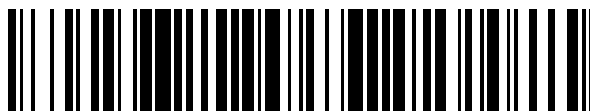


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 383 365**

51 Int. Cl.:
G10L 19/14 (2006.01)
G10L 19/08 (2006.01)
G10L 19/04 (2006.01)
G10L 19/12 (2006.01)
G10L 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07852271 .1**
96 Fecha de presentación: **14.12.2007**
97 Número de publicación de la solicitud: **2132733**
97 Fecha de publicación de la solicitud: **16.12.2009**

54 Título: **Post-filtro no causal**

30 Prioridad:
02.03.2007 US 892667 P

45 Fecha de publicación de la mención BOPI:
20.06.2012

45 Fecha de la publicación del folleto de la patente:
20.06.2012

73 Titular/es:
Telefonaktiebolaget LM Ericsson (publ)
164 83 Stockholm , SE

72 Inventor/es:
BRUHN, Stefan

74 Agente/Representante:
de Elzaburu Márquez, Alberto

ES 2 383 365 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

Post-filtro no causal

Campo técnico

5 La presente invención se refiere, en general, a la codificación y decodificación de señales de audio y/o de voz y, en particular, a la reducción del ruido de codificación.

Antecedentes

10 En general, la codificación de señales de audio y, de manera específica, la codificación de señales de voz, implica establecer una correspondencia entre una señal de entrada de audio o de voz analógica y una representación digital en un dominio de codificación y después otra correspondencia con una señal analógica de salida de audio o de voz. La representación digital va acompañada de la cuantización o discretización de los valores o parámetros que
15 consideran que la cuantización o discretización modifica o perturba los valores verdaderos de los parámetros con un "ruido de codificación". La técnica de la codificación de las señales de voz y audio consiste en realizar la codificación de tal manera que el efecto del ruido de codificación en la voz descodificada a una tasa de bits dada sea tan pequeño como sea posible. Sin embargo, la tasa de bits dada a la cual se codifica el habla define un límite inferior teórico al cual se puede reducir como máximo el ruido de codificación en el mejor de los casos. El objetivo es al menos hacer que el ruido de codificación sea tan inaudible como sea posible.

20 Un punto de vista adecuado acerca del ruido de codificación es suponer que sea un ruido coloreado o blanco aditivo. Hay una clase de métodos de mejora o realce que, tras descodificar la señal de audio o de conversación en el descodificador, modifican el ruido de codificación de tal forma que éste deviene menos audible, lo cual, por lo tanto, da como resultado, que la calidad de la señal de audio o de conversación resulte mejorada. Normalmente se denomina a tal tecnología "post-filtrado", lo cual quiere decir que la señal de audio o de conversación mejorada se deriva de algún procesado posterior después del descodificador real. Hay muchas publicaciones relacionadas con la mejora de las señales de voz mediante post-filtros. Algunos de los documentos fundamentales en ese sentido son
25 los de las referencias (1-4).

30 El principio de trabajo básico de los post-filtros de pitch (o altura tonal) es eliminar al menos partes del ruido de codificación que anega los valles espectrales entre armónicos de la voz. Esto se consigue, en general, mediante la superposición ponderada de la señal de voz descodificada con versiones cuyas desplazadas en el tiempo, de modo que el desplazamiento temporal corresponde al retraso de tono o período del habla. Esto da como resultado una atenuación del ruido de codificación no correlacionado respecto de la señal de voz deseada especialmente entre los armónicos de voz. El efecto descrito se puede obtener tanto mediante de estructuras de filtros repetitivas como no repetitivas. En la práctica se prefieren estructuras de filtros no repetitivas

35 En el contexto de la invención son pertinentes los post-filtros de pitch o de estructura fina. Su principio de trabajo básico es eliminar al menos partes del ruido de codificación que inunda los valles espectrales entre armónicos de la conversación con voces. Esto se consigue, en general, mediante la superposición ponderada de la señal de voz descodificada con versiones cuyas desplazadas en el tiempo, de modo que el desplazamiento temporal corresponde al retraso de pitch o período del habla o de la conversación. Preferentemente, también se incluyen versiones desplazadas en el tiempo en las muestras futuras de la señal de voz. Se describe en la referencia (5) un método de post-filtro de pitch no recurrente más reciente, en el cual los parámetros del pitch en la codificación de la señal se reutilizan en el post-filtrado de la muestra de la señal correspondiente. El método de post-filtro de pitch no iterativo de (5) se aplica también en los estándares de codificación de audio y de voz 3GPP AMR-WB+: 3GPP TS 26.290 "Audio codec processing functions: Extended Adaptive Multi-Rate – Wideband (AMR-WB+) codec: Transcoding functions" ("Funciones de procesado de programa de codificación y decodificación: códec de banda ancha y velocidad variable de adaptación ampliada: funciones de transcodificación") y 3GPP VMR-WB: 3GPP2 C.S0052-A,
40 "Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB), Service Options 62 y 63 for Spread Spectrum Systems" ("Programa de codificación y decodificación de voz de banda ancha multimodal de velocidad variable controlado por la fuente (VMR-WB), opciones de servicio 62 y 63 para sistemas de espectro ampliado"). En la referencia (6) se especifica un método de post-filtro de pitch. Esta patente describe el uso de voz sintetizada pasada y futura dentro de uno y de la misma unidad de información.

50 Un problema con los post-filtros de pitch que evalúan señales de voz futuras es que necesitan acceso a un período de pitch futuro de la señal de audio o de conversación descodificada. Generalmente es posible hacer que esta señal futura esté disponible para el post-filtro amplificando la señal de audio o de conversación descodificada. En aplicaciones del programa de codificación y decodificación (códec) de voz o de audio para conversaciones esto es, sin embargo, indeseable, puesto que aumenta el retraso algorítmico del programa códec y, por tanto, afectaría a la calidad de la comunicación y, en particular a su interactividad.
55

Compendio

Un objeto de la presente invención es proporcionar calidad de voz o audio mejorada a partir de dispositivos

descodificadores. Un objeto adicional de la presente invención es proporcionar montajes o dispositivos de post-filtros eficaces para usarlos con dispositivos descodificadores escalables, que no contribuyen de manera considerable a cualquier retraso adicional de la señal de audio o de conversación.

5 Los objetos anteriores se consiguen mediante dispositivos y métodos según las reivindicaciones de la patente adjuntas. En términos generales, según un primer aspecto, un dispositivo descodificador comprende un circuito de entrada del receptor para parámetros de señales codificadas en grupos o unidades de información y un descodificador conectado al circuito de entrada del receptor, dispuesto para proporcionar bloques de información de señales de audio descodificadas sobre la base de los parámetros. El circuito de entrada del receptor y/o el descodificador se disponen de tal forma que se establezca una diferencia de tiempo entre el momento en que los parámetros de una primera unidad de información están disponibles en el circuito de entrada del receptor y el momento en el que una señal de audio descodificada de la primera unidad de información está disponible en el circuito de salida del descodificador, de tal modo que la diferencia de tiempo corresponde a al menos una unidad de información. Se conecta un post-filtro al circuito de salida del descodificador y al circuito de entrada del receptor. El post-filtro se monta de modo que filtre las unidades de información de las señales de audio descodificadas para proporcionar una señal de salida en respuesta a los parámetros de la unidad de información posterior respectiva. El dispositivo descodificador comprende también un circuito de salida para la señal de salida, conectado al post-filtro.

20 Según un segundo aspecto, un método de descodificación comprende recibir parámetros de señales codificadas en base a unidades de información y descodificar los parámetros en unidades de información de señales de audio descodificadas. La recepción y/o la descodificación provocan una diferencia de tiempo entre el momento en que los parámetros de una primera unidad de información están disponibles tras la recepción y el momento en el que una señal de audio descodificada de la primera unidad de información está disponible tras la descodificación, diferencia de tiempo que corresponde a al menos una unidad de información. Las unidades de información de las señales de audio descodificadas se postfiltran para dar una señal de salida en respuesta a parámetros de una unidad de información posterior respectiva. El método comprende también producir la señal de salida.

25 Una ventaja que tiene la presente invención es que es posible mejorar la calidad de la señal de reconstrucción de los códecs de audio y de voz. Las mejoras se obtienen sin ninguna penalización en forma de retraso adicional, por ejemplo si el códec es un códec de audio y de voz escalable o si se usa en una aplicación VoIP (voz sobre protocolo de internet) con circuito compensador de fluctuaciones o perturbaciones en el terminal receptor. Es posible una mejora particular durante los sonidos transitorios, como por ejemplo en los principios de conversaciones.

30 **Breve descripción de los dibujos**

La invención, junto con objetos adicionales y ventajas de la misma, se puede entender mejor haciendo referencia a la siguiente descripción junto con los dibujos que la acompañan, en los cuales:

La figura 1 es una ilustración de una estructura básica de un códec de audio o voz con un post-filtro.

35 La figura 2 ilustra un esquema de bloques de una realización de un dispositivo descodificador según la presente invención.

La figura 3 ilustra un esquema de bloques de otra realización de un dispositivo descodificador según la presente invención.

La figura 4 ilustra un diagrama de flujo de las etapas de una realización de un método según la presente invención.

La figura 5 es un esquema de bloques de un códec de audio o voz escalable general.

40 La figura 6 es un esquema de bloques de otro códec de audio escalable en el cual las capas superiores soportan la codificación de las señales de audio que no son de voz.

La figura 7 ilustra un esquema de bloques de una realización de un dispositivo descodificador escalable según la presente invención.

45 La figura 8 ilustra un esquema de bloques de otra realización de un dispositivo descodificador escalable según la presente invención.

La figura 9 ilustra un esquema de bloques de otra realización más de un dispositivo descodificador escalable según la presente invención.

La figura 10 ilustra un esquema de bloques de otra realización de un dispositivo descodificador escalable según la presente invención.

50 La figura 11 ilustra un cálculo mejorado del parámetro director de pitch según la presente invención.

Descripción detallada

A lo largo de las presentes descripciones, las características iguales o que se corresponden directamente en diferentes figuras o realizaciones se denotarán con los mismos números de referencia.

5 Con el fin de comprender completamente la descripción detallada, puede ser necesario tener que definir algunos términos de manera más explícita con el fin de evitar confusiones. En la presente descripción, el término “parámetro” se usa como un término genérico, que sirve para cualquier clase de representación de la señal, incluyendo bits o una corriente o flujo de bits.

10 Con el fin de entender las ventajas que se consiguen con la presente invención, la descripción detallada comenzará con una revisión breve del post-filtrado en general. La figura 1 ilustra una estructura básica de un códec de voz o audio con un post-filtro. Una unidad emisora 1 comprende un codificador 10 que codifica una señal de audio o de conversación entrante en forma de una corriente de parámetros 4. De manera típica, los parámetros 4 se codifican y se transfieren a una unidad receptora 2. La unidad receptora 2 comprende un descodificador 20, que recibe los parámetros 4 que representan la señal de audio o de conversación original 3 y descodifica estos parámetros 4 en forma de una señal de audio o de conversación descodificada 5. Se pretende que la señal de audio o de conversación descodificada 5 sea tan similar como sea posible a la señal de audio o de conversación original 3. Sin embargo, la señal de audio o de conversación descodificada 5 siempre comprende ruido de codificación en cierto grado. La unidad receptora 2 comprende, además, un post-filtro 30, que recibe la señal de audio o de conversación descodificada 5 del descodificador 20, lleva a cabo un procedimiento de post-filtrado y produce como salida una señal de conversación o de audio descodificada post-filtrada 6.

20 La idea básica de los post-filtros es modelar la forma espectral del ruido de codificación de tal modo que se haga menos audible, lo cual explota esencialmente las propiedades de la percepción humana del sonido. En general, esto se hace de tal forma que el ruido se mueva hacia regiones de frecuencias menos sensibles perceptualmente en la cuales la señal de voz tiene potencia relativamente alta (picos espectrales) mientras que se retira de las regiones en las cuales la señal de voz tiene baja potencia (valles espectrales). Hay dos enfoques fundamentales en lo que se refiere a los post-filtros: filtros a corto plazo y filtros a largo plazo, también denominados, respectivamente, filtros formante o de conformación y de pitch o de estructura fina. Normalmente con el fin de conseguir buenos rendimientos se usan post-filtros adaptativos.

30 Como se ha mencionado previamente, en la presente invención son útiles los post-filtros de pitch o de estructura fina. La superposición de la señal de voz descodificada con versiones suyas desplazadas en el tiempo da como resultado una atenuación del ruido de codificación no correlacionado respecto de la señal de voz deseada, especialmente entre los armónicos de voz. El efecto descrito se puede obtener tanto con estructuras de filtros iterativas como no iterativas. Una de tales formas generales descrita en (4) está dada por:

$$H(z) = \frac{1 + \alpha z^{-T}}{1 - \beta z^{-T}},$$

donde T corresponde al período de pitch de la voz.

35 En la práctica se prefieren estructuras de filtros no iterativas. Se describe un método de post-filtro de pitch no iterativo más reciente en la solicitud de patente de Estados Unidos publicada número 2005/0165603, que se aplica en los estándares de codificación de audio y de voz 3GPP (3rd Generation Partnership Project, “Proyecto de Colaboración de 3ª Generación”) AMR-WB+ (Extended Adaptive Multi-Rate – Wideband códec, “códec de banda ancha de velocidad variable adaptativo ampliado”) (3GPP TS 26.290) y 3GPP2 VMR-WB (Variable Rate Multi-Mode Wideband codec, “códec de banda ancha multimodal de velocidad variable”) (3GPP2 C.S0052A: “Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB), Service Options 62 y 63 for Spread Spectrum Systems”) (“Programa de codificación y descodificación (códec) de voz de banda ancha multimodal de velocidad variable controlado por la fuente (VMR-WB), opciones de servicio 62 y 63 para sistemas de espectro ampliado”). Aquí, la idea básica es calcular en primer lugar una estimación del ruido de codificación $r(n)$ mediante la siguiente relación:

$$r(n) = y(n) - y_p(n)$$

donde $y(n)$ es la señal de audio o de conversación descodificada e $y_p(n)$ es la señal predicha calculada mediante la expresión:

$$y_p(n) = 0,5 \cdot (y(n-T) + y(n+T)) \tag{1}$$

50 En segundo lugar, se resta de la señal de voz una versión filtrada de paso bajo (es decir, con un filtro de frecuencia cero a frecuencia de un valor dado) (o de paso de banda) de la estimación del ruido, ponderada mediante algún factor α , lo que da como resultado una señal de audio o de conversación mejorada:

$$y_{mej}(n) = y(n) - \alpha \cdot LP\{r(n)\} \quad (2)$$

Una interpretación adecuada de la señal de ruido filtrada de paso bajo, si se invierte el signo, es mirarla como una señal de mejora que compensa la parte de baja frecuencia del ruido de codificación. El factor α adapta en respuesta a la correlación de la señal de predicción y de la señal de voz descodificada, a la energía de la señal predicha y a algún promedio temporal de la energía de la diferencia de la señal de voz y de la señal de la predicción.

Como se ha mencionado, un problema con los post-filtros de pitch de la técnica anterior que evalúan la expresión definida anteriormente $y_p(n) = 0,5 \cdot (y(n-T) + y(n+T))$ es que necesitan un período de pitch futuro de la señal de voz descodificada $y(n+T)$, añadiendo a cambio un retraso algorítmico. AMR-WB+ y VMR-WB resuelven este problema ampliando la señal de conversación o de audio descodificada al futuro, tomando como base la señal de audio o de conversación descodificada disponible y suponiendo que la señal de audio o de conversación se extenderá periódicamente con el período de pitch T . Suponiendo que la señal de audio o de conversación descodificada está disponible, exclusivamente, hasta el índice de tiempo n^+ , el período de pitch futuro se calcula según la siguiente expresión:

$$\hat{y}(n+T) = \begin{cases} y(n+T) & n+T < n^+ \\ y(n) & n+T \geq n^+ \end{cases}$$

Puesto que esta extensión es solamente una aproximación, hay una cierta transigencia en la calidad cuando se compara con lo que podría obtenerse si se usara la verdadera señal de voz descodificada. Debe notarse que (6) no proporciona tampoco ninguna solución conveniente a este problema. Por el contrario, más bien especifica que el post-filtrado con datos de voz sintetizados futuros dentro de la unidad de información presente solo se hace siempre y cuando estén disponibles subunidades de información que sigan a la subunidad que se va a mejorar. En particular, este documento solamente prevé la disponibilidad de las unidades de información de voz hasta la unidad de información de voz presente, pero no unidades de información futuras.

Otro método de post-filtrado relacionado que sin embargo tiene menos relevancia en el contexto de la invención se especifica en (7). Esta patente describe un método de post-filtrado para un códec de voz de velocidad variable en el cual la intensidad del post-filtro se controla en respuesta a la tasa de bits promedio.

Los post-filtros tradicionales (por ejemplo, formante o de conformación, de pitch) no introducen ningún retraso con el fin de mantener el retraso del códec en el mínimo. Esto es así puesto que normalmente el presupuesto de retraso de codificación se emplea de manera más eficaz en el codificador para, por ejemplo, mirar hacia delante. Este hecho provoca los siguientes problemas que disminuyen la capacidad de mejora del post-filtro.

Debe notarse que la ampliación temporal es un problema especialmente en los casos en que el período de pitch de la señal de voz es no estacionario. Este es el caso, en particular en los inicios de conversaciones. Más generalmente, se puede establecer que el rendimiento de los post-filtros convencionales en los períodos transitorios de conversaciones no es óptimo puesto que sus parámetros tienen similares faltas de fiabilidad.

Un parte importante de la idea básica de la invención es por lo tanto aumentar y mejorar el rendimiento del post-filtro por medio de la utilización de información de las unidades de información futuras. Para hacer eso se utilizan retrasos de tiempo inherentes en las operaciones de recepción y de descodificación. La presente invención se basa en una situación en la que una señal descodificada de una unidad de información deviene disponible en conexión con o más tarde que los parámetros de una unidad de información posterior que deviene disponible. En otras palabras, el conjunto constituido por la unidad receptora de entrada y el descodificador se disponen o montan para proporcionar una señal descodificada $y(n)$ de una primera unidad de información, n , esencialmente de manera simultánea con un parámetro $x(n+1)$ de una unidad de información, $n+1$, sucesiva a la primera unidad de información, n . La unidad de información de voz descodificada $y(n)$ se alimenta al post-filtro produciendo una unidad de información de voz de salida mejorada $y_{salida}(n)$. Según la invención, la operación del post-filtro se mejora proporcionando al post-filtro acceso a los parámetros $x(n+1)$ de al menos una unidad de información posterior, $n+1$. Puesto que el retraso de la señal es inherente a las operaciones de recepción y de descodificación, no se provoca ningún retraso adicional.

Una realización comprende un descodificador que opera según un algoritmo que provoca un retraso de la salida de al menos la longitud de una unidad de información L . La unidad de información de voz codificada de índice $n+1$ está entonces disponible en el receptor cuando el descodificador da salida a la unidad de información de voz descodificada $y(n)$ y se puede usar con propósitos de post-filtrado. Tales retardos están disponibles en diferentes dispositivos descodificadores. La figura 2 ilustra un esquema de bloques de tal realización de un dispositivo descodificador según la presente invención. Una unidad receptora 2 comprende una entrada al receptor 40, dispuesta para recibir los parámetros 4 que representan las señales codificadas en base a unidades de información $x(n+1)$, de manera típica señales de audio o de voz codificadas. A la entrada del receptor 40 se conecta un descodificador 20, dispuesto para proporcionar unidades de información $y(n)$ de señales de audio descodificadas 5 basadas en dichos parámetros 4. El descodificador 20 se dispone de tal modo que presente una diferencia de tiempo entre el momento en el que los parámetros 4 de una primera unidad de información estén disponibles en la entrada del receptor 40 y el momento en

que esté disponible en la salida del descodificador 20 una señal de audio descodificada de la primera unidad de información, correspondiendo dicha diferencia de tiempo a al menos una unidad de información. En la presente realización, la operación de descodificación provoca un retraso 51 de la señal de una unidad de información. El conjunto 50 del descodificador 20 y de la entrada al receptor 40 presenta de este modo una señal descodificada $y(n)$ al mismo tiempo que los parámetros de una unidad de información sucesiva $x(n+1)$.

A la salida del descodificador 20 y de la entrada al receptor 40 se conecta un post-filtro 30. El post-filtro 30 se dispone para proporcionar una señal de salida 6 basada en las unidades de información 5 de las señales de audio descodificadas en respuesta a los parámetros $x(n+1)$ de una unidad de información subsiguiente. Por lo tanto, el conocimiento de las señales de información de la señal futuras se puede utilizar en el proceso de post-filtrado, sin embargo, sin añadir retraso adicional de descodificación. Una salida del receptor 60 se conecta al post-filtro 30 para producir la señal de salida 6.

Un elemento esencial de un sistema VoIP (voz sobre protocolo de internet) es el circuito compensador de fluctuaciones en el terminal receptor. Su objetivo es convertir la corriente asíncrona de unidades de información de voz recibidas contenidas en paquetes en una corriente síncrona que posteriormente es descodificada por un descodificador de voz. En consecuencia, el circuito compensador de fluctuaciones puede funcionar como un circuito compensador de parámetros, de acuerdo con las ideas presentadas previamente. En otras palabras, se puede aplicar de manera ventajosa una realización de la invención en una aplicación VoIP, en la que el circuito compensador de fluctuaciones en el terminal receptor proporciona acceso fácilmente a las unidades de información futuras, siempre y cuando el circuito compensador o tampón no esté lleno.

Por lo tanto, otra realización de la presente invención comprende un circuito de entrada receptor que a su vez comprende un circuito compensador o tampón de parámetros, que almacena las unidades de información de voz codificadas recibidas, al menos dos unidades de información. El descodificador descodifica la unidad de información compensada n que proporciona la unidad de información de voz descodificada $y(n)$. Al mismo tiempo, la unidad de información de voz codificada de índice $n+1$ está disponible en el circuito compensador de parámetros y se puede usar con objetivos de post-filtrado. La figura 3 ilustra un esquema de bloques de tal realización de un dispositivo descodificador según la presente invención. Una unidad receptora 2 comprende un circuito de entrada receptor 40, dispuesto para recibir los parámetros 4 que representan señales codificadas a base de unidades de información. El circuito de entrada receptor 40 comprende un compensador de fluctuaciones 41, con posiciones de almacenamiento 42A y 42B para parámetros de al menos dos unidades de información.

A la primera posición 42A del compensador o "búfer" de fluctuaciones 41 se conecta un descodificador 20 y, de este modo, se le proporcionan los parámetros 4A de una primera unidad de información $x(n)$. El descodificador 20 se dispone para proporcionar unidades de información $y(n)$ de señales de audio descodificadas 5 basadas en los parámetros 4A. El circuito de entrada receptor 40 presenta, debido al compensador de fluctuaciones 41, una diferencia de tiempo entre el momento en que los parámetros 4B de una cierta unidad de información está disponible en el circuito receptor de entrada 40 y el momento en el que la señal de audio descodificada 5 de la misma unidad de información está disponible en la salida del descodificador 20, de forma que dicha diferencia de tiempo corresponde a al menos una unidad de información. En la presente realización, la operación de las fluctuaciones provoca el retraso de la señal en al menos una unidad de información. El conjunto 50 formado por el descodificador 20 y el circuito de entrada del receptor 40 presenta de este modo una señal descodificada $y(n)$ al mismo tiempo que los parámetros de una unidad de información sucesiva $x(n+1)$. Luego se dispone el post-filtro 30 de la misma manera que en la figura 2.

La figura 4 ilustra un diagrama de flujo de las etapas de una realización de un método según la presente invención. El método de descodificación comienza en la etapa 200. En la etapa 210 se reciben parámetros de señales codificadas a base de unidades de información. En el paso 212 los parámetros se descodifican en unidades de información de señales de audio descodificadas. Al menos una de las etapas 210 y 212 produce una diferencia de tiempo entre el momento en que los parámetros de una primera unidad de información están disponibles tras la recepción y el momento en el que está disponible una señal de audio descodificada tras su descodificación. La diferencia de tiempo corresponde a al menos una unidad de información. Las unidades de información de las señales de audio descodificadas se postfiltran para dar una señal de salida en la etapa 214 en respuesta a los parámetros de una unidad de información posterior respectiva. En el paso 216, se produce la señal de salida. El procedimiento termina en la etapa 299.

Ejemplos típicos de códecs que tienen retrasos intrínsecos son los códecs escalables o internos ("embedded codecs"). En consecuencia, se presenta a continuación una breve exposición resumen acerca de los códecs escalables. La figura 5 ilustra un esquema de bloques de un sistema de codificación y descodificación (códec) de voz y audio escalable general. La unidad emisora 1 comprende aquí un codificador 10, en este caso un codificador escalable 110 que codifica la señal de audio o de conversación entrante 3 en una corriente de parámetros 4. La codificación completa tiene lugar en dos capas, una capa inferior 7, que en el emisor comprende un codificador primario 11 y al menos una capa superior 8, que en la unidad emisora comprende un codificador secundario 15. El dispositivo códec escalable se puede proporcionar con capas adicionales, pero en la presente descripción se usa como sistema modelo un sistema descodificador de dos capas. Sin embargo, los principios de la presente invención se pueden aplicar también a códecs escalables con más de dos capas. El codificador primario 11 recibe la señal de

audio o de conversación entrante 3 y la codifica en una corriente de parámetros primarios 12. El codificador primario descodifica también los parámetros primarios 12 en una señal primaria estimada 13, que corresponderá idealmente a una señal que se puede obtener a partir de los parámetros primarios 12 en el lado del descodificador. La señal primaria estimada 13 se compara con la señal de conversación o de audio entrante 3 original en un comparador 14, en este caso una unidad de sustracción. La señal diferencia es, en consecuencia, una señal de ruido de codificación primaria 16 del codificador primario 11. La señal de ruido de codificación primaria 16 se proporciona al codificador secundario, que la codifica en una corriente de parámetros secundarios 17. Estos parámetros secundarios 17 se pueden considerar como los parámetros de una mejora preferida de la señal descodificable a partir de los parámetros primarios 12. Juntos, los parámetros primarios 12 y los parámetros secundarios 17 forman la corriente general de parámetros 4 de la señal de audio o de conversación entrante 3.

Típicamente, los parámetros 4 se codifican y se transfieren a una unidad receptora 2. La unidad receptora 2 comprende un descodificador 20, en este caso un descodificador escalable 120, que recibe los parámetros 4 que representan la señal original de voz o de audio 3 y descodifica estos parámetros 4 en una señal de conversación o de audio descodificada 5. La descodificación completa tiene también lugar en las dos capas: la capa inferior 7 y la capa superior 8. En la unidad receptora, la capa inferior 7 comprende un descodificador primario 21. De manera análoga, la capa superior 8 comprende en la unidad receptora un descodificador secundario 25. El descodificador primario 21 recibe los parámetros primarios de entrada 22 de la corriente de parámetros 4. idealmente, estos parámetros son idénticos a los creados en el codificador 10; sin embargo, en algunos casos, el ruido de transmisión puede haber distorsionado los parámetros. El descodificador primario 21 descodifica los parámetros primarios de entrada 22 en una señal de audio o de conversación primaria descodificada 23. De manera análoga, el descodificador secundario 25 recibe los parámetros secundarios de entrada 27 de la corriente de parámetros 4. Idealmente, estos parámetros son idénticos a los creados en el codificador 10; sin embargo, también aquí el ruido de transmisión puede haber distorsionado los parámetros en algunos casos. El descodificador secundario 25 descodifica los parámetros secundarios de entrada 27 para dar una señal de conversación o de audio descodificada mejorada 26. Se desea que esta señal de conversación o de audio descodificada mejorada 26 corresponda tan precisamente como sea posible al ruido de codificación del codificador primario 11 y, por tanto, que sea también similar al ruido de codificación que resulta del descodificador primario 21. La señal de conversación o de audio primaria descodificada 23 y la señal de conversación o de audio descodificada mejorada 26 se añaden en un adicionador 24, produciendo la señal final de salida 5.

Si solamente se reciben los parámetros primarios 22 en la unidad receptora 2, la unidad receptora solo soporta descodificación primaria o por cualquier razón se decide que no se realiza descodificación secundaria, la señal de conversación o de audio descodificada mejorada resultante 26 será igual a cero y la señal de salida 5 devendrá idéntica a la señal de conversación o de audio primaria descodificada 23. Esto ilustra la flexibilidad del concepto de sistemas códec escalables. Cualquier post-filtrado es realizado típicamente según la técnica anterior sobre la señal de salida 5.

El algoritmo de compresión de voz escalable más usado hoy en día es el códec PCM de ley logarítmica A/U de 64 kbps según la Recomendación ITU-T G.711: "Pulse code modulation (PCM) of voice frequencies on a 64 kbps channel", ("Modulación por impulsos codificados (PCM, por sus siglas en inglés) de frecuencias de voz en un canal de 64 kbps"), de noviembre de 1988. El códec G.711 muestreado a 8 kHz convierte muestras PCM lineales de 12 o de 13 bit en muestras logarítmicas de 8 bit. La representación de bits ordenada de las muestras logarítmicas permite eliminar los bits menos significativos (LSB por sus siglas en inglés) en una corriente de bits G.711, haciendo que el códec G.711 sea prácticamente escalable en proporción señal a ruido (o SNR-escalable por las siglas en inglés) entre 48, 56 y 64 kbps. Esta propiedad de escalabilidad del códec G.711 se usa en las redes de comunicación de conmutación por circuitos para propósitos de señalización y control dentro de banda (en frecuencias vocales). Un ejemplo reciente del uso de esta propiedad de escalado G.711 es el protocolo 3GPP-TFO (TFO son las siglas en inglés de "operación sin tándem", según el 3GPP TS28.062 (3rd Generation Partnership Project, o sea Proyecto de Colaboración de 3ª Generación) en su TS (Especificación Técnica por las siglas en inglés) 28.062) que permite el establecimiento y el transporte de conversación de banda ancha sobre enlaces PCM de 64 kbps. Ocho kbps del flujo G711 original de 64 kbps se usan inicialmente para permitir el establecimiento de llamada del servicio de conversación de banda ancha sin afectar de manera considerable la calidad del servicio de banda estrecha. Después del establecimiento de llamada, la conversación de banda ancha utilizará 16 kbps del flujo G.711 de 64 kbps. Otros estándares de codificación de conversación más antiguos que soportan escalabilidad de lazo abierto son los de la Recomendación G.727 de la UIT-T: "Modulación por impulsos codificados diferencial adaptativa y jerarquizada con 5, 4, 3 y 2 bits por muestra (ADPCM por sus siglas en inglés)", diciembre 1990 y en cierto grado la G722 (ADPCM sub-banda).

Un avance más reciente en tecnología de codificación de conversación escalable es el estándar MPEG-4 (MPEG por las siglas en inglés de "Grupo de expertos de imágenes en movimiento"), (ISO/IEC-14496), que proporciona ampliaciones de escalabilidad para MPEG4-CELP. La capa base MPE se puede mejorar mediante transmisión de información de parámetros de filtro adicional o de información de parámetros de información adicional. La unión Internacional de Telecomunicaciones, Sector de Estandarización (UIT-T) ha terminado recientemente la estandarización de un nuevo códec escalable según la recomendación UIT-T G.729.1: "Codificador de velocidad de bits variable interno basado en G.729: un codificador de banda ancha escalable de 8-32 kbit/s de flujo de bits interoperable con G.729", de mayo de 2006, denominada G.729.EV. El intervalo de tasa de bits de este codificador

de voz escalable es de 8 kbps a 32 kbps. El códec proporciona escalabilidad de 8 a 32 kbps. El uso principal de este codificador de voz escalable es el de permitir compartir de manera eficiente un recurso limitado de banda ancha en pasarelas en oficinas o en el hogar, por ejemplo, un enlace de tipo uplink xDSL de 64/128 kbps compartido entre varias llamadas VoIP (de voz sobre protocolo de internet) (DSL por las siglas en inglés de "línea de abonado digital"; xDSL es el término genérico para diversos métodos específicos de DSL).

Una tendencia reciente en la codificación de voz escalable es proporcionar capas superiores que soportan las señales de audio no conversacionales, como música. Tal enfoque se ilustra en la figura 6. En tales códecs, la capa inferior 7 emplea un sistema de codificación de voz meramente convencional, por ejemplo, mediante el paradigma de "análisis por síntesis" (AbS) del cual es un ejemplo destacado CELP ("predicción lineal mediante excitación de códigos", por sus siglas en inglés). En la presente realización, el codificador primario 11 es, de este modo, un codificador CELP 18 y el descodificador primario 21 es un descodificador CELP 28. Como tal codificación es muy adecuada para voz solamente pero no tanto para señales de audio no conversacionales (no de voz), como música, la capa superior 8 trabaja, por el contrario, según un paradigma de codificación usado en códecs de audio. Por lo tanto, en la presente realización, el codificador secundario es un codificador de audio 19 y el descodificador secundario es un descodificador de audio 29. En la presente realización, de manera típica, la capa superior 8 de codificación funciona sobre el error de codificación de la codificación de la capa inferior.

Una realización particular de la invención, ilustrada en la figura 7, es en una aplicación en un descodificador de voz y audio escalable 120 en el cual una capa inferior lleva a cabo una descodificación primaria en un descodificador primario 21 para dar una señal descodificada primaria y_p , mientras que una capa superior lleva a cabo una descodificación secundaria para producir una señal mejorada secundaria y_s en un descodificador secundario 25. La señal mejorada secundaria y_s mejora la señal descodificada primaria y_p para dar una señal descodificada mejorada y_e . En la presente realización se supone que el descodificador 20 funciona sobre unidades de información de voz de una longitud de por ejemplo 20 ms y que el descodificador primario 21 tiene un retraso menor que el descodificador secundario 25 de al menos una unidad de información. En otras palabras, hay un retraso intrínseco 51 en el descodificador secundario 25.

En algunos sistemas de códec especiales, el códec secundario puede funcionar con una longitud de unidad de información diferente que el códec primario. Por ejemplo, el códec secundario puede tener una longitud de la unidad de información mitad que el códec primario y, por lo tanto, descodifica dos unidades de información secundarias mientras que el descodificador primario descodifica una unidad de información. Según como sea el diseño, el retraso intrínseco del descodificador secundario es o bien la longitud de una unidad de información del descodificador primario o bien la longitud de una unidad de información del descodificador secundario.

De manera específica y tal como se visualiza en la figura 7, se supone que el descodificador primario 21 puede descodificar la unidad de información de voz de orden $n+1$, $x(n+1)$ a la unidad de información de salida $y_p(n+1)$ de la señal descodificada primaria 23 sin ningún retraso concreto, es decir, sobre la base de los datos $x(n+1)$ de la unidad de información de voz codificada recibida correspondiente con el índice de unidad de información $n+1$. En contraste, el descodificador secundario 25 necesita incluso los datos de la siguiente unidad de información codificada. Por tanto, con la unidad de información $x(n+1)$ disponible, con índice $n+1$, el descodificador secundario 25 produce la unidad de información descodificada $y_s(n)$ de la señal mejorada secundaria descodificada 26. Con el fin de combinar adecuadamente la señal mejorada secundaria descodificada 26 con la señal descodificada primaria 23, la última tiene que ser retrasada una unidad de información. Esto se lleva a cabo mediante un filtro retardador 53 y da como resultado una señal primaria descodificada retrasada 54.

Este hecho hace posible aplicar la invención sin ninguna penalización de aumentar el retraso incluso además en el descodificador, lo que no sería conveniente. Si el flujo de bits recibido contiene información de capa mejorada, se puede generar la unidad de información $y_s(n)$ de la señal mejorada secundaria descodificada 26. Esta señal 26 se combina con la unidad de información $y_p(n)$ de la señal descodificada primaria retrasada, para formar juntas una unidad de información $y_e(n)$ de la señal descodificada mejorada. Esta unidad de información $y_e(n)$ resulta disponible cuando la unidad de información de parámetros $x(n+1)$ está disponible a partir del conjunto 50B. Posteriormente, se puede alimentar la unidad de información $y_e(n)$ a través de un post-filtro secundario no causal 30B, que puede sacar partido de la invención, como se describe previamente con detalle. De acuerdo con estas ideas, el funcionamiento del post-filtro se puede mejorar utilizando los parámetros codificados de la unidad de información $n+1$. Además, este post-filtro 30B puede sacar partido adicional del uso de la siguiente unidad de información $y_p(n+1)$ de la señal descodificada primaria 23, lo que constituye una aproximación de la unidad de información futura todavía no disponible $y_e(n+1)$. De este modo, en la presente realización, el post-filtro 30B puede mejorar la señal no solo sobre la base de los parámetros de una unidad de información futura sino también a partir de una aproximación bastante buena de la señal real de la unidad de información futura. Por lo tanto, el post-filtro secundario 30B proporciona una señal mejorada postfiltrada 56 como señal de salida 6 del dispositivo de descodificación.

La figura 8 ilustra un diagrama de bloques de otra realización de un dispositivo descodificador escalable según la presente invención. En esta realización, se proporciona un post-filtro primario 30A, conectado a a la salida del filtro retardador 53, esto es, funciona sobre la señal primaria descodificada retrasada 54. El conjunto 50A comprende en esta realización el circuito receptor de entrada 40, el descodificador primario 21 y el filtro retardador 53. El post-filtro primario 30A funciona, de acuerdo con la presente invención, accediendo a los parámetros de una unidad de

información posterior. En esta realización, la señal primaria descodificada 23 de la unidad de información siguiente también está disponible y se puede usar también, de manera ventajosa, en el post-filtro primario 30A. En otras palabras, la unidad de información de voz $y_p(n)$ de la señal primaria descodificada retrasada 54 se puede mejorar mediante un post-filtro primario no causal 30A, que se aprovecha de su acceso a la unidad de información $y_p(n+1)$ de la señal primaria descodificada 23 y a los parámetros 4 de la unidad de información $n+1$.

La unidad de información de salida 55 del post-filtro 30A, es decir $y_p^*(n)$, se usa para combinarla con la señal mejorada secundaria 26 para producir la señal de salida final. Sin embargo, en algunas situaciones, las mejoras proporcionadas por la señal mejorada secundaria 26 pueden, en algunos casos, ser similares a las que se pueden obtener mediante el post-filtro primario 30A y el resultado puede ser una sobrecompensación del ruido de codificación. En tales casos, el post-filtro 30A puede disponerse, de manera ventajosa, para determinar si los parámetros para la descodificación secundaria están disponibles en el circuito receptor de entrada 40. Si los parámetros secundarios están disponibles, se puede anular el funcionamiento del post-filtro, dando de este modo la señal primaria descodificada original como salida de los post-filtros primarios 30A, o al menos cambiar los principios de post-filtrado con el fin de no interferir con el funcionamiento de la señal de mejora secundaria.

La figura 9 ilustra un esquema de bloques de otra realización más de un dispositivo descodificador escalable según la presente invención. En esta realización, el descodificador secundario 25 está seguido, de nuevo, por un post-filtro secundario 30B, como en la figura 7; sin embargo, también se proporciona el post-filtro primario 30A. En tal realización, también se puede mejorar adicionalmente una señal de salida que se proporciona mejorada a partir del descodificador secundario 25, utilizando un post-filtro secundario 30B. También en este caso, el post-filtro secundario 30B puede basar su funcionamiento sobre parámetros de una unidad de información sucesiva. Mientras que este post-filtro 30B no tiene acceso a una unidad de información futura $y_e(n+1)$ de la salida del descodificador mejorada 5, su funcionamiento puede, no obstante, basarse en una unidad de información futura $y_p(n+1)$ de la señal descodificada primaria. Un conjunto primario 50A comprende el circuito de entrada del receptor 40, el descodificador primario 21 y el filtro de retardo 53, mientras que un conjunto secundario 50B comprende el circuito de entrada del receptor 40, el descodificador escalable completo 120 y los post-filtros primarios 30A.

La figura 10 ilustra un esquema de otra realización más de un dispositivo descodificador escalable según la presente invención. En este caso, la señal primaria descodificada retardada sin post-filtrar 54 se proporciona al adicionador 24 para ser combinada con la señal de mejora secundaria 26. Esto evita mezclar las condiciones de ruido de codificación del post-filtro primario 30A y de la mejora procedente del descodificador secundario 25. En lugar de ello, la salida 60 se dispone como un selector 61, montado para producir bien la señal primaria descodificada postfiltrada 55 o la señal mejorada postfiltrada 56 como señal de salida procedente del dispositivo descodificador. Preferentemente, el selector 61 se hace funcionar en respuesta a las señales entrantes, como se indica en el esquema mediante la flecha de rayas 62. Más adelante se discuten más de estas posibilidades con más detalle.

Un aspecto parcial adicional de la presente invención es, como se ha discutido en este texto previamente, aplicar la mejora no causal de los post-filtros dependiendo de las características de la señal de audio o voz. En particular, tal aplicación es beneficiosa durante los transitorios de sonidos. Por ejemplo, un transitorio entre sonidos es la transición de un fonema (elemento fonético) a otro, los cuales son relativamente constantes o estacionarios. Lo típico de tales transitorios es que la señal es no estacionaria y que la estimación de los parámetros que se realiza por el codificador de voz es menos fiable que durante los sonidos estacionarios. Si el post-filtro se basa en dichos parámetros menos fiables, es probable que su rendimiento sea malo. Según la presente invención, el rendimiento del post-filtro durante tales transitorios se puede mejorar utilizando parámetros y preferentemente también la voz sintetizada de una unidad de información futura. Esta mejora se consigue puesto que el sonido durante la unidad de información futura puede haberse hecho más estable lo que permite una estimación de parámetros más fiable.

Esta realización se basa en la detección de transitorios en los cuales es posible el funcionamiento post-filtro no causal específico. Tal detección se puede hacer con un clasificador de sonidos, que en un caso simple puede ser de detector de actividad de voces (VAD por sus siglas en inglés) o, de manera más general, un detector de sonidos, el cual, aparte de la distinción básica conversación (voz) / no conversación, puede por ejemplo distinguir entre diferentes clases de conversación como sonidos de voces, sonidos sordos, comienzo de conversación. Tal detección se puede basar también en la evaluación de la evolución temporal de ciertos parámetros de la señal como la energía o los parámetros LPC (de codificación predictiva lineal) e identificar tales partes de la señal de audio o de voz como transitorios cuando estos parámetros cambian rápidamente. El detector de transitorios se puede montar en el codificador o en el descodificador; en el primer caso, ello requiere transmitir la información de la detección al receptor. Los cambios en las características de audio se pueden cuantificar y medir hasta un cierto grado de significación y se pueden usar para controlar el funcionamiento de un post-filtro. En particular, los post-filtros según la presente invención se pueden disponer para adaptar el grado en el cual el parámetro de pitch usado en el post-filtro de pitch se basa en el parámetro de pitch de una unidad de información posterior. La adaptación se lleva a cabo dependiendo de una medida de la importancia del cambio de las características de audio entre una unidad de información presente y una unidad de información previa o posterior.

Una realización preferida particular para la cual se puede mejorar el rendimiento del post-filtro es una aplicación a inicios de conversaciones con voces después de períodos de inactividad de conversación. En este caso, de manera específica, el post-filtro es un post-filtro de pitch y parámetros usados en él de una unidad de información futura son

los parámetros de pitch de la subunidad de información que pertenece a la unidad de información siguiente a la unidad de información presente.

Según una realización preferida adicional de la invención, relacionada con mejoras en el post-filtro de pitch, el parámetro de pitch se maneja de una forma novedosa y más precisa. Como se ha analizado previamente, los post-filtros de pitch punteros evalúan una expresión basada en las ecuaciones (1) y (2), donde un segmento pasado y un segmento futuro de conversación sintetizada se combinan con un segmento de conversación presente, donde un segmento puede ser una unidad como una subunidad de información o un ciclo de pitch. Estos retardos de los segmentos pasado y futuro conducen respectivamente al segmento presente con el valor de parámetro de pitch T . El uso de T como parámetro de desfase para el segmento de conversación pasado es conceptualmente correcto puesto que está en línea con el paradigma de búsqueda del libro de códigos adaptativo de los códecs de conversación típicos de análisis por síntesis que calculan T como el valor de desfase que maximiza la correlación del segmento desfasado con el segmento de conversación presente.

Sin embargo, utilizar T como el parámetro director para el segmento futuro no es generalmente preciso puesto que implica suponer que el parámetro de desfase de pitch permanece constante incluso para el segmento futuro. Esto es especialmente problemático en transitorios en los cuales el pitch puede cambiar fuertemente. La referencia (6) proporciona una solución a este problema especificando un retardo adicional y determinante conductor y de desfase basado en cálculos de correlación entre los segmentos. Sin embargo, esto presenta desventajas por razones de complejidad.

La solución al problema es como sigue, según la presente invención, tomando como referencia la figura 11. Se supone que el post-filtro de pitch tiene acceso a un vector de parámetros de pitch de una subunidad de información, para la unidad de información presente n y para al menos una unidad de información futura $n+1$. De manera típica, cada unidad de información comprende cuatro subunidades de información. Los parámetros de pitch de las cuatro subunidades de información de la presente unidad de información se denotarán mediante $T(0), \dots, T(3)$ y los de las cuatro subunidades de información de la unidad de información futura se representarán mediante $T(4), \dots, T(7)$. Dado esto, el parámetro conductor para un segmento dado se encuentra buscando que el parámetro de pitch de la subunidad de información con respecto a su posición en la subunidad de información se retrase en el tiempo en el segmento presente. Según el ejemplo de la figura 11, para el segmento presente dado 100 este es el caso para el valor de pitch de subunidad de información $T(4)$. Como se puede ver en esa figura, utilizar el valor de parámetro de pitch del segmento presente $T(1)$ como parámetro director es impreciso puesto que el pitch está cambiando a valores más pequeños. Un algoritmo de ejemplo preferido según el cual se puede encontrar el parámetro director para el segmento dado es como sigue, haciendo referencia a la figura 12. El procedimiento, que puede ser una parte de la etapa 214 en la figura 4, empieza en la etapa 220. Se selecciona en la etapa 222 una primera subunidad de información que sigue el segmento presente. Partiendo de esta primera subunidad de información que sigue el segmento presente, se comprueba en la etapa 224 si el índice de tiempo de la unidad de información menos el valor de pitch de la subunidad de información correspondiente es mayor o igual que el índice de tiempo del segmento presente. Si es este el caso, se toma el valor de pitch de la subunidad de información como parámetro conductor de pitch para el segmento presente en la etapa 226 y se detiene el algoritmo en la etapa 229. En caso contrario, se repite la comprobación con la siguiente subunidad de información. En la etapa 228, se comprueba si hay más subunidades de información disponibles. Si no, el procedimiento termina en la etapa 239; en caso contrario, se selecciona una nueva subunidad de información en la etapa 230 y se repite la comprobación de la etapa 224. En este algoritmo el índice de tiempo de la subunidad de información puede ser por ejemplo el índice de tiempo del comienzo o de la mitad de la subunidad de información. Puede notarse que este algoritmo podría usarse también con alguna ventaja si se usa un determinante de conducción como se describe en la referencia (6), puesto que esto puede ayudar a ahorrar complejidad limitando el intervalo sobre el cual tendrían que realizarse los cálculos de correlaciones.

Las realizaciones que se acaban de describir deben entenderse como unos pocos ejemplos ilustrativos de la presente invención. Las personas expertas en la técnica entenderán que se pueden hacer diversas modificaciones, combinaciones y cambios a las realizaciones sin salirse del alcance de la presente invención. En particular, se pueden combinar diferentes soluciones parciales en las diferentes realizaciones en otras configuraciones, cuando es técnicamente posible. El alcance de la presente invención se define, no obstante, en las reivindicaciones anexas.

Referencias

[1] P.Kroon, B. Atal, "Quantization procedures for 4,8 kbps CELP coders", in Proc IEEE ICASSP, páginas 1650 - 1654, 1987.

[2] V. Ramamoorthy, N.S. Jayant, "Enhancement of ADPCM speech by adaptive postfiltering", AT&T Bell Labs Tech. J., páginas 1465 - 1475, 1984.

[3] V. Ramamoorthy, N.S. Jayant, R. Cox, M. Sondhi, "Enhancement of ADPCM speech coding with backward adaptive algorithms for postfiltering and noise feed-back", IEEE J on Selected Areas in Communications, vol. SAC-6, páginas 364 - 382, 1988.

[4] J. H. Chen, A. Gersho, "Adaptive postfiltering for quality enhancements of coded speech", IEEE Trans. Speech Audio Process., vol 3, número 1, 1995.

[5] B. Besette et al., "Method and device for frequency-selective pitch enhancement of synthesized speech", solicitud de patente de Estados Unidos US20050165603A1.

5 [6] L. Bialik et al., "A pitch post-filter", EP-0807307B1.

[7] Pasi Ojala et al., "A decoding method an system comprising an adaptive postfilter", EP 1 050 040 B1.

REIVINDICACIONES

1. Un dispositivo descodificador que comprende:

una entrada de receptor (40) dispuesta para recibir parámetros (4) de señales codificadas a base de unidades de información;

5 un descodificador (20) conectado a dicho circuito receptor de entrada (40) dispuesto para proporcionar unidades de información de señales de audio descodificadas (5; 54) sobre la base de dichos parámetros;

un post-filtro (30; 30A; 30B) conectado a una salida de dicho descodificador (20) y dispuesto para proporcionar una señal de salida (6) basada en dichas unidades de información de señales de audio descodificadas (5; 54) y

una salida (60) dispuesta para producir dicha señal de salida (6),

10 caracterizado porque:

al menos uno entre la entrada de receptor (40) y dicho descodificador (20) se dispone para establecer una diferencia de tiempo entre el momento en que están disponibles los parámetros de una primera unidad de información en dicha entrada de receptor (40) y el momento en el que está disponible una señal de audio descodificada de dicha primera unidad de información en dicha salida de dicho descodificador (20), correspondiendo la diferencia de tiempo a al menos una unidad de información;

15 estando conectado dicho post-filtro (30; 30A; 30B) a dicha entrada de receptor (40) y

estando dispuesto dicho post-filtro (30; 30A; 30B) para proporcionar un filtrado de dichas unidades de información de señales de audio descodificadas (5; 54) en la señal de salida (6) en respuesta a dichos parámetros (4) de una unidad de información posterior respectiva.

20 2. El dispositivo descodificador según la reivindicación 1, en el que dicho circuito receptor de entrada (40) comprende un almacenamiento (41) para parámetros de al menos dos unidades de información consecutivas, por lo cual se proporcionan a dicho descodificador (20) parámetros (4A) de una primera unidad de información y teniendo acceso dicho post-filtro (30; 30A; 30B) a parámetros (4B) de una segunda unidad de información posterior.

25 3. El dispositivo descodificador según la reivindicación 1, en el que dicho descodificador (20) comprende medios que retardan (51; 53) dichas unidades de información de señales de audio descodificadas antes de que se hagan salir hacia dicho post-filtro (30; 30A; 30B).

4. El dispositivo descodificador según una cualquiera de las reivindicaciones 1 a 3, en el que dicho post-filtro (30; 30A; 30B) comprende un post-filtro de pitch en el que el parámetro de pitch usado en dicho post-filtro de pitch se basa en el parámetro de pitch de dicha unidad de información posterior.

30 5. El dispositivo descodificador según la reivindicación 4, en el que dicho post-filtro de pitch de dicho post-filtro (30; 30A; 30B) se disponen para determinar, para una subunidad de información que sigue a la unidad de información presente, un valor de un índice de tiempo reducido en un valor de pitch para dicha subunidad de información y tomar, si dicho valor determinado es mayor o igual al índice de tiempo de la unidad de información presente, dicho valor de pitch para dicha subunidad de información como parámetro conductor de pitch para dicha unidad de información presente.

35 6. El dispositivo descodificador según las reivindicaciones 4 o 5, que comprende un detector de características de audio, una salida del cual se conecta a dicho post-filtro (30; 30A; 30B); estando dicho post-filtro (30; 30A; 30B) dispuesto para adaptar el grado en el cual dicho parámetro de pitch usado en dicho post-filtro de pitch se basa en dicho parámetro de pitch de dicha unidad de información posterior dependiendo de la medida de la importancia del cambio de las características de audio entre una unidad de información presente y al menos una entre unidades de información previa o posterior.

40 7. El dispositivo descodificador según la reivindicación 6, en el que dicho detector de características de audio es al menos uno entre un detector de actividad de voz o un detector de sonorización y en el que dicho post-filtro se dispone para basar dicho parámetro de pitch usado en dicho post-filtro de pitch en dicho parámetro de pitch de dicha unidad de información posterior en el caso de un comienzo de conversaciones con voces detectado.

45 8. El dispositivo descodificador según una cualquiera de las reivindicaciones 1 a 7, en el que dicho post-filtro (30; 30A; 30B) se dispone para tener también acceso a una señal descodificada de dicha unidad de información posterior.

50 9. El dispositivo descodificador según una cualquiera de las reivindicaciones 1 a 8, en el que dicho descodificador (20) es un descodificador escalable (120) o una parte de un descodificador escalable, en el que un descodificador secundario (25) de dicho descodificador escalable tiene un retardo mayor que un descodificador primario (21) de dicho descodificador escalable.

10. Un dispositivo descodificador que comprende un descodificador escalable (120) y al menos dos dispositivos descodificadores según la reivindicación 7.
11. Un método de descodificación que comprende las etapas de:
- recibir (210) parámetros de señales codificadas basadas en unidades de información;
- 5 descodificar (212) dichos parámetros en unidades de información de señales de audio descodificadas, caracterizado porque al menos una de dichas etapas de recibir y de descodificar provoca una diferencia de tiempo entre el momento en que están disponibles los parámetros de una primera unidad de información tras la recepción y el momento en el que está disponible una señal de audio descodificada de una primera unidad de información tras la descodificación, correspondiendo dicha diferencia de tiempo a al menos una unidad de información;
- 10 post-filtrar (214) dichas unidades de información de señales de audio descodificadas para producir una señal de salida en respuesta a dichos parámetros de la unidad de información posterior respectiva y producir (216) dicha señal de salida.
12. El método de descodificación según la reivindicación 11, que comprende la etapa de:
- 15 almacenar parámetros de al menos dos unidades de información consecutivas en cada instante, por lo que dicha etapa de descodificación se lleva a cabo con parámetros de una primera unidad de información y dicho post-filtrado se lleva a cabo con acceso a los parámetros de una segunda unidad de información posterior.
13. El método de descodificación según la reivindicación 11, que comprende la etapa de:
- 20 retardar dichas unidades de información de señales de audio descodificadas antes de realizar dicha etapa de post-filtrado.
14. El método de descodificación según cualquiera de las reivindicaciones 11 a 13, en el que dicha etapa de post-filtrado (214) comprende realizar un post-filtrado de pitch, de modo que el parámetro de pitch usado en dicho post-filtrado de pitch se basa en un parámetro de pitch de dicha unidad de información posterior.
15. El método de descodificación según la reivindicación 11, en el que dicho post-filtrado de pitch en dicha etapa de post-filtrado (214) comprende:
- 25 determinar (224), para una subunidad de información de una unidad de información presente, un valor de un índice de tiempo reducido en un valor de pitch para dicha subunidad de información y tomar (226), si dicho valor determinado es mayor o igual a un índice de tiempo de la unidad de información presente, dicho valor de pitch para dicha subunidad de información como parámetro conductor de pitch para dicha unidad de información presente.
- 30
16. El método de descodificación según las reivindicaciones 14 o 15, que comprende la etapa de detectar características de audio de dichas señales codificadas basadas en unidades de información;
- 35 de modo que dicha etapa de post-filtrado adapta el grado en el cual dicho parámetro de pitch usado en dicho post-filtro de pitch se basa en dicho parámetro de pitch de dicha unidad de información posterior dependiendo de la medida de la importancia del cambio de las características de audio entre una unidad de información presente y al menos una entre unidad de información previa y unidad de información posterior.
17. El método de descodificación según la reivindicación 16, en el que la etapa de detección comprende la detección de al menos uno entre un detector de actividad de voz o un detector de sonorización y en el que dicha etapa de post-filtrado se basa en dicho parámetro de pitch de dicha unidad de información posterior siguiente solo en el caso de un comienzo de conversaciones con voces detectado.
- 40
18. El método de descodificación según una cualquiera de las reivindicaciones 11 a 17, en el que dicha etapa de post-filtrado (214) se lleva a cabo también en respuesta a una señal descodificada de dicha unidad de información posterior respectiva
19. El método de descodificación según una cualquiera de las reivindicaciones 11 a 18, en el que dicha etapa de descodificación (212) supone descodificar en un descodificador escalable, en el cual un descodificador secundario de dicho descodificador escalable implica un retardo mayor que un descodificador primario de dicho descodificador escalable.
- 45
20. Un método de descodificación que comprende al menos dos métodos de descodificación según la reivindicación 19.

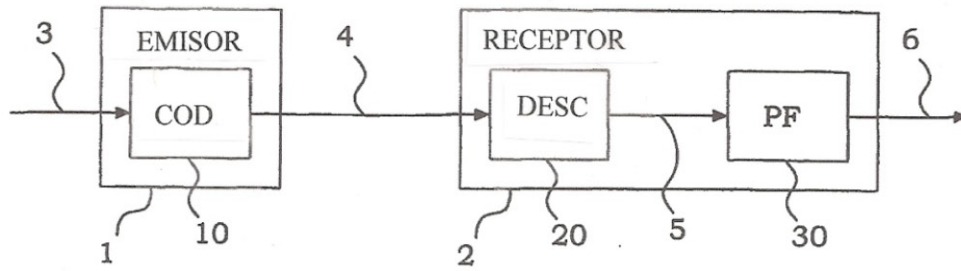


Fig. 1

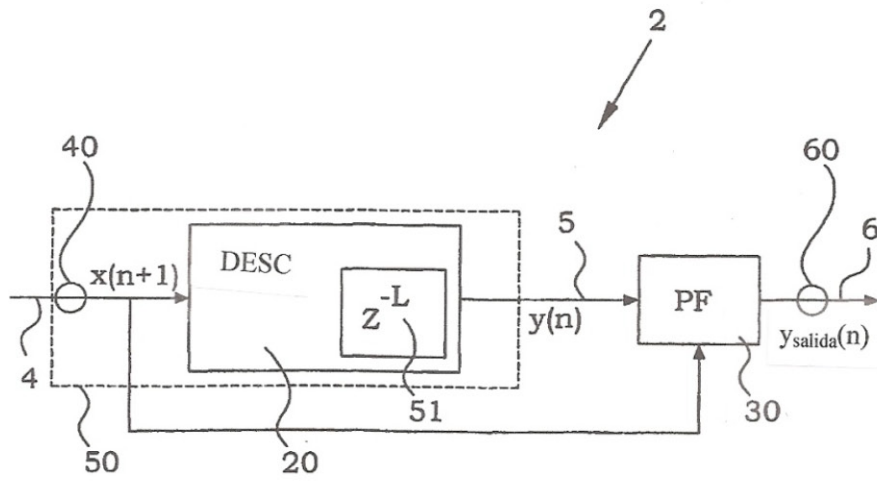


Fig. 2

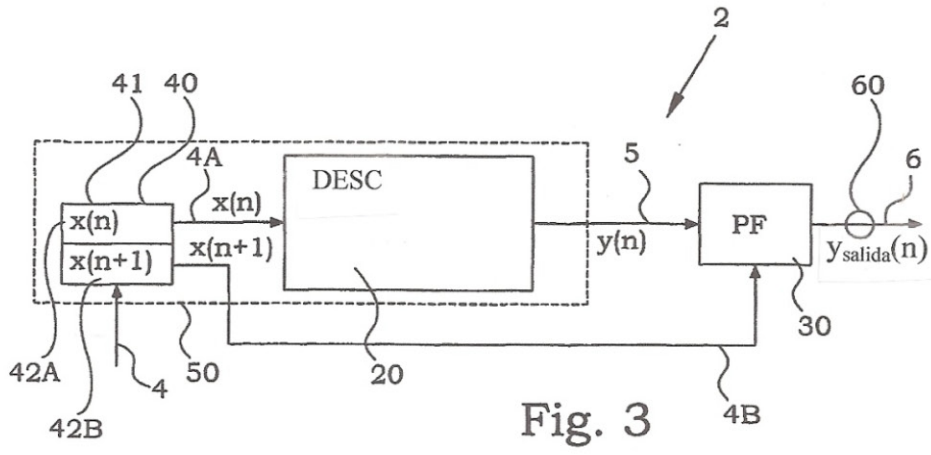


Fig. 3

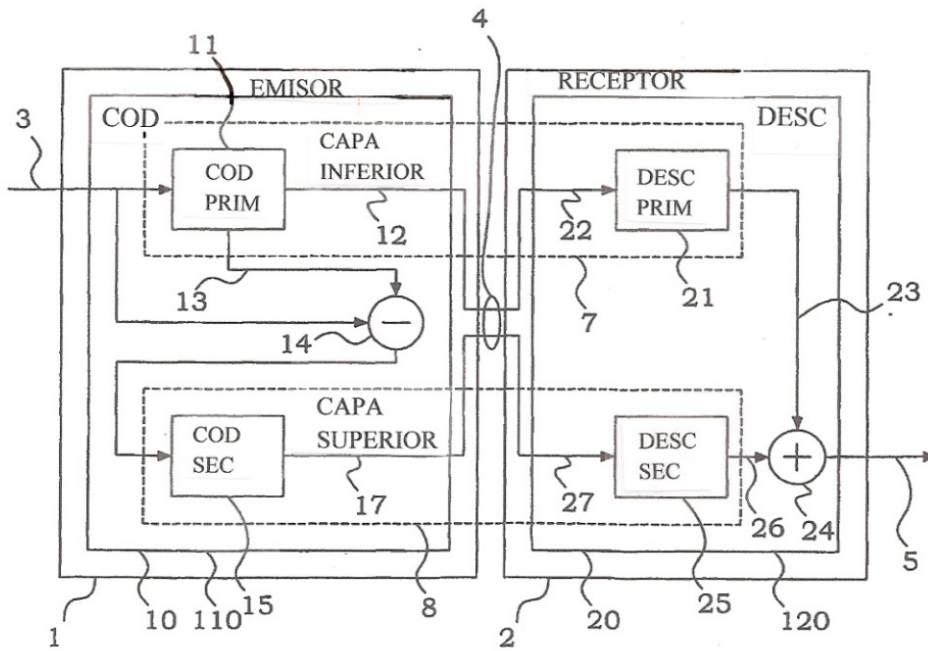


Fig. 5

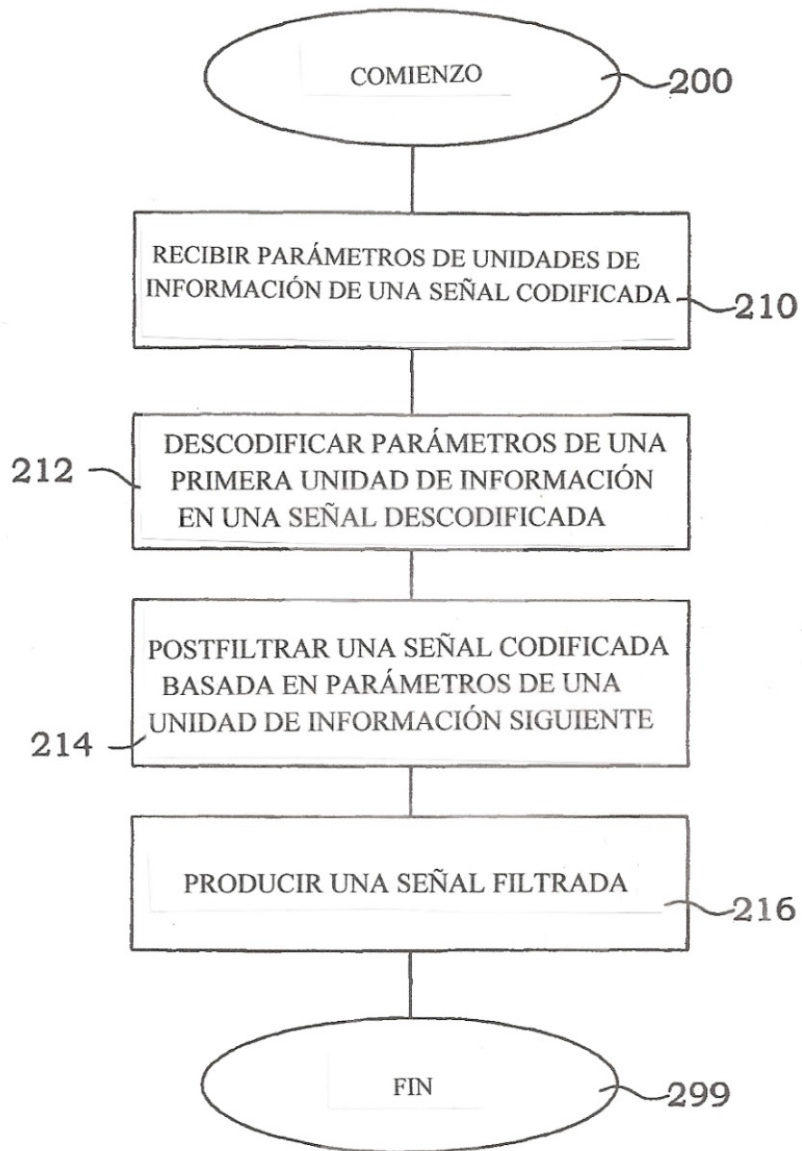


Fig. 4

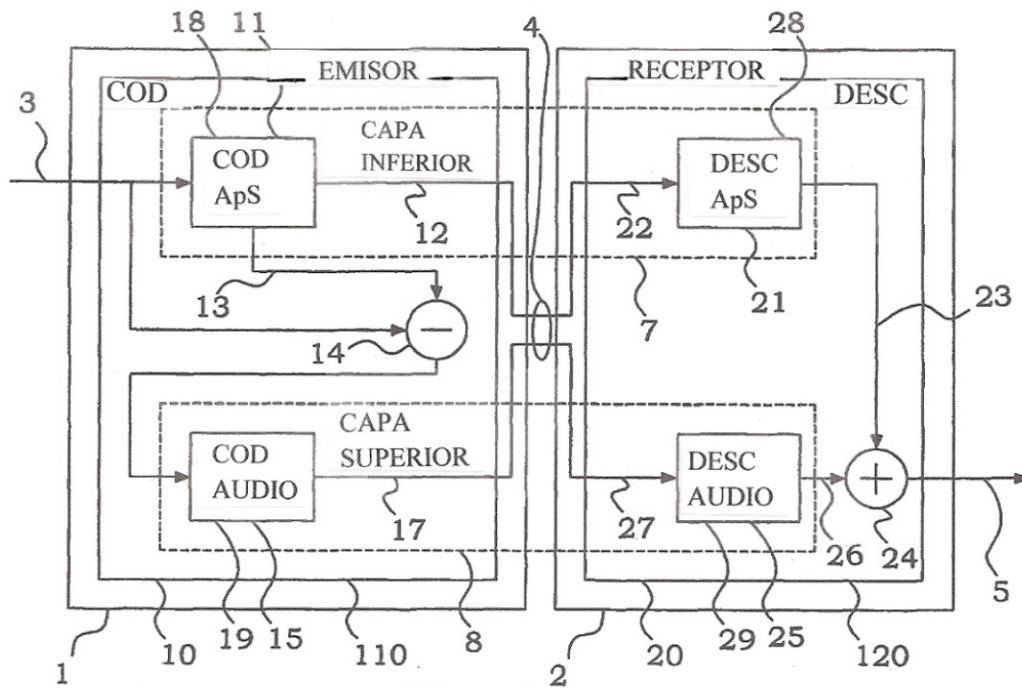


Fig. 6

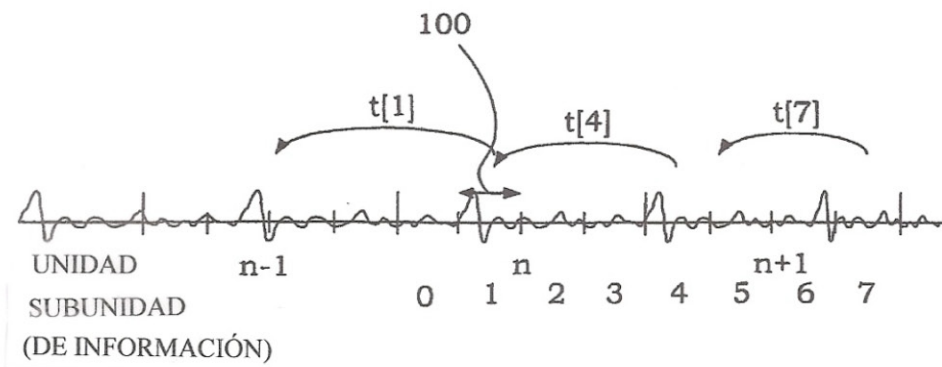


Fig. 11

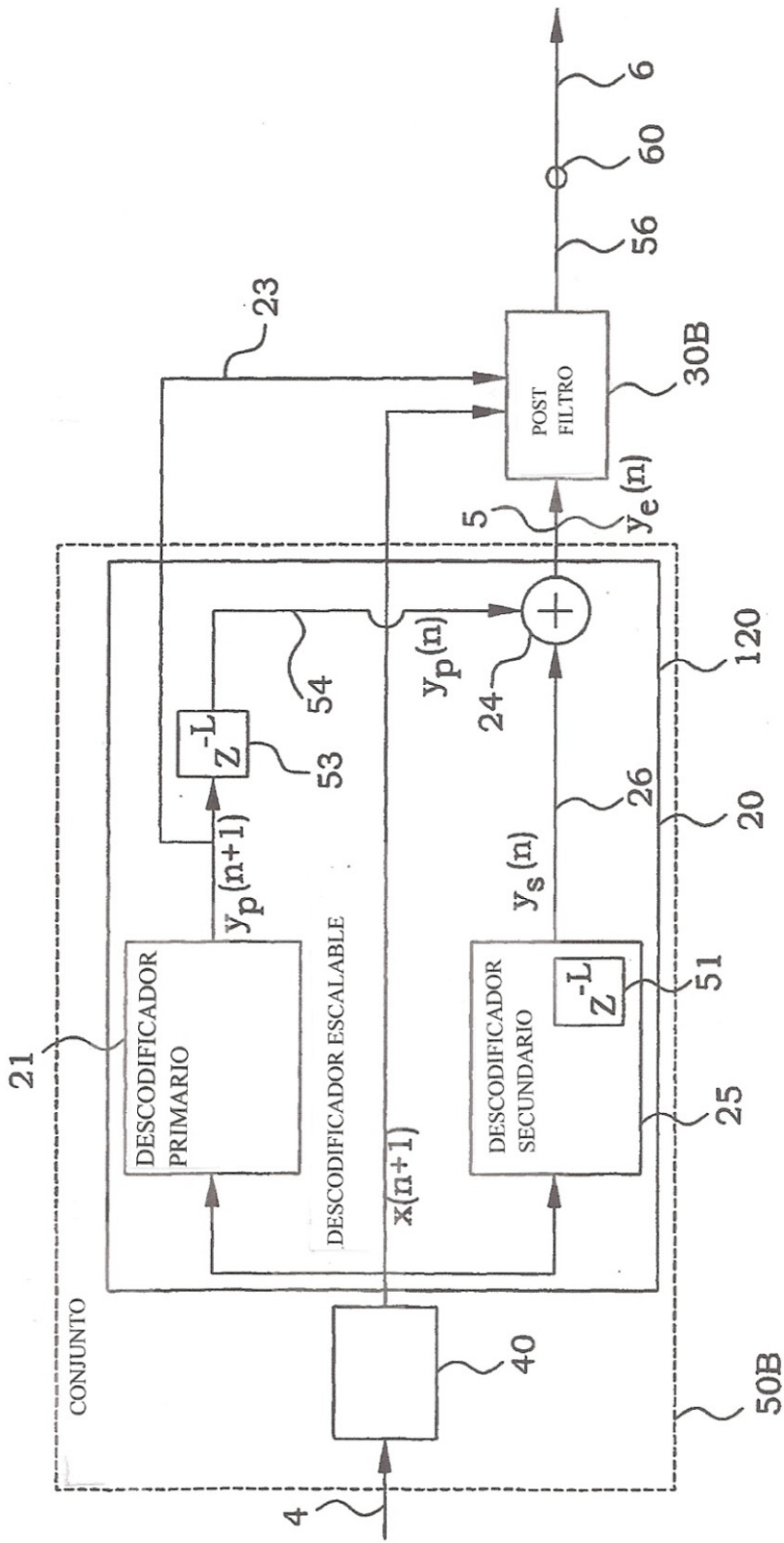


Fig. 7

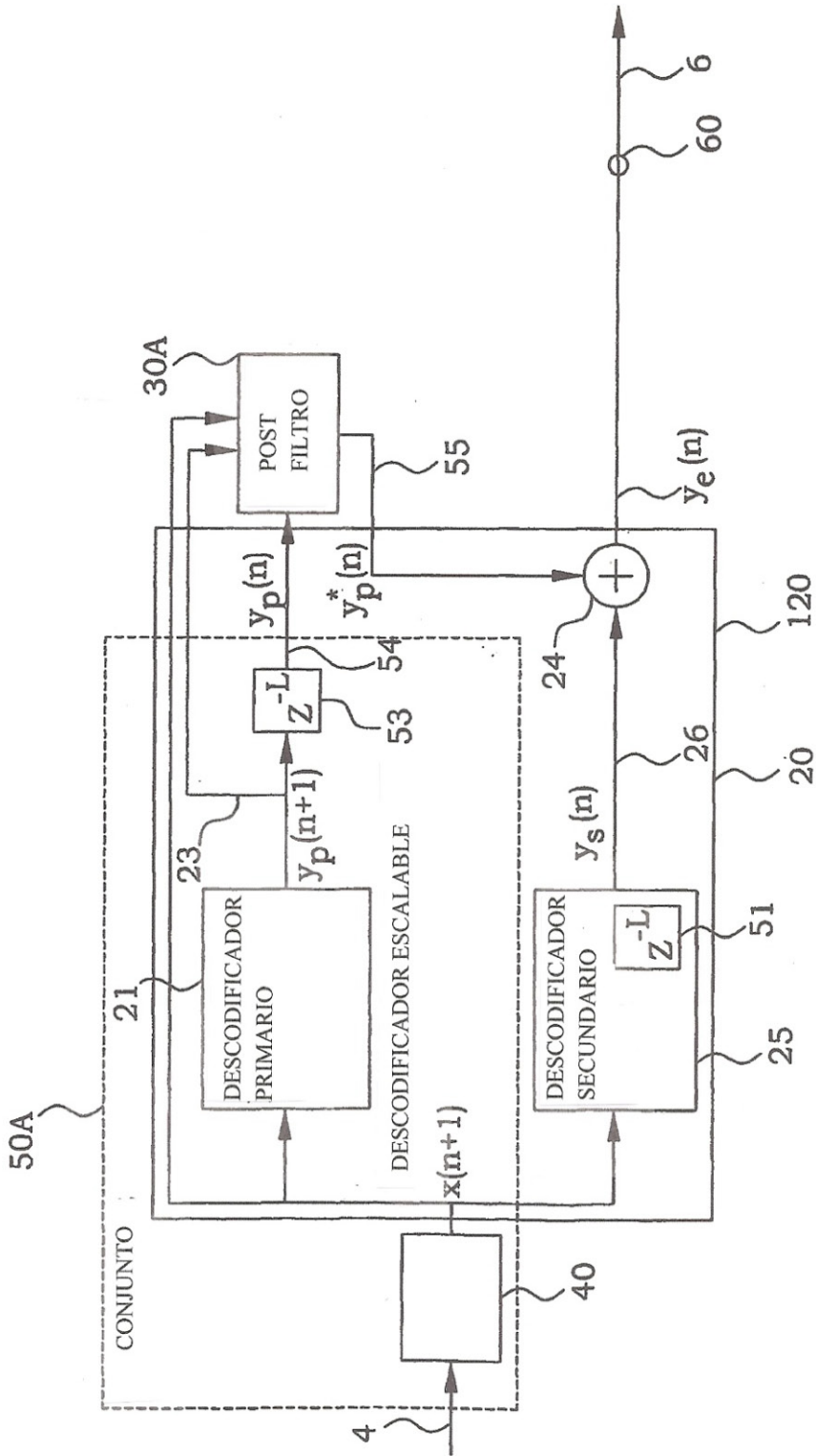


Fig. 8

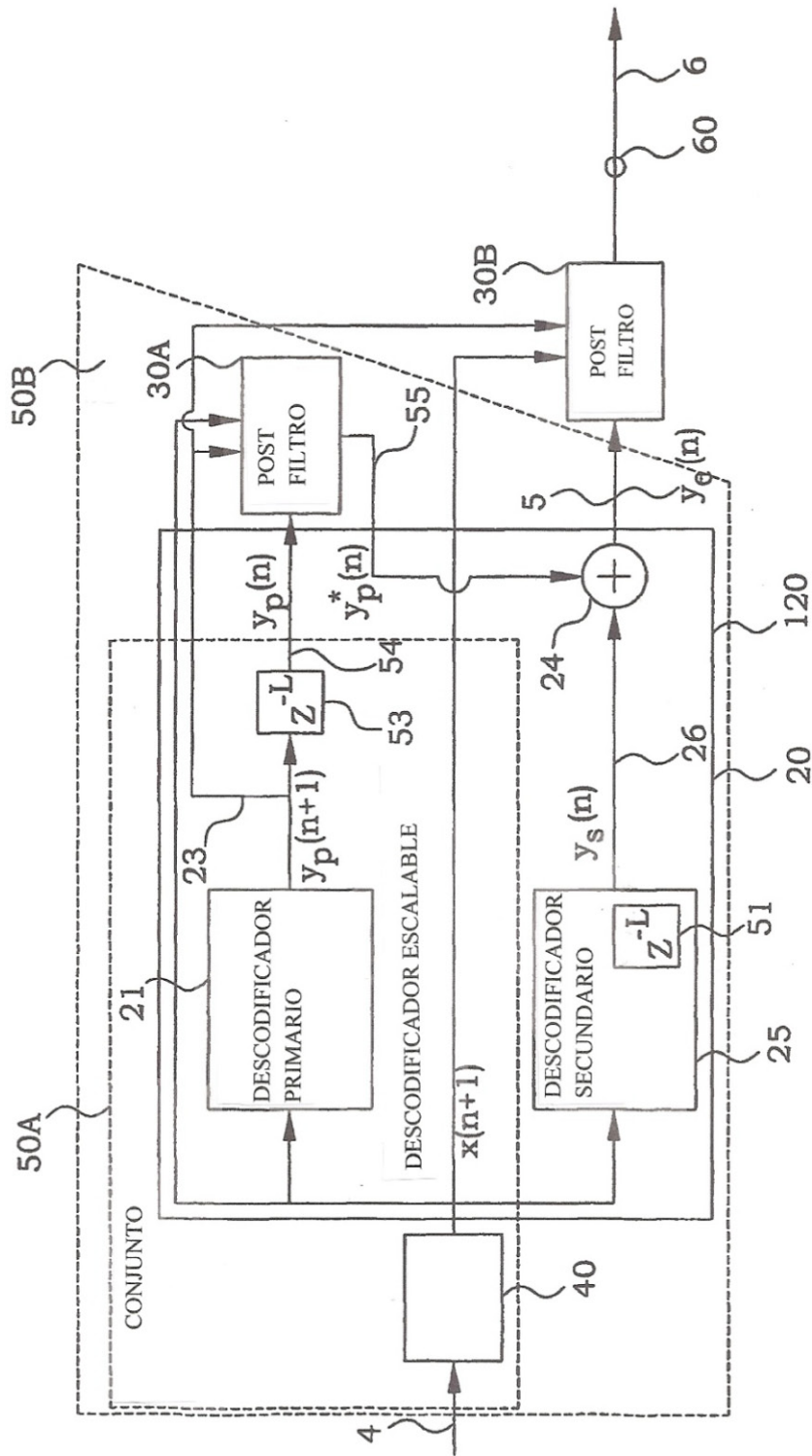


Fig. 9

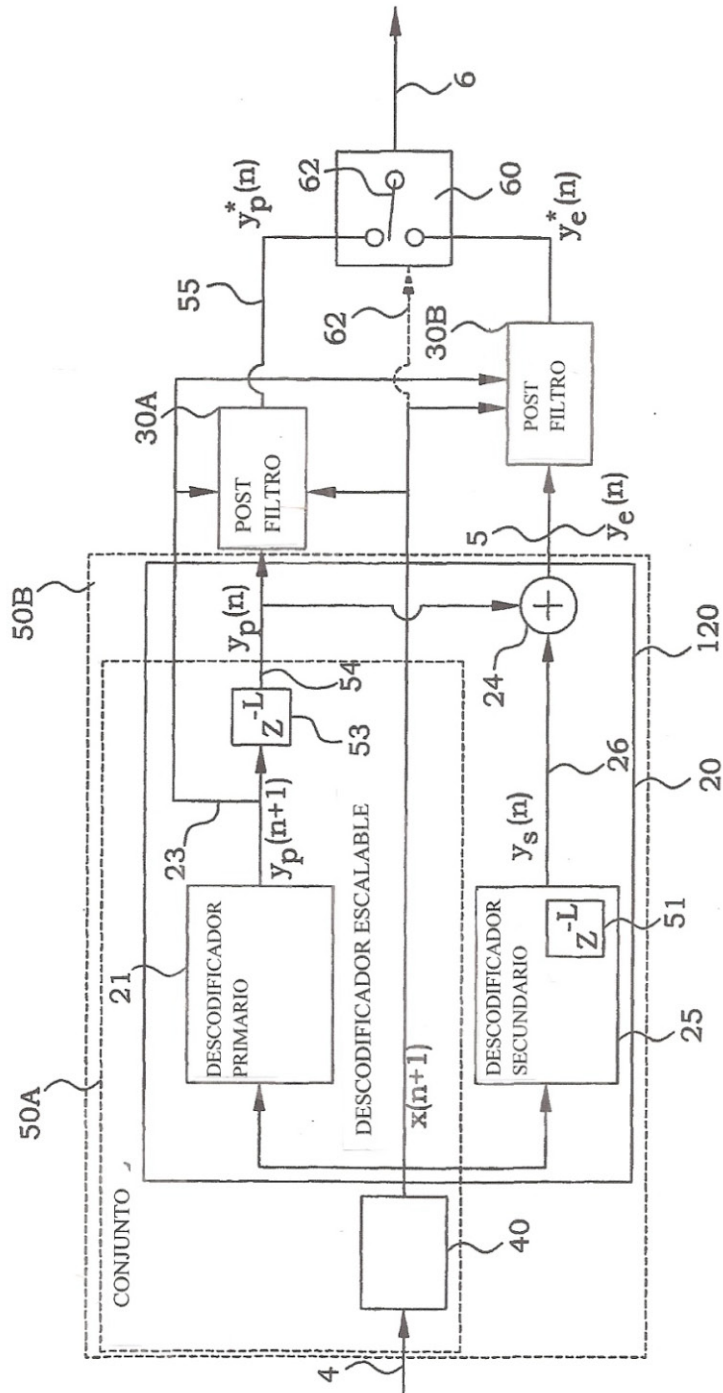


Fig. 10

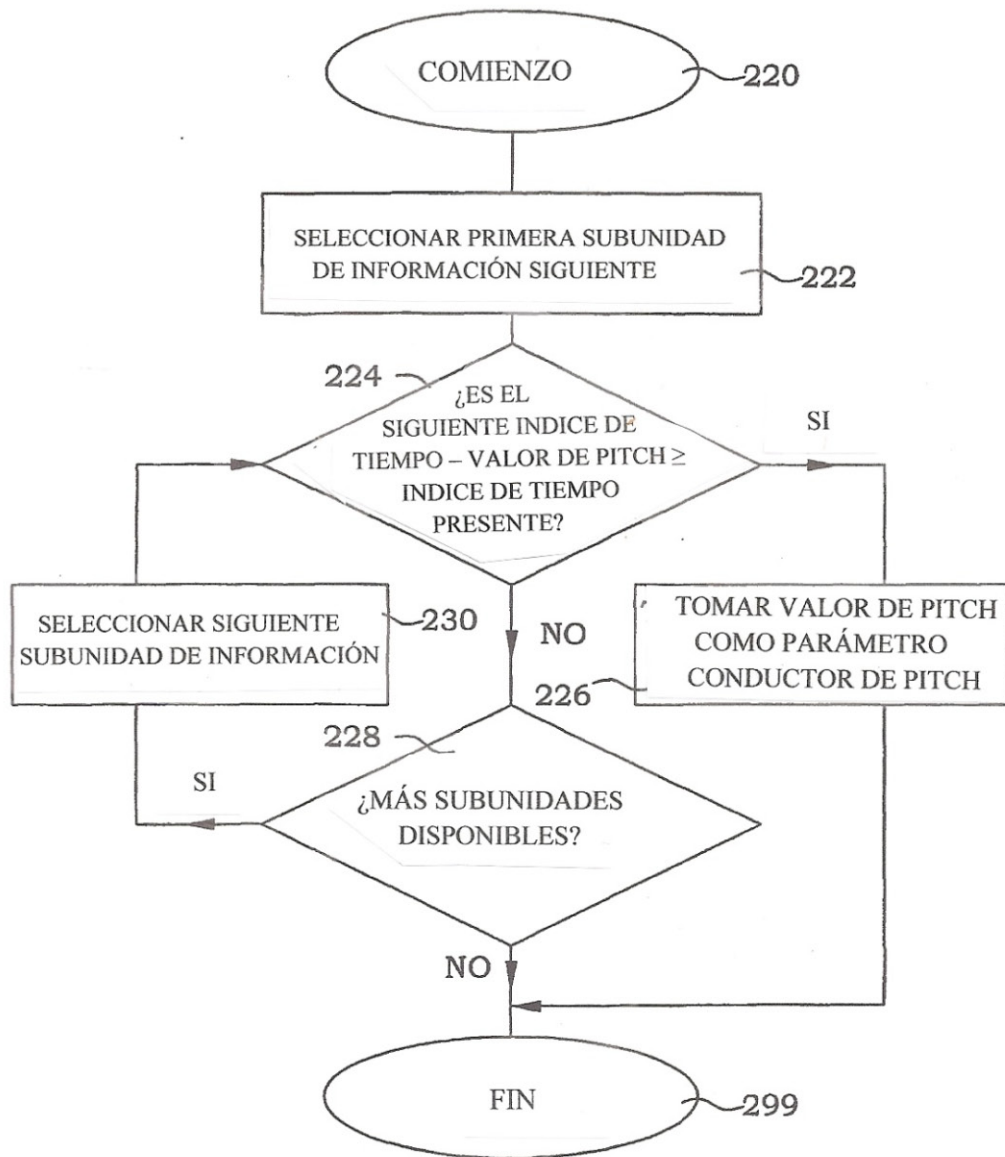


Fig. 12