

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 723 123**

21 Número de solicitud: 201831043

51 Int. Cl.:

G06K 9/62 (2006.01)

G06T 7/30 (2007.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

29.10.2018

43 Fecha de publicación de la solicitud:

21.08.2019

71 Solicitantes:

UNIVERSIDAD DE LEÓN (100.0%)
Avenida de La Facultad, 25
24071 León ES

72 Inventor/es:

BISWAS, Rubel;
GONZÁLEZ CASTRO, Víctor;
ALEGRE GUTIÉRREZ, Enrique y
FIDALGO FERNÁNDEZ, Eduardo

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

54 Título: **PROCEDIMIENTO Y SISTEMA DE DETECCIÓN DE SERVICIOS EN PÁGINAS WEB BASADO EN LA OBTENCIÓN DE FIRMAS DIGITALES DE LA IMAGEN DE SU PÁGINA DE INICIO**

57 Resumen:

La presente invención da a conocer un procedimiento de detección de un servicio asociado a una página web caracterizado porque comprende: obtener una imagen a partir de una captura de pantalla de la página web; procesar la imagen para obtener una firma digital asociada al contenido de la imagen; calcular la similitud entre la firma digital asociada a la imagen y una serie de firmas digitales de referencia; en función de la similitud calculada, correlacionar la firma digital asociada a la imagen con una determinada firma digital de referencia; determinar el servicio de la firma digital de referencia; y asociar el servicio a la página web.

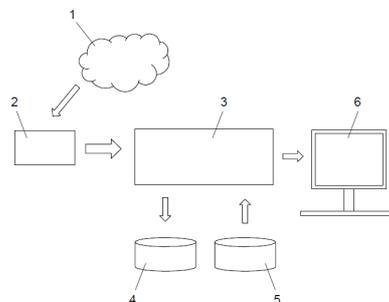


FIG. 1

DESCRIPCIÓN

PROCEDIMIENTO Y SISTEMA DE DETECCIÓN DE SERVICIOS EN PÁGINAS WEB BASADO EN LA OBTENCIÓN DE FIRMAS DIGITALES DE LA IMAGEN DE SU PÁGINA DE INICIO

5

OBJETO DE LA INVENCION

El objeto de la presente invención es un procedimiento y sistema de detección de servicios en páginas web. La presente invención da a conocer un procedimiento y un sistema que permite 10 determinar servicios que se pueden encontrar en internet, en especial en redes de la web oscura, en función de su aspecto visual. Específicamente, la invención permite reconocer si un dominio ofrece alguno de los servicios determinados como de interés en base a su aspecto visual y, en caso afirmativo, determinar de qué servicio se trata. En una realización preferente de la presente invención, el reconocimiento se realiza comparando imágenes asociadas a una 15 captura de pantalla de dichos dominios con imágenes de referencia almacenadas representando tanto la captura como la plantilla mediante técnicas de *hashing* perceptual.

ANTECEDENTES DE LA INVENCION

Originalmente, el propósito de redes como TOR (*The Onion Router*) era el de proporcionar 20 anonimato al usuario de la red y a su localización con el objetivo de, por ejemplo, sortear las restricciones del uso de la red en países sin plena libertad de expresión. En algunos casos, el anonimato que proporcionan estas redes ha motivado su uso por parte de usuarios con intenciones maliciosas o delictivas.

25 En concreto, algunos usuarios pueden ofrecer servicios maliciosos o delictivos que, por razones de seguridad, conviene identificar para poder actuar en consecuencia (M. Spitters, S. Verbruggen and M. v. Staalduinen, "Towards a Comprehensive Insight into the Thematic Organization of the Tor Hidden Services," 2014 IEEE Joint Intelligence and Security Informatics Conference, The Hague, 2014, pp. 220-223). Algunos ejemplos de dichos 30 servicios son los dominios para el intercambio de ficheros o los paneles destinados a que víctimas de algún *Ransomware* paguen dinero para poder desbloquear sus ordenadores. Tal y como se ha mencionado anteriormente, conviene identificar el tipo de servicios que proporciona un dominio determinado para disponer de dominios categorizados y, eventualmente, filtrarlos. Es habitual que esta identificación se realice de manera manual en 35 función de la apariencia visual de cada página.

Las técnicas de clasificación de imágenes utilizan clasificadores automáticos para asignar imágenes a una determinada clase en función de su contenido (E. Fidalgo, E. Alegre, V. González-Castro, L. Fernández-Robles, “Illegal Activity Categorisation in DarkNet Based on Image Classification Using CREIC Method”, International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6--8, 2017, Proceedings, pp. 600-609, 2018). Para que la asignación automática sea fiable, es necesario entrenar el clasificador con un número suficientemente grande de imágenes de cada clase.

10 DESCRIPCIÓN DE LA INVENCION

Normalmente, no es posible contar con un número suficiente de imágenes para entrenar un clasificador que automatice el reconocimiento del tipo de servicio que se proporciona en un dominio determinado utilizando una captura de pantalla del mismo. Por ello, resulta interesante utilizar técnicas de *hashing* perceptual que permiten realizar una clasificación con unas pocas imágenes de referencia con una firma digital, o código *hash*, previamente clasificado, con objeto de comparar con las de aquellas cuyo tipo se quiere averiguar.

El procedimiento y sistema para la detección de servicios en un determinado dominio web permite reconocer, utilizando una captura de pantalla de dicho dominio, si éste proporciona uno de los servicios previamente determinados y de los que también se dispone de una imagen de su página principal. Además, en caso de ser reconocido como uno de los servicios listados, este procedimiento y sistema permite indicar automáticamente el tipo de servicio de que se trata.

El reconocimiento utilizando el procesamiento de la imagen tiene como ventaja frente al reconocimiento manual, realizado por un experto, que el primero anula la subjetividad, los errores debidos al cansancio o falta de atención, la disparidad de criterio entre los expertos, los costes asociados al tiempo del experto y el tiempo necesario para su realización, y aumenta la fiabilidad. La presente invención puede ser utilizada por los organismos policiales o judiciales para reconocer si determinados dominios de una red contienen servicios maliciosos o ilegales y, de ser así, el tipo de servicio.

La presente invención se refiere a un sistema y a un método que, analiza una imagen conteniendo una captura de pantalla de una o varias páginas web asociadas a un dominio y

extrae su firma digital o código *hash* perceptual. Dicha firma digital es posteriormente utilizada para identificar algunos servicios que dicha web pudiera estar ofreciendo.

5 El *hashing* perceptual permite construir una firma digital a partir de una imagen o vídeo derivada de características de su contenido (M. Schneider, S-F. Chang, "A robust content based digital signature for image authentication", Proceedings of 3rd IEEE International Conference on Image Processing, Lausanne, 1996, pp. 227-230 vol.3). De este modo, el *hashing* perceptual se puede utilizar para recuperar imágenes de conjuntos grandes (L. Xie, L. Zhu, P. Pan, Y. Lu, "Cross-Modal Self-Taught Hashing for large-scale image retrieval",
10 Signal Processing, Vol. 124, pp. 81-92, 2016) o para realizar autenticación de imágenes (J. Ouyang, G. Coatrieux, H. Shu, "Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform", Digital Signal Processing, Vol. 41, pp. 98-109, 2015).

15 La presente invención contempla la comparación de esta firma digital o código *hash* perceptual con una serie de datos de referencia correspondientes a firmas digitales de determinados servicios conocidos, por lo que permite indicar si la captura podría corresponder a un servicio determinado. El procedimiento que se propone comprende las siguientes etapas:

- 20 - obtener una imagen a partir de una captura de pantalla de la página web;
- procesar la imagen para obtener una firma digital asociada al contenido de la imagen;
- calcular la similitud entre la firma digital asociada a la imagen y una serie de firmas digitales de referencia almacenadas;
- en función de la similitud calculada, correlacionar la firma digital de la imagen capturada con una de las firmas digitales de referencia;
- 25 - determinar un servicio al que corresponde la firma digital de referencia que es más similar a la firma de la imagen capturada, es decir, un servicio asociado a la firma digital de referencia; y
- asociar el servicio anterior a la página web que se está procesando.

30 En un ejemplo de realización, la imagen es una imagen pre-procesada de la captura de pantalla, por ejemplo, la imagen es una imagen filtrada obtenida a partir de la filtración mediante un filtro gaussiano de la captura de pantalla. Preferentemente, la imagen es una imagen en escala de grises asociada a la captura de pantalla.

35

Más preferentemente, la imagen es una imagen asociada a la captura de pantalla obtenida tras la aplicación de una transformada discreta del coseno.

En una realización, la obtención de la firma digital comprende la realización de un hashing perceptual. El hashing perceptual puede comprender las etapas de:

- 5 - definir una pluralidad de matrices de búsqueda de la imagen;
- definir una serie de sub-matrices vecindario contenidas en la matriz de búsqueda de manera que las sub-matrices vecindario tienen un tamaño menor que la matriz de búsqueda;
- 10 - definir una sub-matriz de referencia contenida en la matriz de búsqueda del mismo tamaño que las sub-matrices vecindario; y
- operar cada sub-matriz vecindario con la sub-matriz de referencia obteniendo una matriz de mapa local (F-DNS) para cada una de las matrices de búsqueda.
- operar las matrices de mapas locales (F-DNS) para obtener una matriz mapa global (F-GNS).
- 15 - vectorizar la matriz de mapa global (F-GNS) obteniendo la firma digital.

Preferentemente, la pluralidad de matrices de búsqueda tiene un tamaño $n \times n$ y las sub-matrices vecindario tienen un tamaño $p \times p$ siendo n múltiplo de p .

Además, la operación entre cada sub-matriz vecindario con la sub-matriz de referencia puede comprender, por ejemplo, vectorizar las sub-matrices vecindario y referencia obteniendo vectores vecindario y referencia respectivamente y calcular la distancia Euclídea entre los vectores vecindario y los vectores referencia.

En una realización, la operación entre las matrices de mapa local (F-DNS), para obtener la matriz de mapa global (F-GNS), es una suma elemento a elemento de todas ellas.

Preferentemente, la similitud entre la firma digital y las firmas digitales de referencia se realiza mediante el cálculo del coeficiente de correlación, según la siguiente ecuación:

$$S(H_1, H_2) = \frac{\sum_{i=1}^n (h_i^{(1)} - \mu_1)(h_i^{(2)} - \mu_2)}{\sqrt{\sum_{i=1}^n (h_i^{(1)} - \mu_1)^2} \sqrt{\sum_{i=1}^n (h_i^{(2)} - \mu_2)^2}}$$

siendo $H_1 = [h_1^{(1)}, h_2^{(1)}, \dots, h_n^{(1)}]$ la firma digital de la imagen a analizar; $H_2 = [h_1^{(2)}, h_2^{(2)}, \dots, h_n^{(2)}]$ la firma digital de referencia; μ_1 la media de la firma digital de la imagen a analizar y μ_2 ; la

media de la firma digital de referencia

En un ejemplo de realización, se correlaciona la firma digital con una firma digital de referencia, de entre todas las evaluadas, si al menos el valor de S es mayor que 0,95. Preferentemente, se correlaciona la firma digital con una firma digital de referencia, preferentemente cuando el valor de S es mayor que 0,987.

En una realización preferente, las firmas digitales de referencia se encuentran almacenadas en una memoria estando cada una de ellas correlacionada con un servicio.

En un ejemplo, las firmas digitales de referencia se pueden obtener utilizando el mismo procesamiento para su obtención que las firmas digitales asociadas al contenido de la imagen.

Además, el procedimiento según la presente invención puede comprender realizar una serie de iteraciones con una serie de imágenes a partir de una serie captura de pantalla de páginas web de un determinado dominio.

Es, además, un objetivo de la presente invención dar a conocer un sistema para la detección de un servicio asociado a una página web caracterizado porque comprende una memoria en la que se encuentran almacenadas una serie de firmas de referencia relacionadas con una serie de servicios y un procesador configurado para:

- obtener una imagen a partir de una captura de pantalla de la página web;
- procesar la imagen para obtener una firma digital asociada al contenido de la imagen;
- calcular la similitud entre la firma digital asociada a la imagen y una serie de firmas digitales de referencia;
- en función de la similitud calculada, correlacionar la firma digital asociada a la imagen con una determinada firma digital de referencia;
- determinar un servicio de a la firma digital de referencia; y
- asociar el servicio a la página web;
- almacenar en una memoria las asociaciones determinadas entre servicios y páginas web.

Preferentemente, la imagen es una imagen pre-procesada de la captura de pantalla. Por ejemplo, la imagen es una imagen filtrada que puede ser obtenida a partir de la filtración

mediante un filtro gaussiano de la captura de pantalla.

Más preferentemente, la imagen es una imagen en escala de grises asociada a la captura de pantalla.

5

En un ejemplo de realización, la imagen es una imagen asociada a la captura de pantalla obtenida tras la aplicación de una transformada discreta del coseno.

En una realización preferente, la obtención de la firma digital comprende la realización, mediante el procesador, de un hashing perceptual.

10

Además, la presente invención da a conocer un medio de almacenamiento no-transitorio con capacidad para ser leído por un procesador caracterizado porque comprende instrucciones para que el procesador ejecute un procedimiento según se ha explicado anteriormente.

15

En un ejemplo de realización especialmente preferente, el procedimiento comprende las siguientes etapas:

1. Adquisición y preprocesamiento de las imágenes:

20

En esta etapa se descargan capturas de pantalla obtenidas automáticamente, por un dispositivo con conexión a Internet, de páginas web alojadas en alguna red de Internet. Preferentemente, las imágenes se transfieren a la memoria de un ordenador donde, en caso de ser capturadas en color se convierten a escala de grises, funcionando de manera análoga si las imágenes se capturan directamente en escala de grises. A continuación, y de acuerdo con una realización preferente de la invención, se suaviza el contraste de la imagen utilizando un filtro gaussiano.

25

2. Extracción del código *hash* perceptual de las imágenes:

30

En esta etapa se extraen las características de la imagen y a partir de ellas se construye el código *hash* perceptual de la misma. En primer lugar, se calcula la Transformada Discreta del Coseno (DCT, del inglés *discrete cosine transform*). A continuación, se calcula, sobre la imagen resultado de la aplicación de la DCT, el descriptor de textura basado en estructuras del vecindario dominante (DNS, del inglés *dominant neighborhood structure*) (F. M. Khellah, "Texture Classification Using Dominant Neighborhood Structure", IEEE Transactions on Image Processing, Vol. 20,

35

No. 11, pp. 3270-3279, 2011) de varias regiones cuadradas o rectangulares de la imagen. Puesto que estas DNS se han calculado sobre la imagen resultante de aplicar la DCT, en vez de sobre la imagen original, en lo sucesivo se llamará F-DNS. De acuerdo a una realización preferente de la invención, dichas regiones tendrán un tamaño de 9x9 píxeles y sus píxeles centrales estarán separados por 9 píxeles tanto en dirección horizontal como en vertical. En un siguiente paso, se suman, elemento a elemento todas las F-DNS anteriormente calculadas para obtener una estructura de vecindario global (GNS, del inglés *global neighbourhood structure*) que, en este caso se llamará F-GNS. En una realización preferente de la invención, se prescinde de la primera fila y de la primera columna para, finalmente, ordenar todos los elementos en un solo vector de 64 números reales.

3. Cálculo de la medida de similitud entre el código *hash* extraído y los códigos *hash* de las plantillas:

Una vez extraído el código hash de la captura de pantalla se compara con los códigos hash de las plantillas de los servicios almacenadas previamente para determinar si la captura de pantalla corresponde a uno de ellos, según el valor de similitud entre dichos códigos hash. De acuerdo a una realización preferente de la invención la similitud entre dos códigos *hash* perceptuales se realiza mediante el coeficiente de correlación entre ellos, aunque se podría utilizar cualquier otra medida de similitud o distancia entre dos vectores de números reales. Además, de acuerdo a una realización preferente de la invención, se indicará que la imagen de la captura corresponde a un servicio si el mayor de los coeficientes de correlación calculados supera el valor de 0,987.

25 BREVE DESCRIPCIÓN DE LOS DIBUJOS

A continuación, se describen de manera muy breve una serie de figuras que ayudan a comprender mejor la invención y que se relacionan expresamente con una realización de dicha invención que se presenta como un ejemplo no limitativo de ésta.

30 La Fig. 1 muestra un esquema simplificado de un ejemplo de sistema según la presente invención

La Fig. 2 muestra un diagrama de bloques que exponen los pasos que comprenden el procedimiento.

35

La Fig. 3 muestra una representación del cálculo de un DNS.

La Fig. 4 muestra una representación del cálculo de un GNS a partir de diversos DNS.

5 REALIZACIÓN PREFERENTE DE LA INVENCION

La Fig. 1 muestra un esquema simplificado de un ejemplo de sistema. En el ejemplo de la Fig. 1 se muestra un dispositivo (2) que obtiene automáticamente capturas de pantalla de una serie de páginas web asociadas a dominios alojados en una red (1), por ejemplo, una red asociada a una “red oscura” o darknet. De acuerdo a un ejemplo de procedimiento de la presente invención, los dominios estarán alojados en la red TOR. En lo sucesivo denominamos “imagen digital” o, simplemente, “imagen” a cada una de las imágenes asociadas a las capturas de pantalla obtenidas.

El dispositivo (2) está conectado a un procesador (3), preferentemente asociado a un ordenador, aunque puede tratarse de un sistema portable como puede ser una Raspberri Pi, mediante una conexión USB, Firewire, o de cualquier otra manera que el dispositivo (2) lo requiera.

El dispositivo de captura (2) realiza una captura de pantalla automática del dominio de la Red (1) que se le indica. A continuación, la imagen capturada se transfiere al procesador (3), donde se puede realizar un pretratamiento de la misma. Dicho pretratamiento puede comprender una primera etapa consistente en la conversión de la imagen a escala de grises, por ejemplo, si dicha imagen capturada es una imagen a color. A continuación, se realiza un filtrado gaussiano para suavizar la imagen. De acuerdo con una realización preferente de la invención este filtro gaussiano se calcula con un valor de σ igual a 1.

El procesador (3) está configurado para recibir una captura de una página web a analizar y puede realizar un pretratamiento de la imagen bien sea para reducir el coste computacional del procesamiento o para resaltar aspectos concretos de la captura que pueden ser relevantes para establecer si ofrece un servicio concreto. En un ejemplo de pretratamiento, el procesador (3) puede aplicar una transformada discreta del coseno a la captura. En cualquier caso, exista pretratamiento o no, se dispone de una imagen que será el objeto del procesamiento para determinar si la página web corresponde a un determinado servicio.

El procesador (3) dispone, adicionalmente, de medios de procesamiento de la imagen para

obtener una firma digital a partir de su contenido. En un ejemplo de la presente invención, la firma digital está relacionada con la información propia de la imagen, es decir, de la información contenida en la misma. Por tanto, el procesador (3) realiza operaciones con la información contenida en la imagen a fin de obtener, a partir de dichas operaciones, la firma digital. En una
5 realización particular, las operaciones en la imagen comprenden un hash perceptual.

Una vez obtenida la firma digital, el sistema objeto de la presente invención dispone de medios de almacenamiento de datos en los que se puede almacenar tanto la firma digital obtenida de la imagen a tratar como las firmas digitales de las imágenes de referencia de los distintos servicios.
10 Por un lado, la firma digital obtenida al procesar la imagen se puede almacenar en la RAM del sistema de procesamiento (3) o en un sistema de almacenamiento específico en forma de un primer archivo (4). Y, por otro, las firmas digitales de referencia se almacenarán igualmente en otro archivo o sistema de almacenamiento en forma de un segundo archivo (5) o, alternativamente, también se puede almacenar en la RAM del sistema de procesamiento (3).

15 Las firmas digitales de referencia están correlacionadas con una serie de servicios, de manera que, si hay una similitud entre la firma digital de la imagen a analizar y una de las firmas de referencia, el procesador asigna el servicio relacionado a la firma de referencia asociada a la imagen y, en consecuencia, a la página web objeto del análisis. En definitiva, las firmas digitales
20 de referencia obtenidas, por ejemplo, a partir de imágenes, sirven de plantilla de los servicios que se quieren reconocer.

Para determinar si existe similitud, el procesador (3) está configurado para comparar dos firmas obteniendo, por ejemplo, el coeficiente de correlación entre ambas tal y como explicará más
25 adelante. Posteriormente, se analiza si la imagen capturada puede relacionarse con uno de los servicios asociados a las firmas digitales de referencia y, en caso afirmativo, se puede mostrar a través de una pantalla de visualización (6) el servicio al que corresponde la captura de pantalla o almacenar esta información en una memoria para su posterior utilización.

30 La Fig. 2 muestra un ejemplo de una serie de transformaciones realizadas para obtener una firma digital en función del contenido de la imagen capturada o tratada. En el ejemplo de la Fig. 2 las transformaciones realizadas se basan en un método de hash perceptual, en concreto, se muestran las transformaciones por las que pasa la imagen desde que se captura la imagen original (7) hasta que se obtiene el código del *hash* perceptual (12).

35

En primer lugar, se puede preprocesar la imagen (8), por ejemplo, mediante un filtrado gaussiano, y pasar la imagen a escala de grises. A continuación, se calcula la Transformada Discreta del Coseno (DCT, del inglés *Discrete Cosine Transform*) (9) de la imagen preprocesada obteniendo una imagen transformada. Si bien otras transformadas como son la de Fourier, Walsh o Hadamard, se emplean en procesamiento digital de imagen, y en especial cuando se quiere reducir la información eliminando aquellas partes de la señal que podrían considerarse menos significativas, hay dos propiedades que hacen que la DCT sea especialmente apropiada para la creación de funciones de hash perceptual. Una de esas propiedades es que la DCT consigue condensar la gran mayoría de la información de la señal original, en este caso una imagen, en pocos coeficientes transformados. La otra propiedad es que el resultado de la transformación se mantiene muy similar cuando las relaciones entre los valores de la señal de entrada son similares, independientemente de que proceda de dos señales muy diferentes entre sí.

A partir de esta imagen transformada (9), se calculan una serie de estructuras del vecindario dominante (DNS, del inglés *Dominant Neighborhood Structure*) (10) que se combinarán en una sola estructura de vecindario global (GNS, del inglés *global neighbourhood structure*) (11) a partir de la cual se obtendrán los componentes del código *hash* perceptual (12) de la imagen original (7). Dicho cálculo de DNS y GNS se explicará con mayor detalle en las figuras 3 y 4.

Según la presente invención, el *hash* perceptual (12) es un código asociado a los contenidos de la captura de pantalla (7). Para ello, en primer lugar, se calcula la transformada discreta del coseno (9) de la imagen asociada a la captura de pantalla (7) obteniendo una imagen transformada. Posteriormente se calculan una serie de estructuras de vecindario dominante (DNS) (10) llamadas en este caso F-DNS, por estar calculadas a partir de la imagen transformada (9) o, en definitiva, de cualquier imagen asociada u obtenida a partir de la captura de pantalla (7).

La Fig. 3 muestra un ejemplo de cómo se calcula una matriz capa local partiendo de una imagen (9) mediante un procedimiento DNS. En primer lugar, se definen una serie de matrices de búsqueda (13), por ejemplo, tal y como se muestra en la Fig. 3 la matriz de búsqueda es una matriz de tamaño 9x9. Alrededor de cada uno de los píxeles (17) contenidos en la matriz de búsqueda (13) se considera una sub-matriz de tamaño 3x3 llamada sub-matriz vecindario (16) y una sub-matriz de referencia (15) del mismo tamaño que la sub-matriz vecindario (16) que se define alrededor del píxel central (14) de la matriz de búsqueda (13). A continuación, se vectorizan las sub-matrices de referencia (15) y vecindario (16), por ejemplo, ordenando

los valores de los píxeles en forma vectorial copiándolos de izquierda a derecha y de arriba abajo obteniendo un vector de referencia (18) y un vector vecindario (19). Finalmente, se calcula la distancia Euclídea (100) entre ambos vectores y se copia este valor (101) en la misma posición (17) en la que se encontraba el píxel central de la ventana de vecindario (16),
 5 dentro de una matriz capa local, o matriz F-DNS (10) en la posición (20) correspondiente a los vectores en la matriz de búsqueda (13). Realizando este proceso sobre todos los píxeles de la ventana de búsqueda (13) se obtendrá la matriz F-DNS (10) completa. En un ejemplo de realización, la matriz de búsqueda (13) es una matriz cuadrada de dimensión $n \times n$ y las sub-matrices de referencia y vecindario son matrices cuadradas de dimensión $p \times p$ siendo n
 10 múltiplo de p .

Siguiendo el mismo proceso para las demás matrices de búsqueda (13) de la serie de imágenes de búsqueda en la imagen (9) centradas en píxeles separados por un número de filas o columnas determinado, y repitiendo este proceso, se obtendrán tantos mapas F-DNS
 15 (10) como matrices de búsqueda (13) se hayan considerado. De acuerdo con una realización preferente de la invención, el número tanto de filas como de columnas de separación entre los píxeles centrales es 9 para optimizar el coste computacional, es decir, corresponde al tamaño n de la matriz de búsqueda (13).

Finalmente se procede a operar la serie de matrices de capas locales para obtener una matriz mapa global, o matriz F-GNS. La Fig. 4 muestra un ejemplo de obtención de una matriz de mapa global, o matriz F-GNS (11). En este ejemplo, se considerarán todos los mapas F-DNS
 20 (10) obtenidos de la imagen (9) y se suman (200) estas matrices para obtener el mapa F-GNS (11), es decir, se suman los valores de los píxeles en una posición de todas las matrices y se copia dicho valor en la misma posición (21) del mapa F-GNS (11).
 25

Una vez calculado el mapa F-GNS de la imagen transformada (11) se puede prescindir de algunos datos, como por ejemplo, de su primera fila y columna y se vectorizarán los valores de la matriz resultante utilizando un procedimiento similar al utilizado con las sub-matrices de
 30 referencia y vecindario, copiando sus valores de izquierda a derecha y de arriba abajo. Los valores contenidos en este vector formarán la firma digital o código *hash* perceptual de la misma (12).

Una vez que se ha calculado la firma digital (12) de la imagen, se procede a compararlo con
 35 una serie de firmas digitales de referencias que, por ejemplo, pueden ser obtenidas a través

de imágenes siguiendo el mismo procedimiento de firma digital y clasificándolas como servicios de forma manual o, simplemente, como firmas digitales correspondientes a imágenes modelo previamente almacenadas en una memoria.

5 Un ejemplo de procedimiento para efectuar esta comparación se realiza calculando una medida de similitud entre las firmas digitales (12) y una serie de firmas digitales de referencia. De acuerdo a una realización preferente de la invención, la medida de similitud calculada es el coeficiente de correlación entre ambos códigos *hash*. Sean $H_1 = [h_1^{(1)}, h_2^{(1)}, \dots, h_n^{(1)}]$ y $H_2 = [h_1^{(2)}, h_2^{(2)}, \dots, h_n^{(2)}]$ los códigos *hash* perceptuales cuya similitud se quiere medir, y sean μ_1 y μ_2 sus respectivas medias. Entonces, el coeficiente de correlación entre el código hash correspondiente a la imagen capturada (H_1) y el código hash correspondiente una firma digital de referencia (H_2) es:

$$S(H_1, H_2) = \frac{\sum_{i=1}^n (h_i^{(1)} - \mu_1)(h_i^{(2)} - \mu_2)}{\sqrt{\sum_{i=1}^n (h_i^{(1)} - \mu_1)^2} \sqrt{\sum_{i=1}^n (h_i^{(2)} - \mu_2)^2 + \varepsilon}}$$

15 donde ε puede ser cero o, en algunos casos, es una constante de valor muy pequeño, para evitar posibles divisiones entre cero. Esta medida está acotada entre -1 y 1 siendo mayor cuanto más similares visualmente son las dos imágenes a las que pertenecen los códigos *hash* H_1 y H_2 .

20 Una vez calculadas las similitudes entre la imagen capturada y cada una de las plantillas se determina si la plantilla corresponde con un servicio y, en caso afirmativo, de qué tipo de servicio se trata.

25 En primer lugar, se considerará que la imagen capturada corresponde a un servicio si la similitud máxima entre su código *hash* perceptual y los códigos *hash* perceptuales de las plantillas almacenadas para imágenes de dicho servicio (5) supera un determinado umbral. De acuerdo a la realización preferente de la invención este umbral deberá ser como mínimo 0,95 y, preferentemente, 0,987.

30 En segundo lugar, cuando la imagen capturada corresponde a un servicio, se considera que este es el mismo servicio que presenta la web con cuya plantilla asociada tiene la máxima similitud. Por lo tanto, se considerará que el dominio del que se capturó la imagen proporciona dicho servicio.

Un segundo aspecto de la presente invención se refiere a un sistema de visión artificial para la extracción de una firma digital o código *hash* perceptual basado en características del contenido de imágenes de capturas de pantalla de dominios de una red. El sistema comprende medios de procesamiento de datos, tales como un ordenador, configurados para realizar una
5 captura de pantalla del dominio; obtener la transformada discreta del coseno, las estructuras DNS y GNS de la misma y, a partir de esta última, la firma o código *hash* de la imagen. Finalmente, comparando esta con una serie de firmas o *hashes* de servicios previamente almacenadas es capaz de indicar el tipo de servicio de la web asociada a la imagen capturada, indicando que es el mismo servicio que el que presta el dominio que tiene una plantilla más
10 similar.

En una realización preferente de la invención, el sistema comprende un dispositivo con conexión a Internet o conectado a un ordenador con conexión a Internet que visita una serie de dominios y que realiza y que almacena capturas de pantalla de los dominios visitados. El
15 sistema también puede comprender unos medios de almacenamiento de datos donde se almacena un archivo de datos de la firma digital o código *hash* de la imagen y un archivo de datos con las firmas digitales o códigos *hash* de las plantillas de los servicios.

Por último, la presente invención también se refiere a un producto de programa que
20 comprende medios de instrucciones de programa para llevar a la práctica el procedimiento anteriormente descrito cuando el programa se ejecuta en un procesador. El producto de programa está preferentemente almacenado en un medio de soporte de programas. Los medios de instrucciones de programa pueden tener la forma de código fuente, código objeto, una fuente intermedia de código y código objeto, por ejemplo, como en forma parcialmente
25 compilada, o en cualquier otra forma adecuada para uso en la puesta en práctica de los procesos según la invención.

El medio de soporte de programas puede ser cualquier entidad o dispositivo capaz de soportar el programa. Por ejemplo, el soporte podría incluir un medio de almacenamiento, como una
30 memoria ROM, una memoria CD ROM o una memoria ROM de semiconductor, una memoria flash, un soporte de grabación magnética, por ejemplo, un disco duro o una memoria de estado sólido (SSD, del inglés *solid-state drive*). Además, los medios de instrucciones de programa almacenados en el soporte de programa pueden ser, por ejemplo, mediante una señal eléctrica u óptica que podría transportarse a través de cable eléctrico u óptico, por radio
35 o por cualquier otro medio.

Cuando el producto de programa va incorporado en una señal que puede ser transportada directamente por un cable u otro dispositivo o medio, el soporte de programa puede estar constituido por dicho cable u otro dispositivo o medio.

5

Como variante, el soporte de programa puede ser un circuito integrado en el que va incluido el producto de programa, estando el circuito integrado adaptado para ejecutar, o para ser utilizado en la ejecución de los procesos correspondientes.

10

REIVINDICACIONES

1. Procedimiento de detección de un servicio asociado a una página web caracterizado porque comprende:
- 5 - obtener una imagen a partir de una captura de pantalla de la página web;
 - procesar la imagen para obtener una firma digital asociada al contenido de la imagen;
 - calcular la similitud entre la firma digital asociada a la imagen y una serie de firmas digitales de referencia;
 - en función de la similitud calculada, correlacionar la firma digital de la imagen capturada
10 con una de las firmas digitales de referencia;
 - determinar un servicio asociado a la firma digital de referencia; y
 - asociar el servicio a la página web.
2. Procedimiento, según la reivindicación 1, caracterizado porque la imagen es una imagen
15 pre-procesada de la captura de pantalla.
3. Procedimiento, según la reivindicación 2, caracterizado porque la imagen es una imagen filtrada obtenida a partir de la filtración mediante un filtro gaussiano de la captura de pantalla.
- 20 4. Procedimiento, según cualquiera de las reivindicaciones 2 ó 3, caracterizado porque la imagen es una imagen en escala de grises asociada a la captura de pantalla.
5. Procedimiento, según cualquiera de las reivindicaciones 2 a 4, caracterizado porque la imagen es una imagen asociada a la captura de pantalla obtenida tras la aplicación de una
25 transformada discreta del coseno.
6. Procedimiento, según cualquiera de las reivindicaciones anteriores, caracterizado porque la obtención de la firma digital comprende la realización de un hashing perceptual.
- 30 7. Procedimiento, según la reivindicación 6 caracterizado porque el hashing perceptual comprende:
- definir una pluralidad de matrices de búsqueda de la imagen;
 - definir una serie de sub-matrices vecindario contenidas en la matriz de búsqueda de manera que las sub-matrices vecindario tienen un tamaño menor que la matriz de
35 búsqueda;

- definir una sub-matriz de referencia contenida en la matriz de búsqueda del mismo tamaño que las sub-matrices vecindario; y
 - operar cada sub-matriz vecindario con la sub-matriz de referencia obteniendo una matriz de mapa local (F-DNS) para cada una de las matrices de búsqueda.
- 5
- operar las matrices de mapas locales (F-DNS) para obtener una matriz mapa global (F-GNS).
 - vectorizar la matriz de mapa global (F-GNS) obteniendo la firma digital.

8. Procedimiento, según la reivindicación 7 caracterizado porque la pluralidad de matrices de búsqueda tiene un tamaño $n \times n$ y las sub-matrices vecindario tienen un tamaño $p \times p$ siendo n múltiplo de p .

9. Procedimiento, según cualquiera de las reivindicaciones 7 u 8, caracterizado porque la operación entre cada sub-matriz vecindario con la sub-matriz de referencia comprende vectorizar las sub-matrices vecindario y referencia obteniendo vectores vecindario y referencia respectivamente y calcular la distancia Euclídea entre los vectores vecindario y los vectores referencia.

10. Procedimiento, según cualquiera de las reivindicaciones 7 a 9, caracterizado porque la operación entre las matrices de mapa local (F-DNS), para obtener la matriz de mapa global (F-GNS), es una suma elemento a elemento de todas ellas.

11. Procedimiento, según cualquiera de las reivindicaciones 7 a 10 caracterizado porque la similitud entre la firma digital y las firmas digitales de referencia se realiza mediante el cálculo del coeficiente de correlación, según la siguiente ecuación:

$$S(H_1, H_2) = \frac{\sum_{i=1}^n (h_i^{(1)} - \mu_1)(h_i^{(2)} - \mu_2)}{\sqrt{\sum_{i=1}^n (h_i^{(1)} - \mu_1)^2} \sqrt{\sum_{i=1}^n (h_i^{(2)} - \mu_2)^2}}$$

siendo $H_1 = [h_1^{(1)}, h_2^{(1)}, \dots, h_n^{(1)}]$ la firma digital de la imagen a analizar; $H_2 = [h_1^{(2)}, h_2^{(2)}, \dots, h_n^{(2)}]$ la firma digital de referencia; μ_1 la media de la firma digital de la imagen a analizar y μ_2 ; la media de la firma digital de referencia.

12. Procedimiento, según la reivindicación 11, caracterizado porque se correlaciona la firma digital con una firma digital de referencia, de entre todas las evaluadas, si al menos el valor de S es mayor que 0,95.

13. Procedimiento, según la reivindicación 11, caracterizado porque se correlaciona la firma digital con una firma digital de referencia, preferentemente cuando el valor de S es mayor que 0,987.
- 5
14. Procedimiento, según cualquiera de las reivindicaciones anteriores, caracterizado porque las firmas digitales de referencia se encuentran almacenadas en una memoria estando cada una de ellas correlacionada con un servicio.
- 10
15. Procedimiento, según cualquiera de las reivindicaciones anteriores, caracterizado porque las firmas digitales de referencia se obtienen utilizando el mismo procesamiento para su obtención que las firmas digitales asociadas al contenido de la imagen.
16. Procedimiento, según cualquiera de las reivindicaciones anteriores, caracterizado porque
- 15
- comprende realizar una serie de iteraciones con una serie de imágenes a partir de una serie de capturas de pantalla de páginas web de un determinado dominio.
17. Sistema para la detección de un servicio asociado a una página web caracterizado porque comprende una memoria en la que se encuentran almacenadas una serie de firmas de
- 20
- referencia relacionadas con una serie de servicios y un procesador configurado para:
- obtener una imagen a partir de una captura de pantalla de la página web;
 - procesar la imagen para obtener una firma digital asociada al contenido de la imagen;
 - calcular la similitud entre la firma digital asociada a la imagen y una serie de firmas digitales de referencia;
- 25
- en función de la similitud calculada, correlacionar la firma digital asociada a la imagen con una determinada firma digital de referencia;
 - determinar un servicio de a la firma digital de referencia; y
 - asociar el servicio a la página web;
 - almacenar en una memoria las asociaciones determinadas entre servicios y páginas
- 30
- web.
18. Sistema, según la reivindicación 17, caracterizado porque la imagen es una imagen pre-procesada de la captura de pantalla.
- 35
19. Sistema, según la reivindicación 18, caracterizado porque la imagen es una imagen filtrada

obtenida a partir de la filtración mediante un filtro gaussiano de la captura de pantalla.

20. Sistema, según cualquiera de las reivindicaciones 18 o 19, caracterizado porque la imagen es una imagen en escala de grises asociada a la captura de pantalla.

5

21. Sistema, según cualquiera de las reivindicaciones 18 a 20, caracterizado porque la imagen es una imagen asociada a la captura de pantalla obtenida tras la aplicación de una transformada discreta del coseno.

10 22. Sistema, según cualquiera de las reivindicaciones 18 a 21, caracterizado porque la obtención de la firma digital comprende la realización, mediante el procesador, de un hashing perceptual.

15 23. Medio de almacenamiento no-transitorio con capacidad para ser leído por un procesador caracterizado porque comprende instrucciones para que el procesador ejecute un procedimiento según cualquiera de las reivindicaciones 1 a 16.

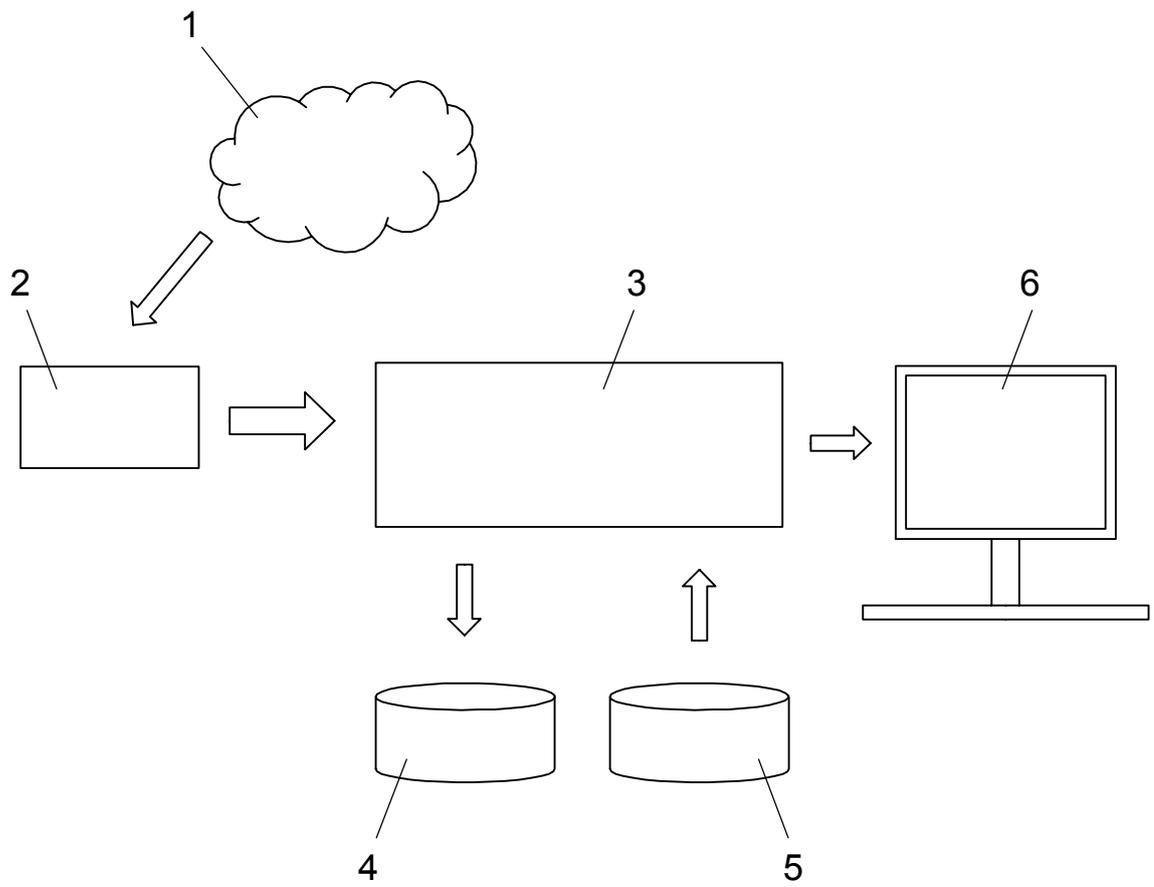


FIG. 1

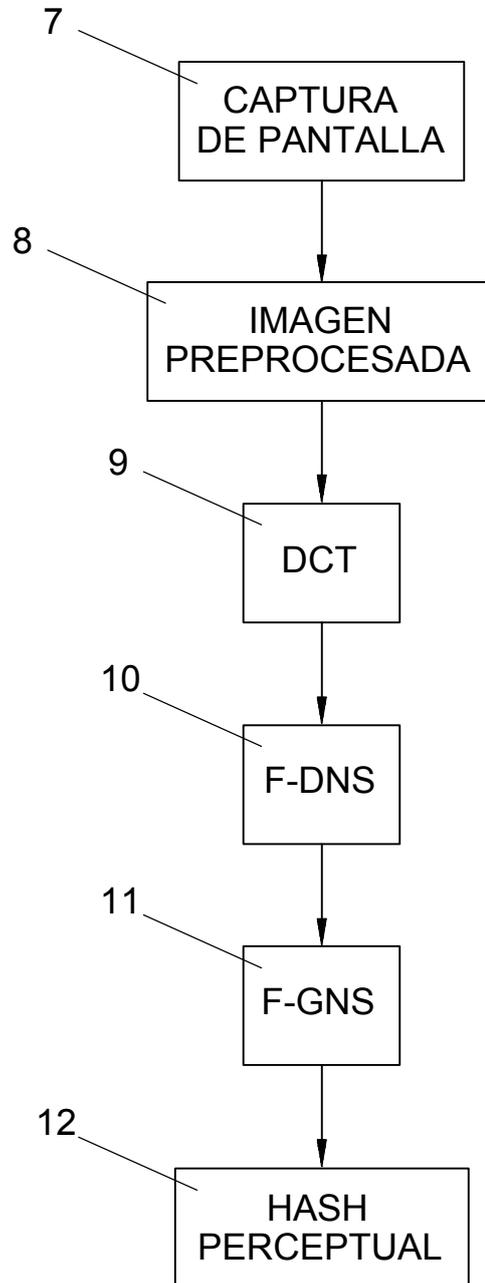


FIG. 2

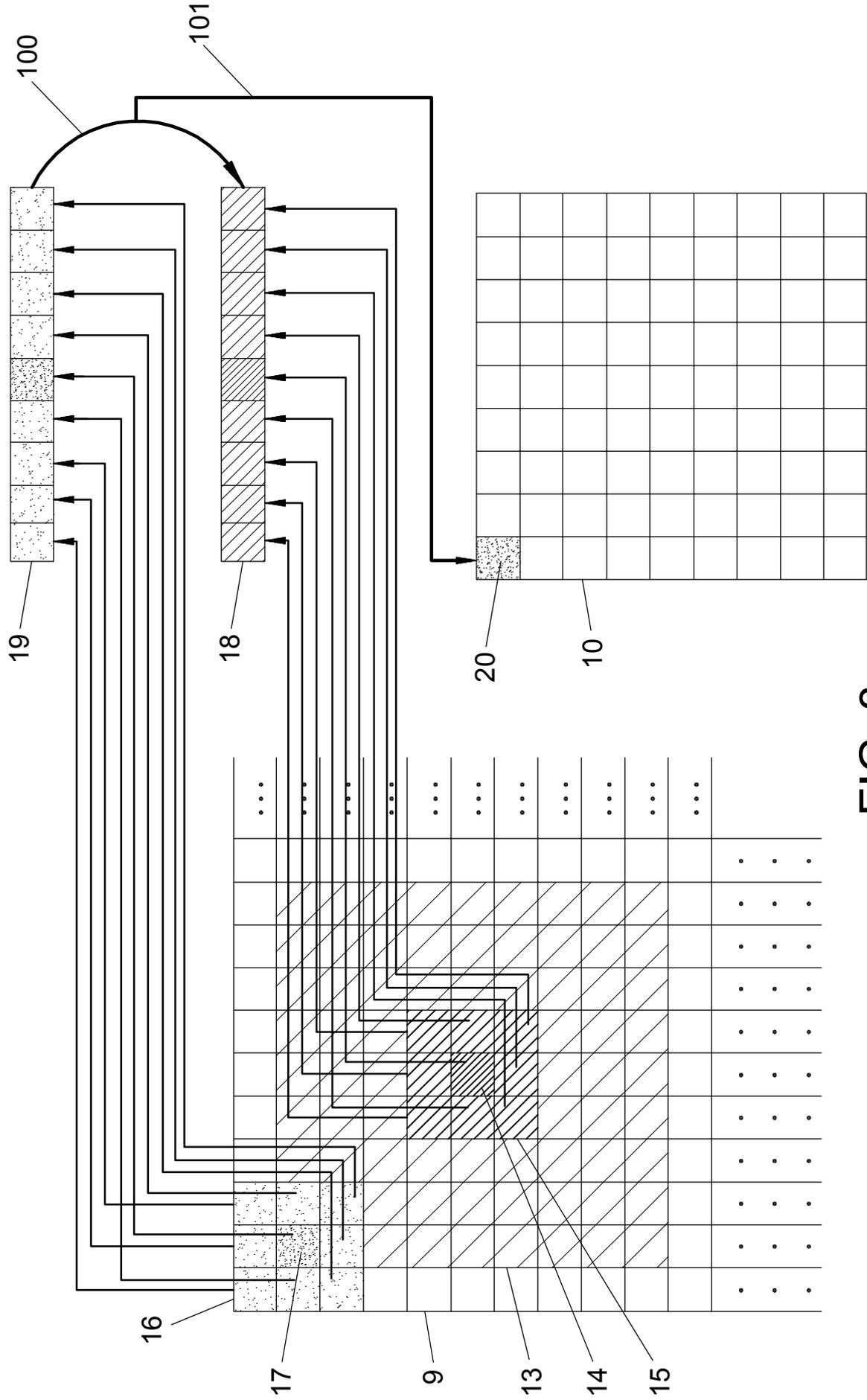


FIG. 3

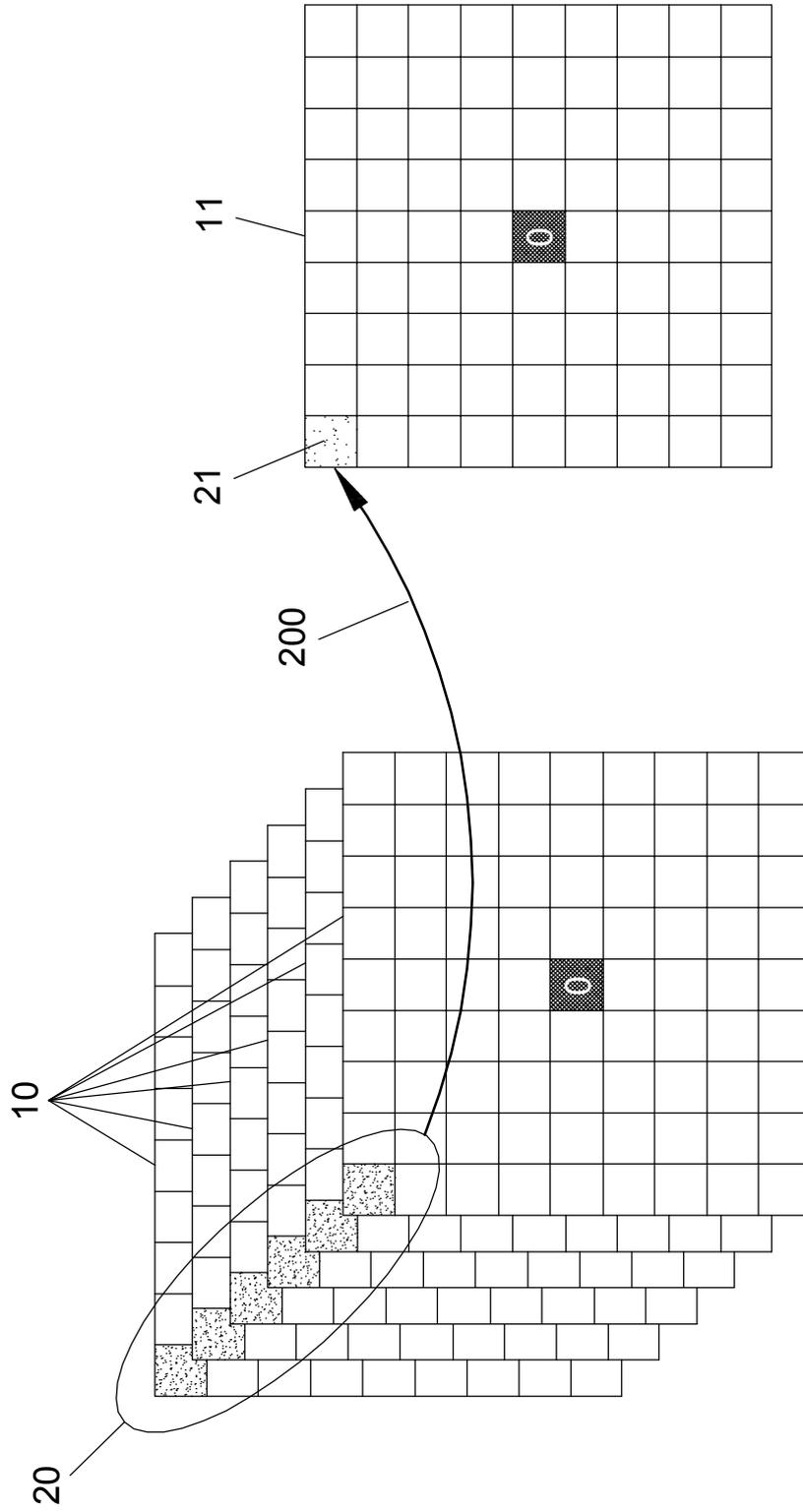


FIG. 4



- ②① N.º solicitud: 201831043
②② Fecha de presentación de la solicitud: 29.10.2018
③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: **G06K9/62** (2006.01)
G06T7/30 (2017.01)

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
X	RUBEL BISWAS, EDUARDO FIDALGO, ENRIQUE ALEGRE. Recognition of Service Domains on TOR Dark Net using Perceptual Hashing and Image Classification Techniques. 8th International Conference on Imaging for Crime Detection and Prevention ICDP-2017, 13/12/2017, Páginas 7-12 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: https://digital-library.theiet.org/content/conferences/10.1049/ic.2017.0041 >. Abstract; pág. 9, párr. 3; col. 2; párr. 1; col. 1; párr. 2; col. 2; pág. 10, párr. 1; col. 1; fig. 1;	1-23
Y	Abstract; párr. 1, col. 1; párr.2, col. 2, pág. 9; párr. 1, col. 1, pág. 10; fig. 1;	7-13
Y	FAKHRY M. KHELLAH. Texture Classification Using Dominant Neighborhood Structure. IEEE Transactions on Image Processing, 01/11/2011, Vol. 20, N° 11, Páginas 3270-3279 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5751691 >; pág. 3270, párr. 4; col. 2; pág. 3271, párr. 1; col. 1; párr. 2; col. 1; párr. 2; col. 2; pág. 3272, párr. 3; col. 2; fig. 2;	7-13
X	MARC SCHNEIDER, SHIH-FU CHANG. A Robust Content Based Digital Signature for Image Authentication. Proceedings of 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 19/09/1996, Páginas 227-230 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=560425 >; pág. 227, párr. 3; col. 2;	3-4, 19-20

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
08.08.2019

Examinador
A. Oropesa García

Página
1/3



- ②¹ N.º solicitud: 201831043
 ②² Fecha de presentación de la solicitud: 29.10.2018
 ③² Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤¹ Int. Cl.: **G06K9/62** (2006.01)
G06T7/30 (2017.01)

DOCUMENTOS RELEVANTES

Categoría	⑤ ⁶ Documentos citados	Reivindicaciones afectadas
A	Gaussian Blur. Wikipedia, 03/08/2018 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: http://web.archive.org/web/20180803202052/https://en.wikipedia.org/wiki/Gaussian_blur >; párr. [1];	3-4, 19-20
A	JAYARAMAN. Digital Image Processing. McGraw-Hill Education, 2009, 01/01/2009, Páginas 723 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: https://books.google.es/books?id=JeDGn6Wmf1kC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false >; apart. 6.3.1;	3-4, 19-20
A	Grayscale. Wikipedia, 17/08/2018 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: http://web.archive.org/web/20180817082525/https://en.wikipedia.org/wiki/Grayscale >; párr. [1]; párr. [1 - 2], apart. 2;	3-4, 19-20
A	Coeficiente de correlación de Pearson. Wikipedia, 17/04/2016 [en línea][recuperado el 08/08/2019]. Recuperado de Internet <URL: http://web.archive.org/web/20160417030416/https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson >; párr. [3];	11-13

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
08.08.2019

Examinador
A. Oropesa García

Página
2/3

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

G06K, G06T

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

INVENES, EPODOC, internet