



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



⑪ Número de publicación: **2 214 115**

⑫ Número de solicitud: 200202362

⑤① Int. Cl.⁷: **G06T 7/60**

⑫

SOLICITUD DE PATENTE

A1

⑫② Fecha de presentación: **15.10.2002**

⑫③ Fecha de publicación de la solicitud: **01.09.2004**

⑫④ Fecha de publicación del folleto de la solicitud:
01.09.2004

⑦① Solicitante/s: **Universidad de Málaga**
Plaza de el Ejido, s/n
29071 Málaga, ES

⑦② Inventor/es: **Banderas Rubio, Antonio;**
Urdiales García, Cristina y
Sandoval Hernández, Francisco

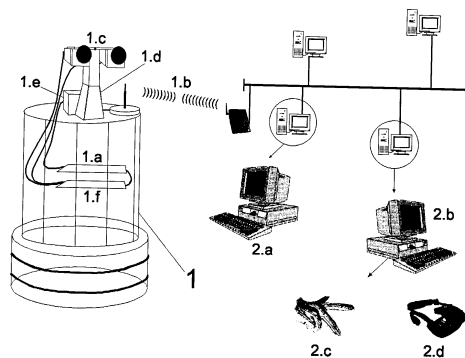
⑦④ Agente: **No consta**

⑤④ Título: **Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual.**

⑤⑦ Resumen:

Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual.

Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual. La invención consiste en un sistema de reconocimiento de objetos que consta de un agente autónomo móvil dotado de sistemas de captura de imágenes y procesamiento, y de una estación externa donde, usando un equipo de realidad virtual formado por gafas y guante, se puede controlar el sistema de percepción visual del agente móvil. El proceso de reconocimiento es, sin embargo, automático y se basa en proyectar la forma del objeto encontrado sobre una base algebraica fija.



ES 2 214 115 A1

DESCRIPCIÓN

Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual.

Sector de la técnica

La invención se encuadra en el sector técnico de la robótica, y más concretamente en el relativo al del control de agentes autónomos móviles empleando herramientas de realidad virtual.

Estado de la técnica

Actualmente, el concepto de robótica ha evolucionado hacia los sistemas móviles autónomos, que son capaces de desenvolverse por sí mismo en entornos desconocidos y parcialmente cambiantes sin necesidad de supervisión. De esta forma, el mundo de la robótica se debate entre modelos sumamente ambiciosos, como es el caso del IT, diseñado para expresar emociones, el COG, también conocido como el robot de los cuatro sentidos, el famoso SOUJOURNER o el LUNAR ROVER, vehículo de turismo con control remoto, y otros mucho más específicos y sencillos (Delgado, L. y Lozano, D., "Robots, el sueño inalcanzable de Pigmalión", *El Semanal*, pp. 24-35, 1999). Entre estos últimos se puede destacar a ROBODOC, empleado en cirugía para operaciones de cadera, CYPHER, un helicóptero robot de uso militar, el guardia de tráfico japonés ANZEN TARO, los robots mascota de Sony, el PIONEER, encargado de la limpieza en la central nuclear de Chernobyl o el DANTE II, que exploró el volcán Mt. Spurr en Alaska (Rosheim, M., "Robot evolution: the development of anthropotics", New York: John Wiley & Sons, Inc., pp. 1-36, 1994). En muchos de estos casos, especialmente en aquellos destinados a aplicaciones que se pueden calificar como críticas por su peligrosidad, resulta obligado contar con una estación de control remoto que permita supervisar el estado del agente autónomo, seguir su evolución por el entorno y evitar en la medida de lo posible, de esta forma, cualquier problema que pudiera surgir, no sólo por la acción del agente sobre el entorno, sino también por el daño que pueda sufrir el agente. En estas situaciones resulta especialmente importante contar con un sistema de control que actúe rápidamente, cuestión difícil si todo el procesamiento se centra en un único computador o si se deja prácticamente en manos del operador humano. De entre las tareas que debe llevar a cabo un agente autónomo móvil, posiblemente sean las relacionadas con la visión artificial las que precisan una mayor capacidad de procesamiento (Levine, M., "Vision in man and machine", New York: McGraw Hill, 1985), especialmente cuando deben de resolver la cuestión del reconocimiento no supervisado de objetos. Es, por ello, normal que hayan aparecido numerosos trabajos relacionados con el tema, algunos de ellos referidos a patentes de invención (ES 2 081 204, 17.12.92, Hendrik Höfer, Procedimiento y dispositivo para el reconocimiento automático de objetos). Pese a todo, el reconocimiento de objetos plantea aún numerosos problemas en las distintas etapas en que se puede dividir dicha tarea. Así, en la primera fase, de representación del propio objeto, puede darse el caso de que el grado de rotación, escala o ruido en la captura no sean siempre los mismos, por lo que algunos autores recurren a usar etapas de preprocesado (Shrikantan, G., Lam, S., y Srihari, S., Gradient-based contour encoding for character recognition, *Pattern Recognition*

7 (29), pp. 1147-1160, 1996; Hsu, J. y Hwang, S. A machine learning approach for acquiring descriptive classification rules of shape contour, *Pattern Recognition* 2(30), pp. 245-252, 1997), pero esto incrementa excesivamente el tiempo de proceso, siendo inviable para situaciones de actuación en las que el tiempo es el factor más importante. La implementación de un sistema de extracción de características que sea tan independiente como sea posible de las distorsiones que pueda sufrir el objeto es posible, pero difícil. Por ello, algunos autores usan enormes bases de datos, en las que guardan información sobre el objeto obtenida del estudio de numerosas versiones distorsionadas. Esto conduce a sistemas en las que el volumen de información a manejar es excesivo y, por tanto, a nuevos problemas de tiempo de procesamiento (Cyganski, D., Orr, J., Cott, T., y Dobson, R. Development, implementation, testing and application of an affine transform invariant curvature function, *Proc. of the First Int. Conf. on Computer Vision*, pp. 496-500, 1987).

Una vez caracterizado el objeto, y dado que la cantidad de información relacionada con el mismo es generalmente muy elevada, surge el problema de determinar qué características emplear finalmente para representar el objeto. Las desventajas que presentan las soluciones adoptadas para ello son, principalmente, el excesivo tamaño del vector de características (Campbell, N., MacKeown, W., Thomas, B. y Troscianko, T. Interpreting image databases by region classification, *Pattern Recognition* 4(30), pp. 555-563, 1997; Romero, R., Touretzky, D. Y Thibadeau, R. Optical chinese character recognition using probabilistic neural network, *Pattern Recognition* 8(30), pp. 1279-1292, 1997), y la necesidad de usar sistemas expertos para extraer el conjunto de características, lo cual implica llevar a cabo un profundo estudio sobre el conjunto de patrones que entran al sistema y, por ello, restringir el sistema de reconocimiento a una aplicación muy específica (Cao, J., Ahmadi, M. y Shridhar, M. A hierarchical neural network architecture for handwritten numeral recognition, *Pattern Recognition* 2(30), pp. 289-294, 1997). En este trabajo se presenta una posible solución a estos problemas, desarrollando una etapa de reconocimiento que actúa de forma rápida y que no presenta limitaciones en el espacio de objetos de entrada. No se debe olvidar, sin embargo, que dicho sistema es sólo una de las partes del sistema que se describe, que consta de un agente autónomo móvil dotado de sistemas de captura de imágenes y procesamiento, pero también de una estación externa donde, usando un equipo de realidad virtual formado por gafas y guante, se puede controlar este sistema de percepción visual del agente móvil.

Explicación de la invención

La invención consiste en un sistema de reconocimiento de objetos que consta de un agente autónomo móvil dotado de sistemas de captura de imágenes y procesamiento, y de una estación externa donde, usando un equipo de realidad virtual formado por gafas y guante, se puede controlar el sistema de percepción visual del agente móvil. El proceso de reconocimiento es, sin embargo, automático y se basa en proyectar la forma del objeto encontrado sobre una base algebraica fija. Esto es, si bien la estación externa permite identificar nuevos objetos y dirigir la atención del agente hacia una determinada región del espacio, el agente realiza la extracción y caracterización de los posibles objetos presentes en su campo visual de ma-

nera no supervisada.

La principal finalidad del sistema es conseguir el control remoto de un agente autónomo con la finalidad de que éste reconozca de forma automática los objetos presentes en su entorno. Como se ha comentado, aunque dicho reconocimiento es llevado a cabo de forma automática, puede ser supervisado externamente. De igual forma, la interfaz externa permite que un operario controle, ayudado por un equipo de realidad virtual, la operación del agente. Inicialmente, este control se realiza con el único objetivo de que el agente autónomo explore su entorno, por lo que sólo se interactúa sobre el movimiento del cuerpo del agente y sobre las cámaras que capturan la imagen. Para llevar a cabo este movimiento de cámaras, el agente incorpora una cabeza con tres grados de libertad (*pan*, *tilt* y *vergencia*). La aplicación de un sistema de este tipo puede resultar especialmente interesante para la exploración o identificación de determinados objetos en entornos de difícil acceso o en situaciones en las que resulte conveniente ejercer un control remoto (reconocimiento de matrículas, trabajo en centrales nucleares, etc.).

Descripción de los dibujos

Figura 1

Arquitectura del sistema

1. Módulo de captura de imagen
2. Módulo de procesamiento
 - 2.1. Submódulo de procesamiento ubicado en el agente
 - 2.2. Submódulo de procesamiento externo
3. Módulo de usuario
4. Módulo de control de movimiento

Figura 2

Esquema de la fase de representación del proceso de reconocimiento de objetos

- C. Vector curvatura del contorno del objeto

A_0, A_1, A_N . Vector resultante de la comparación de C con distintos vectores previamente almacenados en memoria.

Figura 3

Representación del sistema

1. Agente autónomo móvil
 - 1.a. Computador de procesamiento interno
 - 1.b. Enlace inalámbrico para comunicación
 - 1.c. Cámaras
 - 1.d. Cabeza para soporte de cámaras con tres grados de libertad
 - 1.e. Control de la cabeza 1.d
 - 1.f. Capturadora de imagen
2. Red externa de procesamiento y control
 - 2.a. Computador de procesamiento externo
 - 2.b. Computador de control del agente
 - 2.c. Guante de realidad virtual para control de movimiento
 - 2.d. Gafas de realidad virtual y sensor de posición

Descripción detallada

En la Figura 1 se presenta la arquitectura del sistema. En concreto, el sistema se caracteriza porque consta de:

- a) un módulo de captura de imagen (1), residente en el agente autónomo, y que se encarga de capturar las imágenes del entorno. Estas imágenes son enviadas al
- b) módulo de procesamiento (2), que a su vez se divide en dos submódulos, uno de ellos ubicado en el agente autónomo -submódulo 1 (2.1)-, y encargado principalmente de gestionar el paso de comandos e imágenes entre el agente y el exterior, y otro externo -submódulo 2 (2.2)-, donde se lleva a cabo la mayor parte del proceso de tratamiento de imagen. Por lo tanto, una vez que las imágenes llegan al submódulo 1, éste las comprime y transmite vía radio al submódulo 2. Este segundo submódulo se encarga de detectar los posibles objetos, enviando esta información al
- c) módulo de usuario (3), donde se lleva a cabo la clasificación de las entidades detectadas. Este módulo permite al usuario decidir hacia donde seguir la exploración y, para hacerle esta tarea más flexible y ayudar a la integración del mismo en el entorno percibido por el agente autónomo, se emplea un equipo de realidad virtual. En caso de que el usuario decida que debe cambiar el entorno visual percibido, mandará sus órdenes al módulo de procesamiento que, a su vez, las guiará al
- d) módulo de control de movimiento (4), que permite mover el robot por el entorno y girar las cámaras que lleva para que se dirijan hacia el punto seleccionado del mismo. Una vez completada la maniobra, se informa al módulo de procesamiento, que vuelve a comenzar el proceso.

El sistema también se caracteriza porque el proceso de reconocimiento de objetos, ubicado en los módulos de procesamiento y de usuario, se subdivide en:

- a) una fase de detección de objetos, en la cual se detectan, sobre la imagen matricial de entrada, aquellas entidades de puntos que presentan un nivel de gris o color homogéneo, suficientemente distinto del que presentan los puntos que los rodean. Si estas entidades tienen unas características de área y compacidad suficientes se considerarán como objetos potenciales, a expensas de que puedan ser descartados como objetos reales en las siguientes fases del proceso;
- b) una fase de caracterización, en la que se estudia la forma de los objetos detectados en la fase anterior. Aunque en la detección se hayan tenido en cuenta otros factores (área, nivel de gris o color, etc.), la característica que identificará finalmente al objeto será su forma. En particular, la forma del contorno del objeto se cuantifica calculando su curvatura, que se define, en cada punto del contorno, como el cambio de orientación relativa del contorno;

- c) una fase de representación (Figura 2), en la cual el vector curvatura del contorno del objeto (C) se compara con distintos vectores previamente almacenados en memoria. El resultado de estas comparaciones será un vector de características (A), que se asocia al objeto de entrada, y que permitirá posteriormente su clasificación;
- d) una fase de clasificación, en la que utilizando el vector de características asociado al objeto y un algoritmo de agrupación no supervisado se puede clasificar el objeto de entrada en alguna de las clases ya existentes, crear una clase nueva o descartarlo, si el usuario estima que su forma no coincide con ningún objeto real. Hay que destacar que el proceso de clasificación es automático, quedando sólo como tareas del usuario el renombrar cada clase recién creada con el nombre real asociado al objeto, o el ya comentado descarte de objetos no reales.

Las tres primeras fases se encontrarán físicamente ubicadas en el submódulo 2 de procesamiento y la última fase en el módulo de usuario.

En la fase de caracterización, a su vez, hay que señalar el empleo de un algoritmo que calcula la curvatura en cada punto del contorno de manera adaptativa, con lo cual la estimación presenta una relación señal a ruido mayor y permite caracterizar el contorno, sin depender de las escalas naturales que éste presente.

La fase de representación se caracteriza por realizar una proyección del vector curvatura del objeto sobre una base algebraica almacenada en memoria. Para ello, se llevan a cabo las siguientes acciones:

- a) Cálculo de la transformada de Fourier del vector curvatura del contorno del objeto;
- b) Multiplicación de dicha transformada con otras transformadas de Fourier de vectores de curvatura previamente almacenadas en memoria. Estas transformadas almacenadas constituyen una base, sobre la cual se proyecta la transformada de entrada. Los resultados de estas multiplicaciones constituyen el vector de características que identifica al objeto de entrada.

Modo de realización de la invención

A continuación se describe un ejemplo de realización práctica, no limitativo, del presente invento. No se descartan, sin embargo, otros modos de realización en los que se introduzcan cambios accesorios que no desvirtúen su fundamento.

El sistema descrito se muestra en la Figura 3. En este caso, el módulo de procesamiento se reparte entre el Pentium PC 133 MHz (1.a) residente en un agente móvil *Nomad200* (1) de *Nomadic Technologies* y un Pentium PC 200 MHz externo (2.a) conectado a la red local. Al estar el módulo de procesamiento repartido en dos computadoras independientes, se podrá trabajar en paralelo, aumentando con ello la velocidad del sistema. Sin embargo, ya que se debe mantener la autonomía del agente móvil, la comunicación interna a este módulo se resuelve mediante un enlace radio (1.b) que conecta al agente con la red local. En este caso, el enlace es del tipo *RangeLan2* 7510. Las rutinas del módulo de procesamiento que se ejecutan en el interior del *Nomad200* se encargan de controlar la captura de imagen y el movimiento de la cabeza del agente. De cualquier forma, en estas dos misiones el

consumo computacional es mínimo, pues en el control de la cabeza se realiza un simple intercambio de comandos con el exterior, y el control de imagen se reduce a su compresión y transmisión por el enlace radio. De esta manera, la práctica totalidad de las funciones del sistema de reconocimiento se ejecutan en el Pentium PC externo conectado a la red local, que recibirá del agente las imágenes, llevará a cabo las fases de detección de objetos, caracterización y representación, y, con ayuda del módulo de usuario, la de clasificación. Como datos de salida, este PC externo solicitará al residente en el agente que dirija las cámaras hacia una determinada región del campo visual.

Para la captura de imágenes se emplea una cabeza que consta de dos cámaras *KP-D50* (1.c), lo que permite implementar un sistema de visión binocular. Estas dos cámaras obtienen imágenes en color o niveles de gris y formato NTSC/PAL del entorno usando un sistema entrelazado de 2:1. La transmisión de imágenes estéreo en tiempo real por un enlace radio es un problema complejo que requiere, en muchas ocasiones, un enorme ancho de banda.

En el sistema implementado las imágenes recogidas por ambas cámaras presentan una resolución variable en el espacio, de manera que la resolución de la escena se mantiene al máximo sólo en una zona denominada fovea, a partir de la cual la resolución baja progresivamente formando los denominados anillos foveales. Si se mantiene la zona de máxima resolución en el área de atención, no se percibe disminución en la cantidad de información óptima, reduciéndose los requisitos de memoria y la carga computacional necesaria para procesar dichas imágenes.

Las cámaras se sustentan en una cabeza *Zebra Pan/Tilt/Vergence Head* (1.d) que presenta tres grados de libertad (giros horizontal, vertical y vergencia), controlada mediante un circuito de control (*Zebra Controller* (1.e)) que se conecta a través de un puerto serie con el módulo de procesamiento residente en el propio agente autónomo. Las imágenes captadas por ambas cámaras se reciben usando una tarjeta capturadora/digitalizadora de imagen modelo *TMS320C44* (1.f), igualmente conectada con el módulo de procesamiento residente en el agente móvil. Este dispositivo se encargará únicamente de capturar las imágenes a una velocidad lo más constante posible, suministrando éstas al módulo de procesamiento cuando éste lo desee. La conexión entre la tarjeta capturadora/digitalizadora y el módulo de procesamiento presenta el problema de establecerse a través de un bus que siempre presentará un ancho de banda pequeño para la transmisión de imágenes estéreo.

Dadas las limitaciones propias al sistema, se trabaja con imágenes de 256x256 píxeles con 256 niveles de gris, que ocuparían, a resolución uniforme, 65 Kbytes cada una. Para alcanzar una calidad de vídeo de al menos 8 imágenes por segundo, el ancho de banda del enlace debería ser de 520 Mbytes/s, disponiéndose, con el equipo descrito, de tan sólo un 13'56% de dicho valor. Si las imágenes son de resolución variable (fovea centrada de 64x64 píxeles y 2 anillos de resolución), cada imagen se reduce a 10240 bytes. De esta forma, la tarjeta capturadora/digitalizadora transmite hasta 100 imágenes por segundo al módulo de procesamiento interno al agente autónomo. Si se reduce la captura a 9 imágenes por segundo en estéreo, el flujo de información supondría tan sólo 90 Kbytes/s. Aunque podría parecer que el trabajar con una

imagen de resolución variable puede resultar incómodo para el observador interesado en áreas periféricas de la imagen, el desarrollo de fóveas móviles ha permitido solventar dicho problema. Asimismo, se dispone de algoritmos de procesamiento de imágenes multirresolución basados en estructuras jerárquicas que trabajan a velocidades mucho mayores que sus equivalentes en imagen de resolución uniforme con niveles de calidad muy similares.

Finalmente, el usuario podrá supervisar o simplemente seguir el funcionamiento del agente autónomo a través de un módulo de usuario residente en un PC externo al sistema (2.b), y conectado por red al módulo de procesamiento. Adicionalmente, este módulo de usuario cuenta con un equipo de apoyo consistente en un guante y un casco de realidad virtual.

El guante empleado (2.c) es el 5th *Glove* desarrollado por *5DT (Fifth Dimension Technologies)*, que dispone de cinco sensores de flexión para los dedos de la mano y de un detector de posición del guante con dos grados de libertad (inclinación y rotación). La misión del guante es ubicar la fóvea móvil sobre el objeto de interés. Por otro lado, el casco de realidad virtual es el *I-Glasses* de *Virtual I-O* (2.d), que se compone de unas gafas, encargadas de la visión, y de un detector de posición, que transmite al PC los distintos movimientos de la cabeza. Con el casco se controla la posición de las cámaras.

Cada una de las imágenes que finalmente llegan al extremo usuario será analizada, para así detectar, caracterizar e identificar cada uno de los objetos presentes en el campo visual. Para acelerar el proceso de detección de entidades que se distinguen del fondo, la imagen de resolución variable recibida se segmenta empleando una estructura tridimensional denominada polígono foveal. Los sucesivos niveles de esta estructura presentan porciones mayores de la imagen a una resolución progresivamente menor. El algoritmo de detección de objetos trabaja de forma descendente y se basa tanto en criterios descriptivos como proposicionales. El sistema de detección implementado da buenos resultados tanto en imágenes artificiales como con imágenes reales siempre y cuando el fondo sea relativamente homogéneo y es capaz de trabajar con figuras planas y volúmenes, concentrando los recursos del sistema sobre el objeto de estudio. Los tiempos de detección son, además, inferiores al segundo, lo cual permite actualizar el conjunto de posibles objetos a analizar con una frecuencia relativamente elevada.

Una vez detectados los posibles objetos, éstos se caracterizan mediante el cálculo adaptativo de la función de curvatura asociada a su contorno. El hecho de que la curvatura se calcule de forma adaptativa permite eliminar gran parte del ruido de cuantificación sin pérdida de puntos significativos. De forma básica,

el método de cálculo de la curvatura de cada punto del contorno compara iterativamente la distancia euclídea entre el punto en estudio y el ubicado k puntos delante, $d(i, i+k)$, con la distancia en píxeles entre ambos puntos. Si ambas distancias son muy similares se aumenta k , si son muy distintas se decrementa k . Una vez determinado el k asociado a cada punto, se usa una tabla de pendiente- k para obtener la curvatura de cada punto. La simplicidad del método permite que su tiempo de procesamiento sea de tan sólo 4-5 milisegundos, consumiendo escasos recursos del módulo de procesamiento.

Para comprimir la información contenida en una función de curvatura se proyecta la transformada de Fourier (DFT) de dicha función sobre una base previamente calculada del conjunto de transformadas de Fourier de funciones de curvatura (CFDFTs). A efectos de construir la base del subespacio de las CFDFTs se empleó un conjunto de tan sólo 27 objetos. Tras extraer las DFT de sus funciones de curvatura, éstas se agruparon mediante un algoritmo c -medias donde el número de clases c varió desde 2 hasta 15. Cada uno de los prototipos obtenidos por este algoritmo será uno de los elementos de la base buscada. Mientras c fue inferior a 10, los ángulos medios entre vectores presentaron variaciones de aproximadamente 0.0245° . En la transición de 10 a 11 el ángulo medio entre vectores disminuyó en 3.63° . En vista de este decremento se optó por una base de 10 elementos. Si bien el ángulo medio entre vectores está más cerca de 60° que de 90° , los resultados posteriores prueban que el método funciona correctamente.

Una vez se dispone de la base de 10 elementos descrita, para obtener el vector de características de un objeto cualquiera sólo son necesarios tres pasos: a) cálculo de su función de curvatura, b) cálculo de su CFDFT, y c) proyección de dicha CFDFT sobre los vectores de la base. Este proceso lleva un tiempo aproximado de 0.07 s por objeto. Sobre la base de 10 elementos propuesta, cada vector de características ocupa únicamente 40 bytes, dado que son necesarios 4 bytes para representar cada número real en punto flotante.

Finalmente, los resultados de clasificación en pruebas no supervisadas proporcionan errores inferiores al 10%, aún trabajando con espacios muestrales cuyos objetos no formaron parte de los seleccionados para construir la base. Estos datos permiten que el agente autónomo pueda resolver sin supervisión la mayoría de las situaciones que se le presentan, y sólo requiere la intervención del usuario para resolver determinadas situaciones o nombrar los distintos tipos de objetos. Todo ello posibilita que la exploración del entorno sea rápida y prácticamente autónoma.

REIVINDICACIONES

1. Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual, **caracterizado** porque consta de:

- a) un módulo de captura de imagen, que se encarga de capturar las imágenes del entorno y enviarlas a
- b) un módulo de procesamiento, dividido en dos submódulos internamente conectados por un enlace radio, lo que permite la autonomía del agente, así como el disponer de una gran capacidad computacional, de modo que una vez detectados los posibles objetos presentes en el entorno, esta información es transmitida a
- c) un módulo de usuario, en el cual se clasifican las entidades detectadas, que incorpora un equipo de realidad virtual que consta de gafas y guante para permitir al usuario que el control de la exploración sea lo más flexible e interactivo posible, y a través del cual el usuario envía sus órdenes al
- d) módulo de control de movimiento, que mueve al agente autónomo por el entorno y permite que sus cámaras se dirijan al punto seleccionado del mismo.

2. Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual, según reivindicación anterior, **caracterizado** porque el proceso de reconocimiento de objetos se subdivide en:

- a) una fase de detección de objetos, en la que se detectan sobre la imagen matricial de entrada las entidades de nivel de gris o color homogéneo;
- b) una fase de caracterización, en la que se ex-

trae de forma adaptativa la función de curvatura asociada al contorno de los objetos detectados;

- c) una fase de representación, en la cual el objeto se representa mediante un vector de características cuyas componentes resultan del producto de la Transformada de Fourier de la función de curvatura extraída en b) por un conjunto de vectores previamente calculados y que representan una pseudo-base del conjunto de Transformadas de Fourier de las funciones de curvatura del conjunto de objetos;
- d) una fase de clasificación, en la que el vector de características permite identificar el objeto o, si el agente desconoce la naturaleza de éste, proceder a su identificación por el usuario.

3. Sistema de reconocimiento automático de objetos por proyección en una base algebraica y controlado remotamente por realidad virtual, según reivindicaciones anteriores, **caracterizado** porque el proceso de representación de objetos se subdivide en:

- a) Cálculo previo al proceso de reconocimiento de una base del conjunto de Transformadas de Fourier de las funciones de curvatura de los objetos, que se lleva a cabo sólo una vez y de manera externa al proceso de reconocimiento, no siendo necesario que en el proceso de construcción de la base entren todos los posibles objetos que posteriormente se reconocerán, aunque si es necesario que los empleados sean bastante significativos.
- b) Obtención de la Transformada de Fourier de la función de curvatura del objeto.
- c) Proyección de la Transformada de Fourier obtenida en b) sobre la base calculada en a), lo que origina un vector del mismo tamaño que la base, que será el vector de características empleado en el proceso de clasificación.

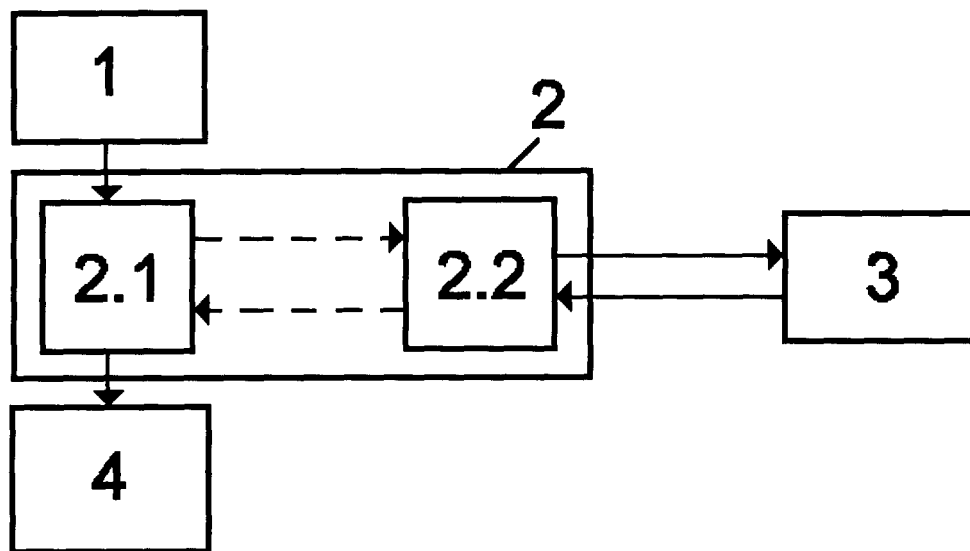


Figura 1.

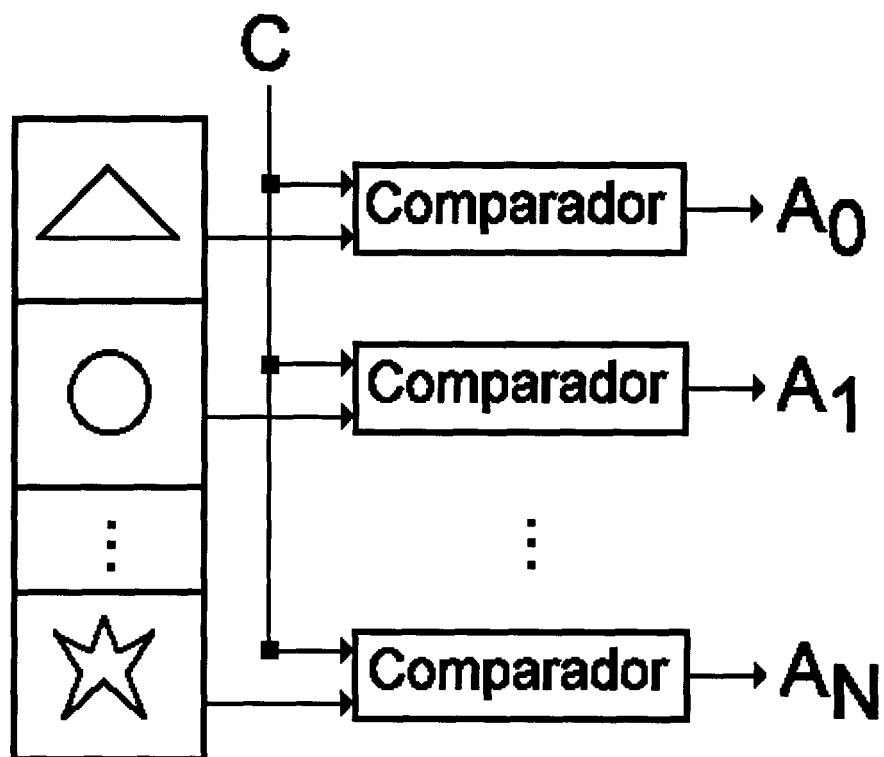


Figura 2.

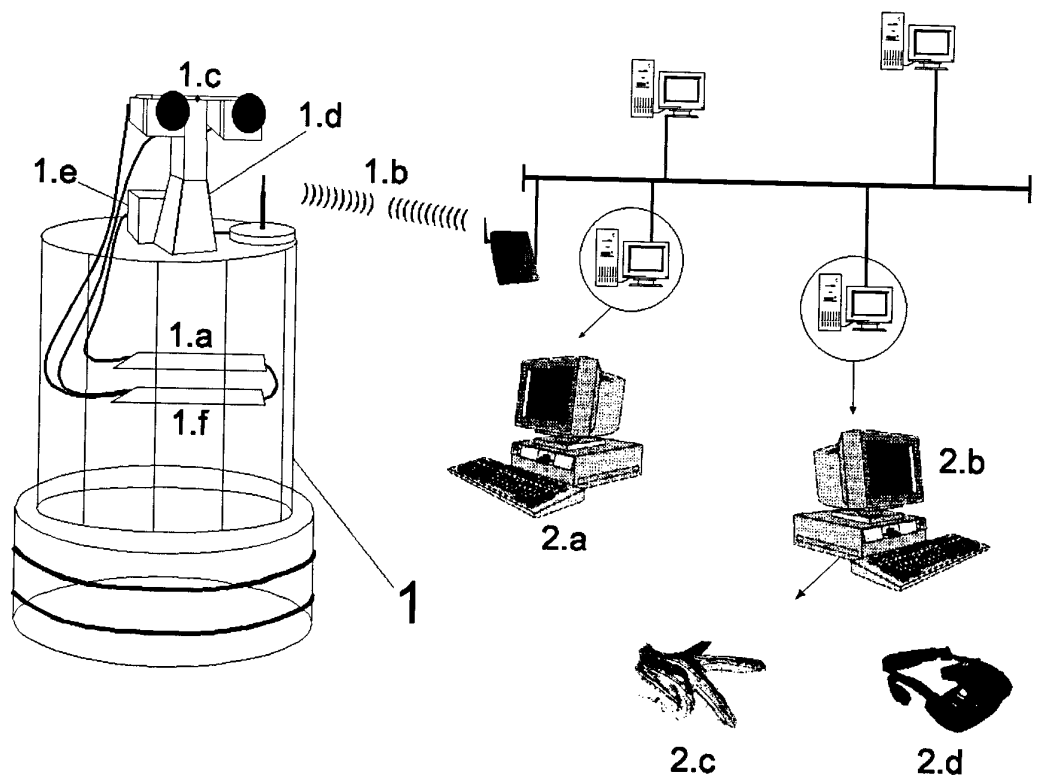


Figura 3.



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

⑪ ES 2 214 115

⑫ Nº de solicitud: 200202362

⑬ Fecha de presentación de la solicitud: 15.10.2002

⑭ Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑮ Int. Cl.7: G06T 7/60

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X	EP 669758 A1 (OH-YOH KEISOKU KENKYUSHO K.K.) 30.08.1995, todo el documento.	1-3
A	US 6016385 A (YEE et al.) 18.01.2000, todo el documento.	1-3
X	JP 08-112449 A (CASIO COMPUT CO LTD) 07.05.1996, todo el documento.	1-3
A	JP 10-040381 A (KUBOTA KK) 13.02.1998, todo el documento.	1-3
X	WO 9416406 A1 (RITCHEY) 21.07.1994, página 59.	1-3
A	WO 0118617 A1 (RUTGERS & N.J. UNIVERSITY) 15.03.2001, todo el documento.	1-3

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:

Fecha de realización del informe

27.07.2004

Examinador

M. Fluvia Rodríguez

Página

1/1