

Writing Web Documents about Films

Wayne Wolf

Department of Electrical Engineering
Princeton University

Abstract

This paper describes our experiences our experience in building and using a Web-based video library designed for educational use. The CAETI Internet Multimedia Library's initial audience is K-12 schools; most of the content of our library comes from news and politics-related historical footage. The video library is a good tool not just for content but also for acquiring visual literacy. Politics and news are very good subjects for video collections because television has come to dominate news coverage and politicians use television to campaign and govern. Beyond the factual content of video/film footage, studying this material can help students understand how politicians, news outlets, and others make use of techniques borrowed from Hollywood and other sources to get across their message.

1 Introduction

This paper describes our experiences our experience in building and using a Web-based video library (www.videolib.princeton.edu) designed for educational use. The CAETI Internet Multimedia Library's¹ initial audience is K-12 schools; most of the content of our library comes from news and politics-related historical footage. We designed the library as a tool in which students can gain both historical/political literacy and visual literacy. Politics and news are very good subjects for video collections because television has come to dominate news coverage and politicians use television to campaign and govern. Beyond the factual content of video/film footage, studying this material can help students understand how politicians, news outlets, and others make use of techniques borrowed from Hollywood and other sources to get across their message.

Video libraries offer some important tools which make moving image material much more useful than is possible with traditional analog videotape. First, video libraries deconstruct the video in order to create summaries. Summaries can be used in either browsing or search mode, but

summaries help users comprehend videos in ways that are difficult using fast-forward and freeze-frame. Second, video libraries allow users to incorporate video material into documents, allowing the reader to comprehend the commentary much more easily than is possible when one reads a traditional paper describing a film.

This paper describes the basic mechanisms of our video library and how they may be used in education. The next section describes the user interface and basic architecture which supports that interface. Section 3 briefly describes the types of material that we have collected for our video library. Section 4 describes the ways in which video and video libraries can be used in education. Section 6 summarizes some issues in the support of multimedia documents on the Web.

2 Library Interface and Architecture

There are two main styles of use in any digital library system: **browsing** and **search**. A browsing system provides some sort of summary of one or a collection of holdings, from which the user can select some section of interest; a search system finds holdings which meet some criterion and presents them to the user for selection. The state-of-the-art in video and image processing is such that image search mechanisms do not provide the accuracy of search that can be found in modern text search systems; as a result, browsing plays a larger role in video libraries than in text-oriented digital libraries.

The basic browsing interface for our system, as for most video libraries, is the **storyboard**, which is a selection of still images (often called **keyframes**) which capture important visual information^{3,2}. The images are arranged from left-to-right and top-to-bottom so that they can be read in a manner similar to text. A sample storyboard from the start of a NASA documentary is shown in Figure 1. The storyboard is commonly used in film production and is a sufficiently natural representation of moving images that most users easily adapt to it. Our *video library* offers a click-and-play interface, whereby users can click on an image in the

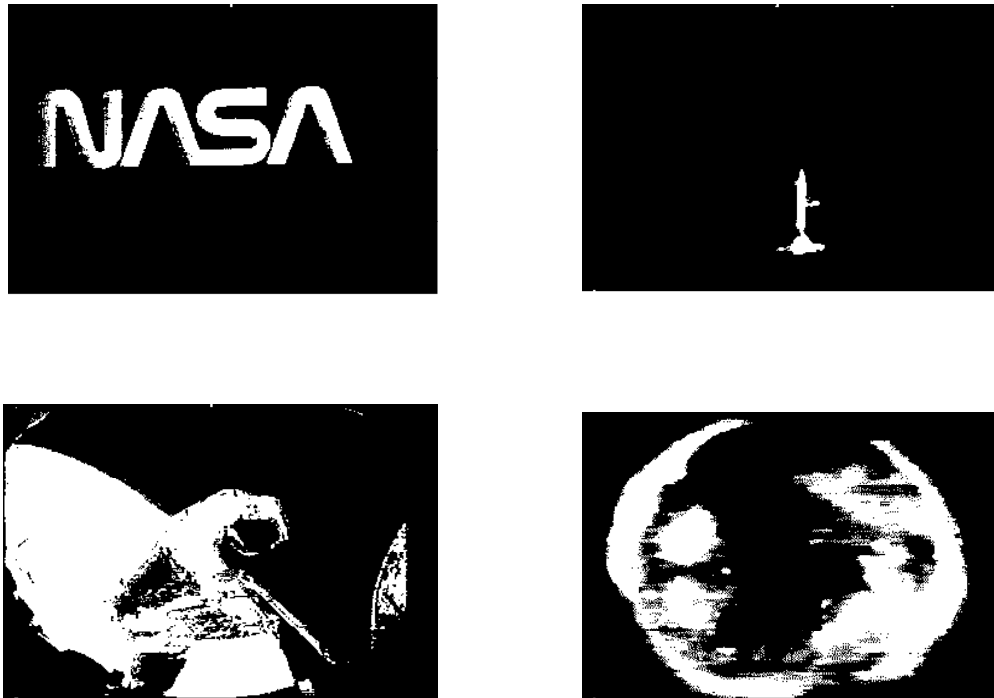


FIGURE 1. A sample storyboard.

storyboard and cause the corresponding video to be downloaded and a video player positioned to that frame in the video for playback.

There are three distinct search interfaces, depending on the type of material to be searched:

- A simple textual interface allows searching of the card catalog for the library. The result of search over keywords is a Web page with hot-linked author/title entries which matched the search.
- A somewhat more sophisticated textual interface is provided for videos which have closed-captioning transcripts available. The transcript is shown as a text page with a search string window; each closed-caption line has a time tag, which is currently hidden from the user. Given a search string, the system returns an HTML page with image/text pairs. The storyboard keyframe closest in time to the time of the caption is displayed along with the text line providing the match; the text line is hot-linked to the point in the storyboard corresponding to that image so that the user can find the position of the quote in the visual timeline.
- Our prototype image feature search system⁴ allows users to search for basic categories such as sky, foliage, flesh tones, etc. This search is conducted over the com-

plete database of keyframes. The result of search is an HTML page of images, each of which is hot-linked back to the storyboard from which it came.

The basic architecture of our video library system, as illustrated in Figure 2, is a basic Web-based client/server system augmented with database generation software. The client end of the system is largely based on standard Web servers and browsers to maximize cross-platform compatibility, with two exceptions: the click-and-play feature of storyboards, requires a Web browser plug-in which currently operates only under Windows 95; and annotation playing may require either Java code or a plug-in.

The structure of our video library allows users to write Web documents which refer to video library objects via URLs. Users can refer both to videos by position, to extracted keyframes, or to the closed-captioned text. The techniques for referring to keyframes and closed-captioning are standard HTML and easy to use; the methods for video reference are not yet standardized and more challenging for novices.

3 Collections

Having a significant collection of useful material is an important component of video library research, since it is difficult to collect meaningful usage statistics or anecdotal results on user interfaces without material that is of practi-

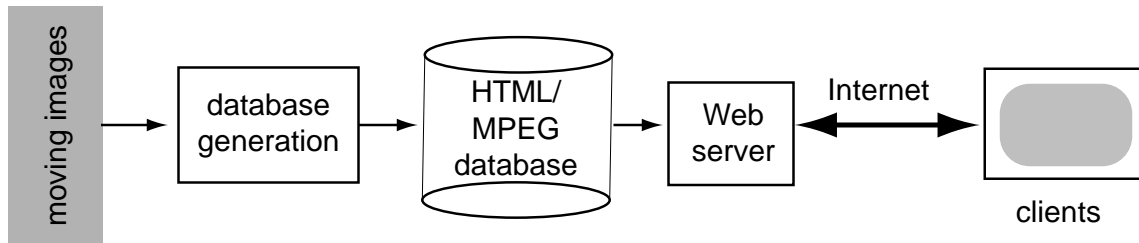


FIGURE 2. Architecture of our video library.

cal interest to real users. Our collections have focused primarily on politics and news, although we have also collected some science material. We have collected material from several sources:

- The Presidential libraries, operated by the National Archives and Records Administration (NARA) provide a wealth of material that is of great historical importance. We have filmed material and radio recordings dating back to FDR, but the amount of moving image coverage of presidents greatly increased starting with Johnson. This material is in the public domain.
- The Universal Newsreel collection at the National Archives provides a wealth of news and entertainment information from the 1930's to the 1960's. The Universal Newsreels were put in the public domain by Universal Studios. NARA has transferred most of the newsreels to video.
- NASA has a significant collection of historical footage and documentaries on space and aviation. This material is in the public domain.
- We have made use of some television news stories donated by CONUS, the largest supplier of television news to independent TV stations, for non-commercial use. Some of this material comprises longer stories, but most are in the 3-5 minute length common for television news.
- We have collected some political advertisements broadcast on television; we make these available for scholarly use only.

4 Writing about films

There are many different ways that moving image material can be used in education:

- English courses may use films to introduce novels and perhaps to compare the book with the film version.
- Historical footage may be used in a number of subjects.

- Experiments may be recorded and reviewed in areas ranging from physics to psychology.
- Of course, films are the subject of film studies courses.

This would seem to give us a very broad range of applications which must be supported by a video library. However, we can divide these uses into two major categories. First, teachers and students may want to study the content of the film, such as a space experiment conducted by astronauts. Second, they may want to study the way in which the film makes its point, as in a political advertisement or a film studies course. Of course, it is possible to consider both aspects at once.

Video libraries are quite capable of supporting both types of educational film usage and with significant advances over analog video. Analog video's disadvantages include:

- It is difficult to find a small clip, since watching the videotape is the only way to skim its contents.
- Videotapes cannot be annotated directly and card-based annotation systems have a number of difficulties.
- Videotapes cannot easily be used by several people at once.

The most significant disadvantage of Internet video libraries is bandwidth. It is possible to download only very short clips in time spans consistent with human attention spans. As a result, use of the video library must be adapted to minimize the amount of video which must be downloaded and viewed. However, that goal is consistent with the typical user's goal for a video library, since users typically do not want to unnecessarily view large amounts of video.

Historical material is one type of material in which we are often concerned with the content as much as the form. We have found that, for current events and historical material, using stills to illustrate points is almost always satisfactory. The visual content can usually best be conveyed by stills. Live playback for this type of material is often most important for sound—consider, for example, the voice of FDR or Reagan. As a result, viewing the storyboard and listening to selected portions of the speech are often more

than sufficient. These steps can be performed with much lower bandwidth requirements and, therefore, much more quickly than watching the full video.

A Presidential speech is an example of an audio-rich historical video. Some amount of information is conveyed by the gestures of the speaker or the reactions of the audience; in particular, Reagan's practice of introducing a typical American with an inspiring story and introducing them to the audience added a cinematic dimension to such speeches. However, most of the content is in the text and the reading of that text. The text of Presidential speeches is released fairly quickly after the speech by the National Archives, and that text can be timed to the video with relatively little effort even in the absence of a real-time closed-captioned transcript. The transcript and the audio track together can provide the student or teacher with a great deal of information at relatively low bandwidth cost.

A physical experiment is a category of film for which motion is usually more important. However, it is usually not necessary to watch a long video to capture the essence of the experiment. For instance, several NASA documentaries contain a series of experiments or explanations of phenomena. Each experiment or explanation takes only a few minutes. A teacher may want to discuss each point separately in order to hold the class's attention and to maximize retention. In that case, the video library should be able to help the teacher or student find the clip of interest quickly and make it possible to download that clip only. The clip can be found using a variety of techniques, including browsing the storyboard, searching keyframes, searching the closed-caption transcript, or by watching a few short segments of the video. All these techniques can be used to minimize the required bandwidth.

In the case of film studies, it is often as important to study the way in which the film is made as it is to consider the overall performances, etc. There are some basic techniques, developed in the early years of the 20th century, which filmmakers of all types and nationalities use to convey information:

- Framing of the subjects conveys information about individuals—for example, a close-up generally conveys an intense emotion. Framing also conveys information about the relationships between individuals through the positioning of the subjects, their relative size in the frame, etc.
- Cuts in film are somewhat analogous to periods in sentences. The positioning of cuts can convey significant information, as can the comparative qualities of the two shots on each side of the cut.

- More complex transitions between shots also signal information—for example, a dissolve implies the passage of time.
- In dramatic material, music plays a critical role in cueing the intended response to the shot.

The analysis tools which are part of the video library are very useful in deconstructing films. Storyboards are extremely useful in studying the composition and visual structure of films, since the storyboard summarizes the shots and the relationships between them. Relationships between shots may be missed when viewing the film: some shots may be short and transitions may be too swift for the viewer to fully analyze them. However, if the storyboard can be browsed and critical sections can be viewed at leisure, the viewer is much more likely to fully understand the techniques used by the filmmaker to convey the message. The same can be said for acting—deconstructing acting can be a valuable tool in understanding the technique used to achieve an effect.

5 Writing about political film

In order to develop content for our video library, we engaged a recent graduate, David Calone, to select and collect important videos from Presidential libraries as well as write a document describing those videos for people who may want to use them. (David worked under the supervision of Prof. Fred Greenstein of Princeton's Woodrow Wilson School.) The resulting document, "The Modern Presidency," which is available on our Web site, illustrates one typical use of videos for education. (We did not have sufficient disk space to digitize all the videos collected before David's departure, so we could not integrate the text and videos as fully as we had hoped at the start of the project.)

In this sort of material, full-motion video is not critical to an understanding of the material—speech carries the primary message. While it may be useful to have an understanding of the milieu, the other people in attendance, etc., much of this information can be gleaned from still images. As a result, we found it sufficient to illustrate the text with keyframes which gave the reader a better feel for the content of the video; we also made a link to the video library's content page for each video to give the reader access to the full-motion version. Having the full-motion video is still important to give the user a full understanding of the content, but the document does not need to be closely coupled to that video.

An advertisement is a good example of a type of material for which analysis of the form is as important as the analysis of the content. Even with advertisements, however, full-

motion video is not essential for understanding. In fact, the graphic design of most ads is best perused in storyboard form. Advertisements have a very dense structure designed to pack a maximum amount of information into 30 or 60 seconds; advertising people often refer to television ads as 60-second feature films. The ads must also send their message in a way that minimizes the viewer's conscious analysis of the point of view being presented. Once again, having the full-motion video is important for completeness, but users can write about the video without making detailed reference to the full-motion version.

An important advantage of such applications is that normal Web techniques are sufficient for developing the annotation and commentary for the video content. Since the details of frame-by-frame content are not necessary, text and still image presentation styles commonly found on the Web are perfectly sufficient in these cases. This simplifies the learning and use of the system for both writers and readers.

However, some types of content and analysis definitely require the detailed study of full-motion video. An example from political science is the study of body language. Some aspects of film studies also require detailed analysis, such as close analysis of acting style or camera movement. In such cases, annotation for full-motion video is important to help the user convey meaning to others. We have experimented with a textual annotation system for full-motion video (which provides subtitles), and we plan to experiment with graphical annotation methods as well.

6 Writing multimedia documents on the Web

The distributed nature of the WWW makes it possible for writers to take direct advantage of documents in digital libraries. Writers should be able to refer either to complete documents or to elements of documents. We have had good success with references to keyframes in Web documents. Since the keyframes are pre-extracted (our library stores keyframes in JPEG format), they can be referred to as discrete entities with their own URLs. Keyframes are often useful in illustrating a point in a document.

However, referring to audio and video clips is still difficult. There are no standards on the Web for referring to sections of streaming media. Our plug-in viewer uses its own conventions for referring to a position in a video. However, it appears that, at this moment, each video library project is defining its own mechanisms for defining clips. Standards for referring to audio and video clips would be relatively simple but an important step in making multimedia documents sharable and usable in academic work.

Web-based interfaces also introduces compromises between interface quality and portability—plug-ins fre-

quently provide much higher performance, more interactive interfaces, but they require different implementations for different platforms. For example, two undergraduate projects implemented different styles of annotation systems: one used Java to display textual annotations with a slide show of keyframes; another used a plug-in to add annotations to MPEG-1 playback. The Java interface was portable across Web browsers and had the advantage of requiring lower network bandwidth, but it was not suitable for annotations which make strong reference to motion. On the other hand, the plug-in annotation player is available only for Windows 95 at this writing.

7 Conclusions

Moving images are useful tools in teaching a variety of subjects from history to physics. Video libraries provide users with two important advantages over videotape: users can find material of interest much more quickly and they can manipulate the material for analysis in ways not possible with analog media. We have had some initial success with using video libraries in academic topics. Along the way, we have identified some important tools that need further development. In particular, better support for distributed annotation and for references to clips are necessary for teachers and students to get the most out of video libraries. However, we believe that video libraries will play an increasingly important role in education.

Acknowledgments

This work was supported by the DARPA CAETI program under a program administered by NRAD.

References

- 1 Wayne Wolf, Yiqing Liang, Michael Kozuch, Heather Yu, Michael Phillips, Marcel Weekes, and Andrew Debruyne, "A digital video library on the World Wide Web," in *Proceedings, ACM Multimedia '96*, ACM Press, 1996, pp. 433-434.
- 2 Wayne Wolf, "Key frame selection by motion analysis," in *Proceedings, ICASSP '96*, IEEE Press, 1996, pp. 1240-1243.
- 3 Michael Philips and Wayne Wolf, "Video segmentation techniques for news," in *Multimedia Storage and Archiving Systems*, SPIE, vol. 2916, 1996, pp. 243-251.
- 4 Hong-Heather Yu and Wayne Wolf, "A visual search system for video and image databases," in *Proceedings, ICMCS '97*, IEEE Computer Society Press, 1997.