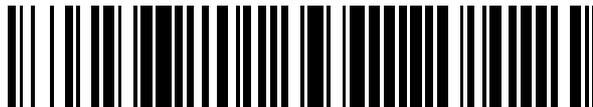


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 530 957**

51 Int. Cl.:

**G10L 21/00** (2013.01)

**G10L 19/12** (2013.01)

**G10L 19/02** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.10.2011 E 11764739 (6)**

97 Fecha y número de publicación de la concesión europea: **03.12.2014 EP 2625688**

54 Título: **Aparato y método para procesar una señal de audio y para proporcionar una mayor granularidad temporal para un códec de voz y de audio unificado combinado (USAC)**

30 Prioridad:

**06.10.2010 US 390267 P**

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

**09.03.2015**

73 Titular/es:

**FRAUNHOFER GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (50.0%)  
Hansastraße 27c  
80636 München, DE y  
VOICEAGE CORPORATION (50.0%)**

72 Inventor/es:

**MULTRUS, MARKUS;  
GRILL, BERNHARD;  
NEUENDORF, MAX;  
RETTELBACH, NIKOLAUS;  
FUCHS, GUILLAUME;  
GOURNAY, PHILIPPE;  
LEFEBVRE, ROCH;  
BESSETTE, BRUNO y  
WILDE, STEPHAN**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 530 957 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

Aparato y método para procesar una señal de audio y para proporcionar una mayor granularidad temporal para un códec de voz y de audio unificado combinado (USAC).

5 La presente invención se refiere a procesamiento de audio y, en particular, a un aparato y a un método para procesar una señal de audio y para proporcionar una mayor granularidad temporal para un códec de voz y de audio unificado combinado (USAC).

10 USAC, como otros códecs de audio, presenta un tamaño de trama fijo (USAC: 2048 muestras/trama). Aunque existe la posibilidad de cambiar a un conjunto limitado de tamaños de transformada más cortos dentro de una trama, el tamaño de trama todavía limita la resolución temporal del sistema completo. Para aumentar la granularidad temporal del sistema completo, para códecs de audio tradicionales se aumenta la tasa de muestreo, lo que conduce a una duración más corta de una trama en el tiempo (por ejemplo, milisegundos). Sin embargo, esto no es posible fácilmente para el códec de USAC:

15 El códec de USAC comprende una combinación de herramientas de códecs de audio generales tradicionales, tales como codificador de transformada de AAC (codificación de audio avanzada), SBR (replicación de banda espectral) y MPEG envolvente (MPEG = Grupo de expertos en imágenes en movimiento), más herramientas de codificadores de voz tradicionales, tales como ACELP (ACELP = Predicción lineal excitada por códigos algebraicos). Ambos, ACELP y el codificador de transformada, se ejecutan habitualmente al mismo tiempo dentro del mismo entorno (es decir, tamaño de trama, tasa de muestreo) y pueden cambiarse fácilmente: habitualmente, para señales de voz claras, se usa la herramienta ACELP, y para señales mixtas, de música, se usa el codificador de transformada.

20 La herramienta ACELP está limitada al mismo tiempo para que funcione sólo a tasas de muestreo comparablemente bajas. Para 24 kbit/s, se usa una tasa de muestreo de sólo 17075 Hz. Para mayores tasas de muestreo, el rendimiento de la herramienta ACELP comienza a disminuir significativamente. Sin embargo, el codificador de transformada así como SBR y MPEG envolvente se beneficiarían de una tasa de muestreo mucho mayor, por ejemplo 22050 Hz para el codificador de transformada y 44100 Hz para SBR y MPEG envolvente. Sin embargo, hasta ahora, la herramienta ACELP limitaba la tasa de muestreo del sistema completo, conduciendo a un sistema subóptimo en particular para señales de música.

25 El objeto de la presente invención es proporcionar conceptos mejorados para un aparato y un método para procesar una señal de audio. El objeto de la presente invención se resuelve mediante un aparato según la reivindicación 1, un método según la reivindicación 15, un aparato según la reivindicación 16, un método según la reivindicación 18 y un programa informático según la reivindicación 19.

30 El RM de USAC actual proporciona un alto rendimiento de codificación por un gran número de puntos de operación, que oscila entre tasas de transmisión de bits muy bajas tales como 8 kbit/s y una calidad transparente a tasas de transmisión de bits de 128 kbit/s y más. Para alcanzar esta alta calidad por una amplia variedad de este tipo de tasas de transmisión de bits, se usa una combinación de herramientas, tales como MPEG envolvente, SBR, ACELP y codificadores de transformada tradicionales. Una combinación de herramientas de este tipo requiere naturalmente de un proceso de optimización conjunto del interfuncionamiento de la herramienta y un entorno común, en el que se sitúan estas herramientas.

35 Se encontró en este proceso de optimización conjunto que algunas de las herramientas tienen deficiencias que reproducen señales, que exponen una estructura altamente temporal en el intervalo intermedio de tasa de transmisión de bits (24 kbit/s - 32 kbit/s). En particular, las herramientas MPEG envolvente, SBR y los codificadores de transformada FD (FD, TCX) (FD = dominio de frecuencia; TCX = excitación de transformada codificada), es decir todas las herramientas, que operan en el dominio de frecuencia, pueden funcionar mejor cuando se operan con mayor granularidad temporal, lo que es idéntico a un tamaño de trama más corto en el dominio de tiempo.

40 En comparación con el estado de la técnica, el codificador HE-AACv2 (codificador AAC v2 de alta eficacia), se encontró que el codificador de calidad de referencia de USAC actual opera a tasas de transmisión de bits tales como 24 kbit/s y 32 kbit/s a una tasa de muestreo significativamente más baja, aunque se usa el mismo tamaño de trama (en muestras). Esto significa que la duración de las tramas en milisegundos es significativamente más larga. Para compensar estas deficiencias, es necesario aumentar la granularidad temporal. Esto puede conseguirse o bien aumentando la frecuencia de muestreo o bien acortando los tamaños de trama (por ejemplo, de sistemas que usan un tamaño de trama fijo).

45 Mientras que el aumento de la frecuencia de muestreo es una manera razonable para que SBR y MPEG envolvente aumenten el rendimiento para señales dinámicas temporales, esto no funcionará para todas las herramientas de codificador de núcleo: se conoce bien que una mayor frecuencia de muestreo sería beneficiosa para el codificador de transformada, pero al mismo tiempo disminuye drásticamente el rendimiento de la herramienta ACELP.

5 Se proporciona un aparato para procesar una señal de audio. El aparato comprende un procesador de señal y un configurador. El procesador de señal está adaptado para recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio. Además, el procesador de señal está adaptado para sobremuestrear la señal de audio mediante un factor de sobremuestreo configurable para obtener una señal de audio procesada. Además, el procesador de señal está adaptado para emitir una segunda trama de señal de audio que tiene un segundo número configurable de muestras de la señal de audio procesada.

10 El configurador está adaptado para configurar el procesador de señal basándose en información de configuración de manera que el factor de sobremuestreo configurable sea igual a un primer valor de sobremuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación. Además, el configurador está adaptado para configurar el procesador de señal de manera que el factor de sobremuestreo configurable sea igual a un segundo valor de sobremuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente. El primer o el segundo valor de relación no es un valor de número entero.

15 Según la realización descrita anteriormente, un procesador de señal sobremuestra una señal de audio para obtener una señal de audio sobremuestreada procesada. En la realización anterior, el factor de sobremuestreo es configurable y puede ser un valor no de número entero. La capacidad de configuración y el hecho de que el factor de sobremuestreo puede ser un valor no de número entero aumenta la flexibilidad del aparato. Cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente, entonces el factor de sobremuestreo configurable tiene un segundo valor de sobremuestreo diferente. Por tanto, el aparato está adaptado para tener en cuenta una relación entre el factor de sobremuestreo y la relación de la longitud de trama (es decir, el número de muestras) de la segunda y la primera trama de señal de audio.

20 En una realización, el configurador está adaptado para configurar el procesador de señal de manera que el segundo valor de sobremuestreo diferente sea superior al primer valor de sobremuestreo, cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras es superior a la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras.

25 Según una realización, se propone un modo de operación nuevo (a continuación denominado "ajuste extra") para el códec de USAC, que mejora el rendimiento del sistema para tasas de transmisión de datos intermedias, tales como 24 kbit/s y 32 kbit/s. Se encontró que para estos puntos de operación, la resolución temporal del códec de referencia de USAC actual es muy baja. Por tanto, se propone a) aumentar esta resolución temporal acortando los tamaños de trama de codificador de núcleo sin aumentar la tasa de muestreo para el codificador de núcleo, y además b) aumentar la tasa de muestreo para SBR y MPEG envolvente sin cambiar el tamaño de trama para estas herramientas.

30 El ajuste extra propuesto mejora en gran parte la flexibilidad del sistema, puesto que permite que el sistema que incluye la herramienta ACELP se opere a mayores tasas de muestreo, tales como 44,1 y 48 kHz. Puesto que estas tasas de muestreo se solicitan normalmente en el mercado, se espera que esto ayude a la aceptación del códec de USAC.

35 El modo de operación nuevo para el elemento de trabajo de codificación de voz y de audio unificada (USAC) de MPEG actual aumenta la flexibilidad temporal de todo el códec, aumentando la granularidad temporal del códec de audio completo. Si (suponiendo que el segundo número de muestras permanece igual) la segunda relación es superior a la primera relación, entonces se ha reducido el primer número configurable de muestras, es decir se ha acortado el tamaño de trama de la primera trama de señal de audio. Esto da como resultado una mayor granularidad temporal, y todas las herramientas que operan en el dominio de frecuencia y que procesan la primera trama de señal de audio pueden funcionar mejor. Sin embargo, en un modo de operación de alta eficacia de este tipo, también es deseable aumentar el rendimiento de herramientas que procesan la segunda trama de señal de audio que comprende la señal de audio sobremuestreada. Puede realizarse un aumento de este tipo en el rendimiento de estas herramientas mediante una mayor tasa de muestreo de la señal de audio sobremuestreada, es decir aumentando el factor de sobremuestreo para un modo de operación de este tipo. Además, existen herramientas, tales como el decodificador de ACELP en USAC, que no operan en el dominio de frecuencia, que procesan la primera trama de señal de audio y que operan mejor cuando la tasa de muestreo de la señal de audio (original) es relativamente baja. Estas herramientas se benefician de un factor de sobremuestreo alto, ya que esto significa que la tasa de muestreo de la señal de audio (original) es relativamente baja en comparación con la tasa de muestreo de la señal de audio sobremuestreada. La realización descrita anteriormente proporciona un aparato adaptado para proporcionar un modo de configuración para un modo de operación eficaz para un entorno de este tipo.

40 El modo de operación nuevo aumenta la flexibilidad temporal de todo el códec, aumentando la granularidad temporal del códec de audio completo. En una realización, el configurador está adaptado para configurar el procesador de señal de manera que el factor de sobremuestreo configurable sea igual al primer valor de relación cuando la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el primer

valor de relación y en la que el configurador está adaptado para configurar el procesador de señal de manera que el factor de sobremuestreo configurable sea igual al segundo valor de relación diferente cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el segundo valor de relación diferente.

5 En una realización, el configurador está adaptado para configurar el procesador de señal de manera que el factor de sobremuestreo configurable sea igual a 2 cuando la primera relación tiene el primer valor de relación y en la que el configurador está adaptado para configurar el procesador de señal de manera que el factor de sobremuestreo configurable sea igual a 8/3 cuando la segunda relación tiene el segundo valor de relación diferente.

10 Según una realización adicional, el configurador está adaptado para configurar el procesador de señal de manera que el primer número configurable de muestras sea igual a 1024 y el segundo número configurable de muestras sea igual a 2048 cuando la primera relación tiene el primer valor de relación y en la que el configurador está adaptado para configurar el procesador de señal de manera que el primer número configurable de muestras sea igual a 768 y el segundo número configurable de muestras sea igual a 2048 cuando la segunda relación tiene el segundo valor de relación diferente.

15 En una realización, se propone introducir un ajuste adicional del codificador de USAC, en la que el codificador de núcleo se opera a un tamaño de trama más corto (768 en lugar de 1024 muestras). Además, se propone modificar en este contexto el remuestreo dentro del decodificador de SBR de 2:1 a 8:3, para permitir que SBR y MPEG envolvente se operen a una mayor tasa de muestreo.

20 Además, según una realización, se aumenta la granularidad temporal del codificador de núcleo reduciendo el tamaño de trama de codificador de núcleo de 1024 a 768 muestras. Mediante esta etapa, la granularidad temporal del codificado de núcleo se aumenta en 4/3 mientras se deja la tasa de muestreo constante: esto permite que la ACELP se ejecute a una frecuencia de muestreo ( $F_s$ ) apropiada.

25 Además, en la herramienta de SBR, se aplica un remuestreo de relación 8/3 (hasta ahora: relación 2), que convierte una trama de codificador de núcleo del tamaño 768 a  $3/8 F_s$  a una trama de salida del tamaño 2048 a  $F_s$ . Esto permite que la herramienta SBR y una herramienta MPEG envolvente se ejecuten a una tasa de muestreo tradicionalmente alta (por ejemplo, 44100 Hz). Por tanto, se proporciona una buena calidad para las señales de voz y música, ya que todas las herramientas van a ejecutarse en su punto de operación óptimo.

30 En una realización, el procesador de señal comprende un módulo decodificador de núcleo para decodificar la señal de audio para obtener una señal de audio procesada previamente, un banco de filtros de análisis que tiene varios canales de banco de filtros de análisis para transformar la primera señal de audio procesada previamente de un dominio de tiempo a un dominio de frecuencia para obtener una señal de audio procesada previamente de dominio de frecuencia que comprende una pluralidad de señales de subbanda, un generador de subbanda para crear y añadir señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia y un banco de filtros de síntesis que tiene varios canales de banco de filtros de síntesis para transformar la primera señal de audio procesada previamente del dominio de frecuencia al dominio de tiempo para obtener la señal de audio procesada. El configurador puede estar adaptado para configurar el procesador de señal configurando el número de canales de banco de filtros de síntesis o el número de canales de banco de filtros de análisis de manera que el factor de sobremuestreo configurable sea igual a una tercera relación del número de canales de banco de filtros de síntesis con respecto al número de canales de banco de filtros de análisis. El generador de subbanda puede ser un replicador de banda espectral que está adaptado para replicar señales de subbanda del generador de señal de audio procesada previamente para crear las señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia. El procesador de señal puede comprender además un decodificador de MPEG envolvente para decodificar la señal de audio procesada previamente para obtener una señal de audio procesada previamente que comprende canales estéreo o envolventes. Además, el generador de subbanda puede estar adaptado para alimentar la señal de audio procesada previamente de dominio de frecuencia al decodificador de MPEG envolvente después de que las señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia se hayan creado y añadido a la señal de audio procesada previamente de dominio de frecuencia.

35 El módulo decodificador de núcleo puede comprender un primer decodificador de núcleo y un segundo decodificador de núcleo, en el que el primer decodificador de núcleo puede estar adaptado para operar en un dominio de tiempo y en el que el segundo decodificador de núcleo puede estar adaptado para operar en un dominio de frecuencia. El primer decodificador de núcleo puede ser un decodificador de ACELP y el segundo decodificador de núcleo puede ser un decodificador de transformada FD o un decodificador de transformada TCX.

40 En una realización, el tamaño de supertrama para el códec de ACELP se reduce de 1024 a 768 muestras. Esto podría realizarse combinando 4 tramas de ACELP de tamaño 192 (3 subtramas de tamaño 64) dando una trama de codificador de núcleo de tamaño 768 (previamente: se combinaron 4 tramas de ACELP de tamaño 256 dando una

55

trama de codificador de núcleo de tamaño 1024). Otra solución para conseguir un tamaño de trama de codificador de núcleo de 768 muestras sería, por ejemplo, combinar 3 tramas de ACELP de tamaño 256 (4 subtramas de tamaño 64).

5 Según una realización adicional, el configurador está adaptado para configurar el procesador de señal basándose en la información de configuración que indica al menos uno del primer número configurable de muestras de la señal de audio o el segundo número configurable de muestras de la señal de audio procesada.

En otra realización, el configurador está adaptado para configurar el procesador de señal basándose en la información de configuración, en la que la información de configuración indica el primer número configurable de muestras de la señal de audio y el segundo número configurable de muestras de la señal de audio procesada, en la que la información de configuración es un índice de configuración.

10 Además, se proporciona un aparato para procesar una señal de audio. El aparato comprende un procesador de señal y un configurador. El procesador de señal está adaptado para recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio. Además, el procesador de señal está adaptado para submuestrear la señal de audio mediante un factor de submuestreo configurable para obtener una señal de audio procesada. Además, el procesador de señal está adaptado para emitir una segunda trama de señal de audio que tiene  
15 un segundo número configurable de muestras de la señal de audio procesada.

El configurador puede estar adaptado para configurar el procesador de señal basándose en información de configuración de manera que el factor de submuestreo configurable sea igual a un primer valor de submuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación. Además, el configurador está adaptado para configurar el procesador de  
20 señal de manera que el factor de submuestreo configurable sea igual a un segundo valor de submuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente. El primer o el segundo valor de relación no es un valor de número entero.

25 A continuación se comentarán realizaciones preferidas de la presente invención con respecto a las figuras adjuntas, en las que:

la figura 1 ilustra un aparato para procesar una señal de audio según una realización,

la figura 2 ilustra un aparato para procesar una señal de audio según otra realización,

la figura 3 ilustra un proceso de sobremuestreo llevado a cabo por un aparato según una realización,

la figura 4 ilustra un aparato para procesar una señal de audio según una realización adicional,

30 la figura 5a ilustra un módulo decodificador de núcleo según una realización,

la figura 5b ilustra un aparato para procesar una señal de audio según la realización de la figura 4 con un módulo decodificador de núcleo según la figura 5a,

la figura 6a ilustra una supertrama de ACELP que comprende 4 tramas de ACELP,

la figura 6b ilustra una supertrama de ACELP que comprende 3 tramas de ACELP,

35 la figura 7a ilustra el ajuste por defecto de USAC,

la figura 7b ilustra un ajuste extra para USAC según una realización,

las figuras 8a, 8b ilustran los resultados de una prueba de escucha según la metodología de MUSHRA y

la figura 9 ilustra un aparato para procesar una señal de audio según una realización alternativa.

40 La figura 1 ilustra un aparato para procesar una señal de audio según una realización. El aparato comprende un procesador 110 de señal y un configurador 120. El procesador 110 de señal está adaptado para recibir una primera trama 140 de señal de audio que tiene un primer número configurable de muestras 145 de la señal de audio. Además, el procesador 110 de señal está adaptado para sobremuestrear la señal de audio mediante un factor de sobremuestreo configurable para obtener una señal de audio procesada. Además, el procesador de señal está adaptado para emitir una segunda trama 150 de señal de audio que tiene un segundo número configurable de muestras 155 de la señal de  
45 audio procesada.

El configurador 120 está adaptado para configurar el procesador 110 de señal basándose en información de configuración de manera que el factor de sobremuestreo configurable sea igual a un primer valor de sobremuestreo

cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación. Además, el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el factor de sobremuestreo configurable sea igual a un segundo valor de sobremuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente. El primer o el segundo valor de relación no es un valor de número entero.

Un aparato según la figura 1 puede emplearse, por ejemplo, en el proceso de decodificación.

Según una realización, el configurador 120 puede estar adaptado para configurar el procesador 110 de señal de manera que el segundo valor de sobremuestreo diferente sea superior al primer valor de sobremuestreo diferente, cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras es superior a la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras. En una realización adicional, el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el factor de sobremuestreo configurable sea igual al primer valor de relación cuando la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el primer valor de relación y en la que el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el factor de sobremuestreo configurable sea igual al segundo valor de relación diferente cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el segundo valor de relación diferente.

En otra realización, el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el factor de sobremuestreo configurable sea igual a 2 cuando la primera relación tiene el primer valor de relación y en la que el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el factor de sobremuestreo configurable sea igual a 8/3 cuando la segunda relación tiene el segundo valor de relación diferente. Según una realización adicional, el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el primer número configurable de muestras sea igual a 1024 y el segundo número configurable de muestras sea igual a 2048 cuando la primera relación tiene el primer valor de relación y en la que el configurador 120 está adaptado para configurar el procesador 110 de señal de manera que el primer número configurable de muestras sea igual a 768 y el segundo número configurable de muestras sea igual a 2048 cuando la segunda relación tiene el segundo valor de relación diferente.

En una realización, el configurador 120 está adaptado para configurar el procesador 110 de señal basándose en la información de configuración ci, en la que la información de configuración ci indica el factor de sobremuestreo, el primer número configurable de muestras de la señal de audio y el segundo número configurable de muestras de la señal de audio procesada, en la que la información de configuración es un índice de configuración.

La siguiente tabla ilustra un ejemplo de un índice de configuración como información de configuración:

Índice	coreCoderFrameLength	sbrRatio	outputFrameLength
2	768	8:3	2048
3	1024	2:1	2048

donde "Índice" indica el índice de configuración, donde "coreCoderFrameLength" indica el primer número configurable de muestras de la señal de audio, donde "sbrRatio" indica el factor de sobremuestreo y donde "outputFrameLength" indica el segundo número configurable de muestras de la señal de audio procesada.

La figura 2 ilustra un aparato según otra realización. El aparato comprende un procesador 205 de señal y un configurador 208. El procesador 205 de señal comprende un módulo 210 decodificador de núcleo, un banco 220 de filtros de análisis, un generador 230 de subbanda y un banco 240 de filtros de síntesis.

El módulo 210 decodificador de núcleo está adaptado para recibir una señal de audio as1. Después de recibir la señal de audio as1, el módulo 210 decodificador de núcleo decodifica la señal de audio para obtener una señal de audio as2 procesada previamente. Entonces, el módulo 210 decodificador de núcleo alimenta la señal de audio as2 procesada previamente, que se representa en un dominio de tiempo, al banco 220 de filtros de análisis.

El banco 220 de filtros de análisis está adaptado para transformar la señal de audio as2 procesada previamente de un dominio de tiempo a un dominio de frecuencia para obtener una señal de audio as3 procesada previamente de dominio de frecuencia que comprende una pluralidad de señales de subbanda. El banco 220 de filtros de análisis tiene un número configurable de canales de banco de filtros de análisis (bandas de banco de filtros de análisis). El número de canales de banco de filtros de análisis determina el número de señales de subbanda que se generan desde la señal de

audio as2 procesada previamente de dominio de tiempo. En una realización, el número de canales de banco de filtros de análisis puede establecerse ajustando el valor de un parámetro c1 configurable. Por ejemplo, el banco 220 de filtros de análisis puede estar configurado para tener 32 ó 24 canales de banco de filtros de análisis. En la realización de la figura 2, el número de canales de banco de filtros de análisis puede establecerse según información de configuración ci de un configurador 208. Después de transformar la señal de audio as2 procesada previamente al dominio de frecuencia, el banco 220 de filtros de análisis alimenta la señal de audio as3 procesada previamente de dominio de frecuencia al generador 230 de subbanda.

El generador 230 de subbanda está adaptado para crear señales de subbanda adicionales para la señal de audio as3 de dominio de frecuencia. Además, el generador 230 de subbanda está adaptado para modificar la señal de audio as3 de dominio de frecuencia procesada previamente para obtener una señal de audio as4 de dominio de frecuencia modificada que comprende las señales de subbanda de la señal de audio as3 de dominio de frecuencia procesada previamente y las señales de subbanda adicionales creadas, creadas por el generador 230 de subbanda. El número de señales de subbanda adicionales que se generan por el generador 230 de subbanda es configurable. En una realización, el generador de subbanda es un replicador de banda espectral (SBR). El generador 230 de subbanda alimenta entonces la señal de audio as4 procesada previamente de dominio de frecuencia modificada al banco de filtros de síntesis.

El banco 240 de filtros de síntesis está adaptado para transformar la señal de audio as4 procesada previamente de dominio de frecuencia modificada de un dominio de frecuencia a un dominio de tiempo para obtener una señal de audio as5 procesada de dominio de tiempo. El banco 240 de filtros de síntesis tiene un número configurable de canales de banco de filtros de síntesis (bandas de banco de filtros de síntesis). El número de canales de banco de filtros de síntesis es configurable. En una realización, el número de canales de banco de filtros de síntesis puede establecerse ajustando el valor de un parámetro c2 configurable. Por ejemplo, el banco 240 de filtros de síntesis puede estar configurado para tener 64 canales de banco de filtros de síntesis. En la realización de la figura 2, la información de configuración ci del configurador 208 puede establecer el número de canales de banco de filtros de análisis. Mediante la transformación de la señal de audio as4 procesada previamente de dominio de frecuencia modificada al dominio de tiempo, se obtiene la señal de audio as5 procesada.

En una realización, el número de canales de subbanda de la señal de audio as4 procesada previamente de dominio de frecuencia modificada es igual al número de canales de banco de filtros de síntesis. En una realización de este tipo, el configurador 208 está adaptado para configurar el número de canales de subbanda adicionales que se crean por el generador 230 de subbanda. El configurador 208 puede estar adaptado para configurar el número de canales de subbanda adicionales que se crean por el generador 230 de subbanda de manera que el número de canales c2 de banco de filtros de síntesis, configurados por el configurador 208, sea igual al número de canales de subbanda de la señal de audio as3 de dominio de frecuencia procesada previamente más el número de señales de subbanda adicionales creadas por el generador 230 de subbanda. De este modo, el número de canales de banco de filtros de síntesis es igual al número de señales de subbanda de la señal de audio as4 de dominio de frecuencia procesada previamente modificada.

Suponiendo que la señal de audio as1 tiene una tasa de muestreo sr1 y suponiendo que el banco 220 de filtros de análisis tiene c1 canales de banco de filtros de análisis y que el banco 240 de filtros de síntesis tiene c2 canales de banco de filtros de síntesis, la señal de audio as5 procesada tiene una tasa de muestreo sr5:

$$sr5 = (c2/c1) \cdot sr1.$$

c2/c1 determina el factor de sobremuestreo u:

$$u = c2/c1.$$

En la realización de la figura 2, el factor de sobremuestreo u puede establecerse a un número que no es un valor de número entero. Por ejemplo, el factor de sobremuestreo u puede establecerse al valor 8/3, ajustando el número de canales de banco de filtros de análisis: c1 = 24 y ajustando el número de canales de banco de filtros de síntesis: c2 = 64, de manera que:

$$u = 8/3 = 64/24.$$

Suponiendo que el generador 230 de subbanda es un replicador de banda espectral, un replicador de banda espectral según una realización puede generar un número arbitrario de subbandas adicionales desde las subbandas originales, en la que la relación del número de subbandas adicionales generadas con respecto al número de subbandas ya disponibles no tiene que ser un número entero. Por ejemplo, un replicador de banda espectral según una realización puede llevar a cabo las siguientes etapas:

5 En una primera etapa, el replicador de banda espectral replica el número de señales de subbanda generando varias subbandas adicionales, en la que el número de subbandas adicionales generadas puede ser un número entero múltiplo del número de las subbandas ya disponibles. Por ejemplo, pueden generarse 24 (o, por ejemplo, 48) señales de subbanda adicionales de 24 señales de subbanda originales de una señal de audio (por ejemplo, puede duplicarse o triplicarse el número total de señales de subbanda).

En una segunda etapa, suponiendo que el número deseado de señales de subbanda es  $c_{12}$  y el número de señales de subbanda disponibles real es  $c_{11}$ , pueden distinguirse tres situaciones diferentes:

Si  $c_{11}$  es igual a  $c_{12}$ , entonces el número  $c_{11}$  de señales de subbanda disponibles es igual al número  $c_{12}$  de señales de subbanda necesarias. No se requiere un ajuste de subbandas.

10 Si  $c_{12}$  es menor que  $c_{11}$ , entonces el número  $c_{11}$  de señales de subbanda disponibles es superior al número  $c_{12}$  de señales de subbanda necesarias. Según una realización, podrían eliminarse las señales de subbanda de frecuencia más alta. Por ejemplo, si están disponibles 64 señales de subbanda y si sólo se requieren 61 señales de subbanda, podrían descartarse las tres señales de subbanda con la frecuencia más alta.

15 Si  $c_{12}$  es superior a  $c_{11}$ , entonces el número  $c_{11}$  de señales de subbanda disponibles es menor que el número  $c_{12}$  de señales de subbanda necesarias.

20 Según una realización, podrían generarse señales de subbanda adicionales añadiendo señales de cero como señales de subbanda adicionales, es decir señales en las que los valores de amplitud de cada muestra de subbanda son iguales a cero. Según otra realización, podrían generarse señales de subbanda adicionales añadiendo señales de subbanda pseudoaleatorias como señales de subbanda adicionales, es decir señales de subbanda en las que los valores de cada muestra de subbanda comprenden datos pseudoaleatorios. En otra realización, podrían generarse señales de subbanda adicionales copiando los valores de muestra de la señal de subbanda más alta, o las señales de subbanda más altas, y podrían usarse como valores de muestra de las señales de subbanda adicionales (señales de subbanda copiadas).

25 En un replicador de banda espectral según una realización, pueden copiarse y emplearse subbandas de banda base disponibles como subbandas más altas de manera que se llenen todas las subbandas. La misma subbanda de banda base puede copiarse dos veces o una pluralidad de veces de manera que todas las subbandas que faltan puedan llenarse con valores.

30 La figura 3 ilustra un proceso de sobremuestreo llevado a cabo por un aparato según una realización. Se ilustran una señal 310 de audio de dominio de tiempo y algunas muestras 315 de la señal 310 de audio. La señal de audio se transforma a un dominio de frecuencia, por ejemplo un dominio de tiempo-frecuencia para obtener una señal 320 de audio de dominio de frecuencia que comprende tres señales 330 de subbanda. (En este ejemplo de simplificación, se supone que el banco de filtros de análisis comprende 3 canales.) Las señales de subbanda de la señal 330 de audio de dominio de frecuencia puede replicarse entonces para obtener tres señales 335 de subbanda adicionales de manera que la señal 320 de audio de dominio de frecuencia comprenda las tres señales 330 de subbanda originales y las tres señales 335 de subbanda adicionales generadas. Entonces, se generan dos señales 338 de subbanda adicionales más, por ejemplo señales de cero, señales de subbanda pseudoaleatorias o señales de subbanda copiadas. La señal de audio de dominio de frecuencia se transforma entonces de vuelta al dominio de tiempo que da como resultado una señal 350 de audio de dominio de tiempo que tiene una tasa de muestreo que es  $8/3$  veces la tasa de muestreo de la señal 310 de audio de dominio de tiempo original.

40 La figura 4 ilustra un aparato según una realización adicional. El aparato comprende un procesador 405 de señal y un configurador 408. El procesador 405 de señal comprende un módulo 210 decodificador de núcleo, un banco 220 de filtros de análisis, un generador 230 de subbanda y un banco 240 de filtros de síntesis, que corresponden a las unidades respectivas en la realización de la figura 2. El procesador 405 de señal comprende además un decodificador 410 de MPEG envolvente (decodificador de MPS) para decodificar la señal de audio procesada previamente para obtener una señal de audio procesada previamente con canales estéreo o envolventes. El generador 230 de subbanda está adaptado para alimentar la señal de audio procesada previamente de dominio de frecuencia al decodificador 410 de MPEG envolvente después de que las señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia se hayan creado y añadido a la señal de audio procesada previamente de dominio de frecuencia.

50 La figura 5a ilustra un módulo decodificador de núcleo según una realización. El módulo decodificador de núcleo comprende un primer decodificador 510 de núcleo y un segundo decodificador 520 de núcleo. El primer decodificador 510 de núcleo está adaptado para operar en un dominio de tiempo y en el que el segundo decodificador 520 de núcleo está adaptado para operar en un dominio de frecuencia. En la figura 5a, el primer decodificador 510 de núcleo es un decodificador de ACELP y el segundo decodificador 520 de núcleo es un decodificador de transformada FD, por ejemplo un decodificador de transformada AAC. En una realización alternativa, el segundo decodificador 520 de núcleo

es un decodificador de transformada TCX. Dependiendo de si una parte de señal de audio asp que llega contiene datos de voz u otros datos de audio, la parte de señal de audio asp que llega se procesa o bien por el decodificador 510 de ACELP o bien por el decodificador 520 de transformada FD. La salida del módulo decodificador de núcleo es una parte procesada previamente de la señal de audio pp-asp.

5 La figura 5b ilustra un aparato para procesar una señal de audio según la realización de la figura 4 con un módulo decodificador de núcleo según la figura 5a.

10 En una realización, el tamaño de supertrama para el códec de ACELP se reduce de 1024 a 768 muestras. Esto podría realizarse combinando 4 tramas de ACELP de tamaño 192 (3 subtramas de tamaño 64) dando una trama de codificador de núcleo de tamaño 768 (previamente: se combinaron 4 tramas de ACELP de tamaño 256 con una trama de codificador de núcleo de tamaño 1024). La figura 6a ilustra una supertrama 605 de ACELP que comprende 4 tramas 610 de ACELP. Cada una de las tramas 610 de ACELP comprende 3 subtramas 615.

Otra solución para conseguir un tamaño de trama de codificador de núcleo de 768 muestras sería, por ejemplo, combinar 3 tramas de ACELP de tamaño 256 (4 subtramas de tamaño 64). La figura 6b ilustra una supertrama 625 de ACELP que comprende 3 tramas 630 de ACELP. Cada una de las tramas 630 de ACELP comprende 4 subtramas 635.

15 La figura 7b expone el ajuste adicional propuesto desde la perspectiva de un decodificador y lo compara con el ajuste de USAC tradicional. Las figuras 7a y 7b exponen la estructura del decodificador tal como se usa normalmente en los puntos de operación como 24 kbit/s o 32 kbit/s.

20 En la figura 7a, que ilustra el RM9 de USAC (modelo de referencia 9 de USAC), ajuste por defecto, una trama de señal de audio se introduce en un banco 710 de filtros de análisis de QMF. El banco 710 de filtros de análisis de QMF tiene 32 canales. El banco 710 de filtros de análisis de QMF está adaptado para transformar una señal de audio de dominio de tiempo a un dominio de frecuencia, en el que la señal de audio de dominio de frecuencia comprende 32 subbandas. La señal de audio de dominio de frecuencia se introduce entonces en un sobremuestreador 720. El sobremuestreador 720 está adaptado para sobremuestrear la señal de audio de dominio de frecuencia mediante un factor 2 de sobremuestreo. Por tanto, se genera por el sobremuestreador una señal de salida de sobremuestreador de dominio de frecuencia que comprende 64 subbandas. El sobremuestreador 720 es un sobremuestreador de SBR (replicación de banda espectral). Tal como ya se mencionó, se emplea una replicación de banda espectral para generar subbandas de frecuencia más alta desde subbandas de frecuencia más baja que se introducen en el replicador de banda espectral.

25 La señal de audio de dominio de frecuencia sobremuestreada se alimenta entonces a un decodificador 730 de MPEG envolvente (MPS). El decodificador 730 de MPS está adaptado para decodificar una señal envolvente de mezcla descendente para derivar canales de dominio de frecuencia de una señal envolvente. Por ejemplo, el decodificador 730 de MPS puede estar adaptado para generar 2 canales envolventes de dominio de frecuencia de mezcla ascendente de una señal envolvente de dominio de frecuencia. En otra realización, el decodificador 730 de MPS puede estar adaptado para generar 5 canales envolventes de dominio de frecuencia de mezcla ascendente de una señal envolvente de dominio de frecuencia. Los canales de la señal envolvente de dominio de frecuencia se alimentan entonces al banco 30 740 de filtros de síntesis de QMF. El banco 740 de filtros de síntesis de QMF está adaptado para transformar los canales de la señal envolvente de dominio de frecuencia a un dominio de tiempo para obtener canales de dominio de tiempo de la señal envolvente.

35 Tal como puede observarse, el decodificador de USAC opera en su ajuste por defecto como sistema 2:1. El códec de núcleo opera en la granularidad de 1024 muestras/trama a la mitad de la tasa de muestreo de salida  $f_{\text{salida}}$ . El sobremuestreo mediante un factor de 2 se realiza implícitamente dentro de la herramienta SBR, combinando un banco de filtros de QMF de análisis de 32 bandas con un banco de QMF de síntesis de 64 bandas que se ejecuta a la misma velocidad. La herramienta SBR emite tramas de tamaño 2048 a  $f_{\text{salida}}$ .

La figura 7b ilustra el ajuste extra propuesto para USAC. Están ilustrados un banco 750 de filtros de análisis de QMF, un sobremuestreador 760, un decodificador 770 de MPS y un banco 780 de filtros de síntesis.

45 A diferencia del ajuste por defecto, el códec de USAC opera en el ajuste extra propuesto como sistema 8/3. El codificador de núcleo se ejecuta a 3/8 de la tasa de muestreo de salida  $f_{\text{salida}}$ . En el mismo contexto, el tamaño de trama de codificador de núcleo se redujo a escala mediante un factor de  $\frac{3}{4}$ . Mediante una combinación de un banco de filtros de QMF de análisis de 24 bandas y un banco de filtros de síntesis de 64 bandas dentro de la herramienta SBR, puede conseguirse una tasa de muestreo de salida de  $f_{\text{salida}}$  a una longitud de trama de 2048 muestras.

50 Este ajuste permite una granularidad temporal bastante aumentada tanto para el codificador de núcleo como para herramientas adicionales: mientras que herramientas tales como SBR y MPEG envolvente pueden operarse a una mayor tasa de muestreo, se reduce la tasa de muestreo de codificador de núcleo y, en cambio, se acorta la longitud de trama. De este modo, todos los componentes pueden funcionar en su entorno óptimo.

En una realización, un codificador de AAC empleado como codificador de núcleo todavía puede determinar factores de escala basándose en una tasa de muestreo de  $\frac{1}{2} f_{\text{salida}}$ , aunque el codificador de AAC opere en  $\frac{3}{8}$  de la tasa de muestreo de salida  $f_{\text{salida}}$ .

5 La tabla a continuación proporciona números detallados sobre tasas de muestreo y duración de trama para el USAC tal como se usa en el codificador de calidad de referencia de USAC. Tal como puede observarse, la duración de trama en el ajuste nuevo propuesto puede reducirse en casi el 25%, lo que lleva a efectos positivos para todas las señales no estacionarias, puesto que la dispersión del ruido de codificación también puede reducirse mediante la misma relación. Esta reducción puede conseguirse sin aumentar la frecuencia de muestreo de codificador de núcleo, que habría movido la herramienta ACELP fuera de su intervalo de operación optimizado.

	Codificador de núcleo de tasa de muestreo	SBR de tasa de muestreo	Duración por trama
USAC por defecto	17075 Hz	34150 Hz	60 ms
Ajuste nuevo propuesto	16537,5 Hz	44100 Hz	46 ms

10

La tabla ilustra las tasas de muestreo y la duración de trama por defecto y un ajuste nuevo propuesto tal como se usan en el codificador de calidad de referencia a 24 kbit/s.

A continuación, se describen en más detalle las modificaciones necesarias para el decodificador de USAC, para implementar el ajuste nuevo propuesto.

15 Con respecto al codificador de transformada, los tamaños de trama más cortos pueden conseguirse fácilmente ajustando a escala los tamaños de transformada y de ventana mediante un factor de  $\frac{3}{4}$ . Mientras que el codificador de FD en el modo convencional opera con tamaños de transformada de 1024 y 128, se introducen transformadas adicionales de tamaño 768 y 96 por el ajuste nuevo. Para la TCX, se requieren transformadas adicionales de tamaño de 768, 384 y 192. Aparte de los tamaños de transformada nuevos que se especifican según coeficientes de ventana, el codificador de transformada puede permanecer sin cambios.

20

Con respecto a la herramienta ACELP, es necesario que el tamaño de trama total se adapte a 768 muestras. Una manera de conseguir este propósito es dejar la estructura global de la trama sin cambios entrando 4 tramas de ACELP de 192 muestras en cada trama de 768 muestras. La adaptación al tamaño de trama reducido se consigue disminuyendo el número de subtramas por trama de 4 a 3. La longitud de subtrama de ACELP no cambia a 64 muestras. Con el fin de permitir el número reducido de subtramas, se codifica la información de tono usando un esquema ligeramente diferente: se codifican tres valores de tono usando un esquema absoluto-relativo-relativo usando 9, 6 y 6 bits respectivamente en lugar de un esquema absoluto-relativo-absoluto-relativo usando 9, 6, 9 y 6 bits en el modelo convencional. Sin embargo, son posibles otras maneras de codificar la información de tono. Los otros elementos del códec de ACELP, tales como los libros de códigos de ACELP así como los diversos cuantificadores (filtros de LPC, ganancias, etc.), se dejan sin cambios.

25

30

Otra manera de conseguir un tamaño de trama total de 768 muestras sería combinar tres tramas de ACELP de tamaño 256 para una trama de codificador de núcleo de tamaño 768.

La funcionalidad de la herramienta SBR permanece sin cambios. Sin embargo, además del QMF de banda de análisis de 32 bandas, es necesario un QMF de análisis de 24 bandas, para permitir un sobremuestreo de factor  $\frac{8}{3}$ .

35 A continuación, se explica el impacto del punto de operación extra propuesto en la complejidad computacional. En primer lugar esto se realiza por herramienta de códec y se resume al final. La complejidad se compara con el modo de tasa de muestreo bajo por defecto y con un modo de mayor tasa de muestreo, tal como se usa por el codificador de calidad de referencia de USAC a mayores tasas de transmisión de bits lo que puede compararse con el ajuste HE-AACv2 correspondiente para estos puntos de operación.

40 Con respecto al codificador de transformada, la complejidad de las partes de codificador de transformada se ajusta a escala con la tasa de muestreo y la longitud de transformada. Las tasas de muestreo de codificador de núcleo propuestas permanecen casi iguales. Los tamaños de transformada se reducen en un factor de  $\frac{3}{4}$ . De este modo, se reduce la complejidad computacional en casi el mismo factor, suponiendo un enfoque de base mixta para las FFT subyacentes. Generalmente, se espera que la complejidad del decodificador basado en transformada se reduzca ligeramente en comparación con el punto de operación de USAC actual y se reduzca en un factor de % en comparación con un modo de operación de alto muestreo.

45

Con respecto a ACELP, la complejidad de las herramientas ACELP se compone principalmente por las siguientes operaciones:

5 Decodificación de la excitación: la complejidad de esa operación es proporcional al número de subtramas por segundo, que a su vez es directamente proporcional a la frecuencia de muestreo de codificador de núcleo (permaneciendo el tamaño de subtrama sin cambios a 64 muestras). Por tanto, es casi igual que el ajuste nuevo.

Filtrado de LPC y otras operaciones de síntesis, incluyendo el posfiltro de bajos: la complejidad de esta operación es directamente proporcional a la frecuencia de muestreo de codificador de núcleo y, por tanto, es casi igual.

10 Generalmente, se espera que no cambie la complejidad esperada del decodificador de ACELP en comparación con el punto de operación de USAC actual y que se reduzca en un factor de  $\frac{3}{4}$  en comparación con un modo de operación de alto muestreo.

Con respecto a SBR, los principales contribuidores a la complejidad de SBR son los bancos de filtros de QMF. En este caso, la complejidad se ajusta a escala con la tasa de muestreo y el tamaño de transformada. En particular, la complejidad del banco de filtros de análisis se reduce en casi un factor de  $\frac{3}{4}$ .

15 Con respecto a MPEG envolvente, la complejidad de la parte de MPEG envolvente se ajusta a escala con la tasa de muestreo. El modo de operación extra propuesto no tiene impacto directo en la complejidad de la herramienta MPEG envolvente.

20 En total se encontró que la complejidad del modo de operación nuevo propuesto es ligeramente más compleja en comparación con el modo de baja tasa de muestreo, pero inferior a la complejidad del decodificador de USAC, cuando se ejecuta en un modo de mayor tasa de muestreo (RM9 de USAC, SR alta: 13,4 MOPS, punto de operación nuevo propuesto: 12,8 MOPS).

Para el punto de operación sometido a prueba, se evalúa la complejidad tal como sigue:

RM9 de USAC, operado a 34,15 kHz: aproximadamente 4,6 WMOPS;

RM9 de USAC, operado a 44,1 kHz: aproximadamente 5,6 WMOPS;

punto de operación nuevo propuesto: aproximadamente 5,0 WMOPS.

25 Puesto que se espera que sea necesario que un decodificador de USAC pueda gestionar tasas de muestreo de hasta 48 kHz en su configuración por defecto, no se espera ningún inconveniente por este punto de operación nuevo propuesto.

30 Con respecto a la demanda de memoria, el modo de operación extra propuesto requiere el almacenamiento de prototipos de ventana de MDCT adicionales, que suman en total una demanda de ROM adicional de menos de 900 palabras (32 bits). A la luz de la demanda de ROM de decodificador total, que es casi 25 kWord, esto parece ser insignificante.

Los resultados de la prueba de escucha muestran una mejora significativa para los elementos de prueba de música y mixtos, sin degradar la calidad de los elementos de voz. Se pretende que este ajuste extra sea un modo de operación adicional del códec de USAC.

35 Se llevó a cabo una prueba de escucha según la metodología de MUSHRA para evaluar el rendimiento del ajuste nuevo propuesto a 24 kbit/s mono. En la prueba se incorporaron las siguientes condiciones: referencia oculta; anclaje de paso bajo de 3,5 kHz; calidad de referencia USAC WD7 (WD7@34,15 kHz); USAC WD7 operado a una alta tasa de muestreo (WD7@44,1 kHz); y calidad de referencia USAC WD7, ajuste nuevo propuesto (WD7\_CE@44,1 kHz).

40 La prueba cubrió los 12 elementos de prueba del conjunto de prueba de USAC y los siguientes elementos adicionales: si02: castañuelas; velvet: música electrónica; y xylophone: caja de música.

Las figuras 8a y 8b ilustran los resultados de la prueba. 22 sujetos participaron en la prueba de escucha. Se usó una distribución de probabilidad t de Student para la evaluación.

45 Para la evaluación de las puntuaciones promedio (nivel de importancia del 95%) puede observarse que WD7 operando a una mayor tasa de muestreo de 44,1 kHz funciona significativamente peor que WD7 para dos elementos (es01, HarryPotter). Entre WD7 y el WD7 que presenta la tecnología, no puede observarse ninguna diferencia significativa.

Para la evaluación de las puntuaciones diferenciales puede observarse que WD7 operando a 44,1 kHz funciona peor que WD7 para 6 elementos (es01, louis\_raquin, tel, WeddingSpeech, HarryPotter, SpeechOverMusic\_4) y se promedia a través de todos los elementos. Los elementos con los que funciona peor incluyen todos los elementos de voz pura y

dos de los elementos de voz/música mixtos. Puede observarse además que WD7 operando a 44,1 kHz funciona significativamente mejor que WD7 para cuatro elementos (twink, salvation, si02, velvet). Todos estos elementos contienen partes significativas de señales de música o se clasifican como música.

5 Para la tecnología a prueba puede observarse que funciona mejor que WD7 para cinco elementos (twink, salvation, te15, si02, velvet), y adicionalmente cuando se promedia a través de todos los elementos. Todos los elementos con los que funciona mejor contienen partes significativas de señales de música o se clasifican como música. No pudo observarse ninguna degradación.

10 Mediante las realizaciones descritas anteriormente, se proporciona un ajuste nuevo para tasas de transmisión de bits de USAC intermedias. Este ajuste nuevo permite que el códec de USAC aumente su granularidad temporal para todas las herramientas relevantes, tales como codificadores de transformada, SBR y MPEG envolvente, sin sacrificar la calidad de la herramienta ACELP. De este modo, puede mejorarse la calidad para el intervalo de tasa de transmisión de bits intermedia, en particular para señales de música y mixtas que presentan una estructura altamente temporal. Además, los sistemas de USAC ganan flexibilidad, puesto que ahora puede usarse el códec de USAC que incluye la herramienta ACELP a un intervalo más amplio de tasas de muestreo, tal como 44,1 kHz.

15 La figura 9 ilustra un aparato para procesar una señal de audio. El aparato comprende un procesador 910 de señal y un configurador 920. El procesador 910 de señal está adaptado para recibir una primera trama 940 de señal de audio que tiene un primer número configurable de muestras 945 de la señal de audio. Además, el procesador 910 de señal está adaptado para submuestrear la señal de audio mediante un factor de submuestreo configurable para obtener una señal de audio procesada. Además, el procesador de señal está adaptado para emitir una segunda trama 950 de señal de audio que tiene un segundo número configurable de muestras 955 de la señal de audio procesada.

20 El configurador 920 está adaptado para configurar el procesador 910 de señal basándose en información de configuración ci2 de manera que el factor de submuestreo configurable sea igual a un primer valor de submuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación. Además, el configurador 920 está adaptado para configurar el procesador 910 de señal de manera que el factor de submuestreo configurable sea igual a un segundo valor de submuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente. El primer o el segundo valor de relación no es un valor de número entero.

Por ejemplo, puede emplearse un aparato según la figura 9 en el proceso de codificación.

30 Aunque se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, en el que un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

35 La señal descompuesta de la invención puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como Internet.

40 Dependiendo de determinados requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tenga señales de control electrónicamente legibles almacenadas en el mismo, que actúan conjuntamente (o pueden actuar conjuntamente) con un sistema informático programable de manera que se realice el método respectivo.

45 Algunas realizaciones según la invención comprenden un soporte de datos no transitorio que tiene señales de control electrónicamente legibles, que pueden actuar conjuntamente con un sistema informático programable, de manera que se realice uno de los métodos descritos en el presente documento.

Generalmente, las realizaciones de la presente invención pueden implementarse como producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un soporte legible por máquina.

50 Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en el presente documento, almacenado en un soporte legible por máquina.

En otras palabras, una realización del método de la invención es, por tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

5 Una realización adicional de los métodos de la invención es, por tanto, un soporte de datos (o un medio de almacenamiento digital o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en el presente documento.

10 Una realización adicional del método de la invención es, por tanto, un flujo de datos o una secuencia de señales que representa el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales puede configurarse, por ejemplo, para que se transfiera a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo de lógica programable, configurado para o adaptado para realizar uno de los métodos descritos en el presente documento.

Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los métodos descritos en el presente documento.

15 En algunas realizaciones, puede usarse un dispositivo de lógica programable (por ejemplo, una disposición de puertas programables en campo) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una disposición de puertas programables en campo puede actuar conjuntamente con un microprocesador con el fin de realizar uno de los métodos descritos en el presente documento. Generalmente, los métodos se realizan preferiblemente mediante cualquier aparato de hardware.

20 Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento resultarán evidentes para otros expertos en la técnica. Por tanto, se pretende que estén limitadas sólo por el alcance de las reivindicaciones de la patente a continuación y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

**REIVINDICACIONES**

1. Aparato para procesar una señal de audio, que comprende:

un procesador (110; 205; 405) de señal que está adaptado para recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio, que está adaptado para sobremuestrear la señal de audio mediante un factor de sobremuestreo configurable para obtener una señal de audio procesada y que está adaptado para emitir una segunda trama de señal de audio que tiene un segundo número configurable de muestras de la señal de audio procesada; y

un configurador (120; 208; 408) que está adaptado para configurar el procesador (110; 205; 405) de señal,

en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal basándose en información de configuración de manera que el factor de sobremuestreo configurable sea igual a un primer valor de sobremuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación y en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el factor de sobremuestreo configurable sea igual a un segundo valor de sobremuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente y en el que el primer o el segundo valor de relación no es un valor de número entero.

2. Aparato según la reivindicación 1, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el segundo valor de sobremuestreo diferente sea superior al primer valor de sobremuestreo, cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras es superior a la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras.

3. Aparato según la reivindicación 1 ó 2, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el factor de sobremuestreo configurable sea igual al primer valor de relación cuando la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el primer valor de relación y en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el factor de sobremuestreo configurable sea igual al segundo valor de relación diferente cuando la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene el segundo valor de relación diferente.

4. Aparato según una de las reivindicaciones anteriores, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el factor de sobremuestreo configurable sea igual a 2 cuando la primera relación tiene el primer valor de relación y en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el factor de sobremuestreo configurable sea igual a 8/3 cuando la segunda relación tiene el segundo valor de relación diferente.

5. Aparato según una de las reivindicaciones anteriores, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el primer número configurable de muestras sea igual a 1024 y el segundo número configurable de muestras sea igual a 2048 cuando la primera relación tiene el primer valor de relación y en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal de manera que el primer número configurable de muestras sea igual a 768 y el segundo número configurable de muestras sea igual a 2048 cuando la segunda relación tiene el segundo valor de relación diferente.

6. Aparato según una de las reivindicaciones anteriores, en el que el procesador (110; 205; 405) de señal comprende:

un módulo (210) decodificador de núcleo para decodificar la señal de audio para obtener una señal de audio procesada previamente,

un banco (220) de filtros de análisis que tiene varios canales de banco de filtros de análisis para transformar la primera señal de audio procesada previamente de un dominio de tiempo a un dominio de frecuencia para obtener una señal de audio procesada previamente de dominio de frecuencia que comprende una pluralidad de señales de subbanda,

un generador (230) de subbanda para crear y añadir señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia y

un banco (240) de filtros de síntesis que tiene varios canales de banco de filtros de síntesis para transformar la primera señal de audio procesada previamente del dominio de frecuencia al dominio de tiempo para obtener la señal de audio procesada, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de

señal configurando el número de un número de canales de banco de filtros de síntesis o el número de canales de banco de filtros de análisis de manera que el factor de sobremuestreo configurable sea igual a una tercera relación del número de canales de banco de filtros de síntesis con respecto al número de canales de banco de filtros de análisis.

5 7. Aparato según la reivindicación 6, en el que el generador (230) de subbanda es un replicador de banda espectral que está adaptado para replicar señales de subbanda del generador de señal de audio procesada previamente para crear las señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia.

10 8. Aparato según la reivindicación 6 ó 7, en el que el procesador (110; 205; 405) de señal comprende además un decodificador (410) de MPEG envolvente para decodificar la señal de audio procesada previamente para obtener una señal de audio procesada previamente que comprende canales estéreo o envolventes,

en el que el generador (230) de subbanda está adaptado para alimentar la señal de audio procesada previamente de dominio de frecuencia al decodificador (410) de MPEG envolvente después de que las señales de subbanda adicionales para la señal de audio procesada previamente de dominio de frecuencia se hayan creado y añadido a la señal de audio procesada previamente de dominio de frecuencia.

15 9. Aparato según una de las reivindicaciones 6 a 8, en el que el módulo (210) decodificador de núcleo comprende un primer decodificador (510) de núcleo y un segundo decodificador (520) de núcleo, en el que el primer decodificador (510) de núcleo está adaptado para operar en un dominio de tiempo y en el que el segundo decodificador (520) de núcleo está adaptado para operar en un dominio de frecuencia.

20 10. Aparato según la reivindicación 9, en el que el primer decodificador (510) de núcleo es un decodificador de ACELP y en el que el segundo decodificador (520) de núcleo es un decodificador de transformada FD o un decodificador de transformada TCX.

25 11. Aparato según la reivindicación 10, en el que el decodificador (510) de ACELP está adaptado para procesar la primera trama de señal de audio, en el que la primera trama de señal de audio tiene 4 tramas de ACELP y en el que cada una de las tramas de ACELP tiene 192 muestras de señal de audio, cuando el primer número configurable de muestras de la primera trama de señal de audio es igual a 768.

12. Aparato según la reivindicación 10, en el que el decodificador (510) de ACELP está adaptado para procesar la primera trama de señal de audio, en el que la primera trama de señal de audio tiene 3 tramas de ACELP y en el que cada una de las tramas de ACELP tiene 256 muestras de señal de audio, cuando el primer número configurable de muestras de la primera trama de señal de audio es igual a 768.

30 13. Aparato según una de las reivindicaciones anteriores, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal basándose en la información de configuración que indica al menos uno del primer número configurable de muestras de la señal de audio o el segundo número configurable de muestras de la señal de audio procesada.

35 14. Aparato según una de las reivindicaciones anteriores, en el que el configurador (120; 208; 408) está adaptado para configurar el procesador (110; 205; 405) de señal basándose en la información de configuración, en el que la información de configuración indica el primer número configurable de muestras de la señal de audio y el segundo número configurable de muestras de la señal de audio procesada, en el que la información de configuración es un índice de configuración.

15. Método para procesar una señal de audio, que comprende:

40 configurar un factor de sobremuestreo configurable,

recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio y

45 sobremuestrear la señal de audio mediante el factor de sobremuestreo configurable para obtener una señal de audio procesada, y que está adaptado para emitir una segunda trama de audio que tiene un segundo número configurable de muestras de la señal de audio procesada; y

50 en el que el factor de sobremuestreo configurable se configura basándose en información de configuración de manera que el factor de sobremuestreo configurable sea igual a un primer valor de sobremuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación y en el que el factor de sobremuestreo configurable está configurado de manera que el factor de sobremuestreo configurable sea igual a un segundo valor de sobremuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene

un segundo valor de relación diferente y en el que el primer o el segundo valor de relación no es un valor de número entero.

16. Aparato para procesar una señal de audio, que comprende:

5 un procesador (910) de señal que está adaptado para recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio, que está adaptado para submuestrear la señal de audio mediante un factor de submuestreo configurable para obtener una señal de audio procesada y que está adaptado para emitir una segunda trama de audio que tiene un segundo número configurable de muestras de la señal de audio procesada; y

un configurador (920) que está adaptado para configurar el procesador de señal,

10 en el que el configurador (920) está adaptado para configurar el procesador (910) de señal basándose en información de configuración de manera que el factor de submuestreo configurable sea igual a un primer valor de submuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación y en el que el configurador (920) está adaptado para configurar el procesador (910) de señal de manera que el factor de submuestreo configurable sea igual a un segundo valor de submuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente y en el que el primer o el segundo valor de relación no es un valor de número entero.

15 17. Aparato según la reivindicación 16, en el que el configurador está adaptado para configurar el procesador (910) de señal de manera que el primer valor de submuestreo sea menor que el segundo valor de submuestreo diferente, cuando la primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras es menor que la segunda relación del segundo número configurable de muestras con respecto al primer número configurable de muestras.

18. Método para procesar una señal de audio, que comprende:

configurar un factor de submuestreo configurable,

25 recibir una primera trama de señal de audio que tiene un primer número configurable de muestras de la señal de audio y

submuestrear la señal de audio mediante el factor de submuestreo configurable para obtener una señal de audio procesada y que está adaptado para emitir una segunda trama de audio que tiene un segundo número configurable de muestras de la señal de audio procesada; y

30 en el que el factor de submuestreo configurable se configura basándose en información de configuración de manera que el factor de submuestreo configurable sea igual a un primer valor de submuestreo cuando una primera relación del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un primer valor de relación y en el que el factor de submuestreo configurable está configurado de manera que el factor de submuestreo configurable sea igual a un segundo valor de submuestreo diferente, cuando una segunda relación diferente del segundo número configurable de muestras con respecto al primer número configurable de muestras tiene un segundo valor de relación diferente y en el que el primer o el segundo valor de relación no es un valor de número entero.

35 19. Programa informático para realizar el método según la reivindicación 15 ó 18, cuando el programa informático se ejecuta por un ordenador o procesador.

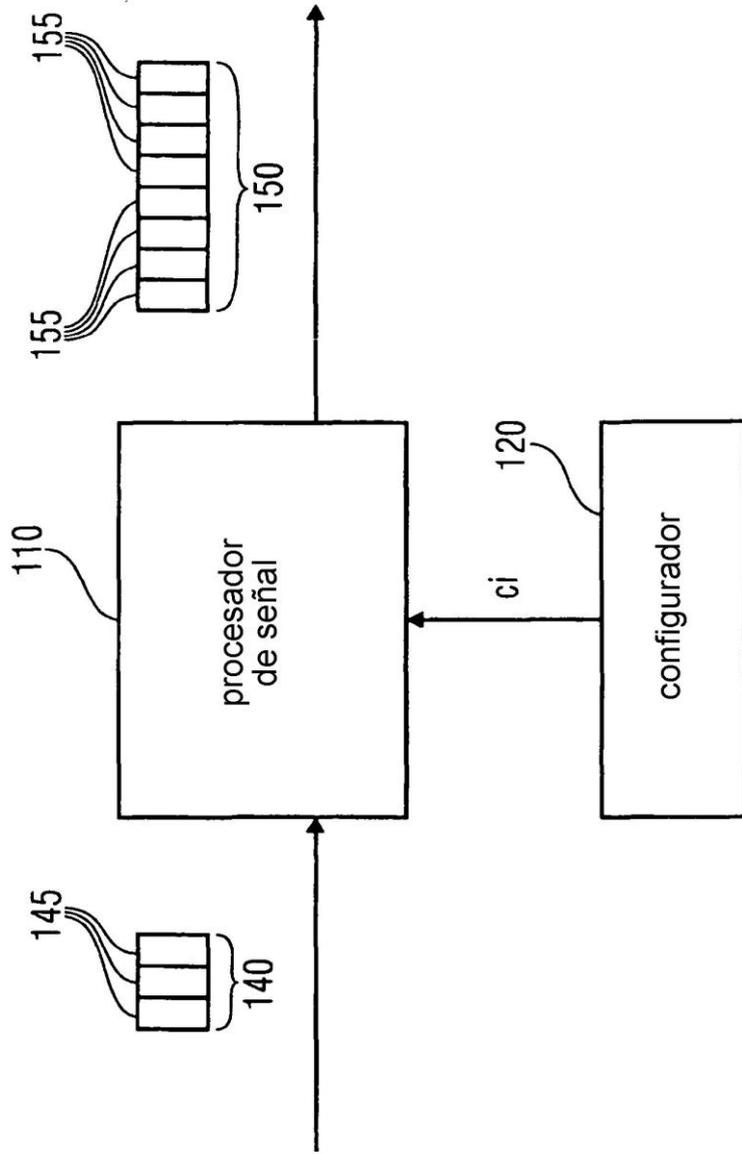


FIG 1

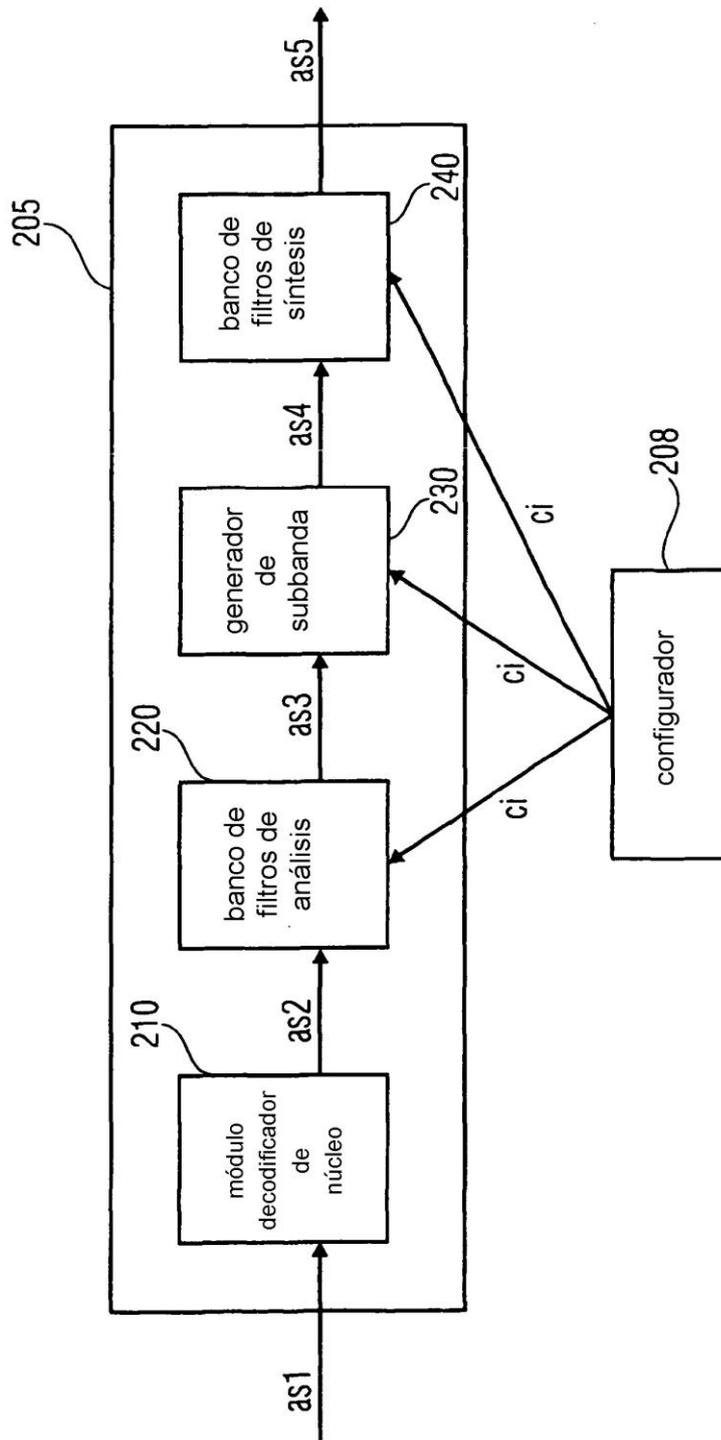


FIG 2

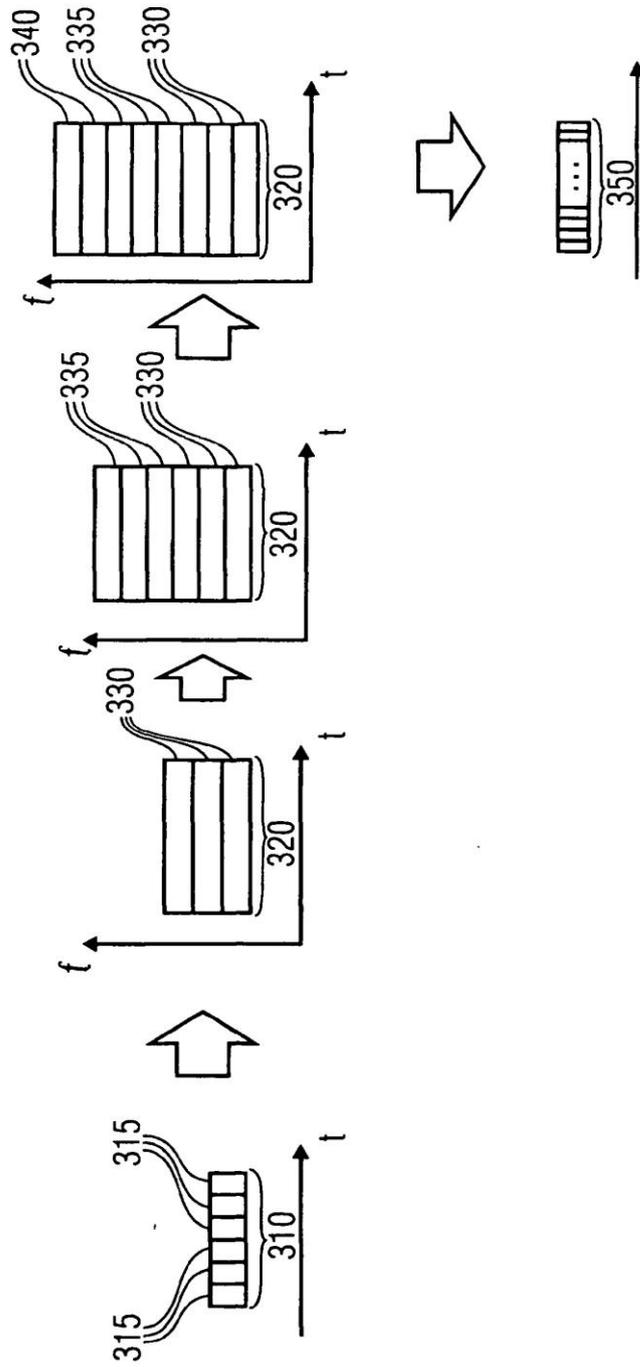


FIG 3

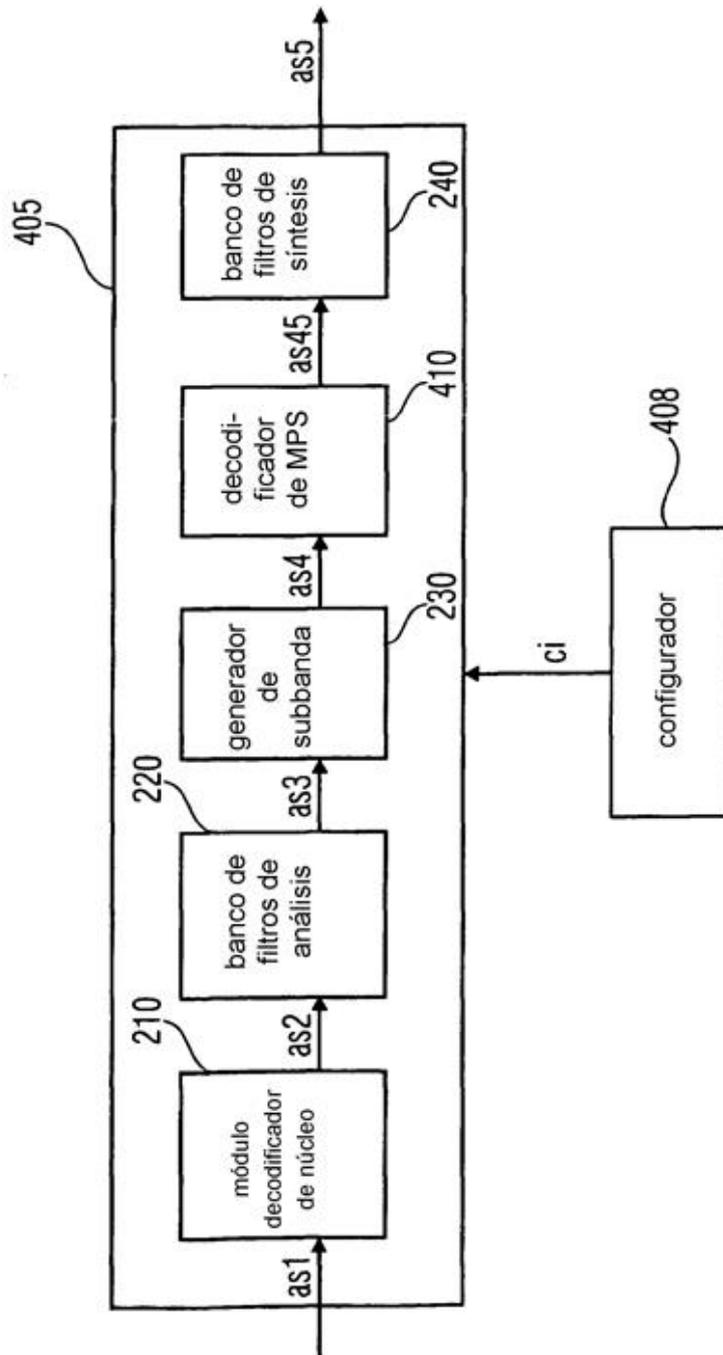


FIG 4

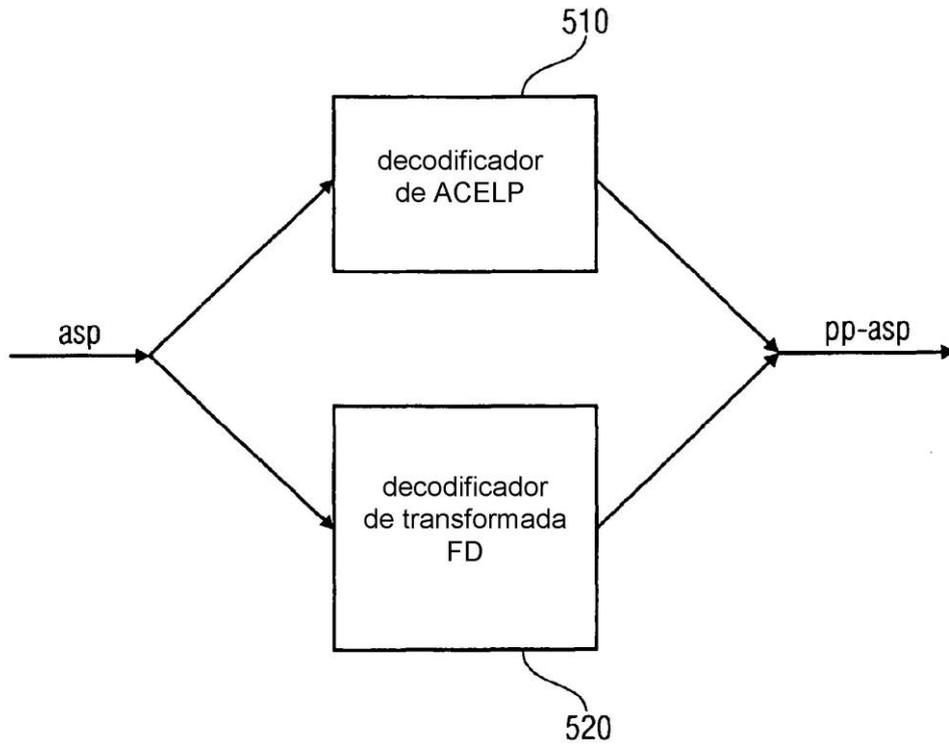


FIG 5A

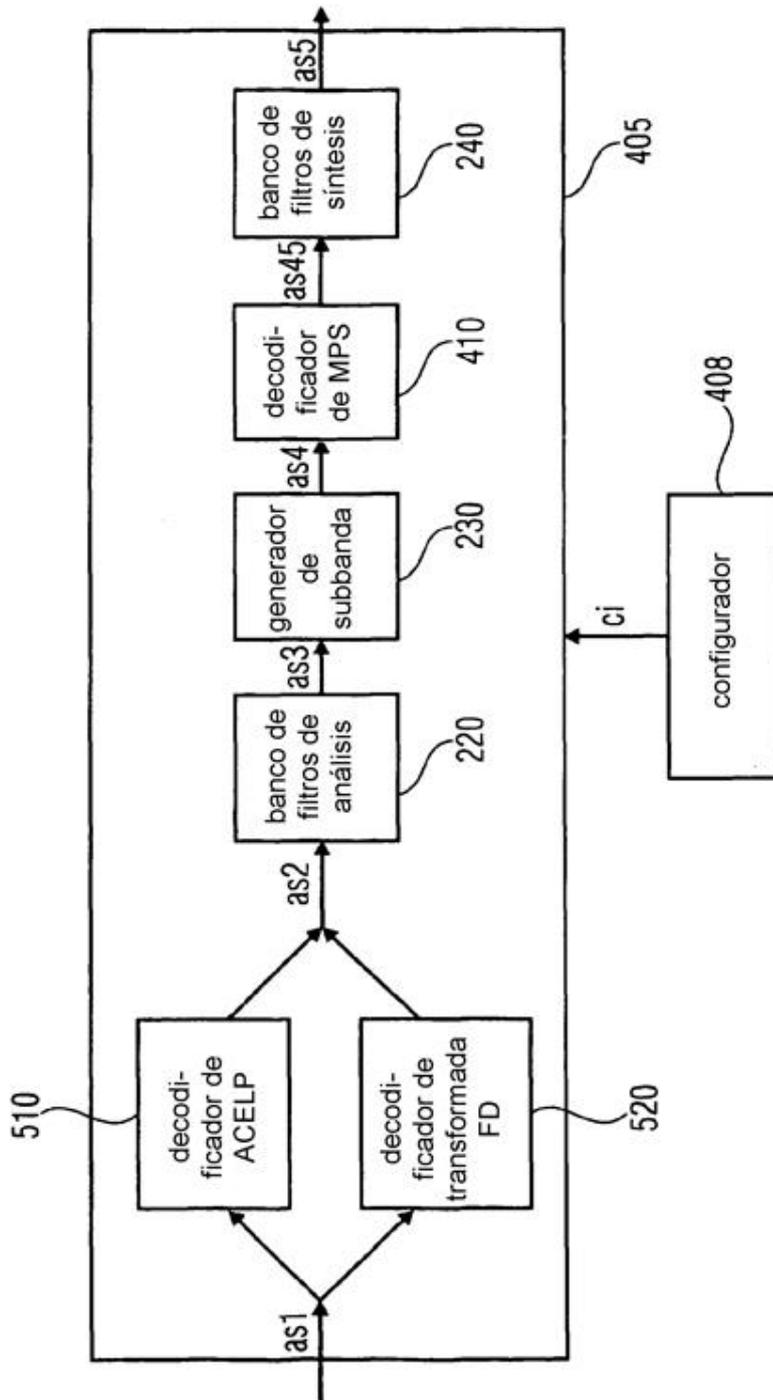


FIG 5B

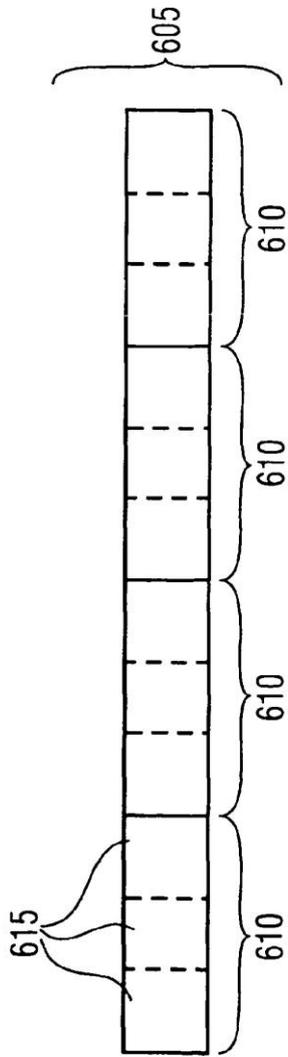


FIG 6A

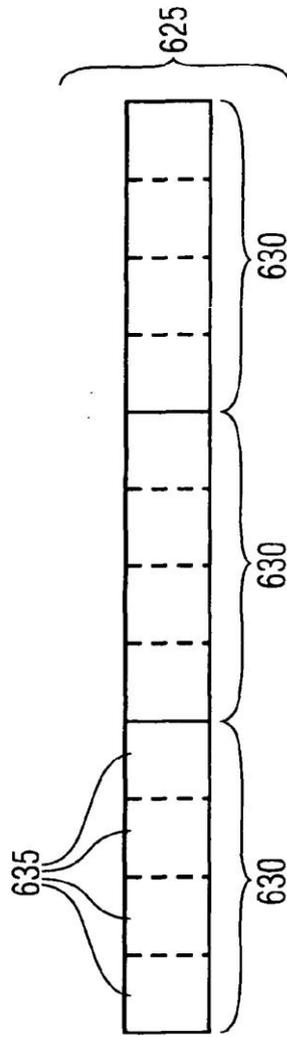


FIG 6B

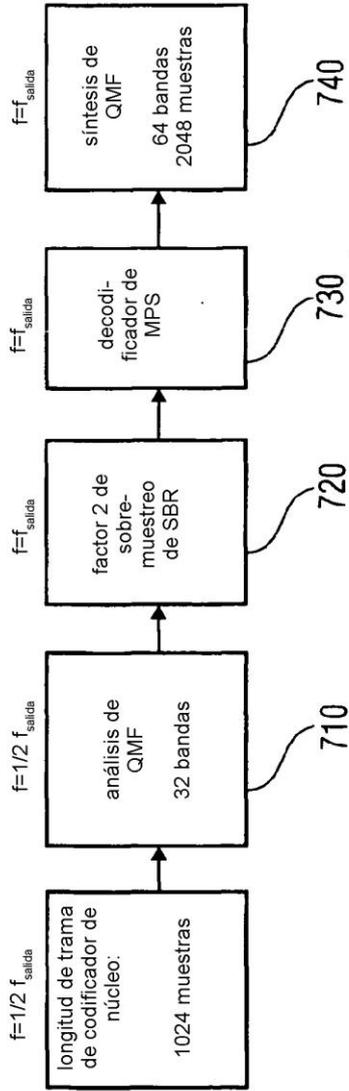


FIG 7A

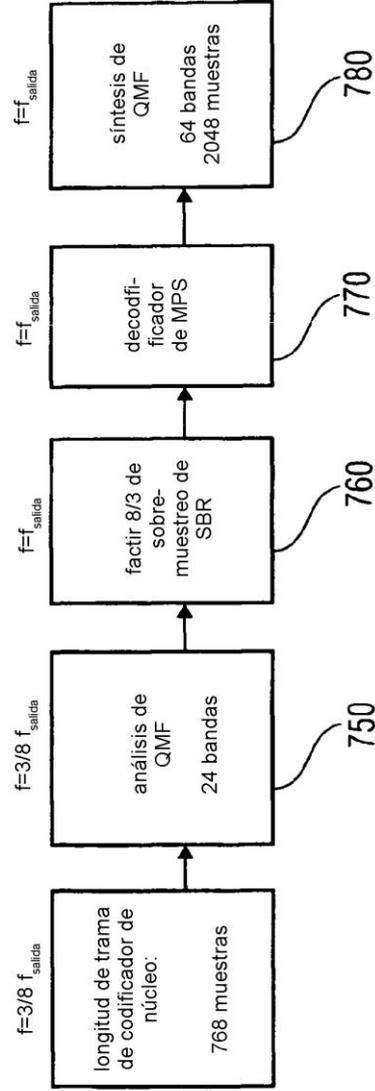
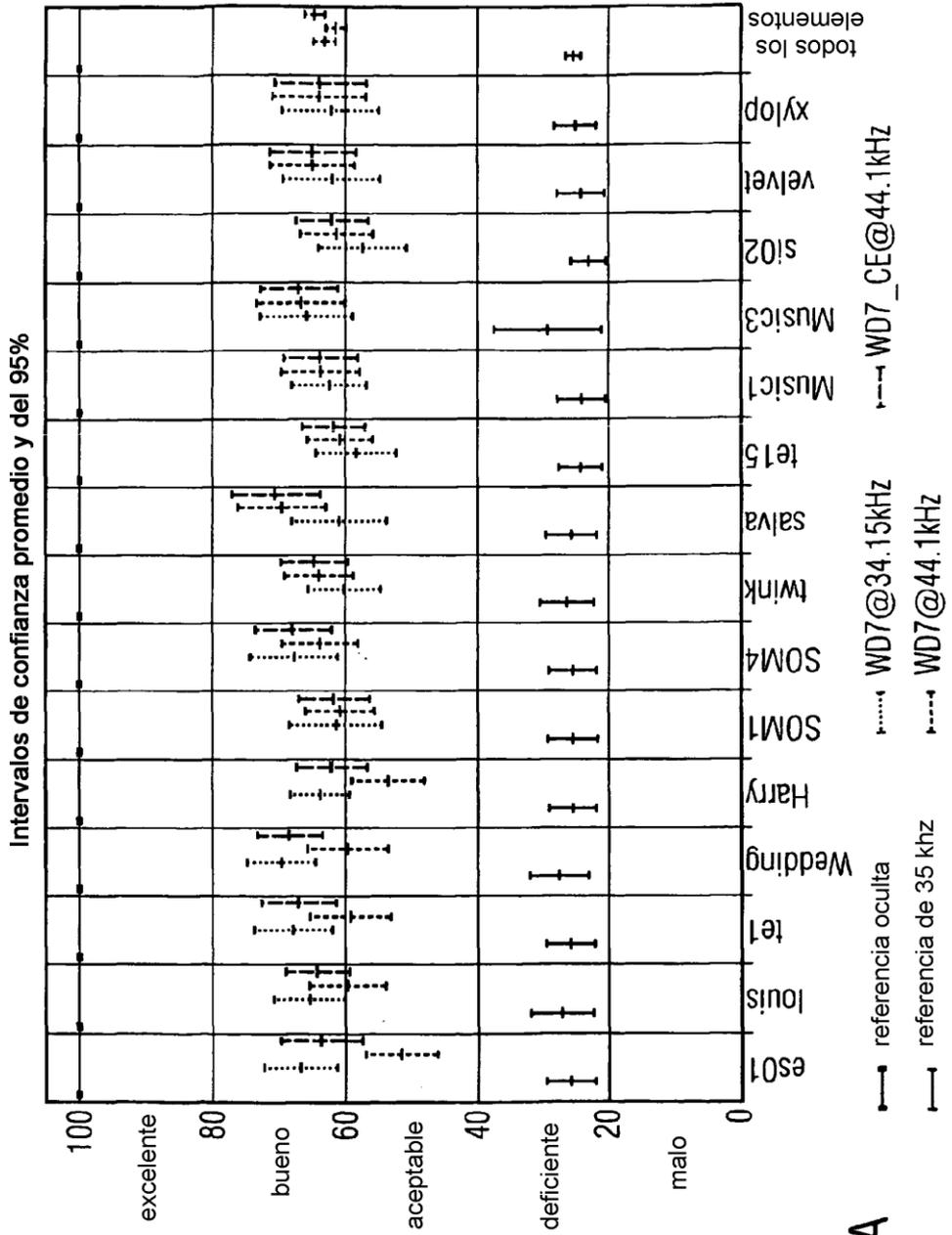


FIG 7B

2011-01-19 12:41:33

sesion i00\_test24m, 22 sujetos, distr. t de Student



**FIG 8A**

2011-01-19 12:41:33

session IQ\_test24m,22 sujetos, distr. t de Student

Intervalos de confianza promedio y del 95%  
(diferencias: Cond.-WD7@34.15kHz)

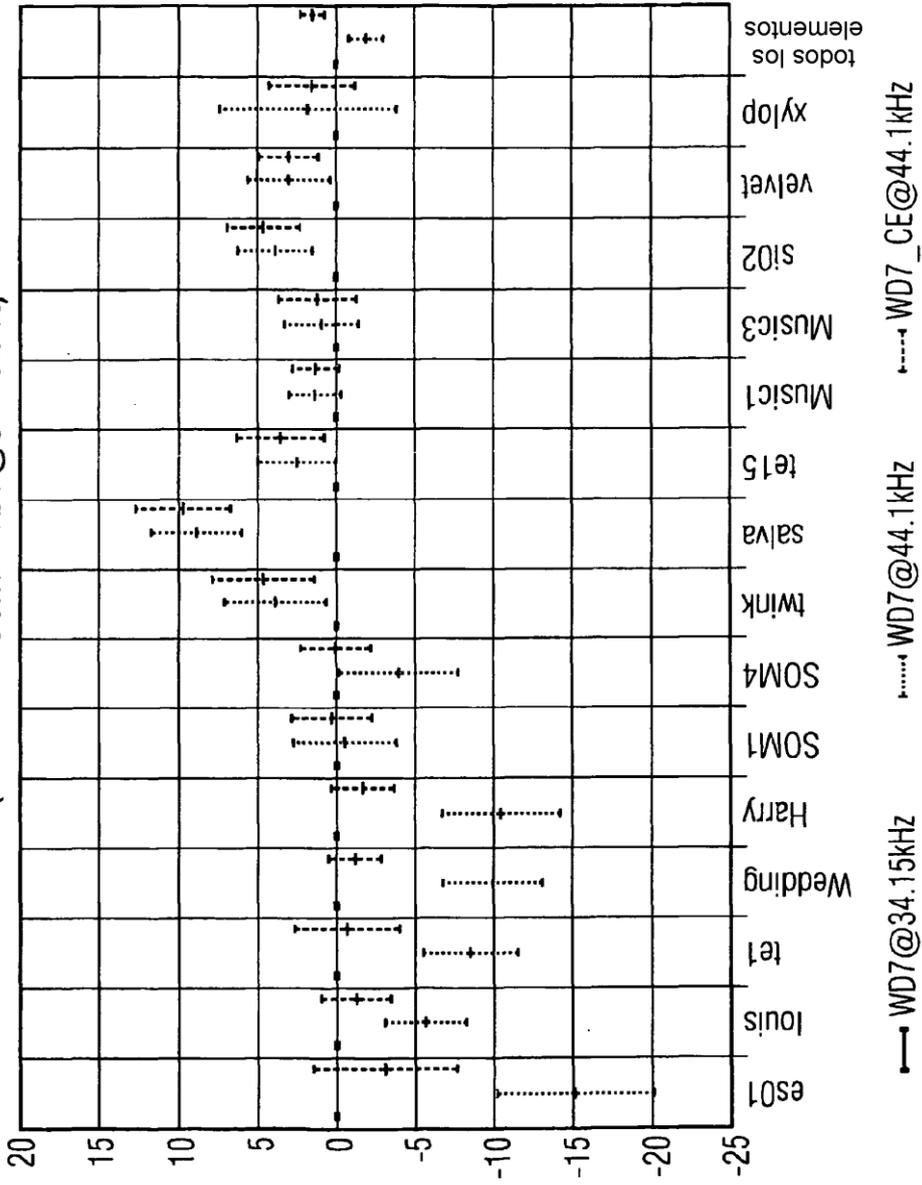


FIG 8B

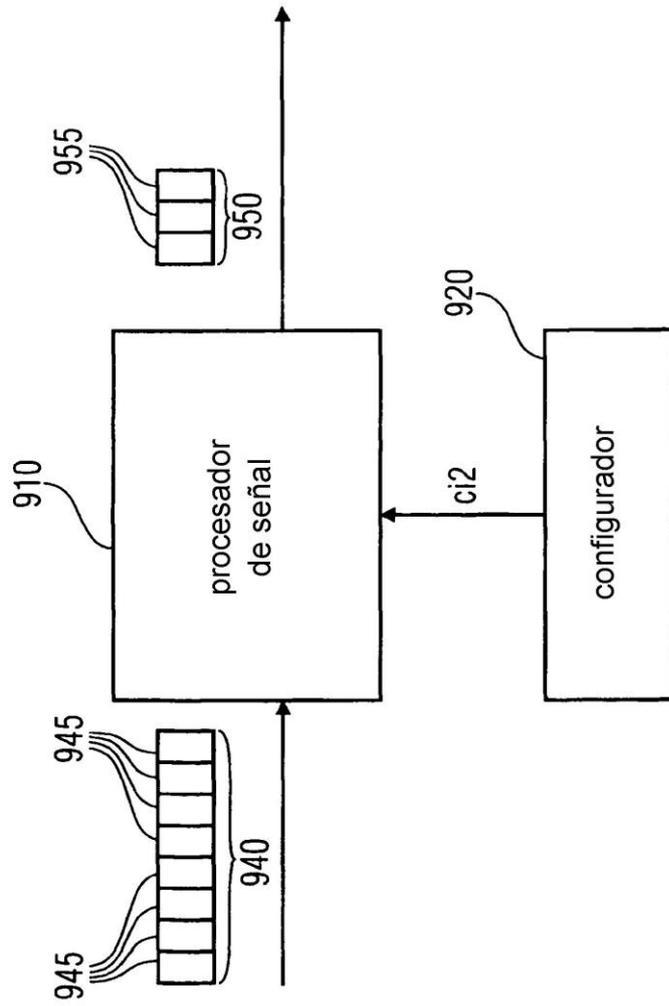


FIG 9