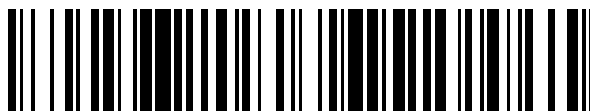


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 574 680**

51 Int. Cl.:

H04M 1/725 (2006.01)
G10L 17/00 (2013.01)
H04M 1/60 (2006.01)
H04M 1/656 (2006.01)
G10L 15/08 (2006.01)
G10L 25/78 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.04.2011 E 11717375 (7)**

97 Fecha y número de publicación de la concesión europea: **02.03.2016 EP 2556652**

54 Título: **Sistema y procedimiento de registro de audio inteligente para dispositivos móviles**

30 Prioridad:

30.03.2011 US 201113076242
08.04.2010 US 322176 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.06.2016

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)
Attn: International IP Administration, 5775
Morehouse Drive
San Diego, California 92121, US

72 Inventor/es:

LEE, TE-WON;
EL-MALEH, KHALED;
YOO, HEEJONG y
SHIN, JONGWON

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 574 680 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

Sistema y procedimiento de registro de audio inteligente para dispositivos móviles

5 ANTECEDENTES

I. Campo

10 La presente divulgación se refiere, en general, a la captura de señales de audio y de habla. Más específicamente, la divulgación se refiere a dispositivos móviles capaces de iniciar y / o terminar operaciones de captura de señales de audio y de habla, o bien, de manera intercambiable, de registrar una operación, en base al análisis de información de contexto de audio.

15 II. Descripción de la técnica relacionada

Gracias al avance de la tecnología del control de energía en los Circuitos Integrados Específicos de la Aplicación (ASIC) y a la potencia informática aumentada de los procesadores móviles tales como el Procesador de Señales Digitales (DSP) o los microprocesadores, un número creciente de dispositivos móviles son ahora capaces de habilitar características mucho más complejas, que no eran consideradas como factibles hasta tiempos recientes, debido a la falta del soporte requerido de potencia de cálculo o de hardware (HW). Por ejemplo, las estaciones móviles (MS) o los teléfonos móviles fueron desarrollados inicialmente para permitir la comunicación de voz o de habla por tradicionales redes celulares inalámbricas basadas en circuitos. Por tanto, la MS fue originalmente diseñada para abordar aplicaciones fundamentales de voz, como la compresión de la voz, la cancelación del eco acústico (AEC), la supresión del ruido (NS) y el registro de la voz.

25 El proceso de implementación de un algoritmo de compresión de voz es conocido como vocodificación y el aparato de implementación es conocido como un vocodificador o "codificador del habla". Existen varios algoritmos estandarizados de vocodificación, como soporte de los distintos sistemas de comunicación digital que requieren la comunicación del habla. El Proyecto 2 de Colaboración de 3ª Generación (3GPP2) es una organización ejemplar de estandarización que especifica tecnología de Acceso Múltiple por División de Código (CDMA), tal como los sistemas de comunicación de IS-95, de Tecnología 1x de Transmisión de Radio (1xRTT) Optimizada para Datos - Evolución de CDMA2000 (EV-DO). El Proyecto de Colaboración de 3ª Generación (3GPP) es otra organización ejemplar de estandarización que especifica el Sistema Global para Comunicaciones Móviles (GSM), el Sistema Universal de Telecomunicaciones Móviles (UMTS), el Acceso de Paquetes de Enlace Descendente de Alta Velocidad (HSDPA), el Acceso de Paquetes de Enlace Ascendente de Alta Velocidad (HSUPA), la Evolución del Acceso de Paquetes de Alta Velocidad (HSPA+) y la Evolución a Largo Plazo (LTE). El Protocolo de Voz por Internet (VOIP) es un protocolo ejemplar usado en sistemas de comunicación definidos en 3GPP y 3GPP2, así como otros. Los ejemplos de vocodificadores empleados en tales sistemas y protocolos de comunicación incluyen el G.729 de la Unión Internacional de Telecomunicaciones (ITU)-T, el códec Adaptativo de Múltiples Velocidades (AMR) y el Códec Mejorado de Velocidad Variable (EVRC), opciones de servicios del habla 3, 68 y 70.

La grabación de la voz es una aplicación para grabar la voz humana. La grabación de la voz es mencionada a menudo como registro de voz o memoria de voz, de forma intercambiable. La grabación de la voz permite a los usuarios guardar alguna parte de una señal del habla, recogida por uno o más micrófonos, en un espacio de memoria. La grabación de voz guardada puede ser reproducida más tarde en el mismo dispositivo, o puede ser transmitida a un dispositivo distinto, a través de un sistema de comunicación de voz. Aunque los grabadores de la voz pueden grabar algunas señales musicales, la calidad de la música grabada, habitualmente, no es magnífica, porque el grabador de voz está optimizado para características vocales emitidas por un conducto vocal humano.

50 La grabación de audio, o el registro de audio, se usa a veces de manera intercambiable con la grabación de la voz, pero se entiende a veces como una aplicación distinta, para grabar cualquier sonido audible, incluyendo la voz humana, los instrumentos y la música, debido a su capacidad de capturar señales de mayor frecuencia que las generadas por el conducto vocal humano. En el contexto de la presente solicitud, la terminología de "registro de audio" o "grabación de audio" será usada extensamente para referirse a la grabación de voz o a la grabación de audio.

El registro de audio permite la grabación de todas, o algunas de, las partes de una señal de audio de interés, que son habitualmente recogidas por uno o más micrófonos en uno o más dispositivos móviles. El registro de audio es mencionado a veces como grabación de audio o memoria de audio, de forma intercambiable. El documento WO 2004 / 057892 A1 describe un procedimiento para organizar la información proporcionada por usuarios, disponible en un dispositivo terminal móvil, con meta-información para permitir la extracción de la información proporcionada por el usuario. Durante la grabación, se obtiene información adicional proporcionada por el usuario e información de contexto referida a la operación de grabación.

65 SUMARIO

La invención está definida en las reivindicaciones adjuntas, a las cuales debería hacerse referencia ahora.

Este documento describe un procedimiento de procesamiento de una señal de audio digital para un dispositivo móvil. Este procedimiento incluye recibir una señal acústica por al menos un micrófono; convertir la señal acústica recibida en la señal de audio digital; extraer al menos una información de contexto auditivo de la señal de audio digital; en respuesta a detectar automáticamente un indicador de inicio de suceso, realizar un registro de audio para la señal de audio digital; y en respuesta a detectar automáticamente un indicador de fin de suceso, finalizar el registro de audio. Esta al menos una información de contexto auditivo puede estar relacionada con la clasificación de audio, la identificación de palabras clave o la identificación del orador. Esta al menos una información de contexto auditivo puede estar basada, al menos en parte, en la energía de señal, la razón entre señal y ruido, el declive espectral o la frecuencia del cruce por el cero. Esta al menos una información de contexto auditivo puede estar basada, al menos en parte, en información no auditiva, tal como información de planificación o información de cronograma. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador relacionados con este procedimiento.

Este documento también describe un procedimiento de procesamiento de una señal de audio digital para un dispositivo móvil. Este procedimiento incluye recibir una señal acústica por parte de al menos un micrófono; transformar la señal acústica recibida en una señal eléctrica; muestrear la señal eléctrica en base a una frecuencia de muestreo y a un ancho de datos para cada dato muestreado, para obtener la señal de audio digital; almacenar la señal de audio digital en un almacén temporal; extraer al menos una información de contexto auditivo de la señal de audio digital; en respuesta a detectar automáticamente un indicador de inicio de suceso, realizar un registro de audio para la señal de audio digital; y en respuesta a detectar automáticamente un indicador de fin de suceso, finalizar el registro de audio. Esta detección de los indicadores de inicio o de fin de suceso puede estar basada, al menos en parte, en información no auditiva tal como información de planificación o información de cronograma. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento de detección de un indicador de inicio de suceso. Este procedimiento incluye seleccionar al menos una información de contexto desde dicha al menos una información de contexto auditivo; comparar la información de contexto seleccionada con al menos un umbral predeterminado; y determinar si el indicador de inicio de suceso ha sido detectado en base a la comparación de la información de contexto seleccionada con al menos un umbral predeterminado. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento de detección de un indicador de fin de suceso. Este procedimiento incluye seleccionar al menos una información de contexto procedente de dicha al menos una información de contexto auditivo; comparar la información de contexto seleccionada con al menos un umbral predeterminado; y determinar si el indicador de fin de suceso ha sido detectado, en base a la comparación de la información de contexto seleccionada con al menos un umbral predeterminado. Esta detección de un indicador de fin de suceso puede estar basada, al menos en parte, en la no ocurrencia de un suceso auditivo durante un periodo de tiempo predeterminado. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento de realización del registro de audio. Este procedimiento incluye actualizar al menos un parámetro relacionado con la conversión, en base, al menos en parte, a dicha al menos una información de contexto auditivo; en respuesta a la determinación de si se requiere un procesamiento adicional, en base, al menos en parte, a dicha al menos una información de contexto auditivo, aplicar el procesamiento adicional a la señal de audio digital para obtener una señal de audio procesada; y almacenar la señal de audio procesada en un almacenamiento de memoria. El procesamiento adicional puede ser procesamiento de realce de señales, tal como la cancelación de eco acústico (AEC), recepción de realce de voz (RVE), cancelación de ruido activo (ANC), supresión de ruido (NS), control de ganancia acústica (AGC), control de volumen acústico (AVC) o control de gama dinámica acústica (ADRC). La supresión del ruido puede estar basada en una solución basada en un único micrófono o en múltiples micrófonos. El procesamiento adicional puede ser procesamiento de compresión de señales, tal como la compresión del habla o la compresión de audio. Los parámetros de compresión, tales como la modalidad de compresión, la velocidad de bits o el número de canal, pueden ser determinados en base a la información de contexto auditivo. El almacenamiento de memoria incluye una memoria local dentro del dispositivo móvil o una memoria remota conectada con el dispositivo móvil a través de un canal inalámbrico. La selección entre la memoria local y la memoria remota puede estar basada, al menos en parte, en la información de contexto auditivo. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento para un dispositivo móvil que incluye detectar automáticamente un indicador de inicio de suceso; procesar la primera parte de una señal de entrada de audio para obtener una primera información en respuesta a la detección de un indicador de inicio de suceso; determinar al menos un parámetro de grabación en base a la primera información; y reconfigurar una unidad de captura de audio del dispositivo móvil, en base a dicho al menos un parámetro de grabación determinado. Esta reconfiguración puede

ocurrir durante una parte inactiva de la señal de entrada de audio. Este al menos un parámetro de grabación incluye información indicativa de una frecuencia de muestreo o de un ancho de datos para un convertidor de A / D del dispositivo móvil. Este al menos un parámetro de grabación incluye información indicativa del número del micrófono activo del dispositivo móvil, o información de temporización, indicativa del intervalo de vigilia o la duración activa de al menos un micrófono. Esta primera información puede ser información de contexto que describe un entorno en el cual el dispositivo móvil está grabando, o una característica de la señal de entrada de audio. Este indicador de inicio de suceso puede estar basado en una señal transmitida por un canal inalámbrico. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento para un dispositivo móvil que incluye detectar automáticamente un indicador de inicio de suceso; procesar una primera parte de una señal de entrada de audio, para obtener una primera información en respuesta a la detección de un indicador de inicio de suceso; determinar al menos un parámetro de grabación en base a la primera información; reconfigurar una unidad de captura de audio del dispositivo móvil, en base a dicho al menos un parámetro de grabación determinado; procesar una segunda parte de la señal de entrada de audio, para obtener una segunda información; realzar la señal de entrada de audio suprimiendo un ruido de fondo, para obtener una señal realzada; codificar la señal realzada para obtener una señal codificada; y almacenar la señal codificada en un almacenamiento local dentro del dispositivo móvil. Esta codificación de la señal realzada incluye determinar un tipo de codificación basado en la segunda información; determinar al menos un parámetro de codificación para la codificación determinada; y procesar la señal realzada en base al tipo de codificación determinado y a dicho al menos un parámetro de codificación determinado, para obtener la señal codificada. En la presente memoria, dicho al menos un parámetro de codificación incluye la velocidad de bits o la modalidad de codificación. Además, este procedimiento puede incluir determinar un grado del realce de la señal de entrada de audio, en base a la segunda información. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

Este documento también describe un procedimiento para un dispositivo móvil que incluye detectar automáticamente un indicador de inicio de suceso; procesar una primera parte de una señal de entrada de audio para obtener una primera información en respuesta a la detección de un indicador de inicio de suceso; determinar al menos un parámetro de grabación en base a la primera información; reconfigurar una unidad de captura de audio del dispositivo móvil, en base a dicho al menos un parámetro de grabación determinado; procesar la segunda parte de la señal de entrada de audio para obtener una segunda información; realzar la señal de entrada de audio suprimiendo un ruido de fondo, para obtener una señal realzada; codificar la señal realzada para obtener una señal codificada; y almacenar la señal codificada en un almacenamiento local dentro del dispositivo móvil. Además, este procedimiento puede incluir detectar automáticamente un indicador de fin de suceso; y, en respuesta a la detección de un indicador de fin de suceso, determinar una ubicación de almacenamiento a largo plazo para la señal codificada, entre el almacenamiento local dentro del dispositivo móvil y un almacenamiento en red conectado con el dispositivo móvil, a través de un canal inalámbrico. Esta determinación de la ubicación de almacenamiento a largo plazo puede estar basada en una prioridad de la señal codificada. Este documento también describe un aparato, una combinación de medios y un medio legible por ordenador, relacionados con este procedimiento.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Los aspectos y las ventajas concomitantes de las realizaciones descritas en la presente memoria devendrán más inmediatamente evidentes por referencia a la siguiente descripción detallada, cuando se considere conjuntamente con los dibujos adjuntos, en los que:

la FIG. 1A es un diagrama que ilustra el concepto de un sistema inteligente de registro de audio.

La FIG. 1B es otro diagrama que ilustra el concepto de un sistema inteligente de registro de audio.

La FIG. 1C es un diagrama que ilustra el concepto de un sistema convencional de registro de audio.

La FIG. 2 es un diagrama de una realización ejemplar del sistema inteligente de registro de audio.

La FIG. 3 es un diagrama de una realización de la Unidad de Procesamiento de Salida 240.

La FIG. 4 es un diagrama de una realización de la Unidad de Procesamiento de Entrada 250.

La FIG. 5 es un diagrama de una realización del Procesador de Registro de Audio 230.

La FIG. 6 es un diagrama que ilustra ejemplos de información de contexto S600.

La FIG. 7 es un diagrama de una realización del identificador de contexto 560.

La FIG. 8 es un diagrama de una realización ejemplar del identificador de contexto 560 y de la información de contexto S600.

La FIG. 9A es una realización del mecanismo de generación de un indicador de inicio de suceso de nivel único.

La FIG. 9B es otra realización del mecanismo de generación de un indicador de inicio de suceso de nivel único.

La FIG. 10 es una realización del mecanismo de generación de un indicador de fin de suceso.

La FIG. 11 es un diagrama de una primera realización ejemplar que ilustra los estados del Procesador de Registro de Audio 230 y la transición de los mismos.

La FIG. 12 es un diagrama de una segunda realización ejemplar que ilustra los estados del Procesador de Registro de Audio 230 y la transición de los mismos.

La FIG. 13 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el estado S1 de monitorización pasiva de audio o el estado S4 de monitorización de audio.

La FIG. 14 es un diagrama de un ejemplo para almacenar entrada de audio digital en el Almacén Temporal 220 en la Unidad de Captura de Audio 215 durante el estado S1 de monitorización pasiva de audio o el estado S4 de monitorización de audio.

La FIG. 15 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el estado S1 de monitorización pasiva de audio.

La FIG. 16 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el estado S2 de monitorización activa de audio.

La FIG. 17 es un diagrama de ejemplo para almacenar entrada de audio digital en el Almacén Temporal 220 en la Unidad de Captura de Audio 215 durante el estado S2 de monitorización activa de audio.

La FIG. 18 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el estado S2 de monitorización activa de audio.

La FIG. 19 es un diagrama de ejemplo de una realización de identificación de contexto en el Procesador de Registro de Audio 230 durante el estado S2 de monitorización activa de audio.

La FIG. 20 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el estado S3 o S5 de registro activo de audio.

La FIG. 21 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el estado S3 de registro activo de audio.

La FIG. 22 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el estado S4 de monitorización de audio.

La FIG. 23 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el estado S5 de registro activo de audio.

La FIG. 24 es un diagrama de flujo de una realización del módulo central de registro de audio durante los estados S3 o S5 de registro activo de audio.

La FIG. 25 es un diagrama de una realización del control de ACTIVACIÓN y DESACTIVACIÓN de micrófono único.

La FIG. 26 es un diagrama de una primera realización del control de ACTIVACIÓN y DESACTIVACIÓN de micrófono único.

La FIG. 27 es un diagrama de una segunda realización del control de ACTIVACIÓN y DESACTIVACIÓN de micrófono único.

La FIG. 28 es un diagrama de una primera realización del control de ACTIVACIÓN y DESACTIVACIÓN de múltiples micrófonos.

La FIG. 29 es un diagrama de una segunda realización del control de ACTIVACIÓN y DESACTIVACIÓN de múltiples micrófonos.

La FIG. 30 es un diagrama de una realización del control de número de micrófono activo.

La FIG. 31 es un diagrama de una realización de la selección de ubicación de almacenamiento, en la cual la selección puede ser controlada de acuerdo a la prioridad pre-definida de la información de contexto S600.

5 La FIG. 32 es un diagrama de una realización de la selección de ubicación de almacenamiento, en la cual la selección puede ser controlada dinámicamente, de acuerdo a la prioridad de la información de contexto S600 durante el Estado de Registro Activo de Audio S3 o S5.

10 La FIG. 33 es un diagrama de una realización de una configuración del tiempo de agotamiento de almacenamiento, en la cual el tiempo de agotamiento puede ser controlado de acuerdo a una prioridad pre-definida de la información de contexto S600.

15 La FIG. 34 es un diagrama de una realización del arranque, etapa por etapa, de bloques dentro del sistema inteligente de registro de audio, en la cual el número de bloques activos y el consumo total de energía de los mismos pueden ser controlados dinámicamente de acuerdo a cada estado.

20 La FIG. 35 es un diagrama de una realización del control de precisión de un convertidor de A / D, en la cual la precisión puede ser configurada como correspondiente a cada estado predeterminado, o controlada dinámicamente de acuerdo a la información de contexto S600.

La FIG. 36 es un diagrama de una realización del control de realce de la señal de entrada de audio, en la cual el realce puede ser dinámicamente configurado de acuerdo a la información de contexto S600.

25 La FIG. 37 es un diagrama de una realización del control de parámetros de compresión de audio, en la cual la compresión puede ser dinámicamente configurada de acuerdo a la información de contexto S600.

30 La FIG. 38 es un diagrama de una realización de la selección del formato de codificación de compresión, en la cual la selección del formato de codificación de compresión, o la ausencia de la misma, puede ser dinámicamente configurada de acuerdo a la información de contexto S600.

DESCRIPCIÓN DETALLADA

La presente solicitud será mejor entendida con referencia a los dibujos adjuntos.

35 A menos que esté expresamente limitado por su contexto, el término “señal” se usa en la presente memoria para indicar cualquiera de sus significados corrientes, incluyendo un estado de una ubicación de memoria (o conjunto de ubicaciones de memoria) según lo expresado en un cable, bus u otro medio de transmisión. A menos que esté expresamente limitado por su contexto, el término “generar” se usa en la presente memoria para indicar cualquiera de sus significados corrientes, tales como computar o producir de otro modo. A menos que esté expresamente limitado por su contexto, el término “calcular” se usa en la presente memoria para indicar cualquiera de sus significados corrientes, tales como computar, evaluar y / o seleccionar entre un conjunto de valores. A menos que esté expresamente limitado por su contexto, el término “obtener” se usa para indicar cualquiera de sus significados corrientes, tales como calcular, deducir, recibir (p. ej., desde un dispositivo externo) y / o extraer (p. ej., desde una formación de elementos de almacenamiento). Allí donde el término “comprender” se usa en la presente descripción y las reivindicaciones, no excluye otros elementos u operaciones. El término “basado en” (como en “A está basado en B”) se usa para indicar cualquiera de sus significados corrientes, incluyendo los casos (i) “basado en al menos” (p. ej., “A está basado en al menos B”) y, si corresponde en el contexto específico, (ii) “igual a” (p. ej., “A es igual a B”).

50 A menos que se indique lo contrario, cualquier divulgación de un funcionamiento de un aparato que tiene una característica particular también está expresamente concebida para divulgar un procedimiento que tenga una característica análoga (y viceversa), y cualquier divulgación de un funcionamiento de un aparato de acuerdo a una configuración específica también está expresamente concebida para divulgar un procedimiento de acuerdo a una configuración análoga (y viceversa). A menos que se indique lo contrario, el término “contexto” (o “contexto de audio”) se usa para indicar un componente de audio o del habla, y lleva información desde el entorno ambiental del orador, y el término “ruido” se usa para indicar cualquier otro fenómeno en la señal de audio o de habla.

60 La FIG. 1A es un diagrama que ilustra el concepto de sistema inteligente de registro de audio. Uno o más micrófonos en un dispositivo móvil pueden ser configurados para recibir señales acústicas continuamente o periódicamente, mientras el dispositivo móvil está en modalidad de reposo. La señal acústica recibida puede ser convertida en señal de audio digital por un convertidor de Analógico a Digital (A / D). Esta conversión puede incluir transformar la señal acústica recibida en una señal eléctrica, en forma analógica o continua, en general, muestreando o cuantizando la señal eléctrica para generar una señal de audio digital. El número y el tamaño de la señal de audio digital pueden depender de una frecuencia de muestreo y de un ancho de datos para cada muestra de audio digital. Esta señal de audio digital puede ser configurada para ser almacenada temporalmente en una memoria o un almacén temporal.

65 Esta señal de audio digital puede ser procesada para extraer información significativa. Esta información es mencionada generalmente como “información de contexto S600” o, de forma intercambiable, “información de

contexto auditivo". La información de contexto puede incluir información acerca un entorno en el cual el dispositivo móvil está grabando y una característica de la señal de entrada de audio recibida por al menos un micrófono. La descripción detallada de la información de contexto S600 será presentada en la divulgación posterior.

- 5 El sistema inteligente de registro de audio puede ser configurado para realizar el inicio inteligente 115 o el fin inteligente 150 del registro de audio. En comparación con un sistema convencional de registro de audio, en el cual un usuario inicia o finaliza manualmente la grabación de la señal de audio, el sistema inteligente de registro de audio puede ser configurado para iniciar o finalizar el registro de audio detectando automáticamente un indicador de inicio de suceso o un indicador de fin de suceso. Estos indicadores pueden estar basados en la información de contexto
- 10 obtenida de la señal de audio; bases de datos situadas dentro del dispositivo móvil o conectadas con el dispositivo móvil mediante conexiones de red cableada o inalámbrica; sensores no acústicos; o incluso una señalización desde otros dispositivos inteligentes de registro de audio. Alternativamente, estos indicadores pueden ser configurados para incluir un comando de voz del usuario, o un comando de tecla asimismo. En una realización, el indicador de fin de suceso puede ser configurado para basarse en la no ocurrencia de un suceso auditivo durante un periodo de tiempo predeterminado. La detección del indicador de inicio de suceso y del indicador de fin de suceso puede incluir las etapas de seleccionar al menos una información de contexto específico, a partir de al menos una información de contexto auditivo; comparar la información de contexto seleccionada con al menos un umbral predeterminado, y determinar si los indicadores de inicio o fin de suceso han sido detectados, en base a la comparación.
- 20 El sistema inteligente de registro de audio puede ser configurado para comprender un cierto número de sub-bloques inteligentes o, de manera intercambiable, cimientos inteligentes basados, al menos en parte, en dicha al menos una información de contexto auditivo. El cimiento inteligente puede ser caracterizado por su capacidad de configurar dinámicamente su propia modalidad operativa, o parámetros funcionales, durante el proceso de registro de audio, a diferencia del registro convencional de audio, en el cual la configuración o la modalidad operativa pueden ser
- 25 predeterminadas o determinadas estáticamente durante el funcionamiento.

Por ejemplo, en una realización del registro inteligente de audio, el bloque inteligente de control de micrófono 120 de la FIG. 1A puede ser configurado para ajustar dinámicamente el número de micrófonos activos o el control de temporización de la ACTIVACIÓN / DESACTIVACIÓN de al menos un micrófono durante el proceso de registro de audio, en base a la información de contexto S600. En otra realización, el bloque inteligente convertidor de A / D 125 de la FIG. 1A puede ser configurado para ajustar dinámicamente sus propios parámetros operativos en base a la información de contexto S600. Tales parámetros pueden incluir la frecuencia de muestreo de la señal de audio capturada desde al menos un micrófono, o el ancho de datos de la muestra de audio digital capturada, en base a la información de contexto S600. Estos parámetros pueden ser mencionados como "parámetro de grabación" porque la selección de estos parámetros afectaría la calidad o el tamaño del registro de audio grabado. Estos parámetros pueden ser configurados para ser reconfigurados, o conmutados, durante una parte inactiva de la señal de entrada de audio, para minimizar el impacto sobre la calidad de audio. La parte inactiva de la señal de entrada de audio puede incluir aún algún nivel de actividad mínima de audio. Pero, en general, "parte inactiva" significa ninguna parte activa, así como una parte relativamente menos activa, de la señal de entrada de audio.

40 En otra realización, el bloque inteligente de realce de audio 130 de la FIG. 1A puede ser configurado para seleccionar dinámicamente, en base a la información de contexto S600, si es necesario el realce de la señal de audio y, en tal caso, qué tipo de realce de señal debería ser realizado. El bloque inteligente de realce de audio 130 puede ser configurado para seleccionar el grado del nivel de realce de señal, por ejemplo, realce agresivo o realce menos agresivo, en base a la información de contexto S600. El realce de señal puede ser configurado para basarse en un único micrófono o en múltiples micrófonos. El bloque inteligente de compresión de audio 135 de la FIG. 1A puede ser configurado para seleccionar dinámicamente el tipo de formato de codificación a usar, o los parámetros de codificación del mismo, tales como la modalidad de compresión, la velocidad de bits o el número de canal de audio / habla, en base a la información de contexto S600. Una descripción más detallada y ejemplos de la característica de configuración dinámica de los sub-bloques inteligentes serán presentados posteriormente. El bloque inteligente de resguardo de audio en almacenamiento 145 de la FIG. 1A puede ser configurado para seleccionar la ubicación en la cual será almacenado el registro de audio capturado, en base a la información de contexto S600. La selección puede ser entre una memoria local del dispositivo móvil y una memoria remota conectada con el dispositivo móvil, mediante un canal cableado o inalámbrico. El bloque inteligente de resguardo de audio en almacenamiento 145 puede ser configurado para almacenar la señal de audio digital en la memoria local, por omisión, durante el proceso del registro de audio, y luego determinar posteriormente una ubicación de almacenamiento a largo plazo entre el almacenamiento local y un almacenamiento en red.

60 Debería observarse que los cimientos inteligentes 120, 125, 130, 135, 145 y el orden de los mismos, divulgados en la FIG. 1A, son solamente con fines ejemplares y, por lo tanto, debería ser obvio para un experto en la técnica que algunos de los cimientos pueden ser reordenados, combinados o incluso omitidos, en su totalidad o en parte, dentro del ámbito de la solicitud. Por ejemplo, en una realización de acuerdo a la presente solicitud, el bloque inteligente de realce de audio 130 puede ser omitido o reemplazado por un bloque tradicional de realce de audio, en el cual la capacidad de reconfigurar dinámicamente su propia modalidad operativa, de acuerdo a la información de contexto S600, no está disponible. Análogamente, el bloque inteligente de compresión de audio 135 puede ser omitido o reemplazado por la compresión convencional de audio.

El sistema inteligente de registro de audio también puede referirse al sistema que puede ser configurado para usar la combinación de alguno de los sistemas convencionales existentes de registro de audio y alguno de los cimientos inteligentes, o bien la característica inteligente del inicio / fin del registro, como fue presentada en la FIG. 1B. Por el contrario, la FIG. 1C es un diagrama que ilustra el concepto de sistema convencional de registro de audio, en el cual no se incluye ni la característica inteligente de inicio / fin de registro de audio ni ninguno de los cimientos inteligentes.

La FIG. 1B muestra tres configuraciones conceptuales ejemplares distintas de un sistema inteligente de registro de audio. La configuración 1 presenta el sistema en el cual están implementados tanto la característica inteligente del inicio / fin del registro de audio 165 como los cimientos inteligentes 175. El sistema en la configuración 1, por lo tanto, es considerado como el más avanzado sistema inteligente de registro de audio. La configuración 2 muestra el sistema que puede ser configurado para reemplazar la característica inteligente de inicio / fin de registro de audio 165 de la configuración 1, con una característica convencional de inicio / fin de registro de audio 160. En una implementación alternativa, la configuración 3 muestra el sistema que puede ser configurado para reemplazar los cimientos inteligentes 175 de la configuración 1 por los cimientos convencionales 170.

La FIG. 2 es una realización ejemplar del sistema inteligente de registro de audio. La Unidad de Captura de Audio 215, que comprende la Unidad de Micrófono 200 y el Convertidor de A / D 210, es la interfaz de usuario del sistema inteligente de registro de audio. La Unidad de Micrófono 200 comprende al menos un micrófono que puede ser configurado para recoger o recibir una señal de audio acústico y transformarla en una señal eléctrica. El Convertidor de A / D 210 convierte la señal de audio en una señal digital discreta. En otra realización, dicho al menos un micrófono dentro de la Unidad de Micrófono 200 puede ser un micrófono digital. En tal caso, puede configurarse la etapa de conversión de A / D para que sea omitida.

El Suceso Auditivo S210 se refiere, en general, a una señal de audio o, en particular, a la señal de audio de interés para un usuario. Por ejemplo, el Suceso Auditivo S210 puede incluir, pero no está limitado a, la presencia de una señal de habla, música, características específicas de ruido de fondo o palabras clave específicas. El Suceso Auditivo S210 se menciona a menudo como una “escena auditiva” en la técnica.

La Unidad de Captura de Audio 215 puede incluir al menos un micrófono o al menos un convertidor de A / D. Al menos un micrófono o al menos un convertidor de A / D podría haber sido parte de un sistema convencional de registro de audio, y puede ser arrancado solamente durante el uso activo del dispositivo móvil. Por ejemplo, una unidad tradicional de captura de audio en el sistema convencional puede ser configurada para ser arrancada solamente durante la llamada entera de voz o la grabación entera de vídeo, en respuesta a la selección del usuario de la realización o la recepción de la llamada, o la pulsación del botón de inicio de la grabación de vídeo.

En la presente solicitud, sin embargo, la Unidad de Captura de Audio 215 puede ser configurada para despertar, o arrancar, intermitentemente, incluso durante la modalidad de reposo del dispositivo móvil, además de durante una llamada de voz o durante la ejecución de otras aplicaciones cualesquiera que pudieran requerir el uso activo de al menos un micrófono. La Unidad de Captura de Audio 215 puede incluso ser configurada para permanecer activada, recogiendo continuamente una señal de audio. Este enfoque puede ser mencionado como “Siempre Activada”. La señal de audio recogida S260 puede ser configurada para ser almacenada en el Almacén Temporal 220 en forma discreta.

La “modalidad de reposo” del dispositivo móvil descrito en la presente memoria se refiere, en general, al estado en el cual el dispositivo móvil no está ejecutando activamente ninguna aplicación en respuesta a la entrada manual del usuario, a menos que se especifique lo contrario. Por ejemplo, los dispositivos móviles típicos envían o reciben señales periódicamente a y desde una o más estaciones base, incluso sin la selección del usuario. El estado del dispositivo móvil que realiza este tipo de actividad es considerado como una modalidad de reposo dentro del ámbito de la presente solicitud. Cuando el usuario está participando activamente en la comunicación de voz o la grabación de vídeo, usando su dispositivo móvil, no se considera como una modalidad de reposo.

El Almacén Temporal 220 almacena datos de audio digital temporalmente antes de que los datos de audio digital sean procesados por el Procesador de Registro de Audio 230. El Almacén Temporal 220 puede ser cualquier memoria física y, aunque es preferible que esté situado dentro del dispositivo móvil, debido a las ventajas del acceso más rápido y a la huella de memoria requerida, relativamente pequeña, desde la Unidad de Captura de Audio 215, el Almacén Temporal 220 también podría estar situado fuera de dispositivos móviles, mediante conexiones de red inalámbricas o cableadas. En otra realización, la señal de audio recogida S260 puede ser configurada para estar directamente conectada con el Procesador de Registro de Audio 230, sin estar temporalmente almacenada en el Almacén Temporal 220. En tal caso, la señal de audio recogida S260 puede ser idéntica a la Entrada de Audio S270.

El Procesador de Registro de Audio 230 es una unidad principal de procesamiento para el sistema inteligente de registro de audio. Puede ser configurado para tomar diversas decisiones con respecto a cuándo iniciar o finalizar el registro, o cómo configurar los cimientos inteligentes. Puede ser adicionalmente configurado para controlar bloques adyacentes, para mantener una interfaz con la Unidad de Procesamiento de Entrada 250 o la Unidad de Procesamiento de Salida 240, para determinar el estado interno del sistema inteligente de registro de audio, y para

acceder a la Unidad de Datos Auxiliares 280, o a bases de datos. Un ejemplo de una realización del Procesador de Registro de Audio 230 está presentado en la FIG. 5. El Procesador de Registro de Audio 230 puede ser configurado para leer los datos discretos de entrada de audio almacenados en el Almacén Temporal. Los datos de entrada de audio luego pueden ser procesados para la extracción de la información de contexto S600, que luego puede ser almacenada en memoria situada bien dentro o bien fuera del Procesador de Registro de Audio 230. Una descripción más detallada de la información de contexto S600 está presentada conjuntamente con la descripción de la FIG. 6 y la FIG. 7.

La Unidad de Datos Auxiliares 280 puede incluir diversas bases de datos o programas de aplicación, y puede ser configurada para proporcionar información adicional, que puede ser usada, en parte o en su totalidad, por el Procesador de Registro de Audio 230. En una realización, la Unidad de Datos Auxiliares 280 puede incluir información de planificación del propietario del dispositivo móvil equipado con la característica inteligente de registro de audio. En tal caso, la información de planificación, por ejemplo, puede incluir los siguientes detalles: "la hora y / o duración de la siguiente reunión de negocios", "participantes invitados", "ubicación del lugar de reunión" o "asunto de la reunión", por nombrar unos pocos. En una realización, la información de planificación puede ser obtenida de una aplicación de cronograma tal como Microsoft Outlook u otras aplicaciones de Calendario cualesquiera, comercialmente disponibles. Al recibir o extraer activamente estos tipos de detalles desde la Unidad Auxiliar de Datos 280, el Procesador de Registro de Audio 230 puede ser configurado para tomar decisiones con respecto a cuándo iniciar o detener el registro de audio, de acuerdo a los detalles, preferiblemente, en combinación con la información de contexto S600 extraída de los datos discretos de entrada de audio, almacenados en el Almacén Temporal 220.

El almacenamiento se refiere, en general, a una o más ubicaciones de memoria en el sistema que está diseñado para almacenar el registro de audio procesado proveniente del Procesador de Registro de Audio 230. El Almacenamiento puede ser configurado para comprender el Almacenamiento Local 270, que está disponible localmente dentro de los dispositivos móviles, o el Almacenamiento Remoto 290, que está conectado remotamente con dispositivos móviles mediante un canal de comunicación cableado o inalámbrico. El Procesador de Registro de Audio 230 puede ser configurado para seleccionar dónde almacenar los registros de audio procesados, entre el Almacenamiento Local 270 y el Almacenamiento Remoto 290. La selección de almacenamiento puede ser hecha de acuerdo a diversos factores que pueden incluir, pero no se limitan a, la información de contexto S600, el tamaño estimado de los registros de audio, el tamaño de la memoria disponible, la velocidad de la red, la latencia de la red o la prioridad de la información de contexto S600. La selección de almacenamiento puede incluso ser configurada para ser conmutada dinámicamente entre el Almacenamiento Local 270 y el Almacenamiento Remoto 290 durante el proceso activo de registro de audio, si es necesario.

La FIG. 3 es un diagrama ejemplar de una realización de la Unidad de Procesamiento de Salida 240. La Unidad de Procesamiento de Salida 240 puede ser configurada para entregar la Señal de Salida S230, generada desde el Procesador de Registro de Audio 230, a diversos dispositivos periféricos, tales como un altavoz, un visor, un dispositivo háptico o dispositivos inteligentes externos de registro de audio. El dispositivo háptico permite al sistema proporcionar una experiencia avanzada de usuario, en base a un mecanismo de retro-alimentación táctil. Puede aprovechar el sentido del tacto de un usuario, aplicando fuerzas, vibraciones y / o movimientos al usuario. El sistema inteligente de registro de audio puede transmitir la Señal de Salida S230, a través de la Unidad de Procesamiento de Salida 240, a al menos otro sistema inteligente de registro de audio. La transmisión de la señal de salida puede ser por un canal inalámbrico y pueden ser usados diversos protocolos de comunicación inalámbrica, preferiblemente, tales como GSM, UMTS, HSPA+, CDMA, Wi-Fi, LTE, VOIP o WiMax. La Unidad de Procesamiento de Salida 240 puede ser configurada para incluir el Demultiplexador (De-Mux) 310, que puede distribuir la Señal de Salida S230 selectivamente a los dispositivos periféricos adecuados. El Generador de Salida de Audio 315, si es seleccionado por el De-Mux 310, genera una señal de audio para altavoz o auricular, de acuerdo a la Señal de Salida S230. El Generador de Salida de Visor 320, si es seleccionado por el De-Mux 310, genera una señal de vídeo para un dispositivo visor, de acuerdo a la Señal de Salida S230. El Generador de Salida Háptica 330, si es seleccionado por el De-Mux 310, genera una señal táctil para un dispositivo háptico. El transmisor, si es seleccionado por el De-Mux 310, genera la señal procesada que está lista para su transmisión a los dispositivos externos, incluyendo otro sistema inteligente de registro de audio.

La FIG. 4 es un diagrama ejemplar de una realización de la Unidad de Procesamiento de Entrada 250. En este ejemplo, la Unidad de Procesamiento de Entrada 250 procesa diversos tipos de entradas y genera la Señal de Entrada S220, que puede ser transferida selectivamente, a través del Multiplexador (Mux) 410, al Procesador de Registro de Audio 230. Las entradas pueden incluir, pero no se limitan a, comandos de voz o de tecla del usuario, la señal desde sensores no acústicos tales como una cámara, un temporizador, un GPS, un sensor de proximidad, un Gyro, un sensor ambiental, un acelerómetro, etc. Las entradas pueden ser transmitidas desde al menos otro sistema inteligente de registro de audio. Las entradas pueden ser procesadas en consecuencia por diversos módulos, tales como el Procesador de Comandos de Voz 420, el Procesador de Comandos de Tecla 430, la Interfaz de Temporizador 440, el Receptor 450 o la Interfaz de Sensor 460, antes de que sean enviadas al Procesador de Registro de Audio 230.

La FIG. 5 es un diagrama ejemplar de una realización del Procesador de Registro de Audio 230. El Procesador de Registro de Audio 230 es el principal motor informático del sistema inteligente de registro de audio, y puede ser implementado en la práctica con al menos un microprocesador, o con al menos un procesador de señales digitales, o con cualquier combinación de los mismos. Alternativamente, algunos de, o todos, los módulos del Procesador de Registro de Audio 230 pueden ser implementados en HW. Según se muestra en la FIG. 5, el Procesador de Registro de Audio 230 puede comprender un cierto número de módulos dedicados a una operación específica, así como un módulo más general llamado "Procesador General de Señales de Audio 595".

El módulo Detector de Actividad Auditiva 510, o "detector de audio", puede detectar el nivel de actividad de audio proveniente de la Entrada de Audio S270. La actividad de audio puede ser definida como clasificación binaria, tal como activa o no activa, o con más niveles de clasificación, si es necesario. Pueden ser usados diversos procedimientos para determinar el nivel de audio de la Entrada de Audio S270. Por ejemplo, el Detector de Actividad Auditiva 510 puede estar basado en la energía de la señal, la razón entre señal y ruido (SNR), la periodicidad, el declive espectral y / o la frecuencia de cruce por el cero. Pero es preferible usar soluciones relativamente sencillas, a fin de mantener la complejidad de cálculo tan baja como sea posible, lo cual, a su vez, ayuda a extender la vida de la batería. El módulo Realizador de Calidad de Audio 520 puede mejorar la calidad de la Entrada de Audio S270, suprimiendo el ruido de fondo activamente o pasivamente; cancelando el eco acústico; ajustando la ganancia de entrada; o mejorando la inteligibilidad de la Entrada de Audio S270 para la señal de habla conversacional.

El módulo Analizador de Señal Auxiliar 530 puede analizar la señal auxiliar proveniente de la Unidad de Datos Auxiliares 280. Por ejemplo, la señal auxiliar puede incluir un programa de planificación tal como un programa de cronograma o un programa cliente de correo electrónico. También puede incluir bases de datos adicionales, tales como un diccionario, un perfil de empleado o diversos parámetros de audio y del habla, obtenidos de terceros, o datos de entrenamiento. El módulo Gestor de Señales de Entrada 540 puede detectar, procesar o analizar la Señal de Entrada S220 proveniente de la Unidad de Procesamiento de Entrada 250. El módulo Gestor de Señales de Salida 590 puede generar la Señal de Salida S230, en consecuencia, para la Unidad de Procesamiento de Salida 240.

El Gestor de Señales de Control 550 gestiona diversas señales de control que pueden ser aplicadas a unidades periféricas del sistema inteligente de registro de audio. Dos ejemplos de las señales de control, el Control de Convertidor de A / D S215 y el Control de Unidad de Micrófono S205, son divulgados en la FIG. 5 con fines ejemplares. El Administrador de Inicios de Sucesos 570 puede ser configurado para gestionar, detectar o generar un indicador de inicio de suceso. El indicador de inicio de suceso es un marcador o señal que indica que el registro inteligente de audio puede estar listo para comenzar. Puede ser deseable usar el indicador de inicio de suceso para el Procesador de Registro de Audio 230, para conmutar su estado interno si su funcionamiento está basado en una máquina de estados. Debería ser obvio para un experto en la técnica que el indicador de inicio de suceso es un marcador o señal conceptual para la comprensión del funcionamiento del Procesador de Registro de Audio 230. En una realización, puede ser implementado usando una o más variables en la implementación de SW, o una o más señales cableadas en el diseño de HW. El indicador de inicio de suceso puede estar en un único nivel, en el cual el Indicador de Inicio de Suceso S910 es activado cuando se satisfacen una o más condiciones, o en múltiples niveles, en los cuales el registro inteligente efectivo del audio es activado cuando más de un nivel de indicadores de inicio de suceso son todos activados.

El Procesador General de Señales de Audio 595 es un módulo de múltiples propósitos para gestionar todos los otros procedimientos fundamentales de procesamiento de audio y de habla, no explícitamente presentados en la presente solicitud, pero aún necesarios para una implementación exitosa. Por ejemplo, estos procedimientos de procesamiento de señales pueden incluir, pero no se limitan a, conversiones de tiempo a frecuencia o de frecuencia a tiempo; filtrados misceláneos; ajuste de ganancia de señal; o control de gama dinámica. Debería observarse que cada módulo divulgado por separado en la FIG. 5 se proporciona solamente con fines de ilustración de la descripción funcional del Procesador de Registro de Audio 230. En una realización, algunos módulos pueden ser combinados en un único módulo, o algunos módulos pueden incluso ser adicionalmente divididos en módulos más pequeños en la implementación en la vida real del sistema. En otra realización, todos los módulos divulgados en la FIG. 5 pueden ser integrados como un único módulo.

La FIG. 6 es un diagrama que ilustra ejemplos de la información de contexto S600. A menos que se indique lo contrario, el término "contexto" (o "información de contexto S600") se refiere a información del usuario tal como la identificación, la emoción, los hábitos, la condición biológica o la actividad de participación; a un entorno físico, tal como una ubicación absoluta o relativa; a información sobre el contenido, tal como la identificación de palabras clave o clases; o al entorno social, tal como la interacción social o la actividad comercial. La FIG. 7 es un diagrama de una realización del Identificador de Contexto 560. El Identificador de Contexto 560 es parte del Procesador de Registro de Audio 230 y extrae la información de contexto S600 desde la Entrada de Audio S270. En una realización, el Identificador de Contexto 560 puede ser configurado para ser implementado en un motor de HW dedicado o en un procesador de señales digitales.

La FIG. 8 es un diagrama de una realización ejemplar del Identificador de Contexto 560 y de la información de contexto S600. El Identificador de palabras clave analiza la Entrada de Audio S270 y reconoce palabras clave

importantes entre el contenido del habla conversacional. El proceso de reconocimiento puede estar basado en una base de datos auxiliar, tal como un diccionario o tablas de consulta que almacenan uno o más vocabularios. El Detector de Música / Habla puede ser configurado para clasificar la señal de Entrada de Audio S270 como en más de una categoría, en base a las características de la señal de entrada. La detección puede estar basada en la identificación de parámetros de audio o de habla y en la comparación de los parámetros identificados de audio o de habla con uno o más umbrales. La clasificación dentro del ámbito de la presente solicitud puede ser considerada como detección, de manera intercambiable.

El Detector de Música / Habla 820 también puede ser configurado para clasificar la señal de entrada en una clasificación de múltiples niveles. Por ejemplo, en una realización del Detector de Música / Habla 820, puede clasificar la señal de entrada en una clasificación de primer nivel, tal como "Música" o "Habla", o "Música+Habla". Posteriormente, puede además determinar una clasificación de segundo nivel, tal como "Rock", "Popular" o "Clásica", para la señal clasificada como "Música" en la etapa de clasificación del primer nivel. De la misma manera, puede también determinar una clasificación de segundo nivel, tal como "Conversación de negocios", "Conversación personal" o "Conferencia", para la señal clasificada como "Habla" en la etapa de clasificación del primer nivel.

El Identificador de Orador 830 puede ser configurado para detectar la identificación del orador para la entrada de señales del habla. El proceso de identificación del orador puede estar basado en una característica de la señal del habla de entrada, tal como la energía de señal o de trama, la relación entre señal y ruido (SNR), la periodicidad, el declive espectral y / o la frecuencia del cruce por el cero. El Identificador de Orador 830 puede ser configurado para identificar una clasificación sencilla, tal como "Orador masculino" u "Oradora femenina"; o para identificar información más sofisticada, tal como el nombre o título del orador. La identificación del nombre, o el título, del orador, podría requerir una extensa complejidad informática. Se torna aún más retador cuando el Identificador de Orador 830 tiene que buscar un gran número de muestras de habla por diversos motivos.

Por ejemplo, supongamos la siguiente situación hipotética. La compañía X tiene en total 15.000 empleados y un usuario Y tiene que asistir a una serie de reuniones de conferencia de audio, relacionadas con el trabajo, cada día, usando su dispositivo móvil equipado con una característica inteligente de registro de audio. El usuario Y quiere identificar a los oradores en tiempo real cuando un cierto número de oradores, empleados de la compañía X, están participando en la conversación. Lo primero: las muestras del habla, o las características del habla, extraídas de las muestras del habla, pueden no estar disponibles en primer lugar para todos los empleados. Lo segundo: incluso si están ya disponibles en la memoria local, o en un sector servidor remoto, conectado mediante un canal inalámbrico, la búsqueda de ese gran número de muestras del habla en tiempo real en el dispositivo móvil puede ser extremadamente retadora. Lo tercero: incluso si la búsqueda puede hacerse en el sector servidor remoto y la potencia informática del servidor puede ser significativamente mayor que la del dispositivo móvil, el procesamiento en tiempo real podría aún ser retador, considerando la latencia de transmisión Rx / Tx. Estos problemas pueden hacerse manejables si se dispone de información adicional procedente de una base de datos auxiliar. Por ejemplo, si la lista de participantes de una conferencia está disponible a partir del programa de cronograma, el Identificador de Orador puede reducir efectivamente el número de personas a buscar de manera significativa, estrechando el espacio de búsqueda.

El Detector de Entorno 850 puede ser configurado para identificar una escena auditiva en base a una o más características de la señal del habla de entrada, tal como la energía de trama, la razón entre señal y ruido (SNR), la periodicidad, el declive espectral y / o la frecuencia del cruce por el cero. Por ejemplo, puede identificar el entorno de la señal de entrada actual como "Oficina", "Coche", "Restaurante", "Metro", "Parque de baile", etc.

El Clasificador de Ruidos 840 puede ser configurado para clasificar las características del ruido de fondo de la Entrada de Audio S270. Por ejemplo, puede identificar el ruido de fondo como "Estático o No Estático", "Ruido callejero", "Ruido de avión" o una combinación de los mismos. Puede clasificar el ruido de fondo en base al nivel de severidad del mismo, tal como "Agudo" o "Medio". El Clasificador de Ruidos 840 puede ser configurado para clasificar la entrada en un procesamiento de estado único o un procesamiento de múltiples etapas.

El Detector de Emoción 850 puede ser configurado para detectar la emoción de un orador, para el habla conversacional, o el aspecto emocional del contenido musical. La música consiste en un cierto número de interesantes parámetros acústicos. Por ejemplo, la música puede incluir ritmos, instrumentos, tonos, trozos cantados, timbres, notas y versos. Estos parámetros pueden ser usados para detectar o estimar la emoción de un orador, para una o más categorías de la emoción, tales como felicidad, ira, miedo, victoria, ansiedad o depresión. El Detector de Actividad Participativa 870 puede ser configurado para detectar la actividad del orador, en base a las características de la Entrada de Audio S270. Por ejemplo, puede detectar que el orador está "Hablando", "Corriendo", "Caminando", "Practicando deportes", "En clase" o "De compras". La detección puede estar basada en parámetros del habla y / o parámetros de señales musicales. La detección también puede ser configurada para obtener la información suplementaria desde la Unidad de Datos Auxiliares 280 o los otros módulos en la FIG. 8. Por ejemplo, el Detector de Emoción 850 puede ser configurado para usar la información procedente del Detector de Entorno 860, el Clasificador de Ruido 840 o cualquier otra combinación de los módulos divulgados en la FIG. 8.

La FIG. 9A y la FIG. 9B son diagramas de una realización ejemplar del mecanismo de generación de indicadores de inicio de sucesos de nivel único y de múltiples niveles, respectivamente. Un indicador de inicio de suceso de nivel único es deseable para una realización relativamente sencilla del mecanismo de inicio, mientras que el indicador de inicio de suceso de múltiples niveles es deseable para una realización algo compleja del mecanismo de inicio, por lo cual es deseable un esquema más agresivo de arranque etapa por etapa, para un consumo eficaz de energía. El Administrador de Inicios de Sucesos 570 puede ser configurado para generar el Indicador de Inicio de Suceso S910, de acuerdo a cualquier combinación de las salidas, o señales de activación interna, desde el Detector de Actividad Auditiva 510, el Analizador de Señales Auxiliares 530 o el Gestor de Señales de Entrada 540. Por ejemplo, el Detector de Actividad Auditiva 510 puede ser configurado para generar una señal de activación interna basada en la actividad de la Entrada de Audio S270, cuando se detectan uno o más sucesos o actividades auditivas interesantes.

El Analizador de Señales Auxiliares 530 también puede generar una señal de activación interna de acuerdo a la planificación del programa de cronograma del usuario. Una reunión específica que el usuario deseaba grabar puede generar automáticamente la señal de activación interna sin ninguna intervención manual del usuario. Alternativamente, el Analizador de Señales Auxiliares 530 puede ser configurado para tomar tales decisiones en base a prioridades explícitas o implícitas de la reunión. La generación de la señal de activación interna puede ser iniciada a partir de entradas distintas al análisis de la Entrada de Audio S270 o la Señal Auxiliar. Tales entradas pueden incluir la voz del usuario o controles manuales de teclas; un temporizador; una señal desde sensores no acústicos tales como una cámara, un temporizador, un GPS, un sensor de proximidad, un Gyro, un sensor ambiental o un acelerómetro; o la señal transmitida desde al menos otro sistema inteligente de registro de audio. La Lógica Combinatoria 900 puede ser configurada para generar el Indicador de Inicio de Suceso S910 en base a ciertos mecanismos de combinación de las señales de activación interna. Por ejemplo, la lógica combinatoria puede ser configurada para generar el Indicador de Inicio de Suceso S910 de acuerdo a la operación lógica O o la operación lógica Y de las señales de activación interna provenientes del Detector de Actividad Auditiva 510, el Analizador de Señales Auxiliares 530 o el Gestor de Señales de Entrada 540. En otra realización, puede ser configurada para generar el Indicador de Inicio de Suceso S910 cuando una o más señales de activación interna han sido fijadas o activadas.

Con referencia de nuevo a la FIG. 9B, el Administrador de Inicios de Sucesos 570 puede ser configurado para generar el Indicador de Inicio de Sucesos de 1er nivel S920 y luego el Indicador de Inicio de Sucesos de 2º nivel S930, antes del inicio del registro efectivo. El mecanismo Indicador de Inicio de Sucesos de múltiples niveles, divulgado en la presente memoria, puede ser preferible para determinar un punto de inicio más preciso del registro de audio, apoyándose en más de un nivel de indicadores. Una implementación ejemplar del Indicador de Inicio de Suceso de múltiples niveles puede ser configurada para adoptar un mecanismo de decisión, relativamente sencillo y de baja complejidad, para el Indicador de Inicio de Suceso de 1er nivel, y para adoptar un mecanismo de decisión, sofisticado y de alta complejidad, para el Indicador de Inicio de Suceso de 2º nivel S930. En una realización, la generación del Indicador de Inicio de Suceso de 1er nivel S920 puede ser configurada para que sea esencialmente similar al procedimiento como el del Indicador de Inicio de Suceso S910 en la FIG. 9A. A diferencia de la FIG. 9A, el Procesador de Registro de Audio 230 no inicia el registro efectivo al activarse el Indicador de Inicio de Suceso de 1er nivel S920, sino que, en cambio, preferiblemente, puede despertar o, de manera intercambiable, arrancar los módulos adicionales necesarios para activar la señal del Indicador de Inicio de Suceso de 2º nivel S930, en base a un análisis adicional en profundidad de la Entrada de Audio S270. Estos módulos pueden incluir el Identificador de Contexto 560 y la Lógica de Evaluación de Contexto 950. El Identificador de Contexto 560 analizará luego la Entrada de Audio S270 de acuerdo a procedimientos divulgados en la FIG. 8, y puede detectar o identificar una cierta cantidad de Información de Contexto S600 que puede ser evaluada por la Lógica de Evaluación de Contexto 950. La Lógica de Evaluación de Contexto 950 puede ser configurada para activar el Indicador de Inicio de Suceso de 2º nivel S930, de acuerdo a diversos procedimientos internos de decisión. Tales procedimientos, por ejemplo, pueden incluir el cálculo de la suma ponderada de prioridades para la salida de algunos de, o todos, los sub-módulos divulgados en la FIG. 8, y la comparación de la suma ponderada con uno o más umbrales. Debería observarse que la Lógica de Evaluación de Contexto 950 puede ser implementada bien con SW o bien con HW, o puede ser implementada como parte del Procesador General de Señales de Audio 595 en la FIG. 8.

La FIG. 10 es una realización del mecanismo de generación de indicadores de fines de sucesos. El Indicador de Fin de Suceso S940 puede ser generado por el Administrador de Fines de Sucesos 580, de acuerdo a cualquier combinación de las salidas, o señales internas de activación, provenientes del Detector de Actividad Auditiva 510, el Analizador de Señales Auxiliares 530 o el Gestor de Señales de Entrada 540. El funcionamiento de los módulos en la FIG. 10 es esencialmente similar al procedimiento explicado bien en la FIG. 9A o bien en la FIG. 9B, pero las señales internas de activación desde cada módulo son habitualmente activadas cuando cada módulo detecta indicaciones para detener el registro efectivo, o indicaciones para conmutar a la modalidad eficaz en términos de energía, desde su modalidad operativa actual. Por ejemplo, el Detector de Actividad Auditiva 510 puede activar su señal interna de activación cuando la actividad de audio de la Entrada de Audio S270 deviene significativamente reducida en comparación o, de manera similar, el Analizador de Señales Auxiliares 530 puede activar su señal interna de activación cuando la reunión ha llegado a su hora planificada de terminación. La Lógica Combinatoria 900 puede ser configurada para generar el Indicador de Fin de Suceso S940, en base a ciertos mecanismos de combinación de las señales internas de activación. Por ejemplo, puede ser configurada para generar el Indicador de Fin de Suceso S940 de acuerdo, por ejemplo, a la operación lógica O o la operación lógica Y de las señales internas

de activación provenientes del Detector de Actividad Auditiva 510, el Analizador de Señales Auxiliares 530 o el Gestor de Señales de Entrada 540. En otra realización, puede ser configurada para generar el Indicador de Fin de Suceso S940 cuando una o más señales internas de activación han sido fijadas o activadas.

- 5 La FIG. 11 es un diagrama de una primera realización ejemplar que ilustra estados internos del Procesador de Registro de Audio 230 y la transición de los mismos para el sistema indicador de inicio de suceso de múltiples niveles. El estado por omisión en el arranque del registro inteligente de audio puede ser el Estado de Monitorización Pasiva de Audio S1, durante el cual el dispositivo móvil que comprende la característica de registro inteligente de audio es esencialmente equivalente al habitual estado de la modalidad en reposo. Durante el Estado de
- 10 Monitorización Pasiva de Audio S1, es crítico minimizar el consumo de energía, porque, estadísticamente, el dispositivo móvil permanece en este estado la mayor parte del tiempo. Por lo tanto, la mayoría de los módulos del sistema de registro inteligente de audio, excepto unos pocos módulos requeridos para detectar la actividad de la Entrada de Audio S270, pueden ser configurados para permanecer en un estado durmiente o en otras modalidades cualesquiera de ahorro de energía. Por ejemplo, unos pocos módulos excepcionales de ese tipo pueden incluir la
- 15 Unidad de Captura de Audio 215, el Almacén Temporal 220 o el Detector de Actividad Auditiva 510. En una realización, estos módulos pueden ser configurados para estar constantemente activados o pueden ser configurados para despertar intermitentemente.

- El estado podría ser cambiado, desde el Estado de Monitorización Pasiva de Audio S1 al Estado de Monitorización Activa de Audio S2, al activarse el Indicador de Inicio de Suceso de 1er nivel S920. Durante el Estado de Monitorización Activa de Audio S2, el sistema de registro inteligente de audio puede ser configurado para despertar uno o más módulos adicionales, por ejemplo, tales como el Identificador de Contexto 560 o la Lógica de Evaluación de Contexto 950. Estos módulos adicionales pueden ser usados para proporcionar la monitorización y el análisis en profundidad de la señal de Entrada de Audio S270, para determinar si se requiere que el Indicador de Inicio de Suceso de 2º nivel S930 sea activado de acuerdo a la descripción presentada en la FIG. 9B. Si el Indicador de Inicio de Suceso de 2º nivel S930 es finalmente activado, entonces el sistema efectúa la transición al Estado de Registro Activo de Audio S3, durante el cual proseguirá el registro efectivo de audio. La descripción detallada del funcionamiento ejemplar en cada estado será presentada en los siguientes párrafos. Si el Indicador de Fin de Suceso S940 es activado durante el Estado de Monitorización Activa de Audio S2, el sistema puede ser configurado
- 20 para poner los módulos adicionales, que fueron arrancados durante el estado, en la modalidad durmiente y conmutar el estado de vuelta al Estado de Monitorización Pasiva de Audio S1. De manera similar, si el Indicador de Fin de Suceso S940 es activado durante el Estado de Registro Activo de Audio S3, el sistema puede ser configurado para detener el registro de audio y conmutar el estado de vuelta al Estado de Monitorización Pasiva de Audio S1.

- La FIG. 12 es un diagrama de una segunda realización ejemplar que ilustra los estados internos del Procesador de Registro de Audio 230 y las transiciones de los mismos para el sistema Indicador de inicio de suceso de nivel único. La realización en la presente memoria es más sencilla que la realización divulgada en la FIG. 11, pues hay solamente dos estados operativos disponibles. El estado por omisión en el arranque del registro inteligente de audio puede ser el Estado de Monitorización de Audio S1, durante el cual el dispositivo móvil que comprende la característica de registro inteligente de audio es esencialmente equivalente al habitual estado de modalidad en reposo. Durante el Estado de Monitorización de Audio S4, es preferible minimizar el consumo de energía porque, estadísticamente, el dispositivo móvil permanece en este estado la mayor parte del tiempo. Por lo tanto, la mayoría de los módulos del sistema de registro inteligente de audio, excepto unos pocos módulos, mínimamente requeridos para detectar la actividad de la Entrada de Audio S270, pueden ser configurados para permanecer en estado
- 30 durmiente o en otras modalidades cualesquiera de ahorro de energía. Por ejemplo, los pocos módulos excepcionales pueden incluir la Unidad de Captura de Audio 215, el Almacén Temporal 220 o el Detector de Actividad Auditiva 510. En una realización, estos módulos pueden ser configurados para estar constantemente activados, o pueden ser configurados para despertar intermitentemente.

- El estado podría ser cambiado desde el Estado de Monitorización de Audio S4 al Estado de Registro Activo de Audio S5, al activarse el Indicador de Inicio de Suceso S910. Durante el Estado de Registro Activo de Audio S5, proseguirá el registro efectivo de audio. La descripción detallada del funcionamiento habitual en cada estado será presentada en los siguientes párrafos. Si el Indicador de Fin de Suceso S940 es activado durante el Estado de Registro Activo de Audio S5, el sistema puede ser configurado para detener el registro de audio y conmutar el estado de vuelta al
- 40 Estado de Monitorización de Audio S4.

- La FIG. 13 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el Estado de Monitorización Pasiva de Audio S1 de la FIG. 11, o el Estado de Monitorización de Audio S4 de la FIG. 12. Se supone que el dispositivo móvil que comprende la característica de registro inteligente de audio está inicialmente en la modalidad de reposo. Dos intervalos se presentan en la FIG. 13. T_1 representa un intervalo de vigilia de micrófono y T_2 representa un lapso en que un micrófono permanece activado. Los diagramas de flujo presentados en la presente memoria son solamente con fines ejemplares y debería ser obvio para un experto en la técnica que algunos de los bloques en el diagrama de flujo pueden ser reordenados de manera intercambiable dentro del ámbito de la presente solicitud. Por ejemplo, en una realización, los bloques dedicados para configuraciones de un
- 50 convertidor de A / D 1315, 1320 en la FIG. 13 pueden ser configurados para ser procesados después del bloque que

enciende un micrófono y / o un convertidor de A / D 1330. En tal caso, los bloques 1315, 1320 pueden ser configurados para ser ejecutados cada intervalo T_1 , en lugar de solamente una vez al comienzo del funcionamiento.

Adicionalmente, la FIG. 13 divulga varios conceptos importantes, fundamentales para la implementación del registro inteligente de audio. El convertidor de A / D puede ser programado para mantener una baja resolución en términos de frecuencia de muestreo y / o de ancho de datos. La configuración de baja resolución ayuda a minimizar el tamaño de los datos a procesar y / o almacenar en el Almacén Temporal 220. La alta resolución puede ser usada para mejorar la precisión de la entrada de audio digitalizada. Sin embargo, en una implementación ejemplar, puede ser preferible usar una configuración de baja resolución, debido al aumento del uso del almacén temporal y del consumo de energía de la configuración de alta resolución. La configuración de baja resolución puede ser deseable, considerando que el propósito de los Estados de Monitorización de Audio S1, S2, S4 es, principalmente, detectar y monitorizar entornos, esperando la temporización correcta para iniciar el registro activo de audio.

Un micrófono puede ser configurado para despertar cada intervalo T_1 , el intervalo de vigilia del micrófono, y recoger la Entrada de Audio S270 durante el lapso T_2 , el lapso de ACTIVACIÓN del micrófono. Los valores de T_1 o T_2 pueden estar predeterminados en un intervalo fijo, o pueden ser adaptados dinámicamente durante el tiempo de ejecución. En una implementación ejemplar del sistema, T_1 puede ser mayor que T_2 , o puede determinarse que T_2 sea más pequeño, pero proporcional a T_1 . Si hay más de un micrófono en la Unidad de Micrófono 200, cada micrófono puede ser configurado para que tenga el mismo intervalo, o algún micrófono puede ser configurado para que tenga intervalos distintos de los otros. En una realización, algunos de los micrófonos pueden no ser activados en absoluto durante el Estado de Monitorización Pasiva de Audio S1 de la FIG. 11, o el estado de Monitorización de Audio S4 de la FIG. 12. En otra realización, uno o más micrófonos pueden estar activados constantemente, lo que puede ser el mero caso especial en el cual T_1 es idéntico a T_2 .

Las entradas de audio digitalizado durante el lapso T_2 pueden ser almacenadas en el Almacén Temporal 220 cada intervalo T_1 , y la entrada almacenada de audio digital puede ser objeto de acceso, y procesada, por el Procesador de Registro de Audio 230 en cada intervalo T_3 . Esto puede ser mejor entendido con la FIG. 14, que muestra un diagrama ejemplar para almacenar entrada de audio digital en el Almacén Temporal 220 en la Unidad de Captura de Audio 215 durante el Estado de Monitorización Pasiva de Audio S1 o el Estado de Monitorización de Audio S4. La entrada de audio digital almacenada 1415, 1425, 1435, 1445 en el Almacén Temporal 220 puede ser analizada por el Detector de Actividad Auditiva 510 dentro del Procesador de Registro de Audio 230. En una implementación ejemplar, el intervalo T_3 puede ser idéntico al lapso T_2 , o puede ser determinado sin ninguna relación con el lapso T_2 . Cuando el intervalo T_3 es mayor que el lapso T_2 , el Detector de Actividad Auditiva 510 puede ser configurado para acceder a, y procesar, más que el tamaño de los datos almacenados en el Almacén Temporal 220 durante un ciclo del intervalo T_1 .

La FIG. 15 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el Estado de Monitorización Pasiva de Audio S1. En este estado, puede ser deseable que la mayoría de los módulos dentro del Procesador de Registro de Audio 230 puedan estar en una modalidad eficaz en términos de energía, excepto un número mínimo de módulos requeridos para el funcionamiento de la FIG. 15. Estos módulos requeridos pueden ser los módulos mostrados en la FIG. 9B. Por lo tanto, el diagrama de flujo en la FIG. 15 puede ser mejor entendido con la FIG. 9B. Si la solicitud de inicio de suceso fue originada a partir de la Señal de Entrada S220, detectada 1515 por el Gestor de Señales de Entrada 540 cuando el dispositivo móvil está en la modalidad de reposo, puede activar el Indicador de Inicio de Suceso de 1er nivel 1540. Si la solicitud de inicio de suceso, originada a partir de la Señal Auxiliar S240, es detectada 1520 por el Analizador de Señales Auxiliares 530, puede activar el Indicador de Inicio de Suceso de 1er nivel 1540. La FIG. 15 también muestra que el Detector de Actividad Auditiva 510 analiza los datos 1530 en el Almacén Temporal 220 cada intervalo T_3 , y puede determinar si ha sido detectada o no cualquier actividad auditiva que indique que puede ser requerido un análisis adicional en profundidad. Las descripciones detalladas de realizaciones ejemplares para estas pruebas fueron previamente divulgadas en la presente solicitud junto con la FIG. 5. Si se detecta la actividad auditiva de interés, puede activar el Indicador de Inicio de Suceso de 1er nivel 1540.

Un experto en la técnica reconocerá que el orden de los bloques en la FIG. 15 es solamente con fines ejemplares en la explicación del funcionamiento del Procesador de Registro de Audio 230 y, por lo tanto, puede haber muchas variaciones que pueden ser funcionalmente equivalentes, o esencialmente equivalentes, a la FIG. 15. Por ejemplo, dicho bloque 1515 y el otro bloque 1520 pueden ser reordenados de modo que el 1520 pueda ser ejecutado primero, o pueden ser reordenados de modo que puedan no ser ejecutados en orden secuencial.

La FIG. 16 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el Estado de Monitorización Activa de Audio S2. El funcionamiento de la Unidad de Captura de Audio 215 en la FIG. 16 es muy similar al funcionamiento divulgado en la FIG. 13, excepto por unas pocas diferencias y, por lo tanto, solamente las partes diferentes pueden ser descritas en la presente memoria. El convertidor de A / D puede ser programado para mantener una mayor resolución, etiquetada como "MEDIA" en la FIG. 16, en términos de frecuencia de muestreo y / o de ancho de datos, que la resolución "BAJA" en la FIG. 13. La configuración de resolución media puede ayudar a obtener datos digitalizados de entrada de audio con mejor precisión, lo que, a su vez, puede ser beneficioso para que el Procesador de Registro de Audio 230 extraiga información de contexto S600 más fiable.

Un micrófono puede ser configurado para despertar cada intervalo T_4 ; el intervalo de vigilia del micrófono, y recoger la Entrada de Audio S270 durante el lapso T_5 ; el lapso de ACTIVACIÓN del micrófono. Los valores de T_4 o T_5 pueden ser idénticos o esencialmente similares, respectivamente, a los valores de T_1 o T_2 . Sin embargo, puede ser preferible fijar T_4 para que sea más pequeño que T_1 , porque puede ser beneficioso para el Procesador de Registro de Audio 230 extraer información de contexto S600 más precisa. En otra realización, los valores de T_4 o T_5 pueden ser predeterminados en un intervalo fijo, o pueden ser adaptados dinámicamente durante el tiempo de ejecución. En otra realización, en la cual hay una pluralidad de micrófonos en la Unidad de Micrófono 200, uno o más micrófonos pueden estar encendidos constantemente, lo que puede ser el mero caso especial en el cual T_4 es idéntico a T_5 .

La FIG. 17 es un diagrama ejemplar para almacenar una entrada de audio digital en el Almacén Temporal 220 en la unidad de Captura de Audio 215 durante el Estado de Monitorización Activa de Audio S2. La entrada almacenada de audio digital 1715, 1725, 1735, 1745 en el Almacén Temporal 220 puede ser analizada por el Identificador de Contexto 560 y la Lógica de Evaluación de Contexto 950 dentro del Procesador de Registro de Audio 230 cada intervalo T_6 . En una implementación ejemplar, el intervalo T_6 puede ser idéntico al lapso T_5 o, alternativamente, puede ser determinado sin ninguna relación con el lapso T_5 . Cuando el intervalo T_6 es mayor que el lapso T_5 , el Detector de Actividad Auditiva 510 puede ser configurado para acceder a, y procesar, los datos almacenados en el Almacén Temporal 220 durante uno o más ciclos del intervalo T_4 .

La FIG. 18 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el Estado de Monitorización Activa de Audio S2. En este estado, el Identificador de Contexto 560 dentro del Procesador de Registro de Audio 230 analiza la Entrada de Audio S270 almacenada en el Almacén Temporal 220 e identifica 1815 la información de contexto S600 en cada intervalo T_6 . La información de contexto S600 puede ser configurada para ser almacenada 1820 en una ubicación de memoria para referencia futura. La Lógica de Evaluación de Contexto 950 puede evaluar 1825 la información de contexto S600 y puede activar el Indicador de Inicio de Suceso de 2º nivel 1835, de acuerdo a diversos procedimientos internos de decisión. Tales procedimientos de decisión, por ejemplo, pueden incluir el cálculo de la suma ponderada de prioridades para la salida de algunos de, o todos, los sub-módulos divulgados en la FIG. 8, y la comparación de la suma ponderada con uno o más umbrales. La FIG. 18 también muestra el mecanismo ejemplar de activación del Indicador de Fin de Suceso S940. El Indicador de Fin de Suceso S940 puede ser activado cuando la Lógica de Evaluación de Contexto 950 no activó el Indicador de Inicio de Suceso de 2º nivel S930 durante el último lapso S que, preferiblemente, puede ser mucho más largo que el intervalo T_6 . En otra realización, el Indicador de Fin de Suceso S940 puede ser generado cuando el Administrador de Fines de Sucesos 580 detecta las señales S1052, S1053 desde el Analizador de Señales Auxiliares 530 o el Gestor de Señales de Entrada 540, según se muestra en la FIG. 10.

La FIG. 19 es un diagrama ejemplar de una realización de identificación de contexto en el Procesador de Registro de Audio 230 durante el Estado de Monitorización Activa de Audio S2. Muestra que el proceso de identificación de contexto, que es realizado por el Identificador de Contexto 560 en cada intervalo T_6 , puede ser configurado para iniciarse asincrónicamente para el intervalo T_4 . El intervalo T_6 puede ser determinado en consideración del tamaño del Almacén Temporal 220 y del equilibrio entre el consumo de energía y la precisión de la decisión. Un proceso de identificación de contexto demasiado frecuente, o un intervalo T_6 demasiado pequeño, pueden dar como resultado un consumo aumentado de energía, mientras que un proceso de identificación de contexto demasiado a menudo, o un intervalo T_6 demasiado grande, pueden dar como resultado la degradación de la precisión de la información de contexto S600.

La FIG. 20 es un diagrama de flujo de una realización de la Unidad de Captura de Audio 215 durante el Estado de Registro Activo de Audio S3, S5. El convertidor de A / D puede ser programado para mantener una mayor resolución, etiquetada como "ALTA" en la presente memoria, en términos de frecuencia de muestreo y / o de ancho de datos, en comparación con las resoluciones "BAJA" o "MEDIA" en la FIG. 13 o la FIG. 16. La configuración de alta resolución puede aumentar el tamaño de los datos de registro de audio, pero también puede ayudar a obtener datos de entrada de audio de mayor calidad. La regulación de la resolución del convertidor de A / D puede ser configurada para que sea ajustada dinámicamente de acuerdo a la señal de control desde el Procesador de Registro de Audio 230. Una descripción más detallada se presenta en una parte posterior de la presente solicitud. En el estado presente, el Procesador de Registro de Audio 230 puede estar implicado en el registro (almacenamiento) de datos de audio en la ubicación de almacenamiento deseada. El almacenamiento deseado puede residir en el dispositivo móvil local o en el sector servidor remoto, mediante una conexión cableada o inalámbrica. El registro de audio puede continuar hasta que el Indicador de Fin de Suceso S940 sea detectado por el Administrador de Fines de Sucesos 580, según se muestra en la FIG. 10.

La FIG. 21 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el Estado de Registro Activo de Audio S3. Si la solicitud de fin de suceso se originó a partir de la Señal de Entrada S220 detectada 2110 por el Gestor de Señales de Entrada 540, puede activar el Indicador de Fin de Suceso 2130. Si la solicitud de fin de suceso originada a partir de la Señal Auxiliar S240 es detectada 2115 por el Analizador de Señales Auxiliares 530, puede activar el Indicador de Fin de Suceso 2130. Si no hay ningún fin de suceso detectado a partir del Gestor de Señales de Entrada 540, o bien del Analizador de Señales Auxiliares 530, entonces el registro efectivo de audio es realizado en el Módulo central de Registro de Audio 2120. Durante el registro de audio, el

Identificador de Contexto 560 puede ser configurado para continuar identificando la información de contexto S600, y la información de contexto más antigua identificada S600, almacenada en la ubicación de memoria, puede ser actualizada por información más reciente de contexto identificada S600. La descripción detallada del funcionamiento interno del Módulo Central de Registro de Audio se presenta en la FIG. 24. Mientras el registro efectivo de audio está en marcha, la Lógica de Evaluación de Contexto 950 puede ser configurada para continuar monitorizando y analizando la Entrada de Audio S270, y activar por ello el Indicador de Fin de Suceso S940 cuando no ha sido detectada ninguna información de contexto S600 interesante durante un periodo de tiempo predeterminado. Una implementación ejemplar para el periodo de tiempo predeterminado puede incluir el uso de los datos de audio durante los S segundos más recientes. Este procedimiento de generación del Indicador de Fin de Suceso S940 puede ser mencionado como un “mecanismo de temporizador agotado”. Tales procedimientos de prueba, por ejemplo, pueden incluir el cálculo de una suma ponderada de prioridades para la salida de algunos de, o todos, los sub-módulos divulgados en la FIG. 8, y la comparación de la suma ponderada con uno o más umbrales.

La FIG. 22 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante la Etapa de Monitorización de Audio S4. El diagrama de flujo en la presente figura puede ser configurado para que sea esencialmente similar al diagrama de flujo en la FIG. 15, excepto porque el último bloque 2240 puede activar el Indicador de Inicio de Suceso, en lugar del Indicador de Inicio de Suceso de 1er nivel 1540. Esta similitud se debe al hecho de que tanto el Estado de Monitorización Pasiva de Audio S1 de la FIG. 11 como el Estado de Monitorización de Audio S4 de la FIG. 12 pueden tener idénticos propósitos: detectar los sucesos auditivos del entorno periódicamente de manera eficaz en términos de energía.

La FIG. 23 es un diagrama de flujo de una realización del Procesador de Registro de Audio 230 durante el Estado de Registro Activo de Audio S5. Debido a que el Procesador de Registro Activo, bien en S3 o en S5, puede realizar operaciones similares, el diagrama de flujo en la presente memoria también puede ser esencialmente cercano, o ser idéntico, al diagrama de flujo en la FIG. 21, con la excepción de los bloques adicionales 2300, 2305 al comienzo del diagrama de flujo. A diferencia de S3, estado donde su estado anterior siempre era el Estado de Monitorización Activa de Audio S2, en el cual el Identificador de Contexto 560 puede ser configurado para identificar la información de contexto S600 periódicamente o continuamente, según la preferencia de diseño, estos bloques adicionales 2300, 2305 pueden ser requeridos en la presente memoria porque el estado anterior de S5 es el Estado de Monitorización de Audio S4, y ninguna etapa de identificación de contexto puede ser realizada en el estado S4. Si la solicitud de fin de suceso se originó a partir de la Señal de Entrada S220 detectada 2310 por el Gestor de Señales de Entrada 540, puede activar el Indicador de Fin de Suceso 2330. Si la solicitud de fin de suceso originada a partir de la Señal Auxiliar S240 es detectada 2315 por el Analizador de Señales Auxiliares 530, puede activar el Indicador de Fin de Suceso 2330. Si no hay ningún fin de suceso detectado desde el Gestor de Señales de Entrada 540, o bien desde el Analizador de Señales Auxiliares 530, entonces el registro efectivo del audio es realizado en el Módulo Central de Registro de Audio 2320. Durante el registro de audio, el Identificador de Contexto 560 puede ser configurado para continuar identificando la información de contexto S600, y la más antigua información de contexto identificada S600, almacenada en la ubicación de memoria, puede ser actualizada por la más reciente información de contexto identificada S600. La descripción detallada del funcionamiento interno del Módulo Central de Registro de Audio se presenta en la FIG. 24. Mientras el registro efectivo del audio está en marcha, la Lógica de Evaluación de Contexto puede ser configurada para continuar monitorizando y analizando la Entrada de Audio S270 y activar por ello el Indicador de Fin de Suceso S940 cuando ninguna información de contexto S600 interesante ha sido detectada durante un periodo de tiempo predeterminado. Una implementación ejemplar para el periodo de tiempo predeterminado puede incluir el uso de datos de audio durante el más reciente lapso S. Este procedimiento de generación del Indicador de Fin de Suceso S940 puede ser llamado “mecanismo de agotamiento temporal”. Tal procedimiento de prueba, por ejemplo, puede incluir el cálculo de la suma ponderada de prioridades para la salida de algunos de, o todos, los sub-módulos divulgados en la FIG. 8, y la comparación de la suma ponderada con uno o más umbrales.

La FIG. 24 es un diagrama de flujo de una realización del módulo central de registro de audio durante los Estados de Registro Activo de Audio S3, S5. En esta realización ejemplar, los tres primeros bloques a partir del extremo superior del diagrama de flujo 2410, 2415, 2420 muestran la característica de configuración dinámica del sistema de registro inteligente de audio, de acuerdo a la información de contexto S600. La frecuencia de muestreo 2410 y / o el ancho de datos 2415 del convertidor de A / D pueden ser dinámicamente reconfigurados durante el proceso de registro de audio, en base a la información de contexto S600. La información de contexto S600, habitualmente, varía gradualmente, o incluso abruptamente, durante el curso entero del registro de audio, que puede durar más que minutos, o incluso horas. Por ejemplo, el tema del habla conversacional puede cambiar a lo largo del tiempo. El ruido de fondo, o el entorno del orador, puede cambiar, por ejemplo, cuando el orador está caminando por la calle o está en tránsito usando el transporte público. Además, el contenido de la Entrada de Audio S270 puede cambiar a lo largo del tiempo, por ejemplo, desde el habla conversacional a la música, o a la música más habla, y viceversa. Puede ser deseable usar una mayor resolución de la frecuencia de muestreo o del ancho de datos para el contenido musical, y una menor resolución de la frecuencia de muestreo o del ancho de datos para una señal principalmente de habla. En otra realización, la resolución puede ser configurada para que sea distinta de acuerdo a la característica del contenido vocal. Por ejemplo, el sistema puede ser configurado para usar una resolución distinta para la comunicación de negocios, en comparación con una conversación personal entre amigos. Los bloques 2410, 2415, 2420 para la regulación dinámica de las configuraciones del convertidor de A / D y la selección dinámica de la

ubicación de memoria, de acuerdo a la información de contexto S600, pueden ser resituados en orden distinto entre los mismos, o a diferencia de otros bloques en el diagrama de flujo, dentro del ámbito del principio general divulgado en la presente memoria.

- 5 El sistema también puede ser configurado para seleccionar dinámicamente la ubicación de memoria 2420 en base a la información de contexto S600. Por ejemplo, el sistema puede ser configurado para almacenar los datos de registro de audio en un almacenamiento que esté remotamente conectado en el sector servidor cuando uno o más oradores durante la conversación resultan coincidir con un cierto perfil, tal como clientes comerciales principales, o cuando la Entrada de Audio S270 incluye esencialmente más música que señales de habla. En tales casos puede ser deseable usar una mayor resolución del convertidor de A / D y, por lo tanto, requerir un mayor espacio de almacenamiento.

- 15 El Procesador del Registro de Audio 230 puede luego ser configurado para leer los datos de audio 2424 desde el Almacén Temporal 220. La nueva Información de Contexto puede ser identificada 2430 a partir de los más recientes datos de audio y la nueva Información de Contexto puede ser almacenada 2435 en la memoria. En otra realización, el proceso 2430 de Identificación de Contexto, o el proceso de resguardo 2434 de la información de contexto S600, puede ser omitido o resituado en un orden distinto, a diferencia de los otros bloques en el diagrama de flujo dentro del ámbito del principio general divulgado en la presente memoria.

- 20 El Procesador de Registro de Audio 230 puede ser configurado para determinar 2440 si es deseable el realce de la señal de Entrada de Audio S270 o, en tal caso, qué tipos de procesamiento de realce pueden ser deseables antes de que la señal procesada sea almacenada en la memoria seleccionada. La determinación puede estar basada en la información de contexto S600, o pre-configurada automáticamente por el sistema, o manualmente por el usuario. Tal procesamiento de realce puede incluir la cancelación del eco acústico (AEC), la recepción de realce de voz (RVE), la cancelación del ruido activo (ANC), la supresión del ruido (NS), el control de ganancia acústica (AGC), el control de volumen acústico (AVC) o el control de gama dinámica acústica (ADRC). En una realización, la agresividad del realce de señales puede estar basada en el contenido de la Entrada de Audio S270 o en la información de contexto S600.

- 30 El Procesador de Registro de Audio 230 puede ser configurado para determinar 2445 si es deseable la compresión de la señal de Entrada de Audio S270 o, en tal caso, qué tipos de procesamiento de compresión pueden ser deseables antes de que la señal procesada sea almacenada en la ubicación de memoria seleccionada. La determinación puede estar basada en la información de contexto S600, estar automáticamente pre-configurada por el sistema, o manualmente por el usuario. Por ejemplo, el sistema puede seleccionar usar la compresión antes de que comience el registro de audio, en base a la duración esperada del registro de audio, preferiblemente en base a la información de cronograma. La selección de un procedimiento de compresión, tal como la codificación del habla o la codificación del audio, puede ser configurada dinámicamente en base al contenido de la Entrada de Audio S270, o a la información de contexto S600. A menos que se especifique lo contrario, la compresión dentro del contexto de la presente solicitud puede significar codificación de origen, tal como codificación / decodificación del habla y codificación / decodificación del audio. Por lo tanto, debería ser obvio para un experto en la técnica que la compresión puede ser usada de forma intercambiable como codificación, y la descompresión puede ser usada de forma intercambiable como decodificación. Los parámetros de codificación, tales como la velocidad de bits, la modalidad de codificación o el número de canal, también pueden ser dinámicamente configurados en base al contenido de la Entrada de Audio S270 o a la información de contexto S600.

- 45 La FIG. 25 es un diagrama de una realización de un control de ACTIVACIÓN y DESACTIVACIÓN de un micrófono único, de acuerdo al control convencional de micrófono. Cuando un dispositivo móvil está en la modalidad de reposo 2550, un micrófono, y los bloques relacionados, requeridos para el funcionamiento del micrófono, tales como un convertidor de A / D, están habitualmente apagados 2510. Un micrófono y sus bloques relacionados están habitualmente encendidos 2520 solamente durante el uso activo de un dispositivo móvil para una aplicación que requiera el uso de un micrófono, tal como una llamada de voz o una grabación de vídeo.

- La FIG. 26 es un diagrama de una primera realización de un control de ACTIVACIÓN y DESACTIVACIÓN de un micrófono único. A diferencia de la FIG. 25, un micrófono puede ser configurado para estar selectivamente ACTIVADO 2520 incluso durante el periodo en que un dispositivo móvil está en la modalidad de reposo 2550. Un micrófono puede ser configurado para estar selectivamente ACTIVADO de acuerdo a la información de contexto S600 de la Entrada de Audio S270. En una realización, esta característica puede ser deseable para el Estado de Monitorización Pasiva de Audio S1, el Estado de Monitorización Activa de Audio S2, o el Estado de Monitorización de Audio S4.

- 60 La FIG. 27 es un diagrama de una segunda realización de un control de ACTIVACIÓN y DESACTIVACIÓN DE micrófono único. A diferencia de la FIG. 26, un micrófono puede estar configurado para estar regularmente ACTIVADO 2700 durante el periodo en que un dispositivo móvil está en la modalidad de reposo 2550. En tal caso, el consumo de energía del sistema puede aumentar mientras un micrófono esté encendido. En una realización, esta característica puede ser aplicable al Estado de Monitorización Pasiva de Audio S1, el Estado de Monitorización Activa de Audio S2, el Estado de Monitorización de Audio S4 o el Estado de Registro Activo de Audio S3 S5.

La FIG. 28 es un diagrama de una primera realización del control de ACTIVACIÓN y DESACTIVACIÓN de múltiples micrófonos. En una realización, uno o más micrófonos pueden ser configurados para funcionar de manera similar al sistema convencional. En otras palabras, uno o más micrófonos solamente pueden ser encendidos durante una llamada activa de voz o durante la grabación de vídeo, u otras aplicaciones cualesquiera que requieran el uso activo de uno o más micrófonos en respuesta a la selección manual del usuario. Sin embargo, los otros micrófonos pueden ser configurados para estar ACTIVADOS intermitentemente. Solamente se presentan dos micrófonos en la figura, con fines ejemplares, pero el mismo concepto de control de micrófono puede ser aplicado a más de dos micrófonos.

La FIG. 29 es un diagrama de una segunda realización del control de ACTIVACIÓN y DESACTIVACIÓN de múltiples micrófonos. A diferencia de la FIG. 28, uno o más micrófonos pueden ser configurados para funcionar de manera similar a un sistema convencional, de manera que uno o más micrófonos solamente puedan ser encendidos durante una llamada de voz activa o durante una grabación de vídeo, u otras aplicaciones cualesquiera que requieran el uso activo de uno o más micrófonos en respuesta a la selección manual del usuario. Sin embargo, los otros micrófonos pueden ser configurados para estar ACTIVADOS constantemente. En tal caso, el consumo de energía del sistema puede aumentar mientras un micrófono esté encendido. Solamente se presentan dos micrófonos en la figura con fines ejemplares, pero el mismo concepto de control de micrófono puede ser aplicado a más de dos micrófonos.

La FIG. 30 es un diagrama de una realización del control de un número de micrófonos activos, de acuerdo a la presente solicitud, en la cual el número activo de micrófonos puede ser controlado dinámicamente de acuerdo a la información de contexto S600. Con fines ejemplares, se supone que el número máximo de micrófonos disponibles es tres y que también es el número máximo de micrófonos que pueden estar encendidos durante el Estado de Monitorización Pasiva de Audio S1, el Estado de Monitorización Activa de Audio S2 o el Estado de Monitorización de Audio S4. Sin embargo, la selección de un número distinto de micrófonos todavía puede estar dentro el ámbito de la presente divulgación. Durante el Estado de Monitorización Pasiva de Audio S1 o el Estado de Monitorización de Audio S4, un micrófono puede ser configurado para ser encendido periódicamente, para que pueda monitorizar un suceso auditivo del entorno. Por lo tanto, durante estos estados, el número activo de micrófonos puede cambiar, preferiblemente entre cero y uno. Durante el Estado de Monitorización Activa de Audio S2, el número activo de micrófonos puede continuar cambiando, preferiblemente entre cero y uno, pero el intervalo entre el periodo de ACTIVACIÓN, T_4 , puede ser configurado para que sea mayor que el del Estado de Monitorización Pasiva de Audio S1 o el del Estado de Monitorización de Audio S4, T_1 .

Durante el Estado de Registro Activo de Audio S3 S5, el número de micrófonos activos puede ser configurado para cambiar dinámicamente de acuerdo a la información de contexto S600. Por ejemplo, el número activo de micrófonos puede ser configurado para aumentar desde uno 3045 hasta dos 3050, al detectar información de contexto S600 específica o información de contexto S600 de alta prioridad. En otro ejemplo, el número de micrófonos puede ser configurado para aumentar cuando las características del ruido de fondo cambian de estáticas a no estáticas, o desde el nivel leve al nivel agudo. En tal caso, un procedimiento de supresión de ruido basado en múltiples micrófonos puede ser capaz de aumentar la calidad de la Entrada de Audio S270. El aumento, o la disminución, del número de micrófonos activos también puede estar basado en la calidad de la Entrada de Audio S270. El número de micrófonos puede aumentar con la calidad de la Entrada de Audio S270; por ejemplo, de acuerdo a la razón entre señal y ruido (SNR) de la Entrada de Audio S270, se degrada por debajo de un cierto umbral.

El almacenamiento del registro de audio puede ser configurado para cambiar dinámicamente entre el almacenamiento local y el almacenamiento remoto durante el proceso de registro efectivo de audio, o después de completarse el registro de audio. Por ejemplo, la FIG. 31 muestra una realización de la selección de ubicación de almacenamiento, en la cual la selección puede ser controlada de acuerdo a la prioridad predefinida de la información de contexto S600. Esta selección puede ser realizada antes del inicio del registro de audio, o después de completarse el registro de audio. Por ejemplo, la información de contexto puede estar pre-configurada para que tenga un nivel distinto de prioridad. Luego, antes del inicio de cada registro de audio, el almacenamiento puede ser seleccionado de acuerdo a la comparación entre las características de la información de contexto S600 durante algún periodo de ventana y uno o más umbrales predefinidos. En otra realización, la selección del almacenamiento a largo plazo puede ser decidida después de completarse cada registro de audio. El registro inicial de audio puede ser almacenado por omisión, por ejemplo, dentro del almacenamiento local con fines de almacenamiento a corto plazo. Al completarse un registro de audio, el registro de audio puede ser analizado por el Procesador de Registro de Audio 230 a fin de determinar la ubicación de almacenamiento a largo plazo para el registro de audio. Cada registro de audio puede tener asignada una prioridad antes o después de completarse el registro de audio. La selección de almacenamiento a largo plazo puede ser configurada para basarse en la prioridad del registro de audio. La FIG. 31 muestra un sistema ejemplar en el cual el registro de audio con información de contexto de menor prioridad es almacenado en almacenamiento local, mientras que el registro de audio con información de contexto de mayor prioridad es almacenado en un almacenamiento en red. Debería observarse que el registro de audio con información de contexto de menor prioridad puede ser almacenado en un almacenamiento de red, o que el registro de audio con información de contexto de mayor prioridad puede ser almacenado en un almacenamiento local dentro del ámbito de la presente divulgación.

La FIG. 32 muestra una realización de la selección de ubicación de almacenamiento, en la cual la selección puede ser controlada dinámicamente de acuerdo a la prioridad de la información de contexto S600 durante el Estado de Registro Activo de Audio S3, S5. A diferencia de la FIG. 31, la selección de almacenamiento puede ser conmutada dinámicamente durante el procesamiento del registro efectivo del audio, de acuerdo a la información de contexto S600, al espacio de memoria disponible o a la calidad del canal entre un dispositivo móvil y un servidor remoto.

La FIG. 33 es un diagrama de una realización de la configuración del tiempo de agotamiento del almacenamiento, en la cual el tiempo de agotamiento puede ser controlado de acuerdo a la prioridad predefinida de la información de contexto S600. El registro de audio almacenado en los almacenamientos puede ser configurado para que sea borrado por selección manual del usuario, o agotado automáticamente por un mecanismo que puede estar basado en el tiempo de agotamiento predefinido. Cuando un registro de audio se agota, el registro de audio agotado puede ser configurado para ser borrado o desplazado a un lugar de almacenamiento temporal, tal como el "Recipiente de reciclaje". El registro de audio agotado puede ser configurado para ser comprimido, si no fue comprimido en el momento de la grabación. En el caso en que ya estuviera codificado en el momento de la grabación, puede ser trans-codificado usando un formato de codificación, o parámetros de codificación, que pudieran permitir una mayor compresión, dando como resultado un tamaño de registro de audio más compacto.

La configuración del tiempo de agotamiento puede ser determinada en el momento del registro de audio, o después de completarse el audio. En una realización, cada registro de audio puede tener asignado un valor de prioridad de acuerdo a las características o las estadísticas de la información de contexto S600 del registro de audio. Por ejemplo, el registro de audio #1 3340 en la FIG. 33 puede tener menor prioridad que el registro de audio #3 3320. En una implementación ejemplar, puede ser deseable fijar el tiempo de agotamiento del registro de audio #1, ET₁, menor que el tiempo de agotamiento del registro de audio #3, ET₃. Como ejemplo, ET₁ puede ser fijado en "1 semana" y ET₃ puede ser fijado en "2 semanas". Es generalmente deseable tener un tiempo de agotamiento para un registro de audio, en proporción a la prioridad del registro de audio. Pero debería observarse que un registro de audio que tenga una prioridad distinta no necesariamente debe tener siempre una configuración distinta del tiempo de agotamiento.

La FIG. 34 es un diagrama de una realización del arranque, etapa por etapa, de los bloques dentro del sistema de registro inteligente de audio, en la cual el número de bloques activos y el consumo total de energía de los mismos pueden ser controlados dinámicamente de acuerdo a cada estado. Durante el Estado de Monitorización Pasiva de Audio S1, uno o más micrófonos pueden ser configurados para despertar periódicamente a fin de recibir la Entrada de Audio S270. A fin de realizar esta operación de recepción, el sistema puede ser configurado para despertar una parte del sistema y, por ello, el número de bloques activos o bien, de forma intercambiable, el número de bloques de arranque, del sistema, aumentados a N1 en la FIG. 34. Durante el Estado de Monitorización Activa de Audio S2, uno o más bloques adicionales pueden ser configurados para despertar, además de N1, lo que hace que el número total de bloques activos sea N2 durante los períodos en que uno o más micrófonos están activos 3420. Por ejemplo, el Identificador de Contexto 560 y la Lógica de Evaluación de Contexto 950 pueden ser configurados para despertar, como fue ejemplificado en la FIG. 9B. Durante el Estado de Registro Activo de Audio S3, es probable que al menos algunos bloques más puedan necesitar despertarse, además de los N2, lo que, a su vez, hace que el número total de bloques activos durante el Estado de Registro Activo de Audio S3 sea N3. El número de referencia de bloques activos 3425 durante el Estado de Monitorización Activa de Audio S2 se fija como N1 en la FIG. 34, que resulta ser el mismo número de bloques activos durante el Estado de Monitorización Pasiva de Audio S1, pero debería ser obvio para los expertos en la técnica que esto puede ser configurado para que sea distinto en otra realización dentro del ámbito de la presente divulgación. El número de bloques activos para el Estado de Monitorización de Audio S4 o el Estado de Registro Activo de Audio S5 puede ser implementado, respectivamente, de manera similar al Estado de monitorización Pasiva de Audio S1 o al Estado de Registro Activo de Audio S3.

La FIG. 35 es un diagrama de una realización del control de precisión del convertidor de A / D, en la cual la precisión puede ser configurada de acuerdo a cada estado predeterminado, o controlada dinámicamente en correspondencia con la información de contexto S600. La unidad convertidora de A / D durante el Estado de Monitorización Pasiva de Audio S1 puede ser configurada para que tenga una regulación de baja resolución, etiquetada como "BAJA" en la FIG. 35, mientras que puede ser configurada para que tenga una regulación de resolución media, regulación "MEDIA", o una regulación de mayor resolución, regulación "ALTA", para el Estado de Monitorización Activa de Audio S2 o el Estado de Registro Activo de Audio S3, respectivamente. Este mecanismo puede ayudar a ahorrar en el consumo de energía o en el uso de memoria, permitiendo configuraciones optimizadas para cada estado. En otra realización, la regulación del convertidor de A / D durante las etapas del Estado de Monitorización Pasiva de Audio S1 y del Estado de Monitorización Activa de Audio S2 puede ser configurada para que tenga la misma resolución. Alternativamente, la regulación del convertidor de A / D durante las etapas del Estado de Monitorización Activa de Audio S2 y del Estado de Registro Activo de Audio S3 puede ser configurada para que tenga la misma resolución.

La regulación de la precisión para la unidad convertidora de A / D puede ser configurada para que cambie dinámicamente durante el Estado de Registro Activo de Audio S3, en base a la información de contexto S600. La FIG. 35 muestra que el cambio dinámico puede ser configurado para que tenga efecto durante el lapso entero, o bien parcial, 3540, durante el proceso de registro activo de audio. Se supone que la configuración de la precisión por omisión para el Estado de Registro Activo de Audio S3 es "Alta" 3520. Cuando hay un cambio significativo en

términos de la prioridad de la información de contexto S600, la regulación de la precisión puede ser reducida a las configuraciones “Media” 3535 o “Baja” 3525. Por ejemplo, el cambio de configuración de precisión puede ser iniciado por el cambio de la clasificación de contenido, que es un subconjunto de la información de contexto S600, de “Música” a “Habla” o de “Habla” a “Música”. Alternativamente, puede ser iniciado por el cambio del nivel del ruido de fondo o el tipo de ruido de la Entrada de Audio S270. En otra realización, puede ser iniciado por el tamaño de memoria disponible en el almacenamiento local, o la calidad del canal entre un dispositivo móvil y un servidor remoto.

La FIG. 36 es un diagrama de una realización del control de realce de señales de entrada de audio, en la cual el realce puede ser configurado dinámicamente de acuerdo a la información de contexto S600. Con fines ejemplares, se supuso que hay varios niveles de realce de señal: ningún realce, realces de nivel bajo, nivel medio y de nivel alto. Durante el Estado de Registro Activo de Audio S3, S5, el nivel de realce de la señal de audio puede ser configurado para que se ajuste dinámicamente de acuerdo a la información de contexto S600. Por ejemplo, las características o el nivel del ruido de fondo pueden ser usados para activar el cambio del nivel de realce de la señal de audio. Cuando el nivel del ruido de fondo es significativamente mayor o las características del nivel del ruido de fondo cambian significativamente desde el ruido de tipo estático al ruido de tipo no estático, la regulación del realce de la señal de audio puede ser configurada para que cambie desde el realce de bajo nivel, o sin realce, al realce de nivel medio, o incluso al realce de alto nivel. Por ejemplo, un usuario puede estar dentro de la estación del Metro esperando que llegue su tren cuando el sistema de registro inteligente de audio podría estar en el Estado de Registro de Audio S3, S5, registrando activamente la Entrada de Audio S270. Cuando el tren está llegando a, o partiendo de, una plataforma, el nivel de ruido a menudo superaba un cierto umbral, más allá del cual el habla conversacional normal es difícil de entender. Al detectar el significativo cambio del nivel o del tipo del ruido de fondo, o al detectar el cambio importante de escena auditiva, el sistema de registro inteligente de audio puede reconfigurar en consecuencia las regulaciones del realce de señales de audio. El cambio de regulación del realce de la señal de audio puede ser seguido o precedido por el número activo de micrófonos.

La FIG. 37 es un diagrama de una realización del control de parámetros de compresión de audio, en la cual la compresión puede ser dinámicamente configurada de acuerdo a la información de contexto S600. Con fines ejemplares, se supuso que hay varios niveles de compresión: ninguna compresión, compresiones “Baja”, “Media” y “Alta”. Durante el Estado de Registro Activo de Audio S3, S5, el nivel de compresión de la señal de audio puede ser configurado para que sea ajustado dinámicamente de acuerdo a la información de contexto S600. Por ejemplo, el cambio de la modalidad de compresión puede ser iniciado por el cambio de la clasificación del contenido, que es un subconjunto de la información de contexto S600, de “Música” a “Habla” o de “Habla” a “Música”. Puede ser deseable usar una mayor velocidad de bits para el contenido de “Música”, mientras que puede ser deseable usar una menor velocidad de bits para el contenido de “Habla”, en el cual el ancho de banda de la señal a codificar es habitualmente mucho más estrecho que el contenido de “Música” habitual. Alternativamente, puede ser iniciado por el tamaño de memoria disponible en el almacenamiento local, o la calidad del canal entre un dispositivo móvil y un servidor remoto.

El formato de codificación puede ser configurado para que cambie asimismo de acuerdo a la información de contexto S600. La FIG. 38 es un diagrama de una realización de la selección del formato de codificación de compresión, en la cual la selección del formato de codificación de compresión, o la falta de la misma, puede ser configurada dinámicamente de acuerdo a la información de contexto S600. Con fines ejemplares, el códec de audio #1 y el códec de habla #1 fueron mostrados en la FIG. 38, pero en general el formato de codificación también puede ser configurado para que cambie entre los códecs de audio o entre los códecs del habla.

Por ejemplo, el presente códec de audio #1 3810 puede ser configurado para que sea cambiado por el códec del habla #1 3820. Al detectar el cambio principal de clasificación de señal, de “Música” a “Habla”. En otra realización, el cambio de formato de codificación, si acaso, puede ser activado solamente después de la “modalidad sin compresión” 3830 o, alternativamente, puede ser activado en cualquier momento al detectar el cambio de la información de contexto S600 predefinida, sin la “modalidad sin compresión” 3830 de por medio.

Se proporcionan diversas configuraciones ejemplares para permitir a cualquier persona experta en la técnica hacer o usar los procedimientos y otras estructuras divulgadas en la presente memoria. Los diagramas de flujo, los diagramas de bloques y otras estructuras mostradas y descritas en la presente memoria son solamente ejemplos, y otras variantes de estas estructuras también están dentro del ámbito de la divulgación. Son posibles diversas modificaciones para estas configuraciones, y los principios genéricos presentados en la presente memoria pueden ser aplicados asimismo a otras configuraciones. Por ejemplo, se subraya que el ámbito de esta divulgación no está limitado a las configuraciones ilustradas. En cambio, está expresamente contemplado, y divulgado por la presente, que las características de las distintas configuraciones particulares, según lo descrito en la presente memoria, pueden ser combinadas para producir otras configuraciones que están incluidas dentro del ámbito de esta divulgación, para cualquier caso en el cual tales características no sean incongruentes entre sí. También está expresamente contemplado y divulgado por la presente que allí donde se describe una conexión entre dos o más elementos de un aparato, pueden existir uno o más elementos intervinientes (tal como un filtro), y que allí donde se describe una conexión entre dos o más tareas de un procedimiento, pueden existir una o más tareas u operaciones intervinientes (tal como una operación de filtrado).

Las configuraciones descritas en la presente memoria pueden ser implementadas, en parte o en su totalidad, como un circuito cableado, como una configuración de circuitos elaborada en un circuito integrado específico de la aplicación, o como un programa de firmware cargado en un almacenamiento no volátil, o un programa de software cargado desde, o en, un medio legible por ordenador, tal como código legible por máquina, siendo dicho código instrucciones ejecutables por una formación de elementos lógicos tales como un microprocesador u otra unidad de procesamiento de señales digitales. El medio legible por ordenador puede ser una formación de elementos de almacenamiento, tales como la memoria semiconductora (que puede incluir, sin limitación, RAM (memoria de acceso aleatorio) dinámica o estática, ROM (memoria de solo lectura) y / o RAM flash), o la memoria ferroeléctrica, polimérica o de cambios de fase; un medio en disco tal como un disco magnético u óptico; o cualquier otro medio legible por ordenador para el almacenamiento de datos. El término "software" debería ser entendido como incluyente de código fuente, código en lenguaje ensamblador, código de máquina, código binario, firmware, macro-código, micro-código, uno o más conjuntos o secuencias cualesquiera de instrucciones ejecutables por una formación de elementos lógicos, y cualquier combinación de tales ejemplos.

Cada uno de los procedimientos divulgados en la presente memoria también puede ser realizado tangiblemente (por ejemplo, en uno o más medios legibles por ordenador, según lo enumerado anteriormente) como uno o más conjuntos de instrucciones legibles y / o ejecutables por una máquina, incluyendo una formación de elementos lógicos (p. ej., un procesador, un microprocesador, un micro-controlador u otra máquina de estados finitos). Por tanto, la presente divulgación no está concebida para limitarse a las configuraciones mostradas anteriormente, sino que, en cambio, ha de concedérsele el más amplio ámbito congruente con los principios y características novedosas divulgados de cualquier modo en la presente memoria, incluidos en las reivindicaciones adjuntas según lo presentado, que forman parte de la divulgación original.

REIVINDICACIONES

1. Un procedimiento de procesamiento de una señal de audio digital para un dispositivo móvil, comprendiendo el procedimiento:
5 recibir (S270) una señal acústica por parte de al menos un micrófono;

 convertir la señal acústica recibida en la señal de audio digital;
10 extraer (S600) al menos una información de contexto auditivo desde la señal de audio digital;

 en respuesta a detectar automáticamente un indicador de inicio de suceso (S910) en base a la información de contexto extraída, o a información de contexto adicional, realizar un registro de audio para la señal de audio digital; y
15 en respuesta a detectar automáticamente un indicador de fin de suceso (S940), finalizar el registro de audio.
2. Un aparato para procesar una señal de audio digital para un dispositivo móvil, comprendiendo el aparato:
20 medios para recibir una señal acústica, por parte de al menos un micrófono (200);

 medios para convertir (210) la señal acústica recibida en la señal de audio digital;
25 medios para extraer (560) al menos una información de contexto auditivo desde la señal de audio digital;

 medios para detectar automáticamente un indicador de inicio de suceso (570) en base a la información de contexto extraída o a información de contexto adicional;
30 medios para realizar un registro de audio (230) para una señal de audio digital, en respuesta a la detección del indicador de inicio de suceso;

 medios para detectar automáticamente (580) un indicador de fin de suceso; y
35 medios para finalizar el registro de audio para la señal de audio digital, en respuesta a la detección del indicador de fin de suceso.
3. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 2, en el que los medios para la conversión comprenden:
40 medios para transformar la señal acústica recibida en una señal eléctrica;

 medios para tomar muestras (210) de la señal eléctrica, para obtener la señal de audio digital; y
45 medios para almacenar la señal de audio digital en un almacén temporal (220),

 en el que la toma de muestras se basa en una frecuencia de muestreo y en un ancho de datos y, preferiblemente, en el que dicha al menos una información de contexto auditivo se refiere a al menos uno entre: clasificación de audio, identificación de palabras clave o identificación de orador y, además, preferiblemente, en el que dicha al menos una información de contexto auditivo está basada, al menos en parte, en uno entre: energía de señal, razón entre señal y ruido, declive espectral o frecuencia de cruce por el cero y, en particular, en el que dicha al menos una información de contexto auditivo está basada, al menos en parte, en información no auditiva.
50
- 55 4. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 3, en el que la información no auditiva comprende información de planificación o de cronograma.
5. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 2, en el que los medios para detectar el indicador de inicio de suceso comprenden:
60 medios para seleccionar (230) dicha al menos una información de contexto entre dicha al menos una información de contexto auditivo;

 medios para comparar (230) la información de contexto seleccionada con al menos un umbral predeterminado; y
65

medios para determinar (230) si el indicador de inicio de suceso ha sido detectado, en respuesta a la comparación.

- 5 6. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 2, en el que la detección del indicador de inicio de suceso está basada, al menos en parte, en información no auditiva, y en el que la información no auditiva comprende información de planificación o de cronograma y, preferiblemente, en el que los medios para realizar el registro de audio comprenden:

10 medios para actualizar al menos un parámetro relacionado con los medios para la conversión, en base, al menos en parte, a dicha al menos una información de contexto auditivo;

medios para determinar si se requiere un procesamiento adicional, en base, al menos en parte, a dicha al menos una información de contexto auditivo;

15 medios para aplicar el procesamiento adicional a la señal de audio digital, para obtener una señal de audio procesada, en respuesta a la determinación; y

medios para almacenar (220) la señal de audio procesada en un almacenamiento de memoria.
- 20 7. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 6, en el que el procesamiento adicional incluye el procesamiento de realce de señal y, preferiblemente, en el que el procesamiento de realce de señal incluye al menos uno entre: la cancelación de eco acústico, AEC, la recepción del realce de voz, RVE, la cancelación activa del ruido, ANC, la supresión del ruido, NS, el control de ganancia acústica, AGC, el control de volumen acústico, AVC, o el control de gama dinámica acústica, ADRC.
- 25 8. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 7, en el que la supresión del ruido está basada en una pluralidad de micrófonos.
- 30 9. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 6, en el que el procesamiento adicional incluye el procesamiento de compresión de señales y, preferiblemente, en el que el almacenamiento de memoria comprende una memoria local dentro del dispositivo móvil, o una memoria remota conectada con el dispositivo móvil a través de un canal inalámbrico.
- 35 10. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 9, en el que el procesamiento de compresión de señales incluye la compresión del habla o la compresión del audio y, preferiblemente, en el que al menos un parámetro de compresión está determinado en base a la información de contexto auditivo.
- 40 11. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 10, en el que dicho al menos un parámetro de compresión incluye la modalidad de compresión, la velocidad de bits o el número de canal.
- 45 12. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 9, en el que el almacenamiento de memoria se selecciona en base, al menos en parte, a la información de contexto auditivo.
13. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 2, en el que los medios para detectar el indicador de fin de suceso comprenden:

50 medios para seleccionar al menos una información de contexto entre dicha al menos una información de contexto auditivo;

medios para comparar la información de contexto seleccionada con al menos un umbral predeterminado;

y

55 medios para determinar si el indicador de fin de suceso ha sido detectado, en respuesta a la comparación.
14. El aparato para procesar una señal de audio digital de acuerdo a la reivindicación 2, en el que los medios para detectar el indicador de fin de suceso están basados, al menos en parte, en una no ocurrencia de un suceso auditivo durante un periodo de tiempo predeterminado y, preferiblemente, en el que los medios para detectar el indicador de fin de suceso están basados, al menos en parte, en información no auditiva, y en el que la información no auditiva comprende información de planificación o de cronograma.
- 60 15. Un medio no transitorio legible por ordenador, que comprende instrucciones para procesar una señal de audio digital para un dispositivo móvil que, cuando son ejecutadas por un procesador, provocan que el procesador:
- 65

al recibir una señal acústica por parte de al menos un micrófono,
convierta la señal acústica recibida en la señal de audio digital;

5 extraiga al menos una información de contexto auditivo desde la señal de audio digital;

detecte automáticamente un indicador de inicio de suceso, en base a la información de contexto extraída
o a información de contexto adicional;

10 realice un registro de audio para la señal de audio digital, en respuesta a la detección del indicador de
inicio de suceso;

detecte automáticamente un indicador de fin de suceso; y

15 finalice el registro de audio en respuesta a la detección del indicador de fin de suceso.

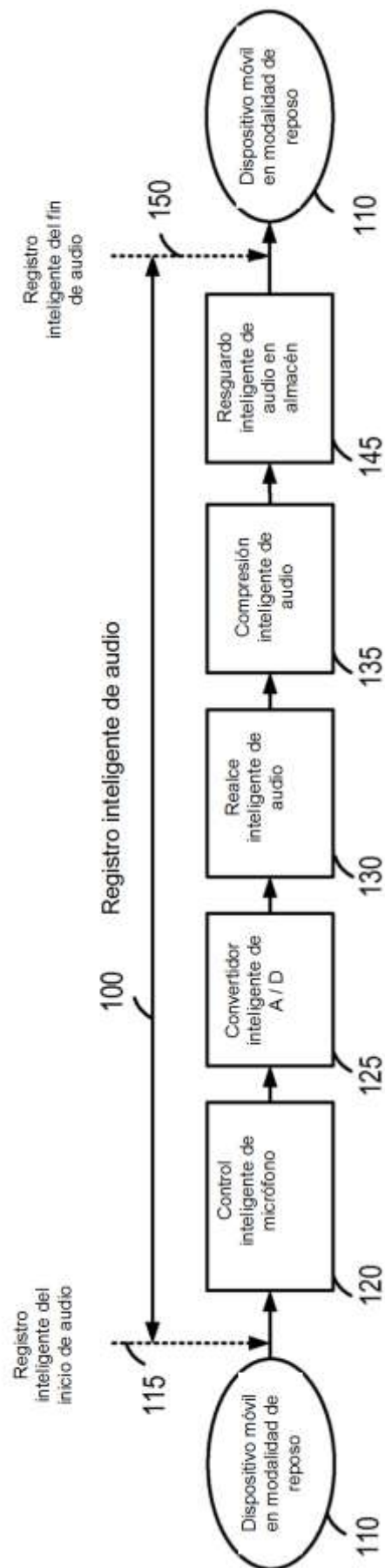


FIG. 1A

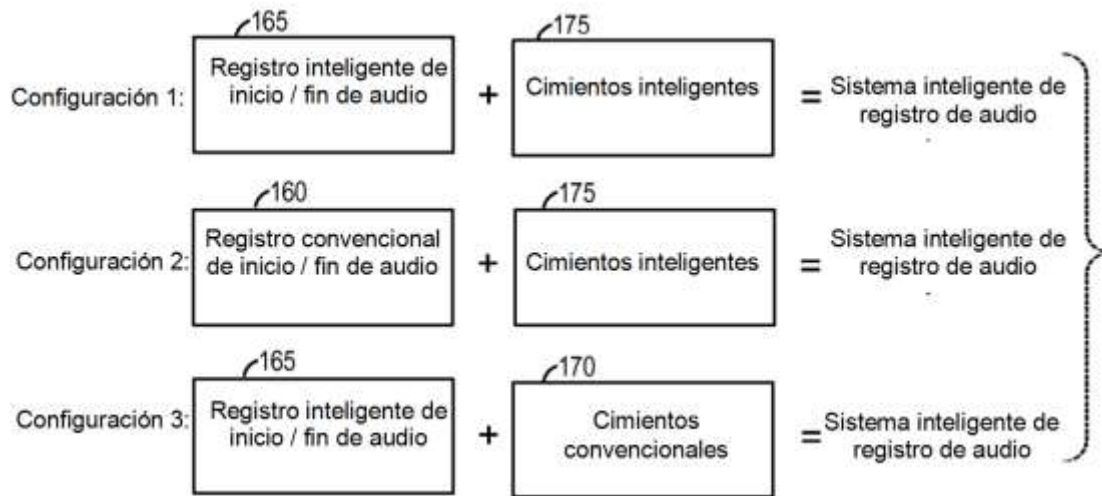
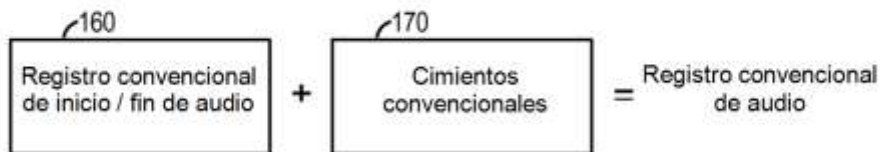


FIG. 1B



TÉCNICA ANTERIOR

FIG. 1C

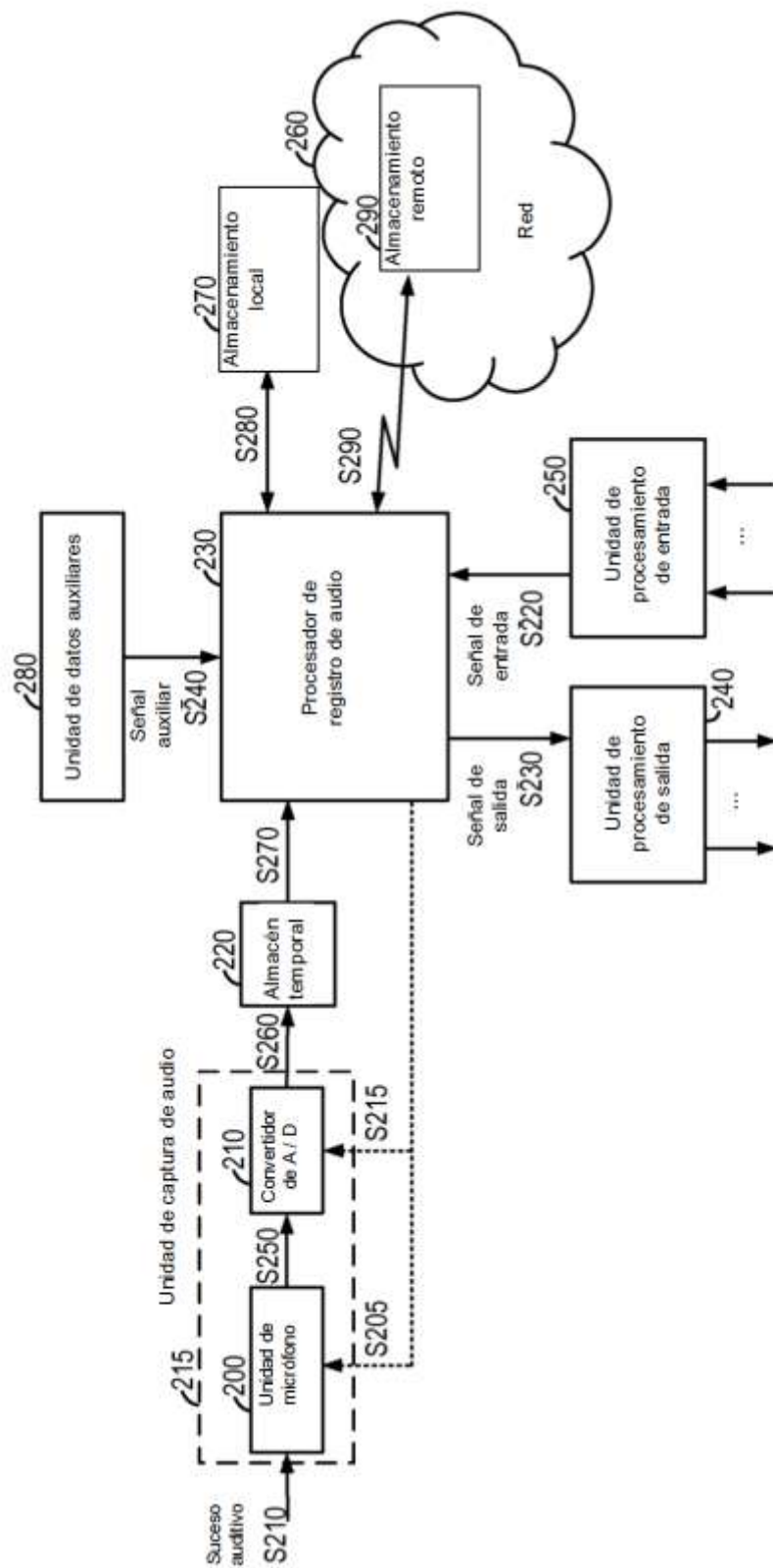


FIG. 2

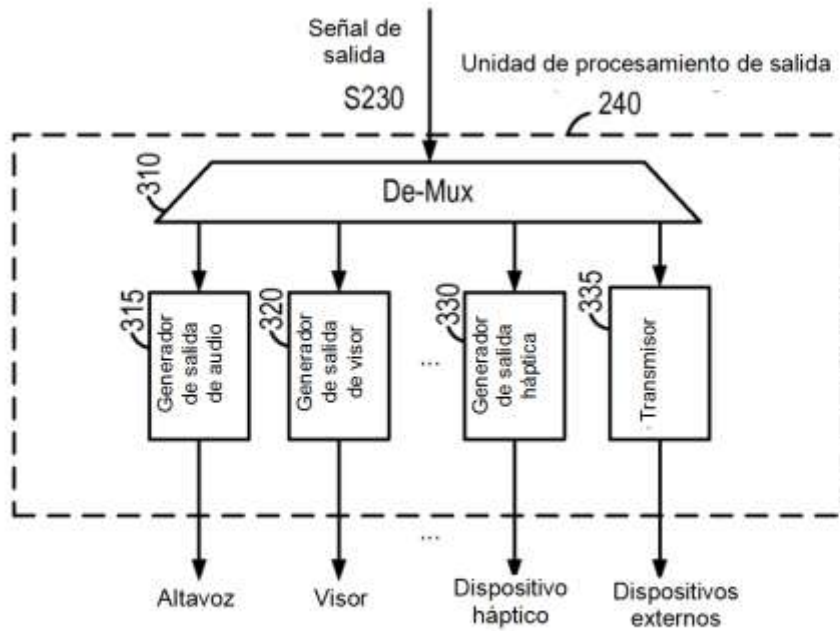


FIG. 3

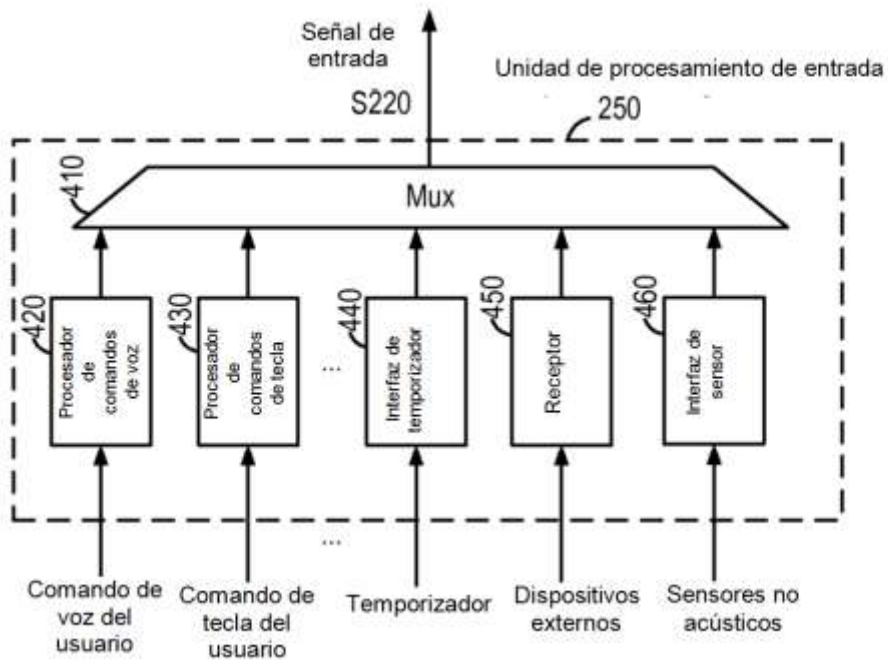


FIG. 4

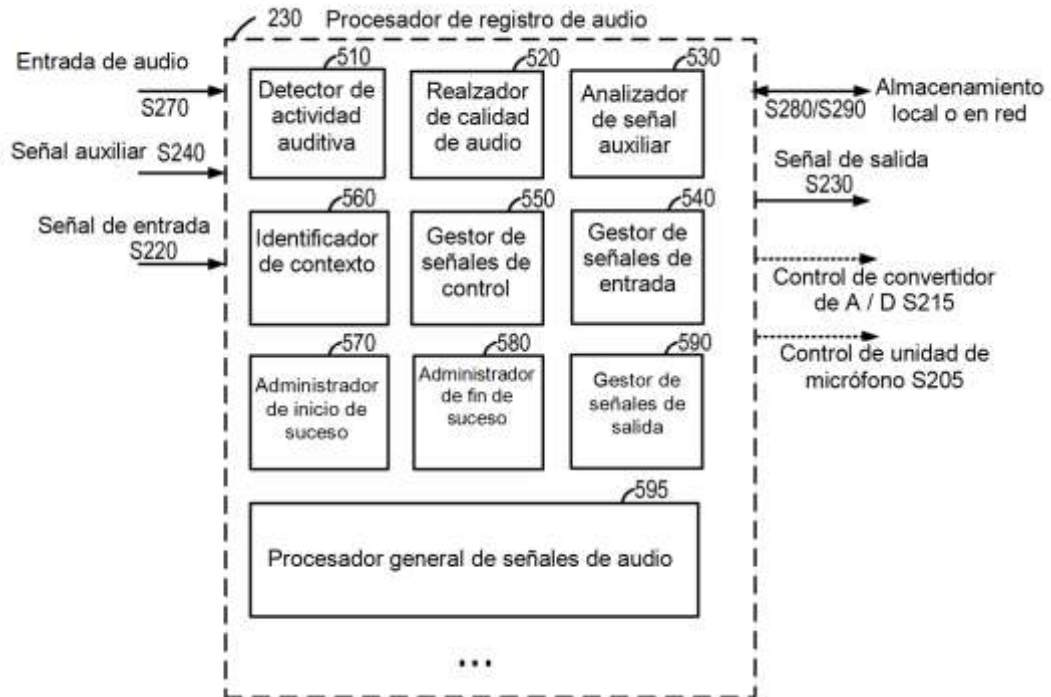


FIG. 5

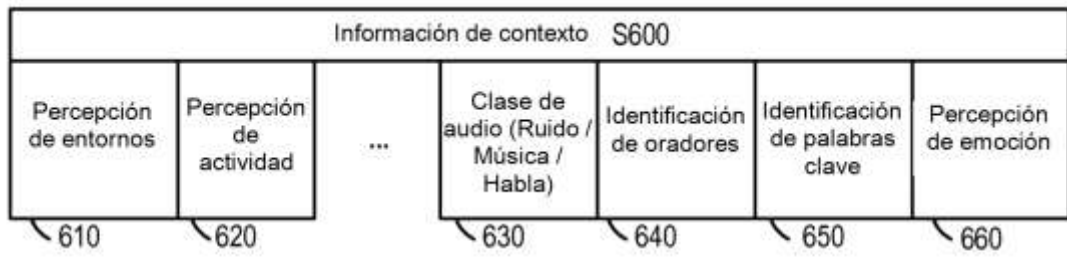


FIG. 6

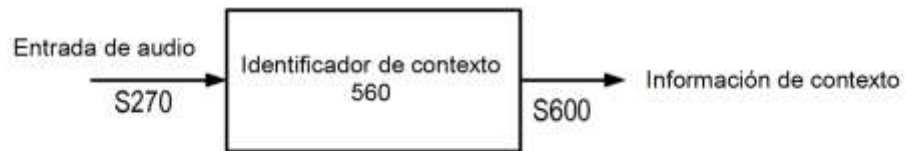


FIG. 7



FIG. 8

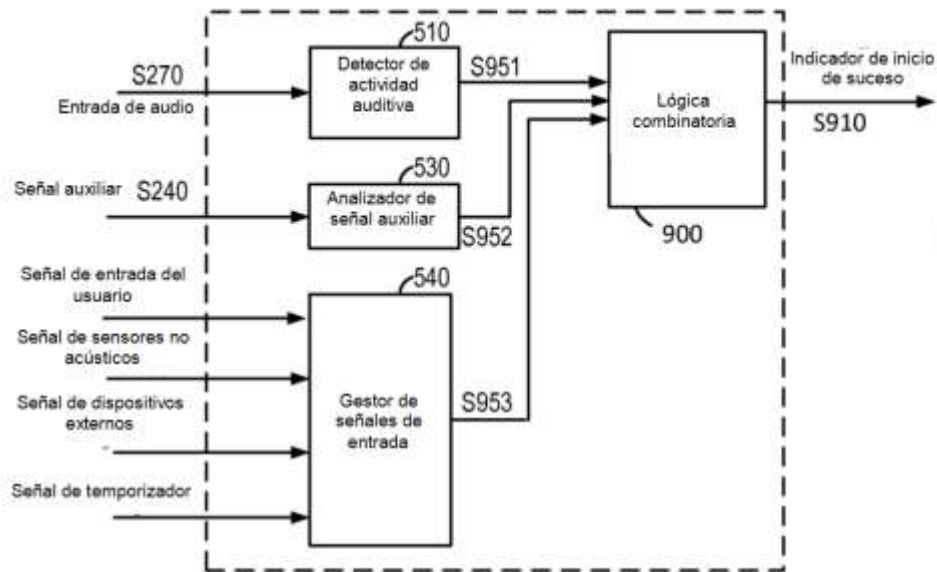


FIG. 9A

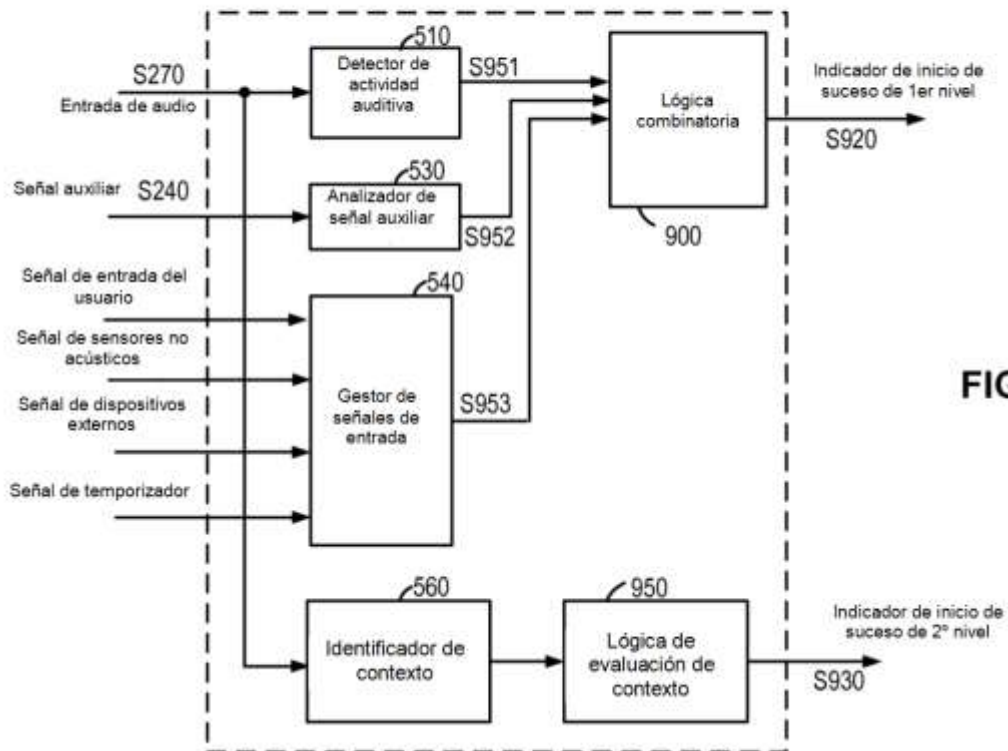


FIG. 9B

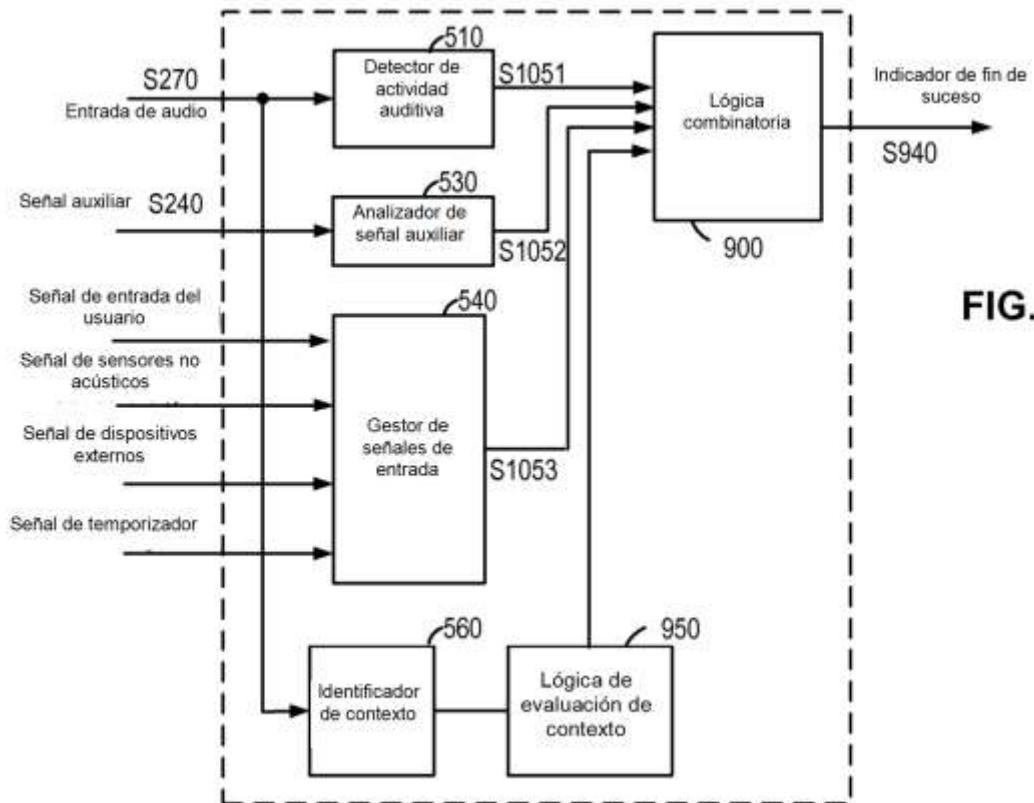
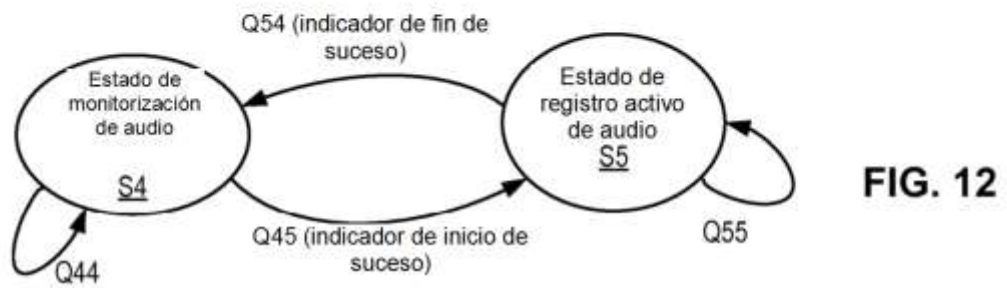


FIG. 10



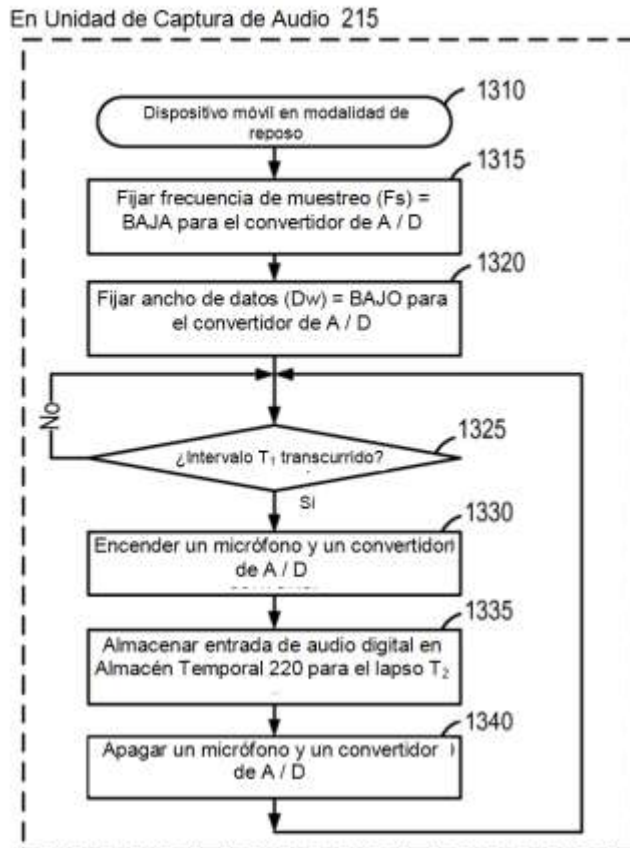


FIG. 13

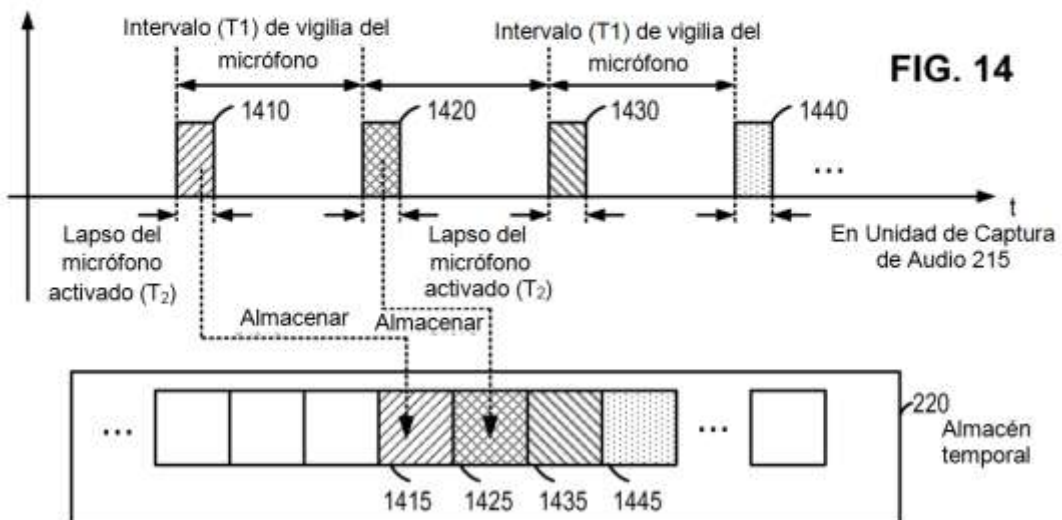


FIG. 14

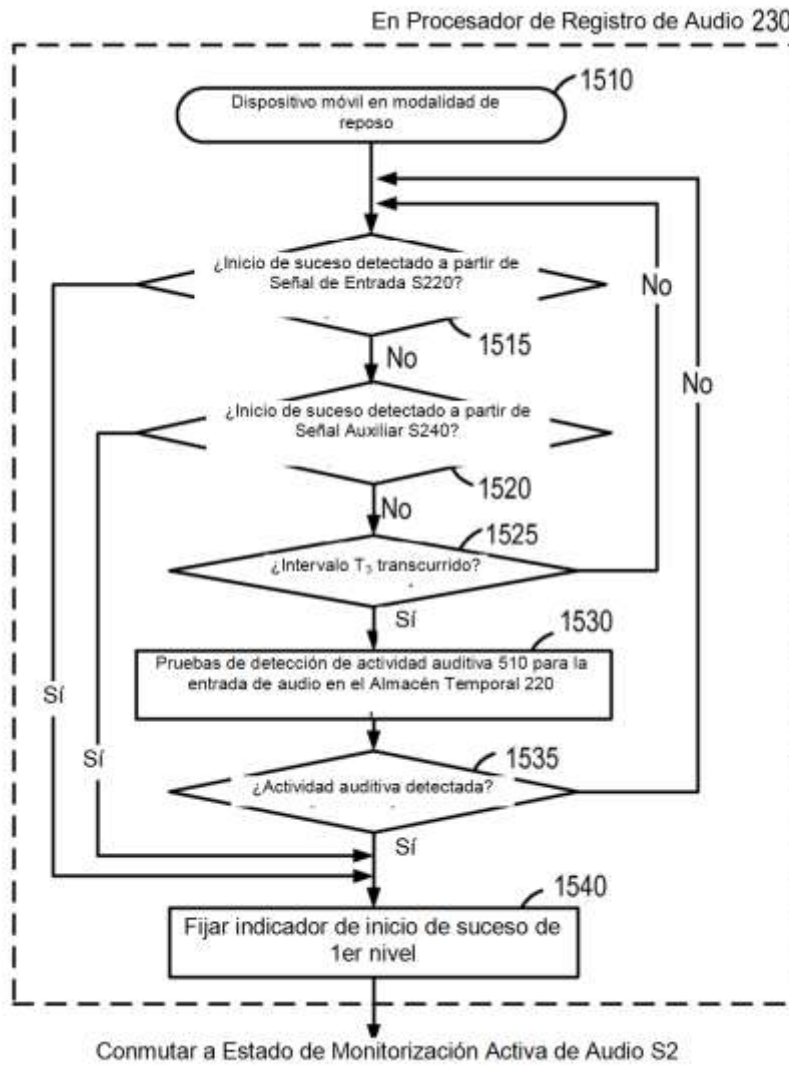
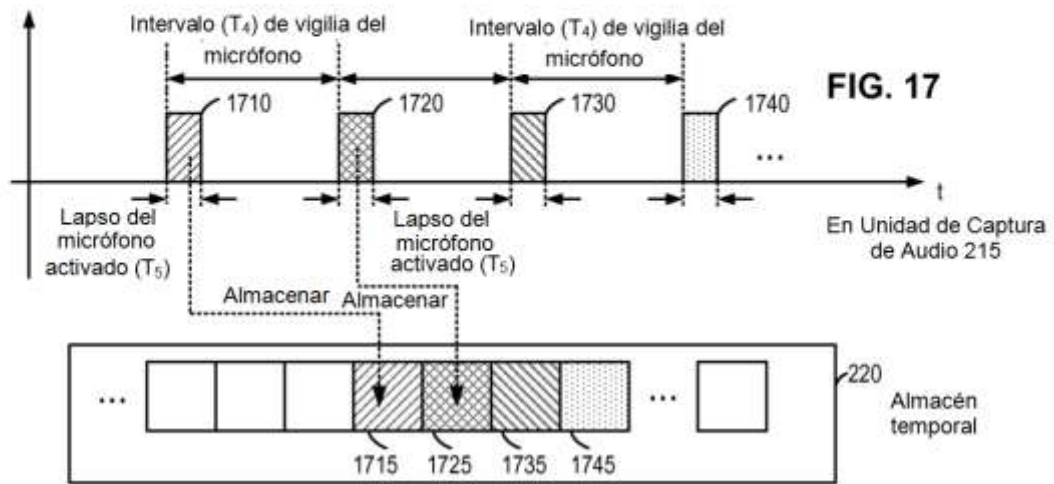
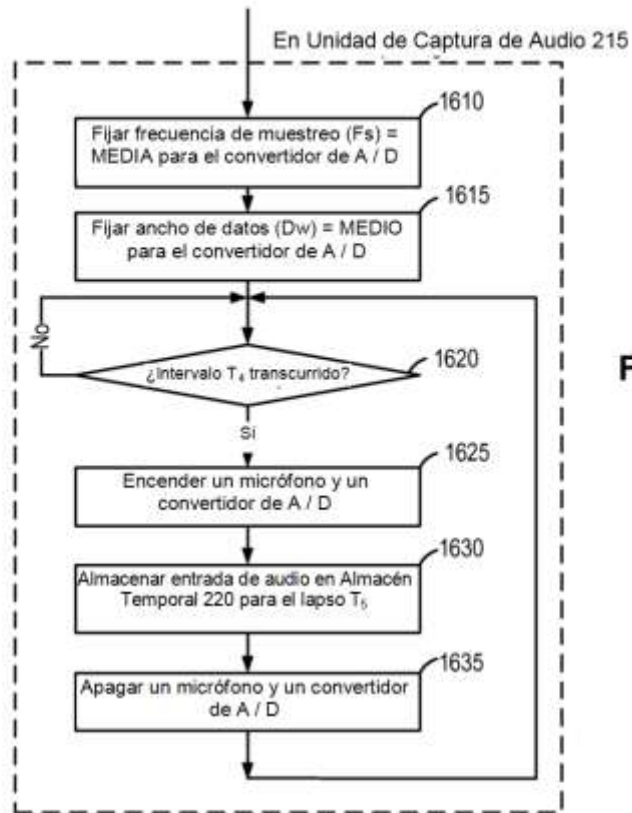
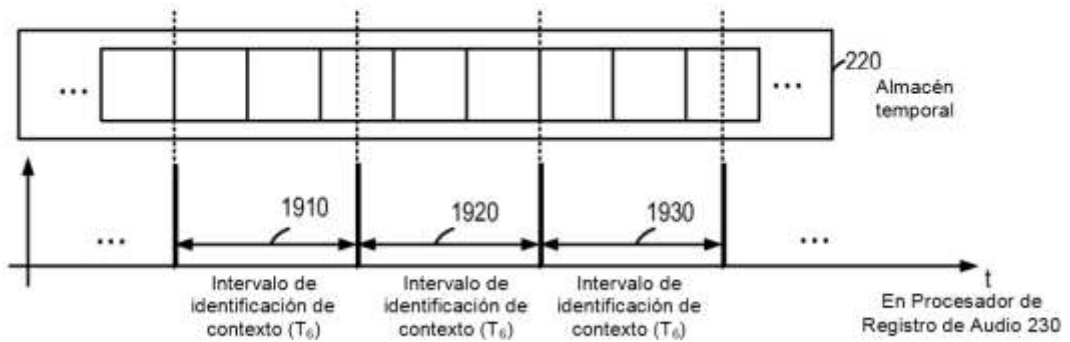
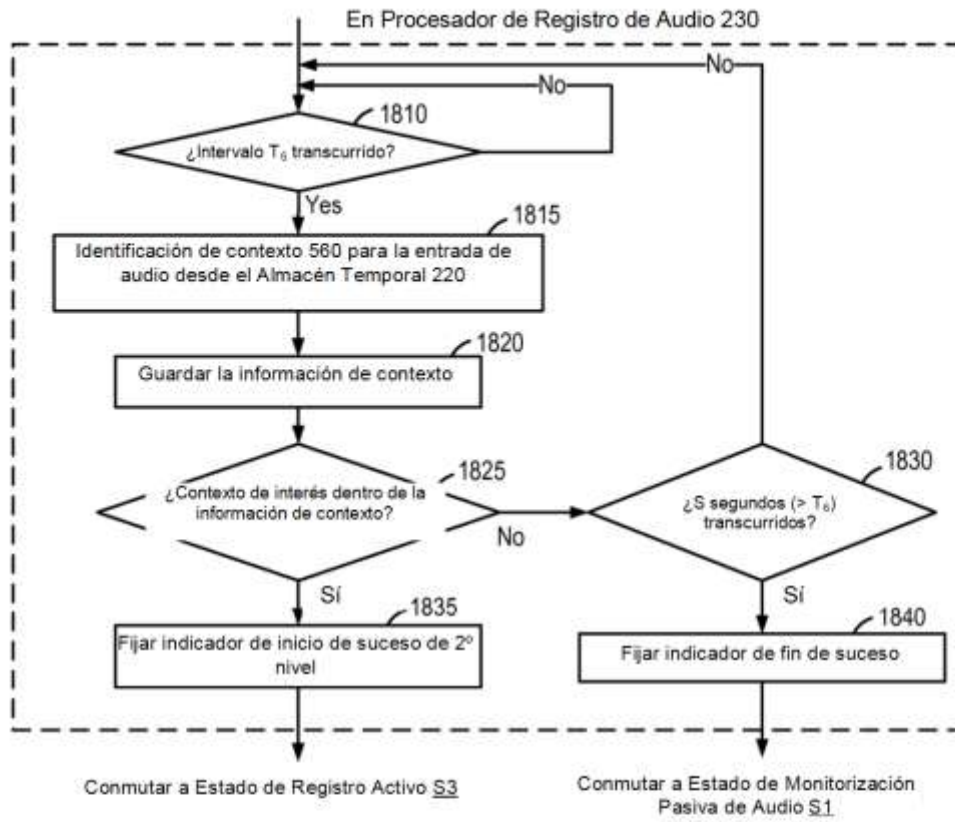


FIG. 15





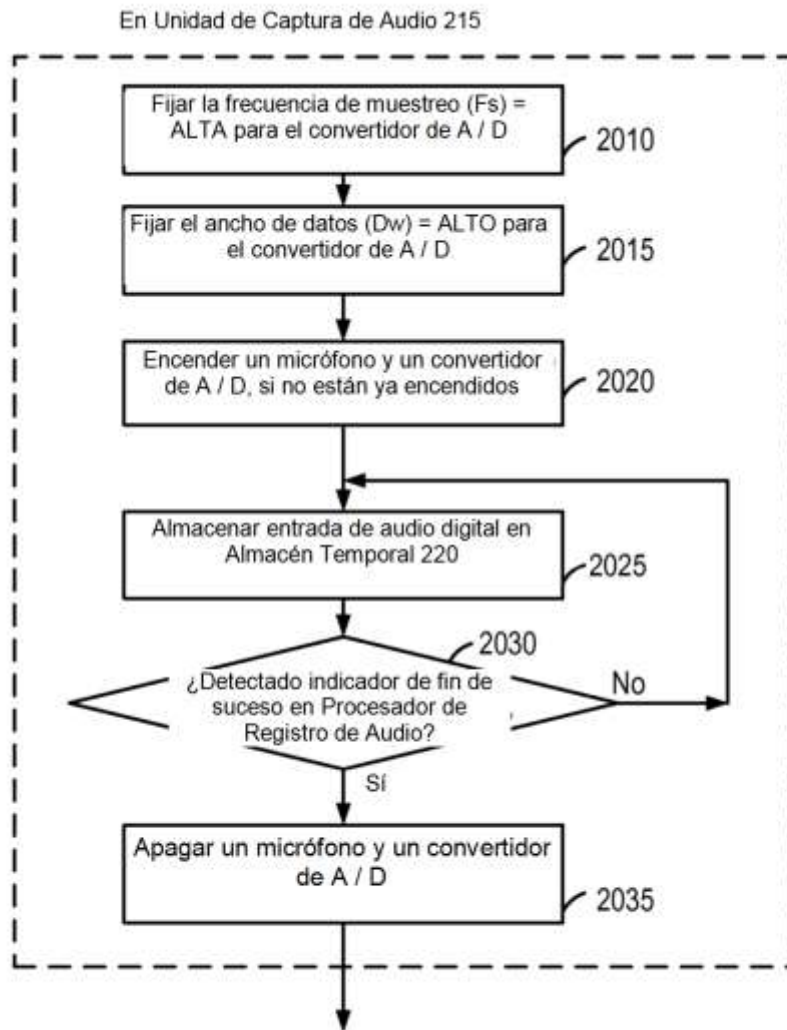


FIG. 20

(Conmutar al estado S1 en el Procesador de Registro de Audio si el estado actual es S3)
 (Conmutar al estado S4 en el Procesador de Registro de Audio si el estado actual es S5)

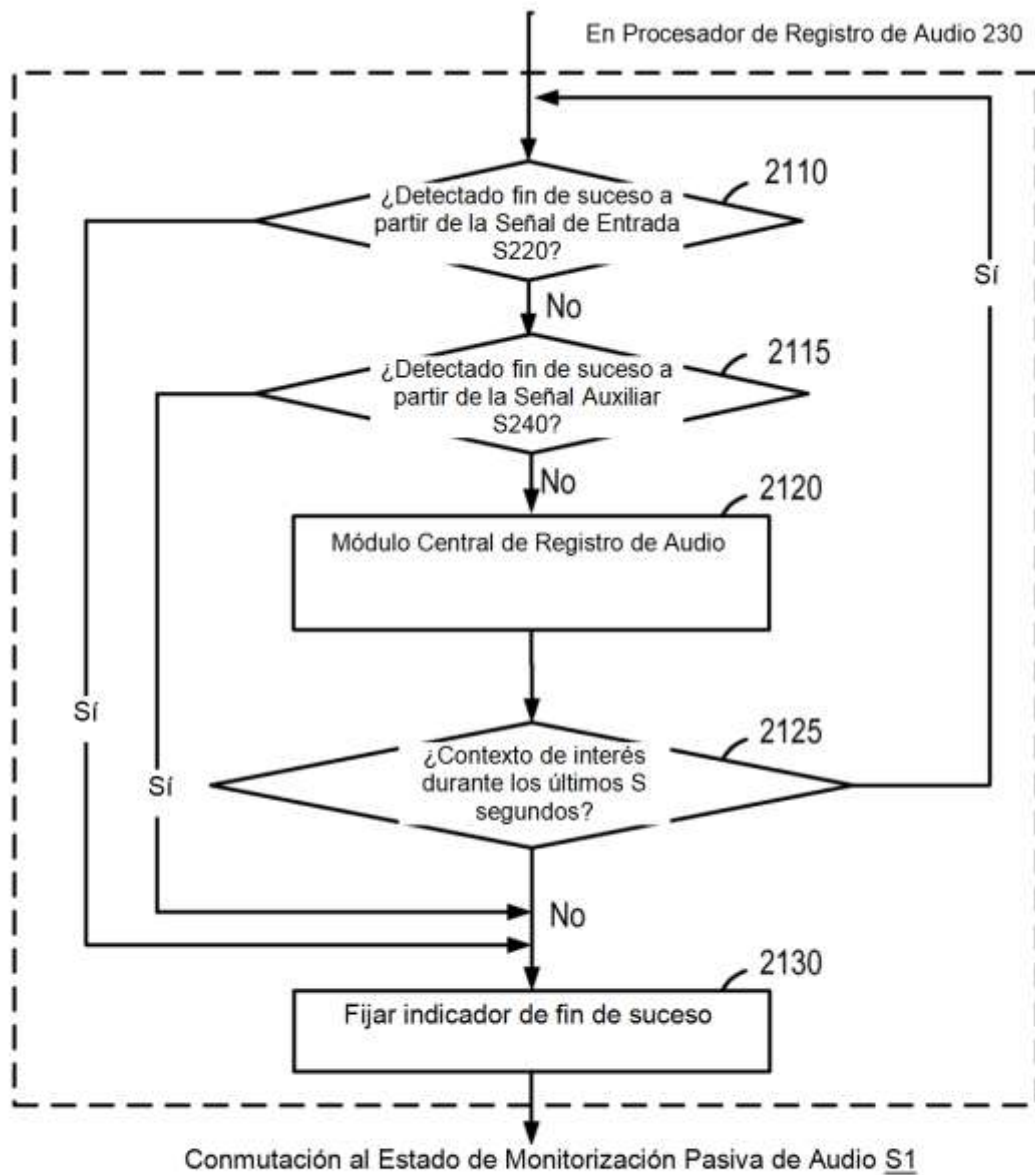


FIG. 21

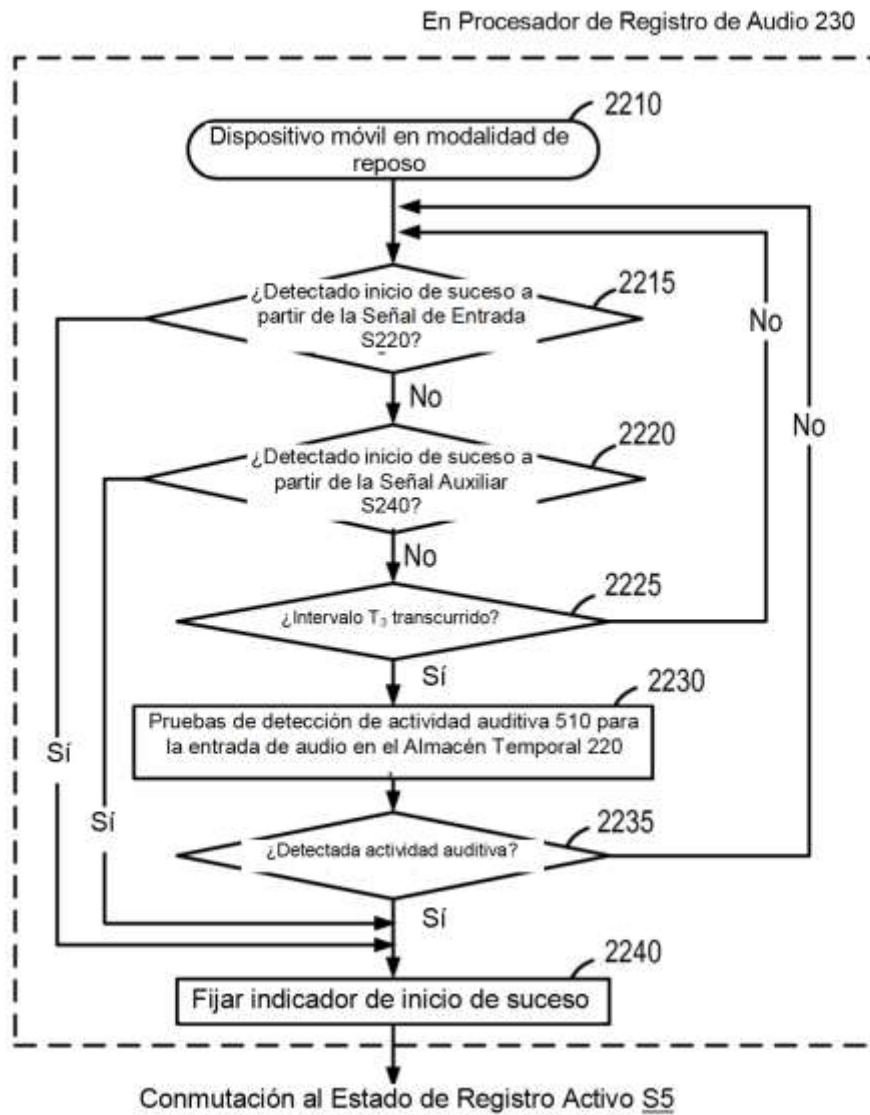


FIG. 22

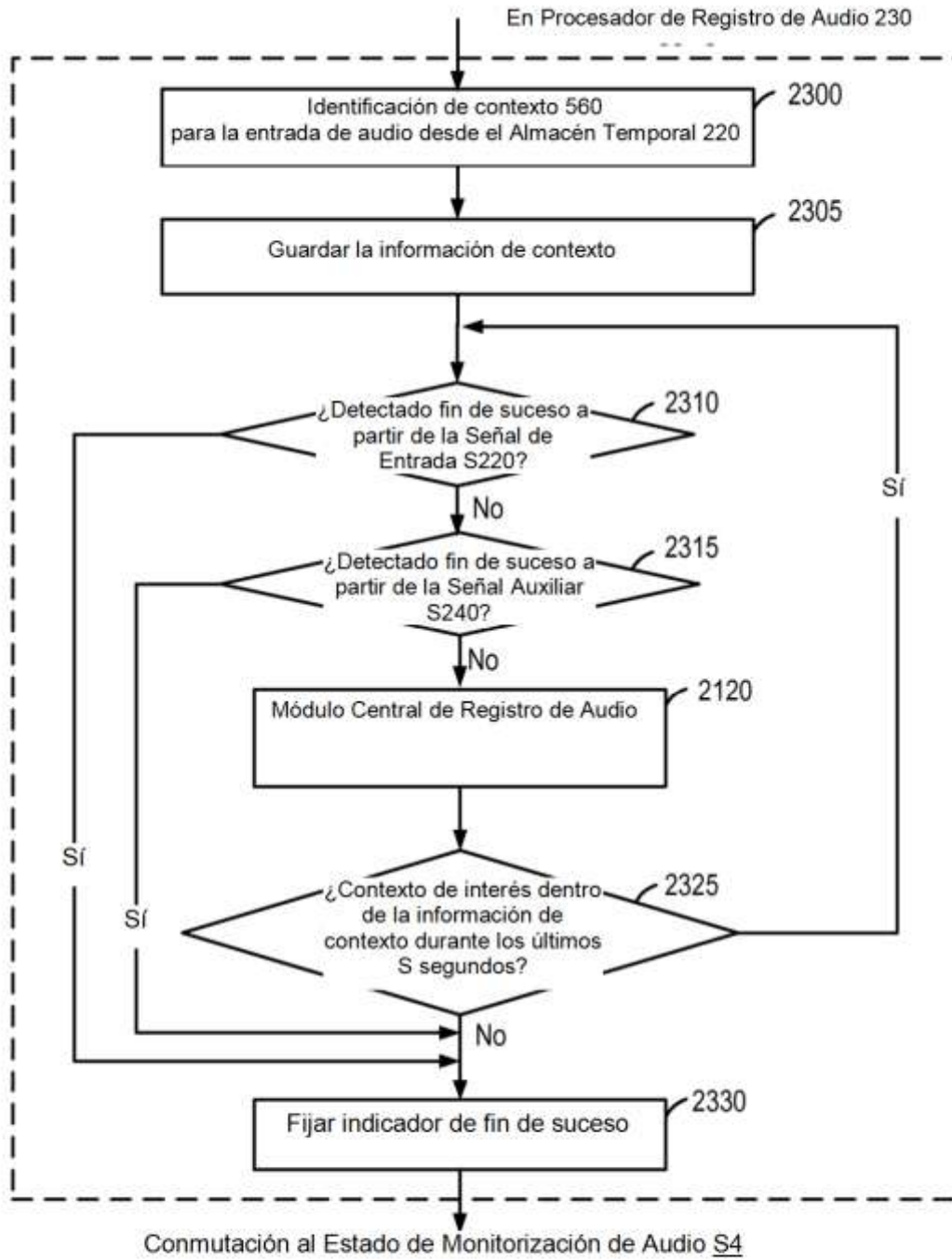


FIG. 23

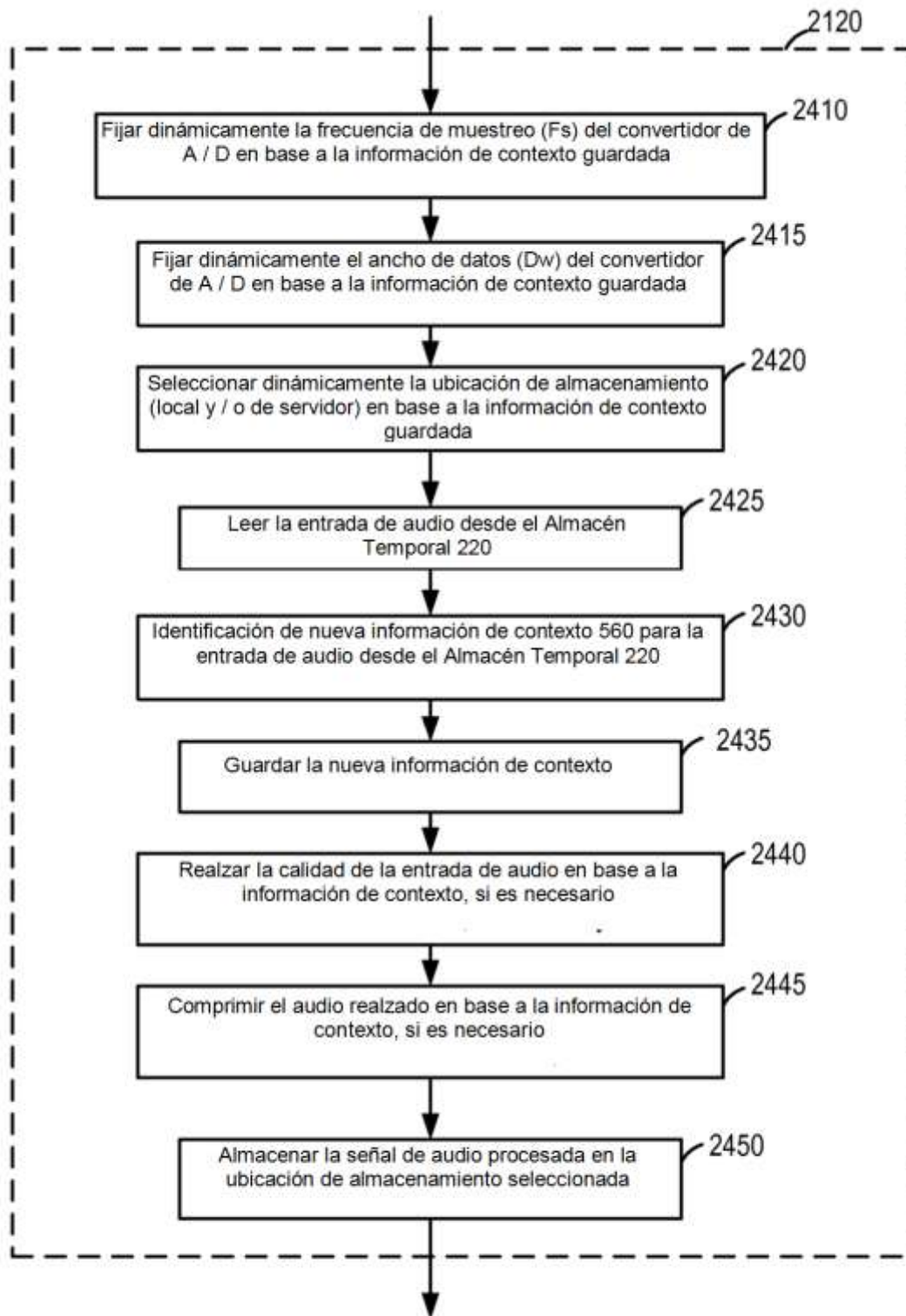


FIG. 24

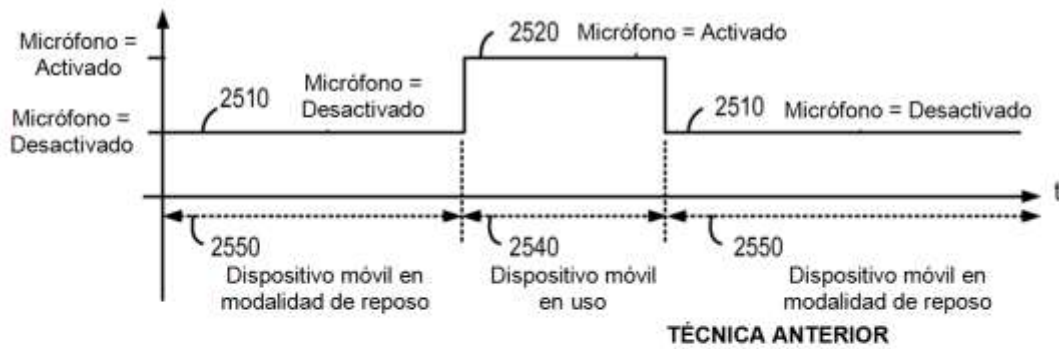


FIG. 25

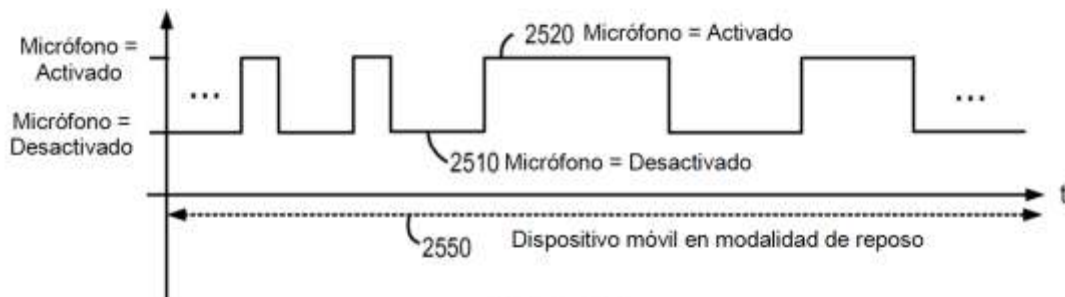


FIG. 26

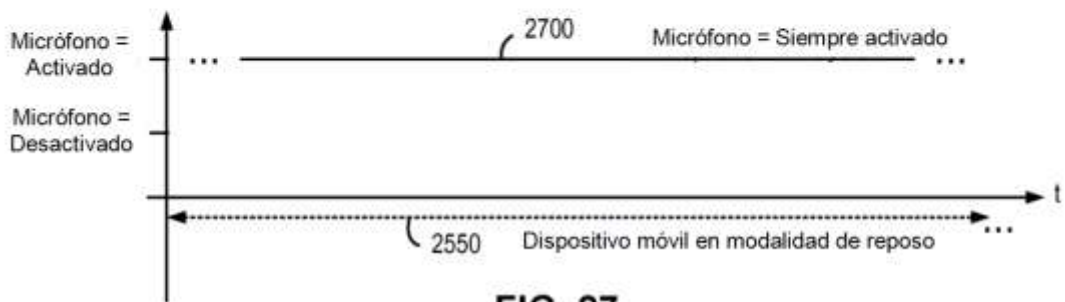


FIG. 27

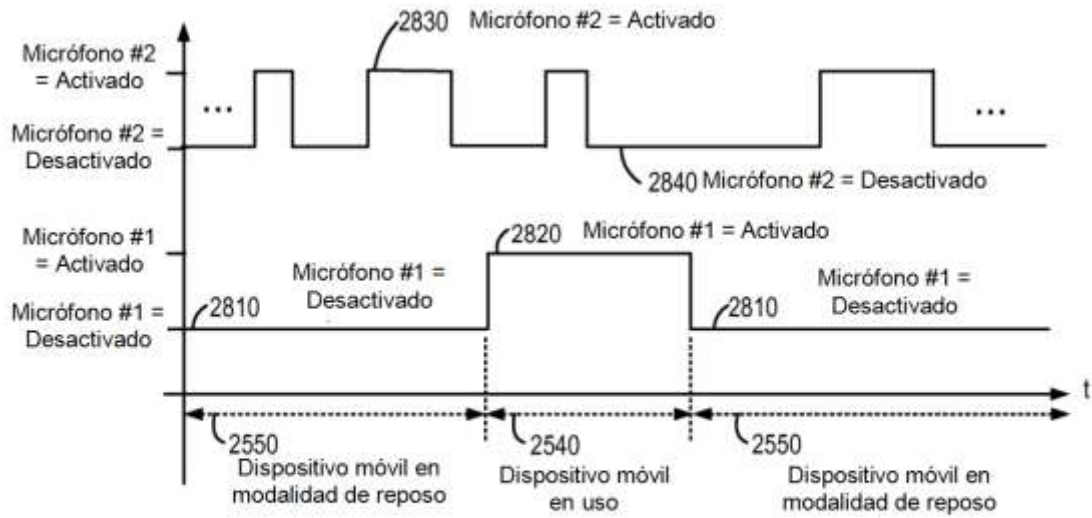


FIG. 28

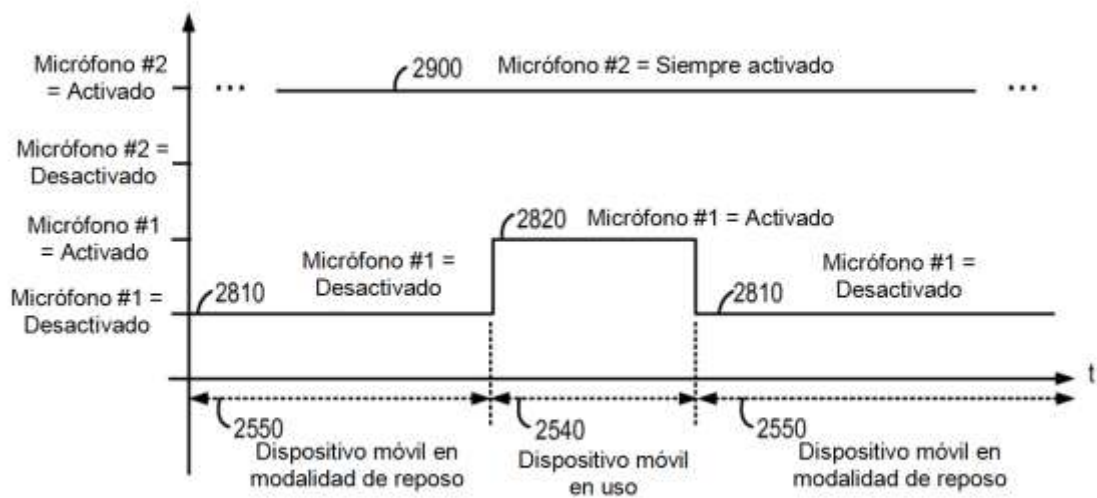


FIG. 29

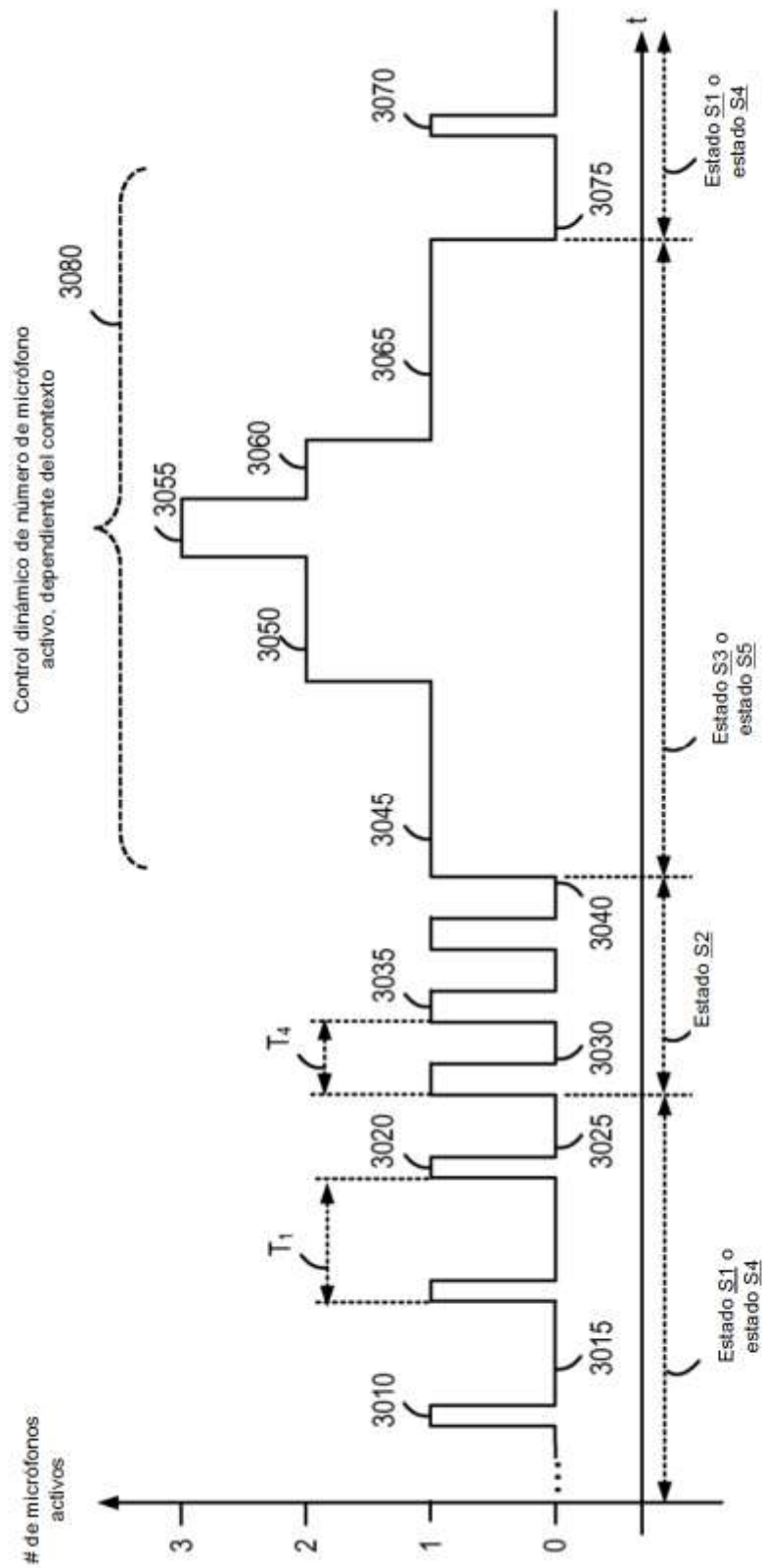


FIG. 30

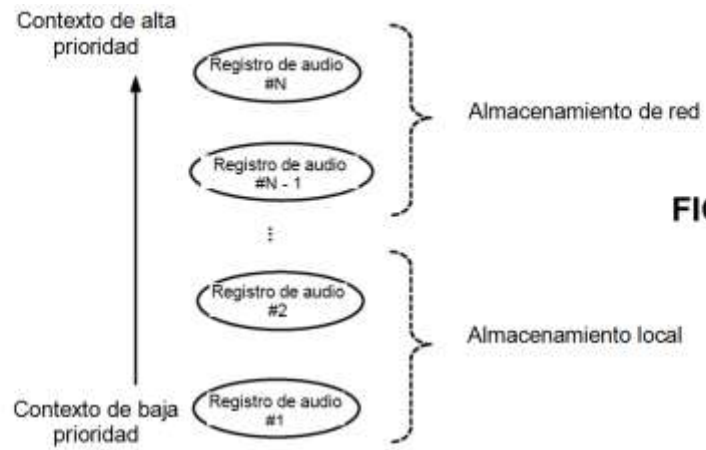


FIG. 31

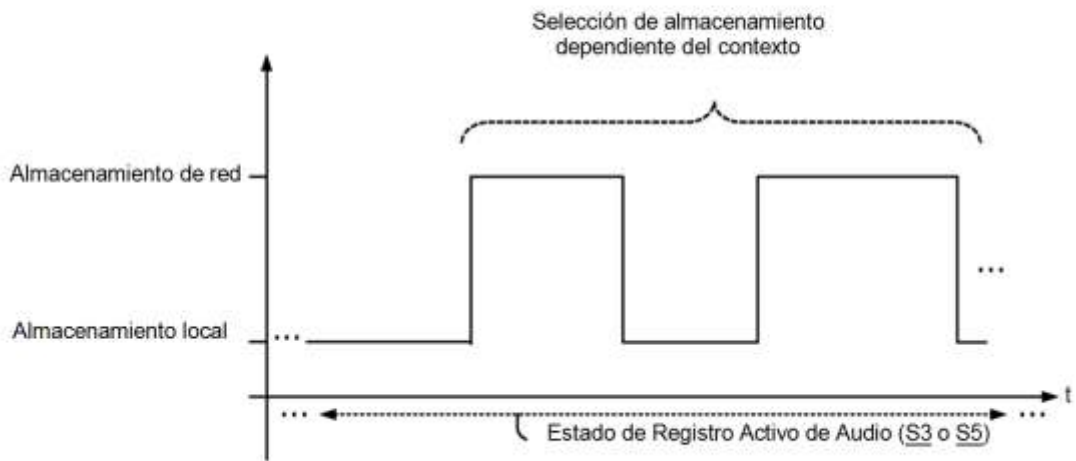


FIG. 32

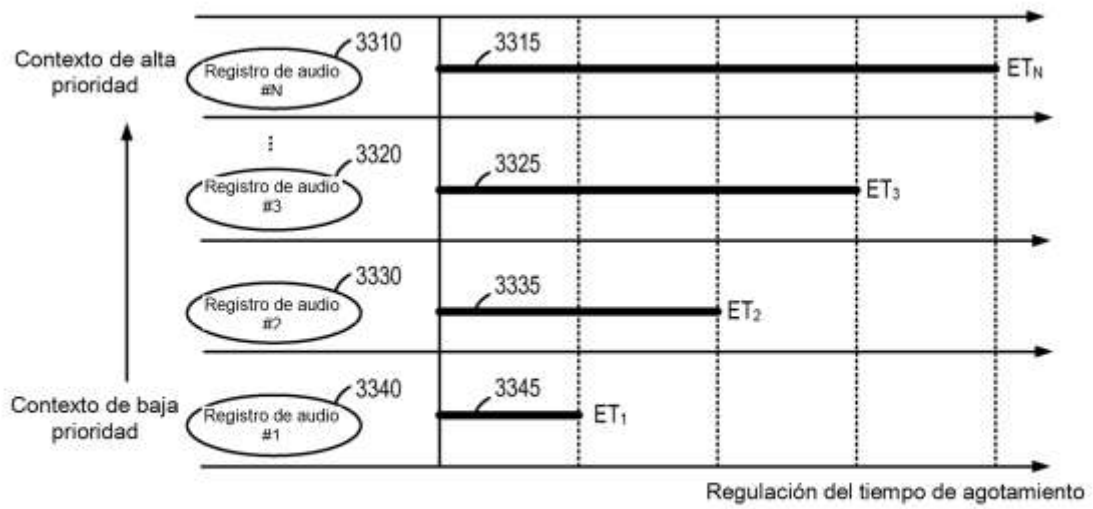


FIG. 33

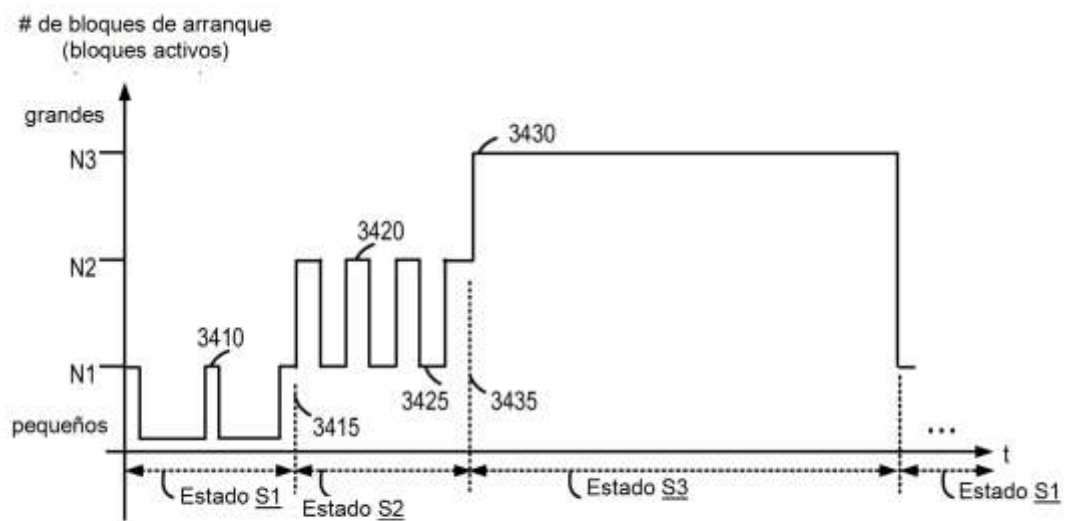


FIG. 34

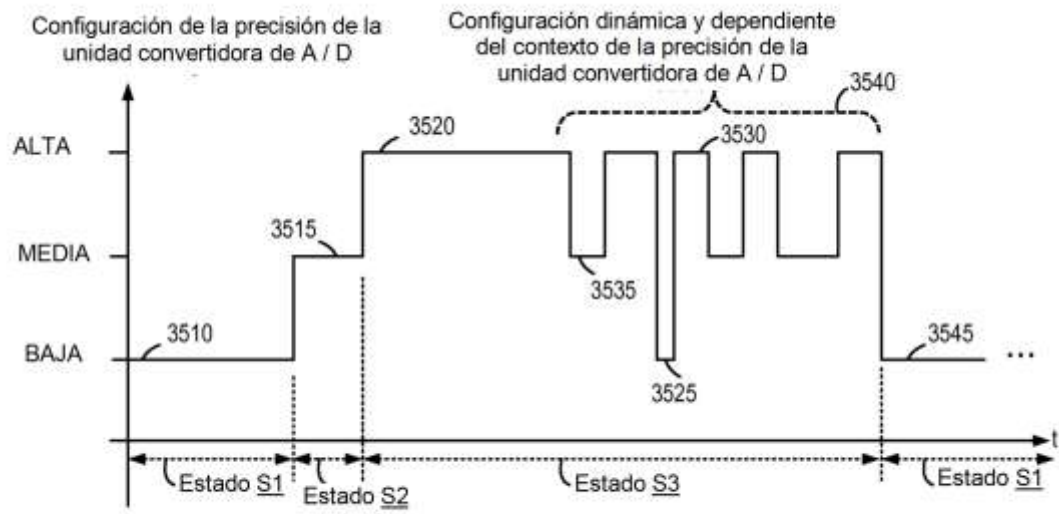


FIG. 35

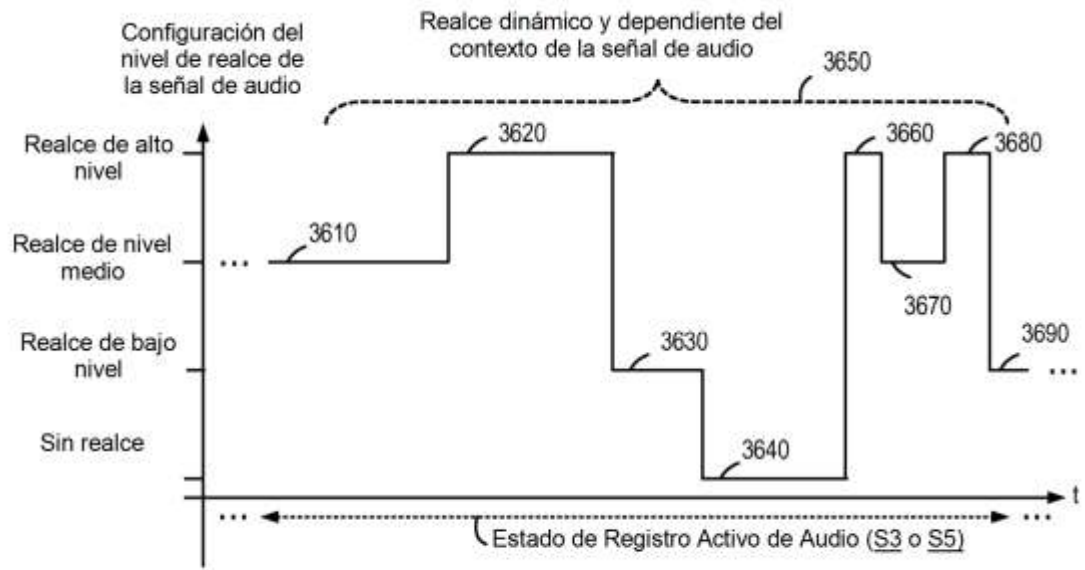


FIG. 36

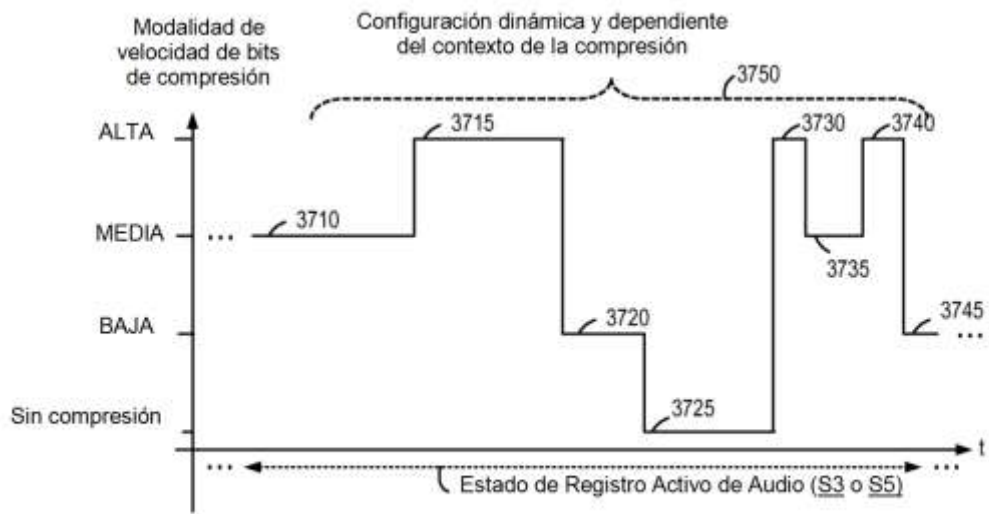


FIG. 37

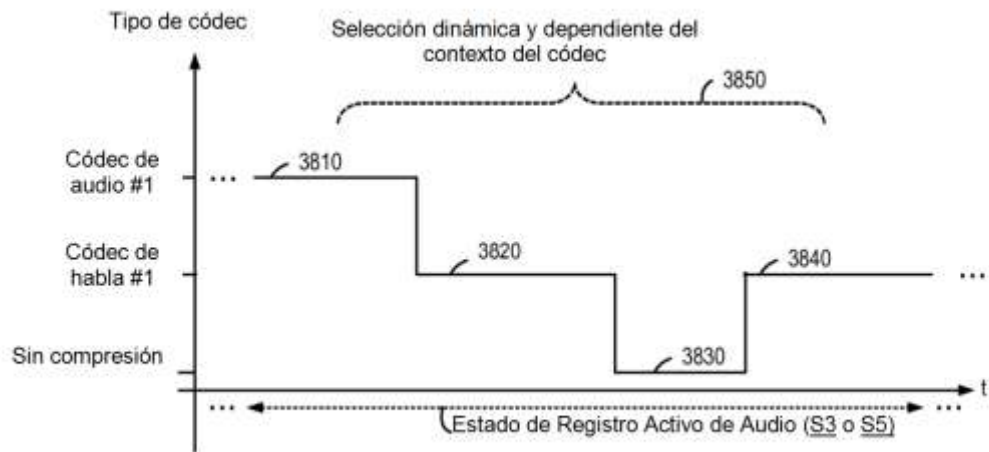


FIG. 38