VideoStore: A system to store, annotate and share video based content

Nuno Correia, Diogo Cabral

IMG (Interactive Multimedia Group), DI and CITI/FCT, New University of Lisbon, Quinta da Torre, 2829 -516 Caparica, Portugal

This paper describes the VideoStore project that has the goal of developing a Web based system to store, transmit, search, and share the contents of courses captured in video. The access will be done using PCs or PDAs for parts of the content, in a wireless broadband network environment. The project is initially tested with contents from courses of the Computer Science Department, and later extended to other areas. The results will also be used in an ongoing national initiative to promote a Distributed Master Degree in Computer Graphics and Multimedia. This degree will have lectures from several universities across the country and will be received, in the same universities. Starting from existing tools, including video streaming servers and capture boards, VideoStore will build new interfaces and access methods to video based lectures stored in digital formats. The paper describes the project objectives, the system architecture, and the interfaces under development.

Keywords video browsing; learning environments; video based interfaces; video annotation

1. Introduction

Integrated multimedia documents, where video plays a dominant role, are becoming increasingly pervasive in different environments and domains. However, video still poses difficult questions, to store and access it properly, when used in environments with many users, especially over the Internet. These questions and problems that must be solved include, on one hand, storage, network and system issues, and on the other hand, require user interfaces and paradigms that enable fast and efficient access to the video streams and associated documents. The work reported in this paper is developed in this context, to enable students to access video based lectures and associated documents, with multiple interfaces, and devices. Furthermore, the system should exploit video as a learning tool and, as such, the interfaces must also be designed with this goal in mind, allowing for user participation and interaction.

Considering these principles, VideoStore has the goal of developing an innovative system with the following main tasks:

1. Annotation of the video materials by a set of users (students and professors). Following our previous work [1,2] we are building a tool for video annotation on the Web that allows sharing annotations that are made with different media (text, images, video).

2. Building novel user interfaces that ease the task of searching for desired materials. This type of interfaces, adapted to each content type, allow adding other materials to the video, in a integrated way, namely animations and references to other Web content. The interfaces will also give direct access to the video content, based on mosaics, micons ("moving icons"), activity indexes, camera movements, and scene changes that are detected automatically.

3. Automatic construction of digests and video based documents, that combine several materials, related to a given lecture or set of lectures. Starting from the original materials (video and slides), combined with annotations and Web hyperlinks, new content is generated. An example is a movie that combines the original video, the slides, and some of the annotations made by the students on the Web site. The three topics above, which are presented in more detail in section 3, are crucial to make video based in-

formation available in a broadband environment. Video information needs efficient and well designed user interfaces, considering the requirements in terms of bandwidth. The ability to access and enrich these materials, e.g. through annotation, enables the wide use of rich multimedia content, that benefits directly from wireless or wired network infrastructures. Also, the ability to transform content between different representations and modalities, given the network or device characteristics (topic 3 above), contributes in a decisive way to the dissemination of the materials and to the effective use of the underlying network and computational infrastructure. Video annotation is also a crucial area in learning environments. Active reading is the combination of reading with critical thinking and learning, and is a fundamental part of education and knowledge work. Active reading uses annotation in physical documents. Active watching is a similar concept applicable to the way we consume what we watch. As text annotation promotes active reading, video annotation promotes active watching. If the user can make annotations, compare and relate with other materials, the audiovisual content will be an effective tool in learning processes, where VideoStore can have a relevant impact.

The paper is organized as follows. The next section presents previous work on video based interfaces and annotation tools that provides background to this project. The section after describes the systems concepts followed by the description of the overall system architecture and the main modules. The paper ends with some conclusions and directions for future work.

2. Previous work on video interfaces and annotation

This sections presents background work on video annotation and video based interfaces. Some of these results will be reused in the VideoStore project, adding support for network access and multiple devices. As such, we are approaching this project considering our previous background and developed work on video annotation. Video annotation can be a major component of learning processes as it stimulates participation from the users, which allows to depart from the traditional and passive way to look at video. As annotation promotes participation, it can be a major drive in learning processes, enabling to share knowledge in a community of users. We have designed and built a video annotator, AntV (Annotations in Video) [2], that can be used to add or edit annotations in video streams. The annotations can be text, images, or even video. Although the annotation files and associated media can be stored anywhere in a network the annotator is a local application. The AntV user interface is depicted in Figure 1.

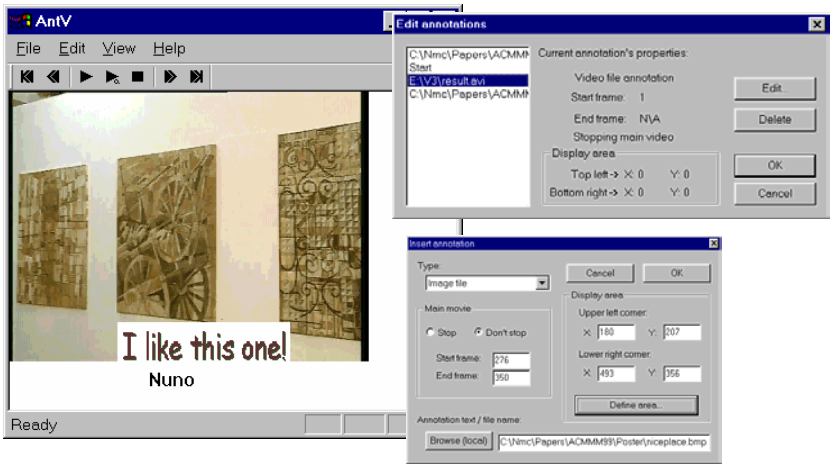

Fig. 1 *AntV (Annotations in Video)*. User interface of the AntV video annotator.

The main purpose of AntV is to enrich existing materials with new content, that is contributed by users. On, another project [1], we have build other annotator application, VAnnotator, that has the goal of de-

scribing the content, by adding metadata to it. To each video stream a set of descriptors can be added, that will describe the different aspects of the content. The layers of annotation are stored with the MPEG-7 standard, to allow interoperability with other annotation tools. In VideoStore we intend to combine the two annotation perspectives that we have in our previous work. Users will be able to add annotations in order to add to the existing materials and metadata will be used to describe the content and generate the different visual representations and interfaces that will be used to interact. The VAnnotator project also introduced the concept of video lenses, to provide interpretations of the multimedia content and give different perspectives of the information. These video lenses will also be reused in VideoStore, allowing for personalized views of the materials. Video lenses can be used to provide alternative representations to the same content or they can be used to filter out information that is not necessary for a given task.

Fig. 2 *Video Streamer and VideoIndex*. Video user interface techniques that show the structure of the video The images in the index are chosen with image processing algorithms for shot detection.

Additional work that is used as background to VideoStore is described in [3]. In this work, several ways to index video are proposed. These include text based, image based, and semantic indexes, timelines and annotation mechanisms. Additional work on testing and building video based interfaces has been done on a course on Multimedia Computing at the New University of Lisbon. As example of such interfaces is depicted on Figure 2, that shows two ways to access video: a time based index that represents the most relevant scenes of a video, chosen by an automatic shot detection algorithm, and an interface for video access based on the pioneer work done by Eddie Elliot and Glorianna Davenport at MIT Media Lab. Both these interfaces provide insights about the structure of the video for better browsing and navigation.

3. VideoStore concepts and features

Video access interfaces usually follow the media player model, which in turn in based on the VCR model, with a set of controls for playing, pausing, and similar operations, and thus do not explore all the dimensions and richness of video content. The VideoStore project will feature a set of user interfaces techniques and visual indexes that give access to the different elements in the video. These interfaces include moving icons or micons, activity level indicators in the video (using automatic video processing), scene change detection and shot representation. A set of syntactic level features are used to map semantic events that are relevant to this domain. As an example, silence detection in audio, camera motion detection, or scene detection can be represented for easier access to the content.

Annotations, made by students or professors, of the materials after being made available on the Video Store Web site, as presented in the section, are a way to enrich these materials and explore them in a process that is usually referred as active reading, for written documents. The goal is to do the same for video – a medium that is usually more experiential and that it does not invite to a more reflexive mode. As such, tools for Web based annotation are being developed within the VideoStore project. These annotations can be text, images, video or references to other Web based content, as in our previous work mentioned above. When adding this content to existing materials, it is possible to have a better browsing experience for all the participants – students and professors. The annotator also includes searching mechanisms for video content and for the annotations that were introduced. It is being developed both for the Web and for PDAs, which allows annotating while the lecture is taking place.

Starting from the lectures videos and other materials, namely annotations, new documents will be generated. These representations can integrate in video (image composition) the textual elements or images or generate alternative representations of all the materials. This requires the use of semi-automatic layout algorithms that allow to format and adapt the content to the final target platform and medium. It is also possible to consider other devices, besides personal computers. This includes PDAs with access to a wireless network and for these devices the contents should be properly adapted. Additional options include generating paper based versions of part of the documents or audio only versions, for access whenever these media are more convenient.

4. VideoStore architecture and main modules

As mentioned previously, the goal of the VideoStore project is to construct a Web-based system for students and professors, giving them additional ways to access these materials and removing some temporal and spatial constraints. It is intended to be a contribution towards more flexible learning technologies and systems. The basic source materials of the VideoStore system are pre-recorded video lectures synchronized with presentations.

We have searched for similar systems and underlying frameworks to integrate in VideoStore, and some of them are available that allow to synchronize video with slide show presentations, for the Web, but most are very restrictive. An additional requirement was to be able to integrate our own user interface techniques and annotation tools. We are currently prototyping with Microsoft Producer (http://www.microsoft.com/office/powerpoint/producer). With this software it is possible to synchronize video with slides and also to include HTML files which is a good way to integrate new features in the system. The final Web presentation made with Microsoft Producer can be designed with Cascading Style Sheets (CSS), giving freedom to the users on the chosen design.

In order to add functionalities to the system we are using PHP, as a way to process the videos and generate the different views, or video lenses as described above, including image based indexes and the annotation tools. We are directly accessing the media files, with our own API and then interfacing with PHP, to display the results in a Web browser. Microsoft Producer manages the synchronization between

media files and we will be adding interaction features to the Web presentations created by the tools. To support display in different environments, and because of the long length of the videos, we are using a streaming based solution to broadcast the lectures. Our current choice is Windows Server 2003, through its Media Services. Figure 3 includes the different modules and media and the relations between them.

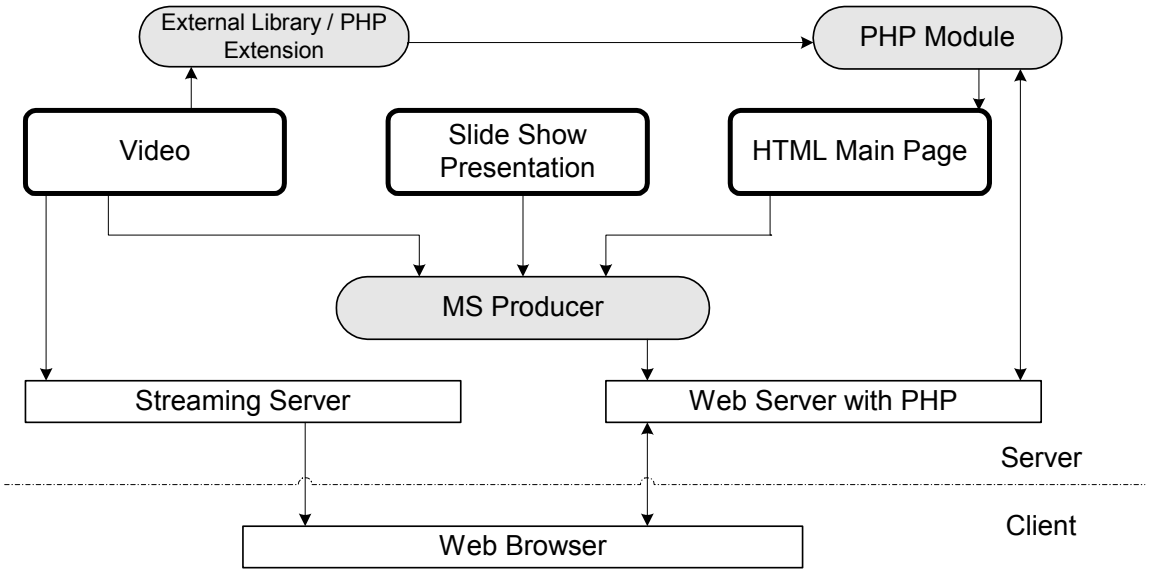

Fig. 3 *System Architecture*.

5. Conclusions and future work

The initial requirements for VideoStore have been identified and based on these requirements the overall architecture was defined. The different modules were specified and whenever possible and available they will be reused. Other modules are currently being built. Further work includes refining the user interface tools and testing them with the users of the system. We expect that additional requirements will appear whenever the system is running and it is used. Wireless networks are now commonly available in most campus and we will also support mobile access and user annotations, with portable devices. We believe that VideoStore can contribute to make learning more effective, by exploiting the full potential of video based content.

Acknowledgements This work is partially funded by POSI (Programa Operacional para a Sociedade de Informação) through the project "A Universalidade dos Saberes", sub-project VideoStore.

References

[1] M. Costa, N. Correia, N. Guimarães, Annotations as Multiple Perspectives of Video Content, ACM Multimedia 2002, Juan Les Pins, France, December 2002

[2] N. Correia, T. Chambel, Active Video Watching using Annotation, ACM Multimedia 1999, Orlando, Florida, USA, November 1999

[3] T. Chambel, T., N. Correia, N. Guimarães, "Hypervideo on the Web: Models and Techniques for Video Integration". International Journal of Computers & Applications, Acta Press, Vol. 23, #2, 90-98, 2001