



GESTIÓN COMPLETA DE VÍDEO

Autor: Antonio Albacete García
albacete@isid.es
Noviembre de 2002

Valle de Tobalina, 40 (Urb. Las Lomas)
28669 Boadilla del Monte (Madrid)
Telf.: 91 632 49 50 - Fax.: 91 632 49 51

1 Introducción

Este documento pretende dar una visión pormenorizada de Videoma como herramienta completa de ingesta, administración y recuperación de activos digitales, siendo el principal objetivo el vídeo.

La principal prioridad de Videoma es dar una solución completa a la gran demanda de todo tipo de necesidad de gestión de vídeo, como **publicación en Internet**, comercialización de contenidos de vídeo pudiendo el usuario hacerse su vídeo a la carta desde el archivo general (**e-commerce**), gestión de **contenidos propios** para distintos cometidos como el de una televisión o el de una constructora para la gestión de todas las grabaciones que tiene de sus obras o para una compañía de seguros para la grabación de siniestros, formación a distancia para distintos procesos de fabricación en la fase de reparación, montaje, etc (**e-learning**), primer paso para la **televisión on-demand**, aplicaciones de **seguridad**, **control publicitario** y por supuesto **auditoria audiovisual**.

Dicha demanda es debida , por un lado, a la necesidad creciente de enriquecer los contenidos en Internet y, por otro, a la problemática que plantea tal necesidad a los propietarios de contenidos audiovisuales: archivos, agencias, productoras, cadenas de televisión, filmotecas y todo aquél que considere a su vídeo un activo importante dentro de su empresa. El tiempo de respuesta, en facilitar la secuencia de video solicitada y en formato digital, constituye el factor crítico origen de la problemática; que se complica en relación directamente proporcional al tamaño o volumen del archivo.

Así mismo, y sin perder de vista Internet, la próxima generación de sistemas de televisión, es decir la televisión digital, estará conectada a nuevas y diversas fuentes de información y entretenimiento provocando una demanda aún mayor de contenidos audiovisuales. Por ello, la correspondiente necesidad de un marco, compuesto por herramientas y sistemas, que permita el procesado, análisis, gestión e intercambio eficiente de la información audiovisual se hace evidente.

Un ejemplo claro de esta demanda de contenidos audiovisuales y del proceso de convergencia de la Televisión e Internet lo tenemos en el hecho que tuvo lugar el 10 de enero de 2000, cuando AOL America Online adquirió a la Time Warner. Este hecho, aunque sonado, no es más que un signo de una tendencia mucho más amplia en el sector de los medios y del entretenimiento. Esta tendencia de fusión de tecnologías dispares convertirá a Internet y la Televisión Interactiva en los entornos de comunicación mas poderosos, dinámicos y extensibles que se hayan conocido.

El vídeo nos intriga y fascina de forma que otros medios no consiguen, ya que contiene 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 mayoría de empresas que hacen uso del vídeo pronto descubren lo voluminoso y el difícil manejo que encierra este medio. Debido al formato analógico en que se encuentran, hoy día, la mayoría de contenidos o activos de vídeo, encontrar un segmento o toma particular significa combinar la búsqueda manual con las prolongadas sesiones de visionado lineal del material seleccionado. Las cadenas de TV, operadoras de cable y empresas de producción son las que más sufren estos procesos, pero también lo comparten empresas de sectores industriales que hacen un uso frecuente de este medio para sus producciones multimedia: sector editorial, agencias de publicidad, sector aeroespacial, industria del automóvil, sector turismo, etc. Fundamentalmente gestionan sus contenidos de vídeo de forma manual, un proceso que requiere mucho tiempo y personal y que resulta lento, costoso y, a veces, frustrante para todos aquellos involucrados.

Aunque la mayoría de archivos de vídeo existentes se encuentran en formato analógico, prácticamente todo el contenido que hoy día se está generando es digital. Esto hace que la conversión de los archivos analógicos existentes se preste, al menos, a una consideración seria, si no inevitable, sobre todo para aquellas empresas u organizaciones que deseen extraer todo el valor posible de sus inversiones en vídeo. La migración a vídeo digital tiene muchas justificaciones, siendo la más relevante la posibilidad de su distribución a través de las redes digitales - incluyendo intranets, extranets e Internet - y directamente a los hogares a través de banda ancha. Esta capacidad abre nuevos canales de distribución para la comunicación y el comercio. Con la tendencia creciente de uso de la Web desde los hogares y entornos de trabajo, que actualmente ya supera la cifra de 350 millones, augura un mercado prometedor para el vídeo B2B en Internet, al ya vigoroso mercado de business-to-customer. Por ello se explica fácilmente el explosivo interés en llevar el vídeo a la Web.

Las implicaciones son claras. Las empresas y organizaciones que quieran explotar al máximo el vídeo, deben habilitar los mecanismos necesarios para que de forma efectiva y eficiente puedan acceder, recuperar y reutilizarlo. Ello ofrecerá enormes beneficios y un potencial considerable de ingresos. El disponer de un acceso flexible e inteligente a los recursos de vídeo, combinado con métodos asequibles y simples de colaboración y distribución, se traduce en ciclos de producción más cortos, mejoras en los procesos de comunicación, sitios web más interesantes y atractivos, establecimiento de un marco idóneo para el e-business y nuevas oportunidades y flujos de ingresos.

El reto para las empresas cuyo negocio es el vídeo, consiste en la eficacia en la gestión del mismo. En como los usuarios, de forma intuitiva, pueden buscar en dichos activos de vídeo, encontrar las secuencias precisas que necesitan, reutilizarlas con creatividad, combinarlas, compilarlas y publicarlas con una facilidad y velocidad sin precedentes. Para las empresas que usan el vídeo como material de formación, promoción o simplemente de archivo de sus proyectos mas relevantes, como puede ser el caso de una constructora o una compañía de seguros, verán que recuperar esa información, para ellas importante, se convierte en algo tan simple como pulsar el ratón del ordenador.

2 Descripción

2.1 Introducción

El objetivo VIDEOMA es implementar un sistema completo que facilite a los usuarios todos los procesos implicados en la gestión de fondos audiovisuales, desde la entrada o ingesta del video hasta su posterior recuperación, una vez pasada la etapa de procesado del mismo.

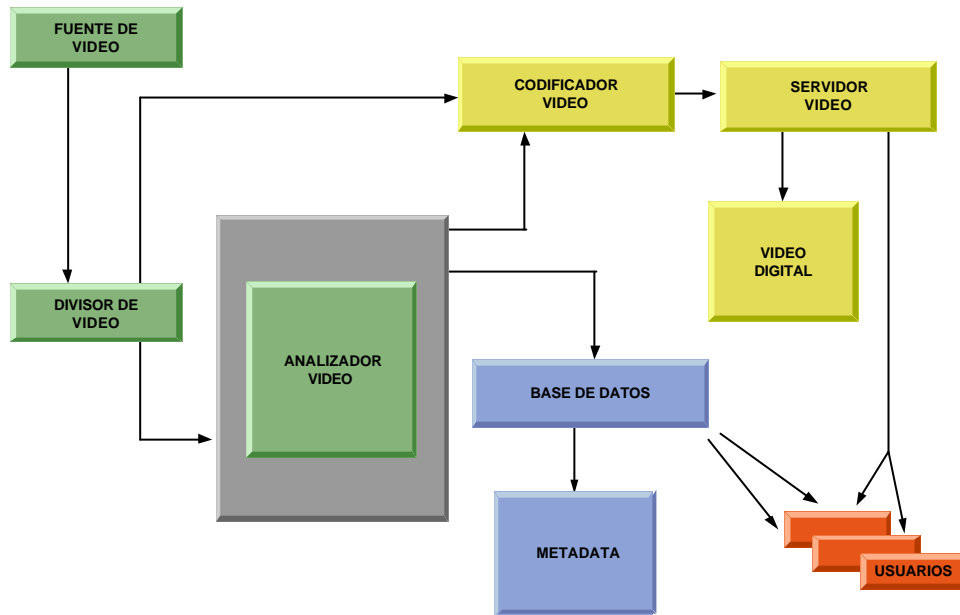


Figura 1. Esquema conceptual del sistema completo

A continuación se describe de un modo conceptual el funcionamiento del sistema en su conjunto.

El sistema está formado por dos grandes bloques: el núcleo o segmentador y las aplicaciones de gestión de los fondos digitales.

El primero de ellos, el núcleo, toma una de las salidas procedentes del divisor y lleva a cabo los siguientes procesos en tiempo real:

- Detecta automáticamente los cambios de escena que se producen en el vídeo y lo segmenta, es decir, extrae los keyframes o imágenes fijas representativas de la escena en curso, creando un índice visual del vídeo.
- Almacena en la base de datos dichos keyframes junto con la referencia temporal extraída del generador de código de tiempo del sistema.

- Al mismo tiempo, pone en marcha los codificadores de vídeo tanto de alta como de baja resolución con la misma referencia temporal que el segmentador (ello permitirá tener una referencia temporal única que permita acceder y recuperar la información audiovisual almacenada). Estos codificadores tienen como entrada la segunda salida del divisor de vídeo.
- Los codificadores de vídeo almacenan la información en los respectivos servidores de vídeo (baja y alta resolución) para que después puedan ser accedidos por los usuarios.

El segundo bloque es el encargado de la gestión (catalogación y recuperación) del archivo audiovisual que va generando el núcleo y proporciona las siguientes funcionalidades:

- Aplicación de catalogación de los clips generados por el núcleo del segmentador. Con el fin de permitir la posterior búsqueda y recuperación del vídeo almacenado en un archivo digital, es necesario disponer de una representación de sus contenidos. Algunos aspectos del contenido pueden extraerse de forma automatizada, pero una representación detallada obliga a que el vídeo sea catalogado o descrito con la suficiente granularidad.

Así pues, esta aplicación tiene la misión de acceder a la base de datos donde se encuentran los keyframes junto con los códigos de tiempo permitiendo al catalogador realizar su tarea de descripción de secuencias mediante la inserción de metadatos. Para ello se pone a su disposición un interfaz que le guía a través de un menú diseñado a tal efecto y que genera los metatags de un modo transparente al usuario.

- Aplicación de recuperación de secuencias. Permite que un usuario busque y recupere (en caso de existir en la base de datos) secuencias de su interés sin más que introducir unos parámetros que describan la imagen deseada. Esta aplicación le mostrará aquellos keyframes que respondan a los parámetros de búsqueda requeridos. El usuario en ese momento tiene la posibilidad de ver las secuencias en forma de stream de vídeo/audio proporcionado por el servidor de vídeo de baja resolución y decidir cuál es la que más le interesa.
- Aplicación de publicación y distribución a través de servidores de streaming y protocolos de distribución de contenidos audiovisuales de Internet.

Para poder entender bien cuál es la novedad que introduce este sistema es necesario remitirse a la dinámica de funcionamiento actual de los archivos audiovisuales de las televisiones, filmotecas y demás agentes del sector.

En la actualidad, los archivos audiovisuales están formados por almacenes que guardan miles de cintas de vídeo en las cuáles se graban de un modo

automático e indiscriminado horas y horas de programación. Podemos poner el ejemplo de un informativo de televisión. La labor del documentalista en este caso, consiste en recuperar de todas esas cintas, aquellas escenas de unos segundos que necesita mostrar en relación con la noticia que quiere desarrollar. Para ello tiene que buscar en las cintas siguiendo la dinámica de introducir cada cinta en un magnetoscopio profesional y usar las teclas de avance y rebobinado hasta encontrar la escena deseada.

2.2 Funcionamiento del sistema

Como ya se ha explicado anteriormente, el sistema tiene dos grandes bloques. La segmentación y la gestión de los contenidos.

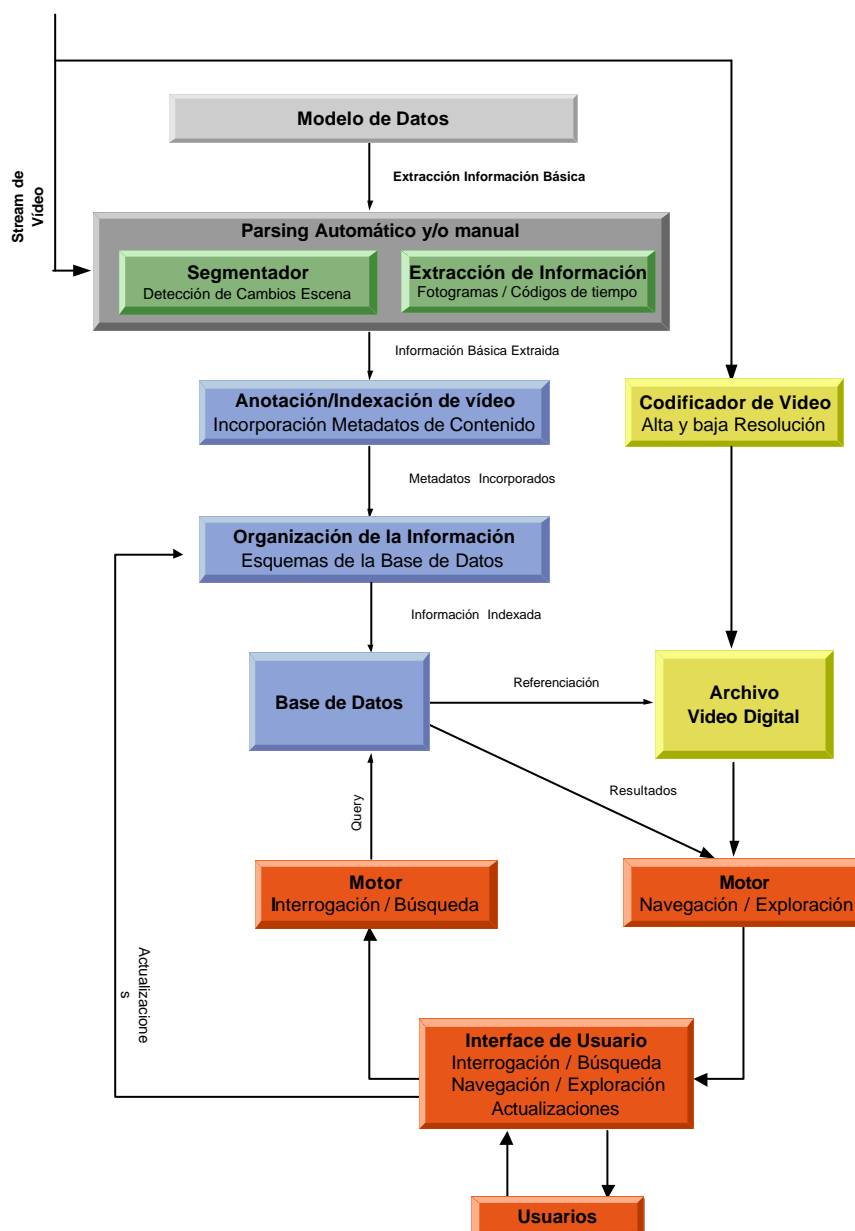


Figura 2. Esquema del sistema completo

1. Segmentación de vídeo.

Por medio del proceso de detección de cambios de escena se consigue la segmentación de un stream de vídeo en un conjunto de tomas (shot). De este modo, el stream de vídeo queda segmentado en tomas (shots) genéricas que constituyen las unidades elementales de la base de datos de vídeo. Cada toma queda representada por los fotogramas (keyframes) más representativos y el código de tiempo (time code) asociado a cada uno de ellos .

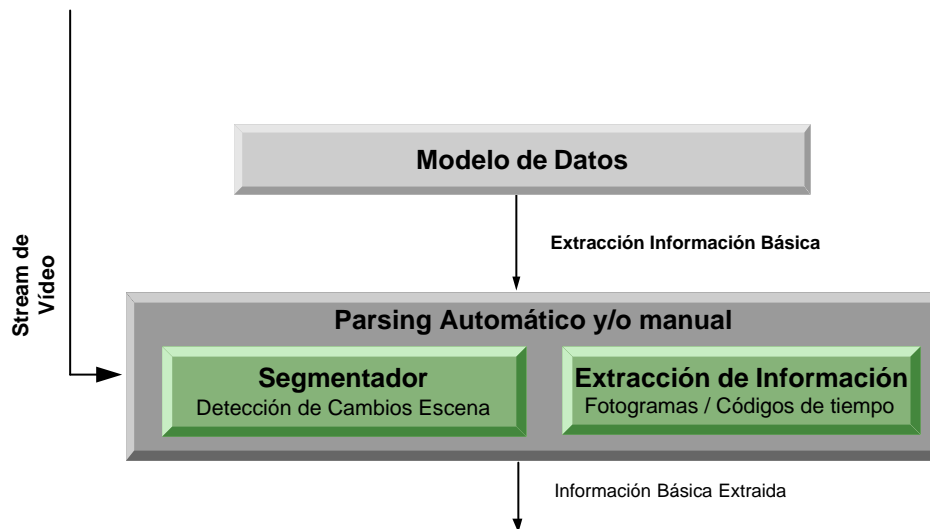


Figura 3. Proceso de segmentación

2. Catalogación de vídeo.

Es un proceso que consiste en asociar o vincular objetos de información descriptiva (etiquetas y/o descriptores semánticos) llamados metadatos, a cada una de las tomas obtenidas por el proceso de segmentación. Este proceso es esencial para facilitar la posterior recuperación de los contenidos del archivo.

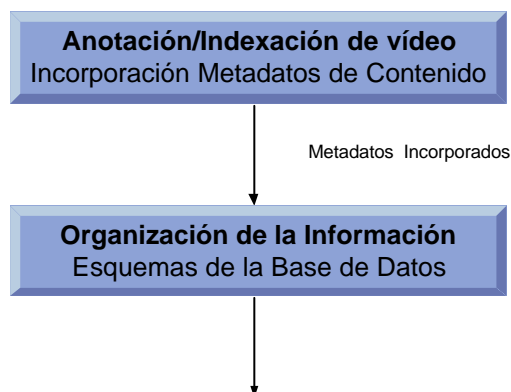


Figura 4. Proceso de catalogación

3. Búsqueda y recuperación.

Son los procesos que facilitan el acceso a las bases de datos mediante consultas basadas en texto y/o ejemplos visuales. Dentro de estos procesos también se contempla la facilidad para navegar o explorar los resultados de una búsqueda o consulta al archivo.

La gran ventaja que incorpora este sistema es la posibilidad, no sólo de realizar consultas y localizar tomas, si no además, poder visualizarlas en el mismo interfaz por medio de un stream de vídeo en baja resolución proporcionado por el servidor. Esto permite que el usuario compruebe de una forma definitiva que lo que ha encontrado es aquello que le interesa sin necesidad de usar una cinta de vídeo o un magnetoscopio en todo el proceso.

El interfaz de indexación y recuperación de secuencias guiará al usuario (navegación por un menú de parámetros descriptivos) en su búsqueda, consiguiendo que esta sea más eficiente.

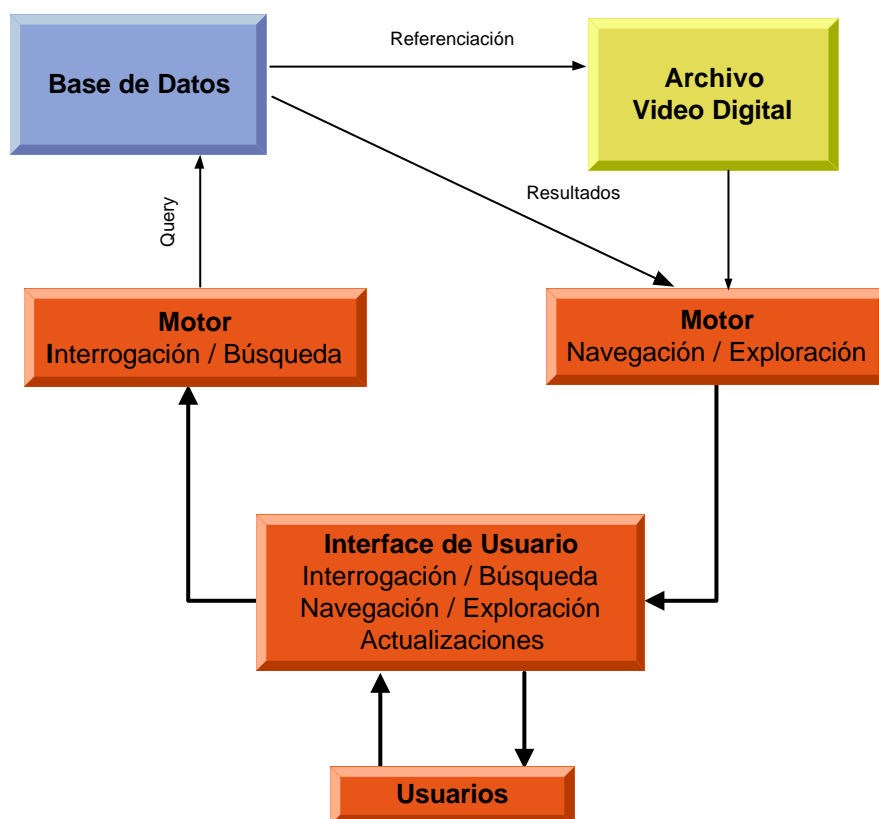


Figura 5. Proceso indexación y recuperación

2.3 Tecnologías empleadas

Entre las más relevantes se encuentran las siguientes.

2.3.1 Tecnologías de tratamiento de imágenes.

El desarrollo clave y propio del proyecto es la segmentación de vídeo en tiempo real, que mediante algoritmos de tratamiento y manipulación de imágenes permite conseguir nuestro objetivo, detectar cambios de escena en un vídeo para segmentarlo adecuadamente.

La detección de transiciones entre shots puede ser una tarea trivial o compleja dependiendo de cómo haya sido combinado el contenido del vídeo y del tipo de transición que haya sido utilizado. Por ejemplo, cuando se junta en un vídeo dos trozos de dos fuentes muy diferentes con ningún frame de transición, resulta relativamente sencillo detectar el cambio de escena. Por otro lado, si dos shots muy similares se combinan en una desaparición gradual de una, superpuesta con la aparición de la otra, los cambios visuales que podremos apreciar serán mucho menores que los esperados en un vídeo con un movimiento de objetos moderado. De esta forma, es muy probable que algunos algoritmos de detección de shots basados en imagen pierdan algunas fracciones de los límites del shot. Afortunadamente, esto no afectará de una forma importante en la calidad de la detección de escenas porque las transiciones graduales entre dos shots son elegidas por los productores cuando los dos shots están relacionados y pertenecen a la misma escena.

La detección de shots con la utilización de estos algoritmos se realiza mediante el análisis, evaluación y combinación de tres características de las imágenes, la media de brillo del vídeo, los cambios en los valores de brillo en cada píxel, de un frame al siguiente y, los cambios en la distribución de color de un frame al siguiente.

2.3.2 Tecnología de intercambio de datos e información

Para la representación interna de datos e intercambio de información y contenidos se incorporan estándares reconocidos internacionalmente y de uso cada vez más frecuente. Estos son el ISO MPEG7, el W3C XML y SMIL.

El ISO/IEC MPEG-7, formalmente conocido como Interfaz de Descripción de Contenidos Multimedia (Multimedia Content Description Interface), se utilizará para la anotación, descripción y categorización de los contenidos audiovisuales.

El W3C XML, Lenguaje de marcado (anotación) extensible (eXtensible Mark-up Language), se empleará para la formalización de esquemas de descripción y descriptores mediante el uso del DDL (Description Definition Language) especificado por el estándar MPEG-7. Esto permitirá definir y formalizar distintos lenguajes de descripción de información audiovisual para la anotación de secuencias de vídeo, por ejemplo: noticiarios, documentales, anuncios, eventos deportivos, etc. y facilitar enormemente el intercambio de datos y contenidos y su reutilización para fines diversos: web broadcast, comercio electrónico, B2B, etc.

El W3C SMIL, conocido como Lenguaje Sincronizado de Integración Multimedia (Synchronized Multimedia Integration Language) y basado en XML,

se empleará para la manipulación y posicionamiento de medios visuales en pantalla, para la sincronización de distintos medios (por ejemplo: una presentación con audio, video y la transcripción a texto del audio), y finalmente para la presentación de medios conforme a las preferencias del usuario, lenguaje, bit-rate, etc.

2.3.3 Tecnologías de codificación.

La codificación de vídeo tiene en este momento un papel importantísimo, incidiendo cada vez más en formatos de fichero lo menos voluminosos posibles y a la vez con la mayor calidad alcanzable. Nuestro criterio es utilizar estándares en codificación y compresión de vídeo. Estos incluyen los estándares de-facto Windows Media, RealVideo y QuickTime así como, los estándares internacionales ISO como MPEG1, MPEG2 y el nuevo MPEG4, según muchos destinado a reemplazar a los dos anteriores.

2.3.4 Tecnologías de Bases de Datos

El Sistema Gestor de Bases de Datos que empleará el sistema estará basado en la tecnología de base de datos Objeto-Relacional con soporte de herencia y polimorfismo. Esto facilitará en gran medida la manipulación de datos, objetos y flujos de información. Así, por ejemplo, los elementos gráficos, tal como un fotograma puede ser tratado como un objeto de información abstracta o "Blob" y ser manejado como cualquier elemento alfanumérico de los que habitualmente utilizamos. Al incluirlo en la base de datos, dentro de un catálogo específico, no será necesario describirlo pues heredará los elementos descriptivos y de categorización ya definidos en el propio catálogo.

2.3.5 Tecnologías Web

Se desarrolla la aplicación en un entorno Web, mediante herramientas de interrogación a Bases de Datos y formateo de páginas WEB estáticas y dinámicas

2.3.6 Tecnologías de Workflow

La información, ya sea antes de ser tratada mediante los distintos elementos del proyecto o después de ello, sigue un curso lógico. En ningún caso podremos realizar una tarea si previamente no hemos realizado la inmediatamente anterior en un cierto grado, definida por el propio sistema.

2.4 ARQUITECTURA

La arquitectura del sistema variará en función de las fuentes de entrada que tengamos, de la cantidad de almacenamiento que queramos mantener en línea, de los formatos a los que hayamos digitalizado nuestros contenidos y de la cantidad de información que vayamos a mover a través de nuestra red de área local.

Hay varias alternativas que podemos contemplar además de lo anteriormente expuesto. ¿Qué Base de Datos queremos utilizar? ¿Vamos a tener algún sistema de backup automático de los datos? ¿Los sistemas de almacenamiento van a poder ser controlados por algún HSM? ¿Qué grado de automatización vamos a poder implementar en nuestra solución a la hora de la ingesta?

VIDEOMA utiliza aplicaciones estándar para la programación de sus módulos, PHP y C++. Como Base de Datos soporta Oracle 9i y MySQL y como Servidor Web IIS y Apache.

El resto de necesidades para poner en marcha el sistema podrán ser aquellas disponibles en el mercado, como diversos sistemas de almacenamiento, diversos servidores de datos, diversos codificadores de vídeo, etc.

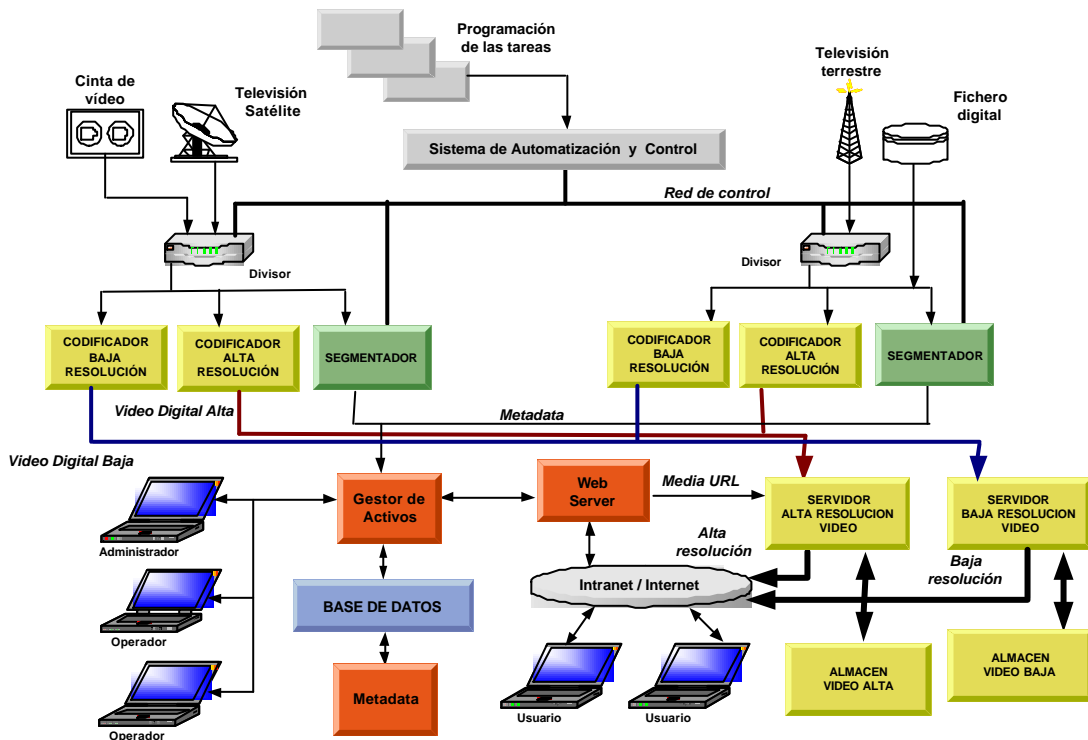


Figura 6. Arquitectura del sistema

3 Segmentador

3.1 CARACTERISTICAS

3.1.1 General.

Comienzo Sesión Programada

Captura Manual

Parada Manual

Control Sensibilidad

Revisión Sesión

3.1.2 Tipo de captura de keyframes y Control Publicitario.

Captura de keyframes automática

Captura de keyframes periódica

Detección automática de anuncios

3.1.3 Control de Sesión y de Código de Tiempo.

Trabajar con Sesión Múltiple.

Trabajar con distintas fuentes de Código de Tiempo

3.1.4 Control de Base de Datos.

Sin conexión a Base de Datos.

Con conexión a Base de Datos.

3.1.4 Control de Codificadores.

Control de codificador Windows Media

Control de codificadores MPEG-1, MPEG-2 y MPEG-4

3.2 IMAGENES

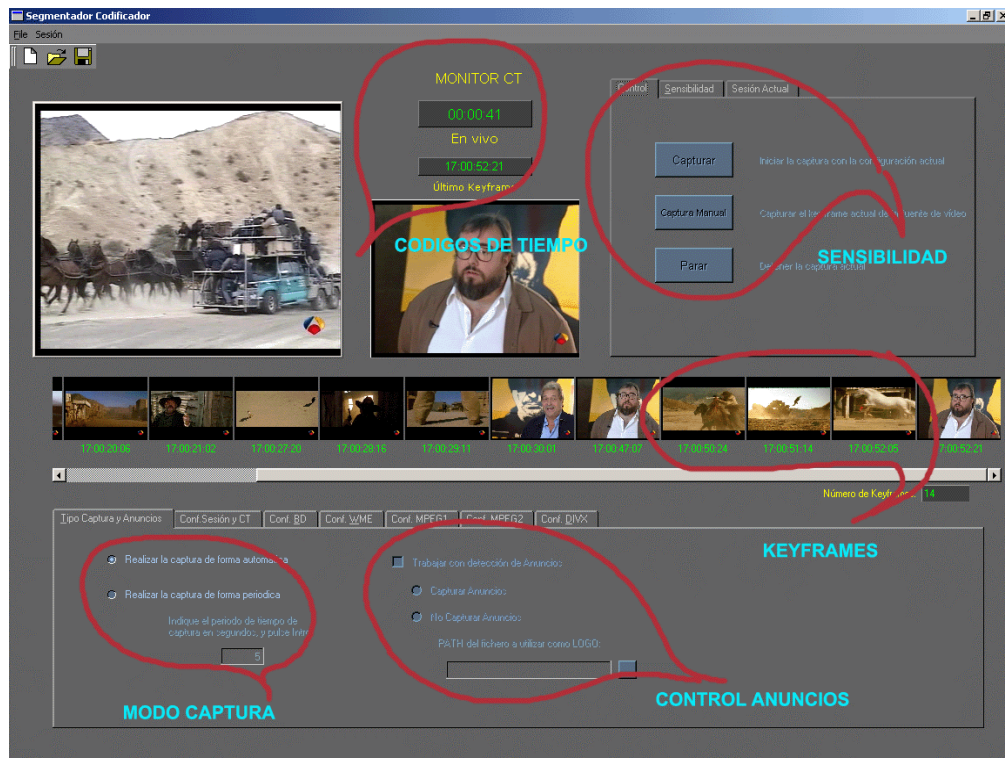


Figura 7. Segmentador Control Principal

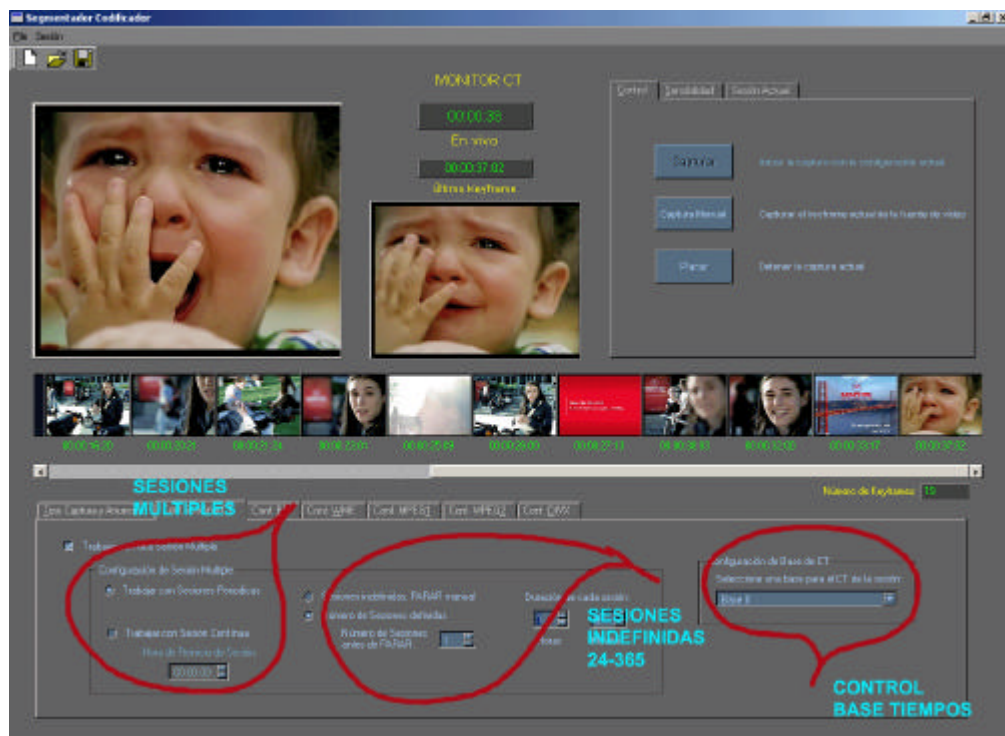


Figura 8. Segmentador Control Sesión

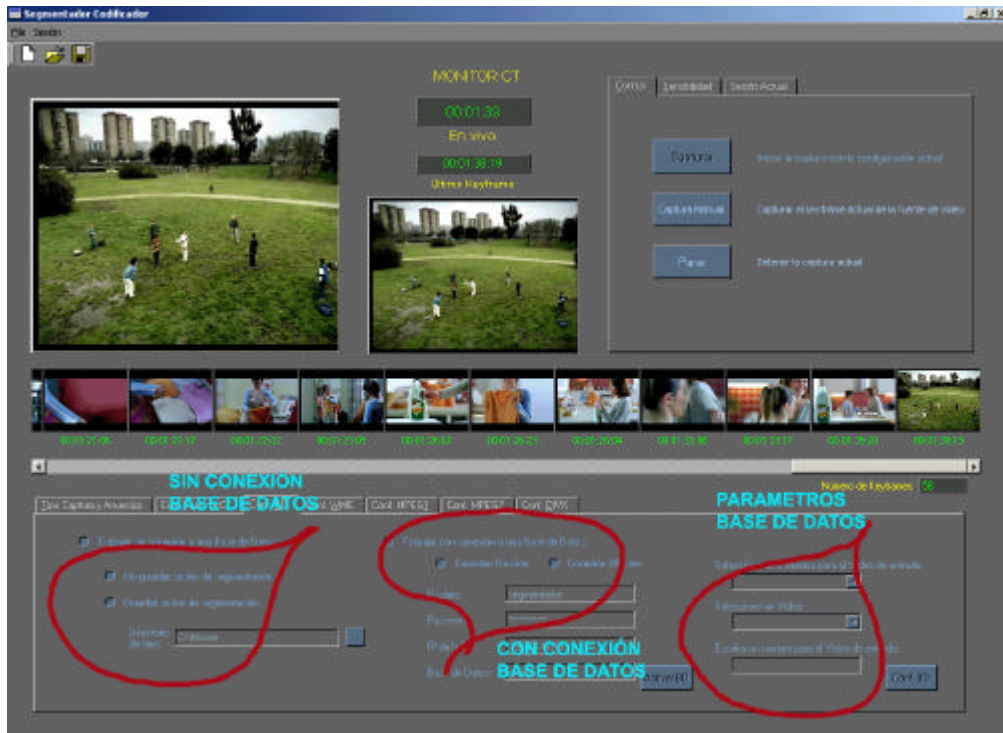


Figura 9. Segmentador Control Base de Datos

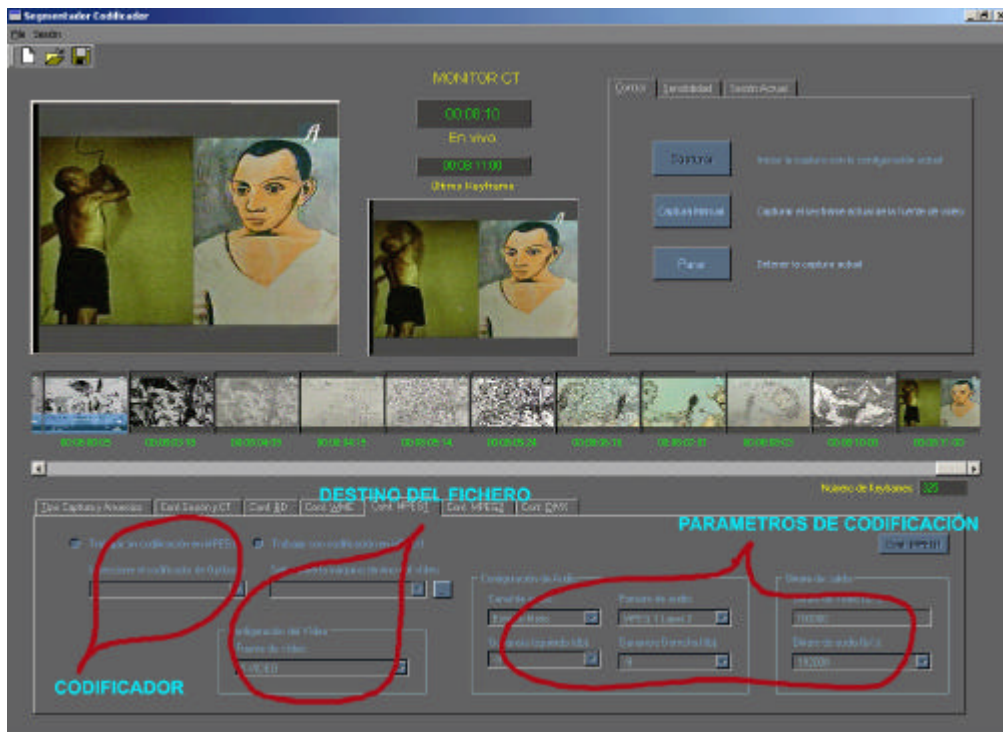


Figura 10. Segmentador Control Codificadores

4 Catalogación y Administración.

4.1 CARACTERISTICAS

4.1.1 Gestión de Usuarios.

Insertar Usuario

Modificar Usuario

4.1.2 Gestión de Esquemas.

Crear Esquema

Modificar Esquema

Ver Esquema

Borrar Esquema

4.1.2 Gestión de Videos.

Insertar vídeo

Control de Estados del vídeo.

Sin Segmentar

Segmentando

Segmentado

En Documentación

Documentado

Publicado

Creación y administración de clips.

Mediante keyframes

Mediante Códigos de Tiempo

Desde el propio vídeo

Inserción de imágenes, documentos, etc.

4.2 IMÁGENES

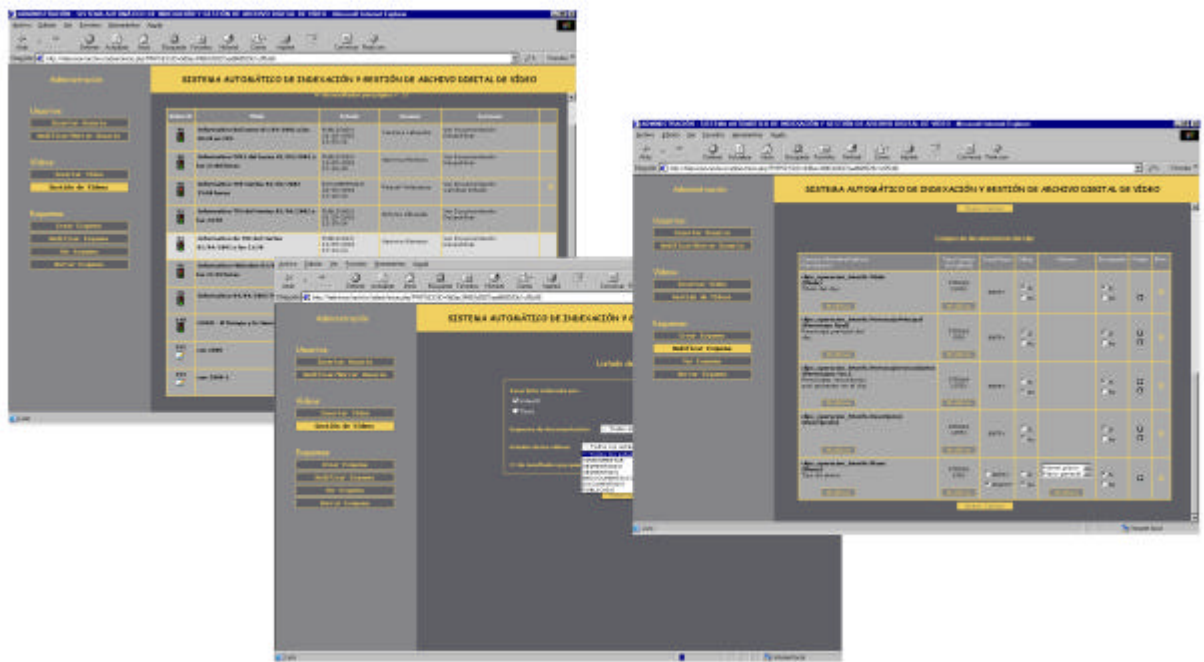


Figura 11. Administración de la aplicación

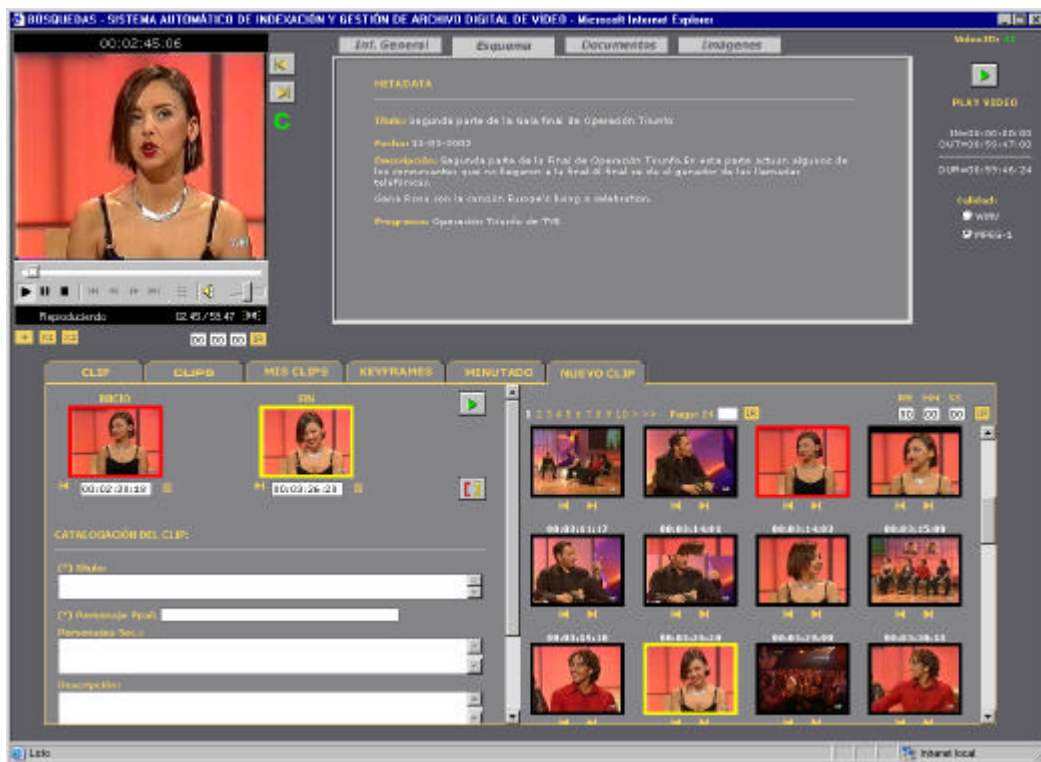


Figura 12. Creación de clips

5 Búsqueda y Recuperación.

5.1 CARACTERISTICAS

Búsqueda de videos y de clips.

Generación de EDL

- Ordenar los distintos clips

- Visionado

- Exportar a distintos formatos de edición

Generación de Carrito de la compra

Visionado de los resultados

- Visor independiente con metadata

- Visor acompañado del resto de documentación asociada.

- Selección de Formatos de visualización.

Información inmediata solo con arrastrar el ratón

Añadir a "Mis Favoritos"

Enviar URL por correo electrónico

Generación de Informe (pdf o txt)

Creación de tus propios clips.

- Mediante keyframes

- Mediante Códigos de Tiempo

- Desde el propio vídeo

Recuperación de información adjunta (Imágenes, documentos, etc).

5.2 IMAGENES

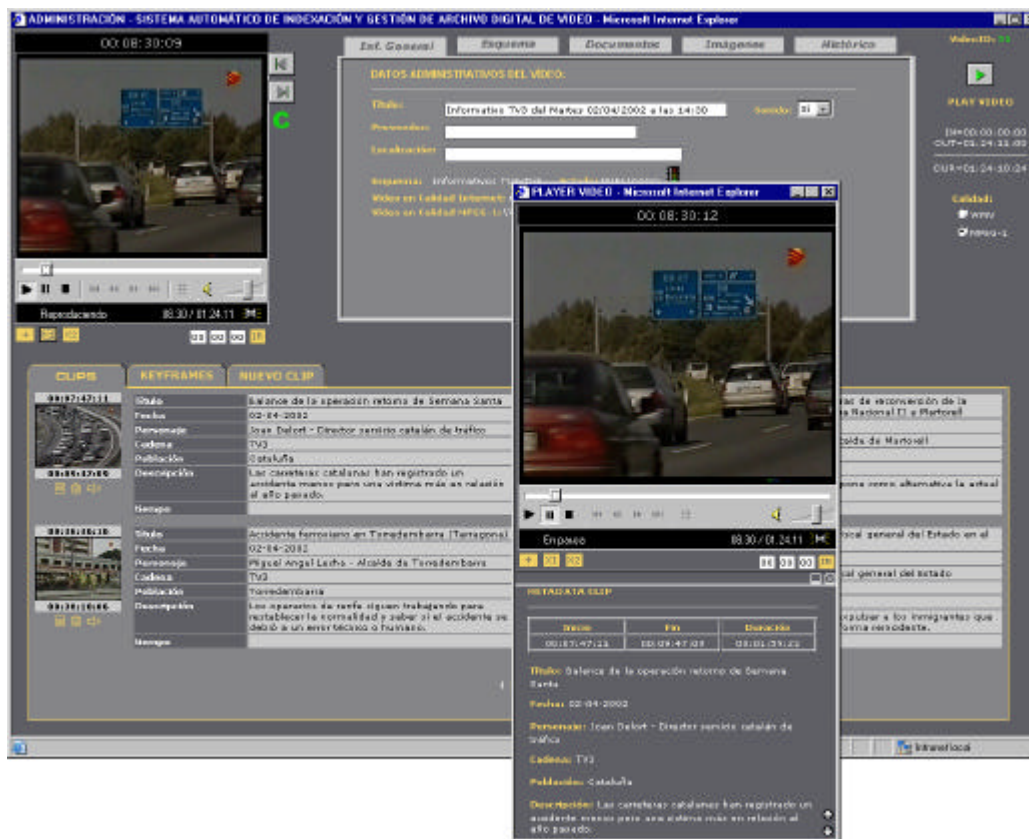


Figura 13. Recuperación y visionado

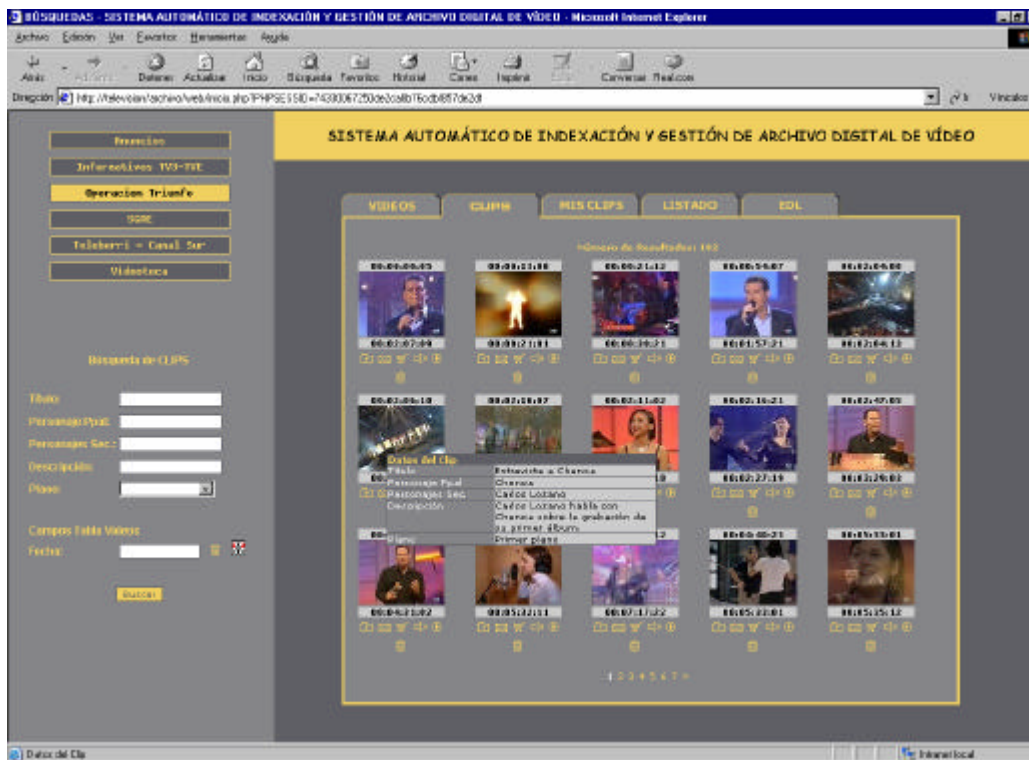


Figura 14. Interfaz de recuperación