(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 597 155**

(21) Número de solicitud: 201530825

(51) Int. CI.:

A61F 9/08 (2006.01)
G01C 11/36 (2006.01)
G08C 23/00 (2006.01)

(12) SOLICITUD DE PATENTE A1

(22) Fecha de presentación:

12.06.2015

(43) Fecha de publicación de la solicitud:

16.01.2017

(56) Se remite a la solicitud internacional:

PCT/ES2016/070441

(71) Solicitantes:

EYESYNTH, S.L. (100.0%)
Ginjols, 1 - Edificio CEEI
12003 CASTELLÓN ES

(72) Inventor/es:

QUESADA HERVÁS, Antonio

(74) Agente/Representante:

TEMIÑO CENICEROS, Ignacio

(54) Título: **SISTEMA PORTÁTIL DE INTERPRETACIÓN SONORA O TÁCTIL DEL ENTORNO PARA UN INVIDENTE**

(57) Resumen:

El diseño presentado es un sistema de comprensión visual para invidentes. Está compuesto por unas gafas especiales con cámaras 3D y un software que convierte información espacial y visual en señales sonoras o táctiles comprensibles para el invidente. Las señales pueden ser sonido sintetizado que permite identificar formas y ubicarlas en el espacio. Ello permite percibir volúmenes y espacios, además con un nivel de detalle sin precedentes. El uso de auriculares cocleares o señales sonoras o táctiles permiten largas sesiones de uso, sin interferir con el oído. En el caso del sonido, es preferiblemente no verbal, lo que elimina la barrera de idioma y facilita el aprendizaje.

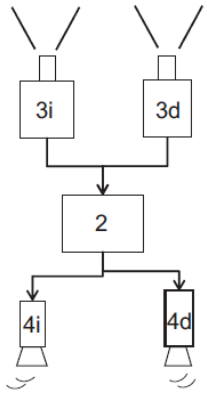

Fig. 1

ES 2 597 155 A1

SISTEMA PORTÁTIL DE INTERPRETACIÓN SONORA O TÁCTIL DEL ENTORNO PARA UN INVIDENTE

DESCRIPCIÓN

Campo técnico de la invención

La invención se relaciona con los dispositivos de asistencia a personas con alguna limitación o minusvalía. En particular, la invención se refiere a un sistema de ayuda orientado principalmente a personas ciegas y deficientes visuales.

Antecedentes de la invención o Estado de la Técnica

Tradicionalmente, los usuarios con discapacidad visual dependen de ayudas básicas, como bastones y perros guía para desplazarse o reconocer su entorno.

Aunque se han desarrollado sistemas que emplean un mayor nivel tecnológico, a menudo resultan invasivos y complejos de manejar. También suelen ser demasiado caros para que su uso deje de ser exclusivo.

Habitualmente, los sistemas actuales suelen medir distancias a un punto, por ejemplo mediante el empleo de un láser, avisando acústicamente si un objeto se interpone en la trayectoria o no. Tales sistemas no proporcionan un análisis volumétrico de la escena, ni su respuesta tiene matices asociados a cambios de posición, tamaño o geometría (curvas, aristas, posición respecto a la línea de horizonte).

Por otra parte, un análisis de una escena a través de la medición distancias con muchos puntos precisa una gran capacidad de computación que generalmente lo hace inviable para emplearse en tiempo real y/o en un dispositivo que sea portátil.

Por lo tanto, existe una necesidad de sistemas efectivos para ayudar a personas con discapacidad visual capaces de facilitar una descripción comprensible del entorno.

Breve descripción de la invención

La invención es principalmente aplicable a personas con problemas de visión. No obstante, podría ser aplicable en otro tipo de escenarios y circunstancias donde no se pueda emplear el sentido de la vista o sea preciso un sistema de guiado alternativo.

La asistencia para personas se facilita con una descripción del entorno mediante una interpretación de los objetos y obstáculos que existen alrededor que se transmite a través de una señal háptica. La generación de la señal háptica se hace a partir de un procesamiento de imágenes estéreo para obtener una representación de zonas más oscuras que se corresponden con regiones más lejanas en tanto que los otras más

claras se asocian a zonas situadas más cerca. Por señal háptica se debe entender una señal sonora o táctil (por ejemplo mediante pulsos de vibración).

Es objeto de la invención presente invención un sistema portátil de interpretación sonora o táctil del entorno para un invidente que comprende:

- dos cámaras, separadas entre sí, para captar simultáneamente una imagen del entorno,

- unos medios de procesamiento que combinan ambas imágenes en tiempo real y establecen al menos una franja vertical con información sobre la profundidad de los elementos de la imagen combinada, donde dichos medios de procesamiento además dividen la franja vertical en una pluralidad de regiones; definen, para cada región, una señal sonora o táctil en función de su profundidad y de su altura en la imagen; definen una señal sonora o táctil de salida a partir de las señales sonora o táctil de cada región de la franja vertical;

- unos medios de reproducción de la señal sonora o táctil de salida.

Preferentemente, en el modo de funcionamiento denominado de rastreo, la franja vertical es central en la imagen combinada y el usuario al moverse escanea el entorno.

Preferentemente, en el modo de funcionamiento denominado de paisaje completo, se establecen una pluralidad de franjas verticales laterales en la imagen combinada a cada lado de la franja vertical central, y se define una señal sonora o táctil lateral izquierda y lateral derecha a partir de las regiones de cada franja lateral izquierda y de cada franja lateral derecha respectivamente, el usuario sin moverse puede escanear el entorno.

Preferentemente, los medios de reproducción reproducen en estéreo combinando una señal sonora o táctil lateral izquierda y una señal sonora o táctil lateral derecha.

Preferentemente, los medios de procesamiento definen una intensidad sonora o táctil de la señal sonora o táctil en función de la altura de la región en la franja vertical.

Preferentemente, los medios de procesamiento definen una frecuencia de la señal sonora o táctil en función de la profundidad de la región.

Preferentemente, la profundidad de una región se determina en función del nivel de gris sobre un mapa de profundidad de la imagen del entorno.

Preferentemente, la región comprende al menos un píxel.

Preferentemente, el sistema comprende una estructura soporte para ser llevada por el usuario donde se pueden ubicar los medios de reproducción y las dos cámaras.

Opcionalmente, la señal sonora o táctil puede ser táctil en unas realizaciones. Por ejemplo, mediante un polímero electroactivo (EAP) o mediante una membrana de un elastómero capaz de modificar su forma en respuesta a voltaje. También de forma mecánica, mediante un pequeño motor generador de vibraciones.

Alternativamente, la señal sonora o táctil puede ser sonora en otras realizaciones.

Preferentemente, la frecuencia de dicha señal sonora se elige dentro del intervalo entre 100 Hz y 18000 Hz.

Preferentemente, los medios de reproducción son unos auriculares cocleares. Ventajosamente, deja libre el oído y se recibe la señal vía ósea. Se consigue que el

usuario pueda conversar a la vez sin que ello interfiera con la señal sonora generada o viceversa.

Preferentemente, la estructura soporte se elige entre al menos unas gafas, una diadema, soporte de cuello, soporte pectoral, soporte en hombro.

Preferiblemente, la señal sonora generada es no-verbal para evitar saturar al usuario con continuos mensajes que, tras un uso prolongado, producen molestias y cansancio. En cambio, un mensaje no-verbal es más rápido de reconocer y puede simultanearse con otras tareas. Una ventaja es que la invención es utilizable sin barreras lingüísticas.

Breve descripción de las figuras

FIG. 1 muestra un diagrama de bloques simplificado.

FIG. 2 muestra la imagen pixelada de un toroide.

FIG. 3 muestra imagen pixelada del toroide procesada.

FIG. 4 muestra un diagrama de flujo simplificado.

FIG. 5 muestra una realización de la invención como gafas.

Descripción detallada de la invención

Para mayor claridad, se describe un ejemplo de realización de la invención con referencia a las figuras sin carácter limitante y enfocado a señales sonora o táctiles sonoras.

En la **FIG. 1** se muestran varios bloques correspondientes al sistema. Las imágenes se adquieren mediante un par de cámaras **3i**, **3d** en estéreo. Preferiblemente, se sitúan a ambos lados de la cara y a la altura de los ojos del usuario para facilitar el enfoque hacia la región de interés con movimientos de cabeza. Las cámaras **3i**, **3d** están alineadas en paralelo.

La propia circuitería de las cámaras **3i**, **3d** hace un pre-procesado de la imagen captada para servir un flujo de imágenes estable, evitando artefactos y aberraciones geométricas o cromáticas. La circuitería de los sensores ofrece una pareja de imágenes sincronizada en el tiempo.

Como resultado, este flujo de vídeo se transmite hacia una unidad de proceso **2**. La unidad de proceso **2** es preferiblemente un diseño de hardware específico que implementa el algoritmo de conversión de imágenes a audio. Para comunicar las cámaras **3i**, **3d** con la unidad de proceso **2** se ha previsto un cable **6**. En otras realizaciones más complejas se contempla la transmisión inalámbrica.

La unidad de proceso **2** convierte las imágenes estereoscópicas en un mapa de profundidad en escala de grises. Previamente, se genera un mapa de disparidad (sin información de escala).

Por mapa de profundidad se entiende imagen en escala de grises en el que el color negro absoluto significa máxima lejanía (en función de la escala que usemos) y el color blanco puro significa cercanía máxima (en función de la escala que usemos). El resto de la gama de grises especifica distancias intermedias.

Por mapa de disparidad se entiende la imagen resultante que se obtiene a partir de la superposición de un par de imágenes estéreo, a las que se somete a un proceso matemático. El mapa de disparidad binocular expresa en una imagen las diferencias a nivel de pixel que hay entre dos imágenes estéreo. Se aplica un algoritmo matemático de disparidad. Teniendo la distancia entre cámaras y unos ficheros de calibración de las mismas, podemos trasladar la diferencia entre píxeles a distancias reales. Gracias a este proceso, se sabe a qué distancia de la cámara se encuentra cada porción (tamaño pixel) de la imagen tomada. Se emplea una escala de grises para expresar esa distancia.

A continuación se hace una conversión a mapa de profundidad. Tras un proceso matemático en el que se aplica escala distancias/nivel de gris, se obtiene un mapa de profundidad.

A partir del mapa de profundidad generado, se aplica un algoritmo de conversión desarrollado a tal efecto que permite que los datos espaciales de profundidad se conviertan a audio.

El resultado es que con una pareja de imágenes iniciales en estéreo, se consigue una señal sonora no-verbal en estéreo que se transmite al usuario a través de unos auriculares cocleares **4i**, **4d**. Así se logra definir un lenguaje audiovisual que traslada intuitivamente al usuario información visual a información auditiva con fidelidad.

En la **FIG. 2** aparece un ejemplo de mapa de profundidad en baja resolución de un toroide. Cada píxel del mapa de profundidad tiene asociada una coordenada (X,Y) que corresponde con las posiciones de pixel capturados por las cámaras. Además cada píxel tiene asociado un nivel de gris (G) que proporciona información sobre la profundidad, es decir la distancia a la que se encuentra la región asociada a dicho píxel.

La **FIG. 3** ilustra simplificadamente una división de la columna o franja vertical central en 3 zonas según su nivel de gris. La zona “A” es negra, la zona “B” es blanca y la zona “C” que es gris. Según lo anterior, se asocian 3 valores de intensidad diferentes a cada zona (silencio para zona “A”, volumen máximo para zona “B” y una intensidad sonora media para la zona “C”. Se ha de entender que habitualmente se definen muchos más rangos de nivel de gris y por tanto de intensidad sonora asociada. La señal sonora se compone sumando las señales individuales correspondientes a los píxeles de cada zona.

Con la información proveniente del mapa de profundidad se construye una matriz o tabla con la información del entorno en ese momento. Esta información debe ser convertida en audio de acuerdo con las siguientes consideraciones.

- Con cada par de fotogramas estéreo se hace un mapeo de disparidad: Dada la diferencia entre píxeles de las imágenes y teniendo los datos de cámaras (FOV, distancia interocular, calibración específica) se pueden establecer triangulaciones, por tanto asociar píxeles a distancias en el mundo real. Con esta información, se procesa la imagen para dar un mapa de profundidad. Es

una imagen de contorno de los objetos y escala de grises expresando sus volúmenes y distancias reales. De esta forma, tenemos una sola imagen conjunta que contiene información espacial de la escena.

- Ejemplo de funcionamiento en Modo Rastreo: Tomamos la FIG.3. Para analizar la imagen habremos de mover la cabeza de izquierda a derecha en gesto de negación. Así, el cursor central (en rojo) rastreará por completo el toroide. El sonido generado se oirá en el centro del panorama estéreo (ya que siempre será el centro del eje de nuestra visión) Con este rastreo se delimitan el tamaño horizontal del objeto (el movimiento de nuestro cuello nos servirá como referencia) y el tamaño vertical nos lo dará el rango de frecuencias.

- Modo Paisaje Completo. Tomemos para analizar la FIG.2. En este caso no se tiene que mover el cuello para interpretar lo que hay delante. La parte derecha del toroide sonará a la derecha en el panorama estéreo. Análogamente sonarán las partes centrales e izquierdas. El nivel de apertura del panorama estéreo nos indicará el tamaño horizontal del toroide. El tamaño vertical vendrá expresado por el rango de frecuencias, tal y como en el modo Rastreo.

- La correspondencia de la imagen con el sonido estéreo es la siguiente: Partiendo de imagen de un paisaje, la señal sonora corresponde a las zonas que analiza. La zona izda. de la imagen sonará en el panorama estéreo izquierdo. La zona derecha en el panorama estéreo derecho. La zona central por tanto, en el centro del panorama estéreo (o lo que es lo mismo, 50% izquierda + 50% derecha)

- El rango de frecuencias que especifican el factor altura tiene un valor de 100Hz a 18000Hz, divididos en 128 fragmentos iguales. Hemos elegido ese rango porque es lo suficientemente ancho para mostrar sonido con detalle y lo suficientemente estrecho para que una persona media los pueda abarcar sin problemas (el rango humano de sonido va de 20Hz a 20000Hz). La frecuencia base (100Hz) se asocia a la primera fila inferior de píxeles en pantalla. La frecuencia superior (18000Hz) a la fila superior de píxeles. Entre medio se asignan los demás fragmentos de frecuencia. Si la imagen tuviese 128 pixeles de altura, cada fila correspondería a un fragmento. Si cambiamos la resolución, se asignarán los fragmentos proporcionalmente a la altura. Este método sirve para sistemas con potencia computacional baja. Si disponemos de potencia bruta en la que la síntesis de sonido se genera en tiempo real, haremos la división del rango de frecuencias entre el número de pixeles de altura y asignaremos cada segmento de frecuencia a cada píxel, sin interpolaciones o promediados.

- El factor distancia espacial respecto al usuario (eje Z) se asocia al factor volumen generado por el algoritmo, de manera que un pixel negro no tendrá volumen perceptible (o sea, -infinito) y un pixel blanco tendrá el máximo volumen (0dB). Esta escala será flexible y adaptativa al uso de distintos rangos de medida (40cm, 2m, 6m)

- La duración del sonido por pixel es directamente proporcional a su “presencia” en pantalla. Si un pixel permanece continuamente blanco, el sonido se repetirá continuamente. Nota: Al ser ruido blanco (de naturaleza aleatoria) no se

perciben patrones de repetición cíclica (“loops”) o puntos de corte y repetición en el sonido.

- El análisis de columna central sólo se usa en el modo rastreo. En principio, se puede emplear una columna central de 1 píxel de anchura. No obstante, a fin de suavizar el sonido y evitar artefactos, se promediarán los valores de pixel de las 3 columnas centrales, o incluso 5, dependiendo de la resolución del mapa de profundidad (dependiente del poder de computación).

Se asocia al valor en la escala de grises de un píxel, una intensidad de volumen (I). Así, el píxel con valores 0.0.0 (modelo RGB) corresponde con una región alejada y la intensidad asociada es silencio (I=0). Un píxel con valores 255.255.255 se corresponde con una región muy próxima y el volumen de la señal es máximo (I= 0 dB). De esta manera, cada píxel puede verse como una “unidad de sonido” con la que hace una composición de audio. Preferiblemente, la frecuencia sonora abarca desde los 100Hz hasta los 18000Hz.

De acuerdo con el modo de funcionamiento, la posición X del pixel podrá interpretarse de dos maneras.

Modo rastreo: Solamente sonarán las señales correspondientes a los pixeles de la columna central (X=1/2 de la resolución horizontal de pantalla, Y=0 a Y=n, donde n es el tamaño vertical de la pantalla en píxeles). Se rastrea de la escena cuando el usuario mueve la cabeza con el gesto de negación. Esto es análogo al rastreo con un bastón.

Modo Paisaje Completo: Sonarán simultáneamente varias columnas de píxeles asociados a la escena. Con este modo, no es necesario rastrear. La imagen se representa (o “suena”) al completo. Por ejemplo, cuanto más a la derecha estén los píxeles, más sonará en la derecha del panorama estéreo. Igualmente para las regiones centrales y de la izquierda.

Nota: El modo Paisaje Completo requiere de potencia computacional alta, por lo que dependiendo de las prestaciones de la unidad de proceso **2**, en vez de sonar todas las columnas de la imagen, se puede optimizar usando 5 columnas: Central, 45º, -45º, 80º, -80º. Pueden usarse más columnas en función de la potencia de proceso.

La posición Y del pixel (altura del objeto) definirá cómo suena en términos de frecuencia: Usaremos un filtro paso-banda (o una frecuencia senoidal generada, o un sample precalculado con un rango específico de frecuencia [alternativas en función de la potencia de cálculo del dispositivo]), con el que los píxeles de la zona alta sonarán agudos y los de la zona baja sonarán graves. El espectro de sonido que abarcará cada píxel vendrá definido por el número de píxeles que tendrá Y.

<u>Ejemplo sencillo:</u>

Para aclarar cómo se hace la generación de sonido a partir de la imagen de profundidad, se presenta este ejemplo. Supongamos se ha seleccionado el modo rastreo y se ha obtenido una imagen de profundidad como la FIG. 3 en la que solamente se distinguen 3 niveles de gris como aproximación. Por tanto, en la columna central hay (de abajo a arriba):

10 píxeles negros, 12 píxeles blancos, 2 píxeles negros, 8 píxeles grises y 15 píxeles negros. Supongamos que se asigna al blanco: 0 dB; al gris: -30 dB y al negro -∞ dB.

La intensidad de la señal que en ese instante sería la mezcla analógica de todas las señales.

El usuario apreciaría diferentes frecuencias según la posición en altura del píxel. Más grave en los píxeles de menor altura y más aguda para los de mayor altura. El sonido generado por esta columna se puede dividir en una componente grave con una intensidad sonora alta (zona B) y una componente de intensidad sonora intermedia de frecuencia más aguda (zona C).

Esta señal se generaría para los dos canales izquierdo y derecho (y se reproduciría respectivamente en los auriculares **4i**, **4d**).

Cuando el usuario cambia la posición de las cámaras al girar la cabeza, se modificará la imagen de profundidad y con ello la señal sonora asociada.

En la **FIG. 4** se ve un diagrama de flujo con algunos de los pasos importantes que se llevan a cabo en la modalidad de rastreo. Un primer paso **P1** de captura de imágenes por las cámaras **3i**,**3d**, un paso de procesado **P2** para generar el mapa de profundidad, un paso de asignación **P3** para asociar frecuencia e intensidad sonora a cada píxel o grupos de píxeles de la columna central del mapa de profundidad, un paso **P4** de generación de la señal sonora resultante correspondiente a la columna central.

En la **FIG. 5** se ilustra una realización de la invención implementada en unas gafas **1**. No obstante, se puede implementar en otro tipo de productos que sirvan de soporte. Por ejemplo, puede implementarse en una gorra, diadema, soporte de cuello, soporte pectoral, soporte en hombro. La ventaja de las gafas es que son cómodas de llevar y permiten por un lado la colocación de los auriculares **4i**, **4d** en la posición deseada y por otro el enfoque preciso de las cámaras **3i**, **3d** a la región de interés. La unidad de proceso **2** se diseña para ser portada por el usuario en un bolsillo o en un cinturón. Se prevé en un futuro reducir su tamaño para integrarla junto con las gafas **1**. Al estar separadas, un cable **6** lleva la información captada por las cámaras **3i**, **3d** a la unidad de proceso **2**. Por otra parte, una vez procesada esta información, la unidad de proceso **2** transmite a los auriculares **4i**, **4d** las señales de audio correspondientes.

La cantidad de información y detalle que presenta el sonido permite identificar formas y espacios con precisión inédita hasta el momento. En los ensayos realizados con invidentes, se ha constatado que permite reconocer formas concretas por cómo es el sonido que llevan asociado tras un corto periodo de entrenamiento. Por ejemplo, botellas, vasos y platos en una mesa tienen características sonoras que permiten distinguirlos.

Para transmitir el sonido se emplea preferiblemente auriculares cocleares que permiten dejar libre el canal auditivo. Con esto mejora la comodidad del usuario, rebajando ampliamente el cansancio auditivo y siendo mucho más higiénicos para sesiones de uso prologado.

En una realización se prevé una interfaz asociada con la unidad de procesamiento **2** con un botón de selección de rango para determinar la Distancia de Análisis. Por ejemplo: cerca, normal y lejos, con distancias de 40cm, 2m y 6m, respectivamente. Al pulsar el botón se irán seleccionando distancias cíclicamente. La selección de rango sirve típicamente para adecuar el alcance a distintos escenarios y circunstancias. Por

ejemplo para ubicar objetos en una mesa: 40cm; para moverse por casa 2m; y para cruzar la calle: 6m.

En una realización se prevé que la interfaz asociada con la unidad de procesamiento **2**, disponga de un botón de Modo de Análisis. La selección entre modos será cíclica.

Modo Rastreo: Análisis solamente en la zona central de la imagen. El usuario girará la cabeza cíclicamente de izquierda a derecha rastreando la escena de un modo análogo a como lo haría con un bastón. El sonido es monoaural.

Modo Paisaje completo: El análisis de se realiza sobre toda la imagen. El sonido es estéreo. De esta manera, el usuario puede percibir las formas y los espacios en todo el campo de visión simultáneamente. Por ejemplo, a la izquierda (panorama estéreo izquierda) se percibe una columna, en el centro (panorama estéreo central) se percibe una mesa baja y a la derecha (panorama estéreo derecha) el paso está libre. Este modo de exploración es más complejo en términos de sonido, al ofrecer más información que el Modo Rastreo. Es sencillo de dominar aunque precisa algo más de entrenamiento.

REIVINDICACIONES

1. Sistema portátil de interpretación sonora o táctil del entorno para un invidente caracterizado por que comprende:

- dos cámaras (3i,3d), separadas entre sí, configuradas para captar simultáneamente una imagen del entorno,

- unos medios de procesamiento (2) configurados para combinar ambas imágenes en tiempo real y para establecer al menos una franja vertical con información sobre la profundidad de los elementos de la imagen combinada, donde dichos medios de procesamiento (2) están además configurados para dividir la franja vertical en una pluralidad de regiones; definir, para cada región, una señal sonora o táctil en función de la profundidad de la región y de la altura de la región; definir una señal sonora o táctil de salida a partir de las señales sonoras o táctiles de cada región de la franja vertical;

- unos medios de generación (4i,4d) de la señal sonora o táctil de salida.

2. Sistema según la reivindicación 1, caracterizado por que la franja vertical es central en la imagen combinada.

3. Sistema según la reivindicación 2, caracterizado por que además los medios de procesamiento (2) establecen una pluralidad de franjas verticales laterales en la imagen combinada a cada lado de la franja vertical central, y por que definen una señal sonora o táctil lateral izquierda y lateral derecha a partir de las regiones de cada franja lateral izquierda y de cada franja lateral derecha respectivamente.

4. Sistema según la reivindicación 3, caracterizado por que los medios de generación (4i,4d) operan en estéreo combinando una señal sonora o táctil lateral izquierda y una señal sonora o táctil lateral derecha.

5. Sistema según una cualquiera de las reivindicaciones anteriores, caracterizado por que los medios de procesamiento (2) definen una intensidad de la señal sonora o táctil en función de la profundidad de la región.

6. Sistema según la reivindicación 5, caracterizado por que los medios de procesamiento (2) definen una frecuencia de la señal sonora o táctil en función de la altura de la región en la franja vertical.

7. Sistema según la reivindicación 6, caracterizado por que los medios de procesamiento (2) determinan la profundidad de una región se en función de la

codificación del color en escala de grises sobre un mapa de profundidad de la imagen del entorno.

8. Sistema según la reivindicación 7, caracterizado por que la región comprende al menos un píxel.

9. Sistema según una cualquiera de las reivindicaciones anteriores, caracterizado por que comprende una estructura soporte (1) para ser llevada por el usuario configurada para ubicar los medios de reproducción (4i,4d) y las dos cámaras (3i,3d).

10. Sistema según una cualquiera de las reivindicaciones anteriores, caracterizado por que la señal sonora o táctil es una señal táctil.

11. Sistema según una cualquiera de las reivindicaciones anteriores, caracterizado por que la señal sonora o táctil es una señal sonora.

12. Sistema según la reivindicación 11, caracterizado por que la frecuencia de la señal sonora tiene se elige dentro del intervalo entre 100 Hz y 18000 Hz.

13. Sistema según la reivindicación 12, caracterizado por que los medios de generación (4i,4d) son unos auriculares cocleares.

14. Sistema según una cualquiera de las reivindicaciones anteriores, caracterizado por que la estructura soporte se elige entre al menos

- unas gafas (1),
- una diadema,
- soporte de cuello,
- soporte pectoral,
-soporte en hombro.

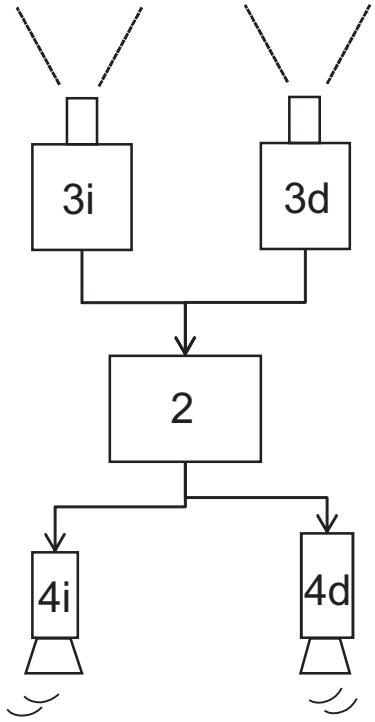

Fig. 1

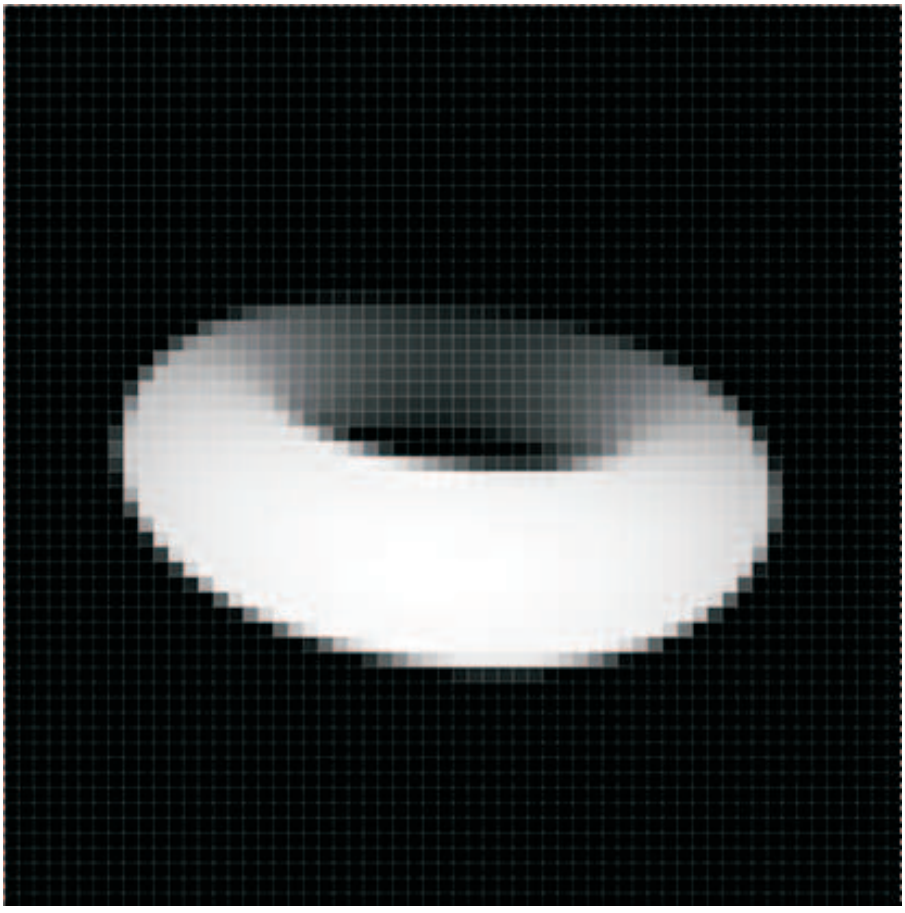

Fig. 2

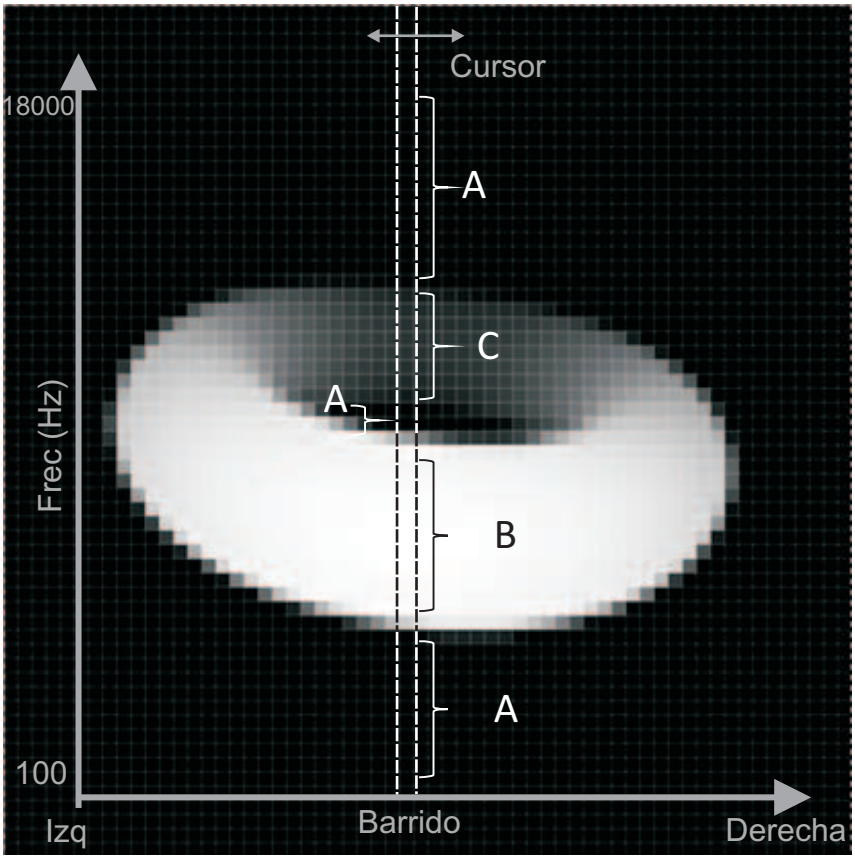

Fig. 3

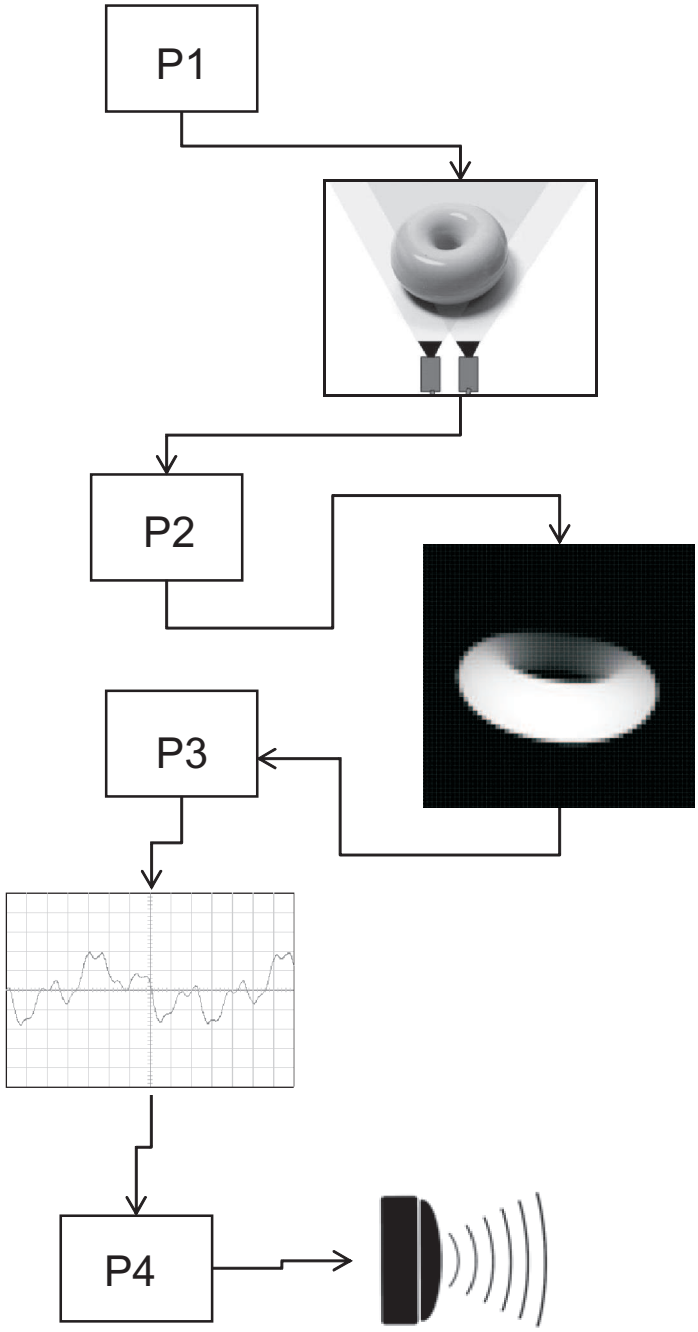

Fig. 4

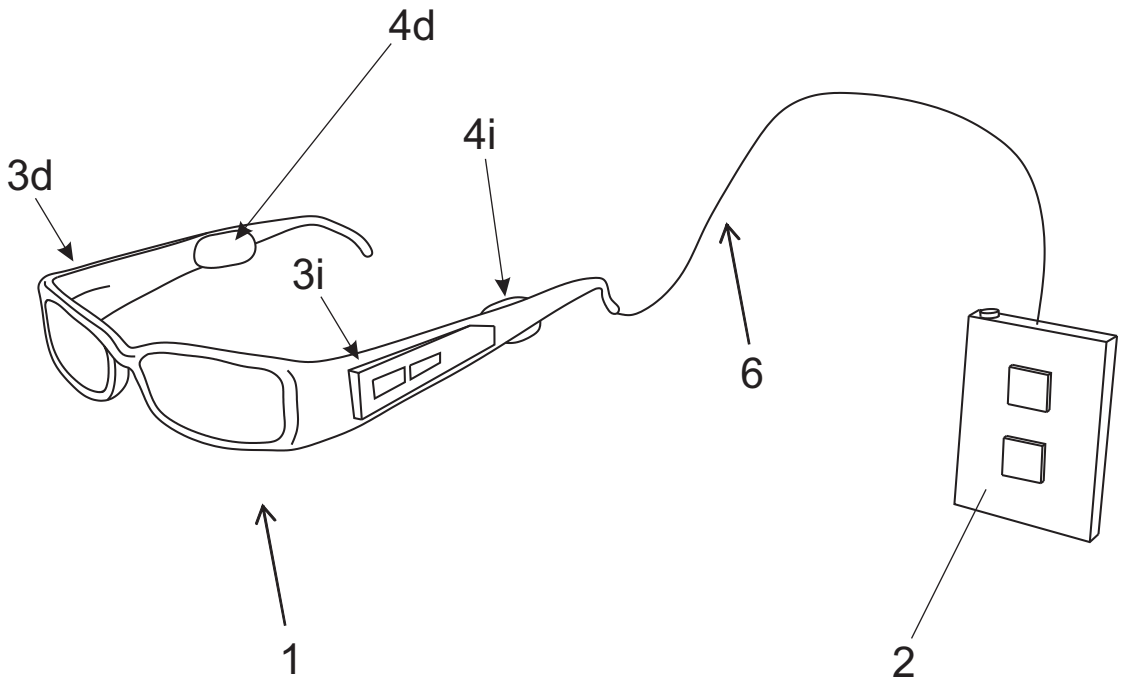

Fig. 5