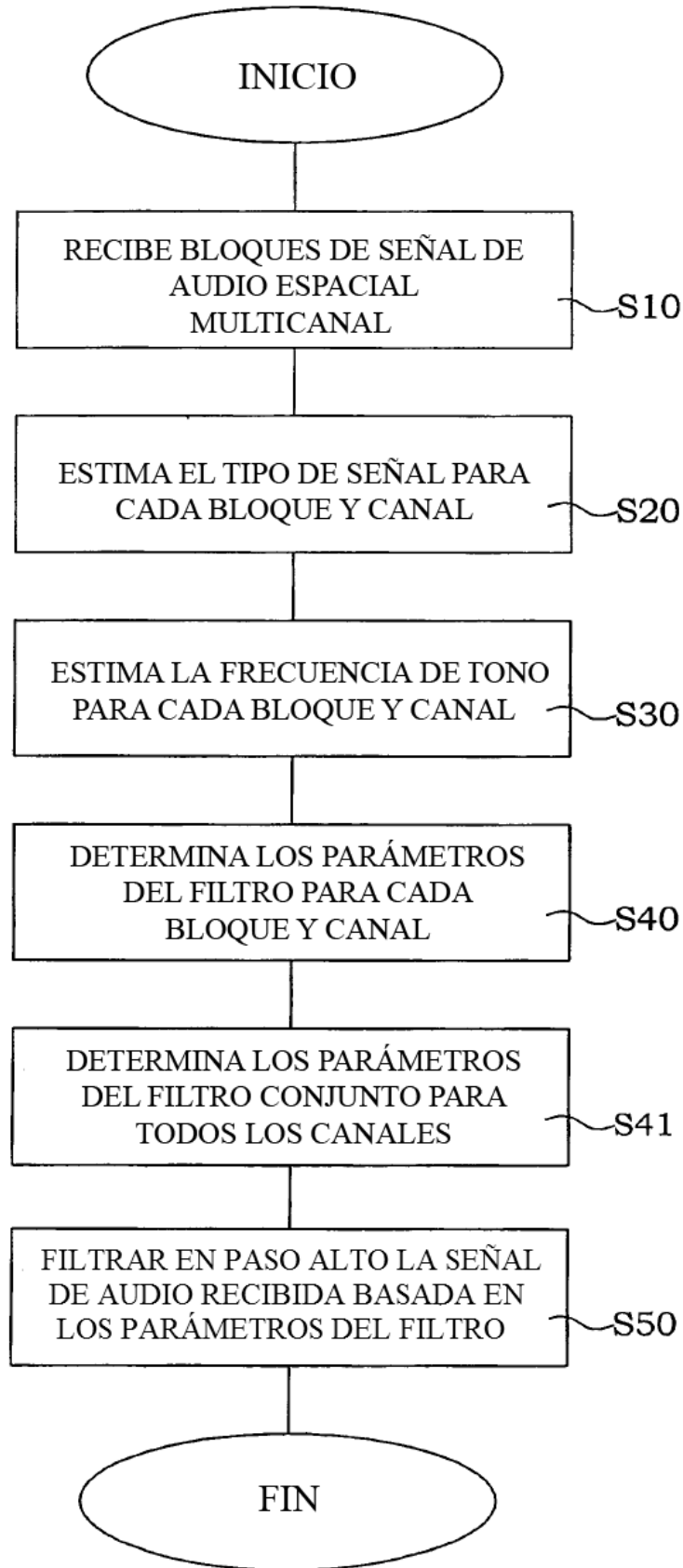


Fig. 2

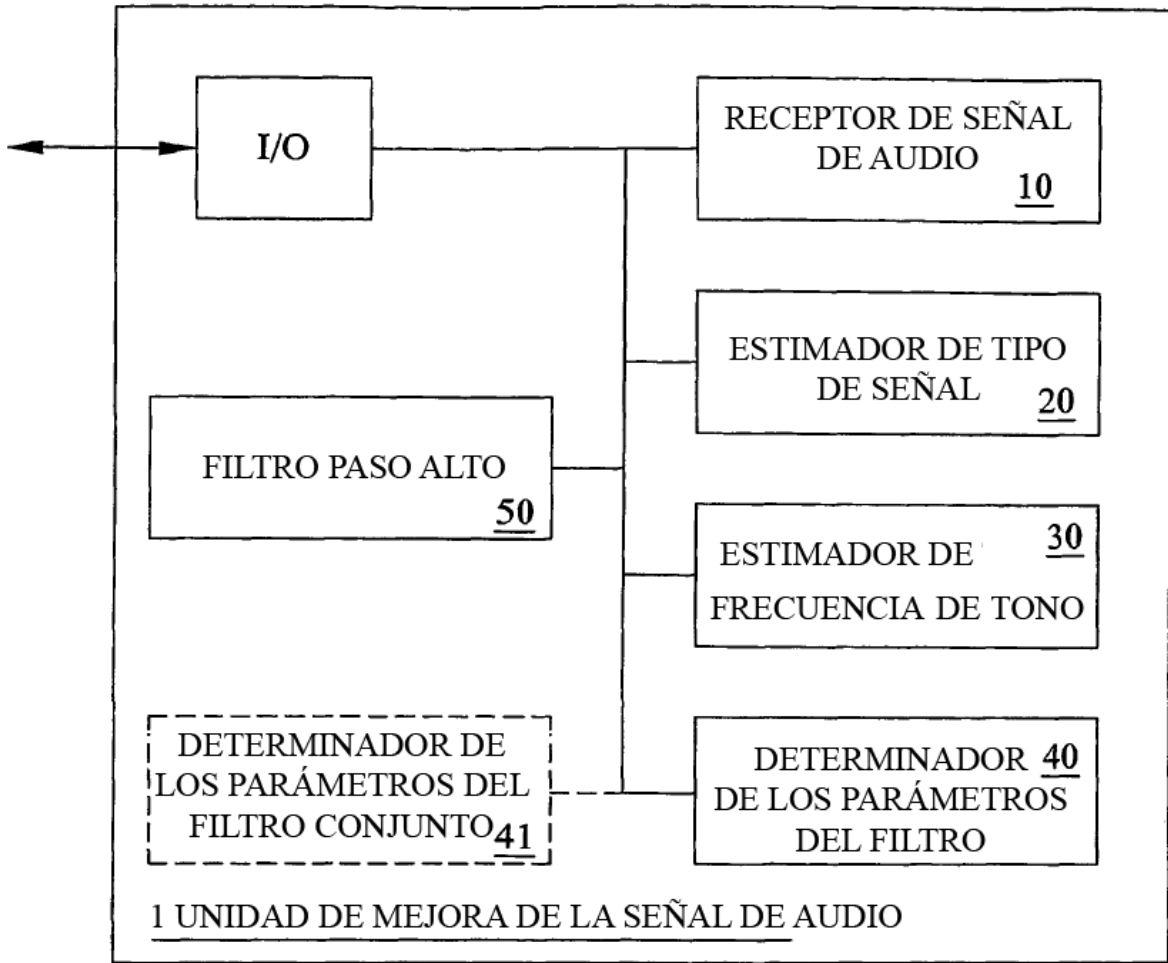


Fig. 3a

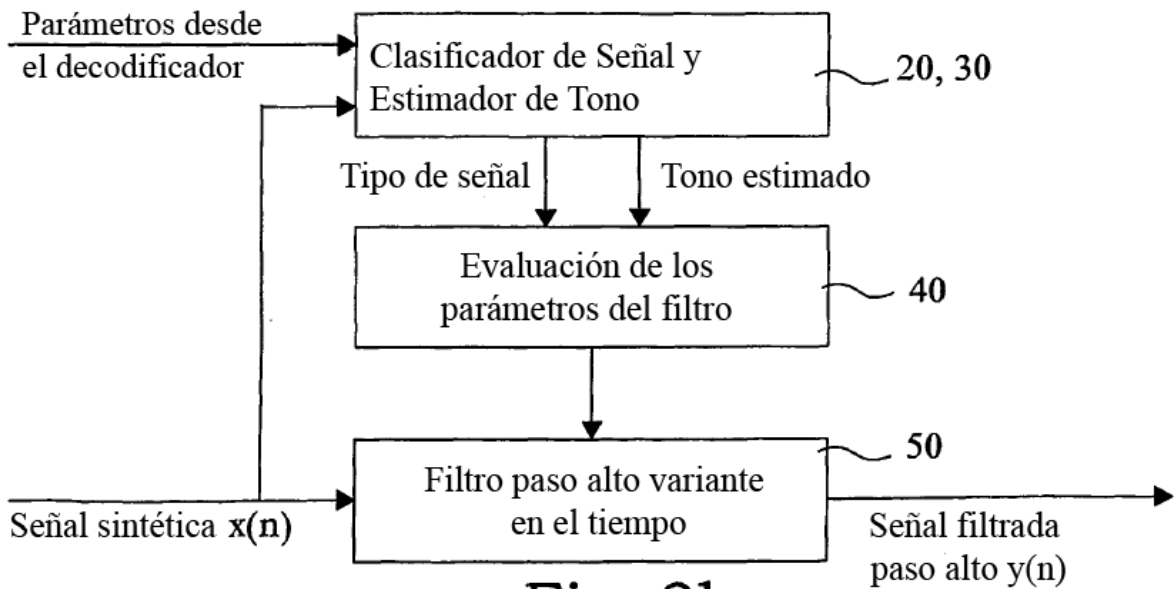


Fig. 3b

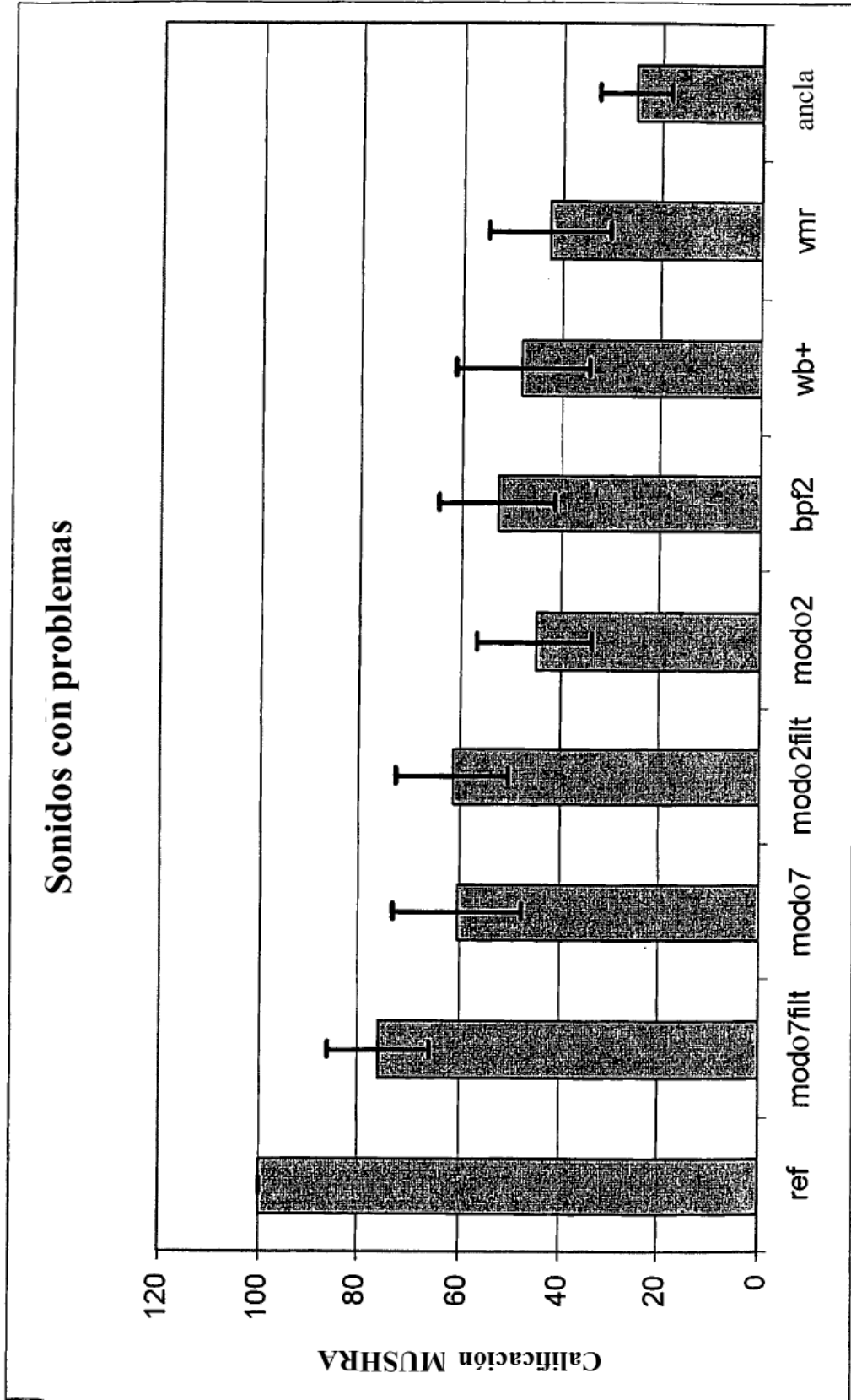


Fig. 4

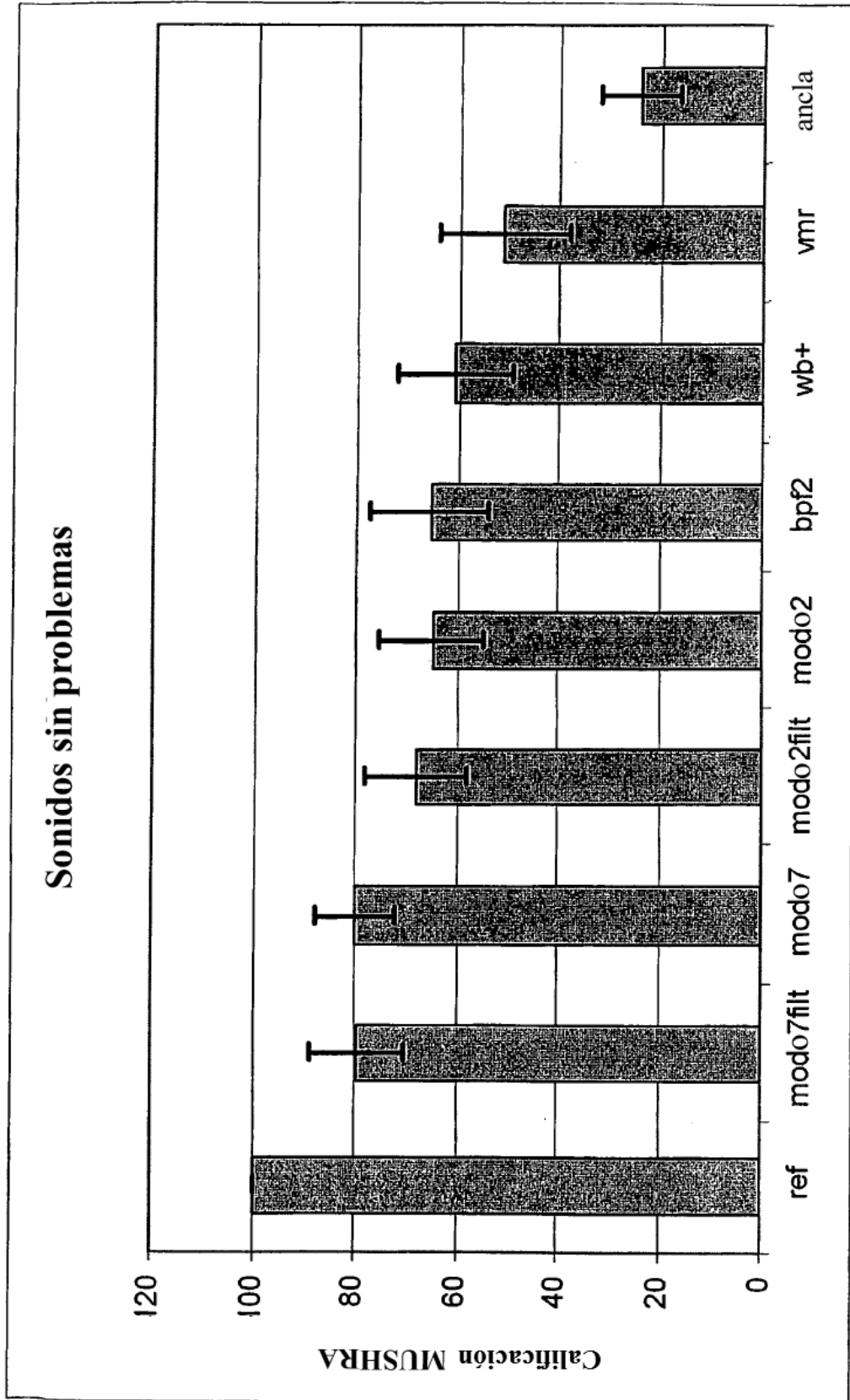


Fig. 5