



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 360 176**

51 Int. Cl.:  
**G10L 21/02** (2006.01)  
**G10L 19/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **09163673 .8**  
96 Fecha de presentación : **18.04.2001**  
97 Número de publicación de la solicitud: **2099028**  
97 Fecha de publicación de la solicitud: **09.09.2009**

54 Título: **Alisamiento de discontinuidades entre tramas de habla.**

30 Prioridad: **24.04.2000 US 557283**

45 Fecha de publicación de la mención BOPI:  
**01.06.2011**

45 Fecha de la publicación del folleto de la patente:  
**01.06.2011**

73 Titular/es: **QUALCOMM Incorporated**  
**5775 Morehouse Drive**  
**San Diego, California 92121-1714, US**

72 Inventor/es: **Manjunath, Sharath;**  
**Huang, Penjung y**  
**Choy, Eddie-Lun, Tik**

74 Agente: **Carpintero López, Mario**

ES 2 360 176 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

Alisamiento de discontinuidades entre tramas de habla

**Antecedentes de la Invención****I. Campo de la Invención**

- 5 La presente invención pertenece en general al campo del procesamiento de habla, y más específicamente a procedimientos y aparatos para compensar los borrados de trama en codificadores de habla de tasa variable.

**II. Antecedentes**

10 La transmisión de voz mediante técnicas digitales ha experimentado una difusión, en particular en aplicaciones de radiotelefonía digital y de larga distancia. Esto, a su vez, ha originado interés en cuanto a la determinación de al menos la cantidad de información que puede ser enviada por un canal mientras se mantiene la calidad percibida del habla reconstruida. Si el habla se transmite simplemente mediante muestreo y digitalización, se precisa una tasa de datos del orden de sesenta y cuatro kilobits por segundo (Kbps) para conseguir la calidad de habla de la telefonía analógica convencional. Sin embargo, mediante el uso de análisis de habla, seguido de la codificación adecuada, transmisión y re-síntesis en el receptor, se puede alcanzar una reducción significativa en la tasa de datos.

15 Los dispositivos para comprimir el habla encuentran uso en muchos campos de las telecomunicaciones. Un ejemplo de campo es el de las comunicaciones inalámbricas. El campo de las comunicaciones inalámbricas tiene muchas aplicaciones incluyendo, por ejemplo, los teléfonos inalámbricos, radiobúsqueda, bucles inalámbricos locales, telefonía inalámbrica del tipo de los sistemas celulares y de telefonía PCS, telefonía móvil de Protocolo de Internet (IP), y sistemas de comunicación por satélite. Una aplicación particularmente importante consiste en la telefonía  
20 inalámbrica para abonados móviles.

Se han desarrollado diversas interfaces sobre-el-aire para sistema de comunicación inalámbrica que incluyen, por ejemplo, acceso múltiple por división de frecuencia (FDMA), acceso múltiple por división de tiempo (TDMA), y acceso múltiple por división de código (CDMA). En relación con estos últimos, se han establecido diversos estándares domésticos e internacionales que incluyen, por ejemplo, el Servicio de Teléfono Móvil Avanzado (AMPS),  
25 el Sistema Global para Comunicaciones Móviles (GSM), y el Estándar Interim 95 (IS-95). Un sistema ejemplar de comunicación de telefonía inalámbrica es un sistema de acceso múltiple por división de código (CDMA). El estándar IS-95 y sus derivados, el IS-95A, ANSI J-STD-008, IS-95B, los estándares US-95C e IS-2000 propuestos de tercera generación, etc. (mencionados en su conjunto en la presente memoria como IS-95), han sido promulgados por la Telecommunication Industry Association (TIA), y otros cuerpos de normas bien conocidos para especificar el uso de  
30 una interfaz sobre-el-aire de CDMA para sistemas de comunicación celular o de telefonía PCS. Sistemas ejemplares de comunicación inalámbrica configurados sustancialmente de acuerdo con el uso del estándar IS-95, han sido descritos en los documentos de Patente U.S. núms. 5.103.459 y 4.901.307.

Los dispositivos que emplean técnicas para comprimir el habla mediante extracción de parámetros que se relacionan con un modelo de generación de habla humana, se denominan codificadores de habla. Un codificador de habla  
35 divide la señal de habla entrante en bloques de tiempo, o tramas de análisis. Los codificadores de habla comprenden típicamente un codificador y un descodificador. El codificador analiza la trama de habla entrante para extraer ciertos parámetros relevantes, y a continuación cuantifica los parámetros en una representación binaria, es decir, un conjunto de bits o paquete de datos binarios. Los paquetes de datos son transmitidos por el canal de comunicación hasta un receptor y un descodificador. El descodificador procesa los paquetes de datos, descuantifica  
40 los mismos para generar los parámetros, y re-sintetiza las tramas de habla utilizando los parámetros descuantificados.

La función del codificador de habla consiste en comprimir la señal de habla digitalizada en una señal de baja tasa de bit, eliminando todas las redundancias naturales inherentes al habla. La compresión digital se consigue representando la trama de habla de entrada con un conjunto de parámetros, y empleando cuantificación para  
45 representar los parámetros con un conjunto de bits. Si la trama de habla de entrada tiene un número de bits  $N_i$ , y el paquete de datos producido por el codificador de habla tiene un número de bits  $N_0$ , el factor de compresión alcanzado por el codificador de habla es  $C_r = N_i/N_0$ . El reto consiste en conservar la alta calidad de voz del habla descodificada mientras se consigue el factor de compresión objetivo. El rendimiento del codificador de habla depende de (1) lo bien que se realice el modelo de habla, o la combinación del proceso de análisis y síntesis  
50 descrito en lo que antecede, y (2) lo bien que se realice el proceso de cuantificación de parámetro a la tasa de bit objetivo de  $N_0$  bits por trama. La meta del modelo de habla consiste por lo tanto en capturar la esencia de la señal de habla, o la calidad de la voz objetivo, con un pequeño conjunto de parámetros para cada trama.

Quizás sea más importante en el diseño de un codificador de habla la búsqueda de un buen conjunto de parámetros (incluyendo los vectores) para describir la señal de habla. Un buen conjunto de parámetros requiere un bajo ancho de banda del sistema para la reconstrucción de una señal de habla perceptivamente precisa. El pitch, la potencia de  
55

señal, la envolvente (o formantes) espectral, los espectros de amplitud, y los espectros de fase, son ejemplos de parámetros de codificación de habla.

5 Los codificadores de habla pueden ser implementados como codificadores en el dominio del tiempo, que pretenden capturar la forma de onda de habla en el dominio del tiempo, empleando procesamiento de alta resolución en el tiempo para codificar pequeños segmentos de habla (típicamente, sub-tramas de 5 milisegundos (ms)) cada vez. Para cada sub-trama, se halla un representante de alta precisión a partir de un espacio de código de claves por medio de varios algoritmos de búsqueda conocidos en el estado de la técnica. Alternativamente, los codificadores de habla pueden ser implementados como codificadores de dominio frecuencial, que pretenden capturar el espectro de habla a corto plazo de la trama de habla entrante con un conjunto de parámetros (análisis), y que emplean un proceso de síntesis correspondiente para volver a crear la forma de onda de habla a partir de parámetros espectrales. El cuantificador de parámetro conserva los parámetros representándolos con representaciones almacenadas de vectores de código, de acuerdo con técnicas de cuantificación conocidas descritas en A. Gersho & R. M. Gray, Cuantificación de Vector y Compresión de Señal (1992).

15 Un codificador de habla en el dominio del tiempo bien conocido, es el codificador Predictivo Lineal Excitado por Código (CELP) descrito en L.B. Rabiner & R.W. Schafer, Procesamiento Digital de Señales de Habla, 396-453 (1978). En un codificador CELP, las correlaciones a corto plazo, o redundancias, en la señal de habla, son retiradas mediante un análisis de predicción lineal (LP), que halla los coeficientes de un filtro formante de corto plazo. La aplicación del filtro de predicción a corto plazo a la trama de habla entrante, genera una señal de residuo de LP, que es además modelada y cuantificada con parámetros de filtro de predicción a largo plazo y un código de claves estocástico consiguiente. De ese modo, la codificación CELP divide la tarea de codificar la forma de onda de habla en el dominio del tiempo, en las tareas separadas de codificación de los coeficientes de filtro a corto plazo de LP y codificación del residuo de LP. La codificación en el dominio del tiempo puede ser llevada a cabo a tasa fija (es decir, utilizando el mismo número de bits,  $N_0$  para cada trama), o a tasa variable (en la que se utilizan diferentes tasas de bit para diferentes tipos de contenidos de tramas). Los codificadores de tasa variable intentan utilizar solamente la cantidad de bits necesarios para codificar los parámetros códec a un nivel adecuado para obtener una calidad objetiva. Un ejemplo de codificador CELP de tasa variable ha sido descrito en la Patente U.S. núm. 5.414.796.

30 Los codificadores en el dominio del tiempo, tal como el codificador CELP, se basan típicamente en un alto número de bits,  $N_0$ , por trama para conservar la precisión de la forma de onda del habla en el dominio del tiempo. Tales codificadores suministran típicamente una voz excelente, calidad proporcionada por el número de bits,  $N_0$ , por trama relativamente grande (por ejemplo, 8 kbps o superior). Sin embargo, a tasas de bit bajas (4 kbps o más bajas), los codificadores en el dominio del tiempo fallan en cuanto a mantener una calidad alta y un comportamiento robusto debido al número limitado de bits disponibles. A tasas de bit bajas, el espacio limitado de código de claves recorta la capacidad de emparejamiento de forma de onda de los codificadores en el dominio del tiempo convencionales, los cuales han sido así desplegados con éxito en aplicaciones comerciales de tasa más alta. Por ello, a pesar de las mejoras con el paso del tiempo, muchos sistemas de codificación CELP que operan a tasas de bit bajas, adolecen de una distorsión perceptivamente significativa caracterizada como ruido.

40 Existe en la actualidad un interés investigador creciente y una fuerte necesidad comercial por desarrollar un codificador de habla de alta calidad que opere con tasas de bit de medias a bajas (es decir, en la gama de 2,4 a 4 kbps y más baja). Las áreas de aplicación incluyen la telefonía inalámbrica, las comunicaciones por satélite, la telefonía de Internet, diversas aplicaciones multimedia y de propagación de voz, correo de voz, y otros sistemas de almacenamiento de voz. Las fuerzas impulsoras son la necesidad de una alta capacidad y la demanda de un comportamiento robusto bajo situaciones de pérdida de paquete. Diversos esfuerzos recientes de estandarización de codificación de habla son otra fuerza impulsora directa que impulsa la investigación y el desarrollo de algoritmos de codificación de habla de baja tasa. Un codificador de habla de baja tasa crea más canales, o usuarios, para que el ancho de banda de la aplicación permisible y el codificador de habla de baja tasa acoplado a una capa adicional de codificación de canal adecuado, puedan ajustar la disponibilidad de bits de las especificaciones del codificador, y proporcionar un comportamiento robusto bajo condiciones de error de canal.

50 Una técnica eficaz para codificar habla eficientemente a bajas tasas de bit, es la codificación multimodo. Un ejemplo de técnica de codificación multimodo ha sido descrita en la Patente U.S. núm. 6.691.084, titulada CODIFICACIÓN DE HABLA DE TASA VARIABLE. Los codificadores multimodo convencionales aplican diferentes modos, o algoritmos de codificación-descodificación, a diferentes tipos de tramas de habla entrantes. Cada modo, o proceso de codificación-descodificación, se personaliza para que represente óptimamente un cierto tipo de segmento de habla, tal como habla con voz, habla sin voz, habla de transición (por ejemplo, entra voz y no voz), y ruido de fondo (silencio, o no habla) de la manera más eficiente. Un mecanismo externo, de decisión de modo en bucle abierto, examina la trama de habla entrante y toma una decisión con relación a qué modo aplicar a la trama. La decisión de modo en bucle abierto se lleva a cabo típicamente mediante extracción de un número de parámetros a partir de la trama entrante, evaluando los parámetros respecto a ciertas características temporales y espectrales, y basando el modo de decisión en la evaluación.

Los sistemas de codificación que operan a tasas del orden de 2,4 kbps, son generalmente de naturaleza paramétrica. Es decir, tales sistemas de codificación operan transmitiendo parámetros que describen el período de pitch y la envolvente espectral (o formantes) de la señal de habla a intervalos regulares. Ilustrativos de estos codificadores conocidos como paramétricos, son el sistema vocoder de LP.

- 5 Los vocoders de LP modelan una señal de habla con voz, con un solo pulso por período de pitch. Esta técnica básica puede ser aumentada para que incluya información de transmisión acerca de la envolvente espectral, entre otras cosas. Aunque los vocoders de LP proporcionan un rendimiento razonable en general, éstos pueden introducir una distorsión perceptivamente significativa, caracterizada típicamente como zumbido.

- 10 En los últimos años, han emergido codificadores que son híbridos de ambos codificadores de forma de onda y codificadores paramétricos. Ilustrativo de estos codificadores denominados híbridos es el sistema de codificación de habla por interpolación de forma de onda de prototipo (PWI). El sistema de codificación de PWI puede ser también conocido como codificador de habla de período de pitch de prototipo (PPP). Un sistema de codificación de PWI proporciona un procedimiento eficaz para codificar habla con voz. El concepto básico de PWI consiste en extraer un ciclo de pitch representativo (la forma de onda de prototipo) a intervalos fijos, para transmitir su descripción, y para reconstruir la señal de habla mediante interpolación entre las formas de onda de prototipo. El procedimiento de PWI puede operar ya sea sobre la señal residual de LP o ya sea sobre la señal de habla. Un ejemplo de codificación de habla de PWI, o de PPP, ha sido descrito en la Solicitud de Patente U.S. Serial núm. 09/217.494, titulada CODIFICACIÓN PERIÓDICA DE HABLA, depositada el 21 de Diciembre de 1998, cedida a la cesionaria de la presente invención. Otros codificadores de habla de PWI, o de PPP, han sido descritos en la Patente U.S. núm. 5.884.253 y en W. Bastiaan Kleijn & Wolfgang Granzow, Procedimientos para Interpolación de Forma de Onda en Codificación de Habla, en Procesamiento de Señal Digital 1, 215-230 (1991).

- 25 En la mayor parte de los codificadores de habla convencionales, los parámetros de un prototipo de pitch dado, o de una trama dada, son cada uno de ellos cuantificado y transmitido individualmente por el codificador. Adicionalmente, se transmite un valor de diferencia para cada parámetro. El valor de diferencia específica es la diferencia entre el valor del parámetro para la trama o prototipo actual y el valor del parámetro para la trama o el prototipo anterior. Sin embargo, la cuantificación de los valores del parámetro y los valores de diferencia, requiere utilizar bits (y con ello ancho de banda). En un codificador de habla de baja tasa de bit, resulta ventajoso transmitir el menor número de bits posible que mantenga una calidad de voz satisfactoria. Por esta razón, en codificadores de habla de baja tasa de bit convencionales, solamente son cuantificados y transmitidos valores absolutos de parámetro. Sería deseable reducir el número de bits transmitidos sin reducir el valor informativo. En consecuencia, un esquema de cuantificación que cuantifica la diferencia entre la suma ponderada de los valores de parámetro para tramas previas y el valor de parámetro para la trama actual, ha sido descrito en una solicitud relacionada núm. EP-A-1 279 167, titulada PROCEDIMIENTO Y APARATO CUANTIFICAR PREDICTIVAMENTE HABLA CON VOZ, cedida a la cesionaria de la presente invención.

- 35 Los codificadores de habla experimentan borrado de trama, o pérdida de paquetes, debido a unas pobres condiciones del canal. Una solución utilizada en codificadores de habla convencionales fue hacer que el decodificador simplemente repita la trama anterior en el caso de que se reciba un borrado de trama. Se ha encontrado una mejora en el uso de un código de claves adaptativo, el cual ajusta dinámicamente la trama inmediatamente después de un borrado de trama. Un refinamiento adicional, el codificador de tasa variable incrementada (EVRC), ha sido estandarizado en el Estándar Interim de la Telecommunication Industry Association EIA/TIA IS-127. El codificador EVRC se basa en que una trama codificada predictivamente baja, altere en la memoria del codificador la trama que no fue recibida, y mejore con ello la calidad de la trama correctamente recibida.

- 45 Un problema del codificador EVRC, sin embargo, consiste en las discontinuidades que se pueden presentar entre un borrado de trama y una trama buena ajustada posteriormente. Por ejemplo, los pulsos de pitch pueden estar colocados demasiado cerca, o demasiado separados, en comparación con sus posiciones relativas en el caso de que no se produzca ningún borrado de trama. Tales discontinuidades pueden provocar un clic audible.

- 50 En general, los codificadores de habla que incluyen una baja previsibilidad (tales como los descritos en el párrafo anterior), se comportan mejor bajo condiciones de borrado de trama. Sin embargo, según se ha discutido, tales codificadores de habla requieren tasas de bit relativamente más altas. A la inversa, un codificador de habla altamente predictivo puede alcanzar una buena calidad de salida de habla sintetizada (en particular, para habla altamente periódica, como el habla con voz), pero se comporta peor bajo condiciones de borrado de trama. Sería deseable combinar las calidades de ambos tipos de codificador de habla. Resultaría además ventajoso proporcionar un procedimiento de alisamiento de discontinuidades entre borrados de trama y tramas buenas alteradas posteriores. De ese modo, existe una necesidad de un procedimiento de compensación de borrado de trama que mejore el comportamiento del codificador predictivo en caso de borrados de trama, y suavice las discontinuidades entre borrados de trama y tramas buenas posteriores.

- 60 Un enfoque seguido por la técnica anterior, ha sido descrito en "Modulación por código de pulso (PCM) de frecuencias de voz; G.711 Apéndice I (09/99); Un algoritmo de alta calidad y baja complejidad para ocultación de pérdida de paquete con G-711; ESTÁNDAR ITU-T EN VIGOR (I), INTERNATIONAL TELECOMMUNICATION UNION, GINEBRA, CH, núm. G.711 Apéndice I (09, 1 de Septiembre de 1999 (01-09-1999)), XP017400851.

**Sumario de la Invención**

La presente invención está dirigida a un procedimiento para alisar la discontinuidad entre una trama de habla perdida y una trama de habla actual en un dispositivo de telecomunicación, según se define mediante la reivindicación 1.

5 En otro aspecto de la invención, se define un aparato correspondiente mediante la reivindicación 6.

Un medio correspondiente de almacenamiento que comprende instrucciones que son ejecutables mediante un procesador para implementar un procedimiento de acuerdo con la invención, se define en la reivindicación 12.

**Breve Descripción de los Dibujos**

La Figura 1 es un diagrama de bloques de un sistema de teléfono inalámbrico;

10 la Figura 2 es un diagrama de bloques de un canal de comunicación que termina por ambos extremos mediante codificadores de habla;

la Figura 3 es un diagrama de bloques de un codificador de habla;

la Figura 4 es un diagrama de bloques de un descodificador de habla;

15 la Figura 5 es un diagrama de bloques de un codificador de habla que incluye porciones de codificador/ transmisor y de descodificador/ receptor;

la Figura 6 es un gráfico de amplitud de señal respecto al tiempo para un segmento de habla con voz;

la Figura 7 ilustra un primer esquema de procesamiento de borrado de trama que puede ser utilizado en la porción de descodificador/ receptor del codificador de habla de la Figura 5;

20 la Figura 8 ilustra un segundo esquema de procesamiento de borrado de trama adaptado a un codificador de habla de tasa variable, que puede ser utilizado en la porción de descodificador/ receptor del codificador de habla de la Figura 5;

la Figura 9 representa la amplitud de señal respecto al tiempo para varias formas de onda de residuo predictivo lineal (LP), para ilustrar un esquema de procesamiento de borrado de trama que puede ser utilizado para alisar una transición entre una trama dañada y una trama en buenas condiciones;

25 la Figura 10 representa la amplitud de señal respecto al tiempo para varias formas de onda de residuo de LP, para ilustrar los beneficios del esquema de procesamiento de borrado de trama representado en la Figura 9;

la Figura 11 representa la amplitud de señal respecto al tiempo para varias formas de onda, para ilustrar una técnica de codificación de prototipo de período de pitch o de interpolación de forma de onda;

la Figura 12 es un diagrama de bloques de un procesador acoplado a un medio de almacenamiento.

**Descripción Detallada de las Realizaciones Preferidas**

35 Las realizaciones ejemplares descritas en lo que sigue, se basan en un sistema de comunicación de telefonía inalámbrica configurado para el empleo de una interfaz sobre-el-aire de CDMA. No obstante, los expertos en la materia comprenderán que un procedimiento y un aparato para codificar predictivamente habla con voz que materializa las características de la presente invención, pueden estar basados en cualquiera de diversos sistemas de comunicación que emplean una amplia gama de tecnologías conocidas por los expertos en la materia.

40 Según se ha ilustrado en la Figura 1, un sistema de teléfono inalámbrico de CDMA incluye por lo general una pluralidad de unidades móviles 10 de abonado, una pluralidad de estaciones de base 12, controladores 14 de estación de base (BSCs), y un centro 16 de conmutación móvil (MSC). El MSC 16 está configurado para hacer de interfaz con una red convencional conmutada de telefonía pública (PSTN) 18. El MSC 16 está también configurado para servir de interfaz con los BSCs 14. Los BSCs 14 están acoplados a las estaciones de base 12 por medio de líneas de retorno. Las líneas de retorno pueden estar configuradas para soportar cualquiera de diversas interfaces conocidas incluyendo, por ejemplo, E1/T1, ATM, IP, PPP, Relé de Trama, HDSL, ADSL, o xDSL. Se comprende que pueden existir más de dos BSCs 14 en el sistema. Cada estación de base 12 incluye ventajosamente al menos un sector (no representado), comprendiendo cada sector una antena omnidireccional o una antena que apunta en una dirección particular radialmente hacia fuera de la estación de base 12. Alternativamente, cada sector puede comprender dos antenas para una diversidad de recepción. Cada estación de base 12 puede estar diseñada ventajosamente para soportar una pluralidad de asignaciones de frecuencia. La intersección de un sector y una asignación de frecuencia, puede ser mencionada como canal de CDMA. Las estaciones de base 12 pueden ser conocidas también como subsistemas transceptores 12 de estación de base (BTSs). Alternativamente, la "estación de base" puede ser utilizada en la industria para referirse en conjunto a un BSC 14 y a uno o más BTSs 12. Los

BTSs 12 pueden ser indicados también como "sitios de célula" 12. Alternativamente, los sectores individuales de un BTS 12 dado pueden ser mencionados como sitios de célula. Las unidades móviles 10 de abonado son típicamente teléfonos celulares o PCS 10. El sistema está configurado ventajosamente para su uso de acuerdo con el estándar IS-95.

5 Durante la operación típica de un sistema de telefonía celular, las estaciones de base 12 reciben conjuntos de señales de enlace reverso desde conjuntos de unidades móviles 10. Las unidades móviles 10 están llevando a cabo llamadas de teléfono u otras comunicaciones. Cada señal de enlace reverso recibida por una estación de base 12 dada, es procesada en el interior de la estación de base 12. Los datos resultantes son enviados a los BSCs 14. Los BSCs 14 proporcionan funcionalidad de asignación de recurso de llamada y gestión de movilidad, incluyendo la orquestación de transferencias suaves entre estaciones de base 12. Los BSCs 14 enrutan también los datos recibidos hasta el MSC 16, el cual proporciona servicios de enrutamiento adicionales para la interfaz con la PSTN 18. De forma similar, la PSTN 18 sirve de interfaz con el MSC 16, y el MSC 16 sirve de interfaz con los BSCs 14, los cuales controlan a su vez las estaciones de base 12 para transmitir conjuntos de señales de enlace directo a conjuntos de unidades móviles 10. Los expertos en la materia podrán comprender que las unidades 10 de abonado pueden ser unidades fijas en realizaciones alternativas.

En la Figura 2, un primer codificador 100 recibe muestras  $s(n)$  de habla digitalizadas y codifica las muestras  $s(n)$  para su transmisión por un medio de transmisión 102, o un canal de comunicación 102, hasta un primer descodificador 104. El descodificador 104 descodifica las muestras de habla codificadas, y sintetiza una señal  $S_{SYNTH}(n)$ . Para la transmisión en la dirección opuesta, un segundo codificador 106 codifica muestras  $s(n)$  de habla digitalizadas, las cuales son transmitidas por un canal de comunicación 108. Un segundo descodificador 110 recibe y descodifica las muestras de habla codificadas, generando una señal de habla de salida sintetizada  $S_{SYNTH}(n)$ .

Las muestras  $s(n)$  de habla representan señales de habla que han sido digitalizadas y cuantificadas de acuerdo con uno cualquiera de diversos procedimientos conocidos en el estado de la técnica incluyendo, por ejemplo, modulación de código de pulso (PCM), ley- $\mu$  compandida, o ley-A. Según se conoce en el estado de la técnica, las muestras  $s(n)$  de habla son organizadas en tramas de datos de entrada, en las que cada trama comprende un número predeterminado de muestras  $s(n)$  de habla digitalizadas. En un ejemplo de realización, se emplea una tasa de muestreo de 8 KHz, comprendiendo 160 muestras de trama de 20 ms cada una. En las realizaciones que se describen en lo que sigue, la tasa de transmisión de datos puede ser cambiada ventajosamente sobre una base de trama-por-trama, desde una tasa completa hasta (media tasa a cuarto de tasa a octavo de tasa). Variar la tasa de transmisión de datos resulta ventajoso debido a que se pueden emplear selectivamente tasas de bit más bajas para tramas que contengan relativamente menos información de habla. Como comprenderán los expertos en la materia, se pueden utilizar otras tasas de muestreo y/o tamaños de trama. También, en las realizaciones que se describen en lo que sigue, el modo de codificación de habla (o la codificación) puede ser variada sobre una base de trama-por-trama en respuesta a la información de habla o a la energía de la trama.

35 El primer codificador 100 y el segundo codificador 110 comprenden en conjunto un primer codificador de habla (codificador/ descodificador), o códec de habla. El códec de habla podría ser utilizado en cualquier dispositivo de comunicación para transmitir señales de habla, incluyendo, por ejemplo, las unidades de abonado, BTSs o BSCs que se han descrito en lo que antecede con referencia a la Figura 1. De manera similar, el segundo codificador 106 y el primer descodificador 104 comprenden en conjunto un segundo codificador de habla. Los expertos en la materia podrán comprender que los codificadores de habla pueden ser implementados con un procesador de señal digital (DSP), un circuito integrado de aplicación específica (ASIC), lógica de puerta discreta, firmware, o cualquier módulo de software programable convencional y un microprocesador. El módulo de software podría residir en memoria RAM, memoria flash, registros, o cualquier otra forma de medio de almacenamiento conocido en el estado de la técnica. Alternativamente, cualquier procesador, controlador o máquina de estado convencional, podría ser sustituido por el microprocesador. Ejemplos de ASICs diseñados específicamente para codificar habla, han sido descritos en la Patente U.S. núm. 5.727.123 y en la Patente U.S. núm. 5.784.432.

En la Figura 3, un codificador 200 que puede ser utilizado en un codificador de habla, incluye un módulo 202 de decisión de modo, un módulo 204 de decisión de pitch, un módulo 206 de análisis de LP, un filtro 208 de análisis de LP, un módulo 210 de cuantificación de LP, y un módulo 212 de cuantificación de residuo. Las tramas  $s(n)$  de habla entrantes son proporcionadas al módulo 202 de decisión de modo, al módulo 204 de estimación de pitch, al módulo 206 de análisis de LP, y al filtro 208 de análisis de LP. El módulo 202 de decisión de modo genera un índice de modo  $I_M$  y un modo M basado en la periodicidad, la energía, la relación señal-ruido (SNR), o la tasa de cruce por cero, entre otras características, de cada trama  $s(n)$  de habla entrante. Se han descrito varios modos de clasificación de tramas de habla conforme a la periodicidad en la Patente U.S. núm. 5.911.128, la cual ha sido cedida a la cesionaria de la presente invención. Tales procedimientos están también incorporados en los Estándares Interim TIA/EIA IS-127 y TIA/EIA IS-733 de la Telecommunication Industry Association. Un ejemplo de esquema de decisión de modo ha sido descrito también en la Patente U.S. núm. 6.691.084 mencionada anteriormente.

El módulo 204 de estimación de pitch genera un índice de pitch  $I_p$  y un valor de retardo  $P_0$  con cada trama  $s(n)$  de habla entrante. El módulo 206 de análisis de LP realiza un análisis predictivo lineal sobre cada trama  $s(n)$  de habla entrante para generar un parámetro  $a$  de LP. El parámetro  $a$  de LP se suministra al módulo 210 de cuantificación de

LP. El módulo 210 de cuantificación de LP recibe también el modo M, llevando con ello a cabo el proceso de cuantificación de una manera dependiente del modo. El módulo 210 de cuantificación de LP produce un índice de LP  $I_{LP}$  y un parámetro de LP  $\hat{a}$  cuantificado. El filtro 208 de análisis de LP recibe el parámetro de LP  $\hat{a}$  cuantificado adicionalmente a la trama  $s(n)$  de habla entrante. El filtro 208 de análisis de LP genera una señal  $R[n]$  de residuo de LP, la cual representa el error entre las tramas  $s(n)$  de habla entrantes y el habla reconstruida en base a los parámetros  $\hat{a}$  pronosticados lineales cuantificados. El residuo  $R[n]$  de LP, el modo M y el parámetro  $\hat{a}$  cuantificado, son suministrados al módulo 212 de cuantificación de residuo. En base a estos valores, el módulo 212 de cuantificación de residuo produce un índice de residuo  $I_R$  y una señal de residuo  $\hat{R}[n]$ .

En la Figura 4, un descodificador 300 que puede ser utilizado en un codificador de habla, incluye un módulo 302 de descodificación de parámetro de LP, un módulo 304 de descodificación de residuo, un módulo 306 de descodificación de modo, y un filtro 308 de síntesis de LP. El módulo 306 de descodificación de modo recibe y descodifica un índice de modo  $I_M$ , generando un modo M a partir del mismo. El módulo 302 de descodificación de parámetro de LP recibe el modo M y un índice de LP  $I_{LP}$ . El módulo 302 de descodificación de parámetro de LP descodifica los valores recibidos para producir un parámetro de LP  $\hat{a}$  cuantificado. El módulo 304 de descodificación de residuo recibe un índice de residuo  $I_R$ , un índice de pitch  $I_P$ , y el índice de modo  $I_M$ . El módulo 304 de descodificación de residuo descodifica los valores recibidos para generar una señal de residuo  $\hat{R}[n]$  cuantificada. La señal de residuo  $\hat{R}[n]$  cuantificada y el parámetro de LP  $\hat{a}$  cuantificado, son suministrados al filtro 308 de síntesis de LP, el cual sintetiza una señal  $\hat{s}[n]$  de habla de salida descodificada a partir de los mismos.

La operación y la implementación de los diversos módulos del codificador 200 de la Figura 3 y del descodificador 300 de la Figura 4, son conocidos en el estado de la técnica y han sido descritos en la Patente U.S. núm. 5.414.796 y en L. B. Rabiner & R. W. Schafer, Procesamiento Digital de Señales de Habla, 396-453 (1978), mencionados en lo que antecede.

En una realización, un codificador de habla 400 multimodo comunica con un descodificador de habla 402 multimodo a través de un canal de comunicación, o medio de transmisión, 404. El canal de comunicación 404 es ventajosamente una interfaz de RF configurada de acuerdo con el estándar IS-95. Los expertos en la materia podrán comprender que el codificador 400 tiene un descodificador asociado (no representado). El codificador 400 y el descodificador asociado forman juntos un primer codificador de habla. Los expertos en la materia podrán entender también que el descodificador 402 tiene un codificador asociado (no representado). El descodificador 402 y su codificador asociado forman juntos un segundo codificador de habla. El primer y el segundo codificadores de habla pueden ser implementados ventajosamente como parte del primer y segundo DSPs, y pueden residir, por ejemplo, en una unidad de abonado y una estación de base de un PCS o sistema de teléfono celular, o en una unidad de abonado y una puerta de enlace de un sistema de satélite.

El codificador 400 incluye un calculador de parámetro 406, un módulo 408 de clasificación de modo, una pluralidad de modos de codificación 410, y un módulo 412 de formateo de paquete. El número de modos de codificación 410 se ha mostrado como  $n$ , lo que cualquier experto podrá entender que podrá significar cualquier número razonable de modos de codificación 410. Por simplicidad, solamente se han mostrado tres modos de codificación 410, indicando con una línea de puntos la existencia de otros modos de codificación 410. El descodificador 402 incluye un módulo 414 de desensamblador de paquete y de detector de pérdida de paquete, una pluralidad de modos de descodificación 416, un descodificador de borrado 418, y un filtro posterior, o sintetizador de habla, 420. El número de modos de descodificación 416 ha sido mostrado como  $n$ , lo que un experto puede entender que podrá significar cualquier número razonable de modos de descodificación 416. Por simplicidad, solamente se han mostrado tres modos de descodificación 416, indicando con una línea de puntos la existencia de otros modos de descodificación 416.

Una señal de habla,  $s(n)$ , se suministra al calculador de parámetro 406. La señal de habla es dividida en bloques de muestras denominados tramas. El valor  $n$  designa el número de trama. En una realización alternativa, se utiliza una señal de error residual de predicción lineal (LP) en lugar de la señal de habla. El residuo de LP es utilizado por los codificadores de habla, tal como por ejemplo, el codificador CELP. El cálculo del residuo de LP se realiza ventajosamente suministrando la señal de habla a un filtro de LP inverso (no mostrado). La función de transferencia del filtro de LP inverso,  $A(z)$ , se calcula de acuerdo con la siguiente ecuación:

$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_p z^{-p},$$

en la que los coeficientes  $a_1$  son grifos de filtro que tienen valores predefinidos elegidos de acuerdo con procedimientos conocidos, según se describe en la Patente U.S. núm. 5.414.796 y en la Patente U.S. núm. 6.456.964 que se han mencionado anteriormente. El número  $p$  indica el número de muestras que utiliza el filtro de LP inverso a efectos de predicción. En una realización particular,  $p$  se ha establecido en diez.

El calculador de parámetros 406 extrae varios parámetros basados en la trama actual. En una realización, estos parámetros incluyen al menos uno de los siguientes: coeficientes de filtro de codificación predictiva lineal (LPC), coeficientes de par espectral de línea (LSP), funciones de autocorrelación normalizadas (NACFs), retardo en bucle abierto, tasas de cruce por cero, energías de banda, y la señal residual de formante. El cálculo de los coeficientes de LPC, de los coeficientes de LSP, del retardo en bucle abierto, de las energías de banda, y de la señal residual de

formante, ha sido escrito con detalle en la Patente U.S. núm. 5.414.796 mencionada en lo que antecede. El cálculo de de las NACFs y de las tasas de cruce por cero, ha sido descrito con detalle en la Patente U.S. núm. 5.911.128 mencionada anteriormente.

5 El calculador de parámetro 406 está acoplado al módulo 408 de clasificación de modo. El calculador de parámetro 406 proporciona los parámetros al módulo 408 de clasificación de modo. El módulo 408 de clasificación de modo está acoplado de manera que conmuta dinámicamente entre los modos de codificación 410 sobre una base de trama-por-trama, con el fin de seleccionar el modo de codificación 410 más apropiado para la trama actual. El módulo 408 de clasificación de modo selecciona un modo de codificación 410 particular para la trama actual comparando los parámetros con valores de umbral y/o de techo predefinidos. En base al contenido de energía de la trama, el módulo 408 de clasificación de modo clasifica la trama como de no habla, o habla inactiva (por ejemplo, silencio, ruido de fondo, o pausas entre palabras), o de habla. En base a la periodicidad de la trama, el módulo 408 de clasificación de modo clasifica a continuación tramas de habla como un tipo particular de habla, por ejemplo, con voz, sin voz, o transitorio.

15 El habla con voz es el habla que presenta un grado de periodicidad relativamente alto. Un segmento de habla con voz ha sido mostrado en el gráfico de la Figura 6. Según se ha ilustrado, el período de pitch es un componente de una trama de habla que puede ser utilizado con ventaja para analizar y reconstruir los contenidos de la trama. El habla sin voz comprende típicamente sonidos consonantes. Las tramas de habla transitorias son típicamente transiciones entre habla con voz y sin voz. Las tramas que son clasificadas como de habla de no de voz o sin voz, son clasificadas como de habla transitoria. Los expertos en la materia podrán comprender que podría emplearse cualquier esquema de clasificación razonable.

20 La clasificación de las tramas de habla resulta ventajosa puesto que se pueden utilizar diferentes modos de codificación 410 para codificar diferentes tipos de habla, dando como resultado un uso más eficiente del ancho de banda en un canal compartido tal como el canal de comunicación 404. Por ejemplo, puesto que el habla con voz es periódica y con ello altamente predictiva, un modo de codificación 410 altamente predictivo, de baja tasa de bit, puede ser empleado para codificar habla con voz. Módulos de clasificación, tal como el módulo de clasificación 408, han sido descritos con detalle en la Patente U.S. núm. 6.691.084 y en la Patente U.S. núm. 6.640.209 mencionadas anteriormente, tituladas CODIFICADOR DE HABLA DE PREDICCIÓN LINEAL EN EL DOMINIO MULTIMODO MIXTO (MDLP) DE BUCLE CERRADO.

25 El módulo 408 de clasificación de modo selecciona un modo de codificación 410 para la trama actual en base a la clasificación de la trama. Los diversos modos de clasificación 410 están acoplados en paralelo. Uno o más de los modos de codificación 410 pueden ser operacionales en cualquier momento dado. No obstante, solamente un modo de codificación 410 opera ventajosamente en cualquier instante dado, y se elige de acuerdo con la clasificación de la trama actual.

30 Los diferentes modos de codificación 410 operan ventajosamente de acuerdo con diferentes tasas de bit de codificación, diferentes esquemas de codificación, o diferentes combinaciones de tasa de bit de codificación y esquema de codificación. Las diversas tasas de codificación utilizadas pueden ser de tasa completa, de media tasa, de cuarto de tasa, y/o de octavo de tasa. Los diversos esquemas de codificación utilizados pueden ser de codificación CELP, codificación de período de pitch prototipo (PPP) (o codificación por interpolación de forma de onda (WI)), y/o codificación por predicción lineal excitada de ruido (NELP). Así, por ejemplo, un modo de codificación 410 particular podría ser CELP de tasa completa, otro modo de codificación 410 podría ser CELP de media tasa, otro modo de codificación 410 podría ser PPP de cuarto de tasa, y otro modo de codificación 410 podría ser NELP.

35 De acuerdo con un modo de codificación 410 CELP, un modelo de tracto vocal predictivo lineal es excitado con una versión cuantificada de la señal residual de LP. Los parámetros cuantificados para la trama previa completa, son utilizados para reconstruir la trama actual. El modo de codificación 410 CELP proporciona así una reproducción de habla relativamente precisa, pero a costa de una tasa de bit de codificación relativamente alta. El modo de codificación 410 CELP puede ser utilizado ventajosamente para codificar tramas clasificadas como habla transitoria. Un ejemplo de codificador de habla CELP de tasa variable, ha sido descrito con detalle en la Patente U.S. núm. 5.414.796 mencionada en lo que antecede.

40 De acuerdo con un modo de codificación 410 NELP, se utiliza una señal de ruido filtrada, pseudo-aleatoria, para modelar la trama de habla. El modo de codificación 410 NELP es una técnica relativamente simple que alcanza una baja tasa de bit. El modo de codificación 412 NELP puede ser utilizado para codificar con ventaja tramas clasificadas como habla sin voz. Un ejemplo de modo de codificación NELP ha sido descrito con detalle en la Patente U.S. núm. 6.456.964, mencionada en lo que antecede.

45 De acuerdo con un modo de codificación 410 PPP, solamente se codifica un subconjunto de períodos de pitch dentro de cada trama. Los restantes períodos de la señal de habla son reconstruidos por interpolación entre estos períodos de prototipo. En una implementación de codificación PPP en el dominio del tiempo, se calcula un primer conjunto de parámetros que describen cómo modificar un período de prototipo previo que se aproxima el período de prototipo actual. Se selecciona uno o más vectores de código que, cuando se suman, aproximan la diferencia entre el período de prototipo actual y el período de prototipo previo modificado. Un segundo conjunto de parámetros



describe estos vectores de código seleccionados. En una implementación de codificación PPP en el campo de la frecuencia, se calcula un conjunto de parámetros que describen la amplitud y los espectros de fase del prototipo. Esto puede hacerse tanto en sentido absoluto como predictivamente. Un procedimiento para cuantificar predictivamente la amplitud y los espectros de base de un prototipo (o de una trama completa), ha sido descrito en la solicitud relacionada EP-A-1 279 167 mencionada anteriormente, titulada PROCEDIMIENTO Y APARATO PARA CUANTIFICAR PREDICTIVAMENTE HABLA CON VOZ. De acuerdo con cualquier implementación de codificación PPP, el descodificador sintetiza una señal de habla de salida mediante reconstrucción de un prototipo actual en base a un primer y un segundo conjuntos de parámetros. La señal de habla es interpolada a continuación sobre la región entre el período de prototipo reconstruido actual y un período de prototipo reconstruido previo. El prototipo es por tanto una porción de la trama actual que será interpolada linealmente con prototipos de las tramas previas, que estaban posicionados de forma similar dentro de la trama, con el fin de reconstruir la señal de habla o la señal residual de LP en el descodificador (es decir, se utiliza un período de prototipo anterior como pronosticador del período de prototipo actual). Un ejemplo de codificador de habla de PPP ha sido descrito con detalle en la Patente U.S. núm. 6.456.964 mencionada en lo que antecede.

La codificación del período de prototipo en vez de la trama de habla completa, reduce la tasa de bit de codificación requerida. Las tramas clasificadas como habla con voz, pueden ser codificadas ventajosamente con un modo de codificación 410 PPP. Según se ha ilustrado en la Figura 6, el habla con voz contiene componentes periódicos, que varían lentamente con el tiempo, que son aprovechados con ventaja por el modo de codificación 410 PPP. Con el aprovechamiento de la periodicidad del habla con voz, el modo de codificación 410 PPP está capacitado para conseguir una tasa de bit más baja que el modo de codificación 410 CELP.

El modo de codificación 410 seleccionado se acopla al módulo 412 de formateo de paquete. El modo de codificación 410 seleccionado codifica, o cuantifica, la trama actual y proporciona los parámetros de trama cuantificados al módulo 412 de formateo de paquete. El módulo 412 de formateo de paquete ensambla ventajosamente la información cuantificada en paquetes para su transmisión por el canal de comunicación 404. En una realización, el módulo 412 de formateo de paquete está configurado para proporcionar codificación de corrección de error y formatear el paquete de acuerdo con el estándar IS-95. El paquete se suministra a un transmisor (no representado), se convierte a formato analógico, se modula y se transmite por el canal de comunicación 404 hasta un receptor (que tampoco se ha representado), el cual recibe, desmodula, y digitaliza el paquete, y suministra el paquete al descodificador 402.

En el descodificador 402, el módulo 414 de desensamblador de paquete y de detector de pérdida de paquete, recibe el paquete desde el receptor. El módulo 414 de desensamblador de paquete y de detector de pérdida de paquete, está acoplado de modo que conmuta dinámicamente entre los modos de descodificación 416, sobre una base de paquete-por-paquete. El número de modos de descodificación 416 es el mismo que el número de modos de codificación 410, y como reconocerá un experto en la materia, cada modo de codificación 410 numerado está asociado a un modo de descodificación 416 respectivo numerado de forma similar, configurado para emplear la misma tasa de bit de codificación y el mismo esquema de codificación.

Si el módulo 414 de desensamblador de paquete y de detector de pérdida de paquete detecta el paquete, el paquete es desensamblado y suministrado al modo de descodificación 416 pertinente. Si el módulo 414 de desensamblador de paquete y de detector de pérdida de paquete no detecta ningún paquete, se declara una pérdida de paquete y el descodificador de borrado 418 realiza ventajosamente el procesamiento de borrado de trama según se describe con detalle en lo que sigue.

La matriz paralela de modos de descodificación 416 y el descodificador de borrado 418, están acoplados al filtro posterior 420. El modo de descodificación 416 pertinente descodifica, o descuantifica, el paquete que proporciona la información al filtro posterior 420. El filtro posterior 420 reconstruye, o sintetiza, la trama de habla, presentando a la salida tramas de habla sintetizadas,  $s(n)$ . Ejemplos de modos de descodificación y de filtros posteriores, han sido descritos en la Patente U.S. núm. 5.414.796 y en la Patente U.S. núm. 6.456.964 que se han mencionado anteriormente.

En una realización, los parámetros cuantificados en sí mismos no son transmitidos. Por el contrario, se transmiten índices de código de claves que especifican direcciones en diversas tablas de búsqueda (LUTs) (no representadas) en el descodificador 402. El descodificador 402 recibe los índices de código de claves y busca en las diversas LUTs de códigos de clave respecto a los valores de parámetro apropiados. En consecuencia, los índices de código de claves para parámetros tales como, por ejemplo, retardo de pitch, ganancia de código de claves adaptativo, y SLP, pueden ser transmitidos, y tres LUTs de código de claves asociado son buscados por el descodificador 402.

De acuerdo con el modo de codificación 410 CELP, se transmiten los parámetros de retardo de pitch, amplitud, fase y LSP. Los índices de código de claves de SLP son transmitidos debido a que la señal de residuo de LP debe ser sintetizada en el descodificador 402. Adicionalmente, se transmite la diferencia entre el valor de retardo de pitch para la trama actual y el valor de retardo de pitch para la trama anterior.

De acuerdo con un modo de codificación PPP convencional, en el que la señal de habla debe ser sintetizada en el descodificador, solamente se transmiten los parámetros de retardo de pitch, amplitud, y fase. La tasa de bit más baja

empleada por las técnicas de codificación de habla PPP convencionales no permite la transmisión de ninguno de los valores de información de retardo absoluto de pitch ni diferencia de retardo relativo de pitch.

De acuerdo con una realización, tramas altamente periódicas tales como tramas de habla con voz, son transmitidas con un modo de codificación 410 PPP de baja-tasa-de-bit que cuantifica la diferencia entre el valor de retardo de pitch para la trama actual y el valor de retardo de pitch para la trama previa en cuanto a su transmisión, y no cuantifica el valor de retardo de pitch para la trama actual respecto a su transmisión. Puesto que las tramas con voz son de naturaleza altamente periódica, la transmisión del valor de diferencia en oposición al valor de retardo de pitch absoluto, permite que se alcance una tasa de bit de codificación más baja. En una realización, esta cuantificación se generaliza de tal modo que se calcula una suma ponderada de los valores de parámetro para tramas previas, en el que la suma de los pesos es uno, y la suma ponderada se resta del valor de parámetro para la trama actual. La diferencia es cuantificada a continuación. Esta técnica ha sido descrita con detalle en la solicitud relacionada EP-A-1 279 167 mencionada anteriormente, titulada PROCEDIMIENTO Y APARATO PARA CUANTIFICAR PREDICTIVAMENTE HABLA CON VOZ.

De acuerdo con una realización, un sistema de codificación de tasa variable, codifica diferentes tipos de habla según se determine mediante un procesador de control con diferentes codificadores, o modos de codificación, controlados por el procesador, o clasificador de modo. Los codificadores modifican la señal residual de trama actual (o como alternativa, la señal de habla) de acuerdo con un contorno de pitch según se especifique mediante el valor de retardo de pitch para la trama previa,  $L_{-1}$ , y el valor de retardo de pitch para la trama actual,  $L$ . Un procesador de control para los descodificadores, sigue el mismo contorno de pitch para reconstruir una contribución de código de claves adaptativo  $\{P(n)\}$ , a partir de una memoria de pitch para el residuo cuantificado o el habla respecto a la trama actual.

Si el valor previo de retardo de pitch,  $L_{-1}$ , se pierde, los descodificadores no pueden reconstruir el contorno de pitch correcto. Esto provoca que la contribución de código de claves adaptativo,  $\{P(n)\}$ , sea distorsionada. A su vez, el habla sintetizada sufrirá una severa degradación incluso aunque no se pierda ningún paquete para la trama actual. Como remedio, algunos codificadores convencionales emplean un esquema para codificar tanto  $L$  como la diferencia entre  $L$  y  $L_{-1}$ . Esta diferencia, o valor delta de pitch, puede ser indicada mediante  $\Delta$ , donde  $\Delta = L - L_{-1}$  sirve para el objetivo de recuperar  $L_{-1}$  si se ha perdido  $L_{-1}$  en la trama anterior.

La realización actualmente descrita puede ser utilizada con mayor ventaja en un sistema de codificación de tasa variable. Específicamente, un primer codificador (o modo de codificación), indicado mediante C, codifica el valor de retardo de pitch de trama actual,  $L$ , y el valor delta de retardo de pitch,  $\Delta$ , según se ha descrito anteriormente. Un segundo codificador (o modo de codificación), indicado mediante Q, codifica el valor delta de retardo de pitch,  $\Delta$ , pero no codifica necesariamente el valor de retardo de pitch,  $L$ . Esto permite que el segundo codificador, Q, utilice los bits adicionales para codificar otros parámetros o para salvar los bits totalmente (es decir, que funcione como codificador de baja tasa de bit). El primer codificador, C, puede ser ventajosamente un codificador utilizado para codificar habla relativamente no periódica tal como, por ejemplo, un codificador CELP de tasa completa. El segundo codificador, Q, puede ser ventajosamente un codificador utilizado para codificar habla altamente periódica (por ejemplo, habla con voz), tal como, por ejemplo, un codificador PPP de cuarto de tasa.

Según se ha ilustrado en el ejemplo de la Figura 7, si el paquete de la trama previa, la trama  $n-1$ , se ha perdido, la contribución de memoria de pitch,  $\{P_{-2}(n)\}$ , tras la descodificación de la trama recibida con anterioridad a la trama anterior, la trama  $n-2$ , se almacena en la memoria de codificador (no representada). El valor de retardo de pitch para la trama  $n-2$ ,  $L_{-2}$ , se almacena también en la memoria del codificador. Si la trama actual, la trama  $n$ , es codificada por el codificador C, la trama  $n$  puede ser denominada una trama C. El codificador C puede restablecer el valor de retardo de pitch anterior,  $L_{-1}$ , a partir del valor delta de pitch,  $\Delta$ , utilizando la ecuación  $L_{-1} = L - \Delta$ . Con ello, un contorno de pitch correcto puede ser reconstruido con los valores  $L_{-1}$  y  $L_{-2}$ . La contribución de código de claves adaptativo para la trama  $n-1$  puede ser reparada dado el contorno de pitch correcto, y se utiliza a continuación para generar contribución de código de claves adaptativo para la trama  $n$ . Los expertos en la materia podrán comprender que un esquema de ese tipo se utiliza en algunos codificadores convencionales tal como el codificador EVRC.

De acuerdo con una realización, la realización de borrado de trama en un sistema de codificación de habla de tasa variable, utilizando los dos tipos de codificadores descritos en lo que antecede (codificador C y codificador Q), se incrementa según se describe en lo que sigue. Según se ha ilustrado en el ejemplo de la Figura 8, un sistema de codificación de tasa variable puede estar diseñado de modo que haga uso tanto del codificador C como del codificador Q. La trama actual, la trama  $n$ , es una trama C y su paquete no está perdido. La trama anterior, la trama  $n-1$ , es una trama Q. El paquete para la trama que precede a la trama Q (es decir, el paquete para la trama  $n-2$ ) fue perdido.

En el procesamiento de borrado de trama para la trama  $n-2$ , la contribución de memoria de pitch,  $\{P_{-3}(n)\}$ , tras la descodificación de la trama  $n-3$ , se almacena en la memoria de codificador (no mostrada). El valor de retardo de pitch para la trama  $n-3$ ,  $L_{-3}$ , se almacena también en la memoria de codificador. El valor de retardo de pitch para la trama  $n-1$ ,  $L_{-1}$ , puede ser recuperado utilizando el valor delta de retardo de pitch,  $\Delta$ , (el cual es igual a  $L - L_{-1}$ ), en el paquete de trama C, de acuerdo con la ecuación  $L_{-1} = L - \Delta$ . La trama  $n-1$  es una trama Q con un valor delta de retarde

pitch codificado asociado que le pertenece,  $\Delta_1$ , igual al  $L_1 - L_2$ . De ahí que, el valor de retardo de pitch para la trama de borrado, la trama  $n-2$ ,  $L_2$ , pueda ser recuperado de acuerdo con la ecuación  $L_2 = L_1 - \Delta_1$ . Con los valores correctos de retardo de pitch para la trama  $n-2$  y para la trama  $n-1$ , los contornos de pitch para estas tramas pueden ser reconstruidos ventajosamente, y la contribución del código de claves adaptativo puede ser reparada en consecuencia. De ahí que, la trama  $C$  tendrá la memoria de pitch mejorada requerida para calcular la contribución de clave de código adaptativo para su señal residual de LP cuantificada (o señal de habla). Este procedimiento puede ser ampliado fácilmente para permitir la existencia de múltiples tramas  $Q$  entre la trama de borrado y la trama  $C$  como podrán apreciar los expertos en la materia.

Según se ha mostrado gráficamente en la Figura 9, cuando se borra una trama, el descodificador de borrado (por ejemplo, el elemento 418 de la Figura 5) reconstruye el residuo de LP cuantificado (o señal de habla) sin la información exacta de la trama. Si el contorno de pitch y la memoria de pitch de la trama borrada fueron restauradas de acuerdo con el procedimiento descrito en lo que antecede para la reconstrucción del residuo de LP cuantificado (o señal de habla) de la trama actual, el residuo de LP cuantificado resultante (o señal de habla) podría ser diferente del que tenía la memoria de pitch dañada que estaba usando. Un cambio de ese tipo en la memoria de pitch del codificador dará como resultado una discontinuidad en los residuos cuantificados (o señales de habla) a través de las tramas. De ahí que, un sonido de transición, o clic, se oye con frecuencia en los codificadores de habla convencionales tales como el codificador EVRC.

De acuerdo con la invención, los prototipos de período de pitch son extraídos de la memoria de pitch dañada con anterioridad a su reparación. El residuo de LP (o señal de habla) para la trama actual se extrae también de acuerdo con un proceso de descuantificación normal. El residuo de LP cuantificado (o señal de habla) para la trama actual se reconstruye entonces de acuerdo con un procedimiento de interpolación de forma de onda (WI). En una realización particular, el procedimiento WI opera de acuerdo con el modo de codificación de PPP descrito en lo que antecede. Este procedimiento sirve ventajosamente para alisar la discontinuidad descrita en lo que antecede, e incrementa además la ejecución de borrado de trama del codificador de habla. Un esquema de WI de ese tipo puede ser utilizado siempre que la memoria de pitch sea reparada debido al procesamiento de borrado con independencia de las técnicas utilizadas para llevar a cabo la reparación (incluyendo, aunque sin limitación, por ejemplo, las técnicas descritas anteriormente en la presente memoria).

Los gráficos de la Figura 10 ilustran la diferencia de apariencia entre una señal residual de LP que ha sido ajustada de acuerdo con técnicas convencionales, produciendo un clic audible, y una señal residual de LP que ha sido alisada posteriormente de acuerdo con el esquema de alisamiento de WI descrito anteriormente. Los gráficos de la Figura 11 ilustran los principios de una técnica de codificación de PPP o de WI.

Los expertos en la materia podrán comprender que los datos, instrucciones, comandos, información, señales, bits, símbolos y chips que puedan haber sido referenciados a través de la descripción que antecede, están representados ventajosamente por tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticas, campos o partículas ópticas, o cualquier combinación de los mismos. Los expertos podrán apreciar además que los diversos bloques lógicos, módulos, circuitos y etapas de algoritmos ilustrativos que se han descrito en relación con las realizaciones divulgadas en la presente memoria, pueden ser implementados como hardware electrónico, software de ordenador, o combinaciones de ambos. Los diversos componentes ilustrativos, bloques, módulos, circuitos y etapas, han sido generalmente descritos en términos de su funcionalidad. Que la funcionalidad se implemente como hardware o software, depende de la aplicación particular y de las limitaciones de diseño impuestas por el sistema en su conjunto. Los expertos reconocen la intercambiabilidad de hardware y software bajo estas circunstancias, y cómo implementar mejor la funcionalidad descrita para cada aplicación particular. Como ejemplos, los diversos bloques lógicos, módulos, circuitos y etapas de algoritmos ilustrativos, que se han descrito en relación con las realizaciones divulgadas en la presente memoria, pueden ser implementados o llevados a cabo con un procesador de señal digital (DSP), un circuito integrado de aplicación específica (ASIC), una batería de puertas programables en campo (FPGA), u otro dispositivo lógico programable, lógica de transistor o puerta discreta, componentes de hardware discretos tales como, por ejemplo, registros y FIFO, un procesador que ejecuta un conjunto de instrucciones de firmware, cualquier módulo de software programable convencional y un procesador, o cualquier combinación de los mismos diseñada para llevar a cabo las funciones descritas en la presente memoria. El procesador puede ser ventajosamente un microprocesador, pero como alternativa, el procesador puede ser cualquier procesador, controlador, microcontrolador o máquina de estado convencionales. El módulo de software puede residir en una memoria RAM, memoria flash, memoria ROM, memoria EPROM, memoria EEPROM, registros, disco duro, un disco extraíble, un CD-ROM o cualquier otra forma de medio de almacenamiento conocido en el estado de la técnica. Según se ha ilustrado en la Figura 12, un procesador ejemplar 500 está acoplado ventajosamente a un medio de almacenamiento 502 a efectos de leer información desde, y escribir información en, el medio de almacenamiento 502. Como alternativa, el medio de almacenamiento 502 puede ser integral con el procesador 500. El procesador 500 y el medio de almacenamiento 502 pueden residir en un ASIC (no representado). El ASIC puede residir en un teléfono (no representado). Como alternativa, el procesador 500 y el medio de almacenamiento 502 pueden residir en un teléfono. El procesador 500 puede estar implementado como una combinación de un DSP y un microprocesador, o como dos microprocesadores junto con un núcleo DSP, etc.

**REIVINDICACIONES**

- 1.- Un procedimiento para suavizar la discontinuidad entre una trama de habla perdida y una trama de habla actual en un dispositivo de comunicación, que comprende:
- 5                    construir una señal de habla para la trama de habla actual;
- extraer prototipos de período de pitch a partir de una memoria de pitch dañada con anterioridad a la reconstrucción de la trama de habla perdida, y
- reconstruir la señal de habla para la trama de habla actual de acuerdo con la una interpolación de forma de onda.
- 10                  2.- El procedimiento de acuerdo con la reivindicación 1, en el que la trama de habla perdida comprende una trama de 20 ms, comprendiendo con preferencia 160 muestras.
- 3.- El procedimiento de acuerdo con la reivindicación 1, en el que la señal de habla comprende una señal residual de Predicción Lineal.
- 4.- El procedimiento de acuerdo con la reivindicación 1, en el que la señal de habla está codificada con una de una pluralidad de tasas de bit variables.
- 15                  5.- El procedimiento de acuerdo con la reivindicación 1, en el que, de acuerdo con un modo de codificación de período de pitch prototipo, solamente se codifica un subconjunto de los períodos dentro de cada trama, y la interpolación de forma de onda de los prototipos de período de pitch reconstruye los restantes períodos de la señal de habla.
- 20                  6.- Un aparato para alisar una discontinuidad entre una trama de habla perdida y una trama de habla actual en un dispositivo de comunicación, que comprende:
- medios para construir una señal de habla para la trama de habla actual;
- medios para extraer prototipos de período de pitch desde una memoria de pitch dañada con anterioridad a la reconstrucción de la trama de habla perdida, y
- 25                  medios para reconstruir la señal de habla para la trama de habla actual de acuerdo con una interpolación de formas de onda.
- 7.- El aparato de acuerdo con la reivindicación 6, en el que la señal de habla comprende una señal residual de Predicción Lineal.
- 8.- El aparato de acuerdo con la reivindicación 6, en el que la señal de habla se codifica con una de una pluralidad de tasas de bit variables.
- 30                  9.- El aparato de acuerdo con la reivindicación 6, en el que el dispositivo de comunicación comprende un teléfono celular.
- 10.- El aparato de acuerdo con la reivindicación 6, que comprende además medios para codificar solamente un subconjunto de los períodos dentro de cada trama de acuerdo con un procedimiento de codificación de período de pitch prototipo, y medios para reconstruir los períodos restantes de la señal de habla mediante interpolación de forma de onda de los períodos de pitch.
- 35                  11.- El aparato de la reivindicación 6, en el que los medios para la construcción de la señal de habla, los medios para la extracción de los prototipos de período de pitch y los medios para la reconstrucción de la señal de habla, comprenden un procesador.
- 40                  12.- Un medio de almacenamiento que comprende instrucciones que son ejecutables por medio de un procesador para implementar el procedimiento de cualquiera de las reivindicaciones 1 a 5.

45

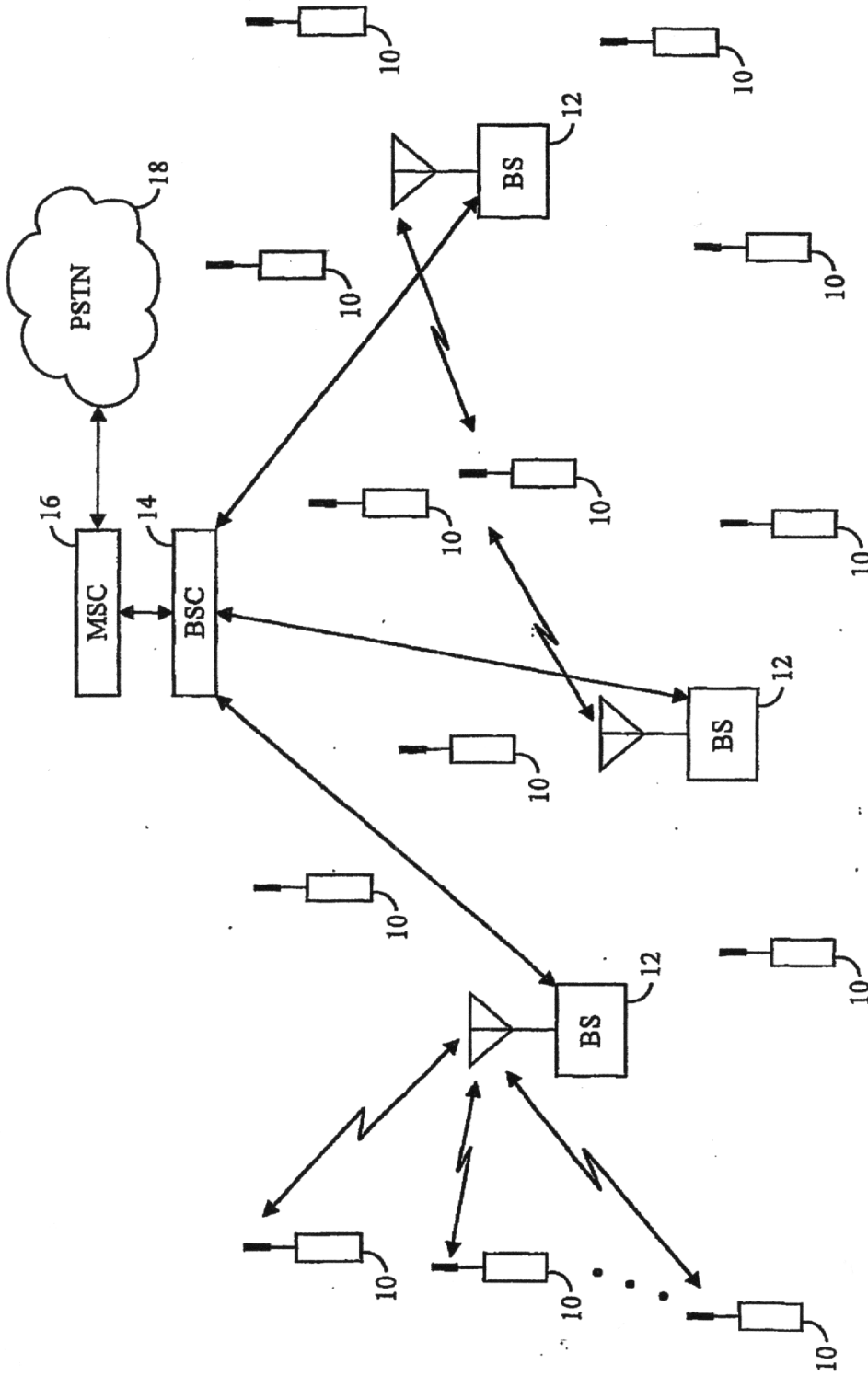


FIG. 1

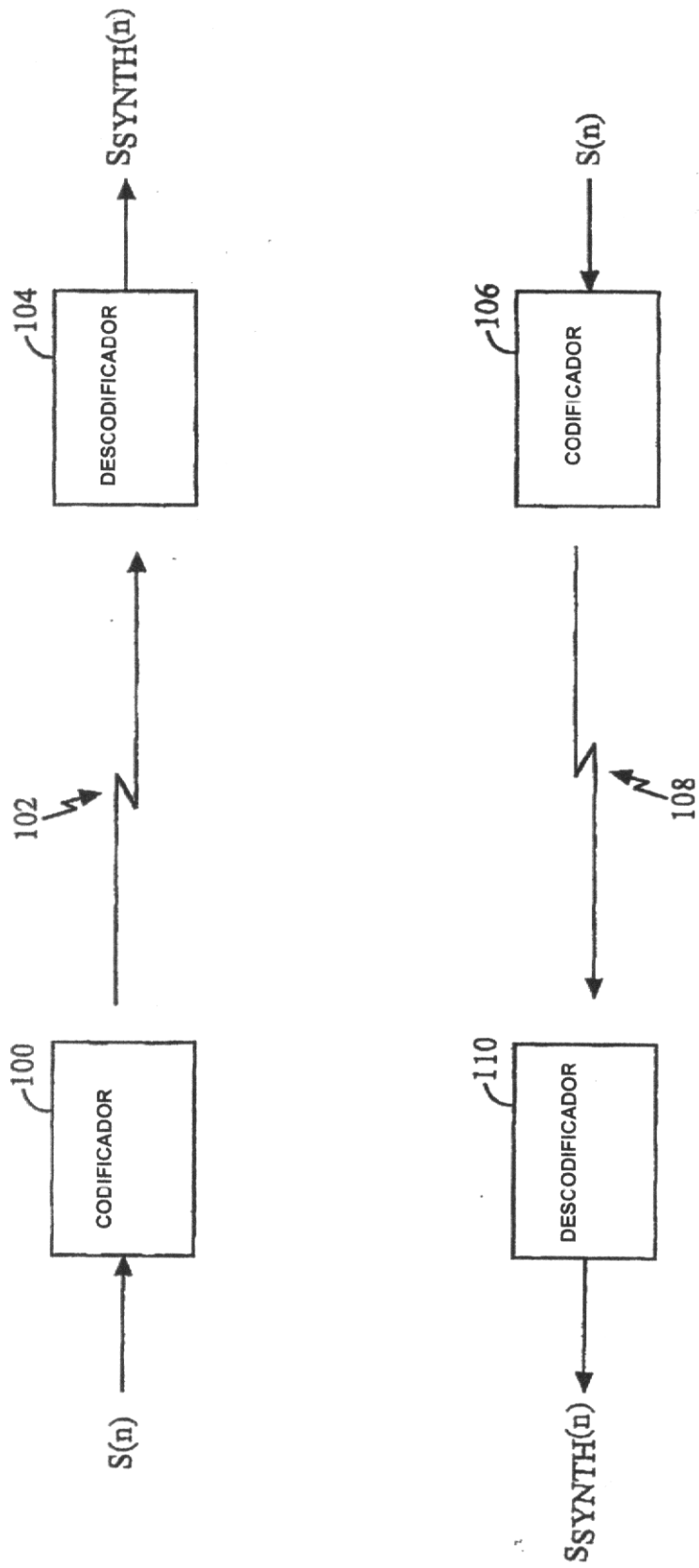


FIG. 2

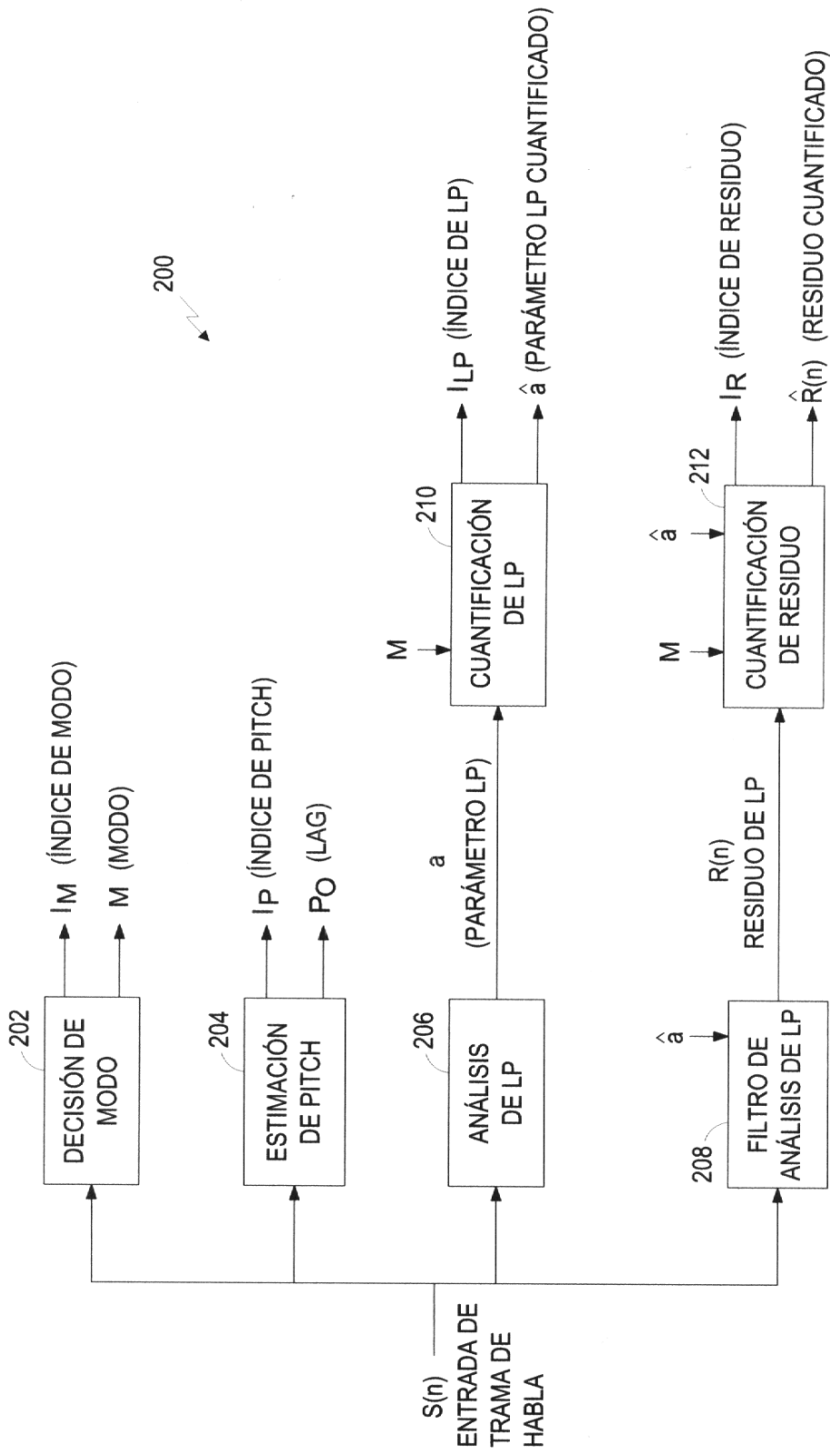


FIG. 3

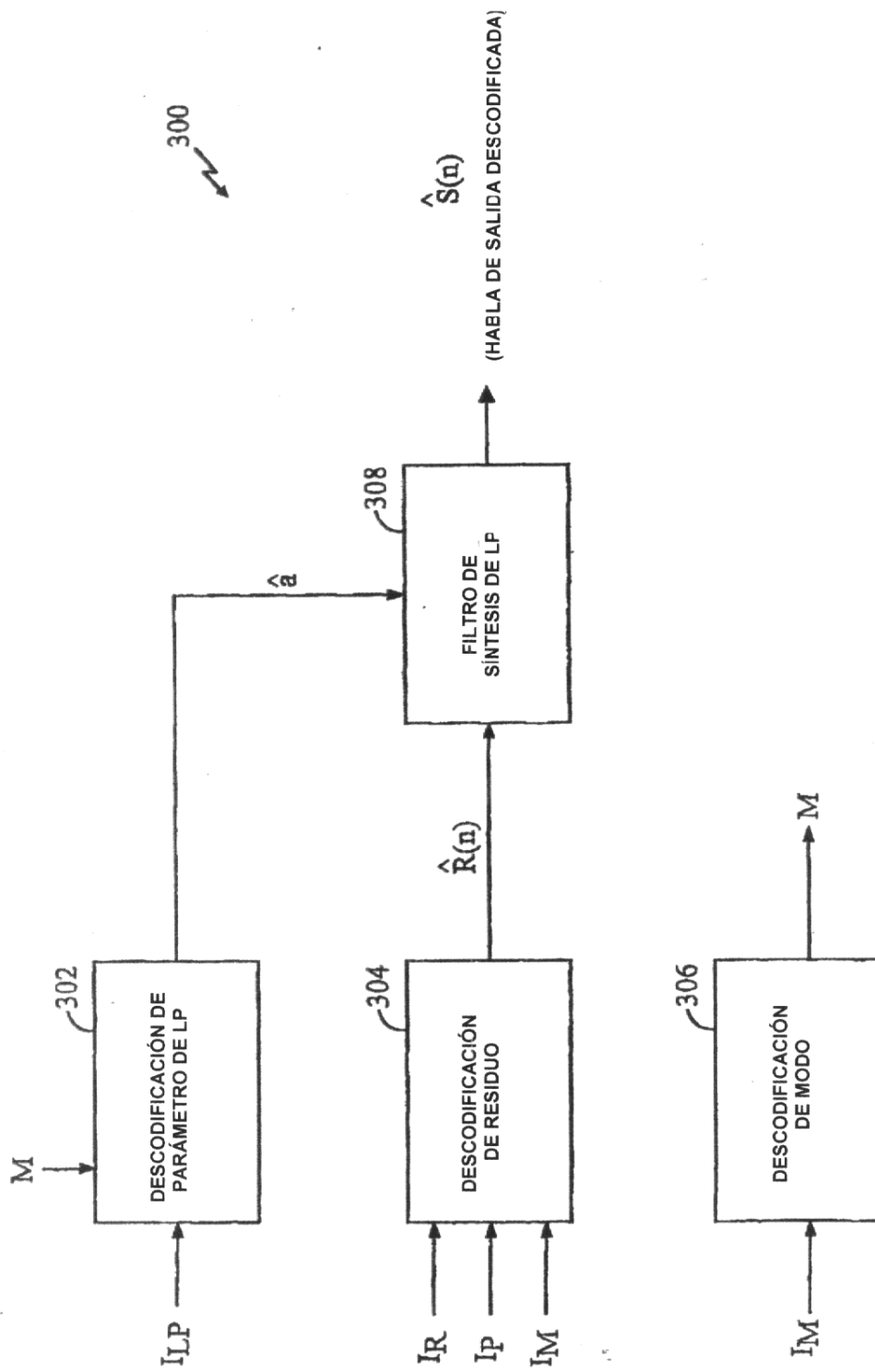


FIG. 4



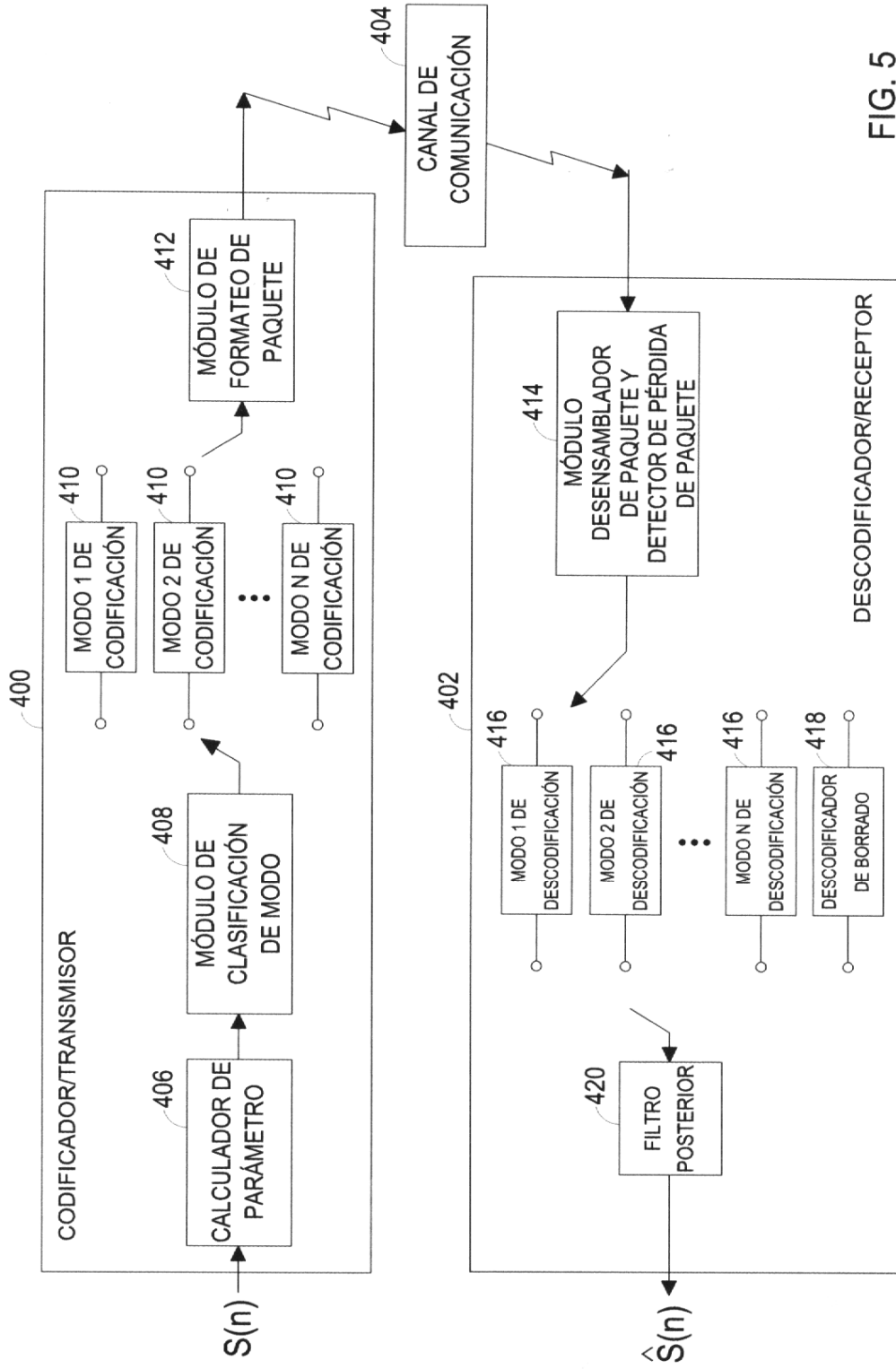


FIG. 5

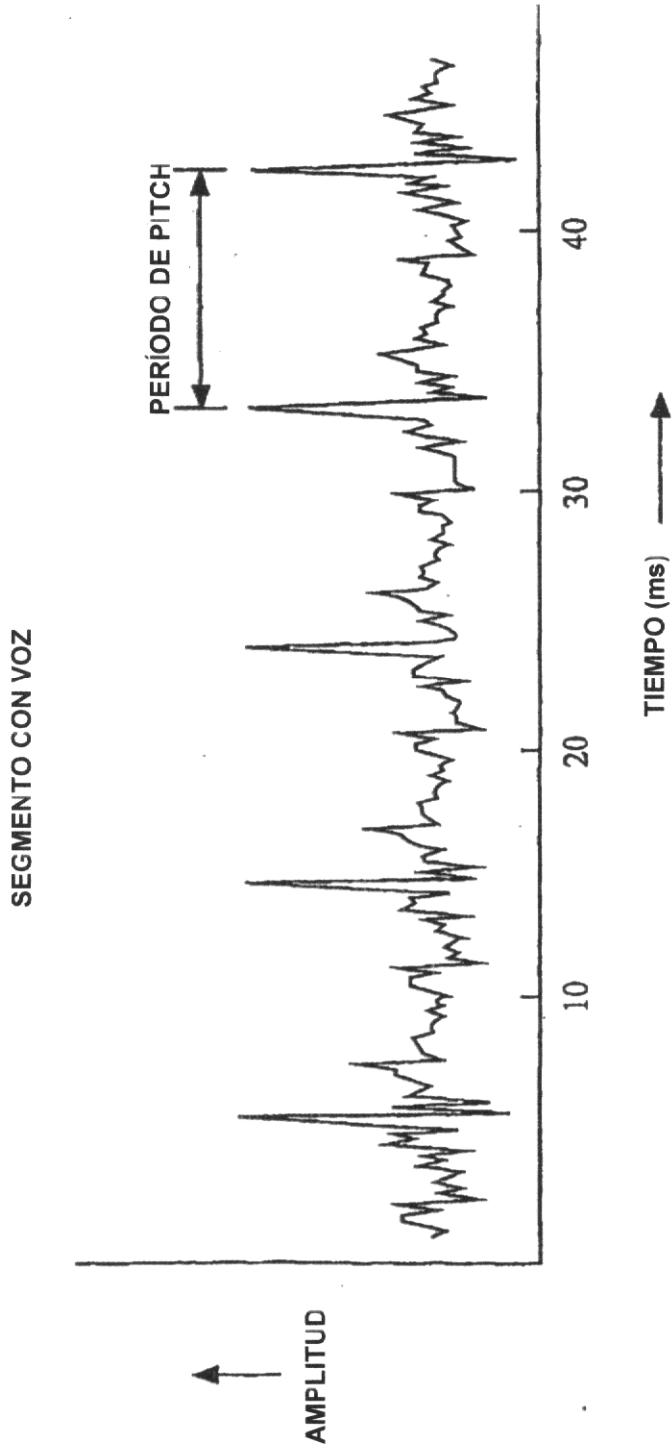


FIG. 6

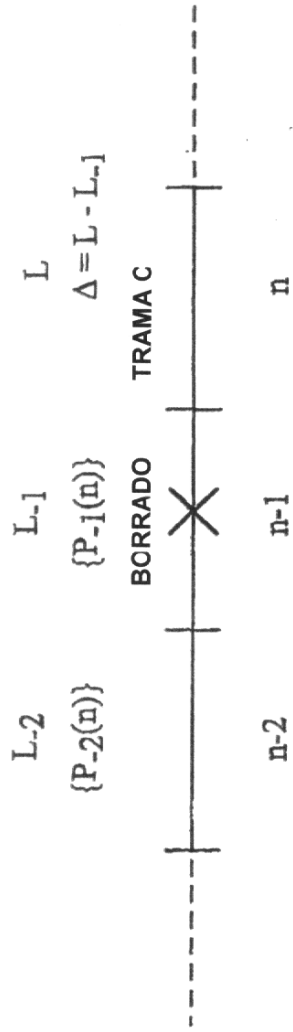


FIG. 7

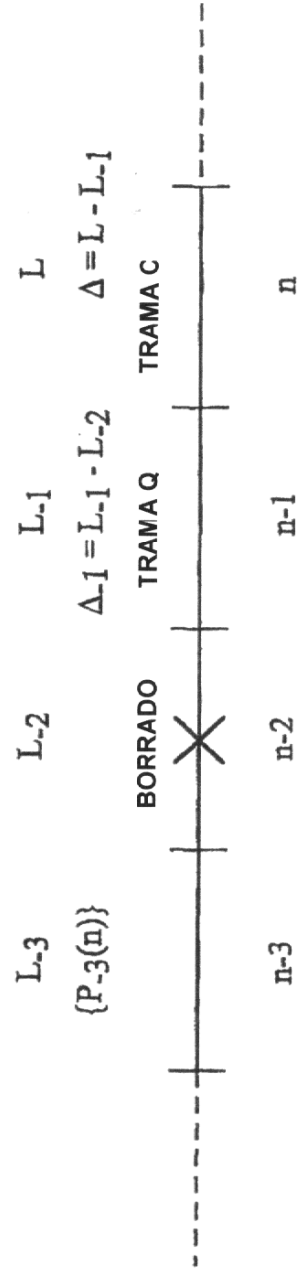


FIG. 8

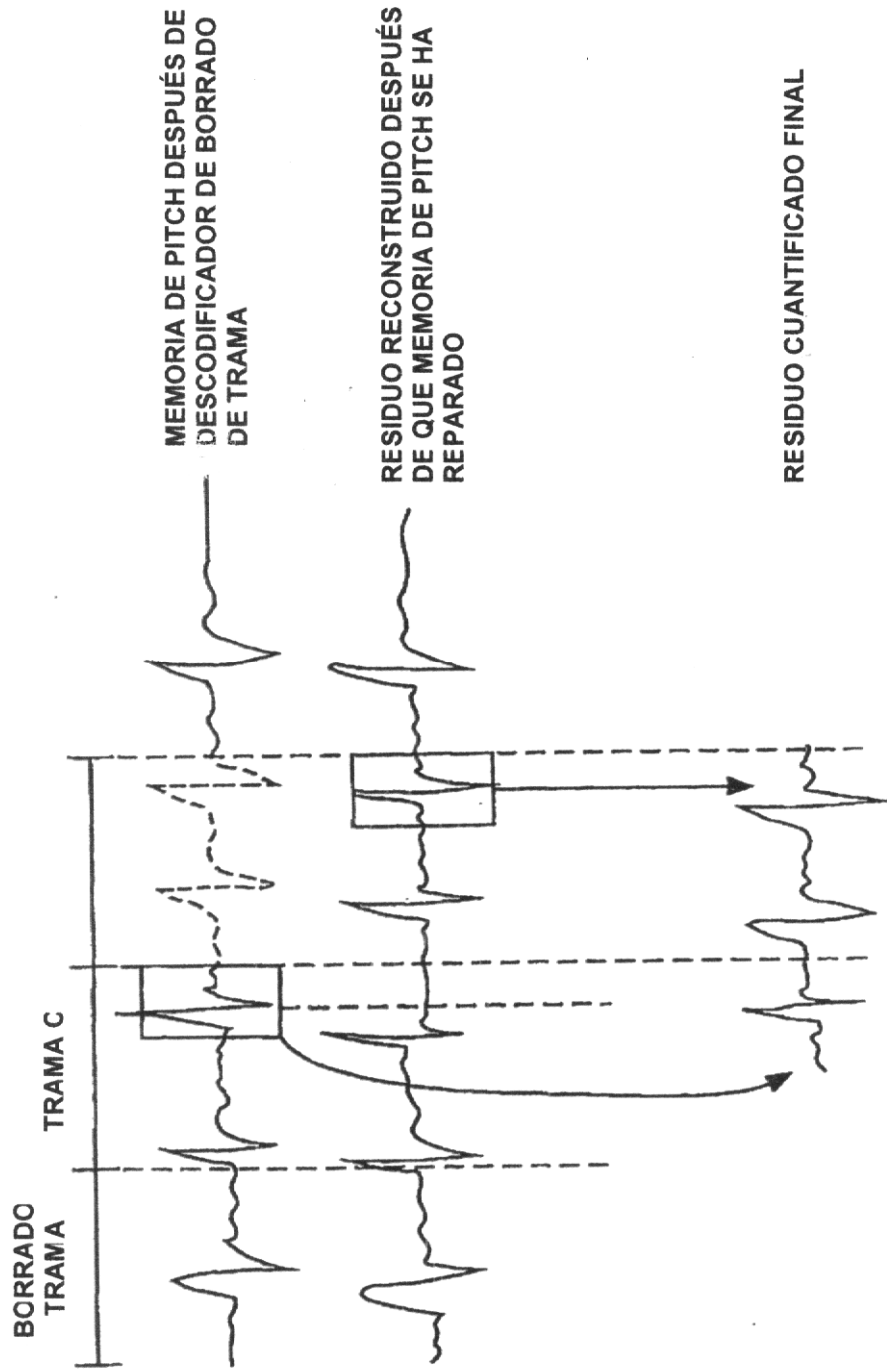


FIG. 9

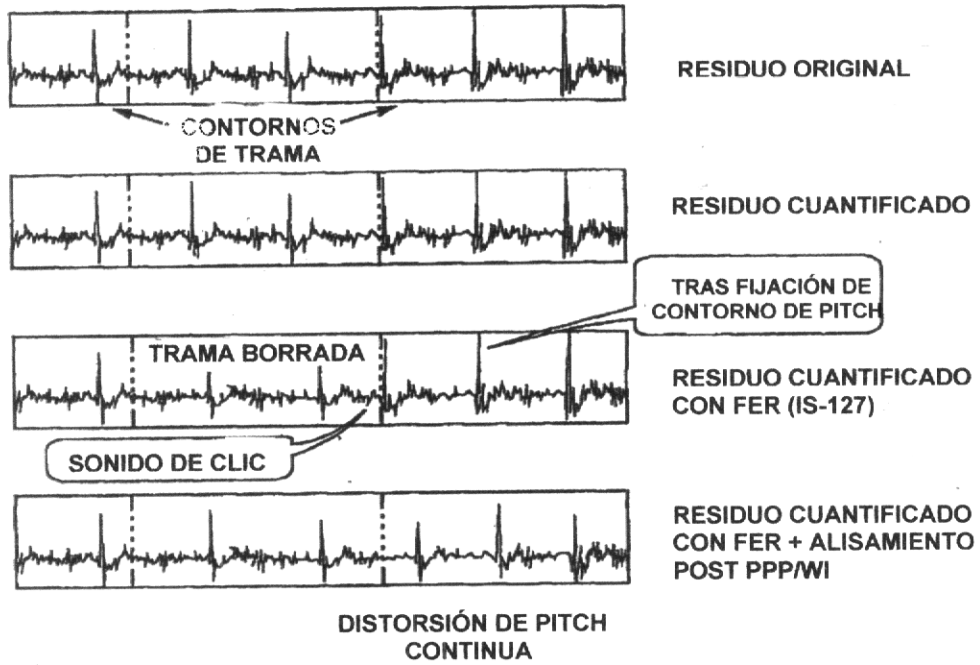


FIG. 10

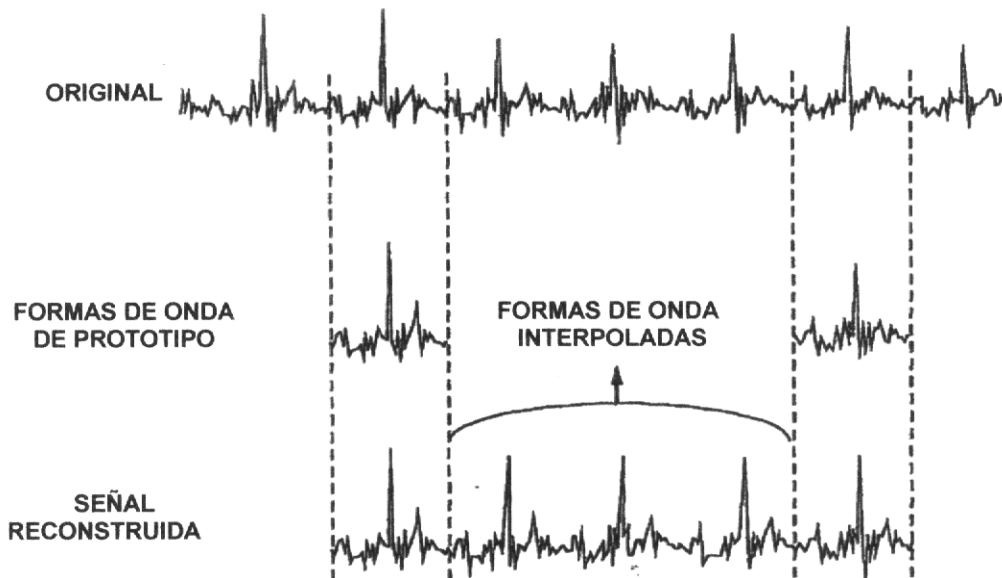


FIG. 11

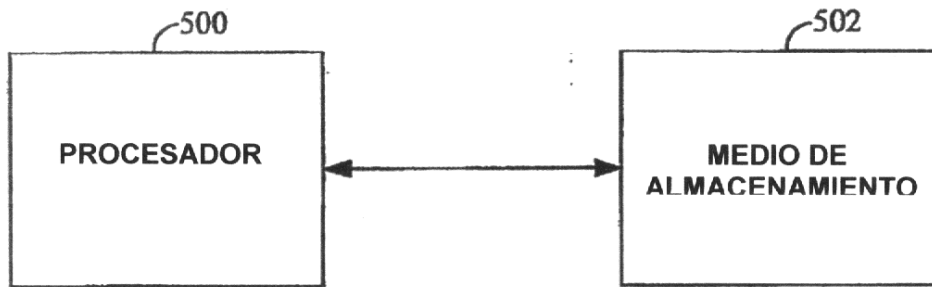


FIG. 12