

## **MODELO PARA UN SISTEMA DE BÚSQUEDA DE VIDEOS DIGITALES POR SU CONTENIDO.**

Ing. Ivis Cabrera Guerra [icabrera@grm.uci.cu](mailto:icabrera@grm.uci.cu)

Lugar de trabajo: Universidad de las Ciencias Informáticas. Facultad Regional Granma. Cuba

Ing. Roexcy Vega Prieto [rprieto@uci.cu](mailto:rprieto@uci.cu)

Lugar de trabajo: Universidad de las Ciencias Informáticas. Cuba

Tutor: Ing. Guillermo Báez Ramos

**Universidad de las Ciencia Informáticas.**

### **Resumen**

Actualmente, la investigación y la práctica han demostrado que los sistemas de búsquedas de archivos audiovisuales son más eficientes si están basados en su contenido. De manera general las búsquedas se hacen tradicionalmente a partir de descripciones de las medias, para almacenar la duración, el título, las extensiones, entre otras características que no logran describir el contenido. Es por ello que la eficiencia de las búsquedas no es la deseada en estos sistemas. El objetivo fundamental del presente trabajo está centrado en crear una propuesta de un modelo para el procesamiento y la búsqueda de medias siguiendo este criterio. Se propone el modelo de un sistema capaz de analizar los archivos audiovisuales y extraer las características que describen su contenido, para usarlas posteriormente en el proceso de búsqueda especificada. La pro-

puesta cuenta con los elementos requeridos en materia de procesamiento de medias y para ello utiliza el estándar de descripción de archivos MPEG-7.

**Palabras claves:**

Archivos audiovisuales, videos digitales, metadatos, estándares.

## Introducción

Con el vertiginoso desarrollo de las TIC (Tecnologías de la Información y las Comunicaciones) el mundo comenzó una nueva era, la Era de la Información. Grandes volúmenes de datos recorren todo el planeta y se hace indispensable su manipulación. El uso generalizado de Internet ha provocado que el flujo de información que transita por la red aumente cada día, y en especial aquella información relacionada con archivos multimedias, entre los que se encuentran los videos digitales.

Debido a sus características, los videos digitales han adquirido un lugar importante en la preferencia de los usuarios de Internet y en general de las personas que tienen acceso a él. Su utilización brinda la posibilidad de ir más allá de las capacidades de la televisión tradicional, pueden adquirirlo, manipularlo y almacenarlo. Además, les brinda no sólo imágenes en movimiento, sino también la palabra hablada, el sonido, las texturas de luz y oscuridad, el color y el tiempo.

Sin embargo la búsqueda de videos digitales no está siendo para los usuarios una tarea fácil, pues aunque se hace sencillo realizarla a través de buscadores disponibles en la red, por palabras claves que describen solo el título, el autor, y otras características del video digital que el usuario domina; se presentan problemas a la hora de responder a las necesidades y preferencias específicas de los clientes, ya que los resultados son ineficaces, debido a que la mayoría de los buscadores se limitan a las características antes citadas y no al contenido del video digital (Amengual Galdón, 2006).

Contar con un método eficaz y sencillo de buscar, gestionar, adquirir y recuperar el video digital que se necesite, es el mayor reto que enfrentan las tecnologías y estándares que dan soporte a la creación, desarrollo, codificación, decodificación y descripción de video independientemente de la plataforma.

La manera común de buscar un video digital actualmente es a través de buscadores disponibles en la red, o sistemas que realizan la búsqueda creados por diferentes compañías que se dedican al manejo de la información, o creados por organizaciones y empresas para las cuales se hace indispensable la gestión de videos digitales.

Existen dos generaciones de buscadores de video según la tecnología que utilizan (Casares, 2003). La primera generación son aquellos en los que la forma de organizar los resultados se basa en los metadatos del video, es decir, los datos que puede incorporar el propio video junto al contenido que hay en la página donde se encuentra (título, descripción, etiquetas, resumen, transcripción, resolución).

La segunda generación aunque también incorpora los elementos anteriores, llevan la tecnología un paso más allá, siendo ellos mismos los que utilizan reconocimiento de voz, análisis visual, reconocimiento del texto que puede aparecer por pantalla.

Un punto de partida esencial para lograr la eficiencia en el funcionamiento de sistemas de búsquedas de videos digitales, son los datos asociados al mismo material, más conocidos como metadatos. En ellos se resume la información del contenido de los videos, lo que permite el desarrollo de sistemas de búsquedas robustos que brinden respuestas rápidas y precisas, sin la necesidad de reprocesar nuevamente el material audiovisual.

Los metadatos de materiales audiovisuales provienen de diferentes fuentes: desarrolladores de contenido, editores, terceras partes o usuarios. Algunos de los metadatos se crean automáticamente y otros se añaden de forma manual; son parte del contenido o externos al mismo.

### **Propuesta d el modelo conceptual del sistema de búsqueda.**

Un sistema eficiente de búsqueda de información audiovisual debe ser capaz de definir elementos de recuperación que sean significativos en el contexto de la aplicación. Además de proporcionar un método de consulta que permita al usuario especificar de forma natural características selectivas.

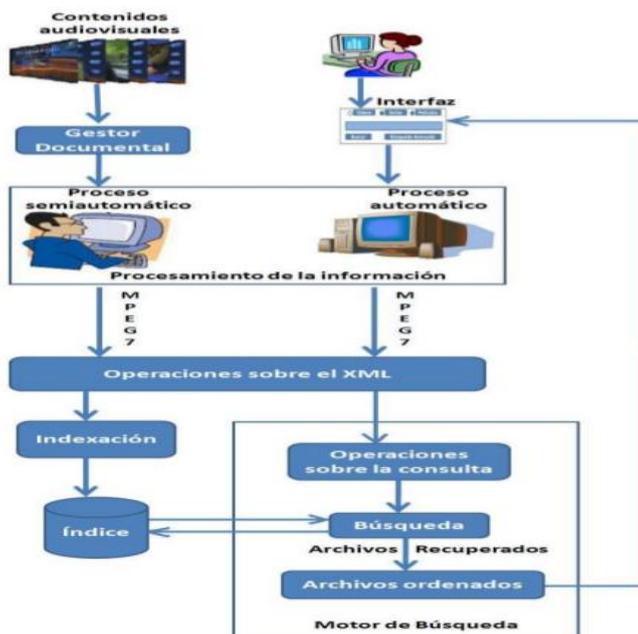
Los sistemas de consultas en bases de datos de videos se basan, tradicionalmente, en que realizan un proceso encaminado a obtener un conjunto de descriptores y utilizar estos para la búsqueda en la base de datos. De esta forma se elimina el coste de realizar el análisis de contenidos de los elementos que componen la base de datos en cada consulta, pero limita las búsquedas a aquellas que se pueden realizar con la combinación de los descriptores utilizados. De manera que este tipo de análisis es efectivo para conjuntos de archivos en los que los descriptores satisfacen cualquier consulta.

En los últimos años, varios sistemas prototipos se han propuesto abordar diferentes aspectos de la información contenida en los videos, como son las texturas, similitud de formas y relaciones semánticas entre objetos. El objetivo de las técnicas de consulta basadas en contenido es encontrar de forma eficiente los materiales audiovisuales en una base de datos que son similares a la búsqueda indicada.

El sistema que se propone debe ser capaz de realizar búsqueda a través de las dos técnicas de búsquedas existentes: las tradicionales (a través de texto) y las búsquedas a través de ejemplos, basada en descriptores que permitan una aproximación mayor al contenido de los archivos multimedia. Pueden ser derivados del análisis estructural de las imágenes, basándose en la utilización de elementos como el color, la forma y la textura que ya se obtienen en la etapa de análisis de la imagen.

La recuperación por contenido cambia la manera de especificar las búsquedas a un paradigma basado en lo que se conoce como búsqueda por ejemplos. Así, la consulta típica no se hace solo mediante una cadena de términos sino que también, introduciendo una imagen o dibujando un esbozo, a partir de lo cual el sistema buscará otras con características similares de color, forma o textura.

Analizando las características de los sistemas de búsquedas descritos anteriormente, basados en el nuevo paradigma de búsquedas por contenido, se propone un modelo arquitectónico donde se representan los principales módulos que integran un sistema de recuperación de información.



En el modelo existen dos procesos fundamentales el proceso de indexación y el de búsqueda el primero comienza cuando el gestor documental decide que material audiovisual se va a procesar, posteriormente pasa al procesamiento de la información que es donde se extraen los metadatos generados en un XML, de allí se pasa a realizarle las operaciones quedando definido los términos de indexación que se guardaran en una base de datos.

El segundo proceso comienza cuando un usuario especifica la búsqueda en una interfaz, de allí pasa al procesamiento de la información porque el usuario puede haber pasado una imagen como parámetro de búsqueda a lo que hay que extraerle todos los metadatos para poderlo comparar con los índices que están almacenados en la base de datos. Posteriormente pasa al motor de búsqueda que este analiza todos los índices existentes en la base datos ordenándolos de acuerdo a su importancia y mostrándolo en la interfaz finalmente. A continuación se explican todos los componentes del modelo conceptual.

**Contenidos audiovisuales:** Es el conjunto de materiales a los cuales se le realizará el proceso de indexación (videos, conferencias, series, películas, imágenes, subtítulos, sinopsis).

**Gestor documental:** Es el encargado de gestionar los archivos audiovisuales que serán procesados.

**Procesamiento de la información:** constará con los procesos automático y semiautomático.

- **Proceso automático:** Este proceso se realiza a través de un sistema el cual se encarga de realizar el análisis sobre texto (sinopsis, subtítulos), los cuales pasarían a formar términos de indexación.
- **Proceso semiautomático:** Se refiere a la detección de metadatos por parte de un usuario, de las escenas automáticamente detectadas así como el proceso de refinamiento de la detección automática. Permite la adición manual de metadatos, agregando información referencial básica como palabras claves, duración, formato, autor, título, breve descripción del video, información descriptiva como personas u objetos que aparecen.

Algunos descriptores que se obtendría del proceso semiautomático son los siguientes:

**Forma de la región** (*Region Shape*)

**Proceso de creación:**

- Definir la estructura.
- Crear las funciones comunes en todos los descriptores.
- Implementar las funciones más importantes del descriptor.
- Calcular el punto central de la imagen de entrada.
- Extraer de cada posición el valor y las coordenadas cartesianas correspondientes de la imagen.
- Calcular el valor de la función base ART.
- Hacer el módulo del coeficiente.
- Cuantificar el valor obtenido.
- Calcular la similitud entre dos descriptores de entrada.
- Dar respuesta.

**Estructura del descriptor de color** (*Color Structure Descriptor*)**Proceso de creación:**

- Definir la estructura denominada *VdColorStructure*.
- Construir las funciones comunes del descriptor.
- Definir el espacio de color.
- Cuantificar el valor no uniforme de cada color.
- Dividir la imagen en bloques de 8x8 píxeles.
- Crear la función de similitud.
- Imprimir la similitud por pantalla.

**Textura Homogénea** (*Homogeneous Textura*)**Proceso de creación:**

- Definir la estructura denominada *VdHomoTexture*.
- Construir las funciones comunes del descriptor.
- Definir memoria para 5 imágenes del mismo tamaño que la de entrada.
- Liberar espacio utilizado.
- Calcular distancia entre vectores.
- Se crea una función *VdHomoTextureDistance*.
- Mostrar resultado.

**Borde de histograma** (*Edge Histogram*)**Proceso de creación:**

- Crear la estructura de tipo descriptor denominada *VdEdgeHistogram*.
- Crear las funciones comunes para un descriptor.
- Crear las funciones específicas del descriptor.
- Calcular las dimensiones que tendrán los 16 bloques.
- Filtrar los bloques con la función denominada *DetectionEdges*.
- Calcular el valor absoluto de cada uno de los píxeles.
- Contar el número de píxeles que forman los contornos de la sub-imagen.
- Normalizar el valor resultante de la suma de píxeles.
- Cuantificar el valor obtenido de la normalización.
- Ejecutar y visualizar los resultados.

**XML:** Es un lenguaje de marcado que ofrece un formato para la descripción de datos estructurados. Esto facilita declaraciones de contenido más precisas y resultados de búsquedas más significativos en varias plataformas. Se puede definir un conjunto ilimitado de etiquetas; además ofrece un marco de trabajo para etiquetar datos estructurados, lo cual se puede realizar de forma exclusiva, lo que permite que un cliente construya búsquedas específicas. Sin XML, es necesario que la aplicación de búsqueda comprenda el esquema de cada base de datos, en el que se describe cómo se ha generado. Esto es prácticamente imposible, pues cada base de datos describe sus datos de una forma distinta. Sin embargo, con el lenguaje los archivos se pueden clasificar fácilmente en categorías estándar por autor, título, u otros criterios. De esta forma, se podría realizar búsqueda de una forma coherente. Se propone la estructura de descripción de una escena, basado en el estándar MPEG-7. (Ver anexo 2)

**Operaciones sobre el XML:** Es donde se hace el análisis del XML creado por el procesamiento de la información, generando así los términos de indexación.

**Índice:** Es la estructura de datos que posibilita un acceso veloz a la colección, una vez procesada, con objeto de obtener los documentos relevantes a una consulta. En este caso se crea un índice por cada descriptor especificado por el estándar MPEG-7 y extraído del documento XML.

**Indexación:** Es la creación de las estructuras de datos adecuadas para permitir un acceso eficiente y eficaz a los documentos. Es un elemento fundamental para los motores de búsqueda y las bases de datos. Tiene como objetivo elaborar un índice que contenga de forma ordenada la información, con la finali-

dad de obtener resultados de forma más rápida y relevante al momento de realizar una búsqueda. Sin un índice el motor de búsqueda debería escanear el contenido de cada archivo de forma individual cada vez que se iniciara una búsqueda, lo cual considerando la cantidad de información existente, sería un proceso lento y tedioso.



- **Análisis de las características:** Proceso por el cual las características extraídas del video quedan convertidas en etiquetas.
- **Etiquetas Meta:** Agrupan características con un significado común.
- **Eliminación de características repetidas:** Se eliminan las características que no aportarán información relevante para la realización de la búsqueda.
- **Etiquetas claves:** Se definen las características que contienen la información relevante, que se utilizarán en el proceso de búsqueda (Etiquetas de color, título, autor, formato, duración, textura, formas, palabras claves, descripción del video, información descriptiva como personas u objetos que aparecen).
- **Términos de indexación:** Lo constituyen las etiquetas claves.
- **Indexación:** Es donde se realiza la construcción del índice.
- **Índice:** Estructura de datos que posibilita un acceso veloz a la colección, una vez procesada, con objeto de obtener los documentos relevantes a una consulta.

**Interfaz:** Es la parte del sistema que permitirá la interacción del usuario con el motor de búsqueda. En ella se especificarán los datos que el usuario necesita para realizar la consulta así como la respuesta.

La interfaz permitirá realizar búsquedas por cadenas de caracteres, tendrá la opción de seleccionar una imagen que constituirá el patrón de búsqueda en el índice, además dará la posibilidad de elegir colores, proponer texturas y formas para realizar la búsqueda.

**Motor de búsqueda:** Aplicación diseñada para gestionar la información que el usuario especifica en la interfaz, buscando la misma en la base de datos. De esta forma cuando se realiza una consulta, el motor de búsqueda se dirige al índice para localizar los elementos deseados, arrojando así resultados precisos y rápidos.

- **Operaciones sobre la consulta**

Se comprueba si la sentencia de la consulta se realizó correctamente.

- **Búsqueda**

Se realiza la búsqueda sobre la consulta especificada en la interfaz utilizando los métodos de búsquedas implementados en la aplicación.

- **Archivos ordenados**

Después de haber recopilado los archivos que tienen relación con la consulta especificada se ordenan de acuerdo a la importancia de los mismos, mostrando este resultado en la interfaz.

## **Conclusiones**

El modelo propuesto surge luego de analizar las tendencias actuales de la búsqueda de contenidos audiovisuales, las particularidades y necesidades de un sistema de búsqueda de videos digitales a través del contenido, y de comprender cuán importante resulta el desarrollo de un sistema de este tipo para la gestión de el gran volumen de información audiovisual manejado en el mundo. Se puede concluir que:

- La búsqueda automática por contenido en materiales audiovisuales no está totalmente desarrollada, pues no es una tarea fácil para los desarrolladores, al ser muy costosa, y computacionalmente compleja.

- Al constituir un archivo de texto, los metadatos son la vía ideal para el acceso al contenido de los videos digitales. La indización de estos metadatos permitirá la obtención de mejores resultados en las búsquedas.

## Bibliografía

**Amengual Galdón, Sebastian. 2006.** “Motores de Búsqueda para contenidos audiovisuales”. [En línea]

[http://www.gti.ssr.upm.es/~eeg/docs/comunicacionesI+D\\_06.pdf](http://www.gti.ssr.upm.es/~eeg/docs/comunicacionesI+D_06.pdf).

**Casares, Javier. 2003.** “Manual de Posicionamiento”. [En línea] 2003. <http://www.manualdeposicionamiento.com/guia-de-referencia-seo/seo-para-videos/>. ISBN 13: 978-84-611-8192-6.

**Abreu Bartomeo, Yanedi, y otros. 2008.** Teleformación. Universidad de las Ciencias Informáticas, 2008. [En línea] <http://teleformacion.uci.cu/mod/resource/view.php?id=14077>.

**Castellano Báez, Maidileydys, Cervantes Rodon, Damian y Espinosa Ronquillo, Dayris. 2007.** Teleformación. Universidad de las Ciencias Informáticas, 2007. [En línea] <http://teleformacion.uci.cu/mod/resource/view.php?id=11402>.

**De Jong, Annemieke. 2003.** “Los metadatos en el entorno de la producción audiovisual. Una introducción”. Federación Internacional de Archivos de Televisión, 2003. [En línea] [http://archivesatrisk.org/restricted/standards/Metadata\\_version\\_Es\\_2003.pdf](http://archivesatrisk.org/restricted/standards/Metadata_version_Es_2003.pdf).

**Delcor Ballesteros, Jordi y Pérez Noriega, Verónica. 2006..** “DESCRIPCIÓN, INDEXACIÓN, BÚSQUEDA Y ADQUISICIÓN DE SECUENCIAS DE VÍDEO MEDIANTE DESCRIPTORES MPEG-7”. [En línea] <http://upcommons.upc.edu/pfc/bitstream/2099.1/3855/1/54955-1.pdf>.

**Echemendía Tourt, Fernando y Rivera Suárez, Yoel. 2008.** “Sistema para la Detección y Extracción de Textos en Videos Digitales”. [En línea] <http://biblioteca.uci.cu>.

**García de León, Alicia y Garrido Días, Adriana. 2001.** “Recuperación de información en Internet: Google, un buscador singular”. [En línea] <http://www.cinfo.cu/Userfiles/file/Cinfo/cinfo2001/v32n3a2001/recuperacion.htm>.

**Orozco García, Alejandra. 2008.** “La construcción de metadatos en el proceso de organización, análisis documental y recuperación de la información en los

archivos de imágenes en movimiento”. [En línea]

<http://www.patrimoniofilmico.org.co/docs/metadatos.pdf>.

**Vetro, Anthony. 2008.** “MPEG-7 Video Browsing and Summarization”. [En línea] <http://www.merl.com/projects/video-browsing/>.

# **SOCIEDAD DE LA INFORMACION**

[www.sociedadelainformacion.com](http://www.sociedadelainformacion.com)

Edita:



Director: José Ángel Ruiz Felipe

Jefe de publicaciones: Antero Soria Luján

D.L.: AB 293-2001

ISSN: 1578-326x