

An Object-Based HyperVideo Authoring System

Han-Bin Chang, Hui-huang Hsu, Yi-Chun Liao, Timothy K. Shih, Chia-Tong Tang

Multimedia Information Network Lab

Department of Computer Science and Information Engineering

Tamkang University, Taiwan, ROC

E-mail: tshih@cs.tku.edu.tw, hhsu@cs.tku.edu.tw, g8190221@tkgis.tku.edu.tw

Abstract

Hypervideo is a new viewing way for films just as hypertext on the WWW. In the literature, many standards and technologies have been developed for the hypertext. However, the present video processing methods cannot provide a proper solution for hyperlinks in video. In this paper, we propose an object-based hypervideo authoring system. Video objects can be described by semantic annotation and multistory movies can be produced. The system offers a manual object-based interface for the film producer to select meaningful video objects and annotate them. It also has a video construction function to facilitate the audience to view the annotation of interesting objects in a scene or to select a particular viewing path. It is very useful to show the annotation of a particular object that the film producer wants to highlight in educational videos or commercials. A hypervideo player is also developed to play the video file made by the authoring system.

1. Introduction

The advancement in the video processing has simplified the process for creating high semantic video data. One of the most active research problems is the video-content analysis and retrieval. The digital form allows processing of the video data to generate appropriate data abstractions that permit flexible video-database organization and enable content-based retrieval of video [1]. It is mentioned in many existing researches that to identify syntactic and semantic components in the video material is desirable [2, 3]. But how can this be achieved? When a commercial is shown, how can the audience get the price or specification of a product right away? How can we extract the information related to a certain video content? In distance learning, it is of great desire that extra information or a link to another short video can be added to the video so that the teacher can explain certain learning materials further.

Hypervideo is a new video processing technique similar to hypertext. In a Web page, useful information can be obtained right on a click without time and space restriction through hyperlinks. The available video producing tools cannot perform such hyperlinks on video.

In [5], the hypervideo methodology is implemented with a framework and hypervideo engine called HyperCafé. Also, some outstanding systems have good performance on object detection and video segmentation issues [4]. However, given a video file, a user might not be interested in the whole film, but rather portions of the film. And the audience might be interested in some video objects of certain frames. In this paper, a novel semantic video

authoring system that focuses on assisting the video author in annotation based on video semantic level is presented.

2. HyperVideo and Multistory Video

There are two major components in our system: video object annotation and video structure construction. Video object annotation is an advanced tool to present an object with extra information. In video structure construction, the user can make a hypervideo movie easily.

In the MPEG specification, a data stream called the user-data stream can be defined [7]. According to codec rules, given a video file, we can put additional information as a data stream. Those additional data can be a text, a video clip, a URL link, or a still image. Furthermore, commands are added to control video playing. In Figure 1, the extension data composes of many data units that enable the functions of hypervideo. These user data units are related to the objects or clips in the original video.

The developed video authoring system can efficiently present semantic entries to the viewer. The film director can easily edit the underlying semantic structure of a video following what the film director wish to depict to the audience. The system has the following features.

- **High Compatibility File Format**

The format of the annotated video by the system is not restricted to common video players. The exported video file can be played in a custom-designed player as well as other players (Figure 2). The original

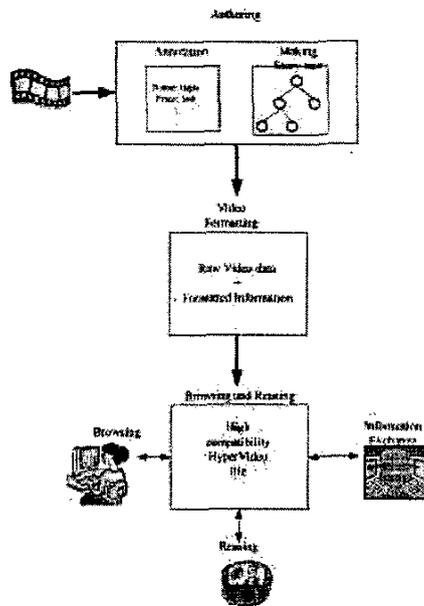


Figure 3: Process in the annotation multistory system

3.1 file format and example

As shown in Figure 4, the raw video data consist of four general clips and one redundant clip and one comment. In distance learning, the instructor

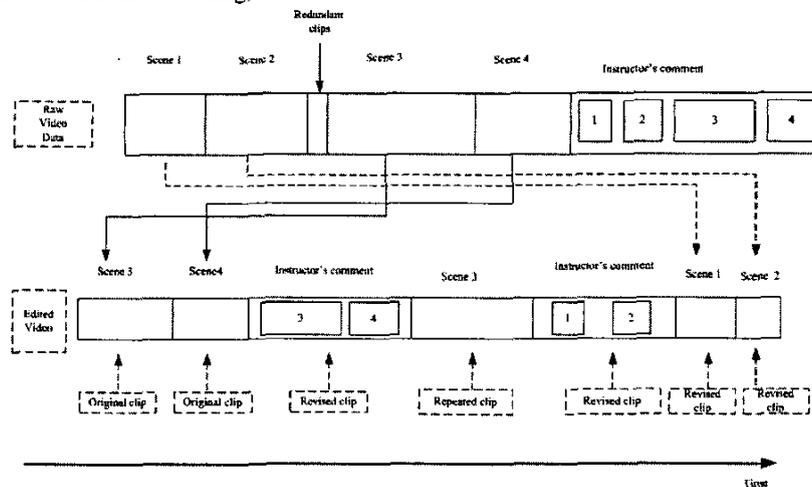


Figure 4: Comparison of the raw video file and the edited video file

3.2 Decoding the video

In our system, the default output video is a MPEG compatible file. Annotation data are regarded as user defined data, and the common video player would not show the extra information unless the player added our control filter (Figure 2). The presentation data and user data are put into one video file.

3.3 Implementation

might notice that the original playing sequence is not suited to certain learners. The instructor can modify the sequence and even delete some learning materials. The editing procedure can be described as follows.

- Step1. Scene 3 and Scene 4 are original, but the presentation order is put in the head.
- Step2. The instructor puts two course illustrations after Scene 3 and Scene 4. The illustrations are cut short to fit time restriction.
- Step3. Because Scene 3 is a key point in this class, the instructor repeats it.
- Step4. The instructor thinks that it is better to make an introduction for the next course. So the original comments of Scene 1 and Scene 2 are edited.
- Step5. Scene 1 and Scene 2 are put in the tail of the video sequence.

There are two advantages in previous description. The instructor can produce many scenarios without cutting the real video file directly. And those scenarios can be attached to the video file instantly.

3.3.1 Authoring tool

Figure 5 shows the user interface of the video editor. Detail description is given in the following.

● Object annotation

Area A is a viewing and drawing area in our system. The user can just load a video file and plays it in the area like using a video player. First the user can use buttons to mark the duration within which the annotated object

will present. Then the user draws a notable region and produces a description for it. This segment is used as a story unit in constructing the multistory and is shown in Area B.

- Story construction

A video storyboard is provided for the user to construct and view the hierarchy of the multistory video in Area D. When a user wants to make a new story tree, he/she can pick pieces of video from Area B or cut new segments from the original file.

How can the user connect these pieces? First, the user needs to choose an annotated region in a segment to be a “branch point.” Then he/she simply drags a line to a destination segment. Notice that the storyboard is not simply a collection of frames. Each image represents a segment of the film. It is impossible for the audience to catch a frame and click it while playing a video. So the annotated region plays on the screen for a period of time. Also, the video producer must define a default playing sequence. Otherwise, the player will play the whole video file.

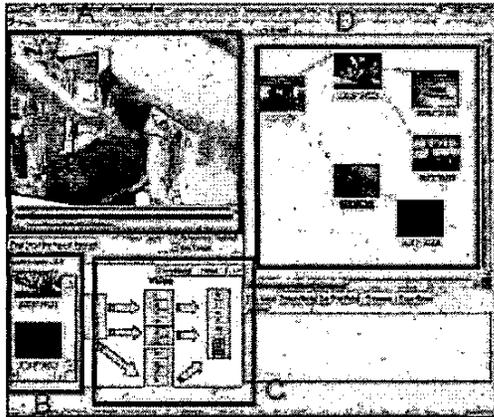


Figure 5: Snapshot of the authoring tool

3.3.2 Player

As mentioned in the previous section, additional video information can be played only in our system. The developed video player is shown in Figure 6. There are a yellow region and a red region that include the additional information. The regions might have been marked as a branch point by the authoring tool. If so, the audience can also click them to jump to other clips. In Area C of Figure 5, an image hyperlink is inserted into the video. So the player presents the inserted image (on the right hand side of Figure 6).

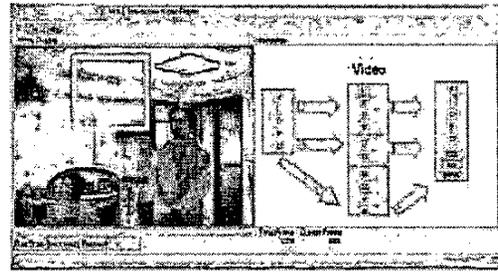


Figure 6: Snapshot of the Player

4. Conclusion

A HyperVideo system that provides a semantic level annotation and constructs the video presentation structure is introduced. The system is different from other video presentation systems. The major focus is on the objects that are interesting and useful to the video producer and the audience. A video player is also developed for the video viewer to view the annotated film efficiently. The prototype of the HyperVideo authoring system can be found at <http://www.mine.tku.edu.tw/video/ivat.htm>

References

1. Dimitrova, N. et al., “Applications of video-content analysis and retrieval,” Vol. 9, Issue 3, IEEE Multimedia, 2002.
2. Wendy E. Mackay and Glorianna Davenport, “Virtual Video Editing in Interactive multimedia Applications”, Vol. 32, Num. 7, Communications of the ACM, July 1989.
3. Yoshiaki Hada, Hiroaki Ogata, and Yoneo Yano, “XML-based Video Annotation system for Language Learning Environment,” Web Information Systems Engineering, Proceedings of the Second International Conference on, Vol. 1, 3-6 Dec. 2001.
4. Correia, P.L. and Pereira, F, “Objective evaluation of video segmentation quality,” Image Processing, IEEE Transactions on, Vol. 12 Issue 2, Feb. 2003.
5. Nitin Sawhney, David Balcom, and Ian Smith, “Authoring and navigating video in space and time,” Multimedia, IEEE, Volume: 4 Issue: 4, Oct.-Dec. 1997.
6. Jim Taylor, “DVD Demystified,” McGraw-Hill Professional (2nd edition), Dec. 22, 2000.
7. J. L. Mitchell, W. B. Pennebaker, C. E. Fogg, and D. J. Legall (Editors), “Mpeg Video: Compression Standard (Digital Multimedia Standards Series),” Kluwer Academic Publishers, October 1996.