

## **Capítulo 3**

### **Diseño**

Este trabajo explora nuevas funcionalidades para U-DL-A: la incorporación de un acervo indexado de video con servicios de búsqueda y recuperación de este tipo de datos vía Internet. Se plantea un esquema de indexado automático de video basado en técnicas de reconocimiento de voz y procesamiento de imágenes así como el desarrollo de las interfaces de usuario necesarias para la utilización del acervo. Los contenidos de los documentos de video pueden variar, pero para fines de esta primera etapa del proyecto todos tienen un fin académico y se refieren a asuntos técnicos y de investigación: cátedras de los profesores, defensas de proyectos de tesis y conferencias.

#### **3.1 Incorporación de Video U-DL-A a la biblioteca digital**

La concepción de Video U-DL-A respeta los niveles definidos en la arquitectura de la biblioteca digital U-DL-A:

- ?? Ofrece un grupo de interfaces disponibles en Internet que le permiten al usuario visualizar los contenidos del acervo, administrar los objetos almacenados, crear anotaciones en un documento particular y efectuar búsquedas.

- ?? Define un nuevo y eficiente servicio de entrega de video en demanda que puede ser utilizado por cualquier aplicación desarrollada en la biblioteca digital.
- ?? El manejo de datos se establece mediante funciones de bajo nivel que proporcionan acceso a los materiales almacenados (videos, imágenes, texto) desde las capas superiores.

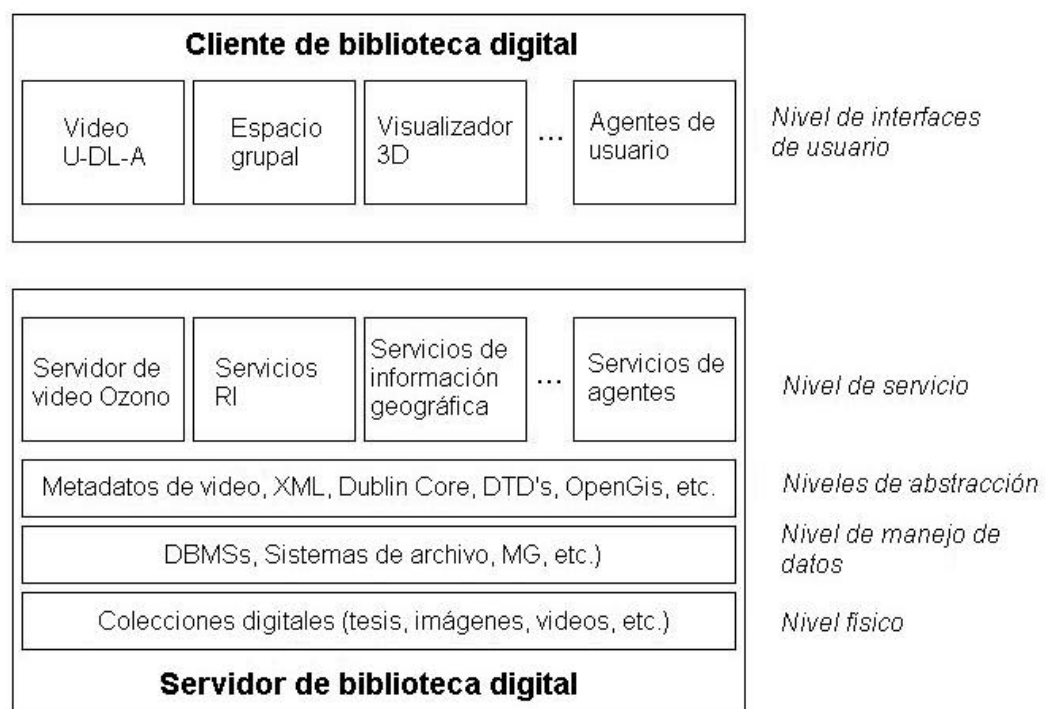


Fig 3.1 Arquitectura de la biblioteca digital con la agregación de Video U-DL-A. Nuevas interfaces y nuevos servicios de manipulación de datos multimediales son definidos.

### 3.2 Modelo de datos

Existen diversos tipos de datos que necesitan manejarse en Video U-DL-A, con sus atributos y sus métodos para accederlos. Su clasificación se ilustra en

figura 3.2, donde podemos visualizar cada medio con su característica fundamental de utilización y la relación entre ellos.

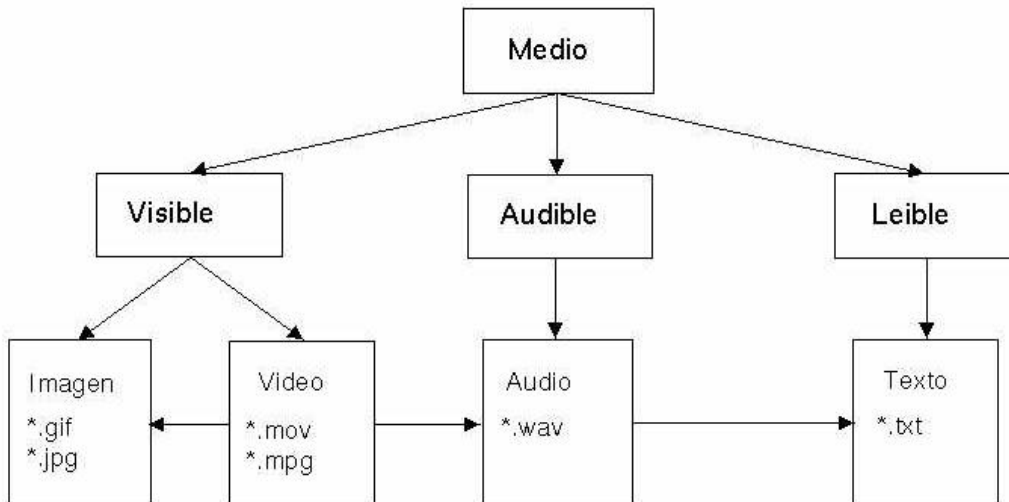


Fig 3.2. Diferentes medios en Video UDL-A. Las líneas horizontales reflejan la dependencia entre los mismos. La información en video es procesada para extraerle el canal de audio y una serie de imágenes representativas. A su vez, el audio es procesado y las palabras encontradas son transcritas.

Los datos de video representan una imagen que cambia continuamente en el tiempo. A diferencia de los tipos de datos tradicionales, el video se accesa de manera incremental como un flujo de tramas en tiempo real.

Existen también meta-datos o descripciones abstractas asociados a los objetos de audio y video en forma de anotaciones de texto e imágenes que representan sus contenidos y son almacenados en forma de estratos.

### Construcción de meta-datos

Los meta-datos relacionan al video con conceptos en un dominio de aplicación y permiten, con algún grado de certeza y en un intervalo de tiempo, describir semánticamente el contenido de los multimedia. Un concepto es una idea abstracta de una entidad en un dominio de aplicación.

El grado de precisión de un meta-dato puede ser representado por un peso  $w \in [0,1]$  que especifique que tan bien se relaciona el contenido del medio con un concepto.

**Definición 1** [Westermann and Klas 1999].

Sea  $M$  el conjunto de datos multimedia. Sea  $C$  un conjunto de conceptos.

La función *duración* regresa la duración del medio  $m \in M$ .

Para todo  $m \in M$ , sea  $I_m$  el conjunto de intervalos de tiempo  $i$  válidos de  $m$ ,  $i = [s,e]$  con  $s \leq e$ ,  $s \geq 0$  y  $e \leq \text{duración}(m)$ .

**Definición 2** [Westermann and Klas 1999].

$a_m = (m, c, i, w)$ ,  $m \in M$ ,  $c \in C$ ,  $i \in I_m$ ,  $w \in [0,1]$  es llamada una anotación o meta-dato de  $m$  con un peso  $w$ .

La figura 3.3 muestra como puede verse un video al que se le ha asociado un conjunto de meta-datos.

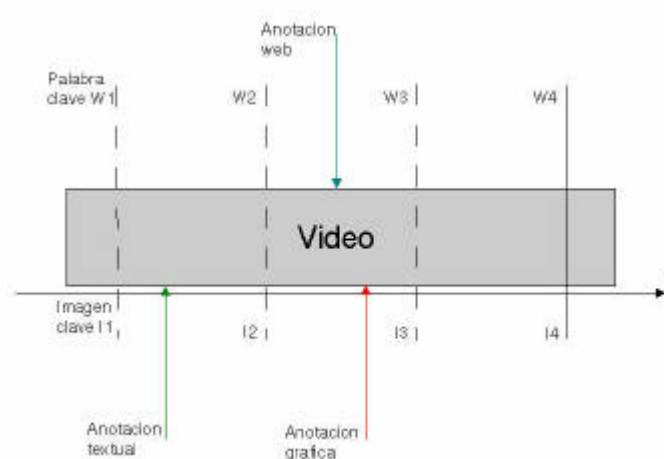


Fig 3.3. Video con meta-datos. Las palabras e imágenes clave definen segmentos temporales mientras que las anotaciones, aunque son válidas en un intervalo específico, son asociadas en cualquier instante de tiempo.

### 3.3 Módulos principales

Para su implementación, Video U-DL-A cuenta con tres módulos funcionales bien diferenciados.

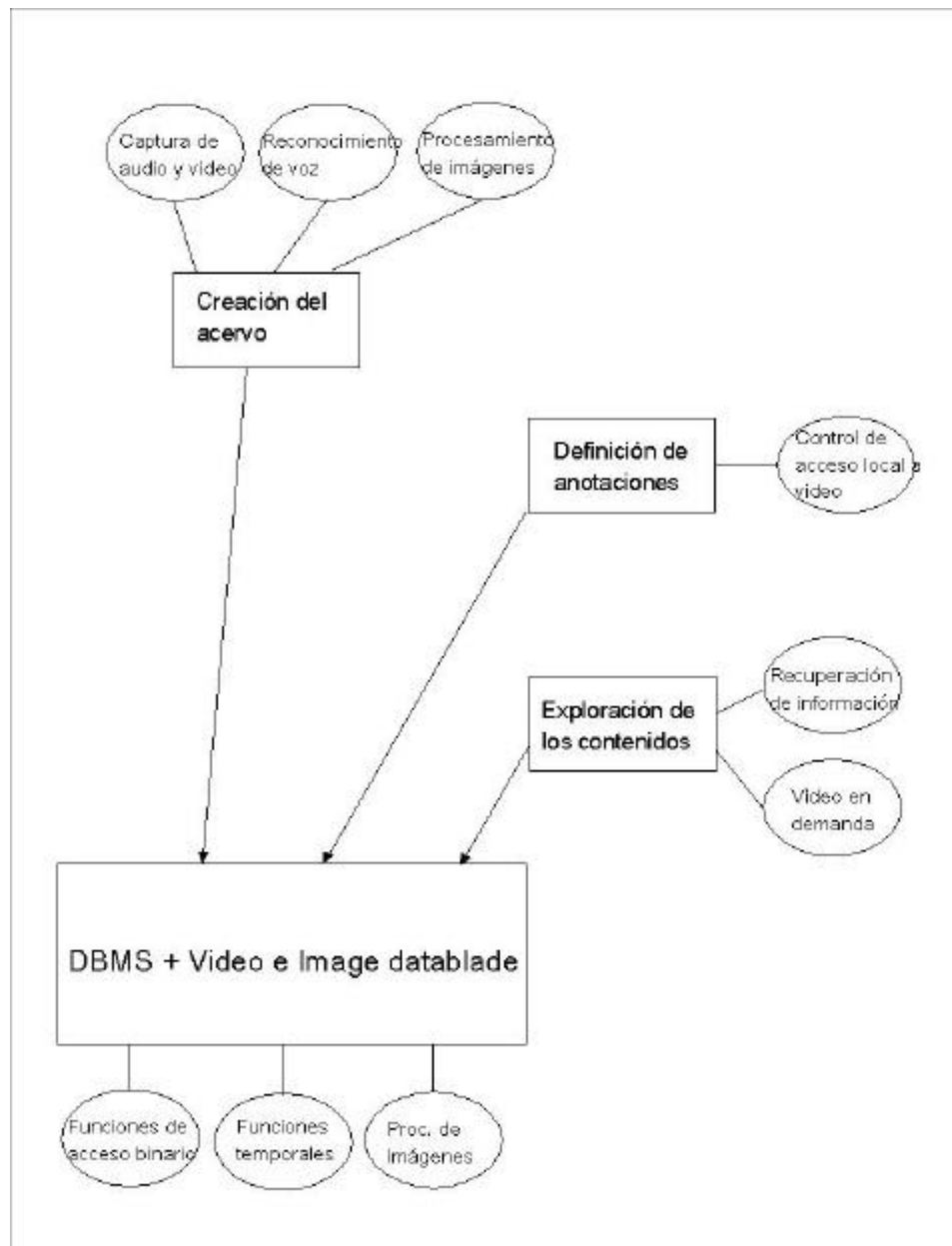


Figura 3.4. Módulos funcionales de Video U-DL-A.

## **Creación e indexado del acervo**

Cada vez que un nuevo material es incorporado al acervo, se somete a un proceso de indexado mediante un sistema de reconocimiento de voz y procesamiento de imágenes. Las palabras clave reconocidas son incorporadas a los documentos en una capa de meta-datos o estrato, mientras que un conjunto de imágenes tomadas de la información de video es procesado y la información resultante es también incorporada al documento en una capa de meta-datos.

## **Definición de anotaciones**

Es necesario ofrecer la facilidad de agregar otros estratos de información a los documentos almacenados como son: descripciones de texto que abunden sobre el contenido multimedial, ligas Web que den referencia a páginas que puedan contener información relevante acerca de lo tratado en los materiales y gráficas o imágenes que complementan lo que se ve o describe en el documento. Esta información adicional debe estar referenciada a un tiempo específico de ocurrencia en el material multimedial.

## **Exploración de la biblioteca**

Un usuario debe ser capaz de consultar los materiales almacenados así como de realizar búsquedas sobre temas específicos. Los resultados le serán proporcionados por un método de recuperación de información y un módulo especializado de transmisión y visualización de multimedios.

El siguiente capítulo describe los detalles técnicos de la implementación de estas funcionalidades.