SIVA Suite

Authoring System and Player for Interactive Non-linear Videos

Britta Meixner
University of Passau
Chair of Distributed Information Systems
Innstraße 43
94032 Passau, Germany
meixner@fim.uni-passau.de

Beate Siegel
University of Passau
Chair of Business Administration II
Innstraße 43
94032 Passau, Germany
beate.siegel@uni-passau.de

Günter Hölbling
University of Passau
Chair of Distributed Information Systems
Innstraße 43
94032 Passau, Germany
hoelblin@fim.uni-passau.de

Franz Lehner
University of Passau
Chair of Business Administration II
Innstraße 43
94032 Passau, Germany
lehner@uni-passau.de

Harald Kosch
University of Passau
Chair of Distributed Information Systems
Innstraße 43
94032 Passau, Germany
kosch@fim.uni-passau.de

ABSTRACT

In this paper, an intuitive authoring system and player for interactive non-linear video called SIVA Suite is presented for demonstration. Such videos are enriched by additional content. Possible forms of additional content are plaintext, richtext, images and videos. Interactivity is implemented based upon selection buttons which allow the user to follow different plotlines. Additional forms of interactivity are realized as clickable objects in the video and a table of contents for the video. The software provides a tool for manually cutting videos and an automated shot detection. The non-linear flow of the video can be designed using a scene graph with fork nodes. Editors for text and images support the user in adding information to the video. A finished video project is exported to an XML file with a specific schema and Flash video (flv) files. The player processes the XML file, plays the interactive video and shows additional contents. It can be customized to the requirements of the presentation of the video and the corporate design of the homepage the video is embedded in.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Interaction styles*; H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia—*Navigation, User issues*

MM'10, October 25–29, 2010, Firenze, Italy.

General Terms

Human Factors, Design

Keywords

Interactive Video, Non-linear Video, Video Annotations, Authoring System, Interactive Video Player

1. INTRODUCTION

The viewer of a traditional video takes a mostly passive role. Traditional videos are linear and cannot provide additional information about objects or scenes. Contrary to linear videos, an interactive video is a digitally enriched form of video materials arranged for an overall concept. It presents additional information beyond the original content. Furthermore, it offers new forms of influence and navigation in the video and additional contents.

Detailed and supplemental images, continuative videos, commentarial texts and interesting audio files can be added to the original video or scenes of the video as additional information. Breaking up the rigid linear structure of a video by cutting it into shots or scenes allows a rearrangement of the video. The structure of the video is no longer linear; it is cut into shots or scenes that can be rearranged as a result of user actions. Branching plotlines can be implemented. It is also possible to access shots and scenes by a table of contents, enabling the viewer to watch only relevant parts of the video. A keyword-search rounds out the navigational potential of an interactive video.

2. USAGE SCENARIOS

Interactive videos can enhance several types of applications. In e-learning applications, an interactive video can support more learners by integrating several options adapted to several types of users with different knowledge levels.

Short quizzes can be integrated in the content, and associated with different scenes depending of the result obtained by the user.

A virtual tour through a city or a building can be rendered more attractive by enabling the users to choose their own path of visit using buttons on junctions or crossroads. Watching a city, additional information can be added about hotels and stores featured in the video. Information about offices, recreation rooms and a cafeteria can be added in a video about a public building. Watching a prefabricated house, information about building materials, price and dimensions can be appended at appropriate spots. This scenario will be demonstrated in the demo section.

In another scenario, a sports athlete shoots a video of a training session and sends it to the coach. The coach then annotates the video with feedback and tips. The athlete can watch the video and analyze faults on his own.

3. DEMONSTRATION

The SIVA Suite is proposed to be demonstrated at ACM Multimedia 2010 showing a complex video tour trough a prefabricated house. Possible walking paths within the house were shot as videos. Choices about where to go at doors, in halls or stairs are reproduced by the scene graph and represented by selection buttons in the video. Viewers of the video are able to choose their own way through the house by clicking one of the buttons. The next scene to be shown is then selected according to the users choice. Each scene provides additional information and clickable overlays containing additional details about furnishings and building materials.

4. RELATED WORK

Many efforts have been made to provide tools for annotation of metadata and navigation in videos and multimedia content. Advene [7] allows navigation via self defined annotations. Vizard [20] and the VideoAnnEx Annotation Tool [22] provide algorithms for shot detection and annotation of the shots. Anvil [14] is a research tool for gesture detection. M-OntoMat-Annotizer [10] is a tool for multimedia analysis, reasoning and retrieval. LongoMatch [6] and ASPOGAMO [8] are tools for analyzing and annotating sports videos. The annotation of many of these tools is based on MPEG-7.

This is a subset of features needed for interactive videos, but for real interactivity, features like non-linearity in video flow and attachment of additional information like pictures or audios are required. Moreover, ease of use of the software, especially for creating non-linear flow and annotations, is necessary to make the authoring software accessible even to amateurs. Many authoring tools provide some of these features. High-end software like Adobe Creative Suite 5 [2], Adobe Flash CS4 Professional [3] and Microsoft Expression Studio 3 [15] allow professionals and Web-programmers to create pleasant interactive videos, but they are too complex to use for non-experts, because of the large set of features they provide. Virtual Campfire [21] provides a lot of interactive features and a graph view for non-linearity, but is designed for collaborative multimedia semantization with mobile social software. Other tools are easier to use, but lack several features needed for interactive videos. 5min Live Videopedia [1], Quick.tv [19], Viddix Beta (Mix Video with the Web) [24], VideoAnt [12], VideoClix [25] and YouTube Video Annotations [26] offer functions for integrating additional information in a video, but do not support non linearity in the video flow. XIMPEL - eXtensible Interactive Media Player for Entertainment and Learning [9] provides non-linearity, but no additional information can be added. ADIVI [13] allows both non-linearity and additional informations, but the player is not adaptable for specific needs, for example the layout cannot be changed. Compared to these tools, SIVA Suite is easy to use, allows many different kinds of annotations, non-linearity in video flow and includes an adaptable player.

Figure 1: SIVA Producer, Player and XML file

Several players capable of rendering the output of the authoring tools are available. They usually take XML files containing instructions and layout information. They must support navigational elements, clickable areas, and integration of additional information. Adobe Flash/Flex [5], Microsoft Silverlight 3 [16] and JavaFX [17] all provide the ability to implement players capable of processing the output of interactive video authoring tools and presenting an interactive video. Adobe Flex is used in our project because of the widespread use of the Flash Player Plugin.

5. SIVA SUITE

SIVA Suite consists of three main parts: SIVA Producer, SIVA Server and SIVA Player (see Figure 1). SIVA Producer is an authoring tool for interactive videos. An export functionality produces a XML file, video files and annotations. The second part is SIVA Server. This Web server hosts all the data needed to play an interactive video. The third part, SIVA Player is a player for interactive videos. It interprets the XML file and displays the interactive video.

5.1 SIVA Producer

SIVA Producer is implemented as an Eclipse Rich Client Platform (RCP) [23] application. It appears in an operating system specific look-and-feel. The layout of the editors, repositories and information areas is arrangable via drag and drop. Our software can be used in German or English. The application is implemented according to the MVC architecture to provide an isolation between data, logic and presentation. SIVA Producer is composed of four major operational parts: project management, video processing, scene arrangement by a scene graph and annotation editors.

The **project management** part covers all functionalities for organizing project-specific data. An export function for the player is implemented; the exported files have to be uploaded to SIVA Server or to another Web server.

The **video processing** part of the software allows the user to define scenes in previously imported videos. A scene can be defined in three ways. The definition can be done via a preview of the video, via a timeline or by quoting the start and end time of the scene. Defined scenes are arranged in an overview, where all scenes of a video are listed in a sorted way. An automated shot detection using Scalable

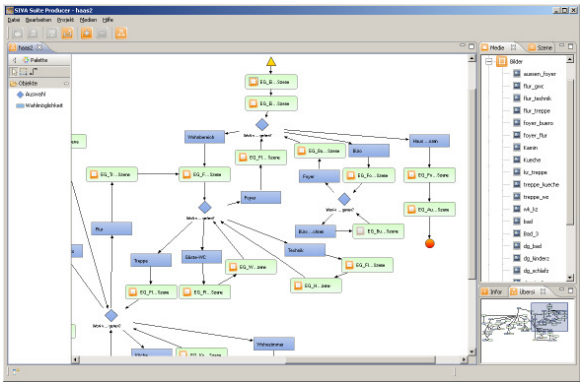

Figure 2: SIVA Producer - scene graph

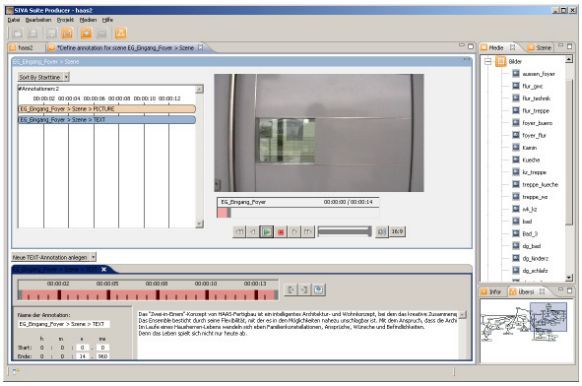

Figure 3: SIVA Producer - annotation editor

Color Descriptor [11] und Edge Histogram Descriptor [18] is also available.

The most important part of the software is the **scene graph** shown in Figure 2. Scenes identified using video processing are made available in a scene repository. They can be added to the scene graph editor by drag and drop from the scene repository. This editor provides an area where graph nodes can be arranged by drag and drop and a toolbar. The scene graph consists of a start and an end node, scene nodes, fork nodes and selection nodes that are connected by edges. The start node (yellow triangle) represents the beginning of the video. It is inserted into the scene graph automatically and cannot be deleted. It must be followed by a scene node. A scene node (green rectangle) represents a scene of a video. It is inserted via drag and drop. It can be followed by either another scene node, by a fork node or by the end node. The scene graph has to contain at least one scene node. A fork node (blue diamond) provides a selection between parallel plotlines. It is inserted via drag and drop from the toolbar and is always followed by at least one selection node (blue rectangle) which defines the caption of buttons. Selection nodes are inserted the same way as fork nodes. They have to be followed by a scene, another fork node or the end node. The end node (red circle) represents the end of the video. It is inserted into the scene graph automatically. It cannot be deleted and cannot be followed by another node. All these conditions are checked when an edge is inserted. The whole graph is checked for consistency during the export. The graph provides tools to align the scenes, undo/redo and a zoom to facilitiate working with larger graphs. An overview with a marker of the area shown in the scene graph window is available for a better navigation in the graph. As shown in Figure 2, complex paths through a video can be constructed. Forks enable the definition of different paths through a video. It is possible to return to a fork node that was already visited and take the same way once more or another way. Scenes can be linked in any order.

After adding a scene to the graph, additional contents can be attached to it using the **annotation editor** as shown in Figure 3. Supported types of content are plaintext, richtext, subtitles, images, videos and audio files. The annotation editor provides an overview of already defined annotations and a selection-button to chose the kind of annotation to add. After selecting which annotation should be added, an appropriate editor, a timeline and several controls are shown to configure the annotation. Start and end time (relative to the scene) and the area where the annotation will be shown in the player have to be defined for all types of annotation. The area where the annotation will be shown can be one of the four side bands. If the height of the upper or lower area or the width of the left or right area is set to zero it cannot be selected for annotation. An annotation can also be shown as an overlay. Size and position of the overlay must be set in a separate editor, where a rectangular area can be resized. The area can then be clicked and moved to the desired position. Position and size of the area can be saved via a recording function. There are text editors for all kinds of text annotation where the text can be added directly. The richtext editor offers basic functionality to format the inserted text similar to HTML. Images have to be placed in the editor area by drag and drop from the repository. Then, the image can be opened in an image editor where basic functionality is provided to mark areas or to write text into the image. Video and audio files are placed in the editor area like images but there is no editing functionality in our program.

5.2 SIVA Player

SIVA Player is implemented in Adobe Flex [5] because of the wide distribution of the Flash Player [4]. The player reads the XML file and loads videos and annotations as instructed by actions and triggers in the XML file. The video can be played in fullscreen mode or a mode with a centered video and annotations positioned around the video in four areas (see Figure 4). Areas that are not needed can be hidden by the author of the video. Several CSS skins are available, which allows the author to adapt the player to corporate design for a website. The player provides typical controls as well as controls for interactive video. Interactivity is implemented manifoldly. First, it is realized via selections with buttons which provide different possibilities for choosing the next scene. In case of a fork in the videos, a button panel is shown, enabling the viewer to choose the further course of the video by clicking on a button. A second form of interactivitiy is implemented in terms of a table of contents. The third form of interactivity are clickable objects in the video. They allow the viewer to jump to another scene in the video.

As shown in this paper, our tool provides non-linearity in video flow defined by a scene graph as well as different types of annotations and several annotation editors. The player is adaptable to specific needs and allows different forms of in-

Figure 4: SIVA Player

teractivity in the video. Future work will focus on a player for mobile devices and on the corresponding export functionality for that in SIVA Producer. Furthermore, features for collaboration will be added in both, SIVA Player and SIVA Producer.

6. ACKNOWLEDGMENTS

This work was partually supported by European Social Fonds and the Bayrisches Staatsministerium für Wissenschaft, Forschung und Kunst (Bavarian State Ministry of Sciences, Research and the Arts) under project name "iVi-Pro".

7. REFERENCES

[1] 5min LTD. 5min Life Videopedia Beta. Website, April 2010. http://www.5min.com/.

[2] Adobe. Adobe Creative Suite 5 Master Collection, April 2010. http://www.adobe.com/products/creativesuite/mastercollection/.

[3] Adobe Systems. Adobe Flash CS4 Professional. Website, April 2010. http://www.adobe.com/de/products/flash/features/?view=topoverall.

[4] Adobe Systems. Adobe Flash Player. Website, April 2010. http://www.adobe.com/de/products/flashplayer/.

[5] Adobe Systems. Adobe Flex. Website, April 2010. http://www.adobe.com/products/flex/.

[6] A. M. Alastruey. *LongoMatch:The Digital Coach*, April 2009. http://longomatch.ylatuya.es/documentation/manual.html.

[7] O. Aubert and Y. Prié. Advene: an open-source framework for integrating and visualising audiovisual metadata. In *MULTIMEDIA '07: Proc. of the 15th international conference on Multimedia*, pages 1005–1008, New York, NY, USA, 2007. ACM.

[8] M. Beetz, N. von Hoyningen-Huene, B. Kirchlechner, S. Gedikli, F. Siles, M. Durus, and M. Lames. ASpoGAMo: Automated Sports Game Analysis Models. *International Journal of Computer Science in Sport*, 8(1), 2009.

[9] W. Bhikharie. XIMPEL overview. Website, April 2010. http://www.cs.vu.nl/~eliens/im/report-im08-ximpel.pdf.

[10] S. Bloehdorn, K. Petridis, C. Saathoff, N. Simou, Y. Avrithis, S. H, Y. Kompatsiaris, and M. G. Strintzis. Semantic annotation of images and videos for multimedia analysis. In *In Proc. of the 2nd European Semantic Web Conference, ESWC 2005*, pages 592–607, Heraklion, Crete (Greece), June 2005.

[11] J.-Y. Chang, C.-J. Lian, H.-C. Fang, and L.-G. Chen. Architecture and analysis of color structure descriptor for real-time video indexing and retrieval. In *PCM (2)*, volume 3332 of *Lecture Notes in Computer Science*, pages 130–137. Springer, 2004.

[12] B. Hosack, C. Miller, and D. Ernst. VideoANT: Extending Video beyond Content Delivery through Annotation. In *Proc. of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009*, pages 1654–1658, Vancouver, Canada, October 2009. AACE.

[13] InnoTeamS GmbH. ADIVI - Hypervideo für Lernen und Support. Website, April 2010. http://www.adivi.net/.

[14] M. Kipp. Anvil - A Generic Annotation Tool for Multimodal Dialogue. In *Proc. of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, pages 1367–1370, Aalborg, September 2001.

[15] Microsoft Corporation. Microsoft Expression Studio 3. Website, April 2010. http://www.microsoft.com/germany/Expression/.

[16] Microsoft Corporation. Microsoft Silverlight 2. Website, April 2010. http://www.microsoft.com/silverlight/.

[17] Oracle Corporation and/or its affiliates. JavaFX. Website, April 2010. http://javafx.com/.

[18] D. K. Park, Y. S. Jeon, and C. S. Won. Efficient use of local edge histogram descriptor. In *ACM Multimedia Workshops*, pages 51–54, 2000.

[19] Quick TV Limited. Make your video dynamic & interactive. Website, April 2010. http://www.quick.tv/.

[20] H. Rehatschek and G. Kienast. VIZARD - An Innovative Tool for Video Navigation, Retrieval, Annotation and Editing. In *Proceedings of the 23rd workshop of PVA 'Multimedia and Middleware'*, May 2001. http://www.video-wizard.com/index-n.htm.

[21] D. Renzel, R. Klamma, Y. Cao, and D. Kovachev. Virtual Campfire - Collaborative Multimedia Semantization with Mobile Social Software. In *Proc. of the 10th International Workshop of the Multimedia Metadata Community on Semantic Multimedia Database Technologies (SeMuDaTe'09)*, Graz, Austria, December 2009.

[22] J. R. Smith and B. Lugeon. Visual annotation tool for multimedia content description. volume 4210, pages 49–59. SPIE, 2000.

[23] The Eclipse Foundation. Rich Client Platform. Website, April 2010. http://wiki.eclipse.org/index.php/Rich_Client_Platform.

[24] VIDDIX B.V. Viddix Beta (Mix Video with the Web). Website, April 2010. http://www.viddix.com/.

[25] VideoClix.tv. VideoClix.tv. Website, April 2010. http://videoclix.tv/.

[26] YouTube, LLC. YouTube-Videoanmerkungen. Website, April 2010. http://www.youtube.com/t/annotations_about.