

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 539 858**

51 Int. Cl.:

H04W 28/06 (2009.01)

H04L 1/00 (2006.01)

G10L 19/005 (2013.01)

H04L 12/70 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **24.06.2011** **E 11727450 (6)**

97 Fecha y número de publicación de la concesión europea: **25.03.2015** **EP 2647241**

54 Título: **Agregación de trama adaptativa de señal de fuente**

30 Prioridad:

03.12.2010 US 419337 P

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

06.07.2015

73 Titular/es:

**TELEFONAKTIEBOLAGET L M ERICSSON
(PUBL) (100.0%)
164 83 Stockholm, SE**

72 Inventor/es:

BRUHN, STEFAN

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 539 858 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

Agregación de trama adaptativa de señal de fuente

5 Campo técnico

La presente solicitud se refiere a un método de procesamiento de una señal para transmisión; a un aparato para procesamiento de una señal para transmisión; y a un medio legible con ordenador.

10 Antecedentes

10 La familia de códec Multi-Tasa Adaptativa (AMR) se utiliza como códec obligatorio en ambos sistemas GSM y WCDMA. La AMR está disponible en ambas formas de banda estrecha (AMR-NB) y de banda ancha (AMR-WB) Los estándares comprenden un número de especificaciones técnicas de múltiples componentes y funciones, por ejemplo códec de Habla (AMR-NB, AMR-WB), Detector de Actividad de Voz (VAD), sistema de Transmisión Discontinua (DTX), Ruido de Confort (CN), Adaptación de Enlace (LA), etc. Todas esas funciones están definidas y descritas en 15 las especificaciones serie 3GPP TS 26. Además, se ha proporcionado una descripción del códec de AMR-NB en "El Codificador de Habla Multi-Tasa Adaptativa", IEEE Taller de Codificación de Habla, Porvoo, Finlandia, p.117-119, 1999, del que son autores Ekudden, E., Hagen, R., Johansson, I., Svedberg, J. Más aún, las descripciones del VAD han sido proporcionadas en "Detección de actividad de voz para el Códec Multi-Tasa Adaptativa de GSM", IEEE Taller de Codificación de Habla, Porvoo, Finlandia, p. 55-57, 1999, del que son autores Vähätalo, A. y Johansson, I.

20 Los codecs de habla de AMR-NB y AMR-WB tienen un número de modos operativos que hacen que sea posible implementar el códec a diferentes tasas de bits, por ejemplo correspondientes a un rendimiento de calidad de habla subjetiva diferente. El códec de AMR-NB puede operar a 8 tasas de bits diferentes comprendidas en la gama de 4,75 kbps hasta 12,2 kbps según se describe en la Tabla 1 que sigue. A través de este documento, "bps" define bits 25 por segundo. Durante períodos de silencio de habla, según sean detectados por el VAD, el sistema genera Ruido de Confort (CN) conformado espectralmente. El CN se describe con 35 bits y una transmisión continua dada de estos parámetros de una vez por trama, siendo la tasa de bits para el CN de 1,75 kbps. En la práctica, la actualización real del sistema se hace normalmente una vez cada 8ª trama, de modo que la tasa de bits para el CN es un octavo de ese valor, de hecho 218,75 bps.

30

Tabla 1: Tasas de bits de códec de fuente para el códec de AMR-NB

Modo de códec	Tasa de bits de códec de fuente
AMR_12.20	12,20 kbps
AMR_10.20	10,20 kbps
AMR_7.95	7,95 kbps
AMR_7.40	7,40 kbps
AMR_6.70	6,70 kbps
AMR_5.90	5,90 kbps
AMR_5.15	5,15 kbps
AMR_4.75	4,75 kbps
AMR_SID	1,75 kbps (218,75 bps)

35 Normalmente, el sistema se configura también usando un sistema de transmisión discontinua (DTX) que incluye un Detector de Actividad de Voz (VAD) y un generador de Confort de Ruido (CN). Estos operan de modo que detectan habla activa y transmiten la señal de habla durante la actividad de voz, y también detectan inactividad de habla a efectos de inhibir la transmisión de señal de habla y activar por el contrario la generación de ruido de confort. La proporción de actividad de voz se denomina Factor de Actividad de Voz (VAF). La combinación de las funciones de DTX, VAD y CN se mencionará como "sistema DTX/VAD/CN" de aquí en adelante.

40 La capacidad total de sistema de un sistema de comunicación celular que hace uso de un estándar tal como GSM y WCDMA está relacionada con el factor de actividad de voz (VAF). Un sistema de comunicación celular tiene por lo general dos enlaces de transmisión, enlace ascendente (UL) y enlace descendente (DL), desde y respectivamente hasta el terminal móvil. El sistema de AMR normalmente empleado utiliza el mismo "sistema DTX/VAD/CN" tanto en UL como en DL.

45 La transmisión de habla con operación DTX puede ser mencionada como método de codificación de tasa de bits variable controlada de fuente simple, donde la tasa puede hacerse variar entre dos niveles, uno para habla activa y el otro para inactividad (y transmisión de ruido de confort). Sin embargo, el término operación de tasa de bits variable de tasa controlada de fuente (SCR VBR) se refiere típicamente a un método en el que la tasa de bits durante el habla activa puede hacerse variar según las necesidades de la señal de fuente, por ejemplo a efectos de mantener 50 un nivel de calidad constante. La codificación SCR VBR persigue por lo tanto un objetivo similar como transmisión de habla con DTX pero puede hacer variar adicionalmente la tasa de bits incluso durante el habla activa. Ejemplos de codecs de habla y de audio con SCR VBR son el códec de 3GPP2 VBR-WB, el Códec 3GPP2 de Tasa Variable Potenciada (EVRC) y el Códec de Audio Avanzado (AAC) de MPEG.

55

La codificación por desviación de trama variable (VFO) ha sido descrita en el documento US20070147314 A1. Éste es un método que suspende la transmisión de aquellos segmentos de habla que el decodificador de habla puede extrapolar apropiadamente a partir del habla recibida. La idea básica consiste en operar un códec de longitud de trama fija de tal manera que una trama de codificación ya no esté limitada a empezar inmediatamente después del final de la trama de codificación anterior. La ganancia proporcionada mediante este método consiste en que la tasa de trama efectiva del códec se reduce a pesar de que la longitud de trama de códec se mantiene constante. Puesto que la tasa de bits de codificación está asociada a cada trama de códec transmitida, la tasa media de bits se reduce. El sistema opera por lo tanto como un códec de tasa variable, incluso cuando se utiliza una tasa de bits de codificación constante.

La transmisión conmutada por paquetes en tiempo real de datos de habla y de audio para aplicaciones de Voz sobre Protocolo de Internet (VoIP) hace uso generalmente del Protocolo de Transmisión en Tiempo Real (RTP) de IETF (según se ha descrito en RFC 3650). Este protocolo proporciona un campo de sello de tiempo que indica el instante de muestreo del primer instante de muestra de la primera muestra codificada para el primer bloque de trama en el paquete. Con servicios de VoIP de forma inalámbrica, sigue siendo importante reducir la tasa de bits a través de los enlaces inalámbricos como en las transmisiones de circuito conmutado. La tasa de bits puede reducirse usando transmisión de habla con DTX, otra operación de SCR VBR, o codificación de VFO según se ha descrito anteriormente. Un método adicional de reducción de tasa de bits consiste en reducir la sobrecarga de los paquetes transmitidos. Un método de ese tipo consiste en compresión de cabecera tal como Compresión de Cabecera Robusta (ROHC). La ROHC ha sido descrita con mayor detalle en IETF RFC3095, RFC3843, y RFC4019.

El objetivo general de las tecnologías descritas en lo que antecede consiste en reducir temporalmente los recursos de transmisión utilizados mientras se mantiene la calidad. Los párrafos que siguen discuten algunos problemas relacionados con esas técnicas.

Aunque la codificación de SCR VBR está capacitada para reducir la tasa de bits media de codificación de fuente, no es siempre deseable utilizar esta característica en cada sistema de comunicación. En el sistema de LTE, por ejemplo, un cambio de tasa de bits de codificación de fuente incluye señalización extra que a su vez puede costar recursos de transmisión adicionales o retardo de la transmisión. Véase 3GPP tdoc S4-100438: Sobre la idoneidad de una codificación de tasa variable para VoIP sobre LTE, para más información. Un problema adicional de la codificación de SCR VBR consiste en que solamente reduce la tasa de bits neta del códec. La sobrecarga relacionada con las transmisiones conmutadas de paquetes como las cabeceras de paquete, podría permanecer sin cambio y con ello la reducción de ancho de banda relativa usando codificación de SCR VBR puede ser pequeña y no justificar los costes y las complicaciones asociadas a la misma, especialmente si se consideran los posibles inconvenientes relacionados con el sistema de transmisión que se han descrito.

La codificación de VFO es una solución que direcciona el problema de la codificación de SRC VBR cuando la sobrecarga de paquetes no se adapta a la tasa de bits seleccionada adaptativamente. Sin embargo, la codificación de VFO adolece de otros problemas. Por ejemplo, cuando los límites del segmento no se emparejan con los límites de la trama de códec de habla, se pueden producir pérdidas de eficiencia cuando se usa la VFO con esquemas de compresión de cabecera robusta como la ROHC. Estas pérdidas de eficiencia se producen debido a que el algoritmo de compresión de cabecera puede perder la capacidad de pronosticar eficazmente partes de las cabeceras de los paquetes de IP. Adicionalmente, la ganancia de la codificación de VFO puede quedar limitada en casos en los que no se requiera ningún sacrificio de calidad o se requiera un sacrificio menor. Esto reduce la probabilidad de encontrar segmentos de habla que puedan ser extrapolados apropiadamente por el decodificador a partir del habla recibida con anterioridad, y de ese modo reduce la probabilidad de ahorros.

La DTX es un método de reducción de tasa muy eficaz para períodos de inactividad de habla puesto que suspende la transmisión durante tales períodos. Idealmente, un sistema de DTX podría transmitir solamente habla activa mientras que la señal inactiva (ruido de fondo) que sea irrelevante para el extremo de recepción podría no ser transmitida en modo alguno. En la práctica, no existe ningún algoritmo de VAD ideal que esté capacitado para distinguir las partes de habla activa de las partes inactivas en una señal de habla de entrada con fiabilidad total. Por ello, un importante objetivo consiste en diseñar un sistema de DTX de tal modo que se ahorren tantos recursos de transmisión como sea posible, mientras que evita posibles artefactos de codificación tal como el recorte de partes de habla activa, que podrían afectar seriamente a la calidad de habla. El recorte se produce con frecuencia para partes de cola del habla (recorte de extremo trasero) o también en partes sin voz del habla con baja energía. Una solución al problema de recorte de extremo trasero consiste en añadir lo que se conoce como período de residuo para el período de transición entre habla activa e inactividad. El período de residuo está siempre codificado como habla activa, con independencia de la indicación de habla activa/inactividad o de una indicación de calidad por parte del VAD. Aunque la adición de un período de residuo es una alternativa segura para evitar el recorte de extremo trasero, ello reduce la ganancia de eficiencia de ancho de banda que la DTX puede proporcionar puesto que, por diseño, una porción grande del período de residuo está igualmente en inactividad que no requiere codificación de habla activa para mantener la calidad de la señal.

Las tecnologías y las técnicas descritas con anterioridad, proporcionan un tipo de marco para optimizar el uso del

ancho de banda en un sistema de comunicación. Sin embargo, según resulta evidente a partir de la explicación anterior, estas técnicas son desventajosas o al menos sub-óptimas de alguna manera y sigue existiendo una necesidad de técnicas adicionales que mejoren la eficiencia de ancho de banda de un sistema de comunicación inalámbrica. Además, cualquier mejora debe ser realizada mientras se conserva un nivel apropiado de calidad de servicio.

Se conoce, según la solicitud de Patente EP-1538806 A2, un método de agregación de trama, estando la trama de capa física formada mediante la variación dinámica de un tamaño de paquete agregado en respuesta a una condición de canal.

También se conoce según la solicitud de Patente Internacional WO 2008/024056 A1, un método para adaptar la transmisión de un medio codificado mediante análisis del rendimiento del medio. Dependiendo de la métrica de rendimiento monitorizado, se puede cambiar la agregación de trama en un paquete.

15 Sumario

La agregación de trama permite la reducción de al menos la sobrecarga relativa en relación con la carga útil de habla transmitida. Esto significa que no solo se transmite una trama de habla codificada por paquete sino que se transmite una pluralidad de tramas de habla codificadas en un paquete. Típicamente, la sobrecarga de paquetización no se incrementa significativamente con el tamaño del paquete y con ello la sobrecarga relativa se reduce si se agregan más tramas. El inconveniente de la agregación de trama es que ello provoca un retardo extra correspondiente al número y a las longitudes de las tramas agregadas. Esto se debe a que el paquete no puede ser transmitido hasta que en el mismo se han recopilado todas las tramas. La agregación de trama puede incrementar también el impacto subjetivo de pérdidas de paquetes puesto que la pérdida de un paquete agregado afecta a todas las tramas que están agregadas en el paquete.

Los métodos y aparatos divulgados en la presente memoria tienen como objetivo incrementar la eficiencia empleando agregación de trama, pero de una forma adaptativa controlada por la señal de fuente. En casos en los que la señal de habla sea completamente estacionaria y donde una trama pueda ser extrapolada con una buena fidelidad desde la información de habla previamente recibida sin degradar significativamente la calidad del habla, esta trama se retarda y se agrega con la trama o tramas subsiguientes. En otro caso, la trama no se retarda ni agrega con tramas posteriores y por el contrario se envía directamente.

En consecuencia, se proporciona un método según la reivindicación 1, de procesamiento de una señal para transmisión. El método comprende determinar una calidad de una primera porción de la señal. El método comprende además, si la calidad satisface un criterio de umbral, agregar entonces una representación de la primera porción de la señal con una representación de una porción subsiguiente de la señal.

La calidad asociada a porciones de la señal puede ser definida de varias maneras. Por ejemplo, podría estar relacionada con la importancia subjetiva de las porciones de la señal de entrada que van a ser transmitidas. Una porción de la señal de entrada que sea ruidosa o que tenga un nivel bajo puede ser considerada como subjetivamente menos importante y tiene una calidad asociada más baja. La calidad podría estar también relacionada con la reconstrucción de la porción de señal de entrada tras la codificación y/o la pérdida de trama potencial. Si la señal reconstruida se parece a la señal de entrada con una fidelidad elevada tiene una calidad alta; en otro caso, la calidad será baja. La calidad puede estar también asociada a la estacionariedad de la señal, donde la alta estacionariedad permita extrapolaciones de la señal con calidad alta.

El criterio de umbral de calidad puede consistir en que la calidad de la primera porción de la señal sea menor que un valor de calidad de umbral o bien esté por encima de un valor de calidad de umbral, dependiendo de la definición del término calidad. El criterio de umbral de calidad puede consistir también, por ejemplo, en que la estacionariedad de la primera porción de la señal sea mayor que un valor de estacionariedad de umbral.

La agregación de representaciones de porciones de una señal reduce la sobrecarga de paquetización en los datos que van a ser transmitidos. Sin embargo, la agregación puede incrementar la probabilidad de que una porción de una señal no sea recibida en el momento en que se necesite. En consecuencia, solamente se seleccionan para la agregación representaciones de porciones de una señal que causen un impacto de calidad relativamente pequeño si no se reciben (o se reciban demasiado tarde), o que sean relativamente poco importantes para la calidad percibida, comparando la calidad de esa porción con un valor de calidad de umbral.

La determinación de la calidad de una primera porción de la señal puede comprender hacer una comparación entre una primera porción de la señal y una estimación de una primera porción de la señal. La estimación de una primera porción de la señal puede ser una estimación basada en al menos una porción previa de la señal. La estimación de una primera porción de la señal puede ser determinada usando síntesis local en base a al menos una porción previa de la señal.

Una representación de una porción particular de la señal puede comprender los parámetros de codificación de la

porción particular de la señal. El método puede comprender además codificar la primera porción de la señal, y codificar la porción subsiguiente de la señal. La tasa o tasas de codificación de las porciones de la señal en un paquete agregado puede(n) reducirse en comparación con un paquete no agregado tal como un paquete agregado que no sea más grande que un paquete no agregado.

5 La tasa de codificación de una porción de la señal puede ser determinada según al menos uno de: la calidad determinada; el número de representaciones de porciones de la señal que estén agregadas entre sí; y la tasa de codificación de otra porción de la señal.

10 Se puede agregar una pluralidad de representaciones de porciones de la señal, y la agregación puede detenerse cuando la representación agregada más recientemente de una porción de la señal satisfaga un número máximo de porciones de umbral de la señal que pueden ser agregadas y paquetizadas en un solo paquete agregado, o bien tenga una calidad que no satisfaga los criterios de umbral. Cuando la agregación se detiene, la agregación para las representaciones normalmente agregadas se detiene. Cuando la agregación se detiene, no se agregan más representaciones de porciones de la señal con las representaciones agregadas normalmente. La agregación adicional de representaciones subsiguientes de porciones de señal es todavía posible.

15 El método puede comprender además paquetizar las representaciones agregadas de porciones de la señal en un paquete agregado. El método puede comprender además transmitir las representaciones agregadas de porciones de la señal en un paquete agregado. Una representación de la primera porción de la señal puede ser agregada con una representación de una porción subsiguiente de la señal si se determina que la primera porción está dentro de un período de residuo.

20 Se proporciona además un aparato según la reivindicación 10, para procesar una señal para transmisión. El aparato comprende un analizador de señal dispuesto para determinar la calidad de una primera porción de la señal. El aparato comprende además un agregador dispuesto para agregar una representación de la primera porción de la señal con una representación de una porción subsiguiente de la señal si la calidad satisface un criterio de umbral.

25 El aparato puede comprender además un codificador dispuesto para codificar porciones de la señal para generar representaciones respectivas de las porciones de la señal. El aparato puede comprender además un paquetizador dispuesto para paquetizar las representaciones agregadas de porciones de la señal en un paquete agregado.

30 El agregador puede estar dispuesto para agregar una pluralidad de representaciones de porciones de la señal, y las representaciones agregadas de las porciones pueden ser paquetizadas por medio de un paquetizador cuando la porción más reciente de la señal: satisface un número de umbral máximo de porciones de la señal que pueden ser agregadas en un único paquete agregado, o bien tiene una calidad que no satisface el criterio de umbral. El aparato puede comprender además un transmisor dispuesto para transmitir el paquete agregado.

35 La primera porción de la señal puede ser agregada con una porción subsiguiente de la señal si se determina que la primera porción está dentro de un período de residuo. La calidad de señal se puede determinar usando síntesis local basada en al menos una trama anterior.

40 La tasa de codificación de la primera porción de la señal puede ser determinada según la calidad determinada. La tasa, o tasas, de codificación de las porciones de la señal en un paquete agregado puede(n) ser reducida(s) en comparación con un paquete no agregado, de tal modo que un paquete agregado no sea mayor que un paquete no agregado.

45 Se proporciona además un medio legible con ordenador según la reivindicación 16, que porta instrucciones, el cual, cuando se ejecuta mediante lógica de ordenador, provoca que dicha lógica de ordenador lleve a cabo cualquiera de los métodos definidos en la presente memoria.

Breve descripción de los dibujos

Ahora se va a describir un método y un aparato para mejorar la eficiencia de ancho de banda en una red de comunicaciones inalámbricas, a título de ejemplo solamente, con referencia a los dibujos que se acompañan, en los que:

La Figura 1 ilustra un período de residuo;

Las Figuras 2a y 2b ilustran agregación de dos tramas;

La Figura 3 ilustra un paquete generalizado;

Las Figuras 4a y 4b ilustran la agregación de una pluralidad de tramas;

Las Figuras 5a, 5b y 5c ilustran dos ejemplos específicos de agregación de trama;

La Figura 6 ilustra un método según se describe en la presente memoria;

La Figura 7 ilustra un método alternativo según se describe en la presente memoria, y

La Figura 8 ilustra un aparato según se describe en la presente memoria.

Se debe apreciar que la disposición de bits dentro de un paquete en las Figuras 2, 3, 4 y 5 es puramente figurativa y no se pretende ilustrar el orden en el que son transmitidos bits particulares o secciones de paquetes. Además, el tamaño de las áreas de bits identificadas no está a escala.

5 Descripción detallada

Una primera realización del método divulgado en la presente memoria se aplica en un sistema que usa transmisión discontinua (DTX). Un Detector de Actividad de Voz (VAD) determina si está presente habla activa que necesite transmisión. Con el fin de evitar recortes (según se ha descrito con anterioridad) se aplica un período de residuo entre cada período de habla activa y cada período de habla inactiva. El período de residuo comprende una pluralidad de tramas que son inactivas según el VAD pero en las que la decisión del VAD puede ser todavía poco fiable o inestable.

15 Se ha ilustrado un período de residuo en la Figura 1. Se determina que un período de habla activa 110 finalice en el extremo de la trama 0. A continuación sigue un período de residuo 120 que comprende 7 tramas. Un período de inactividad 130 se inicia al principio de la trama 8. La tasa de bits requerida para la transmisión de la señal ha sido mostrada en el eje de ordenadas de la Figura 1. Tanto el período de habla activa 110 como el período de residuo 120 tienen la misma tasa de bits. El período de inactividad 130 tiene una tasa de bits inferior a la del período de habla activa 110 y a la del período de residuo 120.

20 Típicamente, durante el período de residuo 120, incluso aunque la decisión del VAD sea errónea y la señal sea realmente habla, la señal de habla es asimismo de baja energía y sin voz. Se ha reconocido que tales sonidos pueden ser normalmente extrapolados desde el habla recibida con anterioridad sin impacto perceptivo significativo.

25 Según una primera realización del método divulgado en la presente memoria, no todas las tramas son paquetizadas y transmitidas inmediatamente. Por el contrario, una pluralidad de tramas que tienen una calidad particular, tal como las que están dentro de un período de residuo, son agregadas en un paquete con anterioridad a la transmisión. El número de tramas agregadas por paquete puede ser un parámetro de diseño o un detalle de implementación. En el ejemplo 2, las tramas son agregadas en un único paquete.

30 A título de ejemplo, se va a describir ahora la aplicación de este método al período de residuo ilustrado en la Figura 1. Este período de residuo tiene una duración de $n=7$ tramas, y las tramas dentro de este período están indexadas desde $k=1$ a $k=7$. La primera trama con índice $k=1$ no está agregada con ninguna trama subsiguiente y es transmitida directamente. Las tramas $k=2$, $k=4$ y $k=6$, sin embargo, son retardadas y agregadas con las tramas subsiguientes respectivas $k=3$, $k=5$ y $k=7$. Estas son paquetizadas a continuación y transmitidas. Esta agregación ha sido ilustrada en la Figura 1 mediante flechas de doble cabeza.

35 El método según esta primera realización proporciona una reducción de la sobrecarga de paquetización durante el período de residuo. Esto puede hacerse sin impacto significativo sobre la calidad de comunicación dado que si no se recibe una trama del período de residuo a tiempo para volver a crear la señal para la reproducción cuando se necesite, entonces la porción de señal contenida en esa trama puede ser extrapolada de manera fiable desde la señal precedente.

40 Un ejemplo específico va a ser descrito ahora con referencia a las Figuras 2a y 2b, para demostrar el ahorro de ancho de banda mediante este método. Una sobrecarga típica que ocurre en redes de comunicación móviles a través de una interfaz de aire como la LTE y que utiliza codecs de habla AMR o AMR-WB con un formato de carga útil RTP como IETF RFC 4867, podría comprender 64 bits por paquete para la sobrecarga de RLP y 10 bits por paquete de trama simple para la cabecera de carga útil de RTP, lo que da un total de 74 bits. La Figura 2a muestra dos paquetes consecutivos de los que cada uno contiene una trama, las tramas 1 y 2 respectivamente, cada una de ellas codificada a una tasa de bits de codificación y cada una de ellas con una cabecera de 74 bits.

50 En el caso de que se agreguen dos tramas, la sobrecarga de RLP por paquete sigue siendo la misma y la sobrecarga de RTP resulta ser de 16 bits, lo que da un tamaño total de cabecera de 80 bits. Con ello, la agregación de las dos tramas reduce la sobrecarga total desde 74 bits por trama hasta 40 bits por trama. Esto ha sido ilustrado en la Figura 2b, donde se ha mostrado un solo paquete que contiene ambas trama 1 y trama 2, con ambas tramas codificadas de modo que tienen el mismo número total de bits como en la Figura 2a, con la cabecera de 80 bits ilustrada abarcando ambas tramas.

55 La agregación de trama de la Figura 2b proporciona de ese modo una reducción de tasa de bits de 34 bits por trama en comparación con la disposición no agregada de la Figura 2a. Existen cincuenta tramas de 20 milisegundos por segundo, de modo que un ahorro de 34 bits por trama equivale a un ahorro de 1700 bits por segundo. La reducción de tasa de bits con esta realización puede ser pequeña pero es significativa. En particular, en casos con habla entre ruido de fondo, es probable que la decisión de VAD alterne frecuentemente entre habla activa e inactividad, y en tal caso la ganancia pueda ser muy meritoria.

65 En una segunda realización, se extiende el método descrito con anterioridad para incrementar la eficiencia de ancho

de banda. Esta realización se basa en la observación de que las tramas durante el período de residuo (con independencia de si se refieren a habla o no) no requieren típicamente que sea transmitida la tasa de bits completa con una calidad aceptable, a diferencia con las tramas durante un período de habla activa. Según esta realización, se aplica una reducción de tasa de bits de codificación de fuente para la señal durante el período de residuo, que no afecta significativamente a la calidad de señal recibida. Esto se consigue usando un códec de SCR VBR para transmitir las tramas que se producen durante el período de residuo a una tasa de bits reducida. Sin embargo, con el fin de subsanar los problemas identificados anteriormente asociados a esquemas de SCR VBR, esta realización utiliza una técnica de ese tipo junto con agregación de trama adaptativa. La aplicación de la agregación de trama adaptativa a codificación de SCR VBR hace que sea posible mantener un tamaño de paquete constante para la transmisión, reduciendo el número de paquetes que son transmitidos cuando se utilizan tasas de bits más bajas para codificación.

La Figura 3 ilustra un paquete generalizado. El número de bits por trama usado para la transmisión de la señal (por ejemplo, habla activa) es n , y éstos se mencionan como bits de fuente. La sobrecarga de paquetización comprende p_1 bits. En consecuencia, el número de bits por paquete con una trama por paquete será $g = n + p_1$.

La Figura 4 ilustra la agregación de a paquetes (Figura 4a) en un solo paquete (Figura 4b). La Figura 4a muestra a paquetes donde cada uno corresponde a una trama y cada uno comprende n bits de fuente y p_1 bits de cabecera. La Figura 4b muestra las a tramas de la Figura 4a agregadas en un solo paquete que ahora tiene un tamaño total de g bits. La sobrecarga de paquete para los paquetes agregados es p_a bits, y por lo tanto los bits de fuente disponibles en el paquete agregado son $g - p_a$ bits. De ello se deduce que, con una distribución uniforme, la tasa de codificación de fuente disponible por trama dentro del paquete agregado será de $n_a = (g - p_a) / a$. Esto reducirá la tasa de bits bruta necesaria para la transmisión tanto como en un factor de $1/a$.

Por lo tanto, conforme a una segunda realización, se realiza la agregación de trama adaptativa de fuente según se ha ejemplificado en lo que antecede para la primera realización descrita, pero además de esto, la tasa de bits de codificación de fuente para cada trama se reduce de tal modo que los paquetes que contienen tramas agregadas son del mismo tamaño que los paquetes que contienen tramas no agregadas. Este tamaño constante de paquete subsana los problemas de señalización encontrados con el uso de un códec de SCR VBR a través de ciertas redes de comunicaciones, tal como por ejemplo, una red de LTE.

Esta segunda realización puede extenderse además al uso de agregación de trama adaptativa durante el período de residuo con el fin de adaptar el número de tramas agregadas. La adaptación puede hacerse dependiendo de la capacidad esperada del descodificador para extrapolar las tramas agregadas para compensar la probabilidad incrementada de que la trama se reciba demasiado tarde. La adaptación puede estar basada, a su vez, en las propiedades de la señal de entrada, tal como por ejemplo la energía, el espectro, la sonoridad, la estacionariedad, etc., o depender de la fiabilidad de la decisión del VAD (es decir, estar basada en parámetros calculados en el algoritmo de VAD).

La segunda realización, no solo reduce la sobrecarga de paquetización durante un período de residuo, sino que también subsana los inconvenientes anteriormente descritos en cuanto a codificación de SCR VBR que resultan de la transmisión de paquetes de tamaño variable en determinadas redes de transmisión tal como la LTE. Según se describe en la presente memoria, el códec de habla es operado con operación de tasa controlada de fuente (SCR). La tasa de codificación se selecciona dependiendo de las propiedades de la señal de habla de entrada, pero con una limitación de tasa y agregación. La limitación es tal que para un grupo de a tramas consecutivas, la tasa de bits por trama puede ser cambiada de una trama algo regular de n bits por trama a alguna otra tasa por trama bajo la provisión de que estas tramas sean agregadas a un tamaño de paquete de g bits. Este tamaño de paquete de g bits es el tamaño total del paquete incluyendo la sobrecarga de paquetización y es idéntico al tamaño de paquete usado para tramas que son transmitidas sin agregación, como paquetes de trama simple. Si el número de tramas agregadas es a y la sobrecarga para el paquete de trama agregada p_a , entonces (como en lo que antecede) la tasa de codificación de fuente máxima disponible por trama será $n_a = (g - p_a) / a$. No obstante, es incluso posible que la tasa de las tramas agregadas pueda variar dentro del paquete agregado, siempre que se cumpla la limitación de que el número de bits de las tramas agregadas (n_1, n_2, \dots, n_a) sume hasta un total que tenga un valor máximo de $g - p_a$ bits. Esta modificación puede ser particularmente útil si la lógica de decisión de SCR asigna diferentes tasas objetivo de bits para las tramas de dentro de un bloque de agregación, por ejemplo con el fin de mantener una cierta calidad de codificación objetiva.

Las Figuras 5a, 5b y 5c ilustran ejemplos de cómo pueden ser agregadas dos tramas en un paquete usando operación de tasa de bits variable. La Figura 5a muestra una tasa de bits de fuente regular (no agregada) que, en este ejemplo, es de 12,2 kbps. Esta tasa corresponde al modo 12.2 de AMR, el cual corresponde a su vez a 244 bits de fuente por trama. La sobrecarga p_1 para cada trama es de 74 bits. El paquete incluye además dos bits de relleno. Por lo tanto, el tamaño total del paquete para paquetes no agregados regulares en este ejemplo es de 320 bits. Ahora, en la situación en que el algoritmo de decisión de SCR detecte un sonido (por ejemplo, sin voz sostenido) que pueda ser codificado con una tasa de bit baja, éste puede determinar agregar dos tramas en un paquete. La sobrecarga de paquetización para el paquete agregado es de 80 bits y por lo tanto el número de bits disponibles

para codificación de fuente de las dos tramas dentro del paquete agregado es de 240 bits. La Figura 5b muestra un primer ejemplo de agregación de trama en el que el algoritmo de decisión de tasa elige asignar el mismo número de bits de fuente a cada una de las dos tramas, es decir 120 bits. Esto hace que sea posible aplicar el modo 5.9 de AMR para ambas tramas, el cual utiliza 118 bits extraídos de los 120 bits disponibles para cada trama. Esto deja 4 bits sin uso en el paquete.

La Figura 5c muestra un segundo ejemplo de agregación de trama en donde el algoritmo de decisión de tasa elige asignar una proporción ligeramente más alta de bits de fuente disponibles dentro del paquete agregado a la primera trama en comparación con la segunda trama. Un motivo para hacer esto consiste en que se determina que la calidad requerida tras la codificación de la primera trama sea mayor que la calidad requerida tras la codificación de la segunda trama. En este caso, se asignan 136 bits de fuente a la primera trama y se asignan 104 bits a la segunda trama. Esto significa que la codificación de la primera trama puede usar el modo 6.7 de AMR, que requiere 134 bits para la primera trama, y que la segunda trama puede usar el modo 5.15 de AMR, que requiere 103 bits para la segunda trama. Esto significa que se usan en el paquete 237 bits de fuente extraídos de los 240 bits disponibles, dejando 3 bits sin uso.

Obsérvese que tanto en el primero como en el segundo ejemplos, el tamaño total del paquete permanece sin cambio a 320 bits. Además, en esos ejemplos, solamente unos pocos bits de fuente posibles podrían mantenerse sin uso: 4 bits en el primer ejemplo y 3 bits en el segundo ejemplo.

Según se ha discutido con anterioridad, la agregación de trama tiene en general el inconveniente de conducir a un retardo de agregación extra, correspondiente al número de tramas agregadas y a sus longitudes. Haciendo de nuevo referencia a los ejemplos de agregación de la Figura 5, la trama 1 se retarda hasta que la trama 2 está lista para la paquetización, lo que necesitará al menos tanto tiempo como la duración de la trama 2, la cual es de 20 milisegundos. Para muchos sistemas de comunicación previstos para llevar a cabo habla conversacional (en tiempo real), los retardos de transmisión deben mantenerse bajos. Por esta razón, la agregación de trama no se utiliza generalmente para transportar habla conversacional debido a que el retardo de comunicación adicional de extremo a extremo que esto provoca es demasiado grande, proporcionando un impacto negativo sobre la calidad de la comunicación. Sin embargo, los métodos y aparatos descritos en la presente memoria subsanan este problema de tal modo que la reducción de sobrecarga y los beneficios del ahorro de tasa de bits de la agregación de trama sacan provecho, sin embargo, del retardo de extremo a extremo y de ese modo la calidad de la comunicación no se ve perjudicada.

Para ilustrar este aspecto de los métodos y aparatos descritos en la presente memoria, se proporcionará un breve análisis del retardo de la cadena de transmisión en un sistema de telefonía de VoIP. El retardo de extremo a extremo está compuesto por varios factores contribuyentes, los cuales incluyen el retardo de códec algorítmico, el retardo de procesamiento de señal para codificación y decodificación, el retardo de paquetización, el retardo de transmisión de enlace de radio, el retardo de transporte de red central, y el retardo de búfer de inestabilidad. Considerando en primer lugar el retardo de paquetización y el retardo de búfer de inestabilidad, el retardo de paquetización incrementado podría ser compensado, por ejemplo, reduciendo el retardo de búfer de inestabilidad, con el fin de mantener el retardo total de extremo a extremo sin cambio. El retardo de inestabilidad compensa las variaciones de la cantidad de tiempo que necesita para que las tramas se desplacen desde un extremo de comunicación hasta el otro extremo; si el retardo de búfer de inestabilidad se reduce, entonces se permite menos tiempo para que una trama se desplace y se considera que una proporción más alta de tramas han llegado demasiado tarde. Esta posibilidad incrementada de pérdida posterior da como resultado una pérdida de trama incrementada.

La funcionalidad de ocultación de pérdida de trama del descodificador tiene como objetivo la ocultación de estas pérdidas de trama. Esta trabaja extrapolando una estimación de la señal representada por la trama perdida a partir de la señal según ha sido representada mediante las tramas previamente recibidas. Por supuesto, tal extrapolación no es típicamente perfecta y existirá alguna diferencia entre la porción de señal extrapolada y la porción de señal codificada originalmente que estaba representada en la trama perdida. Esta diferencia perjudica la calidad del habla reconstruida.

Una ventaja de los métodos y aparatos divulgados en la presente memoria consiste en que la agregación de trama se realiza en función del contenido de la trama, y/o de las propiedades perceptivas de la porción de señal que la trama representa. Esto permite solamente la selección de un subconjunto de tramas que van a ser agregadas. En particular, las tramas que se seleccionan para agregación son aquellas que, tras una pérdida debida a una llegada del paquete demasiado tarde, pueden ser extrapoladas desde la señal de habla más recientemente recibida sin impacto perceptivo significativo. Esto puede determinarse en el extremo de codificación intentando extrapolar una porción de señal representada por una trama actual procedente de la señal anterior. Si la diferencia entre la porción de señal extrapolada y la porción de señal actual, real, está por debajo de un valor de umbral, entonces puede determinarse que un descodificador estará capacitado para reconstruir una estimación aceptable de la porción de señal si la trama que representa la porción de señal no ha sido recibida, o se recibe demasiado tarde. Como tal, el codificador puede correr el riesgo de una probabilidad incrementada de que la trama no sea recibida por agregación de la misma con una trama subsiguiente.

Según el aparato y los métodos divulgados en la presente memoria, una unidad de control de agregación de trama tiene en cuenta propiedades de la señal de habla de entrada y la capacidad de la unidad de ocultación de pérdida de trama del descodificador para ocultar una pérdida potencial de trama. Si, conforme al análisis de esta unidad de control, la señal de habla actual pudiera ser extrapolada apropiadamente desde la señal de habla pasada en un receptor o descodificador, esto puede iniciar la agregación de una trama de habla actual en un paquete con una trama subsiguiente. La unidad de control de agregación de trama puede, en su decisión de adaptación, considerar varias propiedades de la señal de entrada; éstas podrían ser parámetros de la señal tales como la energía, el espectro, la sonoridad, etc. Una propiedad de señal particularmente útil que puede ser considerada es la estacionariedad. Una alta estacionariedad significa que los parámetros de codificación se mantienen relativamente sin cambio o al menos son predecibles con el tiempo. La alta estacionariedad indica que el descodificador estará igualmente capacitado para extrapolar una buena aproximación de la porción de señal actual a partir de la señal precedente.

Considerando la siguiente estimación de retardo sobre la interfaz de aire de LTE, el retardo de paquetización incrementado debido a agregación de trama puede ser compensado en un sistema que use transmisión a través de una interfaz de aire de LTE reduciendo la estimación de retardo sobre la interfaz de aire de LTE. La reducción de la estimación de retardo sobre la interfaz de aire de LTE permite que el retardo general de extremo a extremo se mantenga a un nivel aceptable. Sin embargo, esto no se hace generalmente puesto que reduce la probabilidad de que el protocolo de retransmisión de HARQ transmita con éxito cualesquiera paquetes de datos de habla perdidos en caso de un enlace de radio débil. Sin embargo, si se emplea agregación de trama según se describe en la presente memoria, entonces las únicas tramas que deben ser agregadas (y por tanto retardadas) serán aquellas que es improbable que afecten significativamente a la calidad de habla recibida si llegan demasiado tarde al receptor. La Figura 6 ilustra un método para el procesamiento de una señal para transmisión. En 610 se recibe una porción de señal, pudiendo comprender la porción de señal una trama de 20 milisegundos según se utiliza en muchos algoritmos de codificación de audio. La porción de señal se analiza y se realiza una determinación en 620 respecto a si la porción de señal satisface un criterio de umbral. En este ejemplo, se mide la estacionariedad de la señal para determinar lo bien que la funcionalidad de ocultación de pérdida de trama del receptor estará en condiciones de extrapolar la porción de señal. Si la estacionariedad es demasiado baja, entonces se determina una trama que representa la señal que podría ser transmitida y la porción de señal se codifica en 660 para generar una representación de la porción de señal.

Si se determina que la estacionariedad de la porción de señal es suficientemente alta en la etapa 620, entonces, en la etapa 630, se realiza una determinación respecto a si se ha alcanzado un límite de agregación. En este ejemplo, solamente pueden ser agregadas dos tramas en un paquete, y de ese modo se determinará que el límite de agregación ha sido alcanzado si se determinó que la trama precedente era aceptable para agregación. Si en la etapa 630 se determina que el límite de agregación no ha sido alcanzado, entonces la porción de señal se codifica en 640 y se almacena en una memoria de agregador en 650.

Después de que una porción de señal ha sido codificada en 660, ya sea debido a que la porción de señal actual no satisfaga el criterio de umbral o ya sea debido a que se ha alcanzado el límite de agregación, entonces, en 670, se recuperan cualesquiera representaciones codificadas de porciones previas de señal desde la memoria del agregador. En 680, la representación de la porción de señal actual y cualesquiera representaciones de las porciones de señal recuperadas desde la memoria del agregador, son paquetizadas en un paquete agregado. En 690 se transmite el paquete agregado.

Según la realización ilustrada en la Figura 6, la codificación en 640 y 660 se realiza después de que se hayan realizado determinaciones respecto a si el paquete o los paquetes deberá(n) ser agregado(s). Esto significa que la codificación puede ser realizada a la tasa de codificación apropiada para el número de bits de fuente disponibles dentro de un paquete, tanto si es un paquete agregado como un paquete no agregado.

Un método alternativo para procesar una señal para transmisión ha sido ilustrado en la Figura 7. Se utilizan los mismos números de referencia para referirse a las mismas operaciones en comparación con la Figura 6. Una diferencia entre el método de la Figura 7 comparado con la Figura 6 consiste en que las operaciones de codificación 640 y 660 se sustituyen por una única operación de codificación en 715, inmediatamente después de la recepción de la porción de señal en 710 y con anterioridad a la primera determinación en 720. La pronta codificación en 715 puede ser ventajosa en cuanto a la reducción del cálculo requerido para el análisis de la señal para determinar si la porción de señal satisface los criterios de umbral en 720. Esto puede ser particularmente ventajoso si, según la primera realización descrita en lo que antecede, se realiza agregación de trama y se permite que los paquetes agregados sean más grandes que los paquetes no agregados de tal modo que se pueda usar la misma tasa de codificación para una trama con independencia de si será agregada o no.

Sin embargo, cuando se usan tasas de codificación variables para mantener un tamaño de paquete constante para paquetes agregados y no agregados, entonces la codificación en 715 comprende codificar la porción de señal recibida a una pluralidad de tasas de codificación para crear una pluralidad de representaciones de la porción de

señal, teniendo cada representación un tamaño diferente. Se podrá determinar la representación que se va a usar con anterioridad a la paquetización en 780. Esta determinación se hace durante la agregación en 770.

5 En un ejemplo, el límite de agregación es dos, y un paquete agregado tiene bits de fuente disponibles divididos uniformemente entre las dos tramas. En este caso, la codificación en 715 solamente necesitará que se usen dos tasas de codificación, tal como AMR 12.2 y AMR 5.9 según se ha ilustrado en las Figuras 5a y 5b.

10 La Figura 8 ilustra un aparato 800 adecuado para implementar los métodos descritos en la presente memoria. El aparato comprende un analizador de señal 810, un codificador 820, un agregador 830, un paquetizador 840 y un transmisor 850. El analizador de señal 810 recibe una porción de la señal en forma de trama y analiza al menos una de las propiedades discutidas anteriormente de la señal de entrada, tal como estacionariedad, y determina si la trama debe ser transmitida como un solo paquete de trama o si debe ser agregada con otra trama. El analizador de señal 810 puede comprender también un algoritmo de decisión de tasa que seleccione la tasa de codificación de fuente más adecuada para la trama (según el método de la primera realización, esa determinación no es necesaria).

15 El codificador 820 realiza la codificación de la trama recibida a una tasa o tasas predeterminadas, o a una tasa según indique la unidad de control de tasa controlada de fuente y de agregación.

20 El agregador 830 realiza agregación en base a la determinación realizada en el analizador de señal 810. El agregador 830 contiene una memoria para almacenar tramas codificadas que serán agregadas con representaciones codificadas de tramas recibidas posteriormente. Si la trama codificada no debe ser agregada, se propaga directamente hasta el paquetizador 840. En otro caso, la trama codificada se almacena en la memoria del agregador. La agregación continúa hasta que se reciba una trama que no pueda estar sujeta a retardo de agregación o bien hasta que se alcance un límite del número de tramas que puedan ser agregadas en un único paquete. Cuando la agregación acaba, los datos de trama agregados se alimentan al paquetizador 840. El paquetizador 840 empaqueta los datos recibidos desde el agregador 830. El paquetizador 840 está controlado mediante parámetros de tasa y de agregación de trama obtenidos desde el analizador de señal 810. El parámetro de agregación de trama es típicamente el número de tramas que van a ser agregadas en el paquete. Tras la paquetización, el paquete se envía al transmisor 850 para la transmisión.

30 Debe hacerse hincapié en que ninguna de entre la primera realización (en la que son las tramas agregadas pero la tasa de codificación no cambia) ni de la segunda realización (en la que las tramas son agregadas y se usa una tasa de bits variable para la codificación) necesita estar limitada al período de residuo. Estos métodos pueden ser aplicados a cualquier trama en la que se use una propiedad particular de la señal, tal como estacionariedad, para determinar que un retardo de agregación para esa trama es aceptable. Se puede suponer que, puesto que una trama está en el período de residuo, ésta no es voz activa (un requisito previo para ello es que esté en el período de residuo) y de ese modo resulta aceptable someter la trama a un retardo de agregación. De forma similar, la propiedad de la señal que determina si la trama debe ser agregada puede consistir en si el VAD clasifica la trama como situada dentro de un período de residuo.

40 Como resultará evidente para los expertos en la materia, el orden exacto y el contenido de las acciones llevadas a cabo en el método descrito en la presente memoria pueden ser alterados conforme a los requisitos de un conjunto particular de parámetros de ejecución. En consecuencia, el orden en el que se han descrito las acciones no debe ser entendido como una limitación estricta sobre el orden en el que las acciones deben ser llevadas a cabo. El alcance de la invención está definido solamente por las reivindicaciones anexas.

45 Además, mientras que se han proporcionado ejemplos en el contexto de estándares de comunicaciones particulares, no se pretende que estos ejemplos se limiten a los estándares de comunicaciones a los que pueden ser aplicados el método y el aparato divulgados. Por ejemplo, aunque se han proporcionado ejemplos específicos en el contexto de estándares de 3GPP, los principios divulgados en la presente memoria pueden ser aplicados también a cualquier otro sistema de telecomunicación.

50

un paquetizador cuando la porción más reciente de la señal:

- 5 satisface un número máximo de umbral de porciones de la señal que pueden ser agregadas en un único paquete agregado, o bien,
 tiene una calidad que no satisface el criterio de umbral.
- 14.- El aparato de cualquiera de las reivindicaciones 10 a 13, que comprende además un transmisor dispuesto para transmitir el paquete agregado.
- 10 15.- El aparato de cualquiera de las reivindicaciones 10 a 14, en donde la primera porción de la señal es agregada con una porción subsiguiente de la señal si se determina que la primera porción está dentro de un período de residuo.
- 15 16.- Un medio legible con ordenador, portador de instrucciones, el cual, cuando se ejecuta mediante lógica de ordenador, provoca que dicha lógica de ordenador lleve a cabo cualquiera de los métodos definidos mediante las reivindicaciones 1 a 9.

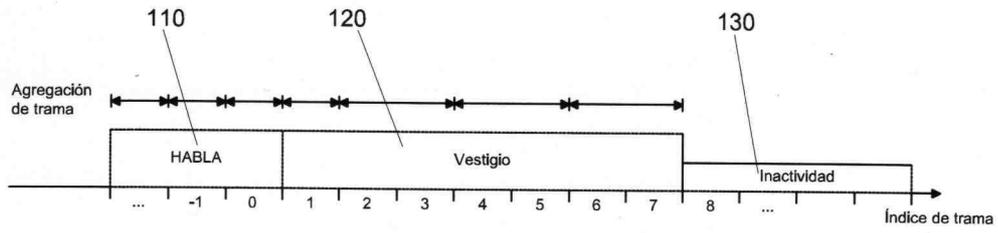


Fig. 1

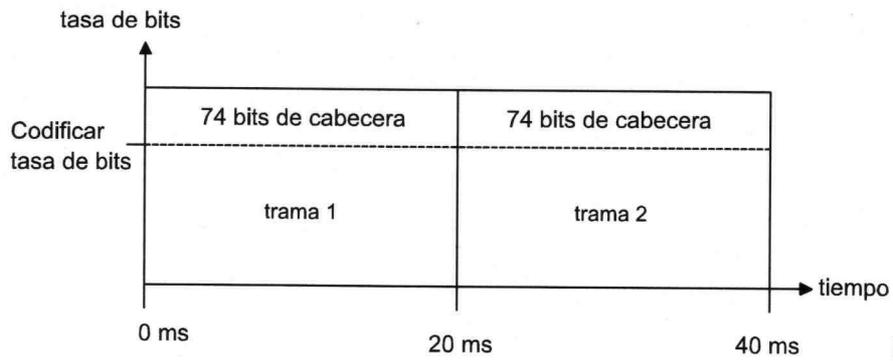


Fig. 2a

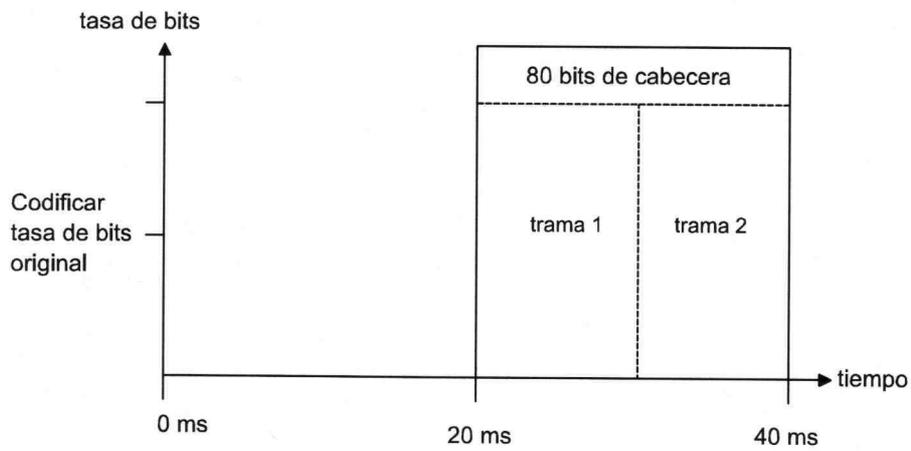


Fig. 2b

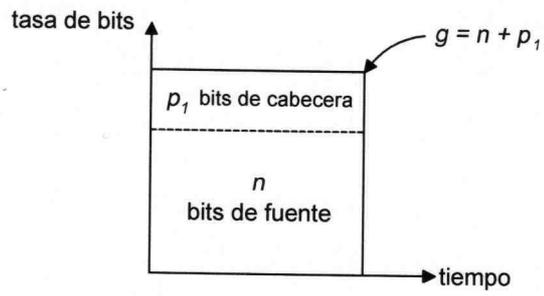


Fig. 3

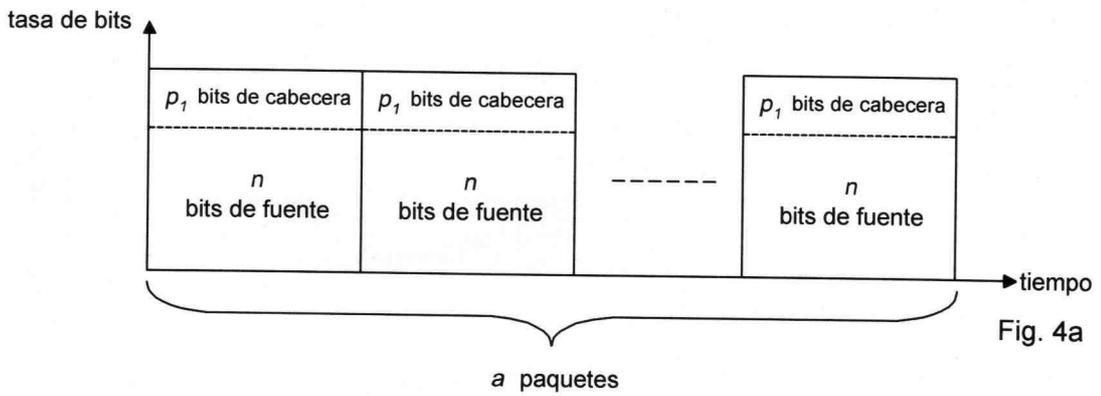


Fig. 4a

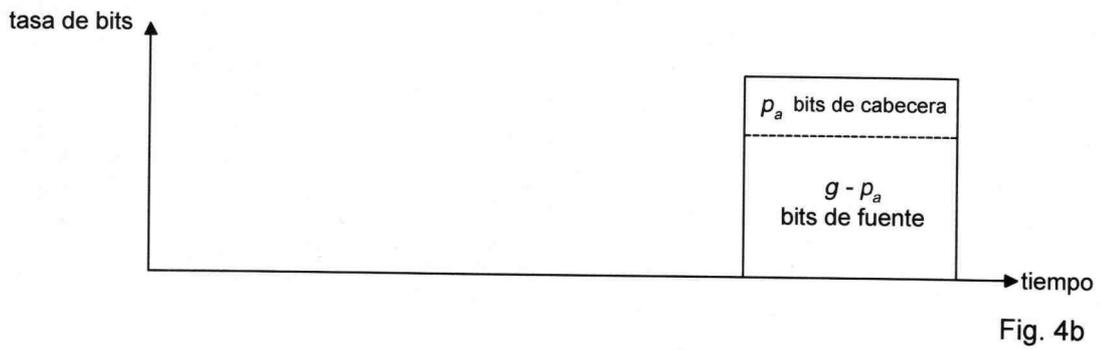
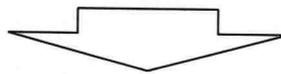
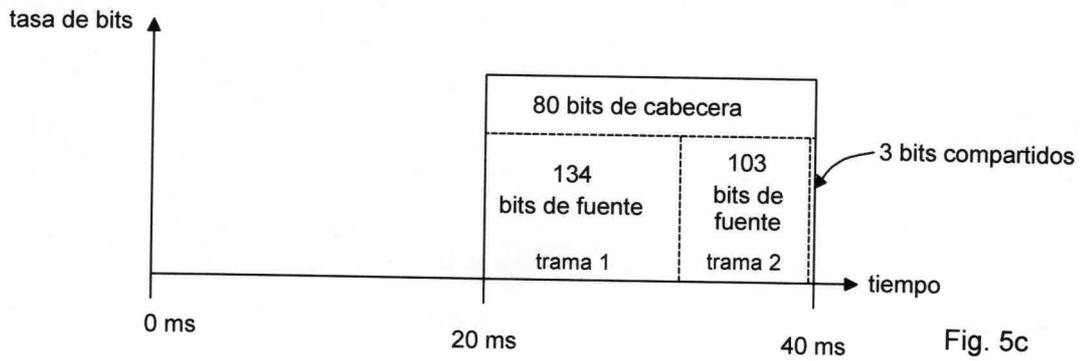
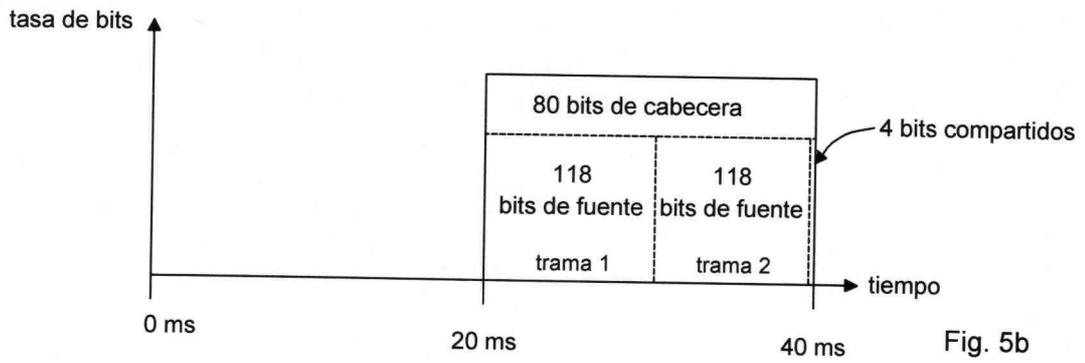
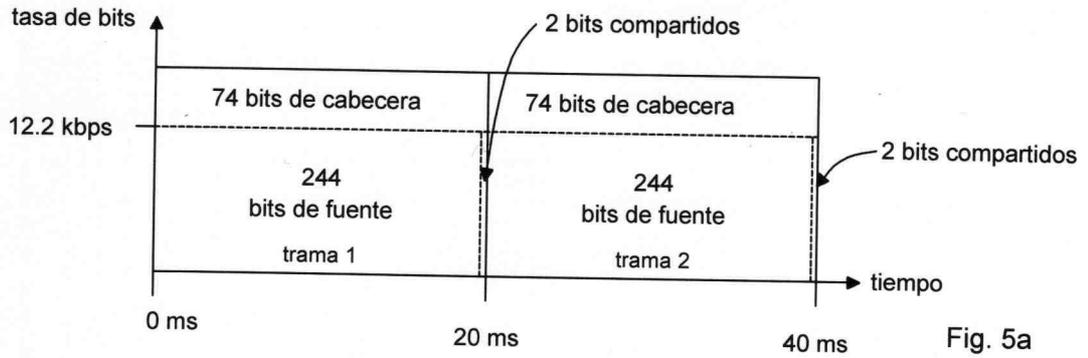


Fig. 4b



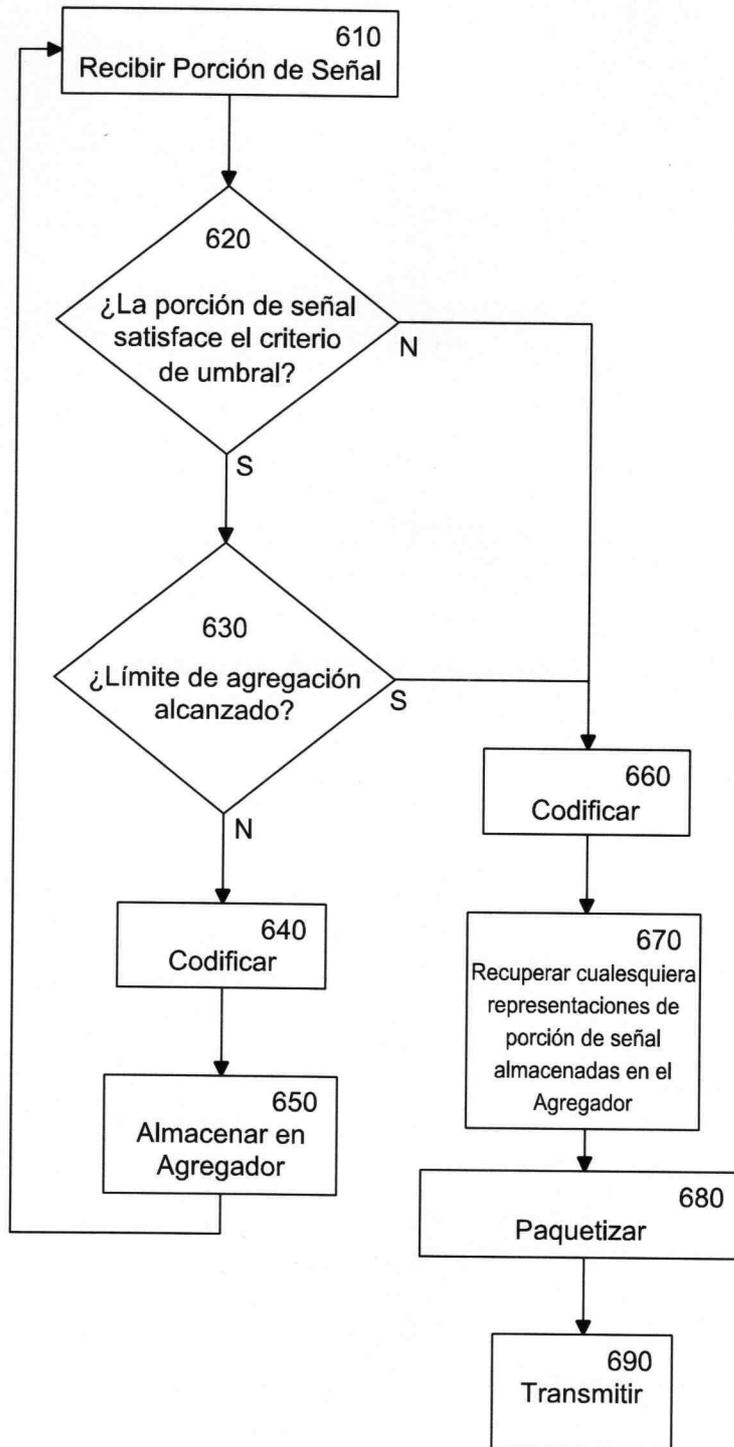


Fig. 6

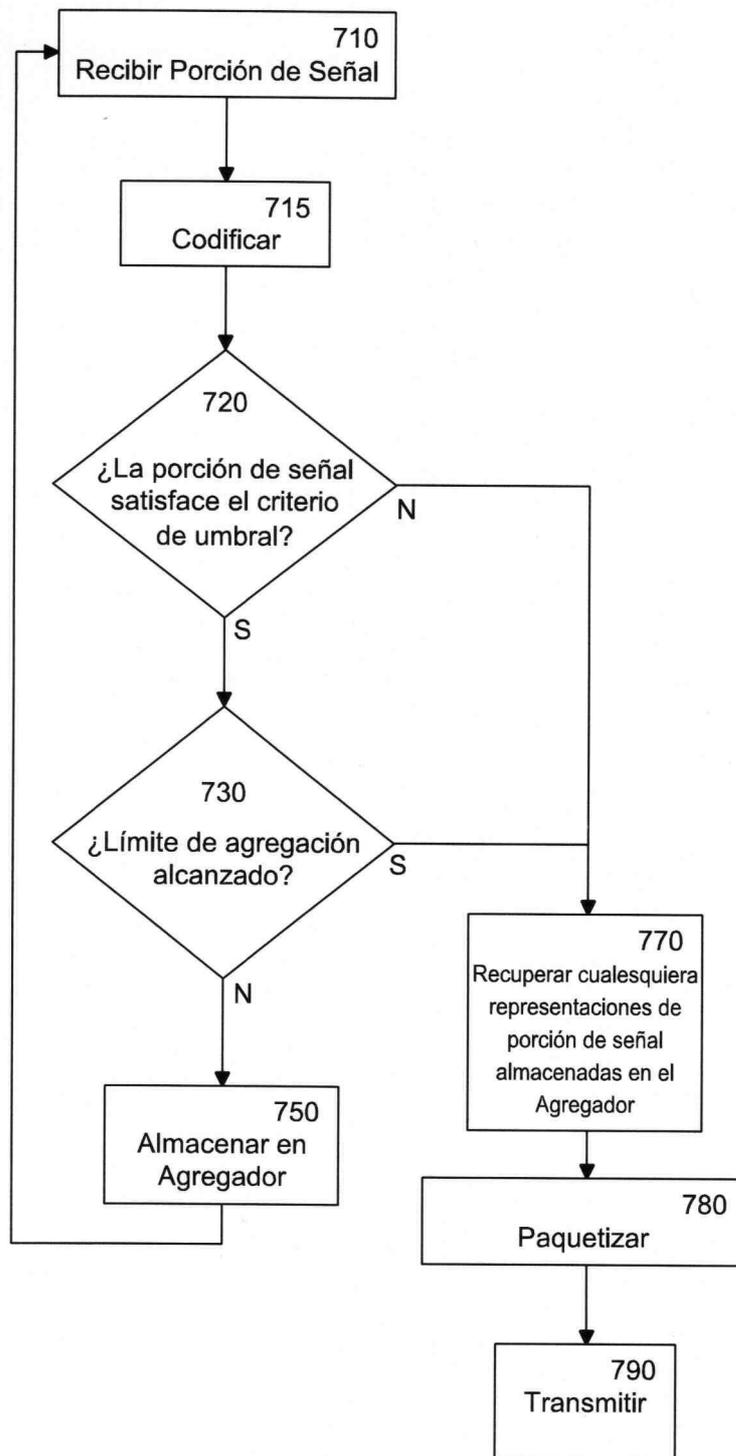


Fig. 7

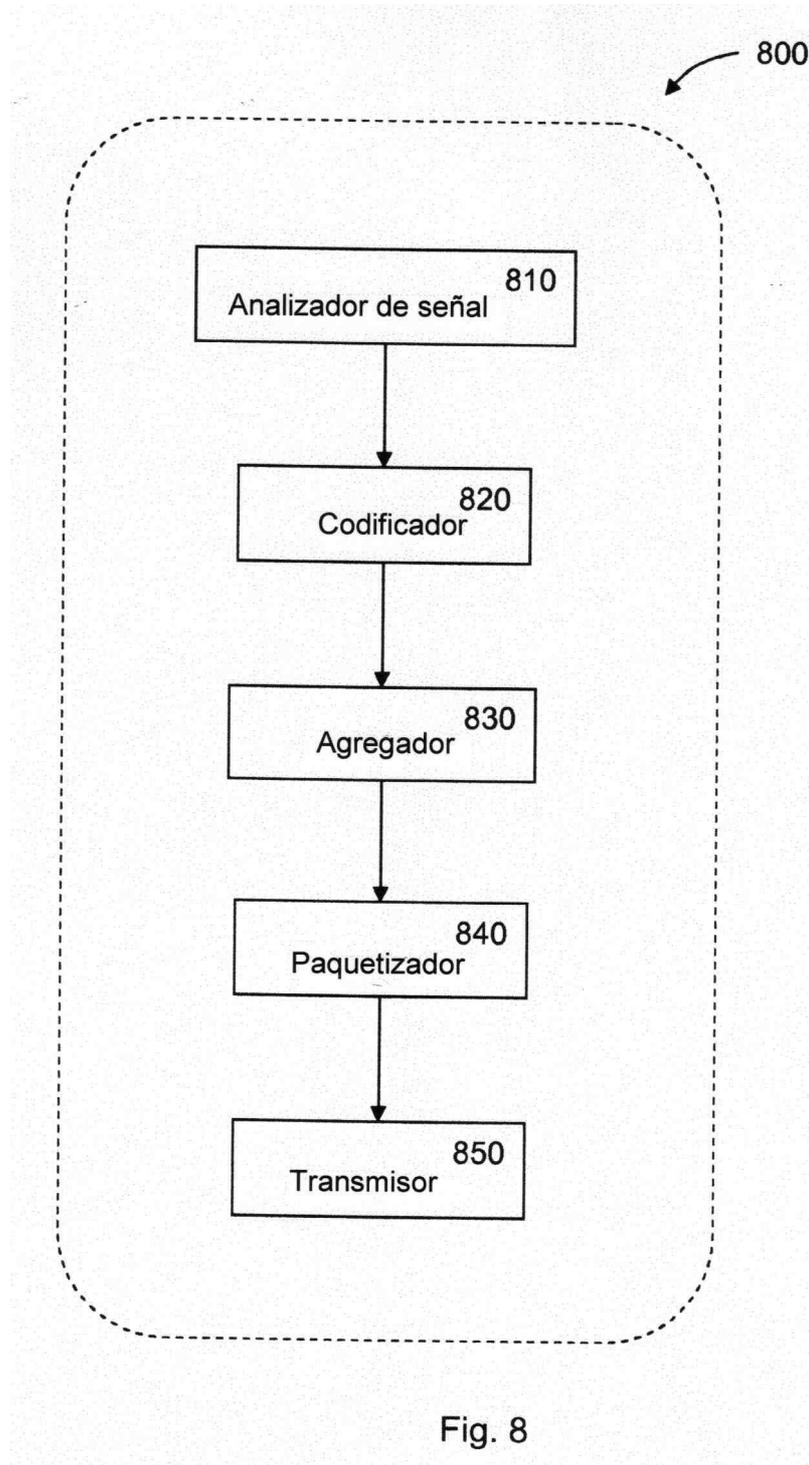


Fig. 8