



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



⑪ Número de publicación: **2 351 330**

⑫ Número de solicitud: 201031108

⑤① Int. Cl.:
G09B 21/00 (2006.01)
A61F 9/08 (2006.01)

⑫

SOLICITUD DE PATENTE

A1

⑫② Fecha de presentación: **20.07.2010**

⑫③ Fecha de publicación de la solicitud: **03.02.2011**

⑫③ Fecha de publicación del folleto de la solicitud:
03.02.2011

⑦① Solicitante/s: **Universidad Politécnica de Madrid
c/ Ramiro de Maeztu, 7
28040 Madrid, ES**

⑦② Inventor/es: **Pardo Muñoz, José Manuel;
Córdoba Herralde, Ricardo de;
Ferreiros López, Javier;
Montero Martínez, Juan Manuel y
San Segundo Hernández, Rubén**

⑦④ Agente: **Carvajal y Urquijo, Isabel**

⑤④ Título: **Método y sistema para representar la presencia de objetos en información acústica binaural.**

⑤⑦ Resumen:

Método y sistema para representar la presencia de objetos en información acústica binaural, que comprende capturar imágenes a partir de las señales de vídeo tomadas en la escena analizada; seleccionar un subconjunto de características a partir de las imágenes capturadas por las señales de vídeo y de las señales captadas por unos sensores de presencia (3) situados en la escena; segmentar cada una de las imágenes capturadas; detectar mediante un reconocimiento de patrones el tipo de cada objeto; estimar el tamaño de cada uno de los objetos; estimar la distancia a la que se encuentra cada uno de los objetos; localizar cada uno de los objetos de la escena; generar señales acústicas binaurales con frecuencia y amplitud variable; capturar señales acústicas en la escena mediante micrófonos omnidireccionales (15); combinar las señales acústicas binaurales generadas con las señales acústicas capturadas en la escena; y emitir las señales acústicas combinadas.

ES 2 351 330 A1

DESCRIPCIÓN

Método y sistema para representar la presencia de objetos en información acústica binaural.

5 Campo de la invención

La presente invención pertenece al campo técnico de las tecnologías de la información y las comunicaciones para el desarrollo de sistemas que mejoran la calidad de vida de las personas con discapacidad.

10 Estado de la técnica

En la actualidad hay una gran cantidad de trabajos relacionados con la percepción binaural que abordan tanto el problema de localización de la fuente de los eventos acústicos, como la construcción de un sistema de realidad aumentada de audio que combina la percepción acústica de la escena con otras señales generadas por el sistema. Sin embargo, las señales generadas por estos sistemas no son más que la reproducción de señales pregrabadas anteriormente por lo que no contienen información útil para el oyente.

La patente WO2010012478 A2 presenta un sistema para generar señales binaurales que simulan la localización de una fuente de eventos acústicos (no de objetos no sonoros). Dicha localización se obtiene mediante arrays de micrófonos. Sin embargo, este sistema sólo simula de forma binaural fuentes de eventos acústicos, es decir, realiza modificaciones a las señales percibidas por un array de micrófonos para acondicionarlas a la percepción binaural. En ningún caso genera señales de forma artificial con información adicional.

Por otro lado, en la actualidad existen numerosos estudios relacionados con la detección automática de objetos a partir de secuencias de vídeos recogidas de una o varias cámaras. Sin embargo, estos estudios no se han integrado en sistemas de realidad aumentada de audio.

Es por tanto deseable un método y un sistema que integre la información que reciba por distintas fuentes de información, de forma que partiendo de la detección de objetos en una escena, pueda formar una realidad aumentada del audio de dicha escena, añadiendo a las señales acústicas que se producen en el medio, unas señales acústicas binaurales generadas automáticamente que simulan la emisión artificial de eventos acústicos por parte de los objetos de la escena.

35 Descripción de la invención

La presente invención resuelve los problemas existentes en el estado de la técnica mediante un método y sistema capaz de representar acústicamente, mediante señales binaurales, la presencia de objetos detectados automáticamente, y de codificar dicha información mediante señales acústicas binaurales. Estas señales se reproducirán a través de unos auriculares con el fin de generar, en el oyente, la sensación de que dichos objetos están emitiendo dichos sonidos desde su localización real. Esta información acústica se añadirá de forma complementaria a la señal binaural de los eventos acústicos que se produzcan en la escena.

El método para representar la presencia de objetos en información acústica binaural, comprende las siguientes etapas:

- a) obtener señales de vídeo de una escena;
- b) capturar imágenes de la escena a partir de las señales de vídeo obtenidas;
- c) captar las señales que emiten unos sensores de presencia situados en la escena analizada;
- d) extraer un conjunto de características de las imágenes capturadas por las señales de vídeo, comprendiendo dicha extracción la detección de contornos de los objetos presentes en las imágenes;
- e) extraer un conjunto de características de las señales captadas por los sensores de presencia;
- f) seleccionar un subconjunto de características a partir de las características extraídas en las etapas d y e;
- g) segmentar cada una de las imágenes capturadas obteniendo los objetos presentes en la escena a partir de la detección de contornos realizado en la etapa d junto con una pluralidad de reglas previamente almacenadas en una base de datos del sistema;
- h) detectar mediante un reconocimiento de patrones el tipo de cada objeto presente en la escena analizada a partir del subconjunto de características seleccionadas y de una pluralidad de patrones previamente almacenados en una base de datos del sistema;

- i) estimar el tamaño de cada uno de los objetos detectados a partir del subconjunto de características seleccionadas y de una relación de tamaños del tipo de objeto detectado previamente almacenada en una base de datos del sistema;
- j) estimar la distancia a la que se encuentra cada uno de los objetos detectados a partir del subconjunto de características seleccionadas, del tamaño estimado de dicho objeto, del tipo de objeto detectado y de la distancia focal y la ampliación de las cámaras; en el caso de que los sensores de presencia estén situados en una buena posición, la información que ofrezcan permitirá mejorar la precisión de la estimación realizada;
- k) localizar cada uno de los objetos de la escena a partir del subconjunto de características seleccionadas, de la distancia a la que se encuentra, de su tamaño y la distancia focal y la ampliación de las cámaras; la localización de un objeto queda fijada al definir la distancia a la que se encuentra el objeto (estimada en el paso anterior) y las posiciones vertical y horizontal de dicho objeto; una vez segmentados los objetos en la escena (etapa g), las posiciones vertical y horizontal se obtienen mediante relaciones trigonométricas utilizando la información sobre la posición del objeto en la imagen, la distancia al objeto estimada y el enfoque (distancia focal y ampliación) de las cámaras; en el caso de que los sensores de presencia estén situados en una buena posición, la información que nos ofrecen permitirá mejorar la precisión de la estimación realizada;
- l) generar señales acústicas binaurales con frecuencia y amplitud variable en función del tipo de objeto detectado, de su tamaño, de la distancia a la que se encuentra y de su localización;
- m) capturar señales acústicas en la escena mediante una pluralidad de micrófonos omnidireccionales;
- n) combinar las señales acústicas binaurales generadas con las señales acústicas capturadas en la escena;
- o) emitir las señales acústicas combinadas.

Preferentemente, los sensores de presencia se seleccionan entre:

- ultrasonidos;
- infrarrojos;
- láser.

En el caso de sensores de presencia de ultrasonidos, se emite un pulso de una señal acústica y dependiendo del retardo de la señal de eco, se estima la distancia del sensor al objeto u objetos de la escena. En el caso de sensores de presencia basados en láser, el funcionamiento es similar pero en este caso se utilizan pulsos de luz en lugar de señales acústicas, consiguiendo un mayor alcance. En el caso de sensores de infrarrojos, se permiten detectar la presencia de objetos que se sitúan entre el emisor y el receptor de infrarrojos. En presencia de un objeto se produce un corte en la señal visual que se detectaría en el receptor de infrarrojos.

Preferentemente, la extracción de un conjunto de características de las señales captadas por los sensores de presencia comprende la extracción del retardo de la señal de eco, de la potencia de la señal de eco y la presencia o ausencia de la señal infrarroja en el caso de que los sensores de presencia comprendan infrarrojos.

Preferentemente, la selección de un subconjunto de características se realiza mediante LDA (Linear Discriminant Analysis) a partir de las características extraídas de las imágenes capturadas por las señales de vídeo y de las señales captadas por los sensores de presencia. También puede realizarse la selección del subconjunto de características mediante árboles de decisión.

Preferentemente, el reconocimiento de patrones comprende calcular la distancia entre el subconjunto de características seleccionadas y cada patrón almacenado en la base de datos del sistema, obteniéndose como tipo de objeto detectado aquel cuyo patrón ofrezca una menor distancia con el subconjunto de características.

El sistema para representar la presencia de objetos en información acústica binaural, comprende:

- una pluralidad de cámaras de vídeo configuradas para obtener señales de vídeo de una escena;
- una pluralidad de sensores de presencia situados en la escena analizada configurados para emitir unas señales;

ES 2 351 330 A1

- un módulo de extracción de características que comprende:
 - un módulo de extracción de características de las secuencias de vídeo configurado para capturar imágenes de la escena a partir de las señales de vídeo obtenidas y para extraer un conjunto de características de las imágenes capturadas por las señales de vídeo, comprendiendo dicha extracción la detección de contornos de los objetos presentes en las imágenes;
 - un módulo de extracción de características de los sensores de presencia configurado para captar las señales que emiten los sensores de presencia y para extraer un conjunto de características de las señales captadas por los sensores de presencia;
 - un módulo de selección de características configurado para seleccionar un subconjunto de características a partir de las características extraídas de las secuencias de vídeo y de los sensores de presencia;
- un módulo de caracterización de los objetos de la escena que comprende:
 - un módulo de segmentación y clasificación del tipo de objetos presentes en la escena configurado para segmentar cada una de las imágenes capturadas obteniendo los objetos presentes en la escena a partir de la detección de contornos y de una pluralidad de reglas previamente almacenadas en una base de datos del sistema, y para detectar, mediante un reconocimiento de patrones, el tipo de cada objeto presente en la escena analizada a partir del subconjunto de características seleccionadas y de una pluralidad de patrones previamente almacenados en una base de datos del sistema;
 - un módulo de estimación del tamaño configurado para estimar el tamaño de cada uno de los objetos detectados a partir del subconjunto de características seleccionadas y de una relación de tamaños del tipo de objeto detectado previamente almacenada en una base de datos del sistema;
 - un módulo de estimación de la distancia a la que se encuentra cada objeto configurado para estimar la distancia a la que se encuentra cada uno de los objetos detectados a partir del subconjunto de características seleccionadas, del tamaño estimado de dicho objeto, del tipo de objeto detectado y de la distancia focal y la ampliación de las cámaras;
 - un módulo de localización de los objetos configurado para localizar cada uno de los objetos de la escena a partir del subconjunto de características seleccionadas, de la distancia a la que se encuentra, de su tamaño y de la distancia focal y la ampliación de las cámaras;
- una pluralidad de micrófonos omnidireccionales configurados para capturar señales acústicas en la escena;
- un módulo que representa acústicamente la presencia de objetos en la escena que comprende:
 - un módulo que codifica información en señales binaurales configurado para generar señales acústicas binaurales con frecuencia y amplitud variable en función del tipo de objeto detectado, de su tamaño, de la distancia a la que se encuentra y de su localización;
 - un módulo de combinación de señales configurado para combinar las señales acústicas binaurales generadas con las señales acústicas capturadas en la escena.
- una pluralidad de auriculares configurados para emitir las señales acústicas combinadas.

Preferentemente, los sensores de presencia se seleccionan entre:

- ultrasonidos;
- infrarrojos;
- láser.

Preferentemente, el módulo de extracción de características de los sensores de presencia comprende la extracción de un conjunto de características de las señales captadas por los sensores de presencia, comprendiendo la extracción del retardo de la señal de eco, de la potencia de la señal de eco y la presencia o ausencia de la señal infrarroja en el caso de que los sensores de presencia comprendan infrarrojos.

Preferentemente, el módulo de selección de características está configurado para realizar la selección del subconjunto de características mediante LDA (Linear Discriminant Analysis) a partir de las características extraídas de las secuencias de vídeo y de los sensores de presencia.

No obstante, la selección del subconjunto de características también puede realizarse mediante árboles de decisión.

Preferentemente, el módulo de segmentación y clasificación del tipo de objetos presentes en la escena está configurado para realizar un reconocimiento de patrones, comprendiendo dicho reconocimiento, calcular la distancia entre el subconjunto de características seleccionadas y cada patrón almacenado en la base de datos del sistema, obteniéndose como tipo de objeto detectado aquel cuyo patrón ofrezca una menor distancia con el subconjunto de características.

El problema que resuelve esta invención es la representación acústica binaural de objetos presentes en la escena combinando las señales acústicas de la propia escena con unas señales acústicas generadas tras el procesado de las secuencias de vídeo recogidas con una o varias cámaras de vídeo y tras el procesado de la información captada por varios sensores de presencia. La invención abre expectativas a personas con discapacidad visual: permitiría desarrollar sistemas de percepción binaural que suplieran parcialmente esta discapacidad, incluso llegando al desarrollo de videojuegos accesibles para personas ciegas.

La invención propuesta podría mejorar la interacción de las personas ciegas con los entornos urbanos. Por ejemplo, en una estación de tren podría ser posible que se perciba acústicamente la presencia de algunos objetos como papeleras, bolsas, mesas o personas.

Breve descripción de los dibujos

A continuación, para facilitar la comprensión de la invención, a modo ilustrativo pero no limitativo se describirá una realización de la invención que hace referencia a una serie de figuras.

La figura 1 muestra el diagrama de bloques de la realización preferida del sistema propuesto.

Descripción detallada de un modo de realización

La realización preferente de la invención presenta el diagrama de bloques mostrado en la figura 1. Los módulos que forman esta realización son los siguientes:

- El primer módulo consiste en un módulo de extracción de características 2 de las señales de vídeo obtenidas con una o varias cámaras de vídeo 1 y de las señales de los sensores de presencia 3. Este módulo comprende:
 - un módulo de extracción de características de las secuencias de vídeo 4. En primer lugar, las secuencias de vídeo se obtienen de una o varias cámaras de vídeo 1 situadas en posiciones similares a la localización de los ojos de una persona. Estas señales de vídeo se digitalizan en secuencias de, al menos, 10 fotogramas por segundo. Para cada uno de los fotogramas obtenidos, se extrae un conjunto de características. Las características principales serían el posicionamiento de las partículas tras realizar un filtrado de partículas y la detección de bordes mediante cambios bruscos en el color, brillo o textura;
 - un módulo de extracción de características de las señales de los sensores de presencia 5. Por otro lado, se recogen varias señales de varios sensores de presencia 3, extrayendo las principales características de dichas señales. Los sensores de presencia 3 pueden incorporar diferentes tecnologías como ultrasonidos, infrarrojos o láser. En el caso de ultrasonidos, el mecanismo de funcionamiento se basa en la emisión de un pulso de una señal acústica y en la recepción del eco de dicha señal. Dependiendo del retardo del eco, se estima la distancia del sensor al objeto u objetos de la escena. Los sensores basados en láser tienen un funcionamiento similar pero en este caso utilizan pulsos de luz en lugar de señales acústicas, consiguiendo un mayor alcance. Para hacer un análisis de una escena y detectar todos los posibles objetos presentes, se debe hacer un barrido en vertical y horizontal. Los sensores de infrarrojos permiten detectar la presencia de objetos que se sitúan entre el emisor y el receptor de infrarrojos. En presencia de un objeto se produce un corte en la señal visual que se detecta en el receptor de infrarrojos,
 - un módulo de selección de características 6. Después de obtener las características de las secuencias de vídeo y de las señales de presencia, se combinan ambos vectores de características con el fin de obtener un único vector. Esta combinación de características tiene el objetivo de reducir la dimensionalidad mediante técnicas de análisis lineal discriminante o árboles de decisión.
- El segundo módulo consiste en un módulo de caracterización de los objetos de la escena 7. Este módulo comprende:
 - un módulo de segmentación y clasificación del tipo de objetos presentes en la escena 8. Según las dimensiones y la forma de cada objeto se puede detectar el tipo de objeto que es (papelera, mesa, o una persona, por ejemplo). La detección del tipo de objeto consiste en un reconocimiento de patrones que, partiendo de las características obtenidas en los módulos de extracción de características 4, 5, decide el tipo de objeto. Este sistema puede estar basado, por ejemplo, en mezclas de gaussianas (modelos Bayesianos) o redes neuronales;

ES 2 351 330 A1

- un módulo de estimación del tamaño 9 de los objetos. Conociendo el tipo de objeto y las relaciones de tamaño entre los diferentes objetos de la escena se puede hacer una estimación del tamaño del objeto. Para este módulo es muy importante la información proporcionada por los sensores de presencia de largo alcance como los láseres;
- un módulo de estimación de la distancia a la que se encuentra cada objeto 10. Conociendo el tipo de objeto, una estimación de su tamaño, la distancia focal y la ampliación (el enfoque) de las cámaras y las dimensiones del objeto percibidas por las cámaras se puede estimar la distancia al objeto detectado mediante relaciones trigonométricas,
- un módulo de localización de los objetos 11. A partir de las características extraídas de las secuencias de fotogramas (mediante técnicas como, por ejemplo, el filtro de partículas o el filtro de Kalman) y de la información de los sensores de presencia se puede localizar la presencia de objetos en la escena. Esta información, junto con la información de la distancia al objeto (obtenida en el módulo anterior), permite posicionar completamente al objeto. La evolución de esta posición a lo largo del tiempo nos pueda dar información sobre el movimiento de dicho objeto. Mientras el análisis de las secuencias de vídeo permiten obtener estimaciones groseras de la localización y de la distancia a cada uno de los objetos de la escena en todas las situaciones posibles, los sensores de presencia permiten hacer estimaciones más fiables pero sólo cuando la localización del objeto, en relación con los sensores de presencia, es apropiada.
- Finalmente, el tercer módulo representa acústicamente la presencia de objetos en la escena 12. Este módulo comprende:
 - un módulo que codifica información en señales binaurales 13. Las características de los objetos se codifican en señales binaurales de la siguiente manera:
 - El tamaño del objeto y el tipo del objeto determinarán la frecuencia o frecuencias de la señal generada, pudiendo asociar segmentos de audio (con voz o música) a la presencia de un determinado objeto.
 - La distancia estimada al objeto se puede codificar mediante la intensidad de la señal binaural generada, y también, mediante la mayor o menor cadencia de los pitidos, en el caso de que la señal generada sea intermitente.
 - La localización aproximada del objeto se codificará mediante la oportuna excitación de las señales binaurales, introduciendo las modificaciones y retardos necesarios entre una y otra señal: la percibida con uno u otro oído.
 - Finalmente, la movilidad del objeto se representa mediante variaciones de la frecuencia en coherencia con los cambios de la localización del objeto.
 - un módulo de combinación de señales 14. Las señales generadas artificialmente se deben añadir a los eventos acústicos que se produzcan en la escena (capturados por dos micrófonos omnidireccionales 15 situados en la posición de los oídos de una persona) dando lugar a un sistema de realidad aumentada de audio que genera señales binaurales por dos auriculares 16.

Una vez descrita de forma clara la invención, se hace constar que las realizaciones particulares anteriormente descritas son susceptibles de modificaciones de detalle siempre que no alteren el principio fundamental y la esencia de la invención.

REIVINDICACIONES

1. Método para representar la presencia de objetos en información acústica binaural, **caracterizado** porque comprende las siguientes etapas:

- a) obtener señales de vídeo de una escena;
- b) capturar imágenes de la escena a partir de las señales de vídeo obtenidas;
- c) captar las señales que emiten unos sensores de presencia (3) situados en la escena analizada;
- d) extraer un conjunto de características de las imágenes capturadas por las señales de vídeo, comprendiendo dicha extracción la detección de contornos de los objetos presentes en las imágenes;
- e) extraer un conjunto de características de las señales captadas por los sensores de presencia (3);
- f) seleccionar un subconjunto de características a partir de las características extraídas en las etapas d y e;
- g) segmentar cada una de las imágenes capturadas obteniendo los objetos presentes en la escena a partir de la detección de contornos realizado en la etapa d junto con una pluralidad de reglas previamente almacenadas en una base de datos del sistema;
- h) detectar mediante un reconocimiento de patrones el tipo de cada objeto presente en la escena analizada a partir del subconjunto de características seleccionadas y de una pluralidad de patrones previamente almacenados en una base de datos del sistema;
- i) estimar el tamaño de cada uno de los objetos detectados a partir del subconjunto de características seleccionadas y de una relación de tamaños del tipo de objeto detectado previamente almacenada en una base de datos del sistema;
- j) estimar la distancia a la que se encuentra cada uno de los objetos detectados a partir del subconjunto de características seleccionadas, del tamaño estimado de dicho objeto, del tipo de objeto detectado y de la distancia focal y ampliación de las cámaras;
- k) localizar cada uno de los objetos de la escena a partir del subconjunto de características seleccionadas, de la distancia a la que se encuentra, de su tamaño y la distancia focal y ampliación de las cámaras;
- l) generar señales acústicas binaurales con frecuencia y amplitud variable en función del tipo de objeto detectado, de su tamaño, de la distancia a la que se encuentra y de su localización;
- m) capturar señales acústicas en la escena mediante una pluralidad de micrófonos omnidireccionales (15);
- n) combinar las señales acústicas binaurales generadas con las señales acústicas capturadas en la escena;
- o) emitir las señales acústicas combinadas.

2. Método para representar la presencia de objetos en información acústica binaural, según la reivindicación 1, **caracterizado** porque los sensores de presencia se seleccionan entre:

- ultrasonidos;
- infrarrojos;
- láser.

3. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones anteriores, **caracterizado** porque la extracción de un conjunto de características de las señales captadas por los sensores de presencia (3), comprende la extracción del retardo de la señal de eco, de la potencia de la señal de eco y la presencia o ausencia de la señal infrarroja en el caso de que los sensores de presencia comprendan infrarrojos.

4. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones anteriores, **caracterizado** porque la selección de un subconjunto de características se realiza mediante LDA (Linear Discriminant Analysis) a partir de las características extraídas en las etapas d y e.

ES 2 351 330 A1

5. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones anteriores, **caracterizado** porque el reconocimiento de patrones comprende calcular la distancia entre el subconjunto de características seleccionadas y cada patrón almacenado en la base de datos del sistema, obteniéndose como tipo de objeto detectado aquel cuyo patrón ofrezca una menor distancia con el subconjunto de características.

6. Sistema para representar la presencia de objetos en información acústica binaural, **caracterizado** porque comprende:

- una pluralidad de cámaras de vídeo (1) configuradas para obtener señales de vídeo de una escena;
- una pluralidad de sensores de presencia (3) situados en la escena analizada configurados para emitir unas señales;
- un módulo de extracción de características (2) que comprende:
 - un módulo de extracción de características de las secuencias de vídeo (4) configurado para capturar imágenes de la escena a partir de las señales de vídeo obtenidas y para extraer un conjunto de características de las imágenes capturadas por las señales de vídeo, comprendiendo dicha extracción la detección de contornos de los objetos presentes en las imágenes;
 - un módulo de extracción de características de los sensores de presencia (5) configurado para captar las señales que emiten los sensores de presencia (3) y para extraer un conjunto de características de las señales captadas por los sensores de presencia (3);
 - un módulo de selección de características (6) configurado para seleccionar un subconjunto de características a partir de las características extraídas de las secuencias de vídeo y de los sensores de presencia (3);
- un módulo de caracterización de los objetos de la escena (7) que comprende:
 - un módulo de segmentación y clasificación del tipo de objetos presentes en la escena (8) configurado para segmentar cada una de las imágenes capturadas obteniendo los objetos presentes en la escena a partir de la detección de contornos y de una pluralidad de reglas previamente almacenadas en una base de datos del sistema, y para detectar mediante un reconocimiento de patrones el tipo de cada objeto presente en la escena analizada a partir del subconjunto de características seleccionadas y de una pluralidad de patrones previamente almacenados en una base de datos del sistema;
 - un módulo de estimación del tamaño (9) configurado para estimar el tamaño de cada uno de los objetos detectados a partir del subconjunto de características seleccionadas y de una relación de tamaños del tipo de objeto detectado previamente almacenada en una base de datos del sistema;
 - un módulo de estimación de la distancia a la que se encuentra cada objeto (10) configurado para estimar la distancia a la que se encuentra cada uno de los objetos detectados a partir del subconjunto de características seleccionadas, del tamaño estimado de dicho objeto, del tipo de objeto detectado y de la distancia focal y ampliación de las cámaras;
 - un módulo de localización de los objetos (11) configurado para localizar cada uno de los objetos de la escena a partir del subconjunto de características seleccionadas, de la distancia a la que se encuentra, de su tamaño y de la distancia focal y la ampliación de las cámaras;
- una pluralidad de micrófonos omnidireccionales (15) configurados para capturar señales acústicas en la escena;
- un módulo que representa acústicamente la presencia de objetos en la escena (12) que comprende:
 - un módulo que codifica información en señales binaurales (13) configurado para generar señales acústicas binaurales con frecuencia y amplitud variable en función del tipo de objeto detectado, de su tamaño, de la distancia a la que se encuentra y de su localización;
 - un módulo de combinación de señales (14) configurado para combinar las señales acústicas binaurales generadas con las señales acústicas capturadas en la escena.
- una pluralidad de auriculares (16) configurados para emitir las señales acústicas combinadas.

7. Sistema para representar la presencia de objetos en información acústica binaural, según la reivindicación 6, **caracterizado** porque los sensores de presencia (3) se seleccionan entre:

- ultrasonidos;
- infrarrojos;
- láser.

8. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones 6-7, **caracterizado** porque el módulo de extracción de características de los sensores de presencia (5) comprende la extracción de un conjunto de características de las señales captadas por los sensores de presencia (3), comprendiendo la extracción del retardo de la señal de eco, de la potencia de la señal de eco y la presencia o ausencia de la señal infrarroja en el caso de que los sensores de presencia comprendan infrarrojos.

9. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones 6-8, **caracterizado** porque el módulo de selección de características (6) está configurado para realizar la selección del subconjunto de características mediante LDA (Linear Discriminant Analysis) a partir de las características extraídas de las secuencias de vídeo y de los sensores de presencia (3).

10. Método para representar la presencia de objetos en información acústica binaural, según cualquiera de las reivindicaciones anteriores 6-9, **caracterizado** porque el módulo de segmentación y clasificación del tipo de objetos presentes en la escena (8) está configurado para realizar un reconocimiento de patrones, comprendiendo dicho reconocimiento, calcular la distancia entre el subconjunto de características seleccionadas y cada patrón almacenado en la base de datos del sistema, obteniéndose como tipo de objeto detectado aquel cuyo patrón ofrezca una menor distancia con el subconjunto de características.

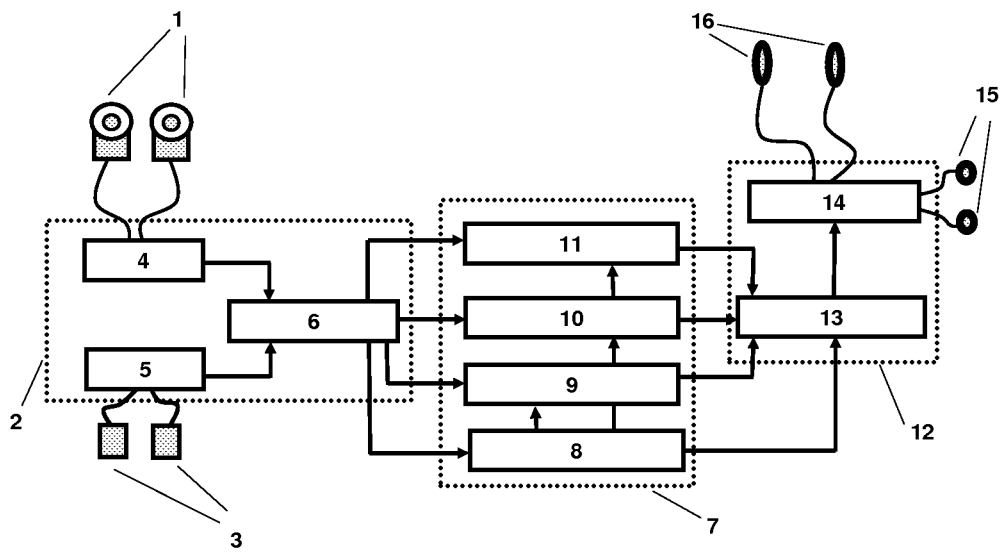


FIG. 1



OFICINA ESPAÑOLA
DE PATENTES Y MARCAS

ESPAÑA

②① N.º solicitud: 201031108

②② Fecha de presentación de la solicitud: 20.07.2010

③② Fecha de prioridad: **00-00-0000**

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: **G09B21/00**(2006.01)
A61F9/08(2006.01)

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X	US 2008058894 A1 (DEWHURST) 06.03.2008 todo el documento	1,5,6,10
A	GB 2378301 A (HEWLETT PACKARD COMPANY) 05.02.2003 todo el documento	1-10
A	US 5097326 A (MEIJER) 17.03.1992 todo el documento 05.02.2003	1-10

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

☒ para todas las reivindicaciones

☐ para las reivindicaciones nº: TODAS

Fecha de realización del informe
16.11.2010

Examinador
J. Botella Maldonado

Página
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G09B, A61F

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, NPL, XPESP, XPAIP, XPI3E, INSPEC.

Fecha de Realización de la Opinión Escrita:

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 2-4, 7-9	SI
	Reivindicaciones 1, 5, 6, 10	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 2-4, 7-9	SI
	Reivindicaciones 1, 5, 6, 10	NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	US 2008058894 A1 (DEWHURST)	06.03.2008
D02	GB 2378301 A (HEWLETT PACKARD COMPANY)	05.02.2003
D03	US 5097326 A (MEIJER)	17.03.1992

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

El documento D01 presenta un método para percibir el entorno mediante sustitución del sentido de la vista por el auditivo y el tátil; representando ciertas características de la imagen captada por una cámara, a través de elementos sonoros variables en volumen y frecuencia y de elementos táctiles. Previamente el usuario elige una serie de parámetros característicos de ciertos objetos (punto 4.1 en página 17 y en especial párrafos [0316] y [0319] en página 18). La imagen captada se segmenta en áreas de igual tamaño (párrafo [0105] página 4) y se analiza identificando contornos, formas y tamaños (párrafo [0374] en página 21, punto 4.4 en página 22, punto 4.5 en página 22 y punto 5.6 en página 27) a partir de los datos sobre objetos almacenados en una base de datos (punto 5.7 página 30). También se estima la distancia de cada objeto a partir de la distancia focal (punto 2.2 página 6) y el movimiento de algún objeto en la escena (punto 5.5 en página 26). La representación sonora es binaural con frecuencia y volumen variables según el objeto detectado, su tamaño, distancia y localización (párrafo [0098] en página 4, párrafo [0115] en página 5, párrafos del [200] en página 11 al [208] en página 12 y del [460] al [466] en página 26).

El objeto de la invención recogido en las reivindicaciones 1ª, 5ª, 6ª y 10ª, deriva directamente y sin ningún equívoco del documento D01 pues en la invención que presenta este documento, la presencia se detecta a partir del movimiento; pudiendo incluir el sistema, un sensor de infrarrojos que permite asignar valores de temperatura a zonas de la imagen captada (párrafo [0131] en página 7).

Por lo tanto las reivindicaciones 1ª, 5ª, 6ª y 10ª, carecen de novedad y actividad inventiva.