

# Tendencias actuales y perspectivas futuras en la recuperación de información gráfica

**Jesús Robledano Arillo**

Universidad Carlos III de Madrid

## 0.1. Resumen

Para la explotación intelectual del cada vez más rico patrimonio gráfico se requiere el desarrollo de modelos de recuperación de documentos gráficos adaptados a la naturaleza formal y semántica de la imagen. Sin embargo, se han venido aplicando de forma masiva modelos de recuperación ideados para documentos texto-lingüísticos, no del todo adecuados a las necesidades de los documentos basados en imágenes. Se distinguen cuatro modelos básicos de recuperación de imagen, concluyendo cuál es el modelo que más se ajusta a los requerimientos de la información gráfica y las implicaciones que ese modelo tiene para la metodología documental aplicada a los fondos documentales gráficos (Autor).

**Palabras clave:** Recuperación de Información. Documentos gráficos.

## 0.2. Abstract

Intellectual exploitation of the increasingly rich graphic heritage requires the implementation of models for graphic information retrieval which are adapted to the formal and semantic nature of the image. However retrieval models aimed at text documents have been widely used, despite their unsuitability to the needs of image-based documents. Four basic models for image retrieval will be stated in this paper, in order to conclude which one meets the requirements of graphic information more suitably. Finally, this model's implications for Information Science methodology applied to image banks will be discussed (Author).

**Keywords:** Information Retrieval. Images.

## 1. Introducción

Un hecho evidente en nuestros días es el crecimiento cada vez más acusado del volumen de información gráfica presente en los centros especializados en la custodia, gestión y difusión de material gráfico. Este mayor volumen no sólo es

explicable por la lógica acumulación progresiva de documentos con el paso del tiempo, o por la falta de unos criterios de selección documental que permitan mantener un volumen equilibrado de documentos, sino que parece obedecer a la necesidad de adaptación de estos centros a una demanda de información gráfica en continuo crecimiento. La imagen cobra cada día mayor importancia como medio de expresión en nuestra sociedad, lo cual provoca un aumento continuo del patrimonio gráfico generado por ésta. Una parte importante de este patrimonio es mantenido en centros especializados —archivos, bancos de imágenes, bibliotecas, museos— de cara a su reutilización. La necesidad de reutilización de un fondo voluminoso y con una tasa de crecimiento acelerada y continua lleva a la búsqueda de sistemas de gestión y recuperación cada vez más eficaces. En esta línea hay que enmarcar la tendencia hacia la digitalización de los sistemas de información empleados para la gestión documental de los fondos gráficos que observamos hoy día. El número de bancos y archivos de imágenes de instituciones, empresas y particulares digitalizados o en proceso de digitalización crece día a día. La evidencia de las mayores ventajas de cara a la gestión, recuperación y difusión que ofrecen los Sistemas de Archivo Digital; el que estas tecnologías sean cada vez más asequibles; y la mayor demanda de imágenes en formato digital son factores que están acelerando este proceso. Los sistemas de información gráficos deben superar, pues, una etapa fundamental para adaptarse a las nuevas circunstancias: la implantación de un sistema de imagen digital que permita una recuperación eficaz y eficiente en un fondo extenso y en continuo crecimiento. A lo largo de este trabajo vamos a plantear cuales son las tendencias tecnológicas actuales y futuras que se están barajando de cara a la resolución de esta cuestión. El objetivo será definir el modelo de recuperación de imágenes más adecuado dentro de las posibilidades técnicas disponibles actualmente o previsibles para un futuro inmediato, valorando el impacto de este modelo respecto a las técnicas documentales.

## **2. Premisas teóricas de partida: las necesidades en la recuperación de información gráfica**

### **2.1. La cuestión tecnológica**

El mayor peso cuantitativo y cualitativo de la información representada a través del texto lingüístico en la memoria colectiva de la sociedad contemporánea se ha reflejado en los métodos y herramientas de almacenamiento, gestión y difusión desarrollados para posibilitar el uso y difusión de todo tipo de información, no sólo la textual. En la evolución de las técnicas y herramientas de análisis y recuperación de información se puede observar un desfase entre los sistemas basados en texto lingüístico y los sistemas de imagen, en detrimento de estos últi-

mos. El crecimiento del volumen de información en formato texto lingüístico se ha ido viendo compensado en las últimas décadas por el desarrollo de productos de recuperación de información adecuados a las necesidades que plantea la recuperación de este tipo de objetos. De este modo se desarrollaron y aplicaron los primeros modelos de gestión de bases de datos, que dieron paso pronto al desarrollo de Sistemas de Recuperación de Información (IRS). La potencia que observamos en los sistemas de recuperación y difusión de documentos texto-lingüísticos está aún a años luz de las posibilidades que presentan los sistemas de recuperación de información gráfica.

Este hecho no es explicable únicamente por la mayor dificultad técnica y consiguiente aumento del coste que plantea la gestión informática de imágenes. Los avances en tecnología de imagen digital y el creciente abaratamiento de la potencia de proceso y de los dispositivos de almacenamiento masivo de la microinformática están difuminando las cada vez menores dificultades técnicas que representa el tratamiento de la imagen digital respecto del texto. Si la cuestión tecnológica parece representar un factor de escasa incidencia en el retraso del desarrollo de sistemas de recuperación adecuados a este tipo de información, habrá que situar las causas de este desajuste más cerca de planteamientos documentales erróneos o poco adecuados que de cuestiones relacionadas con la tecnología informática. Como ha señalado Enser (1995, p. 151), el avance en tecnología de procesamiento de imagen digital no se está viendo acompañado por un avance similar en cuanto a técnicas para el diseño de Sistemas de Información capaces de recuperar eficazmente imágenes.

A diferencia de lo que ocurre con el material documental basado en texto, no podemos hablar del desarrollo de unas bases documentales teóricas y técnicas lo suficientemente sólidas como para permitir una recuperación y difusión pública eficaz. Este factor está detrás de las dificultades con que nos enfrentamos a la hora de obtener un acceso eficaz y eficiente al cada vez más rico patrimonio gráfico disponible en archivos, bibliotecas y bancos de imágenes. Como aparece reflejado en el epígrafe tercero, los modelos de representación y recuperación que se han venido aplicando de forma generalizada a los sistemas de imagen están inspirados en los sistemas que se han venido aplicando a fondos documentales compuestos por documentos que utilizan el texto lingüístico como medio de representación de la información que contienen: libros, artículos de revista, documentos administrativos, etc. Aunque las necesidades funcionales a las que deben enfrentarse los sistemas de imagen no deben diferir en mucho de las de otros tipos de sistemas basados en información texto-lingüística, las bases metodológicas sobre las que debe asentarse el sistema de recuperación han de ser necesariamente distintas, debido a las diferencias que presentan ambos tipos de documentos en cuanto a naturaleza semántica y técnica, y en cuanto a uso.

## **2.2. Implicaciones de la naturaleza de la información gráfica para el diseño de Sistemas de Recuperación**

Uno de los principios básicos para el diseño de un Sistema de Información coherente es el siguiente axioma: el modelo de recuperación debe partir de la comprensión de la naturaleza semántica de los documentos que van a ser representados en ese sistema y de las necesidades de información de sus usuarios. La naturaleza semántica nos orientará sobre el modelo de representación de información adecuado para la recuperación de ese tipo de documentos. Las necesidades de los usuarios nos ayudarán a seleccionar de entre el total de atributos semánticos aquellos que es necesario representar, así como el nivel de exhaustividad requerido en la representación de los atributos seleccionados. Esta forma de actuación permitirá una recuperación viable, es decir, eficaz y eficiente.

En cuanto a las necesidades de usuarios, la bibliografía ofrece buenos ejemplos de estudios de modelos de peticiones en distintos tipos de fondos gráficos (Enser 1995, p. 153; Romer 1995; Nail, 1994, p. 86). La mayor parte de estos estudios concluyen resaltando el hecho de que la casi totalidad de atributos semánticos presentes en las imágenes pueden llegar a convertirse en atributos útiles para la recuperación de éstas en un momento dado. A pesar de que unos atributos cobran mayor importancia que otros en función del tipo de fondo y de usuarios de ese fondo, el acceso a las imágenes por su contenido parece ser el atributo principal para la mayor parte de los tipos de fondos gráficos. Los fondos documentales parecen enfrentarse cada vez más a la satisfacción de demandas de imágenes por contenido: denotado a nivel formal (“quiero imágenes con el aspecto de..., con figuras como..., con predominio de los colores...”), icónico (“quiero imágenes de...”) o conceptual (“quiero imágenes sobre... “); y connotado (“imágenes que sugieran...”, “imágenes que den la sensación de....”). A continuación veremos como esta tipología de demandas está íntimamente relacionada con la naturaleza semántica de la imagen gráfica. De esta tipología de demandas podemos obtener un abanico de características semánticas de la imagen que van desde aspectos puramente formales hasta aspectos conceptuales de gran complejidad. Vamos a definir a continuación estas características, relacionándolas con los sistemas de representación y recuperación más adecuados en función de su naturaleza.

### *2.2.1. Modelo semántico de la imagen*

Aunque pensado exclusivamente para el análisis iconológico de obras de arte, el sistema de análisis iconográfico de Panofsky (1976, p. 13) es útil para el diseño de un modelo estructurado de análisis del contenido semántico de una imagen. Para la identificación de los niveles de análisis que aparecen a continuación nos hemos servido de dicho sistema, aunque previamente adaptado para

mejorar su adecuación a un modelo de representación documental. Hemos distinguido tres niveles semánticos, ordenados en orden creciente de abstracción desde lo puramente objetivo hasta lo puramente subjetivo:

#### 2.2.1.1. Primer nivel: nivel formal

Incluye los aspectos formales del contenido, es decir, el artificio técnico utilizado para representar la imagen que deriva en la forma del contenido. A partir de estos atributos los seres humanos somos capaces de abstraer representaciones gráficas de elementos de la vida real o imaginaria. Podemos reducir los atributos formales fundamentales de cara a la recuperación de imágenes a cuatro entidades: colores, texturas, figuras y relaciones espaciales entre objetos.

Dentro de este nivel podemos aislar un aspecto puramente subjetivo, son las sensaciones estéticas provocadas por la contemplación de los atributos formales de la imagen. A modo de ejemplo podemos citar las sensaciones de belleza, fealdad, armonía, tranquilidad o desasosiego que cualquier sujeto puede sentir mediante la contemplación de este tipo de atributos. Pensemos en las sensaciones que nos pueden llegar a provocar las imágenes no figurativas.

Una vez identificado este nivel como atributo útil para la recuperación en buena parte de los tipos de fondos gráficos, debemos reflexionar sobre las posibilidades de los sistemas de representación que tenemos a nuestro alcance, ya que sólo seremos capaces de recuperar por estos atributos en fondos gráficos voluminosos si somos capaces de representarlos. El sistema de representación lingüístico es el que se ha venido utilizando normalmente en aquellos fondos en los que la recuperación por este tipo de atributos es importante. La representación con texto lingüístico choca con los problemas de trasladar un sistema de expresión a otro. Intentemos expresar con palabras los colores presentes en un fondo gráfico, los tipos de texturas, las formas, las relaciones formales entre estos elementos, o las sensaciones estéticas susceptibles de ser provocadas al contemplar las imágenes. Pronto comprobaremos que es una tarea ardua y abocada al fracaso, pues chocaremos continuamente con lo innombrable, con lo que sólo es comunicable y asimilable visualmente. Sólo de manera muy limitada y después de un esfuerzo considerable será posible representar textualmente estos atributos. El nivel formal requiere, en consecuencia, sistemas de recuperación no basados en texto lingüístico sino en imágenes. Es imposible representar lingüísticamente de manera eficiente las texturas, figuras y colores y relaciones espaciales. La generación de un léxico a este nivel parece algo inasequible.

## 2. Segundo nivel: pre-iconográfico e iconográfico

Incluye el reconocimiento de elementos, acontecimientos y conceptos de la vida real o imaginaria representados en la imagen. Correspondería a los niveles pre-iconográfico e iconográfico del sistema de Panofsky. Para extraer información relevante para la recuperación de imágenes en este nivel podemos estructurar los atributos semánticos según un modelo facetado basado en los famosos interrogantes del paradigma de Lasswell : ¿ Quién ?, ¿ Qué ?, ¿ Donde ?, ¿ Cuándo ?, ¿ Cómo ?, ¿ Por qué ?. Debemos distinguir dos subniveles, uno puramente iconográfico, el contenido visible de la imagen (iconos) y otro conceptual (conceptos representados en la imagen). Dentro del primer subnivel distinguimos dos estadios, un nivel genérico y otro específico, ya que la imagen de un objeto es útil tanto para consultas por la clase del objeto (“quiero ver imágenes de montañas nevadas”) como por ese objeto concreto (“quiero ver imágenes del Everest”). En este primer subnivel la representación será susceptible de contener nombres comunes y propios y verbos que describan lo representado: objetos, seres animados, acciones, emociones, actitudes, aspecto, pose... En el subnivel conceptual entran aquellos conceptos denotados en la imagen, también en dos estadios, genérico y específico. Por ejemplo, una imagen puede ser la representación de una escena bélica cualquiera: guerra, y a su vez la representación de una escena bélica de la Guerra Civil Española.

En la representación de este segundo nivel no sólo es factible, sino imprescindible el empleo de términos del lenguaje que representen conceptos generales del conocimiento humano y nombres propios de acontecimientos, estilos, períodos, etc. Esto es debido a que este tipo de conocimiento se expresa y comunica a través del lenguaje verbal o escrito. Como señala Alonso Erausquín con respecto a la imagen fotográfica, pero perfectamente válido para cualquier imagen figurativa: “La elocuencia autónoma de la información fotográfica [...] queda condicionada a que conozcamos, a través de la palabra, una serie de circunstancias que la imagen no comunica directamente” (Alonso Erausquín, 1995, p.11). La imagen gráfica por sí sola es limitada para la concreción de la información de este nivel, por lo que la única manera de representar estos atributos semánticos con la finalidad de permitir la recuperación es a través de texto lingüístico. Sin embargo un modelo de representación lingüístico en el segundo nivel válido para la imagen debe adecuarse a la naturaleza semántica de ésta, recogiendo los atributos que señalamos en el primer párrafo de este epígrafe.

## 3. Tercer nivel : iconológico

Situamos aquí una interpretación sobre el significado de la imagen. Si la interpretación es subjetiva entra dentro del campo de la connotación: sugerencias, sensaciones y emociones personales obtenidas de la interrelación del conte-

nido de la imagen y su forma de expresión. La utilización del texto para la representación de este nivel es adecuada para la recuperación de conceptos connotados expresables lingüísticamente, por ejemplo : tristeza, soledad, angustia ; aunque difícilmente podremos recuperar por connotación en aquellos casos donde la forma de expresión de un contenido juegue un papel clave para la obtención de una sugerencia, emoción o sensación. En esos casos es ineludible la combinación de recuperación basada en texto lingüístico y en imágenes.

### *2.2.2. Implicaciones del modelo semántico:*

Podemos resumir las consecuencias del modelo semántico definido en el epígrafe anterior en los siguientes puntos :

- La riqueza semántica de la imagen es muy alta y depende de la interrelación de atributos semánticos de los tres niveles estudiados.
- Para posibilitar una recuperación eficaz y eficiente por contenido en fondos gráficos voluminosos es necesario representar esos atributos semánticos.
- El texto lingüístico es adecuado para la representación de una importante parte de los atributos conceptuales de una imagen, siempre y cuando la representación se estructure de una manera adecuada a la naturaleza de la imagen.
- El texto no se adapta a las necesidades de recuperación por el nivel formal. Es imposible trasladar un sistema de representación gráfico a un sistema de representación lingüístico y viceversa sin perder la esencia de la información.
- El uso del texto como único sistema de representación limita la capacidad de recuperación de imágenes, por lo que debemos emplear otras alternativas que permitan el uso conjunto e interrelacionado de dos tipos de recuperación: una recuperación visual y una recuperación conceptual.

## **3. Modelos actuales de Sistemas de Recuperación de Información Gráfica**

Vamos a ver a continuación cómo se han adaptado los principales modelos de recuperación implantados hasta la fecha en cuanto a los requerimientos que plantea la recuperación de imágenes. A lo largo de los siguientes epígrafes constataremos como los modelos más usados han sido aquellos pensados para documentos textuales pero no para la imagen. Podemos seguir una línea que marca distintas etapas que no necesariamente siguen una secuencia cronológica.

### **3.1. Los sistemas manuales con orientación clasificatoria**

Reproducen el sistema de organización intelectual en el sistema de instalación física de los documentos. Se basan en clasificaciones —temáticas, onomásticas o una mezcla de ambas— de tipo jerárquico decimal. En función de la existencia o no de lenguajes documentales clasificatorios, los procesos de recuperación se realizan sobre los índices de materias y a continuación sobre el depósito, o directamente sobre el propio depósito de documentos, que reproduce en la disposición de los documentos una estructura organizativa basada en la ordenación temática, similar a la disposición de los libros en el depósito de una biblioteca ordenada con la CDU. Su origen se remonta a los primeros intentos de organización de fondos gráficos por parte de instituciones encargadas de su custodia. Estos intentos de organización se basan en la aplicación de las técnicas bibliotecarias a la gestión de documentos gráficos. Aún es muy frecuente encontrar estos sistemas, especialmente en archivos fotográficos no digitalizados. Además de la limitación en cuanto a posibilidades de recuperación por contenido temático de los sistemas manuales — muy restringidas al asignar un único índice clasificatorio a cada imagen, a no ser que se duplique cada imagen tantas veces como índices se le asignen —, el principal inconveniente es que fuerzan y limitan al usuario a la lógica de recuperación y de interpretación de los documentos que subyace en la clasificación temática utilizada. Estos sistemas permiten cierto grado de consulta visual a través del ojeo de las imágenes de cada carpeta, pero previamente se ha tenido que preseleccionar la carpeta a través de conceptos temáticos muy rígidos y generales. Estas restricciones suelen provocar graves problemas de recuperación en fondos voluminosos que deben satisfacer una tipología de demandas muy diversificadas.

### **3.2. Los modelos automatizados**

Dentro de este grupo podemos diferenciar, siguiendo el sistema clasificatorio propuesto por Peter Enser (1995, p. 146), cuatro tipos de modelos en función de la relación existente entre el modo de consulta, el sistema de representación utilizado por el usuario para realizar la consulta y el sistema de representación usado por el ordenador para ejecutar esa consulta.

#### *3.2.1. Modelo Lingüístico-Lingüístico*

Tanto el sistema de representación de la consulta por parte del usuario como el sistema de representación de la información sobre la que el ordenador ejecuta la búsqueda son texto-lingüísticos. El texto de la consulta es utilizado para buscar en una base de datos que almacena registros textuales que describen imágenes. El contenido de estos registros puede ser resultado de una indización, o del almacenamiento del título o el pie de foto que acompaña a la imagen.



Normalmente este texto va acompañado de otros textos que contienen datos identificativos y de control de los documentos descritos.

### 3.2.1.1. Los sistemas informáticos con orientación relacional (RDBMS)

Marcan la aplicación de las técnicas de gestión de datos a la gestión de documentos gráficos. Estas técnicas suponen los primeros intentos de digitalización de fondos gráficos con motivos de mejorar su acceso y preservación. El principal avance respecto al modelo manual reside principalmente en que ahora sí es posible describir de forma individual cada unidad documental. Cada documento o grupo relacionado de éstos (un reportaje fotográfico, por ejemplo) merece una ficha descriptiva única. Podemos hablar, pues, de la irrupción de un verdadero catálogo (base de datos referencial), y de las técnicas documentales analíticas que superan la mera clasificación a que se somete cada documento en el sistema anterior como único método analítico y de recuperación. Se recupera ya a partir de textos lingüísticos más o menos normalizados. Este sistema permite una mayor flexibilidad en la recuperación, aunque sus posibilidades de descripción y recuperación son muy limitadas, teniendo en cuenta la gran carga semántica de la documentación gráfica, especialmente la iconográfica, y la limitación del empleo del modelo relacional para la recuperación de este tipo de documentos. Algunos sistemas que se ajustan a este modelo han intentado dar un paso que les acerca a los Sistemas Relacional-Documentales, al soportar un campo en el cual se puede insertar una imagen, como si se tratara de una propiedad más de la entidad documental descrita. Algunas implementaciones basadas en RDBMS sólo suponen un intento de digitalizar la clasificación temática utilizada para la organización de los documentos en el depósito, con la finalidad de superar las grandes limitaciones del manejo manual de vastos ficheros o libros de materias.

### 3.2.1.2. Los sistemas informáticos con orientación documental

Estos son los sistemas que actualmente copan el mercado profesional. Reflejan la adaptación de los sistemas de análisis y recuperación orientados al documento textual a los sistemas de gestión y recuperación de imágenes. Se utilizan Sistemas de Gestión Documentales, más eficaces que los sistemas relacionales a la hora de recuperar documentos en base a sus atributos semánticos. Los motores de recuperación son IRS, algunos de los cuales utilizan campos de tipo relacional — Sistemas Relacional-Documentales — para la gestión de aquellos atributos susceptibles de contener información estructurada, tales como los de identificación del documento. Los más adecuados permiten la visualización de imágenes en modo mosaico o a pantalla completa durante el proceso de recuperación, combinando procesos de selección visual humana con procesos de recuperación de información, basados en lógica booleana y palabras clave, aplicados

contra el texto lingüístico utilizado para identificar y describir la imagen. El usuario realiza una consulta a la base de datos textual y el sistema responde con los registros que se ajustan a esa consulta, mostrando un mosaico de imágenes a baja resolución. A partir de este mosaico es posible acceder al texto descriptivo y a una imagen a mayor tamaño de cada uno de los registros mostrados. Los más avanzados incluyen algoritmos de recuperación que superan la lógica booleana y el *exact match* permitiendo la ordenación de resultados por relevancia.

El peso del proceso de recuperación lo lleva un IRS contra la base textual generada por un indizador humano durante el proceso analítico-descriptivo de cada una de las imágenes que integran la base de datos. Este proceso consiste en la acotación de grupos de imágenes candidatas. El usuario de esa base de datos continúa el proceso de recuperación mediante la visualización de imágenes índice de los candidatos, pudiendo retroalimentar el proceso (generalizar, especificar o cambiar su estrategia de búsqueda) sólo mediante el empleo de texto lingüístico.

Los dos modelos lingüísticos que acabamos de estudiar se muestran ineficaces cuando se usan como único sistema de recuperación en bases de datos voluminosas y con una indización no exhaustiva. El problema se agrava ante necesidades de información poco definidas o no orientadas a la recuperación de imágenes por atributos semánticos muy concretos —temas generales o contenido icónico básico—. A pesar de la ayuda que representa posibilitar técnicas de recuperación visual a través del ojeo de las imágenes en forma de mosaico estos grupos de registros candidatos a satisfacer las necesidades de los usuarios sólo pueden ser creados a partir de consultas basadas en texto lingüístico.

El problema es que se trata de la adaptación de un sistema pensado para un tipo de representación de información, el texto lingüístico, bastante distinto al sistema de representación gráfico. La potencia demostrada por estos sistemas en grandes bases de datos textuales queda pronto en entredicho a la hora de ser aplicada a grandes bancos de imágenes. El sistema de significación de la imagen es distinto al texto, la semántica de una imagen es irreductible a descripciones lingüísticas. No es posible basar el núcleo de la recuperación documental en versiones textuales simples y fragmentarias de esos documentos. Este sistema es inaplicable cuando una parte importante, o la totalidad del proceso de recuperación, debe basarse en la identificación de patrones gráficos no figurados o sensaciones subjetivas del usuario, es decir, en atributos gráficos formales de la imagen o emotivos “indescriptibles”. No podemos llegar a imágenes relevantes por su nivel de significación formal o connotativo a partir de atributos visuales, por lo que en fondos gráficos voluminosos la recuperación por este tipo de atributos semánticos es muy limitada. Esto explica el fracaso de estos sistemas en fondos donde la recuperación visual es importante: bases de datos de diseños industria-

les, de logotipos comerciales, sellos, publicitarias, agencias espaciales, imágenes médicas o imágenes artísticas no figurativas. Otro motivo de fracaso es la evidencia del gran gasto que conlleva la indización humana de la imagen. A diferencia del documento textual, recuperable con técnicas *full text*, sin necesidad de una previa indización, la imagen no tiene texto o es muy escaso, o no está en formato electrónico. Hay abundante bibliografía relativa a este problema (Kanade, 1996; Kato, 1990; Firschein, 1995; Treat, 1996). Para que la recuperación lingüística sea eficaz es necesario que la representación lingüística de esa imagen tenga un contenido semántico adecuado a la naturaleza de la imagen y de las consultas del usuario. El proceso analítico-descriptivo requerido para generar esas representaciones puede llegar a desbordar la disponibilidad de recursos de cualquier sistema documental.

### 3.2.2. Modelo Visual-Visual (Basados en Contenido)

A la vista del fracaso de los sistemas informáticos con orientación documental para las necesidades de recuperación de imágenes en determinados tipos de bancos de imágenes se empieza a pensar en la posibilidad de aplicar técnicas de reconocimiento automático de imágenes, ya utilizadas con cierto éxito en los campos industrial, científico y militar. Estos sistemas son poco conocidos en el ámbito de Documentación, aunque ya hay empresas de desarrollo de software que han iniciado la comercialización de este tipo de aplicaciones. A modo de ejemplo podemos citar el producto *RetrievalWare Image Server* de la empresa Excalibur Technologies Corporation, y el producto *Ultimedia Manager* de IBM (Treat, 1996). Actualmente hay un buen número de equipos de investigación desarrollando este tipo de sistemas, la mayor parte de ellos en campus académicos y en departamentos de Inteligencia Artificial.

Los sistemas basados en contenido obtienen de forma automática atributos visuales del contenido de las imágenes y permiten la recuperación a partir de éstos a través de uso de índices visuales. El software es capaz de recuperar imágenes similares a la del índice visual por atributos formales en cuanto a color, figuras, textura y relaciones espaciales. Un índice visual puede ser un fragmento de imagen, un objeto de la imagen, o un esquema realizado por el usuario sobre un lápiz óptico. Gracias a los índices visuales este modelo permite iniciar un proceso de recuperación visual por atributos gráficos no textuales. El sistema va presentando al usuario grupos de imágenes visualmente similares a las de la imagen índice. A continuación el usuario ojea los grupos predeterminados por la máquina, pudiendo retroalimentar la consulta en cualquier momento.

Para la recuperación por atributos del nivel formal los sistemas basados en contenido son los más apropiados (Eakins, 1996). Las técnicas de este tipo de sis-

temas son cada vez mejor conocidas y los sistemas más potentes, esperándose grandes progresos para un futuro inmediato. Sin embargo, respecto al nivel del reconocimiento icónico (pre-iconografía e iconografía) los progresos son aún muy limitados. Para la recuperación por este nivel se requiere la aplicación de técnicas de Inteligencia Artificial, debido a la necesidad de emplear procesos de razonamiento para la identificación de tipos de objetos y su identidad — nombre comunes y propios — (ibidem). Los sistemas en desarrollo que permiten la interpretación de escenas ofrecen sólo resultados parciales en casos muy concretos, como es el caso de las aplicaciones militares. En cuanto al nivel iconológico, no se ha conseguido ningún progreso en los sistemas basados en contenido, habiendo muy pocas expectativas de futuro. Interpretar la significación y la intencionalidad de las imágenes a este nivel requiere un alto grado de razonamiento y una base de conocimientos muy cercana a la psique humana (ibidem). Hasta un futuro aún muy lejano no se podrá incluir en una máquina el conocimiento necesario para realizar una interpretación conceptual ni en cuanto a denotación ni connotación a partir de una imagen digitalizada.

### *3.2.3. Modelos Lingüístico-Visual y Visual-Lingüístico*

El modelo Visual-Lingüístico se basa en la generación de un tesoro visual, asociando texto a imágenes. El usuario consulta usando las imágenes del tesoro pero la recuperación se realiza con el texto asociado a cada imagen durante la indización. (Seloff, 1990). Este modelo es sólo válido para dominios muy específicos, aunque permite llegar a los niveles pre-iconográfico e iconográfico. El modelo Lingüístico-Visual intenta responder al siguiente problema: la consulta exclusivamente visual de los sistemas automáticos basados en contenido limita mucho las posibilidades de expresión de necesidades de información, pero es muy difícil representar una porción importante de la significación de una imagen textualmente (Enser, 1995, p. 159). Uno de los sistemas implementados más interesantes se basa en la asociación de imágenes a sensaciones o conceptos subjetivos descritos textualmente por los usuarios (Kato, 1990). Cuando el usuario, usando texto, introduce la sensación que le deben provocar las imágenes que desea recuperar en el sistema, éste localiza automáticamente, a través del texto introducido por el usuario, las imágenes índice previamente asociadas y las compara con las de la base de datos hasta dar con otras imágenes similares a las índice. Para ello utiliza un algoritmo de similaridad basado en la presencia y distribución de colores en las imágenes. Este sistema sólo ofrece buenos resultados para la recuperación por atributos semánticos formales y connotativos. Además es excesivamente complejo, siendo necesario la generación de una base de conocimiento para cada usuario de fondo.

### 3.2.4. *Sistemas mixtos*

Estos sistemas integran varios de los modelos anteriores. Generalmente se permite al usuario búsquedas combinadas, utilizando de forma conjunta técnicas del modelo Visual-Visual y una base de datos textual que contiene representaciones textuales producto de la indexación de las imágenes. En algunos casos esta base de datos es relacional, como en el sistema *Ultimedia Manager* (Treat, 1996), y en otros casos documental, como en el sistema utilizado por la empresa Corbis (accesible desde Internet en la URL <http://www.corbis.com>). Unos de los modelos más completos, pero más restrictivo en cuanto a posibilidades de aplicación a distintos tipos de fondos es el Art Museum (Kato, 1990). Este sistema se ajusta a los modelos Visual-Visual y Lingüístico-Visual sustentándose en una base teórica muy sólida.

Según el modelo semántico de la imagen descrito anteriormente, la tecnología que ofrecen los sistemas mixtos es la más adecuada para las necesidades de recuperación de un fondo gráfico voluminoso digitalizado. En la práctica ofrecen las herramientas técnicas que posibilitan el diseño de un sistema de recuperación que permita el acceso por todos los niveles semánticos de la imagen: desde los atributos formales, representados a través de índices visuales, hasta los atributos semánticos complejos representados a través del texto lingüístico.

## **4. Conclusiones: las posibilidades técnicas actuales y perspectivas de futuro**

Hoy día los sistemas más adecuados para permitir una recuperación por contenido de las imágenes son los que hemos denominado mixtos.

Tenemos que superar la idea simplista consistente en pensar que las técnicas de Inteligencia Artificial superarán no dentro de mucho tiempo las barreras de acceso intelectual a la información gráfica sin la necesidad de recurrir a las técnicas documentales de representación. El estado tecnológico actual y futuro a corto plazo obliga al uso de técnicas de representación normalizada de conceptos y contenido icónico basadas en el texto lingüístico: técnicas de indexación, lenguajes documentales y estándares para el intercambio de información descriptiva textual de imágenes que permitan la interconexión de sistemas de recuperación. El éxito o fracaso de un sistema de recuperación de documentación gráfico dependerá de la solidez de estas técnicas.

Sin embargo, los modelos de representación de ese texto deben tener en cuenta la naturaleza especial del documento gráfico, es decir, los atributos de significación útiles para su recuperación, sin caer en la aplicación automática de modelos de representación y recuperación diseñados para documentos no gráficos.

Será necesario seguir indagando en técnicas de recuperación basadas en contenido que superen el estrecho marco de un único sistema. Se debe superar la fase de sistemas *ad hoc* ya que la tendencia a la interconexión de sistemas obliga al diseño de sistemas abiertos, que puedan intercambiar este tipo de información descriptiva en el marco de redes globales de transmisión de datos.

## 5. Referencias

- Eakins, John (1996). Automatic Image Content Retrieval: are we getting anywhere ?..// Abstract of paper presented at the 3rd ELVIRA conference: May 1996.
- Enser, Peter (1995). Progress in Documentation Pictorial Information Retrieval. // *Journal of Documentation*. 51 : 2 (June 1995) 126-171.
- Alonso Erasquín, Manuel (1995). Fotoperiodismo : formas y códigos. Madrid : Síntesis, 1995.
- Firschein, Oscar (ed.) (1995). Image Understanding Program at ARPA. // *IEEE Expert*. 5 : 10 (October 1995) 8-38.
- Kanade, Takeo. Inmersion into Visual Media (1996). New Applications of Image Understanding". // *IEEE Expert*. 1 : 11 (February 1996) 73-80.
- Kato, Toshikazu; Kurita, Takio (1990). Visual interaction with Electronic Art Gallery. // M. Tjoa y R. Wagner (eds.). *Database and Expert Systems Applications. Proceedings of the International Conference : Viena, 1990*. p. 234-240.
- Nail, M. (1994). Multimedia and Image Handling Research Supported by the British Library. // *European Library & Information Research Policy. Proceedings of a seminar held at the Hotel Russell : London, Dec. 7-9, 1994*. p. 84-89.
- Panofsky, Edwin (1976). *Estudios sobre Iconología*. Madrid : Alianza Editorial, 1976.
- Romer, Donna M. (1995). Image and Multimedia Retrieval. // *Research Agenda for Cultural Heritage on Information Networks. The Getty Art History Information Program*. URL : <<http://www.gii.getty.edu/agenda/image.html>>.
- Seloff, G.A (1990). Automated Acces to the NASA-JSC Image Archives. // *Library Trends*. vol. 38 : 4 (1990) 682-696.
- Treat Harold ; Ort De ; Ho Jean (1996). Searching Images Using Ultimedia Manager. // *Informations Services & Use*. 16 (1996) 15-24.