

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **1 175 583**

21 Número de solicitud: 201631302

51 Int. Cl.:

H04N 7/18 (2006.01)

12

SOLICITUD DE MODELO DE UTILIDAD

U

22 Fecha de presentación:

31.10.2016

43 Fecha de publicación de la solicitud:

02.02.2017

71 Solicitantes:

FERMAX DESIGN & DEVELOPMENT, S.L.U.
(100.0%)
AVDA TRES CRUCES, 133
46017 VALENCIA ES

72 Inventor/es:

FERRER ZAERA, Carlos;
GARCIA BORT, Jose Ignacio y
ALBERT PEREZ, Vicente

74 Agente/Representante:

ISERN JARA, Nuria

54 Título: **PORTERO ELECTRONICO ACCESIBLE.**

ES 1 175 583 U

DESCRIPCIÓN

Portero electrónico accesible

Campo de la invención

5 La presente invención se refiere en general a sistemas y equipos de comunicación. En particular, la invención se refiere a placas de calle de portero y vídeo portero electrónicos que facilitan la accesibilidad en general para cualquier persona pero de forma particular para personas discapacitadas.

Antecedentes de la invención

10 Las placas de calle de porteros electrónicos y video porteros se emplean para comunicar a las personas que están en la calle con los residentes de una vivienda y para ello se requiere previamente que se realice una llamada a la vivienda para lo que se suele emplear pulsadores individuales que tienen asociado un rótulo con el número de la vivienda o el nombre del residente para poder localizarlo, o bien teclados tipo telefónico, que pueden ser mecánicos y/ o en una pantalla táctil que permiten componer el número de la vivienda, o
15 bien una agenda de nombres que se puede desplazar hasta localizar al residente y seleccionarlo.

Aunque se trata de sistemas que tienen un largo recorrido y que han ido perfeccionándose con el tiempo, todavía presentan una serie de problemas para ciertos colectivos, tal y como se indica a continuación.

20 Las personas con discapacidad visual tienen problemas para realizar la llamada a una vivienda ya que no pueden localizar el pulsador de llamada a la misma al no poder leer el rótulo asociado al mismo o, en el caso de incorporar la placa de calle una agenda ya sea impresa o en una pantalla electrónica, no pueden igualmente visualizar el nombre de la persona con la que quieren contactar.

25 Las personas con discapacidad física motora, que por ejemplo se desplazan en sillas de ruedas, no llegan al pulsador o teclado correspondiente para realizar la llamada ya que la placa de calle se suele instalar a una altura cómoda para las personas que están de pie.

Personas con dificultad para manejar las extremidades superiores o los dedos no pueden pulsar los botones o introducir códigos de teclado.

En este mismo contexto, los repartidores de mercancías a domicilio suelen tener las manos ocupadas y les dificulta la llamada a la vivienda a la que tienen que realizar el reparto.

Por seguridad o privacidad, en ocasiones no se muestran los nombres en los tarjeteros asociados a los pulsadores de llamada y solo se muestra el número de la vivienda, de manera que si no se conoce el número de la vivienda donde vive no se le puede localizar y llamar.

Se han llevado a cabo intentos para solventar los problemas descritos anteriormente pero no se ha encontrado una solución que solvete todos a la vez.

Por ejemplo, se pueden emplear etiquetas adhesivas transparentes con el número de vivienda grabado en relieve con el código Braille y colocarlas sobre el tarjetero asociado a cada pulsador. Esta solución es compleja cuando hay que rotular nombres y requiere un proceso personalizado que encarece y ralentiza la instalación del equipo. Además se pueden retirar con facilidad o se deterioran con el paso del tiempo al no estar protegidos por el propio porta tarjetas. Otra limitación reside en la poca proporción de invidentes que son capaces de leer el Braille, se estima en un 1%, y más aún ahora que hay disponibles nuevas tecnologías de lectura de pantallas mediante síntesis de voz.

Hay países donde la normativa local de accesibilidad obliga a situar las placas de calle a una altura accesible a las personas en sillas de ruedas, lo que perjudica a las personas que están de pie al tener que agacharse para poder utilizarla o bien requiere la instalación duplicada de la placa de calle, lo que repercute en una incomodidad de uso o un sobre coste respectivamente.

Una solución interesante y que ya se utiliza en dispositivos personales como smartphones o tabletas y también en ordenadores son los sistemas de reconocimiento de voz y asistentes tales como Siri (Apple), Google Now (Google), Alexa (Amazon) o Cortana (Microsoft). Estos sistemas no pueden implantarse en una placa de calle para resolver los problemas anteriores ya que presentan las siguientes limitaciones:

Requieren apretar un botón o pronunciar una palabra clave que ponga en marcha el motor de reconocimiento y así evitar que cualquier palabra que se pronuncie en su proximidad se interprete como una orden o solicitud de información. Por ejemplo, Siri requiere pulsar el botón del Iphone durante un tiempo, al igual que Cortana requiere pulsar un botón específico, Google Now requiere decir la palabra 'Hello Google', Alexa requiere decir la palabra 'Alexa'. Un usuario de una placa de calle no tiene por qué conocer esos métodos de

llamada ya que la va a utilizar esporádicamente. En el caso de los invidentes volvemos al problema de detectar donde se encuentra el pulsador.

5 Estos sistemas de asistencia por reconocimiento de voz se basan en un reconocimiento en la nube mediante servidores externos que procesan la información. Una placa de portero electrónico normalmente no dispone de esa conectividad.

10 Estos sistemas utilizan modelos acústicos de campo próximo, normalmente de no más de 20 cm, ya que están pensados para dispositivos personales que en el caso de una placa de calle no servirían ya que el usuario se suele poner a la distancia que alcanza el brazo, en decir, en torno a 50 cm, y además no son tolerantes al ruido ambiente y por tanto no funcionan bien en la calle.

15 En el campo del portero electrónico se ha hecho alguna tentativa empleando esta tecnología de reconocimiento de voz, pero todos recurren igualmente a la necesidad de pulsar algún botón, que para los colectivos anteriores, discapacitados físicos e invidentes principalmente, se ha visto un inconveniente, o la detección de la persona por medio de un detector de proximidad, donde puede ser fallido en caso de personas en sillas de ruedas.

20 En el estado del arte existen solicitudes de patente como por ejemplo GR20140100122A, EP2448233A1, DE19954844A1 o el modelo de utilidad CN204496627U que requieren de la pulsación del botón de activación del reconocedor de voz. Además, se limitan a reconocer los nombres de los residentes y no permiten realizar la llamada por medio del número de la vivienda, ya que, en ocasiones, no se conoce el nombre del mismo o no se proporciona por privacidad. Esto presenta además un inconveniente en la instalación ya que cuando se completa dicha instalación todavía no hay personas viviendo en el edificio y por tanto no dispone de esa información, por lo que no se podrá llamar hasta que se configure el nombre correspondiente. Otra característica de estas aplicaciones es que no permiten cambiar de manera dinámica el idioma para permitir que personas foráneas puedan utilizarla sin problemas.

A parte de estos problemas genéricos, cada uno de estos documentos presenta los siguientes problemas específicos:

30 En la solicitud GR20140100122A el equipo de reconocimiento de voz es un complemento a la placa de portero electrónico que incorpora su propio micrófono y altavoz por lo que resultará en un coste añadido, más complejo de instalar y no discierne la conversación del

usuario con la vivienda de la entrada del nombre de llamada, dando lugar a malinterpretaciones y además requiere de conexión de banda ancha a internet para ir perfeccionando el modelo acústico para mejorar el reconocimiento.

5 El documento EP2448233A1 no tiene en cuenta la privacidad de las personas ya que cuando se presenta una duda de entre dos nombres los reproduce para que el usuario seleccione uno.

En DE19954844A1 el objetivo es eliminar los pulsadores de llamada por lo que perjudica a las personas con trastorno del habla o que desconocen el idioma local.

10 El modelo de utilidad CN204496627U permite realizar la llamada exclusivamente por el número de vivienda pero no por el nombre y no se distingue si el número vocalizado es para llamar a la vivienda o para abrir la puerta.

15 Por lo tanto, el estado de la técnica actual no antecede un sistema capaz de resolver los problemas descritos con anterioridad en su conjunto y de una manera satisfactoria que permita realizar la llamada a la vivienda empleando comandos de voz sin requerir previamente activar el motor de reconocimiento, cuando se encuentra el usuario enfrente de la placa de calle, mediante un pulsador al efecto y/o un detector de presencia, y empleando indistintamente en la orden de llamada por voz el número de la vivienda o el nombre del residente. Ninguna posee ambas modalidades.

20 Todos estos antecedentes se limitan a realizar la llamada a la vivienda indicando únicamente el nombre del residente sin identificar previamente el comando por lo que no permiten otras opciones de comandos de voz tales como solicitar ayuda en el manejo de la placa de calle o llamar al conserje o incluso configurar el funcionamiento de la misma por parte del instalador o personal de mantenimiento.

Descripción de la invención

25 Es necesario ofrecer una alternativa al estado de la técnica que cubra las lagunas encontradas en la misma, particularmente mejorar la usabilidad actual de las placas de calle de los porteros electrónicos y video porteros como se va a describir a continuación.

30 Con este fin, la presente invención proporciona un portero electrónico accesible que incluye una placa de calle (100) donde dicha placa de calle (100) comprende un micrófono (1) dispuesto para capturar el sonido próximo a dicha placa (100), un interfaz analógico de audio de entrada (2) dispuesto para digitalizar el sonido capturado por dicho micrófono (1),

detector de umbral (3) dispuesto para discriminar si el sonido capturado tiene suficiente calidad para ser tratado, uno o varios modelos acústicos (15) que representan, cada uno de ellos, la pronunciación de fonemas de un idioma, un generador de fonemas (6) dispuesto para detectar fonemas a partir de dicho sonido capturado y discriminado, uno o varios contextos (16) que representan el conjunto de palabras y/o frases y reglas gramaticales que pueden ser reconocidas a partir de dichos fonemas detectados, un reconocedor (7) dispuesto para comparar dichos fonemas detectados con al menos uno de dichos contextos (16) y reconocer una o varias palabras y/o frases, un analizador (8) dispuesto para analizar dichas palabras y/o frases reconocidas y determinar si son aceptadas como comandos y/o identificadores de usuarios o cifras. Un conversor de texto a voz o TTS (*text-to-speech*) (9) dispuesto para reproducir mensajes de voz al usuario de la placa (100), un Interface analógico de audio de salida (12) dispuesto para convertir las señales digitales recibidas del conversor texto a voz o TTS (9) correspondientes al mensaje de voz a reproducir en señales analógicas, un bus de comunicaciones de portero electrónico del edificio (13) dispuesto para establecer las comunicaciones o llamadas entre la placa de calle (100) y lo terminales de las viviendas y/o conserje, un interface de portero electrónico (11) dispuesto para comunicar con dicho bus de comunicaciones (13) y transmitir lo comandados detectados a los terminales de las viviendas y establecer la comunicación audio y/o video con los mismos, un altavoz (14) dispuesto para reproducir las señales de audio en la placa de calle (100), una agenda o lista (18) dispuesta para contener nombres o actividad profesional de los residentes de la vivienda, una memoria RAM (4) dispuesta para almacenar de forma volátil los la información y datos necesarios para el reconocimiento y el software de los distintos elementos que hay en ejecución, una memoria Flash (5) dispuesta para almacenar de forma no volátil los la información y datos necesarios para el reconocimiento y el software de los distintos elementos y una CPU o procesador (10) dispuesto para controlar y gestionar el resto de elementos del dicha placa de calle (100).

Por tanto, al contrario que las soluciones existentes, la placa de calle (100) objeto de la presente invención incorpora un sistema de reconocimiento de voz independiente del usuario para poder reconocer a cualquier persona que lo utilice, además de contar con los medios clásicos de llamada como son los pulsadores, teclado o agenda electrónica. El citado reconocedor está funcionando continuamente cuando la placa de calle (100) está en reposo para atender cualquier solicitud y no requiere ninguna acción por parte del usuario tal como la pulsación de un botón o la detección de presencia mediante un sensor de proximidad.

Dicha placa de portero electrónico (100) acepta comandos de llamada a viviendas mediante

5 cualquier expresión vocal natural que suela utilizar el usuario tales como 'llamar', 'marcar', 'contactar', 'hablar' designando a continuación el nombre del residente, actividad profesional ('dentista', 'abogado') o el número de la vivienda, en diferentes modalidades tales como números naturales ('ciento veintitrés'), dígito a dígito ('uno dos tres'), pares de dígitos ('uno veintitrés' o 'doce tres'), números ordinales ('segundo') y en combinación con letras ('segundo B'). De esta manera el usuario no tiene por qué seguir una regla estricta de llamada y se adapta a las costumbres de cada región a país.

10 Además, dicha placa acepta y discrimina comandos de ayuda a la utilización ('ayuda', '¿Cómo se llama?',...) o de comunicación con un conserje o guardia de seguridad ('conserje', 'guardia',...).

Permite además configurar el equipo mediante comandos de voz que emplea el instalador o personal de mantenimiento ('volumen de audio 4', 'tiempo de apertura de puerta 4 segundos',...) previa identificación como instalador.

15 Dicha placa de calle genera una realimentación al usuario por medio de una voz sintetizada que permite confirmar si la acción ha sido correctamente entendida en caso de duda.

20 La placa de calle puede, opcionalmente, invitar al usuario a utilizar la llamada por voz cuando detecta la presencia de una persona enfrente de la placa de calle, por medio de un detector de presencia, al reproducir un mensaje de voz sintetizada que indica cómo utilizarla por medio de comandos de voz. Dicho detector de presencia no pone en marcha el motor de reconocimiento porque está continuamente activo, simplemente hace reproducir el mensaje de voz. Esta opción es especialmente útil para los invidentes que no pueden leer un mensaje en un display o un cartel que advierta del uso mediante la voz.

25 Por el hecho de utilizar la placa de calle la síntesis de voz para dar feedback de las acciones se adapta a los diferentes niveles de ruido ambiente del entorno para que los mensajes vocales que reproduce se hagan a un volumen adecuado para que se entiendan sin resultar excesivamente altos cuando hay poco ruido ambiente (noche por ejemplo) o excesivamente bajos cuando hay un exceso de ruido (tráfico en una calle congestionada). Emplea un control automático de volumen en función del ruido ambiente.

30 La placa de calle puede ser configurada por el instalador para funcionar con el idioma del país o dialecto para el reconocimiento y la síntesis de voz por defecto, pero también se adapta al idioma del interlocutor al permitir cambiar el idioma de manera dinámica cuando el usuario pronuncia el nombre del idioma que desea en su propia lengua ('Español', 'Catalá',

'Galego', 'English', 'Française', 'Italiano', '中国'(chino),...).

Como se puede deducir, con esta invención la privacidad de los residentes queda asegurada si no se muestran los nombres de los mismos en los rótulos de los pulsadores (solo el nº de vivienda) o en la agenda electrónica. La placa de calle permite la llamada por comando de voz empleando su nombre, si el usuario conoce que vive en ese edificio, no quedando expuesto públicamente. Se puede configurar la placa de calle para que en caso de coincidir con varios nombres el apellido indicado, no pregunte a cual se refiere desvelando los nombres de todos los coincidentes.

Estas y otras ventajas se ven evidentes a la luz de la descripción detallada de la invención.

Breve descripción de los dibujos

Las anteriores y otras ventajas y características se entenderán más completamente a partir de la siguiente descripción detallada de realizaciones, con referencia a la siguiente figura, que deben considerarse de una manera ilustrativa y no limitativa.

La figura 1 muestra, esquemáticamente, los distintos elementos que forma parte del portero electrónico, generalmente implementado como una placa de calle.

Descripción detallada de la invención

Los elementos definidos en esta descripción detallada se proporcionan para ayudar a una comprensión global de la invención. En consecuencia, los expertos en la técnica reconocerán que variaciones y modificaciones de las realizaciones descritas en este documento pueden realizarse sin apartarse del alcance y espíritu de la invención. Además, la descripción detallada de las funciones y elementos suficientemente conocidos se omiten por razones de claridad y concisión.

Por supuesto, las distintas funcionalidades de la invención pueden ser implementadas con diferentes variaciones de arquitectura, protocolos, dispositivos o tipos de servicios y aplicaciones. Cualquier implementación presentada a continuación es incluida con el propósito de ilustrar y hacer comprensiva la invención y no con la intención de limitar aspectos de la misma.

Como se puede ver en la figura 1 la placa de calle (100) incluye diferentes elementos que permite implementar la funcionalidad descrita en la descripción anterior.

Concretamente se distinguen:

- Micrófono (1). Este elemento captura el sonido próximo a la placa de calle (100). Se emplea tanto en el reconocimiento de voz como en la conversación con la vivienda llamada.
- Interface analógico de audio de entrada (2). Realiza la digitalización del sonido que captura el micrófono (1) para poder llevar a cabo el muestro y éste ser tratado de manera numérica. Se trata de un conversor ADC (*analog-to-digital converter*). El sonido capturado se almacena en una memoria RAM (*random access memory*) (4). También se emplea en la conversación con las viviendas cuando hay una comunicación establecida entre la placa y una de esas viviendas.
- Detector de umbral (3). Es un elemento que discrimina el nivel del sonido de entrada para indicar cuando se tiene la calidad suficiente para ser analizado por el reconocedor de voz. Puede ser un módulo software o hardware.
- RAM (4). Memoria de acceso aleatorio volátil que almacena las muestras de voz digitalizada, modelos acústicos activos, aplicaciones de reconocimiento de voz y síntesis de voz y sus datos en tiempo de ejecución, la aplicación del propio portero electrónico y sus datos en tiempo de ejecución y los contextos activos a reconocer.
- Memoria FLASH (5). Memoria no volátil que almacena, entre otras cosas, las aplicaciones de portero electrónico, aplicaciones de reconocimiento y de síntesis de voz, la agenda de nombres (18) de la placa de calle (100), los modelos acústicos (15), contextos (16) y voces de los diferentes idiomas que reconoce y sintetiza así como toda la información que necesita para poder ponerse en marcha.
- Generador de fonemas (6). Módulo, que se puede implementar por software, que detecta patrones de fonemas en las muestras de audio capturadas para poder ser analizadas. Dicha detección se basa en la comparación del sonido detectado y filtrado con al menos uno de los modelos acústicos (15) con los que cuenta la placa (100) y que está cargado en memoria RAM (4).
- Reconocedor (7). Módulo, que se puede implementar por software, que analiza los fonemas detectados y los compara con las palabras y reglas definidas en el contexto de la gramática que debe reconocer. La salida del analizador (8) son las palabras y frases de la gramática que se han detectado junto a una puntuación de semejanza.
- Analizador (8). Módulo software que valora las palabras y reglas detectadas y determina

si son aceptadas dependiendo de la puntuación de semejanza y del contexto de la situación.

- 5
 - TTS (*text-to-speech*) (9). Conversor de texto a voz para realizar realimentación o respuesta al usuario mediante voz; es decir, para reproducir mensajes de voz al usuario de la placa en caso de que se requiera, en el idioma seleccionado.
- CPU (10). Procesador de la placa de calle (100) que controla y gestiona los diferentes elementos hardware y software.
- 10
 - Interface de portero electrónico (11). Comunica con el bus de instalación del edificio para poder transmitir los comandados a los terminales de las viviendas y establecer la comunicación audio y video con los mismos.
- 15
 - Interface analógico audio de salida (12). Convierte las señales de síntesis de voz digitales en señales reproducibles en el altavoz. Típicamente se trata de un conversor DAC (*digital-to-analog converter*) o PWM (*pulse width modulator*). También se emplea en la conversación con las viviendas para adaptar el audio que proviene de los terminales de las viviendas.
- Bus de comunicaciones de portero electrónico del edificio (13). Establece las comunicaciones entre la placa de calle (100) y los terminales de las viviendas.
- Altavoz (14). Reproduce las señales de audio en la placa de calle (100).
- 20
 - Modelo acústico (15). Es una representación de la pronunciación de los fonemas. Depende de cada idioma. El modelo acústico (15) está entrenado para trabajar con ruido ambiente y detección alejada del locutor. Pueden existir varios modelos acústicos (15), uno por idioma, almacenados en la memoria Flash (5), pero al menos uno ellos cargados en memoria RAM (4) para ser usado en un momento dado.
- 25
 - Contexto (16). Representa el conjunto de palabras y frases, con sus reglas, que deben ser reconocidas expresadas en fonemas. Depende de cada idioma. Al igual que los modelos acústicos (15), pueden haber uno o varios almacenados en memoria Flash (5) y al menos uno de ellos se carga en memoria RAM (4) para ser usado en un momento dado en función de la configuración realizada en la placa de calle (100).
- 30
 - Detector de presencia (17). Dispositivo que detecta la presencia de una persona en la proximidad de la placa de calle (100). Cuando se detecta dicha persona próxima se

emite un mensaje de ayuda.

- Agenda o lista (18). Contiene la lista de nombres o actividad profesional de los residentes con información de la vivienda asociada.

5 Con estos elementos, la placa de calle (100) objeto de la invención, cuando recibe alimentación carga los módulos software en la memoria RAM (4) desde la memoria flash (5). Dentro de los módulos que se cargan están todos los relacionados con el reconocimiento de voz, cuya funcionalidad queda preparada para empezar a trabajar. Entre ellos se encuentra el detector de umbral (3), el generador de fonemas (6), el reconocedor (7) y el analizador (8), junto al modelo acústico (15) que debe reconocer el generador de fonemas (6) y 10 contexto (16) que debe reconocer el analizador (8).

El modelo acústico (15) que se aplica en cada momento depende del idioma por defecto que se haya configurado por el instalador para trabajar y se puede cambiar dinámicamente. Es decir, puede haber uno o varios modelos acústicos cargados en memoria RAM (4) de los almacenados en memoria Flash (5) para ser usado en cada momento en función de dicha 15 configuración.

El contexto (16) contiene información de la gramática que debe reconocer, es decir, del conjunto de palabras y reglas para la composición de las frases, y depende también del idioma seleccionado y de la situación del momento. Por tanto hay uno o varios contextos (16) en memoria RAM (4) en cada momento de los almacenados en memoria Flash (5). En 20 el contexto (16) se incluye la agenda de nombres o profesiones que debe reconocer (18). Un cambio en la agenda (18) provoca una actualización del contexto (16) de manera dinámica.

La placa de calle (100) en reposo está a la espera de detectar sonido en su proximidad y analizarlo para determinar si se trata de un comando de voz del repertorio que debe reconocer.

25 Los procesos o modos de funcionamiento del portero electrónico o placa de calle (100) que se suceden son los siguientes:

- El micrófono (1) recoge el sonido ambiente y por tanto del posible locutor y se digitaliza por medio del interface analógico de audio de entrada (2). Las muestras de audio se tratan de forma numérica, es decir, se realiza una cuantización y se almacenan en la 30 memoria RAM (4) para ser analizadas.
- El detector de umbral (3) mide el nivel de sonido y determina si tiene el suficiente nivel

para asegurar una calidad en el reconocimiento. De esta manera se rechaza el ruido ambiente propio del entorno o conversaciones alejadas de gente que circula por la calle y que no contiene información del locutor que quiere utilizar la placa de calle (100). Los sonidos fuera del alcance usual, por ejemplo que provienen de una distancia superior a un metro de la placa (100), se descartan por poseer poca energía. Esta medida del nivel de sonido o ruido se utiliza también para actualizar el valor del ruido ambiente que se tendrá en cuenta a la hora de ajustar el volumen de reproducción de la síntesis de mensajes de voz.

- El generador de fonemas (6) se pone en funcionamiento cuando el detector de umbral (3) notifica que hay sonido con nivel suficiente para considerarse adecuado. Detecta fonemas en las señales de audio adquiridas y los almacena de nuevo en la RAM (4) para ser tratadas por el reconocedor. A la vez rechaza, filtrándolo, ruido ambiente que no corresponde con voz.
- El reconocedor (7) compara las secuencias de fonemas adquiridas con los elementos del contexto (16) de trabajo que debe reconocer. Cada idioma utiliza un contexto diferente con las palabras propias de cada idioma que representan los diferentes comandos ('llamar', 'call', 'appel', 'chiamata',...). Cada cambio de la agenda (18) requiere una modificación del contexto (16) que se realiza de manera dinámica. Si se cambia de estado: reposo, configuración del equipo, etc., también cambia la gramática a emplear y por tanto el contexto (16) para que de esta manera se evite reconocer acciones que no corresponden a la situación actual en la que se encuentra la placa de calle (100). Se pueden tener uno o varios contextos (16) almacenados en memoria Flash (5) pero solo algunos activos en cada momento en RAM (4), en función de la configuración del equipo.
- El reconocedor (7) asigna puntuaciones a las distintas palabras y frases que reconoce en función de la similitud a la frase recogida en el contexto. Esta puntuación determinará si se debe aceptar el comando reconocido o no.
- El analizador (8) decide si la frase reconocida se acepta en función de dicha puntuación y de la situación en la que se encuentra el sistema (reposo, necesidad de confirmar, configuración,...). Para ello se emplean dos valores umbrales: 'confiable' y 'duda', que se determinan de manera pragmática con una serie de experiencias con diferentes usuarios.
- Si la puntuación es superior al umbral 'confiable' se ejecuta el comando reconocido: 'llamada a vivienda 23A', 'hablar con Jesús García', 'contactar con abogado',... y se

vocaliza la acción que se ha entendido, junto a información adicional: nº de vivienda, nombre de la persona que se desea contactar, etc. Para esto el analizador (8) genera un texto con el mensaje a reproducir e invoca al TTS (9).

- 5 • En caso de detectarse varios nombres que contienen el nombre vocalizado, por ejemplo que compartan el apellido, dependiendo del modo de privacidad seleccionado, se solicitará que se indique el nombre correcto (modo privacidad) o se reproducirán los nombres coincidentes y se solicitará que elija el que desea (modo sin privacidad), dependiendo de la configuración del sistema en ese momento.
- 10 • Si la puntuación está entre el umbral 'confiable' y el umbral 'duda' se solicita confirmación mediante un mensaje reproducido en el TTS (9) (como por ejemplo: '¿quiere llamar a Jesús García?'), y si se confirma mediante alguna de las palabras que reconoce el contexto (como por ejemplo: 'sí', 'correcto', 'en efecto', 'vale',...), se acepta el comando entendido.
- 15 • Si la puntuación es inferior al umbral 'duda' se descarta, no generando feedback. De esta manera no es requisito un botón de activación del reconocimiento o detector de presencia ya que se rechaza cualquier sonido que no tenga una puntuación mínima, es decir, cualquier sonido que no tenga un mínimo de calidad en cuanto a ruido y que se corresponda con las palabras y/o frases del contexto activo (16).
- 20 • Cuando el analizador (8) detecta un comando válido se ejecuta la acción requerida:
 - 25 ○ Para realizar la llamada a una vivienda analiza si se emplea un código de llamada correcto o un nombre de residente o actividad profesional incluido en la agenda (18). En caso de una detección positiva, es decir, que se reconozca dicho nombre, código, etc. adecuadamente, se envía el comando de llamada al terminal de la vivienda por medio del bus de comunicaciones (13) y se establece la comunicación audio, y opcionalmente video si el sistema dispone de estas capacidades, con la misma. Durante una conversación con la vivienda el reconocimiento de voz permanece desactivado para impedir la detección de falsos comandos. Cuando finaliza la conversación, es decir, se cierra la comunicación con la vivienda, el reconocimiento de voz se pone en marcha de nuevo.
 - 30 ○ Si se trata de un comando de llamada a conserje se encamina la llamada al terminal del mismo.

- 5

○ Si se trata de un comando de cambio de idioma (por ejemplo: ‘Español’, ‘Catalá’, ‘English’,...) se actualiza el contexto (16) en el reconocedor (7), es decir, se carga desde la memoria Flash (5) donde está almacenado y el modelo acústico (15) en el generador de fonemas (6) de manera que se pueden aceptar comandos de voz en el nuevo idioma. Después del reconocimiento en el nuevo idioma, o transcurrido un tiempo pre-configurado de ausencia de actividad, se vuelve de nuevo al idioma por defecto configurado por el instalador, cargándose el modelo acústico (15) y contexto (16) correspondientes, el de por defecto según la configuración en ese momento.
- 10

○ Si es un comando de configuración se solicita mediante un mensaje reproducido en el TTS (9) un código de acceso y, si es correcto, se accede a un menú vocal de configuración, es decir, también reproducido por el TTS (9), en el que el instalador indica el parámetro a configurar y el valor deseado (como por ejemplo: ‘tiempo de apertura tres segundos’), realizándose el cambio en el parámetro si la puntuación del analizador (8) está por encima del umbral ‘confiable’ o pidiendo confirmación si está por debajo pero por encima del umbral ‘duda’. En cada caso se da feedback mediante la síntesis de voz. También se tiene la opción de consultar el valor actual si se indica el parámetro pero no un valor, dándose la respuesta mediante síntesis de voz (por ejemplo: ‘tiempo de apertura es 5 segundos’). El comando ‘salir’ o una palabra similar devuelve el estado de la placa de calle (100) al modo de funcionamiento normal, preparado para aceptar comandos de usuario. También se dispone de un comando de ayuda que indica los parámetros que son configurables mediante la síntesis de voz.
- 15

○ Un comando de ayuda provoca la reproducción de un mensaje sintetizado que explica como operar mediante comandos de voz.
- 20

○ El conversor de texto a voz TTS (9) cuando es invocado se carga en memoria RAM (4) para ser ejecutado en ese momento y analiza el texto a reproducir, lo transforma en un fichero de audio PCM y lo envía al interface analógico de audio de salida (12) para convertir las señales numéricas a un formato reproducible por el altavoz (14). El volumen de reproducción se ajusta en función del ruido ambiente detectado por el detector de umbral (3). El conversor de texto a voz, cuando se invoca, utiliza el idioma que esté activo en ese momento.
- 25

○ Una vez se ha terminado la reproducción del mensaje el módulo TTS (8) se
- 30

descarga de memoria RAM (4), liberando el espacio que ocupaba.

Opcionalmente se puede activar la detección de presencia de personas. En ese caso, cuando un usuario se acerca a la placa de calle (100) y lo detecta el detector de presencia (17) se genera el mensaje de bienvenida y ayuda mediante TTS (9), de manera que una persona pueda saber cómo realizar la llamada por medio de comandos de voz, de especial
5 utilidad para las personas invidentes. El reconocedor de voz está preparado para detectar cualquier comando que se pronuncie, independientemente de si se realiza la detección de presencia o no. En caso de detectar sonido arranca el proceso de reconocimiento descrito anteriormente.

De forma diferenciadora, en el reconocimiento de voz y la generación de feedback vocal se
10 utiliza el propio micrófono y altavoz de la placa de calle (100), sin requerir por tanto de hardware adicional al que ya disponen para estas funciones las placas de calle habituales.

El sistema de reconocimiento está embebido en la placa de calle (100) de audio o video portero, compartiendo la mayoría de los elementos que ésta ya utiliza para desempeñar su
15 función (micrófono, altavoz, interfaces analógicos, CPU, memoria) e incorpora las nuevas funcionalidades mediante software específico de reconocimiento de voz y de conversión de texto a voz. No se trata por tanto de un sistema añadido ni necesita conectividad a internet y el aspecto de la placa de calle es como la de una placa convencional de portero electrónico o video portero.

La gramática que maneja permite realizar diferentes acciones empleando comandos de voz
20 en lenguaje natural, sin necesidad de seguir una estructura rígida, es decir, permite llamar a una vivienda o al conserje, consultar si vive cierta persona en el edificio, solicitar ayuda para el manejo, configurar los parámetros de funcionamiento por parte del instalador, entre otros.

En el caso de una llamada a vivienda permite diferentes modalidades de llamada: mediante
25 el nombre de la persona a la que se quiere llamar, la actividad profesional o el número de vivienda. En este último caso maneja diferentes formas de indicar el número: un número natural ('ciento veintitrés',...), dígito a dígito ('uno dos tres'), mediante pares o combinación con dígitos ('doce tres' o 'uno veinte tres', mediante ordinales ('quinto') y siempre permite añadir una letra a continuación para poder designar una vivienda de una planta ('primero B'
30 o 'doce C').

REIVINDICACIONES

1. Portero electrónico accesible que incluye una placa de calle (100) **caracterizado** porque dicha placa de calle (100) comprende:

- un micrófono (1) dispuesto para capturar el sonido próximo a dicha placa,

5 - un interfaz analógico de audio de entrada (2) dispuesto para digitalizar el sonido capturado por dicho micrófono (2),

- un detector de umbral (3) dispuesto para discriminar si el sonido capturado tiene suficiente calidad para ser tratado,

10 - uno o varios modelos acústicos (15) que representan, cada uno de ellos, la pronunciación de fonemas de un idioma,

- un generador de fonemas (6) dispuesto para detectar fonemas a partir de dicho sonido capturado y discriminado,

15 - uno o varios contextos (16) que representan el conjunto de palabras y/o frases y reglas gramaticales que pueden ser reconocidas a partir de dichos fonemas detectados,

- un reconocedor (7) dispuesto para comparar dichos fonemas detectados con al menos uno de dichos contextos (16) y reconocer una o varias palabras y/o frases,

- un analizador (8) dispuesto para analizar dichas palabras y/o frases reconocidas y determinar si son aceptadas como comandos y/o identificadores de usuarios o cifras.

20 - un convertor de texto a voz o TTS (*text-to-speech*) (9) dispuesto para reproducir mensajes de voz al usuario de la placa (100),

- un Interface analógico de audio de salida (12) dispuesto para convertir las señales digitales recibidas del convertor texto a voz o TTS (9) correspondientes al mensaje de voz a reproducir en señales analógicas,

25 - un bus de comunicaciones de portero electrónico del edificio (13) dispuesto para establecer las comunicaciones o llamadas entre la placa de calle (100) y los terminales de las viviendas y/o conserje,

- un interface de portero electrónico (11) dispuesto para comunicar con dicho bus de

comunicaciones (13) y transmitir los comandados detectados a los terminales de las viviendas y establecer la comunicación audio y/o video con los mismos,

- un altavoz (14) dispuesto para reproducir las señales de audio en la placa de calle (100),

5 - una agenda o lista (18) dispuesta para contener nombres o actividad profesional de los residentes de la vivienda,

- una memoria RAM (4) dispuesta para almacenar de forma volátil la información y datos necesarios para el reconocimiento y el software de los distintos elementos que hay en ejecución.

10 - una memoria Flash (5) dispuesta para almacenar de forma no volátil la información y datos necesarios para el reconocimiento y el software de los distintos elementos y

- una CPU o procesador (10) dispuesto para controlar y gestionar los elementos de dicha placa de calle (100).

15 2. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho micrófono (1) se utiliza tanto para el reconocimiento de voz como para la conversación con la vivienda a la que se llama.

20 3. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho interfaz analógico de audio de entrada (2) es un conversor ADC (*analog-to-digital converter*) y donde el sonido digitalizado por dicho interfaz (2) se almacena en dicha memoria RAM (4).

4. Portero electrónico accesible según la reivindicación 1 y 3 **caracterizado** porque dicho interfaz analógico de audio de entrada (2) adicionalmente se utiliza en la conversión del sonido capturado cuando hay una comunicación establecida con una vivienda.

25 5. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho detector de umbral (3) discrimina el sonido capturado filtrando, dejando pasar el que tenga suficiente potencia y por tanto se considera que se ha generado a menos de una cierta distancia de dicha placa de calle (100).

30 6. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dichos modelos acústicos (15) están entrenados para trabajar con ruido ambiente y detección alejada del locutor y donde dichos modelos acústicos (15) están almacenados en

dicha memoria Flash (5) y donde al menos uno de ellos está cargado en memoria RAM (4) y es utilizado durante el reconocimiento de voz del sonido capturado y filtrado.

- 5 7. Portero electrónico accesible según la reivindicación 6 **caracterizado** porque dicho generador de fonemas (6) se pone en funcionamiento cuando dicho detector de umbral (3) detecta que hay sonido con suficiente potencia, filtra el ruido de dicho sonido que no corresponde con voz, detecta dichos fonemas utilizando el al menos un modelo acústico (15) cargado en memoria RAM (4) y almacena los fonemas detectados en dicha memoria RAM (4).
- 10 8. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dichos contextos (16) están almacenados en dicha memoria Flash (5) y donde al menos uno de ellos está cargado en memoria RAM (4) y es utilizado durante el reconocimiento de voz del sonido capturado y filtrado junto con dicha agenda (18).
- 15 9. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho reconocedor (7) asigna puntuaciones a dichas palabras y/o frases reconocidas y dicho analizador (8) solo acepta como comando confiable las que están por encima de un umbral en dicha puntuación.
- 20 10. Portero electrónico accesible según la reivindicación 9 **caracterizado** porque cuando dicho analizador (8) detecta un comando confiable ejecuta la acción asociada a dicho comando y se confirma a través del TTS (9).
11. Portero electrónico accesible según la reivindicación 11 **caracterizado** porque la acción asociada a dicho comando consiste en realizar dicha llamada a una vivienda o conserje por medio de dicho bus de comunicaciones (13).
- 25 12. Portero electrónico accesible según la reivindicación 11 **caracterizado** porque si se detecta que existen varios nombres o actividades que pueden coincidir con lo que se ha reconocido en dicho comando, se indica al usuario a través del TTS (9) que indique el nombre completo para garantizar la privacidad de los usuarios del sistema.
13. Portero electrónico accesible según la reivindicación 11 **caracterizado** porque dicho comando para realizar una llamada a una vivienda asociada es un comando reconocido en una expresión vocal natural..
- 30 14. Portero electrónico accesible según la reivindicación 11 **caracterizado** porque durante

el transcurso de dicha llamada a dicha vivienda o conserje se desactiva el reconocimiento de voz de dicha placa (100) y se vuelve a activar cuando finaliza dicha llamada.

- 5
15. Portero electrónico accesible según la reivindicación 9 **caracterizado** porque dicho comando es un comando de cambio de idioma y donde la acción asociada consiste en actualizar el contexto (16) en el reconocedor (7) y el modelo acústico (15) en el generador de fonemas (6) desde la memoria Flash (5) relacionado con dicho idioma.
- 10
16. Portero electrónico accesible según la reivindicación 15 **caracterizado** porque del reconocimiento en el nuevo idioma, o transcurrido un tiempo pre-configurado de ausencia de actividad, se vuelve de nuevo al idioma por defecto configurado por el instalador, actualizándose desde la memoria Flash (5) con el modelo acústico (15) y contexto (16) por defecto.
- 15
17. Portero electrónico accesible según la reivindicación 9 **caracterizado** porque dicho comando es un comando para configurar los parámetros de dicha placa (100) y donde dicha acción consiste en reproducir un menú vocal de opciones de configuración a través del TTS (9) y reconocer a continuación el siguiente comando con el parámetro y el valor a cambiar.
- 20
18. Portero electrónico accesible según la reivindicación 17 **caracterizado** porque previo a reproducir dicho menú vocal de configuración se solicita a través del TTS (9) y se reconoce un código de acceso para poder acceder de forma segura a dicho menú.
- 25
19. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho conversor de texto a voz o TTS (9), cuando es invocado para reproducir un mensaje al usuario de placa, se carga en memoria RAM (4), analiza el texto a reproducir transformándolo en un fichero de audio en formato PCM, lo envía a dicho interface analógico de salida (12) y cuando termina dicha reproducción se descarga de dicha memoria RAM (4), utilizando en cada momento el idioma activo que tiene en activo en momento.
- 30
20. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho interface analógico de salida (12) es un conversor DAC (*digital-to-analog converter*) o un PWM (*pulse width modulator*).
21. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho interface analógico de salida (12) se utiliza tanto para convertir los mensajes del TTS

(9) como el audio digital proveniente de las conversaciones que se realizan con las viviendas o conserje.

- 5
22. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicho interface analógico de salida (12) adapta el volumen del audio de salida en función de un nivel de ruido ambiente detectado por dicho detector de umbral (3).
 23. Portero electrónico accesible según la reivindicación 1 **caracterizado** porque dicha placa (100) adicionalmente incluye un detector de presencia (17) dispuesto para detectar la proximidad de una persona a dicha placa (100) y cuando se detecta dicha presencia reproduce un mensaje de ayuda sobre el uso de dicha placa (100).

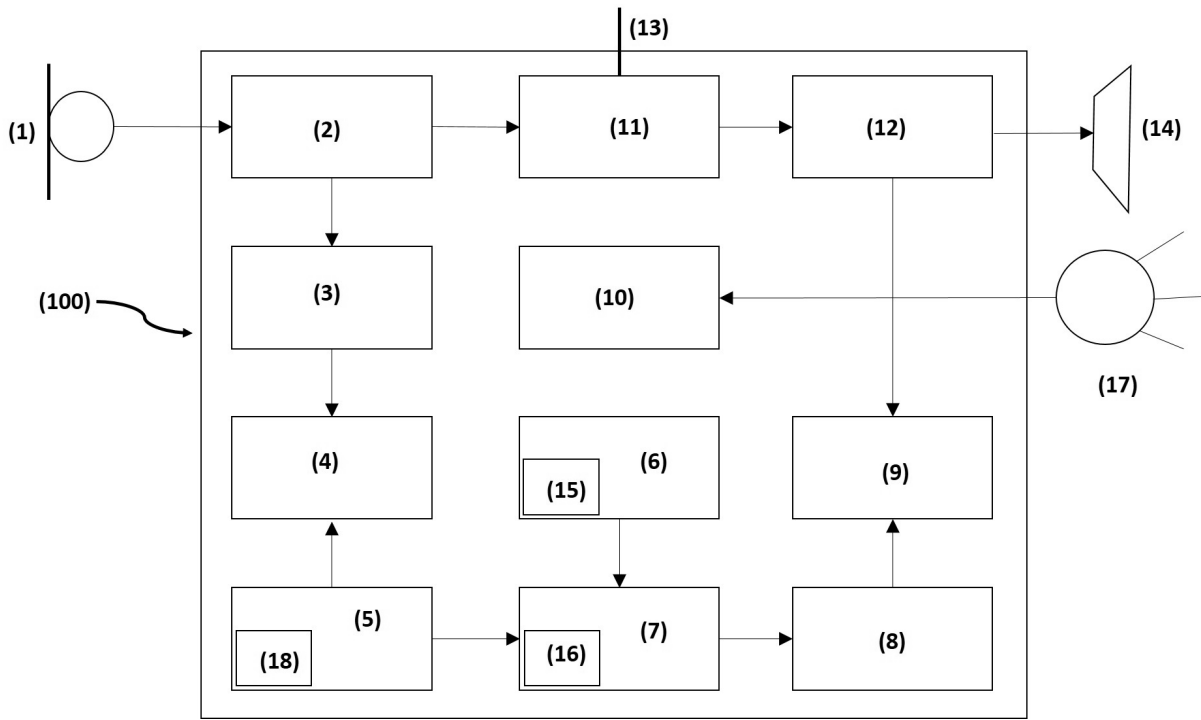


Figura 1