

OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

① Número de publicación: **2 133 078**

② Número de solicitud: 9602283

⑤ Int. Cl.<sup>6</sup>: A61F 9/08

⑫

SOLICITUD DE PATENTE

A1

② Fecha de presentación: **29.10.96**

⑦ Solicitante/s:  
**INSTITUTO DE ASTROFISICA DE CANARIAS**  
C/ Vía Láctea, s/n  
38200 La Laguna, Sta. Cruz Tenerife, ES  
UNIVERSIDAD DE LA LAGUNA

④ Fecha de publicación de la solicitud: **16.08.99**

⑦ Inventor/es: **González Mora, José Luis y Rodríguez Ramos, Luis Fernando**

④ Fecha de publicación del folleto de la solicitud: **16.08.99**

⑦ Agente: **No consta**

⑤ Título: **Sistema para la creación de un Espacio Acústico Virtual, a tiempo real, a partir de la información proporcionada por un Sistema de Visión Artificial.**

⑤ Resumen:

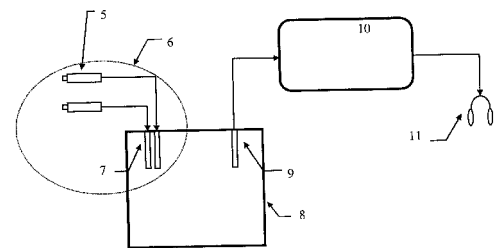
Sistema para la Creación de un Espacio Acústico Virtual, a tiempo real, a partir de la información proporcionada por un sistema de visión artificial.

Sistema óptico-electrónico que capta mediante estereovisión el entorno físico del usuario, y aprovechando las capacidades innatas de las personas para localizar sonidos, genera señales acústicas de modo que el cerebro pueda procesar dicha información sin apoyo visual.

El sistema captará imágenes mediante dos microcámaras (5), calculará la estereovisión y generará sonidos virtuales en una unidad portátil (10), y enviará esa señal acústica a unos auriculares (11), de modo que el sujeto obtenga información de la zona del espacio que se desea identificar.

Se trata de un sistema de ayuda para invidentes, y de un nuevo mecanismo de percepción en general, con el que el entorno que rodea al usuario queda definido por los sonidos generados, y no por la luz reflejada en los objetos.

Dibujo 7.-



ES 2 133 078 A1

## DESCRIPCION

Sistema para la creación de un espacio acústico virtual, a tiempo real, a partir de la información proporcionada por un sistema de visión artificial.

La presente invención se refiere a un sistema óptico-electrónico capaz de generar señales acústicas que contengan información espacial tridimensional, al objeto de que una persona pueda lograr una percepción del espacio que le rodea sin apoyarse en la visión. Se podrá utilizar como ayuda para el movimiento de las personas ciegas, pero también como un nuevo mecanismo de percepción válido para cualquier persona.

El sistema supera a los existentes hasta ahora en que se basa en la información tridimensional del entorno, captada mediante estereovisión a partir de dos microcámaras de vídeo u otro mecanismo similar, y que aprovecha las capacidades innatas de las personas para localizar sonidos y situar su foco de emisión en el espacio, como elemento fundamental sobre el que construir una percepción espacial a nivel cerebral completamente nueva.

### Antecedentes de la invención

De forma general, la investigación en la búsqueda de alternativas de percepción que ayuden a invidentes, o personas con visión residual, a ampliar el ámbito de movilidad e incluso, de una forma más ambiciosa, percibir el entorno que les rodea y orientarse en él, debe estar guiada por tres objetivos básicos.

- 1.- Orientación con respecto al entorno general (p.e ciudades etc.)
- 2.- Orientación en entornos locales (p.e. casas, edificios públicos, parques etc.)
- 3.- Detección de obstáculos peligrosos y búsqueda de un camino seguro.

El primer punto está siendo investigado en varios grupos en todo el mundo, en este sentido, a modo de ejemplo, podemos referirnos al proyecto MOBIC tendente a suministrar información al portador de su posición mediante un sistema de posicionamiento global por satélite o GPS (Global Positioning System), asociado a cartografías informatizadas.

Con respecto al tercer punto, existen muchos sistemas comerciales o en periodo de investigación, tendentes a localizar obstáculos en el camino, mediante diferentes dispositivos como ecosondas, láser etc. e informar a la persona ciega mediante señales acústicas simples o complejas. El más evolucionado de estos equipos es el "Sonic Path Finder", prototipo que fue inicialmente desarrollado por la Universidad de Nottingham en el seno del grupo Blind Mobility Research. En este proyecto el producto final pretende generar un localizador sónico de obstáculos, diseñado para operar en ambientes poco ricos, exteriores, siendo imposible de utilizar en ambientes con gran riqueza de obstáculos, p.e. en el interior de edificios, y dada su pobre resolución espacial, está diseñado para ser utilizado conjuntamente con perros guía, bastones o por personas con visión funcional.

La presente invención está enmarcada en el objetivo número dos, y no se limita a informar al

deficiente visual de su posición en relación con un mapa de una ciudad u otro entorno, ni tampoco se conforma con localizar obstáculos situados en la línea de desplazamiento, sino que de forma mucho más ambiciosa, pretende generar a nivel neuronal un mapa espacial de síntesis con respecto al entorno, obteniendo la información para su percepción mediante imágenes ópticas, proveniente de dos cámaras de vídeo, que luego son procesadas para generar una imagen acústica compleja, que pueda ser interpretada por el cerebro como "percepción espacial del entorno cercano".

Existen precedentes significativos de sistemas que buscan objetivos similares, como la patente belga BE 1.003.022 (UNIVERSITE CATHOLIQUE DE LOUVAIN), que se refiere a un "Dispositivo de Rehabilitación visual utilizando el canal auditivo". Este sistema se limita a generar un equivalente sonoro de una única imagen de vídeo, troceando ésta y asociando a cada pixel un tono de una frecuencia determinada.

También cabe citar la patente francesa FR 2.596.940 (De Tiege,P.) y la PCT WO 8200395 (THALES RESOURCES), que realizan en esencia la misma transformación cambiando el sistema de obtención de sonidos a partir siempre de una imagen única. En el mismo sentido se puede hallar la patente europea EP 410.045 (PHILIPS, inventor Peter Meijer), que realiza una tabla de correspondencia de imagen a sonido que barre horizontalmente la imagen en una escala temporal repetitiva mientras que verticalmente cada pixel contribuye con un tono cuya amplitud es proporcional a su nivel de gris.

Existen también en el estado de la técnica muchos otros sistemas basados en sensores de diversa índole (ultrasonidos, láser,...) que luego codifican sus conclusiones en sonidos que pueden ser bien palabras pregrabadas, bien tonos o similares. También se encuentran patentes como la alemana DE 4218869A1, que se refiere a una conversión de la imagen en información táctil, utilizando para ello diversos mecanismos.

La presente invención se apoya en dos ramas de la técnica radicalmente diferentes, aunándolas de cara a la consecución del objetivo perseguido. Por una parte la obtención mediante ordenador de la información 3D, también conocida como eje Z o información de profundidad, a partir de dos imágenes procedentes de dos cámaras levemente separadas angularmente, constituyendo una de las ramas de lo que se viene denominando visión artificial, de gran aplicación en procesos industriales y en robótica.

Por otra parte, el enorme desarrollo acaecido en el procesamiento digital de las señales, de especial aplicación en el procesado de señales acústicas dadas las velocidades de los procesadores disponibles en el mercado, que permiten incluso su implementación en tiempo real, han conducido a posibilitar la existencia de sistemas de simulación de sonidos tridimensionales, que se basan en el modelado del oído externo mediante un filtro digital, conocido como función HRTF (Head Related Transfer Function), de forma que unos sonidos cualesquiera, procesados mediante dichos filtros y suministrados a una persona a través de auriculares, sean percibidos por ésta como provenientes

de una determinada posición del espacio circundante al sujeto que escucha.

### Descripción de la invención

La idea que caracteriza la presente invención consiste en aprovechar las capacidades de localización del los sonidos en el espacio, innatas en la persona humana, para posibilitar una percepción del mismo, especialmente en una persona ciega. Se trata de conseguir que todos los objetos que rodean al potencial usuario sean "virtualmente" capaces de generar sonidos portadores de suficiente información como para provocar una percepción espacial -a nivel cerebral- de su entorno físico, de la misma forma que la percepción habitual a través de la visión se consigue a partir de la luz que emiten (o reflejan) los objetos que definen el mencionado entorno. Como ilustración podría pensarse en obtener un resultado análogo al que se lograría después de haber recubierto todos los objetos que rodean al deficiente visual con pequeños generadores acústicos, que produjesen sonidos representativos de sus cualidades físicas, posición, color, textura, etc.

El sistema al que se refiere la presente invención permitiría al sujeto detectar a distancia todos los objetos que le rodean, e incluso identificar muchas de sus características, puesto que virtualmente emiten vibraciones sonoras. Con esta información, tras un periodo de entrenamiento, será posible generar un mapa sensorial tridimensional de su entorno que posibilite la percepción del mismo, y por lo tanto permitirle desenvolverse en él de modo análogo a las personas con visión normal.

Desde el punto de vista neurofisiológico, tanto nuestros resultados como los provenientes de otros autores, refuerzan la posibilidad de generar, tanto en individuos ciegos (congénitos o adquiridos) o personas con baja visión, la percepción del entorno espacial mediante escenarios acústicos con información coherente y fisiológicamente adaptada a su sistema sensorial auditivo.

Con todo ello, la presente invención se describe como un sistema que capta las características físicas del entorno que rodea al usuario, (usando visión estereoscópica, por ejemplo) especialmente las distancias hasta los objetos que lo delimitan, para luego enviar a los auriculares que portará el sujeto unos sonidos que se correspondan unívocamente con dichas características, simulando múltiples puntos sonoros alrededor de los objetos, con características diferenciales relacionadas con sus propiedades físicas, y -lo que es determinante para la invención- que hayan sido procesados con las funciones de transferencia identificativas del oído externo (HRTF), de tal forma que el sujeto perciba los sonidos como procedentes de la posición del espacio que se desea identificar.

### Breve descripción de los dibujos

Para una mejor comprensión de la idea a que se refiere la presente invención, se adjuntan varios dibujos ilustrativos, que se describen a continuación.

En el dibujo 1 se representa una descomposición del espacio en "estereopíxeles". Tras haber sido reconocido mediante visión estereoscópica habrá de ser troceado hasta que alcance la

forma de bloques físicos, ya que es necesario discretizarlo para manejarlo electrónica/informáticamente. Trabajamos con un sistema de coordenadas esféricas cuyo centro es la cabeza del sujeto, y del que únicamente reviste interés el casquete esférico que se encuentra inmediatamente delante del individuo. Dicho casquete se descompondrá en estereopíxeles, cada uno de los cuales tendrá un valor de profundidad según sea el espacio que rodea al sujeto, y también podrá tener asociado propiedades físicas de interés, como el nivel de intensidad luminosa, color, textura, etc. El dibujo representa esquemáticamente un ejemplo de distribución de estereopíxeles en un campo rectangular, de 32x8 elementos, los cuales tienen ángulos diferentes desde el punto de vista del sujeto, de tal forma que la resolución en el centro del campo aparece reforzada frente a los laterales.

En los dibujos 2, 3, 4 y 5 se representa esquemáticamente un caso unidimensional en el que el sujeto (4) se encuentra en una habitación normal (2), dirigiendo su campo de visión a una puerta entreabierta (1) que da a un pasillo (3). En el dibujo 2 se esquematiza el entorno citado, y en el dibujo 3 se trazan los estereopíxeles radialmente con centro en el sujeto, tal y como se ha indicado. El dibujo 4 muestra el resultado del cálculo del valor de los estereopíxeles, dependiendo de la distancia relativa al sujeto, y finalmente el dibujo 5 muestra cuál sería la percepción resultante, que ilustra aproximadamente sobre la existencia y ubicación de las características esenciales del entorno descrito, concretamente la existencia aludida de unas paredes y una puerta entreabierta que da a un pasillo. Naturalmente, esta percepción estaría reforzada con otras estructuras similares que se obtendrían a partir de estereopíxeles situados en el eje vertical, los cuales no han sido dibujados al tratarse de una simplificación unidimensional.

El dibujo 6 muestra la situación descrita en los dibujos 2 al 5 pero con todos los estereopíxeles, tanto verticales como horizontales.

El dibujo 7 contiene el diagrama de bloques del modo de realización de la invención que se propone. Se compone de una pareja de microcámaras de vídeo (5), con salida estándar CCIR, con resolución mínima 400 líneas, que podrían ser conectadas a unas tarjetas de adquisición de imágenes de tipo "frame grabber" (7), lo que constituye el subsistema de captación del espacio tridimensional (6). El conjunto es controlado y dirigido desde el Subsistema de Control y Usuario (8) constituido por un ordenador de tipo PC dotado de gran capacidad de procesamiento mediante procesadores de tipo PENTIUM, memoria mayor de 32 Mbytes (70 ns), disco duro tipo IDE, estructura basada en Bus PCI, y monitor en color de resolución superior a 1024x768. El Subsistema de control se conecta mediante una tarjeta de interfase (9) al Subsistema de Generación y Adición de Sonidos (10) con las siguientes características:

- 256 elementos generadores de sonidos
- 64 sonidos diferentes por elemento (mínimo)
- Actualización de los sonidos cada 100 ms
- Carga de los sonidos en modo "off-line", previo a la puesta en marcha

- Muestreo igual o superior a 44,1 KHz
- Reconfigurable por programa
- Salidas de sonido independientes desde cada elemento

Finalmente el conjunto de sonidos, una vez sumados, se reproducen mediante unos auriculares profesionales (11).

El dibujo 8 ilustra sobre la constitución de un esquema idealizado de estereovisión, que sirve de apoyo a la descripción que se realiza en el texto. Puede observarse cómo a partir de dos imágenes (12 y 14) de un mismo punto contenido en un plano de Z constante (15), puede deducirse la distancia de éste hasta el plano origen del sistema (13).

El dibujo 9 muestra un esquema de la disposición en planta de las dos microcámaras que proporcionarán la información de partida para el cálculo de la visión estereoscópica.

El dibujo 10 representa un diagrama de bloques del Subsistema de Generación y Adición de Sonidos, en el que puede apreciarse su naturaleza modular, ya que se construye a partir de múltiples módulos del tipo Generador/sumador. También puede observarse cuál es el mecanismo de direccionamiento y lectura, así como la conversión digital/análogica para cada uno de los canales necesarios, izquierdo y derecho. La leyenda es la siguiente:

- (16): Bus de control y datos
- (17): Amplificador de salida de audio y filtro
- (18): Salida del canal derecho
- (19): Bus de lectura
- (20): Conexión al ordenador de control
- (21): Salida del canal izquierdo
- (22): Módulos Generador-Sumador, descritos en el dibujo 11.

El dibujo 11 muestra un diagrama de bloques un poco más detallado del módulo Generador/Sumador, donde pueden apreciarse las interrelaciones que producirán las prestaciones deseadas, entre las que cabe destacar la cantidad de memoria disponible, el ancho de los buses a utilizar, y la arquitectura global. Se han indicado en el dibujo los anchos de los buses en bits para facilitar la comprensión del mismo, así como la siguiente leyenda:

- (23): Conexión hacia los módulos siguientes, idénticos al dibujado.
- (24): Driver de tipo "tri-estado"
- (25): Memoria principal, con 2 Megamuestras de profundidad
- (26): Bus de lectura, 20 bits de ancho
- (27): Dispositivo lógico programable FPGA que realiza el control del conjunto en cuanto a direccionamiento, identificación, y buferización
- (28): Contador que recorre las muestras de un determinado sonido durante su reproducción.

- (29): Reloj de muestreo para la reproducción de las muestras del sonido
- (30): Bus de control y datos, 24 bits de ancho
- (31) Conjunto de interruptores tipo "DIL-SWITCH" para la identificación del módulo en el sistema
- (32): Amplificador de salida de audio
- (33): Conexión proveniente de los anteriores módulos del subsistema.
- (34): Bus de salida acumulada, 24 bits de ancho más 1 bit de validación
- (35): Bus de salida de sonidos de memoria, 16 bits de ancho
- (36): Bus de direccionamiento de la memoria, 21 bits de ancho
- (37): Líneas de control para los diversos elementos
- $\Sigma$ : Sumador acumulador

### Descripción de una realización propuesta

A la luz de los criterios vertidos en la descripción de la invención, resulta natural la separación de las necesidades en lo que a captación del espacio tridimensional se refiere, por una parte, y la generación de sonidos por otra. Esta división proporcionará además como ventajas añadidas la posibilidad de realizar desarrollos de modo independiente y emplazados simultáneamente en el tiempo, y también el aprovechamiento de las mejoras futuras que se decida adoptar en cada uno de los subsistemas. Además, la necesidad de que el sistema que se pretende desarrollar funcione como un conjunto, aconseja identificar como tarea independiente la inicialización, coordinación, sincronización y control de las actividades antes mencionadas.

De esta forma, el sistema a construir se dividirá desde el punto de vista conceptual en el "Subsistema de Captación", el "Subsistema de Generación de Sonidos" y el "Subsistema de Control y Usuario", teniendo el primero la misión de obtener una evaluación cuantitativa del entorno físico del deficiente visual (ó voluntario), el segundo la de producir, combinar y llevar hasta éste los sonidos que correspondan, y el tercero la organización de los dos primeros, orientado todo ello a que dichos sonidos puedan generar una percepción de dicho entorno.

Sin embargo, dada la dimensión técnica de la invención, y para minimizar su coste, se considera positivo identificar el conjunto de los subsistemas de "Captación" y de "Control y Usuario" como implementados en una única unidad procesadora, que constituirá por tanto el soporte del conjunto, y que será un ordenador del tipo PC pero dotado de una potencia de cálculo extraordinaria, al contar con dos procesadores PENTIUM, lo que permitirá ejecutar los algoritmos necesarios para realizar la captación con la velocidad y resolución deseadas, y también el comando del subsistema de "Generación de Sonidos", el cual sí que poseerá una disposición física independiente. Todo ello queda resumido y esquematizado en el dibujo 7, que presenta el diagrama de bloques del sistema completo.

### El subsistema de captación

El sistema que realiza en la práctica la presente invención necesitará obtener, por medios alternativos, las informaciones descriptivas del entorno del sujeto que éste normalmente recibiría a través de sus ojos. La emulación tecnológica inmediata la proporcionan los sistemas de estereovisión, en los cuales, además de disponerse, como en cualquier cámara, de información de distribución de intensidades en función de los ángulos acimutal y de elevación, se calcula también la profundidad de la materia respecto a las dos cámaras de observación.

Se trata, pues, de la construcción de un sistema de estereovisión aplicable al entorno y morfología descritos y capaz de proporcionar información tridimensional a una frecuencia de, al menos, diez veces por segundo.

El cálculo de la información de tercera dimensión, también conocido como coordenada "z" es en la actualidad un área de fuerte interés académico e industrial, hasta el extremo que las publicaciones especializadas están llenas de propuestas de sistemas que resuelven el problema concreto que se presenta en cada aplicación, en campos tan variados como la robótica, la cartografía o la industria en general (Nitzan, David, "Three Dimensional Vision Structure for Robot Applications", *IEEE Trans. Patt. Analysis & Mach Intell.* 1988) Los principales métodos se engloban bajo el nombre general de *Shape from X*, donde X puede ser *stereo*, *shading*, *motion*, etc. Algunos de estos métodos son activos, es decir, implican la proyección en la escena de algún tipo de información luminosa, mientras que los métodos pasivos se basan únicamente en la información contenida en la escena, sin añadidos externos. La información proporcionada en cada método también es diferente, pudiendo ser más o menos elaborada. Por ejemplo la estereovisión proporciona sólo valores absolutos de profundidad en puntos, mientras que el *shape from shading* da orientaciones relativas de superficies.

Una ventaja en nuestro caso es que, en cuanto al análisis e interpretación de imágenes se refiere, las prestaciones exigibles al subsistema de captación pueden ser bastante reducidas respecto a un sistema equivalente de utilización industrial, debido a que el objetivo final de la información es el cerebro humano, el cual puede fácilmente "aprender" a convivir con las inexactitudes que se produzcan. De hecho, ese aprendizaje se utiliza habitualmente durante la visión normal, y es lo que permite no apreciar los defectos naturales del ojo (zonas ciegas de la retina, etc.). Para nuestro caso se necesitan medidas de profundidad absolutas "en bruto", por lo que el sistema más apropiado es la estereovisión.

La estereovisión, visión estereoscópica o visión binocular, es uno de los métodos comprendidos en *Shape from X*. Su principal ventaja es que no requiere de ningún equipo especial, al contrario que otros métodos, lo que unido a su interés para comprender el sistema binocular humano explica que fuera uno de los primeros en ser estudiado. Su principal inconveniente es que requiere una gran cantidad de computación. Actualmente el empleo de procesadores específicos con cada vez

mayor capacidad de cálculo puede reducir en gran manera este inconveniente.

El funcionamiento de la visión estereoscópica es básicamente el siguiente: Se obtienen dos imágenes de una misma escena, desde dos puntos de vista diferentes (normalmente ligeramente diferentes). Una vez establecida la correspondencia entre puntos de las dos imágenes que corresponden a un mismo punto de la escena, mediante una sencilla triangulación puede hallarse la distancia de ese punto a las cámaras. Esto, que dicho así parece muy sencillo, plantea los siguientes grandes problemas:

1- *Definición del modelo de las cámaras y calibración de las mismas.* Una vez definido el modelo de cámara, se trata de hallar los parámetros que relacionen las coordenadas de la imagen en píxeles con las correspondientes a algún sistema de referencia del mundo real, en unidades métricas.

2- *Obtención de los puntos correspondientes en ambas imágenes (matching).* Es el mayor problema de la estereovisión, a causa de su ambigüedad. En principio, un punto de una imagen puede corresponder a cualquier punto de la otra. Se utilizan una serie de restricciones geométricas y físicas para tratar de reducir la ambigüedad de la correspondencia.

3- *Reconstrucción 3-D.* Una vez realizado el *matching*, el cálculo de la profundidad es inmediato, y sólo plantea algún problema de implementación práctica, ya que se obtiene mediante un sistema de ecuaciones sobredeterminado. El inconveniente puede ser que el número de puntos para el que se halló correspondencia sea escaso, lo que obliga a realizar una interpolación para la obtención de un mapa de profundidades denso.

Consideremos un sistema ideal, con dos cámaras idénticas y perfectamente alineadas (el mismo plano XY). El dibujo 8 muestra la planta de un sistema esquematizado de estereovisión. En estas condiciones se demuestra que la coordenada Z se puede calcular fácilmente mediante la expresión,

$$Z = \lambda - [(\lambda \cdot B) / (x_2 - x_1)]$$

donde  $\lambda$  es la distancia focal de las lentes de las cámaras, B es la distancia entre las cámaras y  $x_1$  y  $x_2$  son las coordenadas del objeto en cada cámara. También se puede calcular las coordenadas XY del objeto mediante expresiones igualmente sencillas. El modelo de cámara utilizado es el *pin-hole*, con un centro óptico o foco situado a una distancia  $\lambda$  del plano imagen. Es un modelo sencillo que se adapta bien a las características ópticas y geométricas de la mayor parte de las cámaras modernas. La diferencia de coordenadas  $x_2 - x_1$  entre las proyecciones de un mismo punto en las dos imágenes recibe el nombre de *disparidad*.

En la práctica nunca se presenta esta situación ideal. Muchas veces interesa que las cámaras formen un cierto ángulo, para definir un campo común más idóneo para la aplicación. Incluso en el caso de que no fuera así, y que mediante un cuidadoso montaje mecánico se consiguiera un alineamiento perfecto de las cámaras, no hay garantías de que los planos imagen (es decir los detectores) y las ópticas correspondientes estén asi-

mismo alineados. Esto supone que las relaciones entre las coordenadas X-Y-Z de un punto de la escena y las coordenadas x-y de su proyección no son las teóricas vistas con anterioridad, sino que deben obtenerse mediante un procedimiento de calibración. En alguno de estos métodos de calibración (método de la matriz de perspectiva) la relación entre coordenadas en el mundo y en la imagen no es función explícita de los parámetros del sistema ( $\lambda$ , B), sino que viene dada por una matriz de calibración que incluye todos los aspectos del mismo (modelo óptico, factores de escala, rotaciones y traslaciones de los planos imagen). La calibración se lleva a cabo tomando imágenes de un cierto número de puntos cuyas coordenadas espaciales son conocidas. El número de puntos suele ser elevado, y se minimiza el error mediante mínimos cuadrados o alguna técnica similar.

El segundo problema que se mencionó anteriormente es la correspondencia o *matching*. Se trata de encontrar cuáles son los dos puntos correspondientes en las dos imágenes diferentes del mismo escenario. El problema es *a priori* ambiguo. Cualquier punto en el segundo plano imagen podría ser el correspondiente de cualquier punto en el primero. Para atacar esta ambigüedad se usan ciertas restricciones, geométricas y físicas. La más utilizada es la restricción epipolar, basada en la geometría relativa del sistema estéreo. Es posible utilizar esta restricción para transformar el sistema en uno ideal de cámaras paralelas, de modo que los puntos correspondientes se hallen en la misma línea en ambas imágenes. Esta transformación se llama rectificación epipolar, y simplifica en gran manera el *matching* posterior.

Los algoritmos de *matching* pueden clasificarse en dos grandes categorías: basados en características, que extraen puntos de interés en la imagen (normalmente bordes), sobre los que llevan a cabo el *matching*, y basados en área, que llevan a cabo la correlación de los niveles de gris en ventanas de las distintas imágenes, considerando que en el entorno de puntos correspondientes los patrones de intensidad deben ser similares. La gran ventaja de los algoritmos basados en área es que en la actualidad existen a precios razonables dispositivos orientados al procesamiento de imagen capaces de llevar a cabo convoluciones y correlaciones en tiempo real, con prestaciones muy superiores a los procesadores de propósito general. También es posible una combinación de ambos sistemas, implementando la correlación sólo en puntos de interés previamente identificados. En nuestro caso el número de puntos de profundidad requerido no es excesivamente elevado, por lo que ésta parece ser la elección óptima. En todo caso debe tenerse en cuenta que, al contrario que la calibración, no puede decirse que el problema de la correspondencia en estereovisión esté resuelto, sino que sigue siendo un problema abierto, en el que las soluciones implementadas suelen ser en gran medida dependientes de la aplicación concreta.

La reconstrucción 3-D no plantea problemas en sí misma, pero en caso de que el número de correspondencias encontradas sea bajo es posible que deba llevarse a cabo una interpolación, que puede ser más o menos sofisticada dependiendo

de si se quieren mantener las discontinuidades en profundidad presentes en la escena. En nuestro caso el número de puntos requeridos no es alto, por lo que la interpolación no sería necesaria.

En resumen, las etapas del subsistema de visión serían:

- Calibración de cámaras. En principio un método sencillo, lineal, sería suficiente.
- Rectificación epipolar de las imágenes
- Correspondencia o *matching*. Un algoritmo de correlación sobre puntos de interés previamente determinados parece adecuado.
- Reconstrucción 3-D (sin necesidad de interpolación).

El desarrollo propuesto se realizaría sobre dos tarjetas digitalizadoras de imágenes de TV del tipo "frame grabber", con capacidad de tiempo real, a las que se asociaría un procesador de propósito general a determinar. La capacidad de cálculo de este procesador debe ser elevada si se quieren conseguir prestaciones del orden de 10 mapas de profundidad por segundo. Un sistema basado en un doble PENTIUM a 133 o 166 MHz. sería un posible candidato.

Se relacionan seguidamente los elementos esenciales que cabría esperar del Subsistema de Captación, en un listado preliminar que será ampliado y/o concretado durante las futuras etapas del diseño.

#### *Información a capturar.*

Debe suministrar un mapa de profundidades, con las características que se describen luego. Además, y aunque ello no sería imprescindible en una primera etapa, debería ser capaz de entregar información adicional sobre las dichas profundidades, tales como intensidad luminosa, color, rugosidad,...

#### *Tiempo real.*

Está expuesto anteriormente que hace falta renovar la información a un ritmo de 10 veces por segundo, lo que define el tiempo admisible para la realización de todas las adquisiciones y cálculos que fueran necesarios, en 0,1 segundos.

#### *Rango de profundidades.*

Aunque el rango ideal de funcionamiento abarcaría desde unos 20 centímetros hasta unos 20 metros (con origen en la cara del sujeto), existe un interés especial en la zona comprendida entre 20 cm y 70 cm (aprox.) que es la zona de la que se puede obtener información táctil usando únicamente el movimiento de los brazos. Por ese motivo, cabría plantear una división del problema global en dos más pequeños que facilite la resolución de los mismos tanto desde el punto de vista óptico como de procesamiento, que serían el entorno próximo (20-70 cm) y el lejano (70 cm -20 m). En general, pensamos que 8 bits de resolución podría ser un punto de partida para caracterizar las profundidades, distribuidos en forma no lineal, de manera que exista más resolución en las distancias más cortas.

#### *Campo de visión.*

La zona de interés aproximada es la situada horizontalmente en  $\pm 45^\circ$ , respecto a la línea de visión, y verticalmente en  $\pm 30^\circ$ , lo que supone

un subconjunto suficiente de lo abarcado por la visión natural humana.

#### *Resolución angular.*

Aunque no está claro cuántos elementos diferenciados pueden ser percibidos mediante audición, este número ha de ser muy inferior a las cifras de píxeles habituales en las imágenes computerizadas que se manejan con las tecnologías actuales. Una cifra tentativa podría ser 32x8 píxeles, aunque sería conveniente que los píxeles disponibles abarcasen el campo de visión de una forma no lineal, en el sentido de que existiese más resolución en la zona central que en los extremos, al estilo -una vez más- de lo que sucede en la visión humana.

#### *Exactitud.*

El sistema que genere el mapa de profundidades podrá utilizar la prerrogativa de que no se necesita una exactitud elevada, con la condición de no introducir contradicciones, ya que su salida no va a ser utilizada por ningún otro algoritmo, sino interpretada directamente por el cerebro humano, el cual es extremadamente flexible y capaz de adaptarse a sistemas con ciertos fallos o carencias, siempre que sean repetitivos. Además, hay que tener en cuenta que existirá una nueva información cada 0,1 segundos, y que el sujeto puede mover su cabeza al objeto de provocar la obtención de mejores informaciones en respuesta a las fallos o incoherencias que se puedan identificar.

#### *Interfase.*

La salida del mapa de profundidades deberá conducirse hacia los demás sistemas mediante alguna interfase adecuada, capaz de admitir el flujo de información involucrado en el caso peor. La información de salida deseable serían dos bytes por pixel, uno con la profundidad y otro con las informaciones adicionales.

#### *Consideraciones físicas.*

El sistema debería idealmente ser portátil, aunque para demostrar inicialmente la viabilidad de la presente invención, se puede utilizar un prototipo que no tendrá limitaciones físicas de ningún tipo, ni de tamaño, ni de consumo de potencia, excepto en lo que se refiere al conjunto de las dos cámaras que han de ser portadas por el sujeto en su cabeza, cuyo peso y dimensiones habrán de minimizarse.

Los elementos del Subsistema de Captación serán: dos cámaras de vídeo montadas en un soporte, dos tarjetas de adquisición de imagen (frame-grabber) en tiempo real y un ordenador de propósito general con elevada potencia de cálculo.

#### *Cámaras*

Se usarán cámaras CCD con salida de vídeo estándar. En la actualidad existen en el mercado microcámaras de tamaño y peso reducidos y bajo consumo de potencia, que serían óptimas para esta aplicación. Las cámaras deben poder ser sincronizadas exteriormente, de modo que la adquisición de imágenes sea simultánea. La disposición de las cámaras seda la mostrada en el dibujo 9.

La separación entre las cámaras (línea base B), el ángulo  $\alpha$ , y la óptica de las cámaras deben elegirse de modo que el campo de visión efectivo (común a las dos cámaras) sea el más amplio po-

sible. Valores óptimos para distancias entre 30 cm. y 10 m. son una línea base de 20 cm. y un ángulo de giro de las cámaras de 7 grados, con objetivos de 8.5 mm. El uso de focales menores aumentaría el campo pero introduciendo una distorsión radial excesiva. Una precisión mayor en los valores óptimos puede lograrse conociendo las características concretas de los detectores empleados.

Las cámaras deberán poder fijarse al soporte con la rigidez suficiente para evitar que su disposición cambie a causa de golpes o movimientos bruscos, ya que el sistema será calibrado para una disposición concreta de las cámaras, y cualquier cambio de disposición implicaría una nueva calibración.

#### *Tarjetas de adquisición de vídeo (frame-grabbers)*

Los frame-grabbers se encargan de digitalizar la señal de vídeo y pasarla al ordenador para su procesamiento. Se usarán *frame-grabbers* capaces de transferir en tiempo real las imágenes a la memoria RAM del ordenador. En la actualidad hay en el mercado gran cantidad de frame-grabbers de tiempo real, en especial para el bus PCI.

#### *Ordenador*

Un ordenador compatible PC con bus PCI es lo más apropiado debido al gran número de frame-grabbers disponibles en el mercado para estos sistemas. La gran capacidad de cálculo requerida por un sistema estéreo en tiempo real hace que un sistema basado en un doble PENTIUM a 166 MHz. sea idóneo. Una posible configuración sería 32 MB de RAM, un disco duro IDE de 1.6 GB, monitor de 17 pulgadas, tarjeta de vídeo SVGA con 4 MB de VRAM, y drives de floppy y CD-ROM.

#### *Software*

El software lleva a cabo las acciones usuales de un sistema estéreo: adquisición de las imágenes, matching y reconstrucción 3-D. Previamente se debe realizar la calibración del sistema de cámaras, para proporcionar la referencia métrica. Los resultados finales del subsistema de captación serán las coordenadas X-Y-Z de cada punto espacial considerado, respecto a un sistema de referencia con origen en el punto medio entre las cámaras.

#### *Calibración*

Se conocen diferentes métodos de calibración de cámaras para estereovisión. La calibración se lleva a cabo tomando imágenes de un plano móvil con puntos cuyas coordenadas espaciales son conocidas. El número de puntos debe ser lo bastante elevado como para permitir un tratamiento estadístico. El método de calibración puede ser más o menos sofisticado según se consideren o no diferentes términos de distorsión radial y tangencial respecto al modelo ideal de cámara. El más sencillo es el método de la matriz de perspectiva, con el que cada cámara (con una determinada posición y orientación espacial) queda caracterizada por una matriz que relaciona píxeles con unidades métricas.

#### *Adquisición de las imágenes*

Una consideración a tener en cuenta es que la adquisición de las imágenes debe ser simultánea en ambas cámaras, por lo que éstas deben estar

sincronizadas. Si las escenas son estáticas esto carece de importancia, pero aun cuando se trate de entornos controlados no debe descartarse la presencia de objetos móviles. El software de adquisición vendrá incluido con el frame-grabber.

#### *Obtención de puntos correspondientes (matching)*

Es el mayor problema de la estereovisión. Se trata de encontrar los píxeles de las dos imágenes que son imagen de un mismo punto real. Utilizando la restricción epipolar, basada en la geometría relativa del sistema de cámaras, pueden modificarse las imágenes de modo que los puntos correspondientes se hallen en la misma línea en ambas imágenes. A continuación se lleva a cabo el matching propiamente dicho.

Existen diversos métodos de matching. El que parece más adecuado para este caso es una correlación sobre puntos de interés previamente identificados, y no sobre toda la imagen. El número de correspondencias obtenidas es bajo en comparación con el número de píxeles de la imagen, pero eso se corresponde con los requerimientos de la aplicación.

#### *Reconstrucción 3-D*

El cálculo de las coordenadas espaciales de los puntos a partir de los píxeles correspondientes es inmediato, aunque requiere cierta computación ya que se obtiene mediante la resolución de un sistema de ecuaciones sobredeterminado.

#### *Interfase con el Subsistema de Control y Usuario*

Los resultados obtenidos por el Subsistema de Captación son tres coordenadas (X-Y-Z) en punto flotante para cada punto considerado. Esta información será entregada al Subsistema de Control, quien la manipulará en la manera que se determine durante las fases posteriores de diseño, para que sirva de base a la actuación del Subsistema de generación de Sonidos.

#### *El subsistema de generación de sonidos*

Se espera de este subsistema que sea capaz de generar y acumular una serie de sonidos simultáneos, que han de responder a la información sobre el entorno que le ha suministrado el Subsistema de Captación, y conducirlos hasta los auriculares que el invidente (o voluntario) llevará puestos.

Se pretende que el subsistema sea capaz de generar un número elevado de sonidos, no completamente determinado a priori pero situado entorno a los 256, aunque debe diseñarse teniendo en mente la fácil ampliación del mismo, es decir, que ésta no implique senas modificaciones de diseño sino ulteriores replicaciones de los módulos y sus interconexiones.

El subsistema estará comandado en el presente diseño conceptual desde un ordenador tipo PC, aunque deberá tener perfectamente definida la interfase hardware que permita su hipotética manipulación futura desde cualquier otro controlador, por ejemplo, una estación de trabajo o incluso un hardware dedicado.

Los sonidos deberán actualizarse en tiempo real, lo que a los efectos del presente proyecto se considera a intervalos de 100 milisegundos (10 veces por segundo), aunque sería deseable que el subsistema permitiese alcanzar velocidades de re-

novación de hasta 33 milisegundos, para así poder llegar a alcanzar el estándar de 30 imágenes por segundo que utilizan muchos sistemas comerciales de captación y tratamiento de imagen.

El número de sonidos diferentes asociados a cada estereopíxel, que acabará correspondiendo con su profundidad, será de un mínimo de 64, aunque sería muy deseable que pudiera configurarse hasta 256, lo que podría ser de utilidad en determinados experimentos.

La duración de cada uno de los sonidos almacenados ha de guardar estrecha relación con el ritmo de actualización, debiendo el subsistema ser capaz de mantener almacenadas las muestras necesarias para que -funcionando al estándar de 44.1 kHz- se puedan almacenar todas las profundidades posibles de cada estereopíxel, suponiendo que cada una de ellas se extiende durante los 100 milisegundos antes citados. Naturalmente, en el caso de que el sistema funcionara a un ritmo más elevado, este requerimiento se suavizaría proporcionalmente.

La programabilidad de los sonidos debe ser total, debiendo utilizarse las muestras en modo "raw", es decir, sin compresión ni procesado alguno durante el almacenamiento/reproducción, al objeto de garantizar la total controlabilidad de las características de los sonidos que se reproduzcan.

Se diseñará el subsistema para que opere con muestras de 16 bits, aunque sería una característica deseable que pudiera flexibilizarse para trabajar con 8 bits y duplicar así la capacidad de memoria, ya que la pérdida de calidad que ello implicaría podría ser aceptable en determinadas situaciones.

La velocidad de muestreo, a emplear en la reproducción, será una cualquiera de las siguientes: 44.100, 22.500, ó 11.250 KHz. Deberán existir filtros paso bajo en la salida que conformen la señal, y que se sintonicen adecuadamente a la frecuencia de muestreo.

Los sonidos del subsistema se cargarán en su totalidad con antelación a la utilización, y quedarán almacenados en memoria del tipo que se decida, entre RAM EEPROM o FLASH, según las consideraciones técnicas y económicas que procedan. Esta operación no necesitará de una gran velocidad, pero tiempos superiores a 5 minutos no serán admisibles.

El subsistema ofrecerá un mecanismo de comprobación que permita la verificación del correcto almacenamiento de los sonidos en las memorias, mediante lectura de las mismas por un bus independiente. Además, se desarrollarán los programas necesarios para la verificación automática del correcto estado de las memorias.

Se dispondrá una salida analógica independiente para cada uno de los estereopíxeles, obtenida directamente por conversión digital/analógica y con potencia de salida fija a un nivel regulable hasta 1 Watio. Estará también dotada de un filtro paso bajo analógico adecuado a la velocidad de muestreo que se esté utilizando.

En la medida de lo posible, las ideas de flexibilidad y reconfigurabilidad estarán presentes durante el diseño detallado, de forma que sean admisibles un máximo de posibilidades de funcionamiento, concretamente en lo que a la distribución



de la memoria disponible se refiere, dada la naturaleza de sistema prototipo que preside el presente desarrollo. A tal efecto, se sugiere la utilización de dispositivos FPGA (Lógica reprogramable) en la mayor medida posible, así como memorias modulares que puedan reubicarse en determinadas ocasiones para ofrecer diferentes posibilidades de reproducción de sonidos.

#### *Módulo Generador/Sumador*

Las consideraciones anteriores aconsejan la construcción del Subsistema de Generación de Sonidos como combinación de múltiples módulos Generador/sumador, es decir, según la estructura de bloques que aparece en el dibujo 10, y que se describe a continuación.

Al objeto de que puedan añadirse o eliminarse módulos sin modificaciones de diseño, se propone una solución basada en la concatenación de módulos y la repetición de las señales, de forma que puedan obviarse las dificultades de presentaría la manipulación de un número elevado de módulos (256) desde un único controlador. A este efecto, las señales de control y datos han de ser interpretadas en cada módulo para responder a las órdenes de direccionamiento, pero también son repetidas íntegramente hacia el siguiente módulo, de manera que los circuitos de cada módulo únicamente han de alimentar al siguiente. Naturalmente, este método de realización está limitado por el tiempo de retraso que introduce cada repetición, cuestión que quedará asegurada cuidando de que la diferencia entre el primero y el último módulo sea inferior al 10 % del tiempo de muestreo en el caso peor.

El elemento base del subsistema, que se repite tantas veces como estereopixeles se desee implementar, responde al diagrama de bloques del dibujo 11.

El direccionamiento se realizará en respuesta al estado del bus correspondiente, en comparación con el dil-switch que identifica unívocamente el módulo, y que debe tener al menos 9 bits, permitiéndose así que el subsistema contemple hasta 512 módulos simultáneamente (256 por canal).

Los sonidos se almacenarán en un conjunto modular de chips de memoria, del tipo que se determine en el diseño detallado (con excepción de la DRAM, que complicaría la circuitería), que tendrá un ancho de 16 bits para alojar las muestras con comodidad, y una longitud configurable de hasta 2 Megamuestras.

La configuración de la longitud absoluta se realizará mediante el puro insertado/desinsertado de los circuitos integrados, en los zócalos que han de existir en el circuito impreso, unido a la reprogramación de la lógica y a la reconfiguración del software.

La memoria será direccionada por una combinación configurable de líneas procedentes del registro de profundidad actual (RPA), de ocho bits, y de un registro contador de 16 bits. El registro de profundidad indicará cuál de los sonidos almacenados debe extraerse, mientras que el contador será accionado por un reloj a la frecuencia de muestreo de manera que vaya extrayendo las diversas muestras del sonido. La combinación configurable de las salidas de ambos dispositivos permitirá variar la cantidad de sonidos diferentes

generables frente a la longitud de los mismos.

El registro de profundidad siguiente (RPS) se preparará con antelación de manera que cuando el contador concluya la extracción del sonido actual, el salto al siguiente se produzca de inmediato con intervención únicamente de hardware, dejando así un colchón de la longitud de un sonido (100 ms) al ordenador de control para que programe los cambios de profundidad.

Además, durante la escritura de la memoria, tanto los registros como el contador habrán de poder funcionar en modo de carga directa e inmediata, para direccionar la célula de memoria que se desea escribir, con los datos que vendrán por el bus de datos de 16 bits de ancho. Para poder confirmar la correcta realización de la operación de escritura, el contenido de cada una de las células de memoria podrá ser leído -tras el direccionamiento directo aludido- a través de un bus específico de lectura que se extiende por la totalidad de los módulos, y que será manejado en cada momento por aquél que esté convenientemente direccionado. Dicho bus de lectura también informará del estado interno del módulo, mediante el estado combinado de 4 líneas.

La salida de la memoria será conducida hasta un sumador acumulador de 24 bits de salida, que podrá acumular sin desbordamiento la salida de los 256 módulos contemplados como número máximo por canal. Este acumulador también funcionará en modo concatenado entre todos los módulos, recibiendo cada módulo los 24 bits de las sumas anteriores cuando la línea de control indique su validez, añadiendo las 16 muestras de la salida de la memoria, y transmitiendo el resultado hacia el siguiente módulo al activar la línea de validez.

Además, la salida de la memoria se llevará hasta la entrada de un conversor Digital/Analógico, y seguidamente a un amplificador con filtrado, de forma que pueda obtenerse una salida analógica independiente en todo momento.

#### *Interfase con el Ordenador de Control*

La interconexión con el ordenador de control se realizará mediante los dos buses siguientes: "Bus de Lectura" y "Bus de Control y Datos". Físicamente las líneas de los mismos habrán de ser llevadas desde el primer módulo hasta las tarjetas de interfase de las que se dote al ordenador, en un recorrido inferior a 1.5 metros, y con los cables que se decida en el diseño detallado.

El "Bus de Lectura" devolverá al ordenador el contenido de la célula de memoria que haya sido direccionada, en el módulo que se seleccione, así como el estado interno del mismo (4 bits). Será por lo tanto siempre una entrada digital en el lado del ordenador, de 20 bits de ancho, que podrá ser leída de modo extremadamente simple a través de una tarjeta I/O de dichas dimensiones.

La velocidad de funcionamiento del bus de lectura habrá de suavizarse debido a que éste se extenderá a lo largo de todos los módulos, lo que supone unas longitudes y cargas apreciables. A este efecto, se dimensionarán los componentes del conjunto que afectan a este bus (drivers, tarjeta I/O...) de manera que se garantice un tiempo de establecimiento inferior a 250 Nanosegundos.

El "Bus de Control y Datos" tendrá un an-

cho de 24 bits, dividido conceptualmente en 16 líneas de datos y ocho de control. Las tareas que durante el diseño detallado han de realizarse mediante la manipulación de dichas líneas serán:

- Inicialización del subsistema
- Direccionamiento del módulo
- Escritura del registro RPS
- Escritura del valor inicial del contador
- Carga directa del Contador y del RPA para el direccionamiento de la memoria
- Escritura de la memoria direccionada
- Lectura de la memoria direccionada
- Lectura del estado del módulo direccionado

Además, las líneas de control del “Bus de Datos y Control” deberán conducir hasta los módulos la señal de reloj necesaria para el funcionamiento del contador, por lo que también será necesario contar con una tarjeta contadora en el interior del Ordenador de Control.

#### *El Subsistema de Control y Usuario*

Este subsistema está compuesto por los programas necesarios para realizar la gestión del conjunto de los dos subsistemas anteriores, programas que serán ejecutados por el ordenador PC que constituye la base del sistema. No contiene por tanto ningún hardware especialmente asignado.

Se le asignan también las actividades de interfase con el usuario, las cuales deberán estar lo bastante elaboradas como para que permitan que un usuario entrenado, como por ejemplo uno cualquiera de los miembros del equipo de inves-

tigación, sea capaz de operar correctamente el sistema. No será por tanto necesario contemplar la utilización del mismo por parte de un usuario cualquiera, lo que simplificará en gran medida la realización de la interfase.

Concretamente, deberá permitir que el usuario realice las siguientes tareas:

- Inicialización del sistema
- Verificación del correcto funcionamiento de los subsistemas
- Determinación de los parámetros de actuación para los subsistemas de Captación y de Generación de Sonidos
- Visualización de la actividad del sistema, mediante la representación gráfica de los parámetros más significativos
- Registro de la actividad del sistema, mediante la introducción en una base de datos de los nombres del personal que desarrolla el experimento, fechas y horas, comentarios, etc.
- Puesta en marcha y parada del sistema

Cuando el sistema esté puesto en marcha, este subsistema debe responsabilizarse de la gestión de los programas que constituyen el Subsistema de Captación, así como de la correcta actuación de la interfase con el Subsistema de Generación de Sonidos, transmitiéndole la información resultado de la captación, y haciendo uso de los elementos de verificación del correcto funcionamiento que éste pueda ofrecer.

## REIVINDICACIONES

1. Sistema óptico-electrónico para la creación de un espacio acústico virtual, a tiempo real, a partir de la información proporcionada por un sistema de visión artificial, **caracterizado** por la generación de dos señales acústicas, conteniendo información suficiente para crear la percepción tridimensional de un espacio físico cualquiera, mediante el procesado de dos señales de vídeo.

2. Sistema según la reivindicación anterior, **caracterizado** por la conversión de una imagen estereoscópica en una representación tridimensional de sonidos a través de unos auriculares estereofónicos, al objeto de que una persona (con o sin deficiencia visual) pueda lograr una percepción del espacio que le rodea sin apoyarse en la visión.

3. Sistema según la reivindicación 1 o 2 **caracterizado** porque permite, mediante dos micrófonos separadas angularmente y situadas en una montura de unas gafas, conocer la profundidad (eje Z) de cada pixel o grupo de ellos, generándose una imagen tridimensional del entorno, con las especificaciones necesarias para servir de soporte de cara al posterior procesado acústico y una velocidad de generación de cada imagen estereoscópica de 1/10 de segundo.

4. Sistema según reivindicaciones preceden-

tes, **caracterizado** porque tomando la imagen estereoscópica producida según se **caracteriza** en la reivindicación 3, se genera un código de sonidos combinando frecuencias, amplitudes y timbres determinados para cada posición, azimutal, elevación y profundidad (X, Y y Z) de cada pixel o grupo de píxeles.

5. Sistema según cualquiera de las reivindicaciones 1 a 4, **caracterizado** porque permite la simulación de sonidos localizados tridimensionalmente, que se basa en el modelado del oído externo en un ambiente acústico enriquecido mediante un filtro digital, conocido como función HRTF, de forma que un sonido monoaural obtenido según el método descrito en la reivindicación 4, procesado mediante dichos filtros y suministrados a través de auriculares a una velocidad de 1/10 segundo, sean percibidos como externos y provenientes de una determinada posición del espacio que circunda al sujeto que escucha.

6. Sistema como el descrito en todas las reivindicaciones anteriores, **caracterizado** porque el dispositivo es portable, de bajo consumo, y operado a baterías que permita ser transportado por una persona, con la posibilidad añadida de contener información tomada por sistemas de posicionamiento GPS.

5

10

15

20

25

30

35

40

45

50

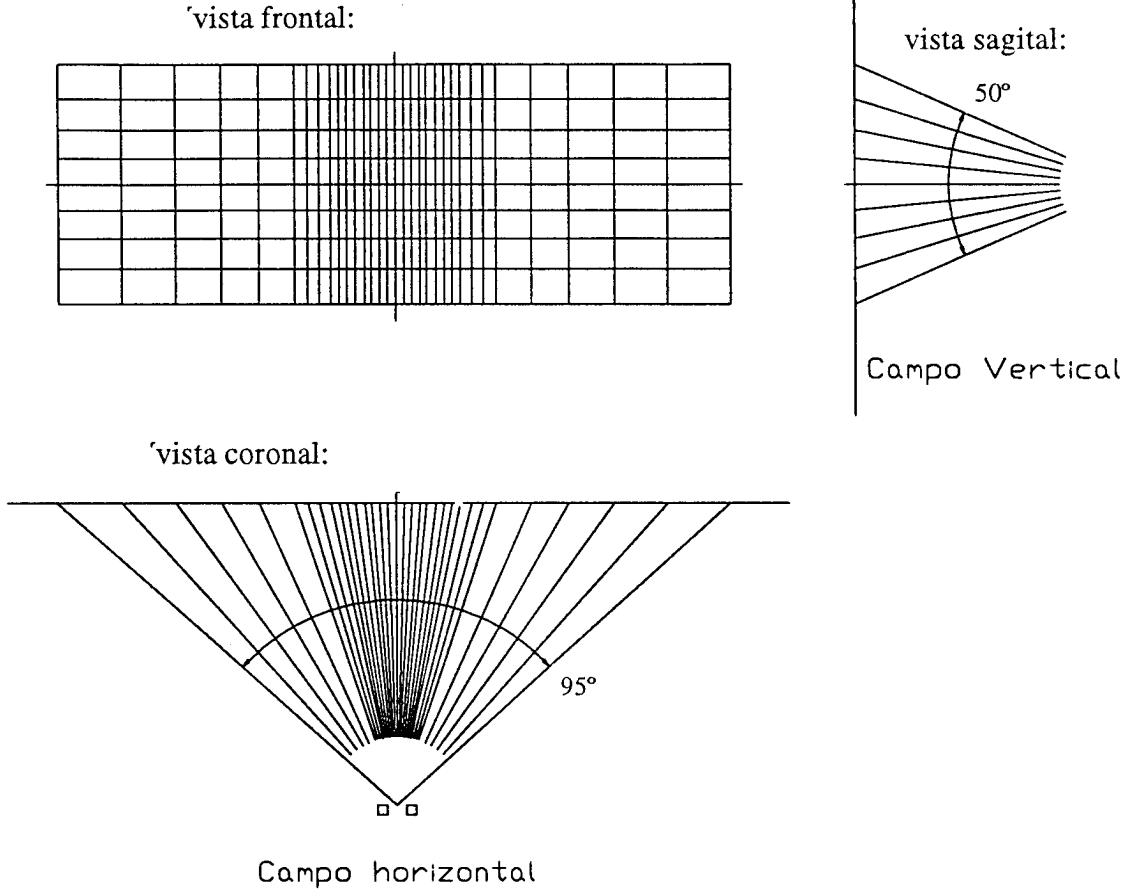
55

60

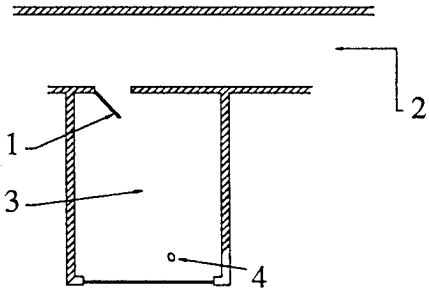
65

# DIBUJOS

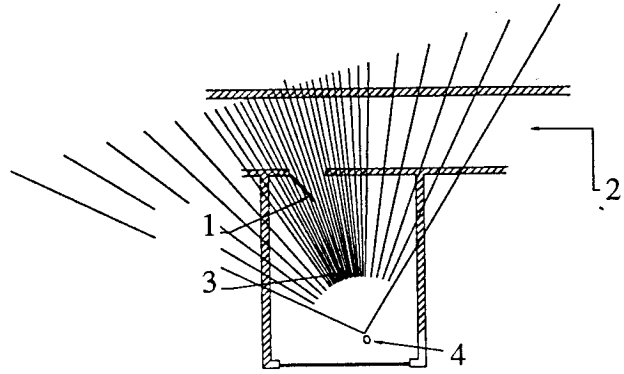
Dibujo 1.-



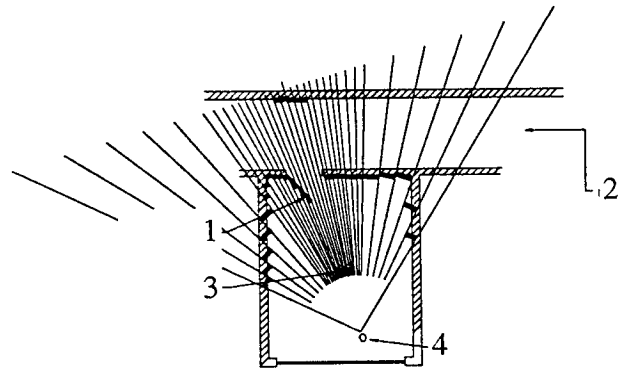
Dibujo 2.-



Dibujo 3.-



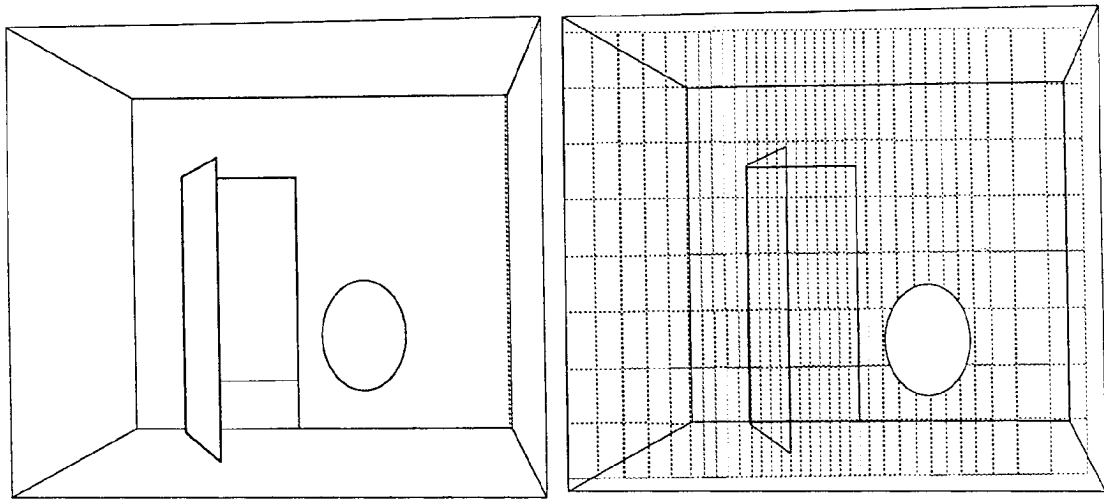
Dibujo 4.-



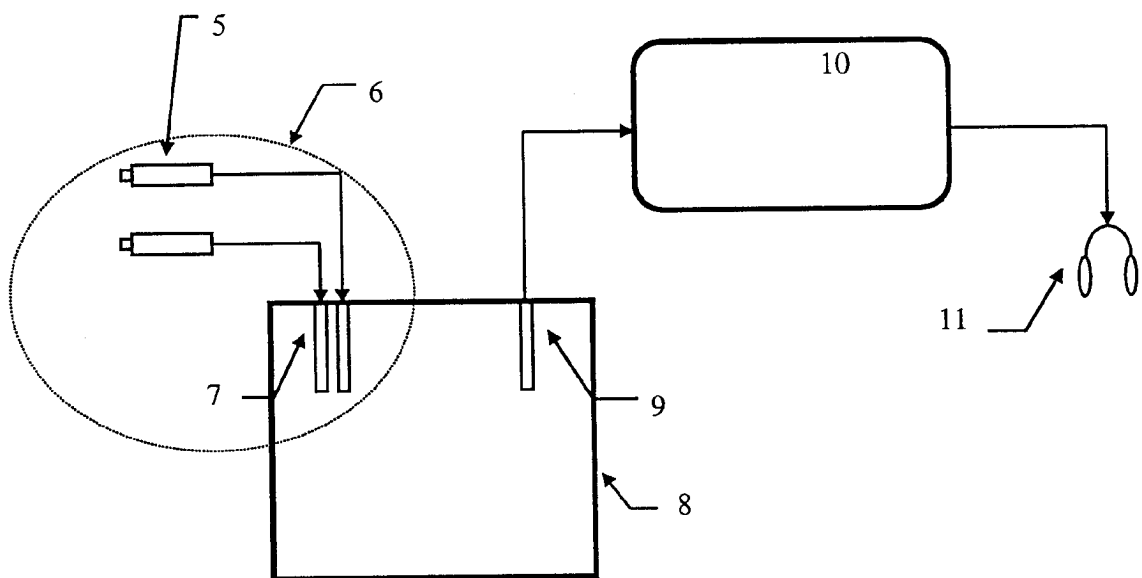
Dibujo 5.-



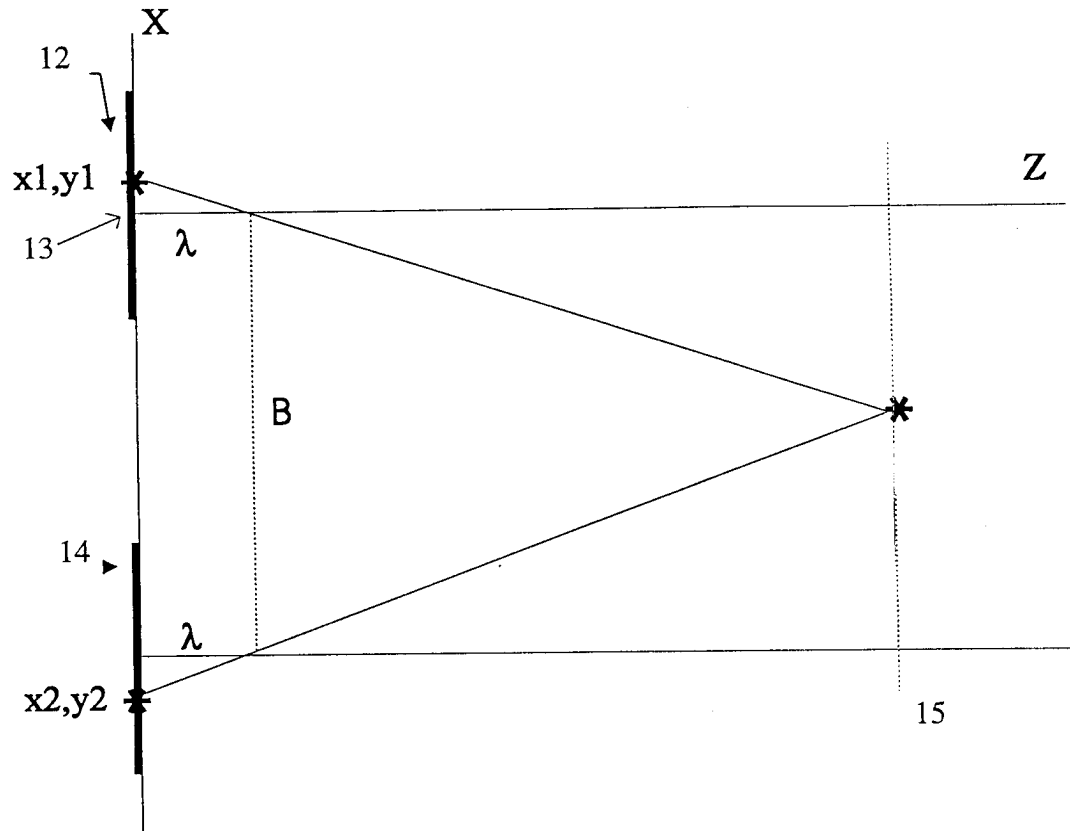
Dibujo 6.-



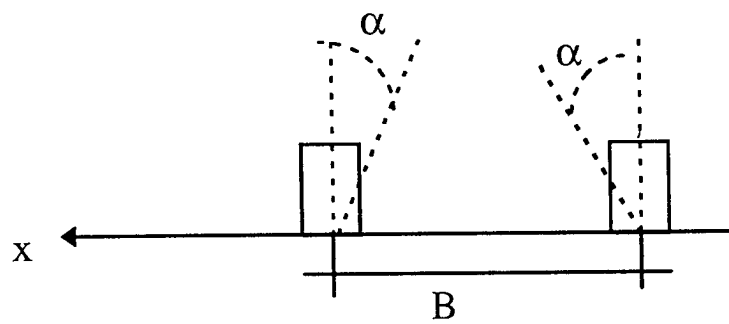
Dibujo 7.-



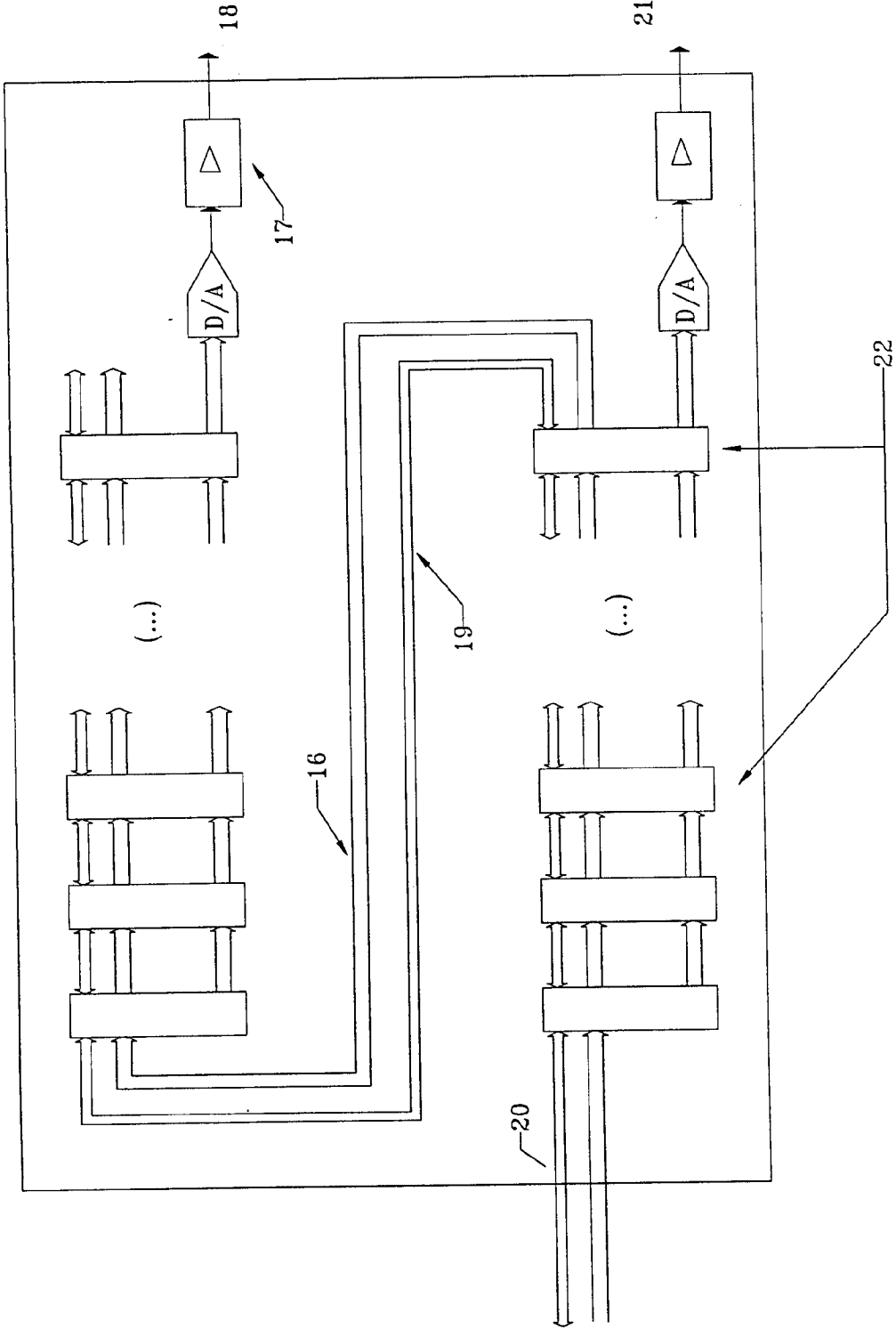
Dibujo 8.-



Dibujo 9.-

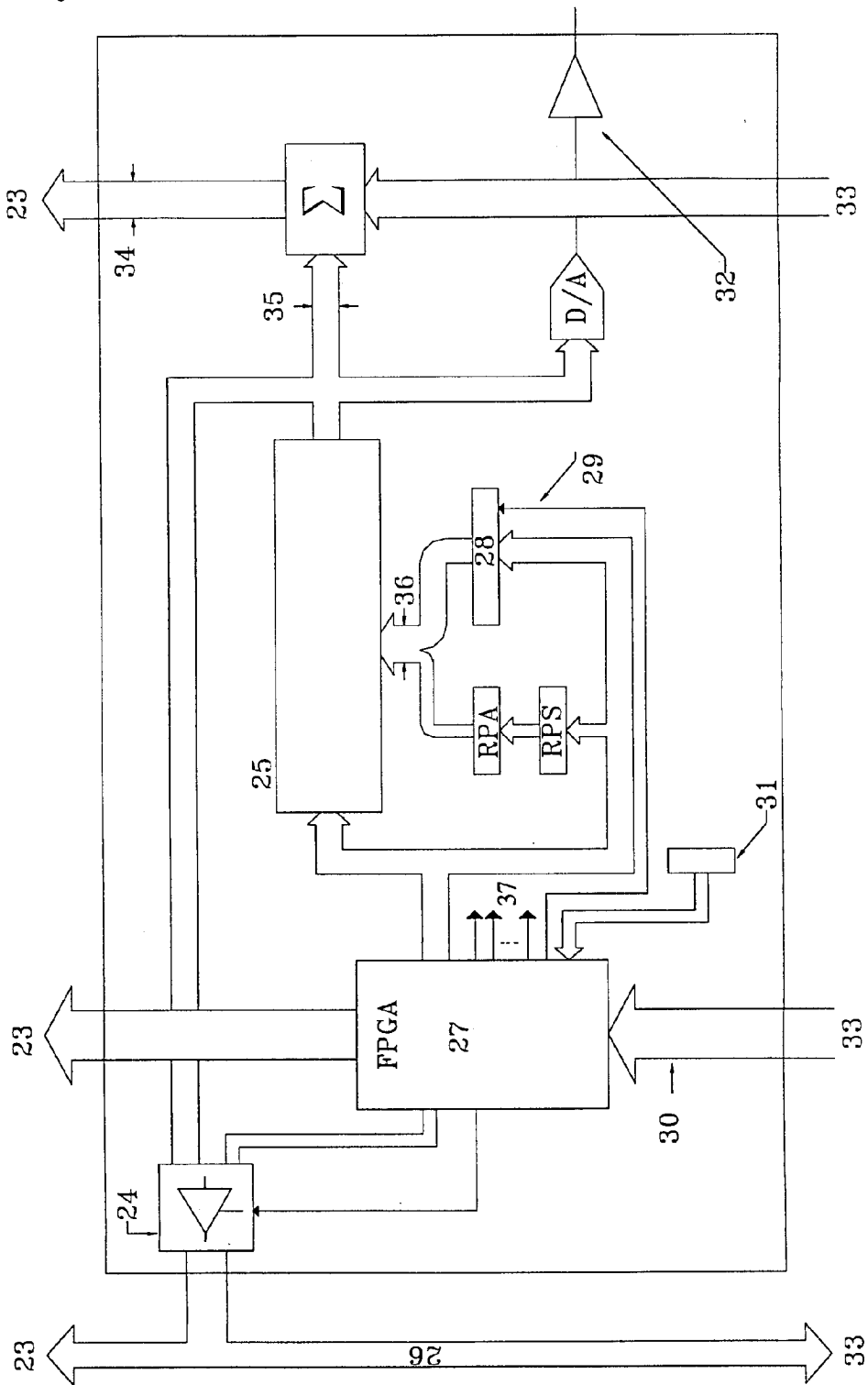


Dibujo 10.-





Dibujo 11.-





INFORME SOBRE EL ESTADO DE LA TECNICA

⑮ Int. Cl.<sup>6</sup>: A61F 9/08

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
Y	FR 2598316 A1 (CORNUEJOLS, GEORGES MARC) 13.11.1987, resumen; página 1, línea 26 - página 3, línea 22; página 4, líneas 1-18; página 5, líneas 2-16; reivindicaciones 1,3,10; figura 1.	1-4,6
Y	BE 1003022 A3 (UNIVERSITE CATHOLIQUE DE LOUVAIN) 29.10.1991, página 4, línea 32 - página 8, línea 28; reivindicaciones 1-10; figura 3.	1-4,6
A	US 5521981 A (GEHRING) 28.05.1996, resumen; columna 1, línea 1 - columna 2, línea 35; figura 2.	5
A	FR 2596940 A1 (DE TIEGE, PATRICK) 09.10.1987, resumen; página 1, línea 1 - página 4, línea 5; figura 1.	1,2,6
A	WO 8200395 A1 (THALES RESOURCES, INC.) 04.02.1982, resumen; página 6, línea 1 - página 8, línea 9.	1,6
A	EP 410045 A1 (N.V. PHILIPS' GLOEILAMPENFABRIEKEN) 30.01.1991, resumen; página 1, líneas 1-53.	1,6
A	WO 9600401 A1 (ROSCOE C. WILLIAMS Ltd) 04.01.1996	
A	BASE DE DATOS WPI en EPOQUE, semana 9322, Londres: Derwent Publications Ltd., AN 93-176309 Clases P32, S05, W04 y W06 & DE 4135444 A1 (NICKEL, W) 27.05.1993	

**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 n.º:

**Fecha de realización del informe**  
07.06.99

**Examinador**  
A. Figuera González

**Página**  
1/1