A Case Study of Technologies for a Digital Video Archive System

Hsiang-An Wang, Guey-Ching Chen, Chih-Yi Chiu, Yen-Chun Lin*

Institute of Information Science, Academia Sinica, Taipei, 115, Taiwan
{sawang, ching64, cychiu}@iis.sinica.edu.tw
*Department of Computer Science and Information Engineering,
National Taiwan University of Science and Technology, Taipei, 106, Taiwan
yclin@et.ntust.edu.tw

Abstract. A digital video archive system differs from a regular digital archive system in that it can manage multimedia resources, such as video and audio contents, as well as metadata textual documentation. Thus, a complete digital video archive system must combine various technologies of multimedia searching and presentation. This paper takes the Development of Value-Added Application on Video Archive, a sub-project of the National Science Council's digital archive application project, as a case study to demonstrate the role of information science technologies in developing digital video archive systems and digitizing video and audio resources. By sharing our experiences and the technologies developed in our research, we expect to provide digital content providers and researchers with guidelines for the design and development of digital video archive systems and value-added video/audio data. We also highlight some difficulties and suggest possible solutions to assist in the future development of digital video archive systems.

Keywords: nonlinear online editing, video archive, video digitization

1   Introduction

The Digital Museum of Taiwan's Social and Humanities Video Archive is an applied research project of video and audio archives [1][3]. Its main purpose is to offer free public access to a digital library of 6,000 volumes (3000+ hours) of 16mm and Beta cam video footage. This video content was produced or collected by Dr. Daw-Ming Lee, an associate professor of Taipei National University of the Arts (TNUA). The Development of Value-Added Application Video Archive extends this digital library of excellent video footage and integrates it with a commercial process by adding membership management, online trading and editing of video/audio resources, and data integration across web sites. Users can browse the video/audio database and purchase segments of the content, which is our goal of adding innovative value to a regular video/audio digital database.

  The Development of Value-Added Application on Video Archive is funded by the National Science Council (NSC) Digital Museum project. This is a collaborative project between the TNUA, the Institute of Information Science (IIS) of Academic

Sinica, and Wordpedia Co. Ltd. The TNUA is responsible for digitalizing video/audio data, constructing metadata, designing data content and user interfaces, and the visual presentation of information. Meanwhile, the IIS is responsible for providing and integrating information technologies, and building metadata databases and management systems. It is also responsible for developing sub-systems for the following: video/audio data format transformation, shot detection, metadata searching, audio searching, online editing of video and audio data, data exchange over the Internet, integration of encryption and security certification of members' data, and the environment for the integration and distribution of data streaming. Finally, Wordpedia is responsible for the membership management system, the design of trading processes, and the integration of commercial models.

In this paper we present an overview of the information technologies for video digital archives that we developed in this research project. In particular, we focus on the processes of integrating current video/audio digital libraries with external commercial websites in order to add value to the libraries. We describe these processes and provide owners of digital content with information about how to add value to their content and how business models can be integrated with such content. We also provide some guidelines for future research and the development of techniques related to digital libraries.

The remainder of this paper is organized as follows. Section 2 describes the video digitization process and technologies. Section 3 details the system architecture, including the search engine, database management, and e-commerce trading. Section 4 describes implementation issues, together with our experimental results. Finally, Section 5 concludes this paper and gives future research directions.

2   The Digitization Process and Technologies

The first challenge we face in constructing a video digital archive system is how to establish a digitization process that preserves consistency along the workflow. In this section we describe the process and technique of digitization, and describe our approach by comparing it with the Open Video Digital Library [7], which is similar to our system.

2.1   Video/Audio Digitization Overview

Our digitization process is illustrated in Fig. 1. First, TNUA digitizes Betacam tapes into MPEG-2 files via a video capture card. The files are then transferred to the computer center at Academia Sinica as backup copies. The research team at TNUA then uses the Digital Video Archive System (DVAS) to store the metadata in the database. In the next step, the team at IIS downloads the MPEG-2 files for further processing.

To process visual data, we convert MPEG-2 to WMV (Windows Media Video), a streaming file format compatible with DVAS. In addition, the Shot Detection System performs shot detection of the MPEG-2 files and outputs the results as an XML file containing the time points where scene contents change drastically. The Video

Abstracts Extracting System then extracts a 2-second segment of content from every shot detected. It combines and converts these segments into a video abstract in MPEG-1 format, which allows users to preview the video content. Meanwhile, based on the shot detection results, the Key Frames Extraction System extracts the frame from every shot-change to construct a JPEG image file for static display.

To process audio data, we apply the "Audio Extraction System" to filter out MP2 audio data from the MPEG-2 files produced in the previous step. The re-sampled MP2 audio data is converted into the WAV format. With these WAV audio files and the shot detection result, the Voice Recognition System splits the audio data into segments for the voice recognition services supported by DVAS.
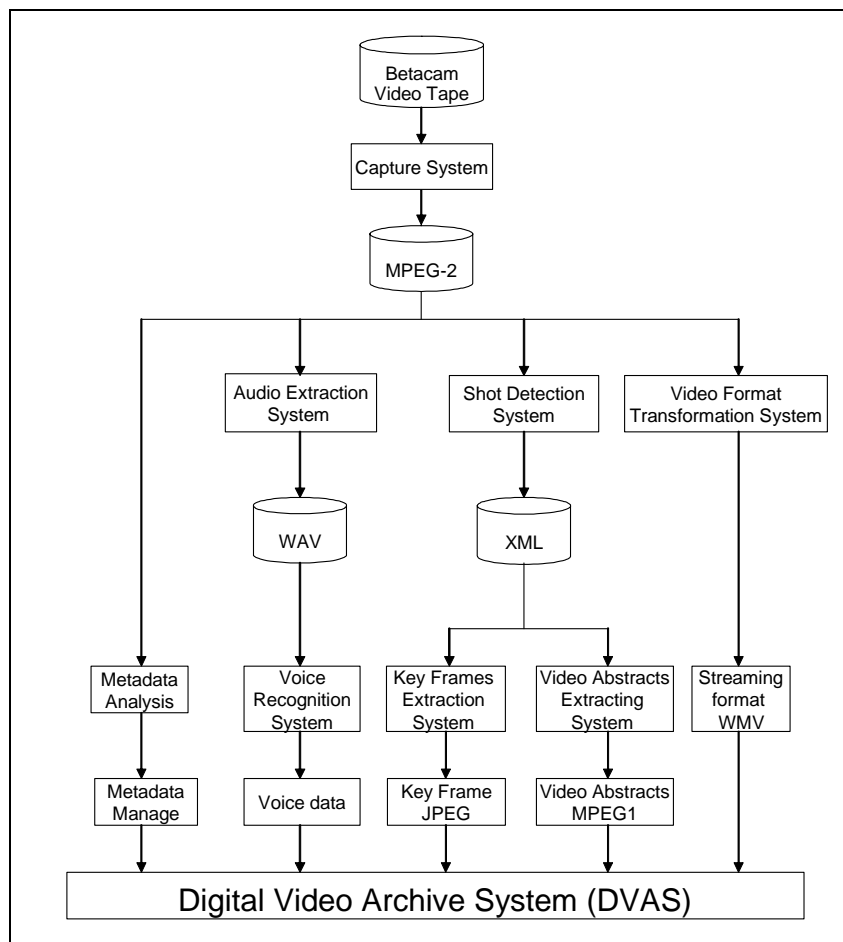

Fig. 1. The video/audio digitization process

2.2　Digitization technologies

The technologies used to digitize video/audio are: video format transformation, video shot detection, video abstract extraction, key frame extraction, streaming data transformation, audio data extraction, voice recognition, and metadata management. The IIS has developed systems for all of these techniques. Due to space limitations, we briefly describe the video format transformation system that we obtained from the Development of Value-Added Application on Video Archive project.

As its name suggests, this system transforms video data into different formats. For example, it can convert MPEG-2 audio files into MPEG-1, WAV, MP2, WMV, or RM format. In addition, if the frame size and bit rate are configurable, it can transform data to a multi-bit rate format. The program can support multi-stage transformation of multiple files into different formats simultaneously, which is very useful for tasks involving multi-staged transformation of a large number of files. The Video Format Transformation System was developed with Borland C++ Builder 6.0 and Microsoft Visual C++ 6.0, using open source software packages FFMPEG [6], Helix DNA Producer SDK [14], and Microsoft Media Encoder 9 Series SDK [15].

2.3　Open Video Digital Library and our method

In 2002, Marchionini and Geisler published the Open Video Digital Library (OVDL) [7], an integrated system for processing data in digital video archives. The digitization flow of this system consists of the following steps:
(1) Digitize NTSC or BetaSP tapes.
(2) Extract key frames via MERIT software [13].
(3) Prune key frames manually to identify representative key frames.
(4) Annotate keywords manually for video and audio content.
(5) Store digitized video files into a disk array on the server side.

OVDL combines a number of key frames to present images of video content. It supports three kinds of search: (1) attribute search, based on the attributes of genres, duration, color, etc. predefined by classification; (2) full text search of bibliographic records, as well as transcripts with text field entries; and (3) a pull-down menu with manually created keywords for keyword searches.

Our method differs from OVDL in two respects: (1) we use a technique for voice recognition that allows users to perform a voice search; and (2) OVDL uses multiple key frames to preview video content, whereas our approach combines a single key frame with WMV streaming of the complete video content to play a video.

3　System Architecture

We have developed two major systems for this project. The first is DVAS, which preserves video metadata and digital video data. The second system is the Member Management System (MMS) for membership management, member certification, online trading, and a payment system at the backend. We have also integrated a mechanism that enables a rights management mechanism. An E-editing Tool

(Nonlinear online editing) is also integrated with DVAS and MMS to facilitate the sale of multimedia video data segments online. We now describe the two systems in detail.

3.1 Digital Video Archive System

DVAS is an environment for video storage and applications. It integrates the mechanisms of data management and the technology of information searching to manage and present video data. The system is divided into two sub-systems based on their functionalities:

(1) The video metadata management system. This is for data entry and the management of metadata of digitized videos, including information about videotapes, continuous scenes, shots, and audio, visual and text content.

(2) The video search system (VSS). In addition to supporting metadata searching, this sub-system integrates the functions of voice data searching developed by IIS. To support real-time online viewing of videotapes, reduce the need for high network bandwidth by users, and protect intellectual property rights (e.g., to prevent illegal copies), this system employs a streaming technique to play the video/audio content of videotapes. The video search system is shown in Fig. 2.
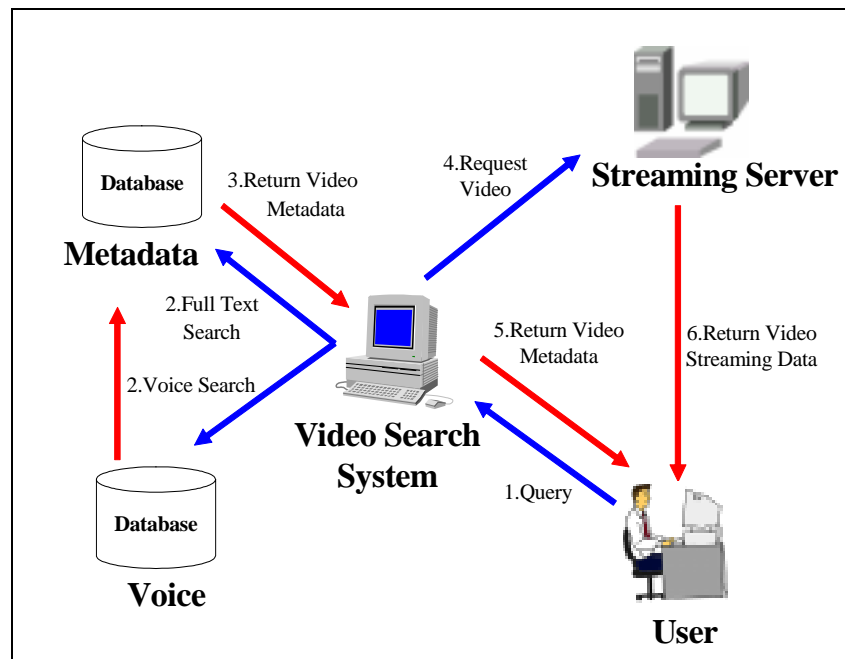

Fig. 2. The architecture of the video search system

3.2 E-Commerce Trading System

This system plays the role of an e-commerce retail store. It combines digital content providers and acts as an agent to sell digital content online. In addition, it supports online trading among users, and mediates between users and digital content providers. The system is comprised of three parts (as shown in Fig. 3):

(1) The member management system (MMS). This is responsible for the management of members' profiles, control of members' rights, shopping cart, and a personal database of videotapes.

(2) The trading management system (TMS). This manages online trading. Its primary function is to provide members with services, such as online shopping (e.g., shopping cart), the cost of merchandise, and online payment facilities. TMS is integrated with financial institutions (e.g., banks and credit card companies.) to facilitate online payments.

(3) The cross-system security certification mechanism. DVAS and MMS have totally different functions and operate independently. We use a mechanism, called Cross-System Security Certification, to integrate the two systems and provide a common mechanism of rights management that allows users to operate on both systems without having to login twice. We use SSL to protect information transmitted between the two systems, and MD5 encoding to protect members' data. MD5 also acts as a guideline for security verification. The architecture of the mechanism is shown in Fig. 4.
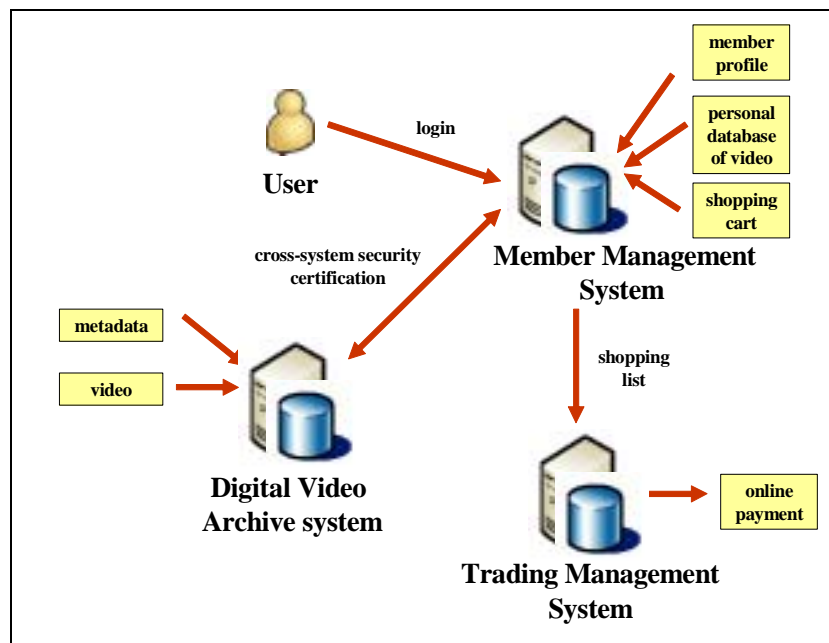

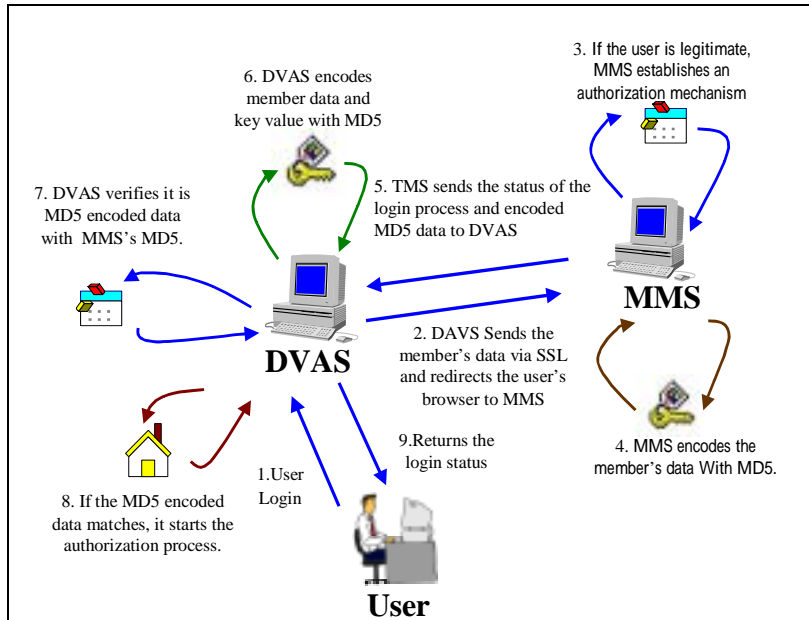**Fig. 3.** The E-Commerce trading system framework

Fig. 4. Cross-system security certification

As mentioned above, the cross-system security mechanism allows users to login from either system. Although the systems' workflow and login processes differ, the dissimilarity is not significant. The login sequence of DVAS is described below.

(1)   A user logs into DVAS, and enters his membership data.

(2)   DVAS sends the data via the SSL mechanism to the MMS and redirects the user's browser page to the MMS

(3)   Upon receipt of the data, the MMS checks it against the database and determines if the user is legitimate. The MMS then establishes an authorization mechanism if the user's legitimacy is confirmed.

(4)   The MMS encodes the member's data and key value with the MD5 algorithm.

(5)   The MMS sends the status of the login process and encoded MD5 data back to DVAS.

(6)   DVAS encodes the member's data and key value with MD5 for verification purpose in later steps.

(7)   When DVAS receives the login status, it checks if the user is legitimate. If this is confirmed, it verifies that the MD5 encoded data in both systems match. If the data does not match, or if the user is not legitimate, go to Step (9).

(8)   If the data matches and the user is legitimate, DVAS starts the authorization process.

(9)   DVAS sends the login status (succeed or fail) back to the user.

3.3 The nonlinear online E-editing Tool

To allow users to preview the video content stored in DVAS and clip the segments they want to purchase, we developed a software tool called the E-editing Tool. This allows users to clip segments of a video online and save them in the MMS, the situation of which is shown in Fig. 5.
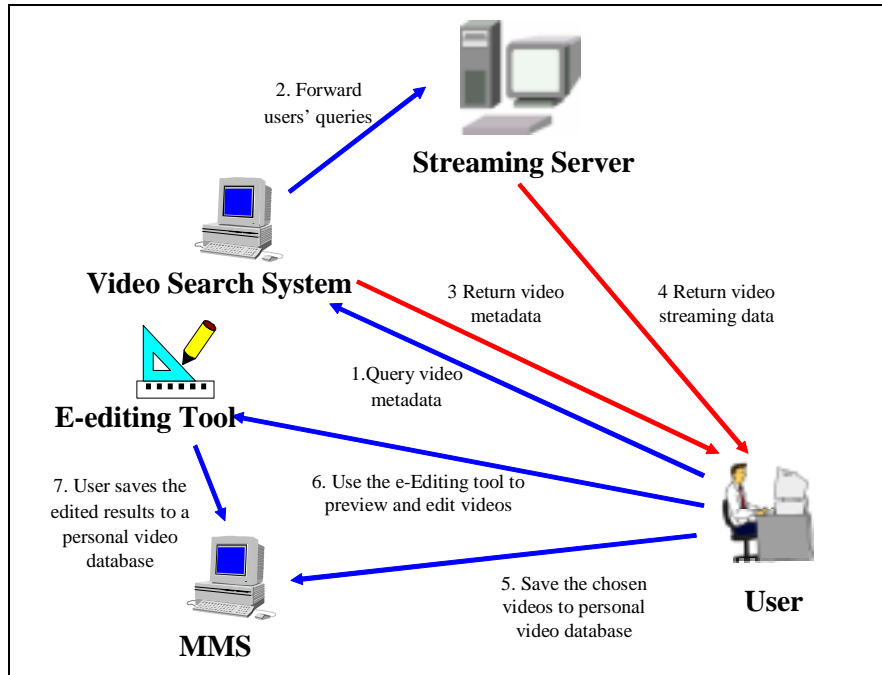

Fig. 5. The situation for using the E-editing Tool

We now describe the sequence of the E-editing tool.
(1) After logging in, the member submits a query about the metadata of videotapes to the Video Search System (VSS) through a browser.
(2) VSS forwards the queries to the streaming server with a request to send the data back to the user.
(3) VSS sends the video metadata back to the user.
(4) The streaming server sends the video data back to user.
(5) The user saves the requested videos in a database in the MMS.
(6) The user selects the videos he wants to edit from the personal video database, and uses the E-editing Tool to preview and edit the videos.
(7) After Step 6 is completed, the user saves the edited videos in his personal video database in the MMS.

4 System Implementation

4.1 The Digital Video Archives System

DVAS has a 3-tier architecture: Apache and Tomcat Web serve as application server-tiers, and Oracle 8.1.6 serves as the database-tier. We use the Linux Red Hat 7.3 operating system on the server-tiers. The web pages were developed with JSP technology and Java Beans. We also integrated the search engine with a streaming server. DVAS uses the Microsoft Windows 2000 Server operating system. We use the Microsoft Media Server as the streaming server and the streaming format is WMV. The hardware comprises two 1U servers with Intel Xeon processors to run DVAS and the streaming system. There is also a disk array that stores video abstracts and key frames. The total volume is 840 GB. Original videotapes are backed up with large tapes.

4.2 The E-Commerce Trading System

This system also has a 3-tier architecture: Microsoft Internet Information Services as the application server-tiers and Microsoft SQL Server as the database-tier. Servers are run on the Microsoft Windows 2000 platform. The web pages were developed with ASP technology.

4.3 The Cross-System Security Certification Mechanism

This mechanism consists of two systems developed on the Java/JSP/Linux platform and the ASP/Windows platform, respectively. Although the system platforms and programming technologies are different, they are both operated on a web-based architecture. Thus, to integrate our major systems, we used web-based methodologies to design the cross-system security certification mechanism.

We tested the following two methods for data transfer and protection:

(1) We used the RC4+Base64 encoding algorithm to encrypt a member's data and transferred it with the HTTP protocol. The receiver then decrypted the data with the RC4+Base64 decoding algorithm to obtain the member's data.

(2) Combining additional information, we encoded a member's data with the MD5 algorithm and applied the SSL protocol to protect and transfer the data to the receiver. Upon receipt of the member's data containing the additional information, the receiver also encoded it with the MD5 algorithm. This procedure allowed the receiver to check that the data was the same as that transmitted by the sender, and thereby verify that it had not been captured and modified during transfer.

With method (1), it is necessary to forward a member's data from the login web page to the page that actually does the encoding. However, the data is transferred as a string without encoding, so it can be hijacked easily. We therefore chose method (2)

to protect data during transfer. Table 1 lists the results of the comparison of the two methods.

Table 1. Comparison of cross-system security certification mechanisms

Method to protect data during transfer	Encode with RC4 + Base64, send with HTTP	Encode and transfer with SSL, verify with MD5 encoding
Has methodology to verify data	No	Yes
Overall security level	Low	High
Remarks	Needs to forward member's data to encoding web page; thus there is no protection during the process	Member's data protected during transfer and verified with MD5

The mechanism can be performed on Windows or Linux platforms. Since the ASP and JSP technologies supported by these two platforms differ, we found that MD5 encoding produced different results on the two platforms. The reason was that the text processors of the platforms were configured to encode Big-5 or Unicode (UTF-7, UTF-8). After setting both platforms to Big-5 encoding, the problem was resolved.

4.4 The E-editing Tool

The primary function of the E-editing Tool is to enable a user to mark time frames on the video he wants and save segments of the video to his personal shopping cart in the database. Based on these time frames, video content providers can clip the segments from the original videotape, edit and process them, and sell them to the user.

The E-editing Tool was developed with Visual C++ 6.0. Currently it supports Microsoft Windows 2000 XP and runs on both the Internet Explorer and Netscape web browsers. The program requires 3.5MB memory for execution. We have defined a new file type with file extension .edl and MIME Type application/edit, which has to be started up from the web browser. Fig. 6 shows a screen shot of the E-editing interface.
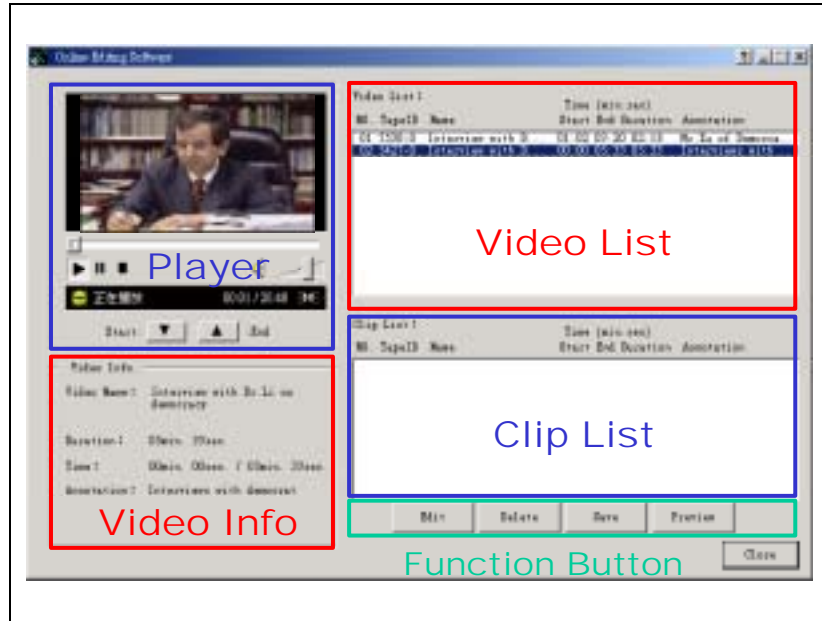
Fig. 6. The E-editing interface

5   Conclusions and Future Work

The Development of Value-Added Application on Video Archive project integrates the research results of the Digital Museum of Taiwan's Social and Humanities Video Archive project and incorporates technologies of online trading processes, member management, cross-system security certification, and the E-editing tool. This project, which has been in existence for 3 years, developed from digitizing original negatives; to the formulation, entry and management of metadata; to video searching; and finally to our goal of value-added applications and online trading. A complete workflow of digital archive applications has been established and verified. Furthermore, the project has yielded productive research results, especially in the development of technologies applied in value-added applications, which we have described in detail.

We now indicate some future research directions.

(1)   Improve the workflow of automatic digitization. Currently, certain steps, such as the selection of key frames, are sometimes performed manually. In the future, we will integrate different systems to reduce manual processing, which will reduce errors and improve the system's overall performance.

(2)   Improve the technology for voice searching, as the accuracy rate of voice recognition in video files is not very good.

(3)   Integrate content-based retrieval techniques. Content-based visual retrieval has received a great deal of attention from researchers in recent years [10-12]. Users can use color, texture, shape, and motion as visual cues to search perceptually

similar video clips. Therefore, the integration of text and content-based retrieval would provide a more flexible way for users to make queries.

(4) Integrate watermark technologies for digital rights management. Currently, we often add watermarks to protect digital rights, but this is very costly in terms of computing time. Also, watermarks spoil parts of a frame. Currently the robustness of watermarks against attack is not good enough to guarantee security. Also, the technology of digital rights management has been discussed and developed. We will continue research in these two directions and develop relevant technologies.

(5) Extended functions of the E-editing Tool. A hypermedia concept will be integrated into our E-editing Tool [9]. After selecting a video segment, users will be able to edit the hyperlink of the video segment, such as other detailed video or description text files about the video segment. When users browse a sequence of video segments, the mechanism will provide a convenient way to view related media content.

(6) Establish a business model. We have finished a simulation test and demonstrated the feasibility of integrating DVAS with a business retail model. Digital content providers and business executives should consider integrating E-commerce operations to establish a business model for digital content retailing.

The complete project plan of the Development of Value-Added Application on Video Archive and the Digital Museum of Taiwan's Social and Humanities Video Archive has a vast amount of high quality content and employs several techniques to process and present it. Due to space limitations, we have only described the technologies and system architecture of DAVS. Other topics, including video content, e-Learning systems, the design of metadata, and the development of information science technologies are not discussed in this paper. In the future, we will continue in-depth research and development of these areas in order to construct a more advanced digital museum.

Acknowledgements

References

[1] H. A. Wang, C. W. Fann, and J. M. Ho, Digital Video Archive System: a Case Study on Digital Museum of Taiwan's Social and Humanities Video Archive (Chinese), *The Second Workshop on Digital Archives Technologies*, pp 57-62, July 2003.

[2] D. M. Lee, Development of Value-Added Application on Video Archive, *NSC Project Report*, April. 2003.

[3] Taipei National University of the Arts and Academia Sinica, Digital Museum of Taiwan's Social and Humanities Video Archive, web site: http://twemovie.iis.sinica.edu.tw/

[4] Carnegie Mellon University, Informedia, digital video understanding research, http://www.informedia.cs.cmu.edu/

[5] D. Green, Beyond word and image: networking moving images: more than just the 'movies', D-Lib Magazine, 3(7/8).

[6] FFMPEG, FFMPEG Multimedia System, http://ffmpeg.sourceforge.net/

[7] G. Marchionini and G. Geisler, The Open Video Digital Library, *D-Lib Magazine*, 8 (12), http://www.open-video.org/

[8] M. Christel, Visual Digests for News Videos Libraries, *ACM International Conference of Multimedia*, Orlando, FL, Nov. 1999.

[9] F. Shipman, A. Girgensohn, and L. Wilcox, Generation of interactive multi-level video summaries, *ACM International Conference on Multimedia*, pp. 392-401, Nov. 2-8, 2003.

[10] S. W. Smoliar and H. J. Zhang, Content-based video indexing and retrieval, *IEEE Multimedia*, 1(2), pp. 62-72, 1994.

[11] S. F. Chang, W. Chen, H. J. Meng, H. Sundaram, and D. Zhong, A fully automated content-based video search engine supporting spatiotemporal queries, *IEEE Transactions on Circuits and Systems for Video Technology*, 8 (5), pp. 602-615, 1998.

[12] A. K. Jain, A. Vailaya, and X. Wei, Query by video clip, *Multimedia Systems*, 7(5), pp. 369-384, 1999.

[13] Maryland Engineering Research Internship Teams, http://www.ece.umd.edu/MERIT/

[14] Helix DNA Producer SDK, https://producersdk.helixcommunity.org/

[15] Microsoft Media Encoder 9 Series SDK, http://www.microsoft.com/windows/windowsmedia/mp10/sdk.aspx