The retrieval tool (cf. figure 3, right) allows even naive users (e.g. journalists or program directors that have not been trained to use a database) to retrieve video material from the archive, using the available metadata from the catalog server. The database can be queried using both textual and visual data. Textual queries address specific fields of a video entity defined by the documentalist during the archival process. Visual queries address visual metadata extracted during the preprocessing phase. For instance, one may specify an example image from those already available on the user’s desktop, and ask for all video entities whose keyframes match its visual content. Alternatively, one may define the desired type of camera motion, for instance in terms of pan, tilt, and zoom. Furthermore, all these aspects may be combined into a unique query by means of weighting coefficients. Once the desired video entities are selected, it is possible to export them to the editing tool in order to compose a new video.

2. Dissemination of results

The system prototype has been demonstrated to several private companies and we are currently negotiating technology transfer agreements with two of them. The system architecture, the database schema, and the video and image processing algorithms have been presented at the Fribourg meeting of the MPEG-7 consortium (Oct. 1997), which aims at standardizing multimedia data representations for archival and retrieval.

Overall, the DVP project has allowed the vision group of the CUI to start a research group on image and video databases. This group has produced a number of scientific publications, and its activities are currently being continued through new research grants.

We would like to thank the Télévision Suisse Romande (Geneva) for their help in defining the user requirements. TSR has joined in 1996 the DVP consortium as a sponsoring partner.

1.2 The Video Archival and Retrieval Subsystem

Catalog Server

The role of the catalog server is to compute and to provide access to metadata extracted from video clips retrieved from the archive server. Metadata are of two types: descriptions of the video visual content, and textual descriptions. Visual descriptions are automatically extracted by processing the video clip in the MPEG-1 format, without requiring decompression.

![Example of MPEG-1 motion vectors.](image)

First, a clip is decomposed into smaller segments, and transitions between shots are detected by analyzing each frame's motion vectors. For each shot, still images (keyframes) are extracted for display purposes. Moreover, a compact, numerical representation is extracted from each keyframe using a wavelet decomposition, in order to enable browsing and search through similar images. Finally, camera and camera lens motion (pan, tilt, zoom, stationary) properties are also computed from the motion vectors (cf. figure 2).

The second type of metadata consists of textual descriptions introduced by a documentalist at various abstraction levels of the clip, as well as close-caption data retrieved from the original video.

Graphical User Interfaces

The CUI has developed two tools that run on the end-user workstation: the archival tool and the retrieval tool. Both have been implemented as Java applets in order to maximize portability. The archival tool (cf. figure 3, left) allows a documentalist to analyze the video's content, using the metadata automatically extracted from the clip (keyframes, temporal segmentation). The user can display/edit these results, playback the video using random access through the keyframe display, and may define group of shots that, albeit visually different, may share a common semantic content. Once a video entity (shot, group of shots, clip) has been selected, the user can enter additional textual annotation.

![Main panels of the (left) Archival Tool and (right) Retrieval Tool.](image)
Distributed Video Editing, Archival and Retrieval

R. Milanese, A. Jacot-Descombes, T. Pun
F. Deguillaume, L. Petrucci, M. Cherbuliez, A. De Giacomi

Within the Distributed Video Production project (European Union ACTS A089), the goal of the Distributed Video Editing, Archival and Retrieval (DVER) application is to provide broadcasters with a complete solution for distributed video post-production. This system should combine together archival, retrieval, and editing functionalities in order to increase the accessibility and reuse of archive material. Moreover, the system should be geographically distributed, guarantee a high degree of portability to different platforms, and employ digital video using standard compression formats. All these objectives have been achieved, and a complete prototype has been integrated and put in operation by our end user (MegaChannel TV) for news post-production.

1. Work done by the CUI

The Centre Universitaire d’Informatique (CUI) of the University of Geneva has played a major role in this application, by contributing to the assessment of the state of the art in the field, by participating to the functional specifications of the complete DVER system, by completing the technical specifications of the archival subsystem, and finally by implementing the archival subsystem.

1.1 DVER Architecture Design

The system architecture includes an archive server, an editing server, a catalog server, and a client station for the end user (see figure 1). The archive server offers video streaming services through ATM, as well as file transfer through FTP/IP. Compressed digital videos are stored at two quality levels. The low-bitrate version (MPEG-1, 1.5 Mb/s) is used mainly for archival/retrieval/browsing purposes. This type of videos can also be edited on a low-cost client station. The high-bitrate version (MPEG-2, 8-50 Mb/s) is stored for producing the final program, of suitable quality for broadcasting.

Sun Microsystems’ MediaCenter has been employed as the basic platform for the archive server. On its top, an API has been developed for Java clients. Connectivity with the other system components is provided through an FTP server, and an ATM interface.

The catalog server, running on a Sun UltraSparc workstation, downloads new video clips from the archive server, and preprocesses them in order to extract metadata. These metadata are stored and indexed in a relational database built on the Illustra DBMS. Database access is provided to Java clients through an HTTP server and a database connectivity server/driver (Wedji).

The client station allows remote users to perform archival and retrieval operations, using both the catalog and the archive servers. In order not to overload the client station and the network, only low-bitrate video material is used to this end. Once some video segments of interest have been identified, a user can export them to an editing tool, in order to compose a new program. The editing list created by the editing tool is then transmitted to the editing server, which applies it to the corresponding high-bitrate material, in order to produce the ready-to-broadcast final video.
TECHNICAL REPORT
VISION

Distributed Video Editing, Archival and Retrieval

Ruggero Milanese, Alain Jacot-Descombes, Thierry Pun
Frédéric Deguillaume, Lori Petrucci, Michel Cherbuliez, André De Giacomi

Computer Vision Group
Centre Universitaire d'Informatique, University of Geneva
24 rue du Général Dufour, CH - 1211 Geneva 4 SWITZERLAND
Phone: +41 (22) 705-7660, Fax: +41 (22) 705-7780
E-mail: FirstName.LastName@cui.unige.ch