

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 706**

51 Int. Cl.:

G08B 25/14 (2006.01)

G06F 3/16 (2006.01)

G10L 15/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.03.2015 E 15162105 (9)**

97 Fecha y número de publicación de la concesión europea: **27.11.2019 EP 2933789**

54 Título: **Sistema de alarma de seguridad con procesamiento del habla adaptativo**

30 Prioridad:

15.04.2014 US 201414253165

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

23.06.2020

73 Titular/es:

**HONEYWELL INTERNATIONAL INC. (100.0%)
115 Tabor Road
Morris Plains, NJ 07950, US**

72 Inventor/es:

**OH, ERIC;
ADDY, KENNETH L.;
KHOT, BHARAT BALASO y
ZAKREWSKI, DAVID S.**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 768 706 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 de alarma de seguridad con procesamiento del habla adaptativo

La solicitud se refiere a sistemas de monitorización regional. Más en particular, la solicitud se refiere a ciertos sistemas que proporcionan una interfaz sencilla de utilizar para facilitar unas interacciones expandidas o más complejas del usuario con dichos sistemas.

Antecedentes

Los sistemas de alarma de seguridad tradicionales no son intuitivos de utilizar por parte de los usuarios finales. Los típicos teclados numéricos de iconos fijos no proporcionan demasiada asistencia para ayudar al usuario a interactuar con el sistema. De manera habitual, los usuarios deben memorizar un conjunto fijo de pulsaciones o botones presionados en función de un flujo del menú para introducir órdenes al sistema. La mayoría de los usuarios promedio terminan utilizando únicamente unas pocas de las órdenes básicas y pueden estar intimidados por la interfaz de usuario del sistema y no tender a utilizar otras características avanzadas del sistema.

El documento de patente número EP2083499A1 describe un sistema que tiene dispositivos de comunicación dotados de un micrófono y un altavoz. Los dispositivos de comunicación están conectados con un sistema de control de secuencias central mediante un sistema de bus y una alimentación sobre el interruptor de Ethernet que suministra la energía. El sistema de control está conectado a un centro de control y procesa la información de los dispositivos de comunicación. El centro de control está diseñado como un ordenador. El centro de control y el sistema de control están dotados de un dispositivo de identificación para una identificación precisa de los dispositivos de comunicación.

El documento de patente número US2006/090079A1 describe un sistema de seguridad que incluye una autenticación biométrica, tal como una autenticación de voz con detección de etiquetas inalámbricas para identificar tanto al usuario como a una etiqueta inalámbrica que lleva el usuario. Cuando el usuario se acerca a la entrada de un edificio que está protegido mediante el sistema de seguridad, un detector de etiquetas de una interfaz de usuario recibe un identificador desde la etiqueta. La interfaz de usuario transmite el identificador de etiqueta a una estación base para su comparación con unos identificadores de etiquetas almacenados previamente en una base de datos. Si existe coincidencia se solicita al usuario que proporcione una entrada biométrica, tal como vocalizando una frase acordada previamente. La entrada biométrica se compara con las plantillas almacenadas en la base de datos para determinar si existe coincidencia. Si existe coincidencia se desarma el sistema de seguridad. La plantilla se puede seleccionar en función del identificador de etiquetas para mejorar la velocidad de procesamiento.

El documento de patente número US2003/229500A1 describe un detector de condiciones ambientales que incorpora circuitos de reconocimiento de voz para la recepción de órdenes verbales en un lenguaje seleccionado. Las órdenes se pueden utilizar para iniciar funciones, tales como comprobaciones de diversos componentes del detector o del funcionamiento del detector en general. Los resultados de las comprobaciones se pueden comunicar de manera audible, por ejemplo, mediante tonos, de manera verbal u óptica. Las órdenes y los mensajes identificables se pueden registrar previamente y no se pueden alterar por parte del usuario. Se pueden iniciar de manera verbal otro tipo de funciones. Como alternativa, el usuario puede ampliar el conjunto de órdenes registradas previamente o puede registrar un conjunto de órdenes específico del usuario.

Sería deseable proporcionar una interfaz más sencilla de utilizar para dichos sistemas. Se puede esperar que la sencillez de utilización dé como resultado una mayor utilización de las características avanzadas de dichos sistemas.

La presente invención en sus diversos aspectos es tal como se presenta en las reivindicaciones adjuntas.

Descripción breve de los dibujos

La figura 1 ilustra un diagrama de bloques de un sistema de acuerdo con la presente.

Descripción detallada

Aunque las realizaciones expuestas pueden adoptar múltiples formas diferentes, en los dibujos se muestran sus realizaciones específicas y se describirán con detalle en la presente sobreentendiendo que la presente exposición se debe considerar como una ejemplificación de sus principios, así como también de la mejor manera de llevarla a la práctica, y no pretende limitar la solicitud o las reivindicaciones a la realización específica ilustrada.

En un aspecto de la presente, se utiliza el reconocimiento del habla con tecnología de filtrado inteligente en los sistemas de monitorización regional para interpretar órdenes audibles u orales para proporcionar unas interacciones fluidas e intuitivas entre el usuario y el sistema. Por ejemplo, el usuario puede invocar de manera fluida e intuitiva funciones de seguridad básicas y avanzadas tales como ignorar, armar, obtener estado, establecer modo de funcionamiento.

- Las realizaciones de la presente utilizan tecnología de conversión de habla a texto para procesar órdenes audibles o de voz en forma de frases de texto. Las órdenes audibles o de voz recibidas se filtran a través de unas palabras clave configuradas previamente para interpretar las órdenes de seguridad específicas al sistema de seguridad. A continuación se ejecutan esas órdenes. Algunas órdenes orales no requerirán ninguna respuesta oral y otras sí.
- 5 Para las órdenes que requieren respuestas de estado audibles, el sistema podría utilizar una combinación de archivos de audios de voz registrados previamente y de respuestas de conversión de texto a habla. Algunas órdenes, a modo de ejemplo, incluyen “armar sistema”, “desarmar sistema” y código, “salir del hogar”, “cámaras”, “mostrar tiempo” y “estado de la vivienda”. Se pueden proporcionar muchas otras órdenes.
- 10 En otro aspecto más, las realizaciones de la presente pueden aprovechar cualquier solución de conversión de habla a texto que procese una frase de texto recibida y analice toda la frase para extraer las palabras clave con el fin de compararlas con una lista de órdenes de seguridad programadas previamente y adaptativas en tiempo real. La lista de palabras programada previamente se puede almacenar en la lista de órdenes del sistema y en las tablas de descriptores del dispositivo. Algunos ejemplos de estas palabras programadas previamente incluyen “refugio”, “puerta”, “ventana”, “armar”, “ignorar”, “comprobar estado”, ...etc.
- 15 Los instaladores o usuarios podrían crear o expandir la lista de palabras adaptativas en tiempo real tecleando las palabras o vocalizando las palabras a un dispositivo de interfaz de usuario, tal como el teclado o un dispositivo móvil. Un instalador podría añadir palabras específicas para una instalación particular, de modo que no se requiera necesariamente el reconocimiento de voz, aunque aún funcionaría el reconocimiento del habla. De modo que, por ejemplo, “habitación” y “ventana” podrían estar en la base de datos de palabras fijas cargada previamente, pero el
- 20 instalador podría añadir de manera local (o por medio de una descarga) p. ej., “de Johnny”, de modo que sea reconocible una frase como “ignorar ventana de la habitación de Johnny”. La lista de vocabulario en tiempo real también se podría adaptar a la preferencia del habla, la gramática y el acento de cada usuario.
- 25 En consecuencia, la lista de vocabulario en tiempo real adaptativo puede crecer dentro de cada sistema individual en función del número de dispositivos conectados y de la frecuencia de utilización de las órdenes orales. Podría existir una base de datos local de palabras fijas y una base de datos local de palabras personalizadas del instalador o usuario final en la base de datos con la que se podría combinar para personalizar la instalación.
- 30 Las realizaciones expuestas también pueden proporcionar respuestas de voz de retorno de información y del estado de la seguridad de vuelta al usuario por medio de una combinación de frases grabadas previamente y respuestas de conversión de texto a habla. Las frases grabadas previamente se pueden almacenar previamente en los sistemas de seguridad respectivos. Algunos ejemplos incluyen, sin carácter limitante, “desarmar sistema”, “listo para armar” y “fallo en puerta delantera”.
- 35 Las capacidades de conversión de texto a habla proporcionan unas respuestas de voz mejoradas a los usuarios, cuando el sistema necesita responder, en función de una lista de vocabulario en tiempo real adaptativo. Como la lista de vocabulario en tiempo real se construye mediante un aprendizaje automático de palabras y frases adaptativas, el vocabulario se podría utilizar para construir frases de respuesta de conversión texto a habla adecuadas.
- 40 En realizaciones de la presente, la asistencia de voz automática e interactiva proporciona una ayuda oportuna para que los usuarios completen una función avanzada tal como ignorar una zona. Por ejemplo, cuando un usuario tiene la intención de ignorar una ventana pero no está seguro de cómo dirigir el sistema en una oración completa, el usuario puede comenzar preguntando al sistema “ignorar ventana”. En respuesta, el sistema puede preguntar “¿cuál?”. El usuario puede responder diciendo “la ventana de la habitación de Johnny”. El sistema en respuesta ejecuta ignorar y proporciona una confirmación de voz al usuario.
- 45 En otro aspecto más, la asistencia de voz se puede integrar en una estación y un servicio de monitorización central de seguridad desplazado para enviar mensajes a un departamento de atención al cliente y recibirlos desde este. Dichas realizaciones pueden proporcionar un procesamiento automático de las solicitudes de usuario para actualizaciones, pagos de facturas u otros servicios. De manera adicional, dichas estaciones/servicios desplazados pueden notificar a los usuarios de problemas en el servicio, problemas en la torre de comunicaciones local y similares, todos sin ninguna limitación.
- 50 La figura 1 ilustra un sistema 10 de acuerdo con la presente, El sistema 10 incluye una monitorización del sistema y una unidad de control 12. La unidad 12 se puede implementar, al menos en parte, mediante un procesador programable 12a y software de control ejecutable 12b. La unidad 12 incluye una interfaz de usuario 14, circuitos de reconocimiento y filtrado del habla, que en parte se podrían implementar mediante el procesador 12a y las instrucciones 12b.
- 55 La unidad 12 también puede incluir una base de datos 20. La base de datos 20 puede incluir palabras y frases almacenadas previamente que forman la lista de vocabulario adaptativo 20a. En la unidad 12 también se pueden incluir circuitos de voz de retorno de información 22.

- 5 Se puede instalar una pluralidad de sensores 26 en una región R y se pueden acoplar de manera cableada o inalámbrica a la unidad 12. Se puede colocar una pluralidad de actuadores 28 en la región R y se pueden acoplar de manera cableada o inalámbrica a la unidad 12. Aquellos que son expertos en la técnica sobreentenderán que los soportes respectivos 26a, 28a pueden incluir una o más redes informáticas inalámbricas, tales como Internet o una intranet.
- Una pluralidad de dispositivos de comunicación inalámbricos 34, tales como teléfonos inteligentes, tabletas y similares, pueden estar en comunicación inalámbrica por medio del soporte 34a. El soporte 34a puede incluir una o más redes informáticas inalámbricas, tales como Internet o una intranet.
- 10 Una estación o un servicio de monitorización desplazado 36 puede estar en comunicación con la unidad 12 por medio del soporte 34a. La pluralidad 26 puede incluir detectores de seguridad, tales como sensores de movimiento, detectores de rotura de cristales, así como también sensores de las condiciones ambiente, tales como sensores de humo, gas o incendio, y similares, todos sin ninguna limitación. La pluralidad 28 puede incluir dispositivos de control de equipos para controlar ventiladores, iluminación o el suministro eléctrico, por ejemplo, o dispositivos de salida de aviso de alarma o dispositivos de control de acceso a puertas, todos sin ninguna limitación.
- 15 En resumen, las realizaciones de la presente, tal como el sistema 10, utilizan tecnología de conversión de habla a texto para procesar órdenes audibles o de voz recibidas, por medio de la interfaz 14 o los dispositivos 34, en forma de frases de texto, tal como en los circuitos 18, 20. Las órdenes audibles o de voz recibidas, por medio de la interfaz 14 o los dispositivos 34, se filtran a través de palabras clave configuradas previamente para interpretar las órdenes de seguridad específicas al sistema de seguridad, por medio de los circuitos 18 y la base de datos 20. A
- 20 continuación, se ejecutan esas órdenes por medio de los circuitos de control 12 y los actuadores 28. Algunas órdenes orales no requerirán ninguna respuesta oral y otras sí. Para las órdenes que requieren respuestas de estado audibles el sistema 10 podría utilizar una combinación de archivos de audios de voz grabados previamente y respuestas de conversión de texto a habla y circuitos de voz de retorno de información 22.
- 25 En otro aspecto más, la asistencia de voz se puede integrar en una estación y un servicio de monitorización central de seguridad desplazado 36 para enviar mensajes a un departamento de atención al cliente y recibirlos desde este. Dichas realizaciones pueden proporcionar un procesamiento automático de las solicitudes de usuario para actualizaciones, pagos de facturas u otros servicios. De manera adicional, dichas estaciones/servicios desplazados 36 pueden notificar a los usuarios de problemas en el servicio, problemas en la torre de comunicaciones local y similares, todos sin ninguna limitación.
- 30 A partir de lo anterior, se observará que se pueden efectuar numerosas variaciones y modificaciones dentro del alcance de las reivindicaciones adjuntas. Se debe sobreentender que no se pretende limitar o se debería inferir limitación alguna con respecto al aparato específico ilustrado en la presente. Obviamente, se pretende que las reivindicaciones adjuntas abarquen todas esas modificaciones que se encuentran dentro del alcance de las reivindicaciones. Asimismo, los flujos lógicos representados en las figuras no requieren el orden particular mostrado,
- 35 u orden secuencial, para lograr unos resultados deseables. Se pueden proporcionar otros pasos, o se pueden eliminar pasos, de los flujos descritos, y se pueden añadir otros componentes a las realizaciones descritas o eliminarlos de estas.

REIVINDICACIONES

- 5 1. Un sistema de monitorización (12) que comprende un dispositivo de interfaz de usuario que puede operar de manera manual (14) acoplado, al menos de forma intermitente, a un componente de conversión de habla a texto (18) y a un componente de análisis de frases de texto (18) en combinación, para extraer las palabras clave del habla del usuario que coinciden con una lista de órdenes de seguridad en tiempo real programadas previamente y adaptativas, donde el sistema de monitorización está adaptado de modo que ejecute las órdenes, y donde la orden de seguridad en tiempo real adaptativa comprende una combinación de al menos una palabra de una base de datos de palabras fijas (20) cargada previamente y una palabra adaptativa en tiempo real de una lista de vocabulario en tiempo real (20a), donde el sistema de monitorización (12) está adaptado para que los usuarios añadan palabras adaptativas en tiempo real a la lista de vocabulario en tiempo real (20a) tecleando o vocalizando las palabras al dispositivo de interfaz de usuario (14).
- 10 2. Un sistema según la reivindicación 1, donde el dispositivo de interfaz comprende al menos un teclado o dispositivo móvil.
- 15 3. Un sistema según la reivindicación 2, donde la palabra adaptativa en tiempo real se puede adaptar a cada preferencia, gramática o acento del habla del usuario.
4. Un sistema según la reivindicación 3, donde la lista de vocabulario en tiempo real (20a) se puede expandir correspondientemente dentro de cada sistema individual, en función del número de dispositivos conectados y de la frecuencia de utilización de las órdenes orales.
- 20 5. Un sistema según la reivindicación 1, donde la asistencia de voz automática interactiva (22) puede proporcionar un retorno de información al usuario final para completar una orden seleccionada o una función seleccionada.
6. Un sistema según la reivindicación 5, donde se puede seleccionar una orden o función a partir de una categoría que incluye al menos, armar sistema, desarmar sistema y código, salir de la vivienda, cámaras, mostrar tiempo, mostrar estado de la vivienda, ignorar zona o ignorar ventana.
- 25 7. Un sistema según la reivindicación 5, donde cuando un usuario dirige verbalmente el sistema para ignorar una ventana, el sistema pregunta cuál, y el usuario responde especificando una ventana, y donde el sistema repite y reconoce la solicitud.
8. Un método que comprende:
 proporcionar a un sistema de monitorización (12) un dispositivo de interfaz de usuario que puede operar de manera manual (14) y que incluye un componente de conversión de habla a texto (18) y un componente de análisis de frases de texto (18) en combinación, para extraer palabras clave del habla del usuario que coinciden con una lista de órdenes de seguridad en tiempo real programadas previamente y adaptativas, donde la orden de seguridad en tiempo real adaptativa comprende una combinación de al menos una palabra de una base de datos de palabras fijas (20) cargada previamente y una palabra adaptativa en tiempo real de una lista de vocabulario en tiempo real (20a), donde los usuario añaden las palabras adaptativas en tiempo real a la lista de vocabulario en tiempo real (20a) tecleando o vocalizando las palabras al dispositivo de interfaz de usuario (14), donde el sistema de monitorización ejecuta las órdenes.
- 30 9. Un método según la reivindicación 8, donde se analizan las frases recibidas para extraer las palabras clave.
- 35 10. Un método según la reivindicación 9, que incluye comparar las palabras clave con las palabras almacenadas en la base de datos de palabras fijas (20) cargada previamente y en la lista de vocabulario en tiempo real (20a).
- 40

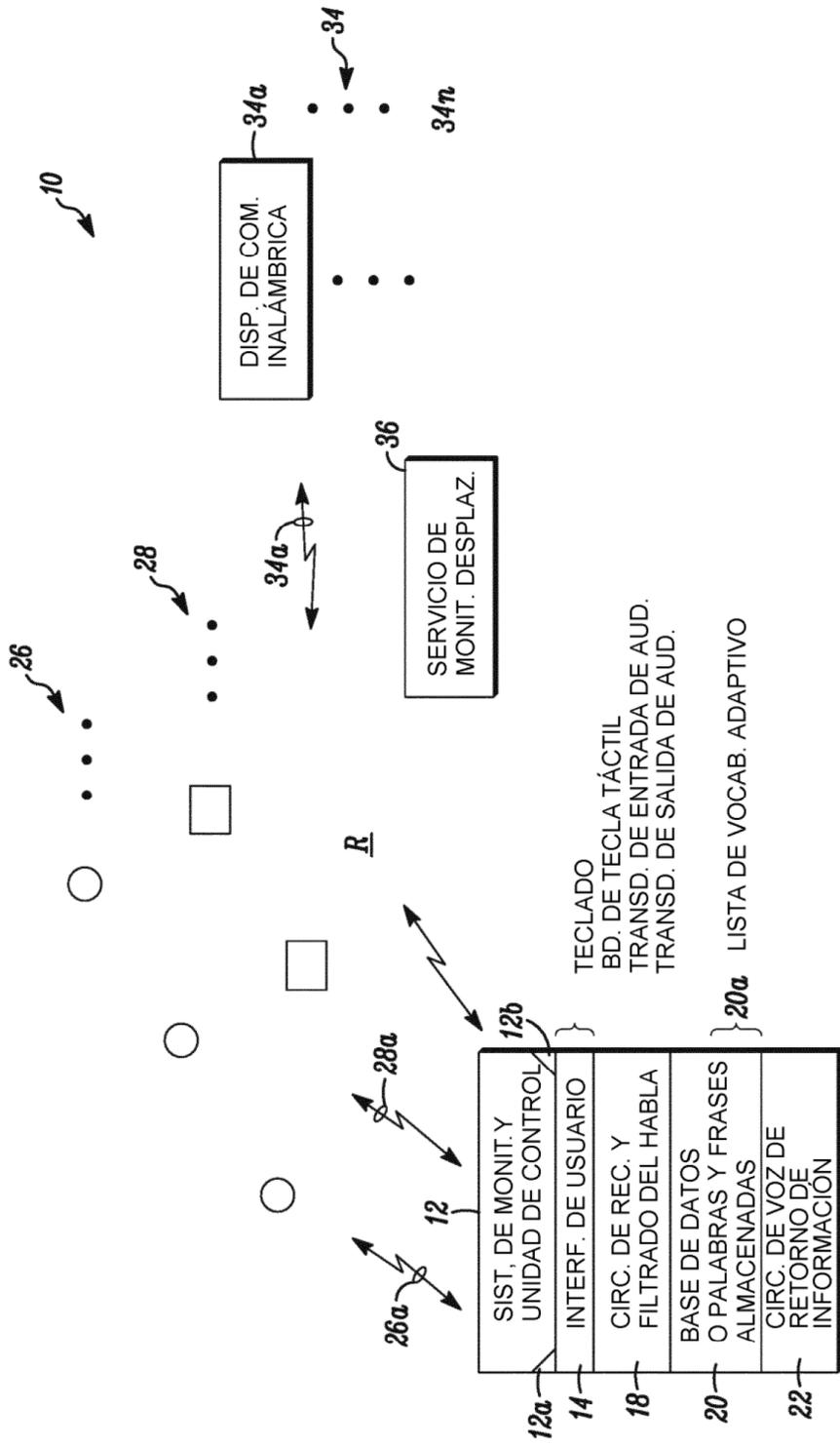


FIG. 1