

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 409 533**

21 Número de solicitud: 201132071

51 Int. Cl.:

G09B 21/00 (2006.01)

G06T 7/00 (2006.01)

G06K 9/00 (2006.01)

G06K 9/46 (2006.01)

G06K 9/62 (2006.01)

12

PATENTE DE INVENCION CON EXAMEN PREVIO

B2

22 Fecha de presentación:

21.12.2011

43 Fecha de publicación de la solicitud:

26.06.2013

Fecha de la concesión:

03.10.2013

45 Fecha de publicación de la concesión:

15.10.2013

73 Titular/es:

**UNIVERSIDAD DE ALCALÁ (100.0%)
Plaza de San Diego, s/n
28801 Alcalá de Henares (Madrid) ES**

72 Inventor/es:

**LÓPEZ SASTRE, Roberto Javier;
GIL JIMÉNEZ, Pedro;
ACEVEDO RODRÍGUEZ, Francisco Javier;
LAFUENTE ARROYO, Sergio;
GÓMEZ MORENO, Hilario y
MALDONADO BASCÓN, Saturnino**

74 Agente/Representante:

GUTIÉRREZ DE MESA, José Antonio

54 Título: **Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje**

57 Resumen:

El sistema tiene como objetivo la aplicación de algoritmos de reconocimiento de clases de objetos en imágenes para mejorar la capacidad comunicativa en personas con alteraciones del lenguaje. Para ello se emplea un sistema de captura de imágenes (1), encargado de obtener la imagen que será procesada por un módulo que implementa un sistema de localización de clases de objetos en imágenes (2), el cual está basado en técnicas de visión artificial. Este módulo (2) ofrecerá una salida visual (3) en la que se recuadrarán las clases de objetos reconocidas en la imagen. El sistema interactuará con el usuario (4), ofreciéndole la información del objeto localizado, o interrogándole sobre el mismo para que sea él quien introduzca la información relativa a la clase reconocida. En este segundo modo, el sistema informará al usuario si el ejercicio se completó correctamente.

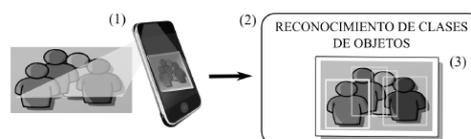


Figura 1

ES 2 409 533 B2

DESCRIPCIÓN

**SISTEMA DE RECONOCIMIENTO DE CLASES DE OBJETOS POR VISIÓN
ARTIFICIAL PARA LA MEJORA DE LA CAPACIDAD COMUNICATIVA EN
PERSONAS CON ALTERACIONES DEL LENGUAJE**

5 **SECTOR DE LA TÉCNICA**

La presente invención se enmarca dentro de las tecnologías de la información y las comunicaciones, y dentro del campo del reconocimiento de clases de objetos mediante técnicas de visión artificial. También, dentro del campo de metodologías y terapias para el tratamiento de alteraciones del lenguaje.

10

ESTADO DE LA TÉCNICA

La aplicación de las tecnologías de la información y las comunicaciones (TIC) para mejorar la capacidad comunicativa en personas con alteraciones del lenguaje ha despertado, recientemente, un gran interés en la comunidad científica. Existen numerosas tecnologías de ayuda para el apoyo de las personas con dificultades de expresión y comprensión del lenguaje [1]. La mayoría de estas ayudas técnicas, se basan en el reconocimiento de dibujos y/o fotografías, bien sobre tableros o fichas, bien sobre sistemas electrónicos, como la pantalla del ordenador. Así pues, esta tecnología permite a los usuarios realizar un ejercicio de asociación entre imágenes y texto, con la única limitación de que las imágenes deben haber sido seleccionadas previamente para ser utilizadas en el sistema. El acceso a la palabra se suele realizar mediante un interfaz que puede ser táctil para seleccionar el pictograma correspondiente, controlado por la voz, o mediante editores predictivos, que se integran en el elemento de apoyo más recomendado, como los comunicadores personales, que permiten una comunicación funcional básica. Estas tecnologías son utilizadas con propósito terapéutico, rehabilitador, o como elementos de apoyo para la compensación de alteraciones. Ejemplos de estos sistemas son: PLAPHOONS [2], PREDWIN [3] o C-SPEAK APHASIA [4], entre otros.

En esta línea, se encuentran también algunas patentes relacionadas [5-6]. Mientras que en [6] se describe un comunicador controlado por medio de la voz, en [5] se detalla un sistema terapéutico específico para el tratamiento de la afasia, que presenta por pantalla una imagen al

30

usuario, para que éste, por medio de su voz, decida de qué imagen se trata.

5 Cabría destacar que ninguno de los sistemas descritos en [1-6] utiliza ningún sistema para la
captura de imágenes, ni ninguna técnica de reconocimiento de clases de objetos para realizar
el reconocimiento en tiempo real de los objetos que aparecen en imágenes capturadas por el
mismo usuario. Ésta es una diferencia fundamental con respecto a la invención objeto de esta
patente: por primera vez se aplican técnicas de visión artificial, para que el sistema pueda
reconocer automáticamente clases de objetos en las imágenes que el usuario captura mientras
interactúa con el sistema, es decir, que éstas no tienen que estar incluidas en el sistema con
10 antelación.

Así pues, la utilización de técnicas de visión artificial para el reconocimiento de clases de
objetos en imágenes, para la mejora de la capacidad comunicativa en personas con
alteraciones del lenguaje, no presenta ningún precedente en la revisión del estado de la técnica.

15

EXPLICACIÓN

Breve descripción

El sistema objeto de esta invención aplica algoritmos de visión artificial para el
reconocimiento de clases de objetos en imágenes, con el objetivo de mejorar la capacidad
20 comunicativa en personas con alteraciones del lenguaje. El sistema propuesto debe disponer
de una cámara (1) que le permita al usuario del mismo capturar imágenes. El sistema realizará
un reconocimiento automático de las clases de objetos que estén presentes en la imagen (2),
localizando las clases detectadas (p. ej. la clase persona) en la misma imagen (3). Por último,
el sistema dispondrá de una interfaz de usuario (4) que le permitirá al mismo interactuar con la
25 aplicación, de modo que pueda mejorar su capacidad comunicativa. Este interfaz, ofrecerá la
información relativa al objeto detectado, tanto de forma visual como auditiva, y además podrá
utilizarse para interrogar al usuario sobre el objeto reconocido, de modo que sea él mismo el
que introduzca el nombre de la clase detectada (4). El sistema informará al usuario de si el
ejercicio se completó correctamente.

30

El sistema puede implementarse tanto en dispositivos móviles con cámara, como en sistemas no móviles que también dispongan de un sistema de captura de imágenes.

5 La principal diferencia de este sistema con respecto al estado de la técnica, es la utilización de imágenes reales, capturadas por el sistema y procesadas en tiempo real mediante algoritmos de visión artificial para la localización de clases de objetos, así como la aplicación de estas técnicas a la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje.

Descripción detallada

10 El sistema que se describe permite la aplicación de técnicas de visión artificial de reconocimiento y detección de clases de objetos en imágenes, para mejorar la capacidad comunicativa en personas con alteraciones del lenguaje. En la mayoría de las TIC aplicadas a este problema [1-6], los usuarios no pueden trabajar con imágenes capturadas en el momento por ellos mismo, sino con bancos de imágenes que son precargados en las aplicaciones, y que
15 en muchas ocasiones son principalmente pictogramas. Los usuarios realizan un ejercicio de asociación entre la imagen y la palabra que les evoca la misma. Este ejercicio es de gran utilidad a la hora de trabajar con personas que presenten alteraciones del lenguaje, como podrían ser la afasia, la anartria o la disartria.

20 El sistema objeto de la invención puede ser implementado tanto en un dispositivo móvil, como no móvil. El sistema debe disponer de un sensor para la captura de imágenes, una cámara (1). El usuario utilizará la cámara para capturar las imágenes que serán procesadas por el módulo de reconocimiento de clases de objetos (2). Este módulo recibe como entrada una imagen capturada por el usuario, y será el encargado de localizar las clases de objetos para las que ha
25 sido entrenado. La localización se realizará recuadrando en la imagen dónde aparecen las clases detectadas (3). Es importante destacar que el sistema trabaja en el nivel de las clases, las categorías, es decir, que no sólo reconocerá un objeto concreto, sino la clase a la que el objeto pertenece. Algunos ejemplos de clases para las que puede ser entrenado el sistema son: persona, animal, coche, moto, perro, etc. Una vez el módulo de localización (2) identifica las
30 clases en la imagen, el interfaz de usuario del sistema (4) será el encargado de transmitir ésta

información. Así, el interfaz (4) podrá directamente identificar las clases detectadas e informar e interactuar con el usuario de forma visual y/o auditiva, es decir, visualizando el nombre en la pantalla del sistema de la clase reconocida, o emitiendo de forma sonora dicha información. También contemplamos que el interfaz (4) pueda interrogar al usuario acerca de la clase reconocida, de modo que sea él mismo quien tenga que proporcionar esta información, y que sea el sistema quien valide la información e informe al usuario de si la información fue introducida de forma correcta.

Así pues, el sistema descrito en esta patente resulta novedoso, y las principales características del mismo que lo diferencian de forma fundamental con respecto al resto de sistemas que abordan el mismo problema son: a) el dispositivo para la captura de imágenes, que permite al usuario trabajar con imágenes reales, de modo el conjunto de imágenes que maneja el sistema es dinámico y no tiene que estar precargado en el mismo; b) la utilización de técnicas de visión artificial para el reconocimiento y localización de clases de objetos en imágenes, que permiten al sistema localizar los objetos con los que el usuario podrá trabajar.

DESCRIPCIÓN DE LOS DIBUJOS

El sistema objeto de esta invención se ilustra en las Figuras 1 y 2.

Figura 1 - El sistema debe incorporar al menos un sistema de captura de imágenes (p. ej. una cámara) representado en (1), donde se observa cómo el usuario puede capturar con el sistema una imagen de su entorno. La imagen capturada será procesada por el módulo de localización de clases de objetos, representado en (2). El resultado del módulo de detección de clases de objetos (2) se representa en (3), donde observamos cómo el sistema enmarca las clases que es capaz de identificar en la imagen proporcionada por el usuario. En el ejemplo representado, el módulo ha detectado cuatro objetos de la clase “persona” en la imagen.

Figura 2 - Se muestra un posible diseño para la interfaz gráfica (4), a través del la cual el sistema interactuará con el usuario, tanto de forma visual como auditiva.

30

MODO DE REALIZACIÓN

Una implementación práctica de la invención detallada se corresponde con un sistema móvil que disponga de al menos una cámara, un subsistema para el procesamiento de los algoritmos de reconocimiento de clases de objetos y una pantalla o visor. El sistema dispondrá de un interfaz 5 táctil para la interacción con el usuario, además de un altavoz para la emisión de señales auditivas. El módulo de detección de clases de objetos tomará la imagen capturada por la cámara del sistema, y ejecutará un procesamiento de reconocimiento de clases objetos. Este procesamiento se realizará mediante la implementación de un software específico que pueda ser ejecutado en el sistema operativo que controla el sistema móvil. La pantalla del sistema será 10 utilizada para visualizar tanto las detecciones de objetos como la imagen capturada y parte del interfaz de usuario. El interfaz de usuario desarrollado para el sistema permitirá al usuario interactuar con la aplicación. La alimentación del sistema será obtenida de la batería incorporada en el sistema móvil.

15

Referencias:

- [1] Torres, S. Sistemas alternativos de comunicación: Manual de comunicación aumentativa y alternativa: Sistemas y estrategias. Archidona, 2001.
- [2] www.xtec.cat/~jlagares/f2kesp.htm
- 20 [3] http://www.ceapat.es/ceapat_01/index.htm
- [4] <http://www.mayer-johnson.com/c-speak-aphasia/>
- [5] Linebarger, M. C., Romani, John F. Aphasia therapy system. Número de Solicitud: US19980060313 19980415.
- [6] O'Brien, Daniel C.; Buchholz, ED. Communication system for aphasia patients comprising 25 wireless transceiver for transmitting and receiving communications from other devices. Número de solicitud: NZ20080581184 2008042.

REIVINDICACIONES

1. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, que consta de:
 - 5 a. Un subsistema de captura de imágenes encargado de obtener las imágenes que serán procesadas por el sistema.
 - b. Un subsistema para el procesado digital de imágenes, computador, procesador o cualquier otro sistema que permita operar con las imágenes capturadas.
 - 10 c. Un subsistema de reconocimiento y localización de clases de objetos en imágenes, que contiene el software necesario para realizar estas tareas, aplicando técnicas de visión artificial.
 - d. Un interfaz que permitirá al usuario controlar e interactuar con la aplicación.
 - e. Una pantalla donde el sistema visualizará el resultado de la detección y parte del interfaz de usuario.
- 15 2. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, según reivindicación 1, caracterizado por estar implementado en una arquitectura móvil, smartphone, tablet u ordenador portátil, donde se integren todos los subsistemas y
- 20 elementos descritos en la reivindicación 1.
3. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, según reivindicaciones 1 y 2, caracterizado por disponer de múltiples cámaras. Éstas serán
- 25 utilizadas para realizar las tareas de localización de clases de objetos en las imágenes que capturen, y/o funcionar como sensores de un interfaz natural de usuario para interactuar y controlar el sistema.
4. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la
- 30 capacidad comunicativa en personas con alteraciones del lenguaje, según

reivindicaciones 1 y 2, caracterizado por disponer de una o múltiples cámaras de tiempo de vuelo, o por cualquier sistema de cámaras que permita recuperar, para cada píxel de la imagen, una medida de profundidad o distancia. La información proporcionada por estas cámaras podrá ser utilizada para realizar las tareas de localización y reconocimiento de clases de objetos en las imágenes que capturen, y/o funcionar como sensores de un interfaz natural de usuario para interactuar y controlar el sistema.

5

5. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, según reivindicaciones 1 y 2, caracterizado por disponer de un interfaz táctil que permita al usuario interactuar con el sistema.

10

6. Sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, según reivindicaciones 1 y 2, caracterizado por disponer de un micrófono que permite al usuario controlar e interactuar con el sistema, controlándolo por voz.

15

20

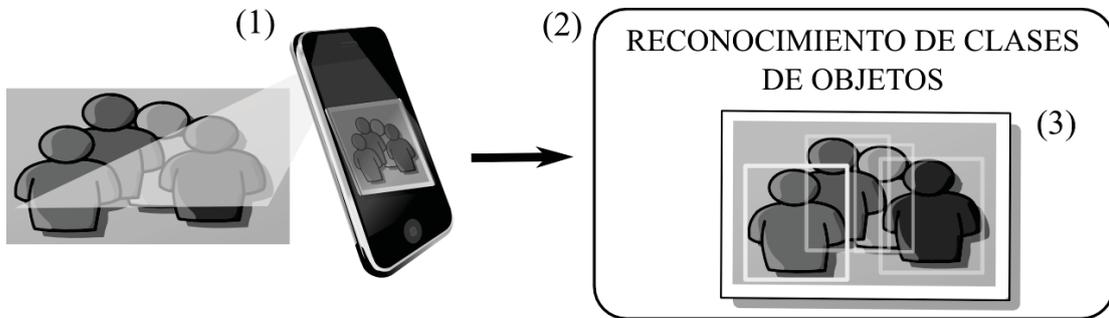


Figura 1

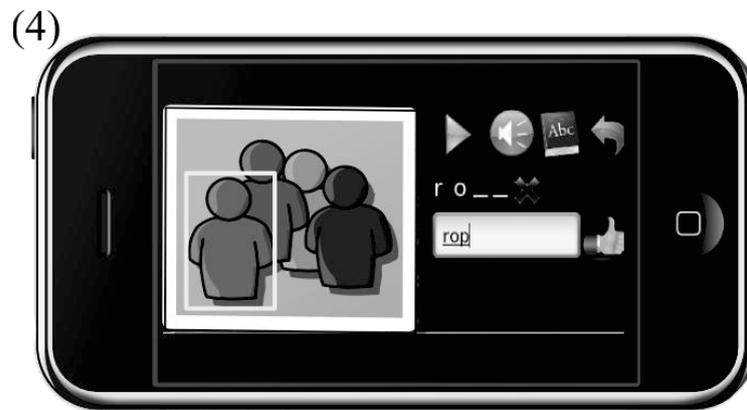


Figura 2



- ②① N.º solicitud: 201132071
②② Fecha de presentación de la solicitud: 21.12.2011
③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	WO 2005017820 A1 (SCAPE AS et al.) 24.02.2005	1
A	US 2003059106 A1 (GUTTA SRINIVAS) 27.03.2003	1
A	US 2009316988 A1 (XU NING et al.) 24.12.2009	1
A	US 5666441 A (RAO KASHI et al.) 09.09.1997	1

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
11.06.2013

Examinador
M. C. González Vasserot

Página
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

G09B21/00 (2006.01)

G06T7/00 (2006.01)

G06K9/00 (2006.01)

G06K9/46 (2006.01)

G06K9/62 (2006.01)

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

G09B, G06T, G06K

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

INVENES, EPODOC, WPI

Fecha de Realización de la Opinión Escrita: 11.06.2013

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 1-6	SI
	Reivindicaciones	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 1-6	SI
	Reivindicaciones	NO

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

Base de la Opinión.-

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

1. Documentos considerados.-

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

Documento	Número Publicación o Identificación	Fecha Publicación
D01	WO 2005017820 A1 (SCAPE AS et al.)	24.02.2005
D02	US 2003059106 A1 (GUTTA SRINIVAS)	27.03.2003
D03	US 2009316988 A1 (XU NING et al.)	24.12.2009
D04	US 5666441 A (RAO KASHI et al.)	09.09.1997

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

Los documentos citados solo muestran el estado general de la técnica, y no se consideran de particular relevancia. Así, la invención reivindicada se considera que cumple los requisitos de novedad, actividad inventiva y aplicación industrial.

1.- El objeto de la presente solicitud de patente está dentro de las tecnologías de la información y las comunicaciones, y dentro del campo del reconocimiento de clases de objetos mediante técnicas de visión artificial. También, dentro del campo de metodologías y terapias para el tratamiento de alteraciones del lenguaje. El sistema tiene como objetivo la aplicación de algoritmos de reconocimiento de clases de objetos en imágenes para mejorar la capacidad comunicativa en personas con alteraciones del lenguaje. Para ello se emplea un sistema de captura de imágenes, encargado de obtener la imagen que será procesada por un módulo que implementa un sistema de localización de clases de objetos en imágenes, el cual está basado en técnicas de visión artificial. Este módulo ofrecerá una salida visual en la que se recuadran las clases de objetos reconocidas en la imagen. El sistema interactuará con el usuario, ofreciéndole la información del objeto localizado, o interrogándole sobre el mismo para que sea él quien introduzca la información relativa a la clase reconocida. En este segundo modo, el sistema informará al usuario si el ejercicio se completó correctamente.

2.- El problema planteado por el solicitante es utilizar algún sistema para la captura de imágenes, alguna técnica de reconocimiento de clases de objetos para realizar el reconocimiento en tiempo real de los objetos que aparecen en imágenes capturadas por el mismo usuario. Se aplican técnicas de visión artificial, para que el sistema pueda reconocer automáticamente clases de objetos en las imágenes que el usuario captura mientras interactúa con el sistema, es decir, que éstas no tienen que estar incluidas en el sistema con antelación. El sistema realizará un reconocimiento automático de las clases de objetos que estén presentes en la imagen, localizando las clases detectadas (p. ej. la clase persona) en la misma imagen. Por último, el sistema dispondrá de una interfaz de usuario que le permitirá al mismo interactuar con la aplicación, de modo que pueda mejorar su capacidad comunicativa. Este interfaz, ofrecerá la información relativa al objeto detectado, tanto de forma visual como auditiva, y además podrá utilizarse para interrogar al usuario sobre el objeto reconocido, de modo que sea él mismo el que introduzca el nombre de la clase detectada. El sistema informará al usuario de si el ejercicio se completó correctamente.

El documento D1 puede considerarse como el representante del estado de la técnica más cercano ya que en este documento confluyen la mayoría de las características técnicas reivindicadas.

Análisis de las reivindicación independiente 1

D1 se diferencia del documento de solicitud de patente en que no es un sistema de reconocimiento de clases de objetos por visión artificial para la mejora de la capacidad comunicativa en personas con alteraciones del lenguaje, que consta de:

- Un subsistema de captura de imágenes encargado de obtener las imágenes que serán procesadas por el sistema.
- Un subsistema para el procesado digital de imágenes, computador, procesador o cualquier otro sistema que permita operar con las imágenes capturadas.
- Un subsistema de reconocimiento y localización de clases de objetos en imágenes, que contiene el software necesario para realizar estas tareas, aplicando técnicas de visión artificial.
- Un interfaz que permitirá al usuario controlar e interactuar con la aplicación.
- Una pantalla donde el sistema visualizará el resultado de la detección y parte del interfaz de usuario.

La reivindicación 1 es nueva (Art. 6.1 LP 11/1986) y tiene actividad inventiva (Art. 8.1 LP11/1986).

Análisis del resto de los documentos

De este modo, ni el documento D1, ni ninguno del resto de los documentos citados en el Informe del Estado de la Técnica, tomados solos o en combinación, revelan la invención en estudio tal y como es definida en las reivindicaciones independientes, de modo que los documentos citados solo muestran el estado general de la técnica, y no se consideran de particular relevancia. Además, en los documentos citados no hay sugerencias que dirijan al experto en la materia a una combinación que pudiera hacer evidente la invención definida por estas reivindicaciones y no se considera obvio para una persona experta en la materia aplicar las características incluidas en los documentos citados y llegar a la invención como se revela en la misma.