



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 503**

51 Int. Cl.:
H03G 3/20 (2006.01)
G10L 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06817327 .7**
96 Fecha de presentación : **23.10.2006**
97 Número de publicación de la solicitud: **1997101**
97 Fecha de publicación de la solicitud: **03.12.2008**

54 Título: **Método y sistema para reducir los efectos de los artefactos que producen ruido.**

30 Prioridad: **20.03.2006 US 385553**

45 Fecha de publicación de la mención BOPI:
08.04.2011

45 Fecha de la publicación del folleto de la patente:
08.04.2011

73 Titular/es: **MINDSPEED TECHNOLOGIES, Inc.**
4000 MacArthur Blvd., East Tower
Newport Beach, California 92660-3095, US

72 Inventor/es: **Gao, Yang y**
Shlomot, Eyal

74 Agente: **Curell Aguilá, Marcelino**

ES 2 356 503 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

ANTECEDENTES DE LA INVENCION**1. CAMPO DE LA INVENCION**

5 La presente invención se refiere en general a la codificación del habla. Más particularmente, la presente invención se refiere a la reducción de los efectos de los artefactos que producen ruido en un códec (codificador-decodificador) de voz.

2. ANTECEDENTES DE LA TÉCNICA

10 La compresión del habla se puede utilizar para reducir el número de bits que representan la señal de habla, reduciéndose así el ancho de banda necesario para la transmisión. No obstante, la compresión del habla puede producir el deterioro de la calidad del habla descomprimida. En general, una mayor velocidad binaria producirá una calidad superior, mientras que una menor velocidad binaria producirá una calidad inferior. No obstante, las técnicas actuales de compresión del habla, tales como las técnicas de codificación, pueden producir un habla descomprimida de calidad relativamente alta a velocidades binarias relativamente bajas. En general, las técnicas de codificación actuales intentan representar las características de la señal del habla, importantes desde el punto de vista perceptivo, sin conservar la forma de onda real del habla. Los sistemas de compresión del habla, denominados comúnmente códecs, incluyen un codificador y un decodificador, y se pueden utilizar para reducir la velocidad binaria de señales de habla digitales. Se han desarrollado numerosos algoritmos para códecs de habla, que reducen el número de bits requeridos para codificar digitalmente el habla original al mismo tiempo que se intenta mantener un habla reconstruida de alta calidad.

20 La figura 1 ilustra un sistema convencional de decodificación del habla 100, que incluye un decodificador de excitación 110, un filtro de síntesis 120 y un posprocesador 130. Tal como se muestra, el sistema de decodificación 100 recibe un flujo continuo codificado de bits de habla 102 a través de un medio de comunicación (no representado) desde un codificador, en el que el sistema de decodificación 100 puede formar parte de un dispositivo de comunicaciones móviles, una estación base u otro dispositivo de comunicaciones inalámbricas o por cable que pueda recibir el flujo continuo codificado de bits de habla 102. El sistema de decodificación 100 funciona de modo que decodifica el flujo continuo codificado de bits de habla 102 y genera una señal de habla 132 en forma de una señal digital. A continuación, la señal de habla 132 se puede convertir en una señal analógica por medio de un conversor digital-a-analógico (no representado). La salida analógica del conversor digital-a-analógico puede ser recibida por un receptor (no representado) que puede ser un oído humano, un grabador de cintas magnéticas, o cualquier otro dispositivo que pueda recibir una señal analógica. Alternativamente, la señal del habla 132 puede ser recibida por un dispositivo de grabación digital, un dispositivo de reconocimiento del habla, o cualquier otro dispositivo que pueda recibir una señal digital.

30 El decodificador de excitación 110 decodifica el flujo continuo codificado de bits del habla 102 según el algoritmo de codificación y la velocidad binaria del flujo continuo codificado de bits del habla 102, y genera una excitación decodificada 112. El filtro de síntesis 120 puede ser un filtro de predicción a corto plazo que genera habla sintetizada 122 basándose en la excitación decodificada 112. El posprocesador 130 puede incluir filtrado, aumento de la señal, reducción del ruido, amplificación, corrección de la inclinación (*tilt correction*) y otras técnicas similares que pueden mejorar la calidad de percepción del habla sintetizada 122. El posprocesador 130 puede reducir el ruido audible sin deteriorar perceptiblemente el habla sintetizada 122. La reducción del ruido audible se puede lograr realzando la estructura de los formantes del habla sintetizada 122 ó suprimiendo el ruido en las regiones de frecuencia que no son relevantes desde el punto de vista de la percepción, para el habla sintetizada 122.

40 En los codificadores del habla de velocidad variable, las partes del habla importantes desde el punto de vista de la percepción (por ejemplo, voz sonora, oclusivas, o comienzos sonoros) se codifican con un número mayor de bits, y las partes menos importantes del habla (por ejemplo, partes sordas o silencio entre palabras) se codifican con un número menor de bits. La supresión del ruido hace que mejore la calidad de la señal de voz reconstruida y ayuda a que los codificadores del habla de velocidad variable distingan partes de voz con respecto a partes de ruido. La supresión del ruido ayuda también a que los codificadores del habla de baja velocidad binaria produzcan una salida de mayor calidad mejorando la calidad de percepción del habla. Típicamente, las técnicas de supresión del ruido eliminan el ruido por métodos de sustracción espectral en el dominio de la frecuencia. Un detector de actividad vocal (VAD) determina en el dominio del tiempo si una trama de la señal incluye habla o ruido. Las tramas de ruido se analizan en el dominio de la frecuencia para determinar las características de la señal de ruido. A partir de estas características, los espectros de tramas de ruido se restan de los espectros de las tramas de habla, proporcionando una señal de habla limpia en las tramas de habla.

55 Además, se puede aplicar una atenuación del ruido en el dominio del tiempo para mejorar la calidad de una señal de habla. Por ejemplo, en un sistema de codificación del habla con atenuación del ruido en el dominio del tiempo, descrito en el documento US 2002/00354 70 A1, las ganancias de la codificación del habla con predicción lineal se ajustan por medio de un factor de ganancia para suprimir el ruido de fondo. Tal como se describe en dicho documento, el sistema de codificación del habla utiliza una supresión del ruido en el dominio de la frecuencia junto con una atenuación de la voz en el dominio del tiempo para reducir adicionalmente el ruido de fondo. Después de que una señal analógica se haya convertido en una señal digitalizada, un preprocesador suprime el ruido en la señal digitalizada

5 utilizando un VAD y supresión del ruido en el dominio de la frecuencia. Cuando el VAD identifica una trama asociada a solamente ruido (sin habla), una trama con ventana que incluye la trama identificada de aproximadamente 10 ms se transforma al dominio de la frecuencia. A continuación, se modifican magnitudes espectrales de la señal del habla con ruido para reducir el nivel de ruido según una SNR estimada, y las magnitudes espectrales modificadas se combinan con las fases espectrales sin modificar. A continuación, el espectro modificado se transforma de nuevo al dominio del tiempo. Un esquema de análisis-por-síntesis selecciona la mejor representación para varios parámetros, tales como una ganancia ajustada de libro de códigos fijo, un índice de libro de códigos fijo, un parámetro de retardo, y el parámetro de ganancia ajustada del predictor a largo plazo. Las ganancias se pueden ajustar con un factor de ganancia antes de la cuantificación. El factor de ganancia G_f puede suprimir el ruido de fondo en el dominio del tiempo al mismo tiempo que manteniendo la señal del habla, en el que G_f es definida por $G_f = 1 - C \cdot \text{NSR}$, en el que NSR es la relación ruido/señal basada en tramas, y C es una constante en el intervalo de entre 0 y 1 y controla el grado de reducción del ruido. NSR representa un valor de aproximadamente 1 cuando en la trama se detecta únicamente ruido de fondo, y cuando se detecta habla en la trama, NSR es la raíz cuadrada de la energía de ruido del fondo dividida por la energía de la señal en la trama. Los documentos US5550924 y US6453289 dan a conocer métodos de reducción de ruido para la mejora del habla.

10 No obstante, los planteamientos existentes no plantean correctamente la reducción de efectos de los artefactos que producen ruido en las zonas de silencio de una señal de habla, y especialmente cuando no se encuentra disponible un VAD. En las zonas de silencio, los valores de las muestras son bastante pequeños y sus errores de cuantificación son relativamente muy elevados. Los efectos de estos errores son agravados adicionalmente por la implementación de coma fija de un algoritmo de codificación del habla con su precisión limitada por el uso de valores enteros, lo cual da como resultado que los errores de cuantificación relativos de las zonas de silencio resulten mucho mayores, provocando así una mayor energía de salida en comparación con la energía de entrada original de silencio, y un ruido más audible.

20 Por consiguiente, resultan muy necesarios en la técnica métodos y sistemas de codificación del habla de baja complejidad que puedan reducir eficazmente los efectos de los artefactos en una señal del habla, especialmente en zonas de silencio de la señal del habla y/o sin utilizar un VAD.

SUMARIO DE LA INVENCION

La presente solicitud proporciona métodos y sistemas según las reivindicaciones adjuntas.

Otras características y ventajas de la presente invención se pondrán más claramente de manifiesto para los expertos en la materia a partir de la descripción detallada siguiente y los dibujos adjuntos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las características y ventajas de la presente invención se pondrán más claramente de manifiesto para los expertos en la materia a partir de la descripción detallada siguiente y los dibujos adjuntos, en los que:

la figura 1 ilustra un diagrama de bloques de un sistema de decodificación convencional para la decodificación y el posprocesado de un flujo continuo codificado de bits del habla;

35 la figura 2 ilustra un diagrama de bloques de un posprocesador del habla, según una forma de realización de la presente solicitud; y

la figura 3 ilustra un diagrama de flujo de un método de posprocesado para ser utilizado por el posprocesador del habla de la figura 2, según una forma de realización de la presente solicitud.

DESCRIPCIÓN DETALLADA DE LA INVENCION

40 Aunque la invención se describe con respecto a las formas de realización específicas, los principios de la invención, según se definen mediante las reivindicaciones adjuntas a la presente memoria, se pueden aplicar evidentemente más allá de las formas de realización descritas específicamente de la invención descrita en la presente memoria. Por otra parte, en la descripción de la presente invención, se han omitido ciertos detalles con el fin de no dificultar la apreciación de los aspectos inventivos de la invención. Los detalles omitidos resultarán evidentes para el experto en la materia.

45 Los dibujos de la presente solicitud y su descripción detallada adjunta se refieren a las formas de realización de la invención únicamente ejemplificativas. Para mantener la concisión, otras formas de realización de la invención que utilizan los principios de la presente invención no se describen específicamente en la presente solicitud y no se ilustran específicamente mediante los presentes dibujos. Debería considerarse que, a no ser que se indique lo contrario, los elementos iguales o correspondientes entre las figuras se pueden indicar mediante referencias numéricas iguales o correspondientes.

50 La figura 2 representa un diagrama de bloques del posprocesador del habla 220, según una forma de realización de la presente solicitud. Tal como se muestra, el posprocesador del habla 220 recibe una señal entrante (sig_in) 210 y genera una señal saliente (sig_out) 230 después del posprocesado de sig_in 210 para reducir los efectos

audibles de los artefactos en las zonas de silencio de sig_in 210. En referencia a la figura 3, que ilustra un diagrama de flujo ejemplificativo del método de posprocesado 300 para ser utilizado por el posprocesador del habla 220, un calculador de niveles de energía de subtramas 222 recibe sig_in 210, en la etapa 310, y calcula una suma de nivel de energía absoluto de cada muestra de una subtrama de sig_in 210, que se puede definir mediante:

$$L = \sum_{i=0}^n |\hat{S}(n)| \quad \text{Ecuación 1,}$$

en la que L es el nivel de energía de subtrama, $\hat{S}(n)$ designa sig_in 210, y (n) es el número de muestras.

A continuación, el comparador de niveles de energía de subtramas 224 recibe el nivel de energía de subtrama (L) del calculador de niveles de energía de subtramas 222, y en la etapa 320, el comparador de niveles de energía de subtramas 224 compara el nivel de energía de subtrama (L) con un umbral predeterminado (TH), por ejemplo, 1.024, para obtener una determinación de si el nivel de energía de subtrama (L) es indicativo de una zona de silencio.

A continuación, un modificador del nivel de energía de subtrama 226 recibe la salida del comparador de niveles de energía de subtramas 224. Si el modificador del nivel de energía de subtrama 226 determina que el nivel de energía de subtrama (L) es mayor que el umbral predeterminado (TH), en la etapa 320, el método de posprocesado 300 se desplaza a la etapa 330, siendo identificativo de una zona de no silencio del habla. En la etapa 330, un valor de ganancia (g) se fija a uno (1). Por otro lado, si el modificador del nivel de energía de subtrama 226 determina que el nivel de energía de subtrama (L) no es mayor que el umbral predeterminado (TH), en la etapa 320, el método de posprocesado 300 se desplaza a la etapa 340, lo cual es indicativo de la zona de silencio del habla. En la etapa 340, se fija el valor de ganancia (g) según el resultado del nivel de energía de subtrama (L) dividido por el umbral predeterminado (TH), en donde $0 < g < 1$, tal como se muestra a continuación:

$$g = \begin{cases} 1 & L > TH \\ L/TH & L \# TH \end{cases} \quad \text{Ecuación 2.}$$

A continuación, el método de posprocesado 300 se desplaza a la etapa 350, en la que el modificador del nivel de energía de subtrama 226 modifica el nivel de energía de subtrama (L), para reducir los efectos de los artefactos en las zonas de silencio de la señal saliente (sig_out) 230 del posprocesador, por ejemplo, multiplicando el nivel de energía de subtrama (L) por el umbral predeterminado (TH), tal como se muestra en la etapa 350, lo cual se define mediante:

$$g * \hat{S}(n) \quad \text{Ecuación 3.}$$

En una forma de realización preferida de la presente invención, las formas de realización de la figura 2 y la figura 3 se implementan en un decodificador del habla; no obstante, los aspectos dados a conocer también se pueden implementar mediante un codificador. Además, aunque la ecuación 2 muestra que g es una función de L para zonas de silencio, g también puede ser una función L en zonas de no silencio ($L > TH$) en otras formas de realización. Por ejemplo, en las formas de realización de la figura 2 y la figura 3, g se fija a uno (1) en zonas de no silencio ($L > TH$), de tal manera que $\hat{S}(n)$ se mantiene sin modificar después de la actuación de la ecuación 3 en zonas de no silencio. Adicionalmente, aunque la ecuación 2 muestra que g se define mediante la función de L/TH , otras formas de realización pueden utilizar otras funciones de L .

Los apéndices A y B muestran una implementación de una forma de realización de la presente invención utilizando el lenguaje de programación "C", respectivamente, con coma fija y coma flotante. En una forma de realización, en el lado del decodificador, la energía de la señal se reduce después de detectar una señal de silencio de bajo nivel. Por ejemplo, el nivel de la señal antes del posprocesado del habla se puede definir como:

$$\text{gain_in} = \sum_{i=0}^{L_{\text{sub}}-1} |\text{sig_in}(i)| \quad \text{Ecuación 4,}$$

en la que L_{sub} es el tamaño de subtrama o el número de muestras de habla para cada subtrama, y $\text{sig_in}(i)$ es la señal antes de ejecutar el posprocesado del habla. A continuación, el nivel alisado de sig_in se calcula mediante:

$$\text{Level_in_s m} = 0,75 * \text{Level_in_s m} + 0,25 * \text{gaing_in} \quad \text{Ecuación 5,}$$

en la que el valor inicial de $Level_in_sm$ es cero (0). En una forma de realización, las ecuaciones 4 y 5 las puede ejecutar el calculador de niveles de energía de subtramas 222 del posprocesador 200 en la figura 2.

En dicha forma de realización, la detección del silencio se basa en el nivel de la señal y el parámetro de inclinación espectral, al que se hace referencia también como primer coeficiente de reflexión ($parcor0$), tal como se muestra a continuación:

5

$$Sil_Deci = (Level_in_sm < 1024) \&\&(gain_in < 2 * Level_in_sm) \&\&(parcor0 < 512./32768);$$

Ecuación 6.

en la que, $Sil_Deci=1$ es indicativo de la detección de silencio de bajo nivel. En una forma de realización, el comparador de niveles de energía de subtramas 224 puede determinar Sil_Deci , según la ecuación 6.

10

Además, si se cumple la siguiente condición, $gain_in$ (o el nivel de señal de referencia) se sustituye por el nivel de la señal alisada, lo cual puede ser ejecutado por el modificador del nivel de energía de subtrama 226:

$$si \ ((Sil_Deci) == 1) \&\&(Level_in_sm < gain_in) \\ gain_in = Level_in_sm;$$

Ecuación 7.

En la que, $(Level_in_sm < gain_in)$ es indicativo de la no existencia de picos grandes en la señal. A continuación, considerando que $sig_out(i)$, $i=0, 1, \dots, L_{sub}-1$, es la señal posfiltrada antes del ajuste de ganancia que se describe a continuación, el nivel inicial de la señal posfiltrada se puede calcular por medio del calculador de niveles de energía de subtramas 222, de la manera siguiente:

15

$$gain_out = \sum_{i=0}^{L_{sub}-1} |sig_out(i)|$$

Ecuación 8,

en la que, la ganancia inicial para ajustar la energía de la señal posfiltrada se puede determinar por medio del modificador del nivel de energía de subtrama 226 como:

$$g0 = \frac{gain_in}{gain_out}$$

Ecuación 9.

20

Una vez que se ha detectado el silencio de nivel bajo, es decir, $Sil_Deci=1$, el modificador del nivel de energía de subtrama 226 puede reducir adicionalmente la ganancia inicial aplicando el siguiente factor de ganancia ($gain$), tal como se muestra a continuación:

$$gain = k1 * (Level_in_sm/1024) + (1 - k1);$$

Ecuación 10,

en la que $0 \leq gain \leq 1$, y $k1$ ($0 \leq k1 \leq 1$) es una función de $parcor0$, y:

25

$$k1 = (512./32768) - parcor0; \\ si \ (k1 > (2047/32768)) \\ k1 = (2047/32768); \\ k1 = k1/(2047/32768);$$

Por último, el modificador del nivel de energía de subtrama 226 aplicará un ajuste de ganancia a la señal posfiltrada, de la manera siguiente, en la que g_sm es la ganancia alisada:

```

si (Sil_Deci == 1) {
    g0 = g0 * gain
    para (i = 0; i < Lsub; i++) {
        g_sm = g_sm * 0,95 + g0 * 0,05;
        sig_out(i) = sig_out(i) * g_sm;
    }
}

```

5 En otras palabras, *sig_out* se modifica multiplicando *sig_out* para cada una de la pluralidad de muestras salientes por una ganancia alisada (*g_sm*), en la que *g_sm* se obtiene utilizando iteraciones desde 0 a n-1 de (*g_sm* previa *0,95 + g0*0,05), en la que n es el número de muestras o el tamaño de la subtrama, y *g_sm* previa es cero (0) antes de la primera iteración.

10 En una forma de realización preferida de la presente invención, la reducción de ganancia de silencio descrita anteriormente se realiza únicamente para la señal de habla de banda estrecha (0 a 4 KHz) en el decodificador, pero no para la señal de habla de banda ancha (4 a 8 KHz). No obstante, otros ejemplos pueden incluir implementaciones en codificador y/o de banda ancha. Además, cuando se detecta una transición desde/hacia la señal de habla de banda estrecha hacia/desde la señal de habla de banda ancha, la ganancia se puede cambiar o ajustar gradualmente en lugar de una aplicación brusca (transición de banda ancha a banda estrecha) o una no aplicación (transición de banda estrecha a banda ancha) de la ganancia para reducir los efectos de los artefactos en las zonas de silencio.

15 A partir de la descripción anterior de la invención, se pone de manifiesto que se pueden utilizar varias técnicas para implementar los conceptos de la presente invención sin apartarse de su alcance. Por otra parte, aunque la invención se ha descrito haciendo referencia específicamente a ciertas formas de realización, los expertos en la materia apreciarán que se pueden introducir cambios en cuanto a forma y detalle sin apartarse del alcance de la invención. Por ejemplo, se contempla que la circuitería dada a conocer en la presente memoria se puede implementar en software, o viceversa. Las formas de realización descritas deben considerarse en todos los aspectos como ilustrativas y no limitativas. Debería apreciarse asimismo que la invención no se limita a las formas de realización particulares descritas en la presente memoria, sino que pueden introducirse muchas reestructuraciones, modificaciones, y sustituciones sin apartarse del alcance de la invención, tal como es definido en las reivindicaciones.

APÉNDICE A

```

/*****/
/*****/
25 /* Limpieza de silencios con coma fija */
/*****/
/*****/

Word16Level_in_sm=1024; /* temporalmente poner esta variable aquí */
Word16PostNB=0; /* fijar temporalmente a 0; valor real: 0<=PostNB<=1 */
30 /*-----
* G729EV_G729_scale_st – control de la ganancia de subtrama
* gain[n] = G729EV_G729_AGC_FAC * gain[n-1] + (1 – AGC_FAC) g_in/g_out
*-----
*/

35 static void G729EV_G729_scale_st(Word16 *sig_in, /* entrada: señal de entrada del posfiltro */
    Word16 *sig_out, /*entrada/salida: señal de salida del posfiltro */
    Word16 *gain_prec, /* entrada/salida: último valor de la ganancia para la subtrama */
#ifdef SILENCE_CLEANING

```

```

                Word16 parcor0,
                Word16 PostNB,
                Word32 *Level_in_sm

#endif
5   )
   {
   Word32     L_acc, L_temp;
   Word16     i;
   Word16     scal_in, scal_out;
10  Word16     s_g_in, s_g_out, temp, sh_g0, g0;
   #ifdef SILENCE_CLEANING
   Word16     gain;
   Word16     Cond;
   #endif
15         /* calcular entrada de nuevo */
           L_acc= 0L;
   #ifdef WMOPS
           move32();
   #endif
20     for (i=0; i<G729EV_G729_L_SUBFR; i++)
       {
           L_temp=L_abs(L_deposit_l(sig_in[i]));
           L_acc=L_add(L_acc, L_temp);
       }
25 #ifdef SILENCE_CLEANING
       /* Nivel de alisado */
       *Level_in_sm=L_add(L_shr(*Level_in_sm, 1), L_shr(*Level_in_sm, 2));
       *Level_in_sm=L_add(*Level_in_sm, L_shr(L_acc, 2));
       /* Detectar silencio*/
30     Cond = (*Level_in_sm<1024) && (L_acc<L_shl(*Level_in_sm, 1)) && (parcor0<512);
           /* Si se detecta silencio, sustituir el nivel original por un nivel alisado */
       if (Cond ==1)
           L_acc = *Level_in_sm;
   #endif
35 #ifdef WMOPS
           test();

```

```

#endif
    if (L_acc== 0L)
    {
        g0 =0;
5  #ifdef WMOPS
        move16();
    #endif
    }
    else
10  {
        scal_in = norm_l(L_acc);
        L_acc = L_shl(L_acc, scal_in);
        s_g_in = extract_h(L_acc); /*normalizado */
        /* Calcular salida de nuevo */
15  L_acc =0L;
    #ifdef WMOPS
        move32();
    #endif
    for (i = 0; i<G729EV_G729_L_SUBFR; i++)
20  {
        L_temp = L_abs(L_deposit_l(sig_out([i]));
        L_acc = L_add(L_acc, L_temp);
    }
    #ifdef WMOPS
25  test();
    #endif
    if (L_acc == 0L)
    {
        *gain_prec = 0;
30  #ifdef WMOPS
        move16();
    #endif
        return;
    }
35  scal_out = norm_l(L_acc);
    L_acc = L_shl(L_acc, scal_out);

```

```

s_g_out = extract_h(L_acc); /* normalizado */
sh_g0 = add(scal_in, l);
sh_g0 = sub(sh_g0, scal_out); /* scal_in - scal_out + 1 */

#ifdef WMOPS
5      test();
#endif

      if (sub(s_g_in, s_g_out) < 0)
      {
10      g0 = div_s(s_g_in, s_g_out); /* s_g_in/s_g_out en Q15 */
      }
      else
      {
15      temp = sub(s_g_in, s_g_out); /* suficiente por normalizado */
      g0 = shr(div_s(temp, s_g_out), 1);
      g0 = add(g0, (Word16) 0x4000); /* s_g_in/s_g_out en Q14 */
      sh_g0 = sub(sh_g0, 1);
      }
      /* L_gain_in/L_gain_out en Q14 */
      /* sobredesbordamientos si L_gain_in > 2 * L_gain_out */
20      g0 = shr(g0, sh_g0); /*sh_g0 puede ser >0, <0, ó =0 */

#ifdef SILENCE_CLEANING
      if ( Cond == 1 )
      { /* Aplicar una reducción de ganancia para el silencio; la ganancia se define como gain =
25      (Level_in_sm/MAX_SILENCE_LEVEL)*k1 + (1-k1); k1 (0=<k1<=1) es una función de
      PARCOR0 */
      /* k1 en Q15*/
      temp=sub(512, parcor0);
      if (temp>2047) temp=2047;
      temp=shl(temp, 4);
30      /* gain =(Level_in_sm/MAX_SILENCE_LEVEL) en Q15 */
      if (*Level_in_sm>1023) gain= 1023;
      else gain = extract_l(*Level_in_sm);
      gain = shl(gain, 5);
      /* gain = gain*k1 + 1-k1*/
35      gain = mult_r(gain, temp);
      gain = add(gain, sub(32767, temp));
      gain = mult_r(gain, sub(32767, PostNB));

```

```

        gain= add(gain, Post_NB);
        g0 =mult_r(g0, gain);
    }
#endif
5          g0 = mult_r(g0, G729EV_G729_AGC_FAC1); /* L_gain_in/L_gain_out *
AGC_FAC1 */
    }
    /* gain(n) = G729EV_G729_AGC_FAC gain(n-1) + G729EV_G729_AGC_FAC1
gain_in/gain_out */
10    /* sig_out(n) = gain(n) sig_out(n) */
    gain = *gain_prec;
    for (i = 0; i < G729EV_G729_L_SUBFR; i++)
    {
        temp = mult_r(G729EV_G729_AGC_FAC, gain);
15        gain = add(temp, g0); /* en Q14 */
        L_temp = L_mult(gain, sig_out[i]);
        L_temp =L_shl(L_temp, 1);
        sig_out[i] = round(L_temp);
    }
20    *gain_prec =gain;
#ifdef WMOPS
    move16();
    move16();
#endif
25    return;
}

```

APÉNDICE B

```

/*****
/*****
30 /*          Limpieza de silencios con coma fija          */
/*****
/*****
REAL Level_in_sm=1024.; /* temporalmente poner esta variable aquí */
REAL PostNB=0.; /* fijar temporalmente a 0; valor real: 0<=PostNB<=1 */
35 /**-----
* Función G729EV_G729_scale_st – control de la ganancia de subtrama

```

```

*   gain[n] = G729EV_G729_AGC_FAC * gain[n-1] + (1 - G729EV_G729_AGC_FAC)
g_in/g_out
*-----
*/
5 void
G729EV_G729_scale_st (REAL *sig_in,    /*<entrada : señal de entrada del posfiltro */
                    REAL *sig_out,    /*<entrada/salida: señal de salida del posfiltro */
                    REAL *gain_prec /*<entrada/salida: último valor de la ganancia para la subtrama */

#ifdef SILENCE_CLEANING
10
                    ,
                    REAL parcor0,
                    REAL PostNB,
                    REAL *Level_in_sm

#endif
15 )
{
    int i;
    REAL gain_in, gain_out;
    REAL g0, gain;
20 #ifdef SILENCE_CLEANING
    short Cond;
    REAL k1;
#endif
    /* calcular ganancia de nuevo */
25 gain_in =(REAL) 0.;
    for (i = 0; i<G729EV_G729_L_SUBFR; i++)
    {
        gain_in += (REAL) fabs (sig_in[i]);
    }
30 #ifdef SILENCE_CLEANING
    /* Nivel de alisado */
    *Level_in_sm =0,75*(*Level_in_sm) + 0,25*gain_in;
    /* Detectar silencio */
    Cond = (*Level_in_sm<1024.) && (gain_in<*Level_in_sm*2.) &&
35 (parcor0<512./32768);
    /* Si se detecta silencio, sustituir el nivel original por nivel suavizado */

```

```

if (Cond == 1)
    gain_in = *Level_in_sm;
#endif
if (gain_in==(REAL) 0.)
5   {
    g0 = (REAL) 0.;
    }
else
    {
10  /*Calcular salida de nuevo */
    gain_out =(REAL) 0.;
    for (i = 0; i < G729EV_G729_L_SUBFR; i++)
        {
15      gain_out += (REAL) fabs (sig_out[i]);
        }
    if (gain_out == (REAL) 0.)
        {
            *gain_prec=(REAL) 0.;
            return;
20      }
    g0 = gain_in / gain_out;
#ifdef SILENCE_CLEANING
    if ( Cond == 1 )
25      { /* Aplicar una reducción de la ganancia para el silencio; la ganancia se define como gain =
        (Level_in_sm/MAX_SILENCE_LEVEL)*k1 + (1-k1);
        k1 (0=k1<=1) es una función de PARCOR0 */
        /*k1*/
        k1=(512./32768) – parcor0;
        if (k1>(2047./32768))k1=(2047./32768);
30      k1 /= (2047./32768);
        /* gain = (Level_in_sm/MAX_SILENCE_LEVEL)*/
        if (*Level_in_sm>1023) gain =1.;
        else gain =*Level_in_sm/1024.;
        gain = gain*k1 + 1-k1;
35      gain = gain *(1.-PostNB) + PostNB;
        g0 *= gain;
    }

```

```

#endif
    g0 *= G729EV_G729_AGC_FAC1;
}
/*      calcular gain(n)      =      G729EV_G729_AGC_FAC gain(n-1)      +      (1-
5  G729EV_G729_AGC_FAC)gain_in/gain_out */
/* sig_out(n) = gain(n) sig_out(n) */
gain = *gain_prec;
for (i = 0; i < G729EV_G729_L_SUBFR; i++)
{
10  gain *= G729EV_G729_AGC_FAC;
    gain += g0;
    sig_out[i] *= gain;
}
*gain_prec = gain;
15  return;
}

```

REIVINDICACIONES

- 5 1. Método de reducción del efecto de los artefactos que producen ruido en las zonas de silencio de una señal de habla para su utilización por un sistema de decodificación del habla, comprendiendo el método: obtener una pluralidad de muestras entrantes de una subtrama del habla; sumar un valor absoluto de un nivel de energía para cada una de la pluralidad de muestras entrantes con el fin de generar un nivel de entrada total ($gain_in$); alisar el nivel de entrada total para generar un nivel alisado ($Level_in_sm$); determinar que la subtrama del habla está en una zona de silencio sobre la base del nivel de entrada total, el nivel alisado y un parámetro de inclinación; definir una ganancia utilizando $k1 * (Level_in_sm/1024) + (1-k1)$, en el que $k1$ es una función del parámetro de inclinación; y modificar un nivel de energía de la subtrama del habla utilizando la ganancia.
- 10 2. Método según la reivindicación 1, en el que el alisado se realiza utilizando: $Level_in_sm = 0,75 * Level_in_sm + 0,25 * gain_in$ y/o la determinación se realiza utilizando: $(Level_in_sm < 1024) \&\& (gain_in < 2 * Level_in_sm) \&\& (parcor0 < 512/32768)$.
- 15 3. Método según la reivindicación 1, que comprende además: asignar $Level_in_sm$ a $gain_in$ ($gain_in = Level_in_sm$) si $Level_in_sm < gain_in$, comprendiendo además opcionalmente: sumar un valor absoluto de un nivel de energía para cada una de la pluralidad de muestras salientes, antes de la modificación, con el fin de generar un nivel de salida total ($gain_out$); determinar una ganancia inicial utilizando ($gain_in / gain_out$); y modificar la ganancia utilizando la ganancia inicial para generar una ganancia modificada ($g0$), opcionalmente además
- 20 en el que la modificación comprende multiplicar sig_out para cada una de la pluralidad de muestras salientes por una ganancia alisada (g_sm), en el que g_sm se obtiene utilizando iteraciones desde 0 a $n-1$ de (g_sm previa $* 0,95 + g0 * 0,05$), en el que n es el número de muestras, y g_sm previa es cero (0) antes de la primera iteración.
- 25 4. Sistema de decodificación del habla para reducir el efecto de los artefactos que producen ruido en las zonas de silencio de una señal de habla, comprendiendo el sistema de decodificación del habla: un calculador del nivel de energía de subtrama configurado para obtener una pluralidad de muestras entrantes de una subtrama del habla, y configurado para sumar un valor absoluto de un nivel de energía para cada una de la pluralidad de muestras entrantes con el fin de generar un nivel de entrada total ($gain_in$), y configurado además para alisar el nivel de entrada total con el fin de generar un nivel alisado ($Level_in_sm$); un comparador de nivel de energía de subtrama configurado para determinar que la subtrama del habla se encuentra en una zona de silencio sobre la base del nivel de entrada total, el nivel alisado y un parámetro de inclinación; y un modificador del nivel de energía de subtrama configurado para definir una ganancia utilizando $k1 * (Level_in_sm/1024) + (1-k1)$, en el que $k1$ es una función del parámetro de inclinación, y configurado además para modificar un nivel de energía de la subtrama del habla utilizando la ganancia.
- 30 5. Sistema de decodificación del habla según la reivindicación 4, en el que:
- el calculador de nivel de energía de subtrama alisa el nivel de entrada total utilizando: $Level_in_sm = 0,75 * Level_in_sm + 0,25 * gain_in$.
- 35 6. Sistema de decodificación del habla según la reivindicación 4, en el que el comparador de nivel de energía de subtrama determina que la subtrama del habla está en la zona de silencio utilizando:
- $(Level_in_sm < 1024) \&\& (gain_in < 2 * Level_in_sm) \&\& (parcor0 < 512/32768)$.
- 40 7. Sistema de decodificación del habla según la reivindicación 4, en el que el modificador del nivel de energía de subtrama asigna $Level_in_sm$ a $gain_in$ ($gain_in = Level_in_sm$) si $Level_in_sm < gain_in$.
- 45 8. Sistema de decodificación del habla según la reivindicación 7, en el que el calculador del nivel de energía de subtrama está configurado además para sumar un valor absoluto de un nivel de energía para cada una de la pluralidad de muestras salientes, antes de la modificación por parte del modificador del nivel de energía de subtrama, con el fin de generar un nivel de salida total ($gain_out$), y el modificador del nivel de energía de subtrama está configurado además para determinar una ganancia inicial utilizando ($gain_in / gain_out$) y modificar la ganancia utilizando la ganancia inicial con el fin de generar una ganancia modificada ($g0$).
- 50 9. Sistema de decodificación del habla según la reivindicación 8, en el que el modificador del nivel de energía de subtrama modifica el nivel de energía de subtramas del habla multiplicando sig_out para cada una de la pluralidad de muestras salientes por una ganancia alisada (g_sm), en el que g_sm se obtiene utilizando iteraciones desde 0 a $n-1$ de (g_sm previa $* 0,95 + g0 * 0,05$), en el que n es el número de muestras, y la g_sm previa es cero (0) antes de la primera iteración.

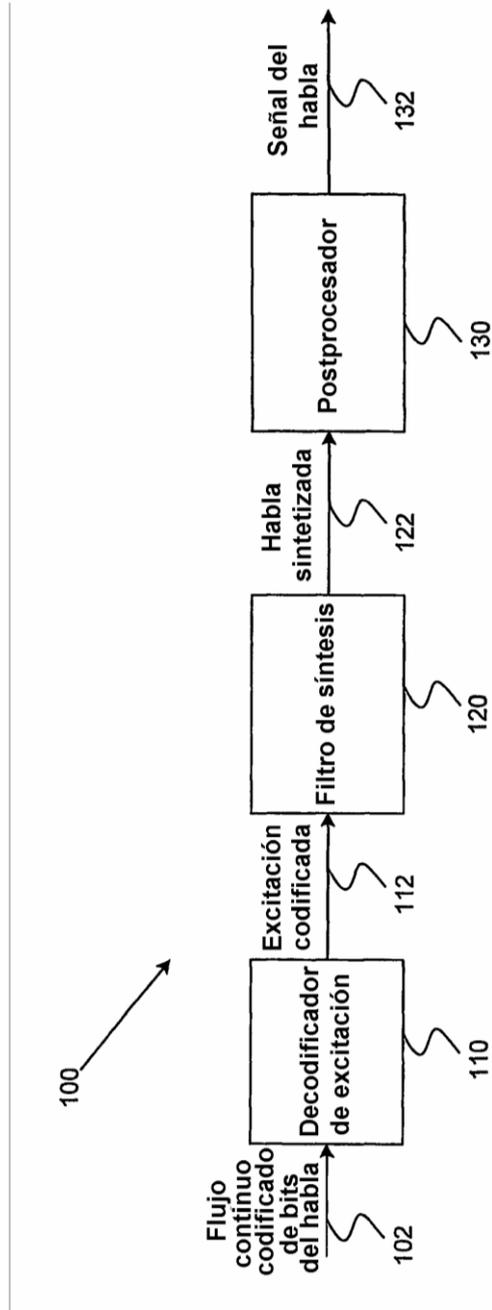


FIG. 1
(TÉCNICA ANTERIOR)

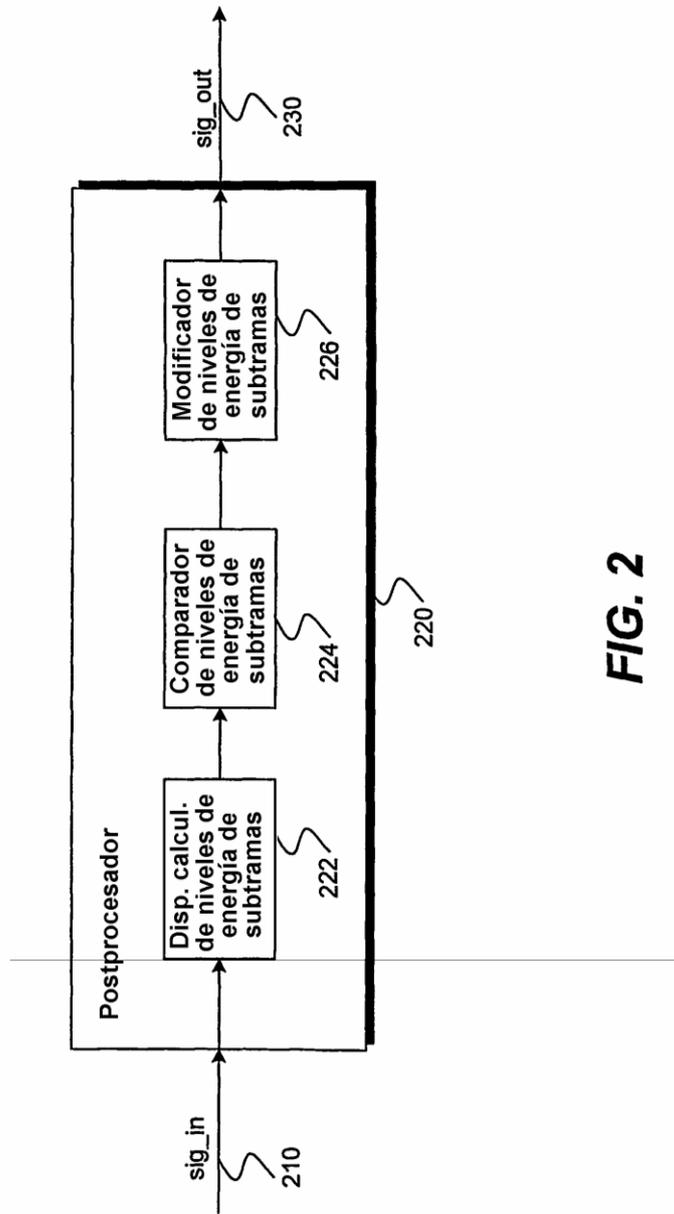


FIG. 2

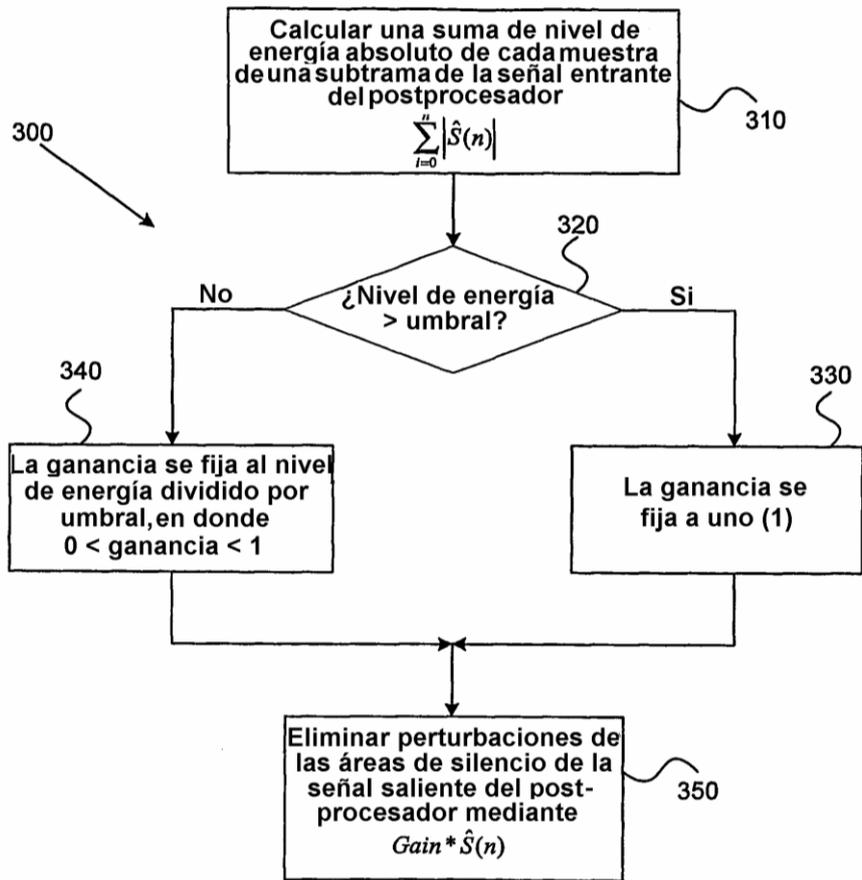


FIG. 3