



GOS: búsqueda visual de imágenes

Silvia Cortés Yuste

Licenciada en Comunicación Audiovisual (UAB)

Ingeniera Técnica de Telecomunicaciones, especialidad Sonido e Imagen (EUETIT –UPC)

Xavier Giró i Nieto

Grupo de Procesado de la Imagen, Teoría del Señal y Comunicaciones, UPC

Ferran Marqués Acosta

Grupo de Procesado de la Imagen, Teoría del Señal y Comunicaciones, UPC

Las imágenes televisivas utilizadas en este artículo son propiedad de TVC, Televisió de Catalunya, SA, y contienen copyright.

RESUMEN

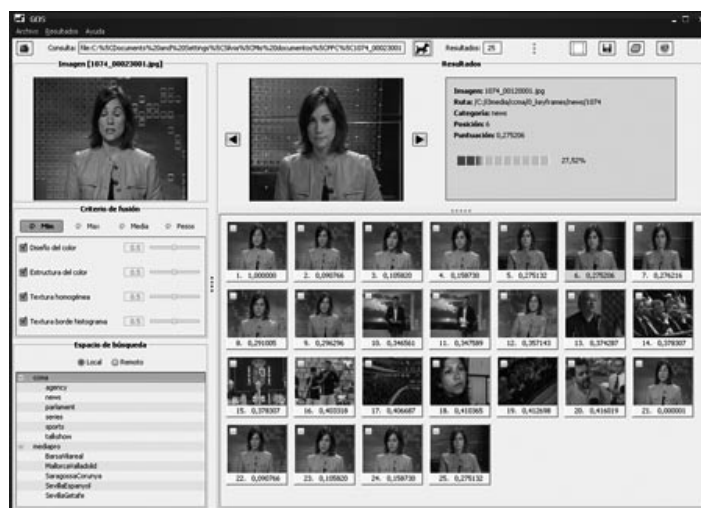
No es fácil encontrar aquello que buscamos. Los sistemas de búsqueda automatizados son máquinas que deben interpretar la información introducida por el usuario para ejecutar una petición y recuperar la información deseada. La tarea del “traductor” es esencial en el proceso, por tanto, una buena interfaz de usuario (GUI) es determinante en el éxito de la búsqueda. Os presentamos el GOS (Graphic Object Searcher), una aplicación que utiliza un sistema de búsqueda visual, para recuperar imágenes similares a otra imagen usada como ejemplo en la consulta. La industria del sector audiovisual está especialmente interesada en el desarrollo de este tipo de herramientas de gestión de contenidos que han de facilitar su trabajo diario.

INTRODUCCIÓN

Dicen que “una imagen vale más que mil palabras”. Cuántas veces hemos tratado de describir algo muy concreto que buscamos y no hemos encontrado las palabras acertadas para hacer entender a alguien qué demonios queremos. Qué fácil sería tener una fotografía, esa idea hecha imagen y mostrarla: “Esto, esto es exactamente lo que busco”. Si ese “alguien” es una máquina, todavía es más complicado.

En el siglo XXI la imagen le ha ganado el pulso a la palabra. Tanto en el ámbito profesional como en el doméstico, la generación de contenido audiovisual ha aumentado de forma vertiginosa en los últimos años gracias a la digitalización, dificultando el acceso a una información que no da tiempo a ordenar ni catalogar, y por lo tanto, no es fácil de encontrar. A los consumidores de información nos resulta imposible absorber todos estos contenidos, y se evidencia la necesidad de desarrollar nuevas técnicas y herramientas de gestión de información audiovisual, que nos permitan acceder de forma rápida y eficiente a cualquier contenido que nos interese.

Con este objetivo nace el **GOS (Graphic Object Searcher)**, una interfaz gráfica (GUI - *Graphic User Interface*) destinada a realizar búsquedas de imágenes alojadas en grandes bases de datos a partir de una imagen ejemplo y de unos criterios de búsqueda establecidos por el usuario. El GOS es una aplicación enmarcada dentro de la iniciativa **i3media** [1], proyecto dedicado a la investigación y desarrollo de tecnologías para la creación y la gestión automatizada de contenidos audiovisuales inteligentes. Como herramienta de soporte para estas tecnologías, el GOS proporciona un entorno gráfico amigable para la utilización de nuevos sistemas de indexación y recuperación de contenido audiovisual. Empresas líderes del sector media participan en el consorcio i3media con el objetivo de impulsar diversas áreas de investigación asociadas al contenido audiovisual y mejorar el papel de la industria audiovisual española en el mercado global. Para ello cuentan con la colaboración de grupos de investigadores expertos de universidades y centros tecnológicos que llevan años trabajando con éxito en este campo.



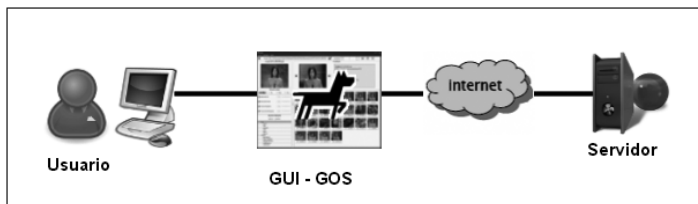
Interfaz del GOS

DIÁLOGO HOMBRE-MÁQUINA: PAPEL DE LA INTERFAZ GRÁFICA DE USUARIO

Hoy en día, personas y ordenadores están condenados a “entenderse”. Las interfaces de usuario son las herramientas clave para establecer esa comunicación entre el hombre y la máquina. El usuario debe transmitir a la máquina lo que desea conseguir, la máquina debe entender la orden y ejecutarla, y finalmente responder al usuario con el resultado del proceso o acción ejecutada.

Todo este diálogo se realiza a través de la interacción que se establece entre ambos actores y se gestiona a través de un elemento intermedio, **la interfaz de usuario (GUI)**. En cualquier sistema de recuperación de información o contenido audiovisual, este diálogo es imprescindible, y de la precisión del “traductor” depende el éxito de los resultados.

El GOS es una GUI desarrollada por el Grupo de Procesado de Imagen (GPI) de la Universidad Politécnica de Cataluña (UPC), pensada para facilitar la utilización de métodos de búsqueda inteligente de imágenes desarrollados por el grupo. La GUI se encarga de realizar la tarea de intermediario entre el usuario y el sistema de búsqueda de imágenes. Por un lado, la interfaz recoge los datos de entrada facilitados por el usuario y necesarios para realizar la consulta, y se pone en contacto con el motor de búsqueda, alojado en un servidor remoto, para que la ejecute. Por otro, una vez realizada la búsqueda,



la interfaz presenta al usuario los resultados obtenidos.

Funcionamiento del GOS

La historia de las GUI va unida indiscutiblemente a la evolución de la tecnología. Antes de la explosión tecnológica que se dio en los años 70, los ordenadores eran manejados por personas muy especializadas. El hombre se adaptaba en su totalidad a la máquina, y no al revés (la antítesis de la usabilidad). A partir de los años 80, la proliferación del ordenador personal dirigido a un usuario final no experto en informática (para uso comercial, administrativo y empresarial), supone una gran revolución. Nace la necesidad de crear herramientas que faciliten el trabajo con el ordenador: *“el usuario no quiere utilizar una aplicación, quiere*

hacer su trabajo de la forma más sencilla y rápida posible” [2].

De ahí, que la disciplina **Interacción entre Humanos y Máquinas (HCI-Human Computer Interaction)** esté adquiriendo cada vez más importancia en el desarrollo de GUI. Enmarcada dentro de las ciencias documentales, la HCI nace con la voluntad de ayudar a mejorar la comunicación entre los usuarios y los sistemas de documentación (informatizados), y se ocupa del análisis y diseño de interfaces entre hombre-máquina, estudiando la creación de productos informáticos que ayuden en la realización de tareas a sus usuarios atendiendo a la facilidad de uso, al tiempo de ejecución, a la evitación de los posibles errores y, en consecuencia, a su satisfacción [3]. Así, la HCI tiene un carácter interdisciplinar, que abarca aspectos humanos, tecnológicos y la comunicación entre ambos.

¿CÓMO BUSCAMOS?

La interfaz del GOS utiliza un sistema de búsqueda inteligente para recuperar imágenes similares a la utilizada en la consulta. Este tipo de técnicas de búsqueda pertenecen al conjunto de sistemas conocidos como sistemas CBIR (*Content-Based Image Retrieval*) y permiten recuperar imágenes digitales a partir de atributos visuales, como los colores, las formas o texturas. Dichos atributos se extraen y se representan automáticamente a través de estructuras de datos numéricas, y las imágenes son indexadas sin necesidad de tener asociada ningún tipo de anotación manual. Gracias a estos métodos de anotación e indexación automática, es posible realizar búsquedas sin necesidad de utilizar palabras, usando en su lugar paletas de colores, dibujando o seleccionando una imagen a partir de la cual el sistema pueda recuperar otras similares. La mayoría de estos sistemas utilizan esta última técnica, la **consulta mediante ejemplo (QbE - Query by Example)**, que consiste en realizar la búsqueda a partir de una imagen inicial de consulta y una serie de criterios establecidos por el usuario para encontrar imágenes parecidas en grandes bases de datos.

Los sistemas CBIR surgieron a comienzos de la década de los '90, pero no es hasta finales de esa misma década que se produce un auge en los trabajos de investigación dedicados a este ámbito. Sin embargo, la mayoría trata el tema desde el punto de vista técnico y son pocos los trabajos dedicados a aspectos como la evaluación de su eficacia, el diseño de interfaces o estudio de los usuarios [4]: “Mientras las tecnologías base de los sistemas CBIR evolucionan, desarrolladores e investigadores ignoran la importancia de la interacción entre humanos y máquinas, así como el rol crucial de la interfaz de usuario” [5]

El proyecto del GOS se centra en el diseño de una GUI para un sistema CBIR, para ampliar ese pequeño conjunto de trabajos dedicados a este tema. El método de diseño se centra en el usuario (UCD - *User-centered design*), puesto que el objetivo final de todo sistema de recuperación de imágenes es conseguir que el usuario acceda al contenido que busca. La HCI afirma que la forma de presentar la información al usuario en un sistema de recuperación de contenido hace variar su manera de interactuar [3], y por lo tanto influirá en la utilidad de la herramienta y el grado de satisfacción final que se genere. En este sentido, una aplicación de búsqueda de imágenes como el GOS, debe trabajar con minuciosidad la presentación de la consulta a realizar y la presentación de los resultados obtenidos.

Tipología de GUI para sistemas CBIR

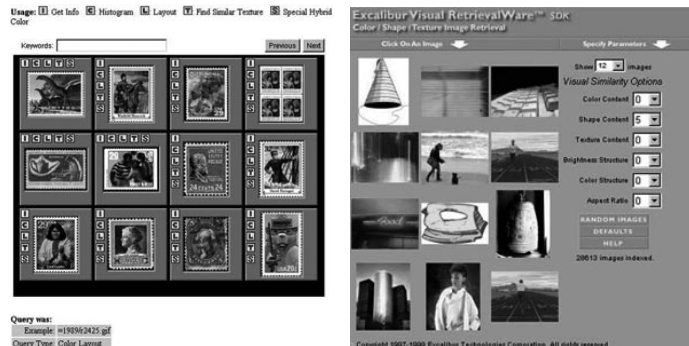
Es precisamente el tipo de usuario al que se dirige la aplicación, lo que marca las dos tendencias existentes en tipos de GUI de sistemas CBIR desarrollados hasta el momento. Por un lado, encontramos las interfaces dirigidas a los académicos e investigadores de este campo, que utilizan interfaces pobres en diseño y usabilidad, pero complejas en su manejo y en sus funcionalidades. Por otro, encontramos aplicaciones destinadas a un público más amplio, productos más comerciales enfocados a un usuario no experto, con diseños más trabajados y atractivos, en general más usables y a su vez más sencillos. Esta diferencia de enfoque en el diseño según el usuario tipo de la aplicación nos sirve para establecer una clasificación de las interfaces de sistemas CBIR, que debido a la coetaneidad de la mayoría de ellas (todas se desarrollaron en los últimos 15 años), es un criterio más interesante que el clásico análisis cronológico.

· Interfaces comerciales de sistemas CBIR

De manera genérica, vamos a aplicar el calificativo de **interfaces comerciales** a aquellas aplicaciones que están desarrolladas por empresas y destinadas a su comercialización. La mayoría de estas empresas inicialmente se preocupa de implementar los sistemas de búsqueda que posteriormente los clientes adaptarán a sus productos.

El **sistema QBIC** (1995) desarrollado por IBM fue uno de los sistemas CBIR pioneros [6]. El objetivo de IBM era vender simplemente el sistema de búsqueda, por lo que no se preocupó demasiado en diseñar una buena GUI para complementarlo. El punto fuerte de QBIC es su potente motor de búsqueda, que se sigue utilizando en múltiples aplicaciones,

como el buscador de la web del State Hermitage Museum de St. Petersburg de Russia (integrado a través de applets de Java permite realizar búsquedas sobre la colección digital del museo a través del espectro de color (*QBIC Colour Search*) o de formas geométricas (*QBIC Layout Search*) de las obras) [7]. Otras empresas siguieron el ejemplo de IBM, como Excalibur Technologies Corporation con su **Excalibur Visual RetrievalWare** (1997-1999), que desarrolló software para crear aplicaciones de manipulación de imágenes digitales y su contenido visual, extracción de características, indexación y recuperación basada en contenido. Al igual que IBM, Excalibur priorizó las técnicas de búsqueda en detrimento del diseño de una GUI de soporte para su tecnología.



Interfaces QBIC, Excalibur e ImageFinder

La compañía Attrasoft desarrolló un sistema CBIR para Windows que comercializó en tres productos diferentes (*ImageFinder*, *Internet ImageFinder* e *ImageHunt*). La tecnología utilizada era la misma para los tres, pero el diseño de sus interfaces era diferente. Actualmente solo sobrevive **ImageFinder** [8], cuya primera versión data de 1998, y hoy se comercializa la versión 7.0. ImageFinder es una herramienta compleja, no utilizada solamente para realizar búsquedas, sino que también permite el tratamiento y el procesado de imágenes, por lo que el destinatario final es un usuario experto. Así, en su GUI prima la funcionalidad por encima del diseño y la usabilidad, descuidando la estética, que se ha quedado bastante anticuada.

Actualmente una gran parte de estas aplicaciones proliferan en la web, presentadas como valor añadido a los tradicionales buscadores de imágenes que utilizan como método de búsqueda la recuperación basada en texto. Este método anota las imágenes a través del texto HTML del documento que las contiene, basándose en la suposición de que una imagen en una página web está semánticamente relacionada con el texto que la rodea. Las GUI en la web utilizan un formulario para los datos de entrada y una lista o tabla de resultados con las imágenes obtenidas como datos de salida. Las más complejas aumentan las funcionalidades de la aplicación con lenguajes de programación que permiten asociar programas o procesos ejecutables con los componentes de una página.

Google ha lanzado **Similar Images** (2009) en su Google Labs

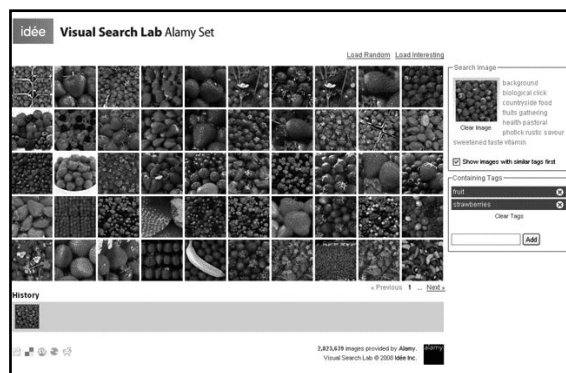


[9], como funcionalidad extra a su clásico buscador de imágenes en la web. Su aparición supone un gran paso en la incorporación de estos sistemas de búsqueda inteligentes en buscadores de imágenes convencionales, por la enorme repercusión de cualquier producto de la marca Google. La gran ventaja de esta herramienta es su fácil manejo, ya que el usuario está familiarizado con la estética y el tipo de buscador creados por Google (el diseño de la aplicación no ha sufrido ninguna variación respecto al buscador tradicional).

Sin embargo, paralelamente a *Similar Images* se están desarrollando proyectos mucho más interesantes e innovadores como **Picollator** (2008), **Idée Labs** (2008) o **GazoPa** (2008), aunque de momento, con menor repercusión.

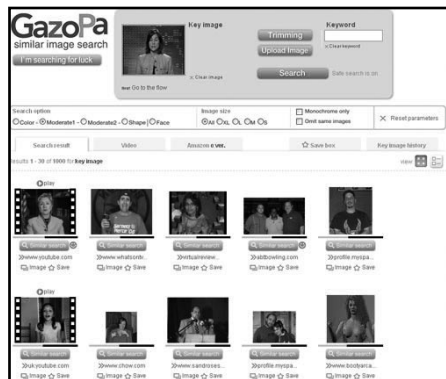
Picollator es un buscador de imágenes especializado en el reconocimiento facial, está desarrollado por la empresa de origen ruso Recognition LLC y se encuentra en versión beta en la web [10]. **Piximilar**, **PixID** y **TinEye** son productos creados por la compañía Idée Inc., empresa especializada en software de reconocimiento avanzado de imágenes y búsqueda visual [11]. *PixID* es un servicio de monitorización para identificar las imágenes que se utilizan en las publicaciones impresas y en Internet, especialmente dedicado a la detección de copias (cumplimiento de licencias, malos usos, reclamar pagos por usos no autorizados, etc.). *TinEye* es una herramienta de búsqueda de imágenes en la web

que es capaz de detectar dónde y cómo una imagen es utilizada (incluso si ha sido modificada) [11]. *Piximilar* es una herramienta especializada en búsqueda de imágenes pertenecientes a una colección. Trabaja con el método de búsqueda por imagen similar o por selección de varios colores. Idée es otra muestra del creciente interés que los sistemas CBIR están adquiriendo en el mercado. Utilizan un sistema de búsqueda propietario llamado Visual Search.



Interfaces de Picollator e Idée Labs.

GazoPa, desarrollado por la empresa Hitachi, es otra de las aplicaciones que utiliza QBIC de IBM como sistema de búsqueda inteligente. Se encuentra en fase beta, pero explota el potencial de este tipo de buscadores en el mercado y la clara tendencia actual a proporcionar herramientas cada vez más elaboradas y orientadas al consumidor de contenidos de la web (un público masivo) [12]. La GUI de GazoPa recuerda ligeramente a Google, pero con un diseño muy superior, elaborado y atractivo. Incorpora una herramienta de dibujo y recorte/retoque de imágenes en Flash, además de la posibilidad de buscar dentro del portal Amazon con un filtro de resultados que permite visualizar productos a la venta en esta tienda on-line con imágenes similares a la utilizada en la consulta.



Interfaz de Gazopa

El negocio en la web viene de la mano de compañías especializadas en proveer de contenido visual a determinados consumidores (generalmente profesionales del sector audiovisual). Estas empresas disponen de grandes colecciones de imágenes, generalmente catalogadas manualmente, y utilizan sistemas propietarios para indexar y recuperar estas imágenes a través de palabras clave (las colecciones suelen ser actualizadas periódicamente). *Getty Images*, *Corbis* o *Masterfile* [13] son algunos ejemplos, y algunas de ellas como *Masterfile*, ya incorporan en sus buscadores el método de consulta por imagen ejemplo. Pero este tipo de empresas es solo una pequeña muestra de las posibilidades de negocio de estos sistemas de búsqueda. Tiendas on-line como **Like.com** [14] son propuestas innovadoras, que apuestan por el uso de este tipo de tecnología orientado a un público masivo, el consumidor particular. El portal Like.com ha acuñado el término de *Visual Shopping* (compra visual) y se describe como “el primer motor de búsqueda realmente visual, donde el contenido de las fotos se utiliza para buscar y recuperar objetos similares”. Su GUI está completamente orientada a facilitar la compra, incluyendo fotografías de personajes famosos, donde aparecen estrellas (icono del portal) sobre la ropa que visten y que permiten con un solo clic, iniciar la búsqueda de la prenda (puesta a la venta).

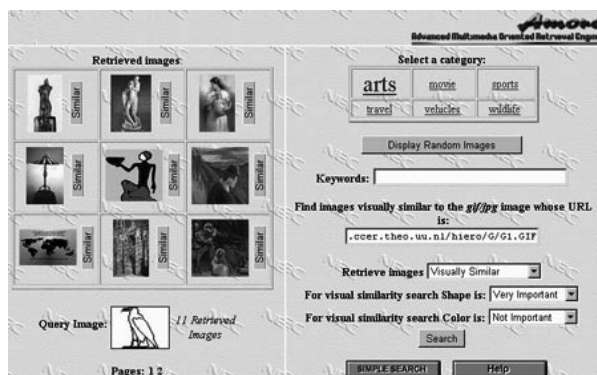
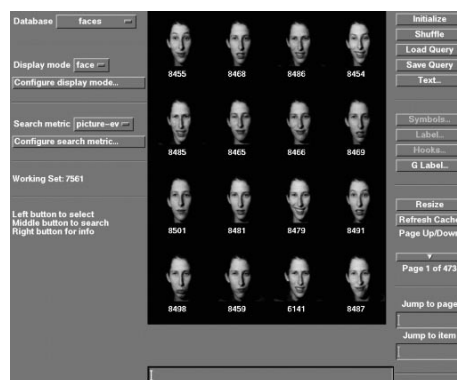


Portal Like.com

· Interfaces académicas de sistemas CBIR

En contraposición de las interfaces comerciales, otorgamos el calificativo de **interfaces académicas** a aquellas aplicaciones desarrolladas por universidades y centros de investigación, que en principio, no tienen ningún tipo de interés en su comercialización. Generalmente, las interfaces de estos sistemas no son especialmente atractivas en diseño, preocupa más su funcionalidad. Para este tipo de sistemas, la web es el banco de pruebas preferido, porque permite una mayor difusión del trabajo realizado. Además, la mayoría de sistemas CBIR académicos elaboran proyectos completos, que abarcan desde herramientas para la anotación e indexación de las imágenes, hasta las técnicas de búsqueda.

Uno de los pioneros fue **Photobook** (1994), desarrollado en el MIT Media Laboratory de Cambridge [15] como un conjunto de herramientas interactivas para la búsqueda y navegación de imágenes y de secuencias de imágenes. Photobook es el primer sistema que subraya la importancia de la interactividad entre el motor de búsqueda y el usuario para conseguir resultados satisfactorios. Su interfaz utiliza la tecnología Motif [16] para conseguir una interactividad realmente eficaz con el usuario, aunque no presta ninguna atención al diseño de la aplicación. El sistema de búsqueda de Photobook se complementa con la herramienta de anotación **FourEyes**, que utiliza el mismo tipo de interfaz.



Interfaces de Photobook y Amore

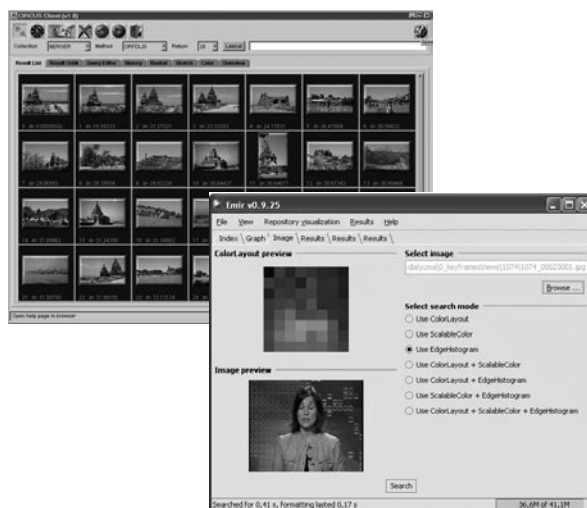
El sistema **AMORE** (*Advanced Multimedia Oriented Retrieval Engine*) (1999), desarrollado por los C & C Research Laboratories NEC USA, es uno de los primeros buscadores de imágenes para web que se preocupó por implementar un método de visualización de resultados más cuidado, tendencia que se extendería a partir del año 2000, cuando los sistemas de búsqueda de imágenes empiezan a mostrar más interés por el uso más eficiente del espacio de la pantalla [17].

Otros proyectos destacados del ámbito académico son **MARS** (Multimedia Analysis and Retrieval Systems) de las Universidades de Illinois y California (1997-2003) [18] y **CIRES** (Content Based Image REtrieval System) de la Universidad de Texas (2002-2007) [19]. MARS destaca por ser el primero en introducir técnicas de *relevance feedback* en este tipo de aplicaciones (sistemas de feedback que utilizan el resultado de una primera consulta para refinar las búsquedas con consultas posteriores) y su interfaz supone una evolución en el tratamiento visual de la información en pantalla [17]. Los desarrolladores de MARS también son los responsables de **ImageGrouper**, un sistema CBIR basado en consultas de grupos de imágenes. CIRES fue junto con MARS, uno de los sistemas CBIR de referencia por su implementación de métodos *relevance feedback* en sus interfaces, pero no ha evolucionado su propuesta, a diferencia de MARS con su **ImageGrouper**. CIRES ha sorprendido con un cambio de su GUI hacia un diseño más elaborado estéticamente pero más simple en funcionalidades (ha eliminado la funcionalidad de *relevance feedback* en su aplicación).



Interfaces de ImageGrouper y CIRES

También podemos encontrar proyectos interesantes de código abierto, que promueven la colaboración entre diversos centros de investigación e universidades. Un ejemplo es **CIRCUS** (*Content-based Image Retrieval and Consultation User System*), proyecto dedicado al desarrollo de un sistema para la recuperación de imágenes de colecciones distribuidas, heterogéneas y anotadas, que destaca por su elaborada arquitectura cliente/servidor a través del protocolo abierto de comunicación MRML (*Multimedia Retrieval Markup Language*), disponible bajo licencia pública GNU [20]. El trabajo realizado en CIRCUS (colaboración entre Laboratoire de communications audiovisuelles (LCAV), Ecole Polytechnique Fédérale de Lausanne (EPFL), y el Computer Vision Group University of Geneva) tiene su continuidad en VIPER, un proyecto para la recuperación de información multimedia que ha desarrollado el GIFT, un paquete de código abierto para implementar sistemas CBIR utilizando el método QBE (*Query-by-Example*) y muy fácil de integrar en interfaces web. Y bajo el auspicio de SourceForge.net, se ha desarrollado **Caliph & Emir** [21], un kit de herramientas en Java para la anotación y recuperación de imágenes y fotografías digitales basándose en los descriptores visuales definidos en el estándar MPEG-7. Aunque SourceForge.net es propiedad de la empresa SourceForge Inc, los proyectos que se gestionan en este sitio web no pueden ser considerados comerciales, ya que contribuyen a la difusión de nueva tecnología. Caliph & Emir utilizan la librería **LIRE** (Lucene Image REtrieval), desarrollada también por SourceForge.net dentro del proyecto Apache Lucene, dedicado al desarrollo de software de búsqueda. LIRE proporciona un sistema para crear índices de imágenes, realizar búsquedas, navegar por estos índices y crear mosaicos de imágenes para su visualización. Todo es código abierto.



Interfaces de CIRCUS y Emir

GOS

El GOS quiere iniciar una nueva tendencia en la tipología de las GUI de sistemas CBIR, un nuevo tipo de GUI donde convergen características de interfaz comercial e interfaz académica, con aplicaciones destinadas a dar servicio tanto a expertos de la materia como a clientes (empresas del sector audiovisual o consumidor particular).

Empresas como **Mediapro S.L.**, líder del proyecto i3media y dedicada a la producción de contenidos y provisión de servicios media, y la **Corporació Catalana de Mitjans Audiovisuals (CCMA)**, empresa de desarrollo de tecnologías de gestión de contenidos, son los principales destinatarios del GOS. La herramienta permite realizar búsquedas con imágenes de ejemplo en sus archivos de vídeo para crear nuevos contenidos, agilizando el proceso de documentación y recuperación de contenido, al visualizar rápidamente los resultados de la búsqueda. A su vez, el GOS también sirve de soporte a la UPC para experimentar y avanzar en el desarrollo de técnicas de gestión y recuperación de contenido audiovisual.

¿Qué puede hacer el GOS? El GOS es la GUI de un buscador de imágenes basada en ejemplos, y como tal, permite formular una consulta, ejecuta la búsqueda y visualiza los resultados. Éstas son las funciones básicas que todo buscador implementa. Sin embargo, el GOS pretende aportar un valor añadido que diferencie la aplicación del resto de productos disponibles hasta el momento. Con el objetivo de facilitar el trabajo de búsqueda y recuperación de imágenes al usuario final, el diseño de la GUI ha cuidado enormemente los tres aspectos fundamentales que dan forma a una aplicación: el diseño visual o gráfico, la usabilidad y la tecnología. La mayor dificultad es aunar funcionalidad y estética, dotar a los elementos gráficos de una “personalidad” propia, un *look&feel* característico de la aplicación, pero sin perder de vista las necesidades reales del usuario, y paralelamente ir sorteando las limitaciones de la tecnología.

Una buena distribución de los elementos en la pantalla es la base para una utilización óptima de la herramienta. ¿De qué nos sirve tener potentes funcionalidades disponibles en la GUI, pero no fáciles de utilizar o de localizar? Como requisito imprescindible, un buscador de imágenes mediante consulta por ejemplos debe diferenciar siempre dos zonas en pantalla, la zona destinada a la formulación de la consulta y la zona destinada a la visualización de los resultados, creando un flujo de acciones coherentes y lógicos en su ejecución.



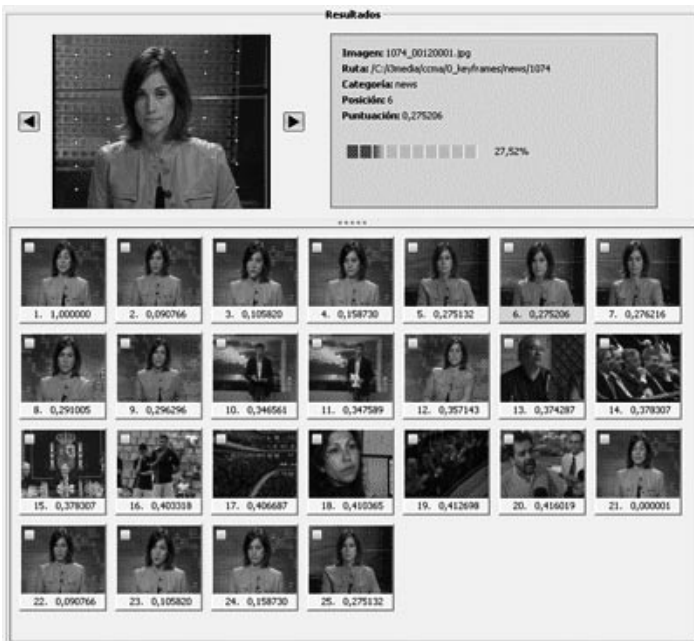
Organización de elementos en pantalla del GOS

En el GOS, se ha destinado la zona superior de la aplicación al menú principal y la barra de herramientas (zona roja de la figura). De esta manera, todas las funcionalidades son accesibles al usuario de forma rápida. En la zona izquierda ubicamos el área de consulta (zona amarilla de la figura), donde el usuario podrá configurar la petición de búsqueda. Siguiendo los pasos verticalmente, la GUI permite cargar la imagen de consulta, seleccionar el criterio de fusión y los descriptores visuales a tener en cuenta en la búsqueda, y el ámbito de la búsqueda, escogiendo la categoría dónde buscar. El resto de pantalla (zona verde de la figura) se utiliza para la visualización de los resultados, que vuelve a reproducir el esquema anterior: en la parte superior del área de resultados se coloca una zona destinada a la presentación de información destacada (una imagen resultado junto a un área de texto con la información asociada a esta imagen), y en la parte inferior se visualizan todos los resultados de la búsqueda en forma de parrilla de imágenes, ordenados según el grado de similitud con la imagen consulta (en orden decreciente).

Además, la aplicación dispone de un método para refinar las búsquedas, ya que cualquier imagen resultado puede ser utilizada para lanzar una nueva consulta haciendo simplemente doble clic sobre ella, y permite navegar cómodamente a través de la lista de resultados obtenidos. Para obtener más información sobre una imagen resultado, el usuario puede darle foco haciendo clic sobre la diapositiva. Cuando una diapositiva obtiene el foco, su fondo cambia de color, de blanco a azul, y su borde de gris a un azul más oscuro, y automáticamente pasa a ocupar el lugar de la imagen destacada del panel superior. La interfaz dispone de dos métodos alternativos

CONCLUSIONES

para navegar por los resultados y visualizarlos como imágenes destacadas. Por un lado, encontramos dos típicos botones de navegación a la izquierda y derecha de la imagen destacada, para recorrer los resultados (anterior y posterior respectivamente) con un solo clic, sin necesidad de desplazarse a la zona de la parrilla para clicar sobre la imagen deseada. Por otro, a través de la rueda central del ratón, el usuario puede desplazarse por los resultados de forma rápida y sin necesidad de ir realizando clics.



Visualización de los resultados en el GOS

El GOS también permite al usuario seleccionar aquellos resultados que sean de su interés y guardarlos en un archivo XML. Esta selección puede realizarse de modo individual, haciendo clic sobre el checkbox de la diapositiva o en grupo, a través de la herramienta de selección rectangular.

La elección de Java para desarrollar la aplicación, nos permite crear una herramienta muy versátil, con más posibilidades de evolución e integración con el resto de aplicaciones y herramientas, de la UPC y del proyecto i3media. Además, la utilización del estándar MPEG-7 en ficheros XML favorece la interoperabilidad, entre las diferentes piezas que componen el sistema, y la manipulación de la información visual que contienen las imágenes.

El GOS está disponible en línea a través del software de Java Web Start, que permite la descarga y ejecución de la aplicación desde su web [22], para jugar un poco con él y descubrir sus posibilidades.

A juzgar por el resultado final, podemos avanzar que el GOS supone un paso más en la evolución de las GUI para sistemas CBIR. La interfaz es fácil de utilizar, bien estructurada, pensada para simplificar el trabajo de recuperación de imágenes al usuario, y a la vez incorpora técnicas de búsqueda y recuperación de imágenes pioneras. Como valor añadido, la aplicación permite su ejecución en entorno remoto y local, aunque el objetivo final es acabar ofreciendo el servicio de búsqueda de manera remota, siguiendo la tendencia de la gran mayoría de proveedores de contenido media. A partir de servicios web, cualquier usuario en red puede tener acceso a múltiples aplicaciones, de forma cómoda y sencilla. El enorme volumen de información disponible, especialmente en la industria audiovisual, precisa de herramientas que se adapten a este tipo de arquitecturas, modulares e integrables.

Pero esto es solo el principio. El GOS se encuentra en una primera versión beta, y su desarrollo continua dentro del proyecto i3media, para ir incorporando nuevas técnicas de recuperación de contenido audiovisual que trabaja la UPC. El siguiente paso que se plantea el GOS es la ampliación de los métodos de búsqueda que soporta la interfaz. Esta ampliación deberá realizarse en dos fases: la primera, y más inmediata, es la implementación de la búsqueda de imágenes basada en regiones, y la segunda, dar la posibilidad de utilizar y combinar más de un tipo de consulta (basada en texto, imágenes y regiones). Será interesante seguirle el rastro. Podéis seguir la evolución del proyecto en el blog: <http://bitsearch.blogspot.com/search/label/GOS>.

REFERENCIAS

PFC: Cortés Yuste, Silvia. **Interfaz Gráfica de Usuario para la Búsqueda de Imágenes basada en Imágenes**. EUETIT-UPC, Junio-2009.

(http://gps-tsc.upc.es/imatge/_Xgiro/teaching/thesis/2008-2009/SilviaCortes/memoria.pdf)

[1] Web i3media: <http://www.i3media.org>

[2] Roe, Benjamin. **Diseño de interfaces de usuario usables**. Publicado en Mundo Geek. Diciembre 2004

(<http://mundogeek.net/traducciones/interfaces-usuario-usables/gui.html>)

[3] Marcos, Mari-Carmen. “**HCI (Human computer interaction): concepto y desarrollo**”. En: El profesional de la información, 2001, junio, v. 10, n. 6, pp. 4-16.

(<http://www.elprofesionaldelainformacion.com/contenidos/2001/junio/1.pdf>)

[4] Pérez Álvarez, Sara. **Análisis de usabilidad de sistemas CBIR | User friendliness of CBIR systems analysis** -. Abril 2008

(<http://biblioteca.universia.net/ficha.do?id=34103694>)

[5] Colin C. Venters. **User Interface Design & Evaluation for a Content-Based Image Retrieval System**. Department of Information & Library Management, University of Northumbria at Newcastle, England.

(<http://www.bcs-hci.org.uk/hci1998/C51/>)

[6] Sistema QBIC: <http://www.qbic.almaden.ibm.com/>.

[7] State Hermitage Museum: <http://www.hermitagemuseum.org/cgi-bin/db2www/qbicSearch.mac/qbic?selLang=English>.

[8] ImageFinder de Attrasoft: <http://www.attrasoft.com/imagefinder70/>

[9] Similar Images de Google: <http://similar-images.googlelabs.com/>

[10] Picollator: <http://www.picollator.com/>.

[11] Labs Idée: <http://labs.ideeinc.com/>; Tineye: <http://tineye.com/>

[12] Gazopa de Hitachi: <http://www.gazopa.com>

[13] Getty Images (<http://www.gettyimages.com>); Corbis (<http://pro.corbis.com/>); Masterfile (<http://www.masterfile.com/info/products/simsearch.html>)

[14] Like.com: <http://www.like.com/aboutus.py>

[15] Photobook: <http://vismod.media.mit.edu/vismod/demos/photobook/index.html>

[16] Pentland, R.W.; Picard, S.; Sclaroff, S. **Photobook: Content-Based Manipulation of Image Databases**. The Media Laboratory, Massachusetts Institute of Technology. June 1995

[17] Nakazato, Munehiro and Manola, Ljubomir and Huang, Thomas S. (2003) **ImageGrouper: a group-oriented user interface for content-based image retrieval and digital image arrangement**. Journal of Visual Languages & Computing, 14 (4). pp. 363-386. ISSN 1045926X

[18] MARS: <http://www.ifp.illinois.edu/~qitian/MARS.html>

[19] CIRES: <http://cires.matthewwiley.com/>

[20] CIRCUS: http://vipser.unige.ch/doku.php/demos#content-based_image_retrieval

[21] Caliph&Emir de Sourceforge.net: <http://sourceforge.net/projects/caliph-emir/>, <http://caliph-emir.sf.net>

[22] Web del GOS: <http://gps-tsc.upc.es/imatge/i3media/gos/>

AUTORES



Silvia Cortés Yuste, es licenciada en Comunicación Audiovisual e Ingeniera Técnica de Telecomunicaciones, especialidad Sonido e Imagen. Trabaja en proyectos relacionados con el desarrollo de herramientas de gestión de información y de nuevas aplicaciones y servicios para contenido audiovisual en diferentes medios (televisión, IPTV, web), en materia de diseño, usabilidad y desarrollo de software.



Xavier Giró i Nieto, investigador del Grupo de Procesado de la Imagen (GPI) y profesor del grado en Ingeniería de Sistemas Audiovisuales en la Escuela de Ingeniería de Terrassa de la UPC. Es titulado en Ingeniería de Telecomunicaciones por la UPC en 2000 y sigue desde el 2002 el programa de Doctorado del Departamento de Teoría del Señal y Comunicaciones. Ha trabajado sobre codificación de imágenes volumétricas en la Vrije Universiteit de Bruselas (VUB), televisión digital en el Sony Development Center Brussels, y indexación automática de imágenes en el GPI y en la Columbia University de Nueva York. Su investigación puede seguirse en el blog <http://bitsearch.blogspot.com>.



Ferran Marqués Acosta, realiza su investigación en el Grupo de Procesado de Imagen (GPI) del Departamento de Teoría de la Señal y Comunicaciones (TSC) y su docencia en la Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona (Telecom BCN).