Second-Layer Navigation in Mobile Hypervideo for Medical Training

Britta Meixner[1,2](✉) and Matthias Gold[1]

[1] University of Passau, 94032 Passau, Germany
meixner@fim.uni-passau.de
[2] FX Palo Alto Laboratory, Inc., Palo Alto, CA, USA
meixner@fxpal.com

Abstract. Hypervideos yield to different challenges in the area of navigation due to their underlying graph structure. Especially when used on tablets or by older people, a lack of clarity may lead to confusion and rejection of this type of medium. To avoid confusion, the hypervideo can be extended with a well known table of contents, which needs to be created separately by the authors due to an underlying graph structure. In this work, we present an extended presentation of a table of contents for hypervideos on mobile devices. The design was tested in a real world medical training scenario with the target group of people older than 45 which is the main target group of these applications. This user group is a particular challenge since they sometimes have limited experience in the use of mobile devices and physical deficiencies with growing age. Our user interface was designed in three steps. The findings of an expert group and a survey were used to create two different prototypical versions of the display, which were then tested against each other in a user test. This test revealed that a divided view is desired. The table of contents in an easy-to-touch version should be on the left side and previews of scenes should be on the right side of the view. These findings were implemented in the existing SIVA HTML5 open source player (https://code.google.com/p/siva-producer/ (accessed February 06, 2015)) and tested with a second group of users. This test only lead to minor changes in the GUI.

Keywords: Hypervideo · User interface · Navigation · Preview · Table of contents

1 Introduction

With new technologies like HTML5 and high internet bandwidths, the embedding of high quality video into homepages is nothing special anymore. Easily portable end user devices like tablets or smart-phones can be used to watch online videos. Using these technologies, appealing hypervideos can be created while providing a good quality of experience to the viewers even on mobile devices. Hypervideos in general consist of different types of media (like video, audio, images, and text) which are linked with each other. They provide interactive elements (links and other clickable elements). Using these interactive

Q. Tian et al. (Eds.): MMM 2016, Part I, LNCS 9516, pp. 382–394, 2016.
DOI: 10.1007/978-3-319-27671-7_32

elements, a viewer can chose her/his own path through the hypervideo. Accordingly, the shown elements of one whole hypervideo vary from viewer to viewer. A detailed explanation and definition of the term hypervideo is given in [13]. The hypervideos described in this work have video as a main medium, which are cut into scenes and linked to a scene graph. Additional information in form of text, images, audio files, or other videos are linked to these main video scenes. Hypervideos are very suitable for the transfer of knowledge due to their structure and the creative potential they offer [3,6,19]. They can be used to support learning in an every day work live with increasing mobility. Benefits of a video based learning portal like reduced overall costs, better learning results, improved employee satisfaction, and a higher reachability of the employees[1] can also be applied to training videos in healthcare or fitness scenarios. With further savings in the healthcare sector, the rehabilitation treatments in clinics become shorter and the patients have to do their training at home after leaving the clinic. Thereby a hypervideo with different trainings may be helpful for a proper execution of the exercises.

In this paper, we propose interfaces for medical training scenarios. The target group were people above 45. This group is more likely to need medical treatment which often results in longer recovery times than for people of younger age. Their computer skills are ranging from nonexistent to expert. A more precise description of the scenario and the target group can be found in Tonndorf et al. [18].

1.1 Problem Statement

This works focuses on navigation in a whole hypervideo scene graph, not in one scene or a very limited part of the scene graph, like for example by clicking on a tracked object (which may invoke the display of an annotation) or different levels of detail for a scene and different camera angles (which are usually implemented as parallel strands of scenes with time synchronization between scenes).

Three types of navigation may exist in a hypervideo (see also Fig. 1):

- ***Graph-based navigation*:** navigation in the underlying graph from scene to scene; the follow-up scene at the end of an already watched scene depends on the user interaction; for example, the user clicks on a button in a button panel
- ***Graph-independent navigation*:** jumps in the hypervideo from one scene to another which are not necessarily connected by an edge; for example, by the selection of a result from a search on the whole graph
- ***Second-layer navigation*:** navigation in a second navigation layer that leads to specific scenes in the graph structure; for example, a table of contents (tree structure) which provides entry points to certain nodes in the scene graph

The graph-based and the graph-independent navigation can be easily and intuitively implemented as lists of buttons or clickable areas, because the lists are usually very short (3–8 items). The more complex tree structure of a table of

[1] http://site.kaltura.com/rs/kaltura/images/TheStateofEnterpriseVideo2014-KalturaReport-Final.pdf (accessed August 04, 2015).

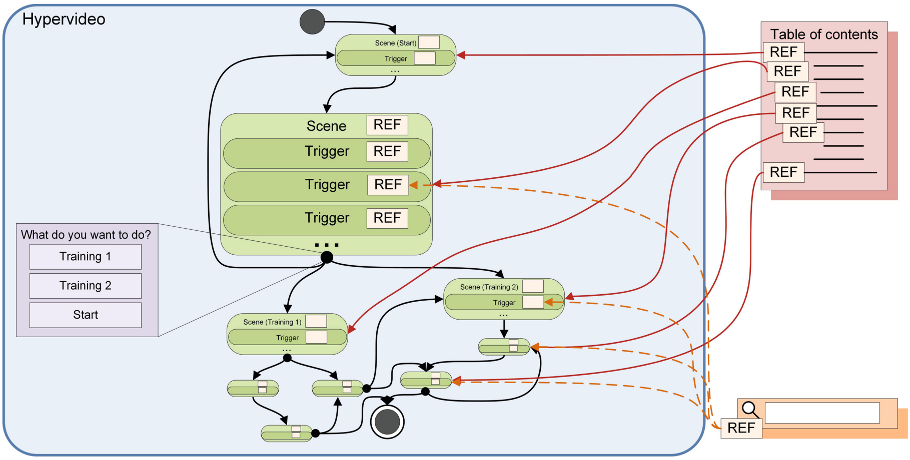

Fig. 1. Graph-based navigation (purple color), graph-independent navigation (orange color), and second-layer navigation (red color) in hypervideos (Color figure online).

contents contains some more challenges in contrast, because they contain many entries which afford a structure. Traditionally, tables of contents are presented as indented lists of formatted links. The indentation indicates the hierarchy between the links. These formatted links are comparatively small but easy to handle with a mouse on a PC. Using the same traditional representation on a tablet with touch input appears to be much more difficult due to the fat finger problem. Accordingly, another appearance is needed. It should be easy to interact with, especially on smart-phones and tablets. Furthermore, it should provide an overview of all levels with simple interactions (like scrolling). Design patterns for menus/link structures representing different levels of contents exist for web pages, like the "Accordion Menu"[2], but they are unusable on smart-phones or tablets. The iOS design guidelines provide a similar working pattern called "Table View"[3] which also could be used for a table of contents, but only provides an overview of one level at a time. The underlying level has to be folded out to see its contents. Navigation from the top level of the structure to lower levels is very laboriously because the lower levels have to be expanded one by one and no quick overview is provided which can be accessed with simple gestures. To the best of the author's knowledge, no design pattern suitable for smart-phones or tablets could be found which represents an indented list with different levels and gives an overview of all contents. Used with the underlying graph structure of the hypervideo, an intuitive linking between the hypervideo structure and the table of contents is necessary. Furthermore, a traditionally designed table of

[2] http://ui-patterns.com/patterns/AccordionMenu (accessed August 04, 2015).

[3] https://developer.apple.com/library/ios/documentation/UserExperience/Conceptual/MobileHIG/ContentViews.html#//apple_ref/doc/uid/TP40006556-CH13-SW1 (accessed August 04, 2015).

contents does not provide any preview of the linked video contents which further complicates making the right selection in a hypervideo. When a video is paused or stopped and the player is closed and reopened, the viewer usually has to start from the beginning which can be annoying in larger trainings.

1.2 Research Contributions

Dealing with the problems previously described, this work proposes the following solutions:

- A design for a table of contents which is usable on smart-phones or tablets and provides a preview from the table of contents into the hypervideo structure.
- Additional orientation in the table of contents to find entry points after a break.

In this paper we firstly present related work. Then, our usability evaluation consists of two parts. A concept is created and evaluated, then the GUI is implemented, and afterwards the implementation is tested and refined.

2 Related Work

A study on mobile multimedia summarized “that usability aspects, like an intuitive UI, are strongly related to the users? desire for being effective and ambitious” [10]. Keeping that finding in mind, we now try to give an overview of related work for navigation in hypervideos on mobile devices. It has to be noted, that by the best of the authors knowledge, no work exists which unites the terms “hypervideo”, “navigation”, and “mobile device”. Related work can mainly be found in two areas: Navigation in hypervideos and video navigation on mobile devices. We will present and discuss papers from both areas.

Implementations for navigation in linear videos are examined by Cunha et al. [2] and Hurst et al. [7]. Hurst et al. evaluate methods of “pen-based navigation of videos on PDAs” [7]. They implement gestures to “skim a video along the timeline on different granularity levels” [7] and to manipulate the replay speed. However, this navigation is very basic and provides no additional information to the viewer to help her/him to find the desired contents. Cunha et al. explore “the generation of textual annotations on [linear] videos played on mobile device” [2]. Their “approach is to offer an application that allows associating annotations to a navigation line decorated with frames that are representative of the points of interest” [2]. Users should be able to find interesting points in the video intuitively using a combination of thumbnails and text as points of reference.

Related work regarding the navigation in hypervideos can be found in two forms. Either the GUI is evaluated or waiting times should be reduced. Shipman et al. [17] and Girgensohn et al. [4] introduce the Hyper-Hitchcock detail-on-demand video player. Thereby, detail-on-demand videos are a restricted form of hypervideo consisting nothing but videos and providing only one link at a time. This type of video is used to “watch short video segments and to follow hyperlinks to see additional detail" [4]. This desktop player provides no table of contents. The proposed interface, which should "provide users with the appropriate affordances to understand the hypervideo structure and to navigate it effectively" [4] was evaluated for its usability and refined in this process. A successful navigation in this type of video was tested afterwards. Girgensohn et al. summarize that "the user interface needs to present users with an intuitive view of the hypervideo structure. Such a view should be suitable for different tasks and guide the users towards the most appropriate interaction" [4]. Grigoras et al. [5] use a form of hypervideo which is composed of video scenes that are linked in a graph structure. Their work focuses on the streaming and reduction of latency through prefetching after analyzing the user behavior, but not on the creation of the user interface. Their findings are not especially tailored for mobile devices, but can be applied nevertheless.

Online course platforms usually provide linear videos or sequences of videos which can be navigated by lists of entry points. Furthermore, it is possible to jump from one video to its successor or predecessor. Watched videos are marked with symbols. The Coursera[4] and Udacity[5] apps provide lists of video parts which can be selected for viewing. A more complex structure of the video can only be provided by structuring the labels, for example with numbers or a repetition of the main headline followed by a separator and the sub-headline. The Khan Academy app[6] provides lists with categories, which results in a menu-like structure, but does not provide a good overview for navigation to certain contents or explanations. The Lynda app[7] has a contents list with two layers. Thereby, headlines are formatted differently. None of these apps provides a table of contents-like representation or has a scene graph as an underlying structure.

Other related work can be found in the areas of multi-view video on mobile devices and on direct manipulation video navigation. Apostu et al. [1] study navigation in linear multi-view videos on mobile devices based on spatial information. Miller et al. [14] explore multi-view video with hyperlinks on mobile devices. They "offer several mechanisms for viewing hypermedia and perspective selection" [14]. No user study can be found about the prototypes in this work. In both works, the only used medium is video, none provides a table of contents or any other navigational structures. The research is more focused on the switching between the different view instead of navigation in the whole video. Karrer et al. [8] and Nguyen et al. [15] examined direct manipulation video navigation with gestures in linear videos on mobile devices. These works reveal important findings on the fat finger problem, which is also an issue in our work.

3 Usability Evaluation in the Design Phase

The design phase for the extended table of contents was separated into tree major parts. An expert group was conducted to get first hints on how the table

[4] https://www.coursera.org/ (accessed August 05, 2015).
[5] https://www.udacity.com/ (accessed August 05, 2015).
[6] https://www.khanacademy.org/ (accessed August 05, 2015).
[7] http://www.lynda.com/ (accessed August 05, 2015).

of contents should be presented to the users and which information should be displayed. In a second step, a survey was composed to find out how people above 45 use technologies and media, as well as to find out about habits while searching for information. The third step was the implementation of a prototype, which was used to make a first usability test.

3.1 Expert Group

The expert group was conducted to find answers to the following questions: Which information should be presented to users considering the technical constraints of hypervideos? How should the table of contents be presented?

Study Method and Participants. Five experts ($N = 5$) from different disciplines were invited: a media and communication scientist, a web developer, a programmer, a multimedia researcher, and an expert for media philology and media analysis. All of them had experience with hypervideos and/or the usability of web and mobile GUIs. At least three of them had experience in experiments with the target age group. A discussion guideline as well as first paper prototypes from a brain storming session were presented to the group.

Result and Discussion. As a result it can be noted that a history is considered necessary but should be implemented separately from the table of contents. According to the experts, most hypervideos that will be used for learning or training will have a table of contents and not a structure which is created from the graph structure automatically. A result of an automated creation has the disadvantage that it is dependent from the graph structure and the used algorithms. Unfitting combinations thereof may result in an unusable table of contents which may need a lot of manual restructuring. This lead to the conclusion that no algorithms are needed to analyze the whole scene graph and bring it into a tree-like structure. In the remainder of this work, we focus on the table of contents created by the author of the video. The expert group agreed that the view of the table of contents should be divided. There should be a structure inspired by a table of contents on the left side and a preview on the selected scene and follow-up scenes on the right side. The upright table of contents should be on the left side due to the reading direction and common practices for websites. A video preview or representative thumbnail image is preferred to a simple textual description of the scene.

3.2 Survey

Our survey consisted of several sections which tried to give hints for the answers of the following questions: Which knowledge about tablets does the target group have? Which attitude towards learning and training videos does the target group have? Which devices are used for learning and training videos by the target group? What do users expect from a table of contents in a hypervideo?

Study Method and Participants. The process of creating the survey consisted of several steps. A first version of the survey was created and then tested for its functionality by two people with knowledge about empirical social research. After the proposed changes were made, the survey underwent a pre-test with twenty people with knowledge about surveys. A second revision of the survey with the comments of the pre-test followed and lead to the final survey which was available in two versions, on paper and online.

The survey was propagated in three different ways: in a forum where it was seen 200 times, in a social network where it was seen 113 times, and on paper where it was printed 125 times. The survey had 193 participants, whereby 130 people used the online version and 63 people filled out the paper form. 104 of the 193 participants were over 45 years old. In the remainder of this section, we focus on the over 45 year olds ($N = 104$). In this group, 47 were male and 55 female (2 did not answer the question). 38 participants had a degree from a university or university of applied sciences, 29 have completed an apprenticeship.

Result and Discussion. The following statements can be made about the results of the survey: Most of the devices are used daily/several times a week (like TV, computer, notebook, and smart-phone) or rarely/not regularly (like TV with internet access, tablets, mobile phone without internet access, MP3-player). Only 5.8 % of the participants above 45 do not use the internet at all. Regarding devices with touch screens, the most used device is an ATM with a touchscreen. Smart phones or mobile phones are the mainly used devices in the private area followed by tablets and notebooks with touch screen. 74 participants claimed that they have watched learning and training videos at some time. With multiple answers possible, 39 participants had watched tutorials, 38 participants had watched repair instructions, and 26 participants had watched assembly instructions. Besides those types of materials, videos about language courses, sport exercises, or styling tips were watched. Especially sport exercises were watched repeatedly. The question about the devices used to watch the learning and training videos revealed the following findings (multiple answers possible): 44 participants used the PC, 41 used a laptop, 25 used the TV, 25 used a tablet, and 25 used a smart-phone.

According to the fact that currently neither exists a commonly recognized structure for hypervideos, nor are hypervideos widespread and thus not known to many people, we used a more analog scenario to find out how the participants would navigate in contents. The scenario of searching the name of an unknown bird in an encyclopedia was introduced. This type of book is usually not read in a linear way and provides links between different entries, which is very similar to hyper-linked structures on the Web. The participants were asked how they would search for the bird. The most preferred structure was searching in an index, secondly they would just browse the book and thirdly they would use the table of contents. Using the table of contents for this task, the participants were asked how it should look like. They considered images of the bird with a short description as most useful. The grouping of the birds in their categories as well as images of the birds without text were also contemplated as very advantageous. Textual descriptions were acknowledged as less helpful. In another scenario, namely the trip to a foreign city, the participants were asked how they would note down the sightseeing route. Thereby the options of marking the sights and marking the taken path were preferred to numbering the sights consecutively, marking the taken path at crossroads, and noting down the street names.

Summarizing the results of the survey, it can be noted that 90 % of the participants know touch screens and smart-phones. Tablets are widespread but not always used regularly. Learning and training videos are well known among the participants, about 1/3 of them has watched these videos on a tablet. Regarding the table of contents, a structure of the entries is important, a combination of image and text are favored, and already visited contents should be marked somehow. These results match the findings of the expert group and will be implemented in the prototypes for the user test.

3.3 User Test with Prototype

The user test with the prototypes was conducted to find answers to the following questions:

- Is the table of contents used for navigation by the users?
- Is a structured or a graphical (menu-like) approach more clear for the users?
- Is the separation between a structured table of contents and the follow-up scenes comprehensible for every user?
- Are the used markers for already visited scenes understood correctly?
- Are the symbols large enough?
- How comfortable is the user handling the device (tablet)?

Study Method and Participants. The results from the expert group and the survey were implemented in software supported paper-like prototypes created with Balsamiq[8]. Two versions, the structured version and the graphical (menu-like) version were designed. An example of the structured prototype can be found in Fig. 2, an example of the graphical (menu-like) version is illustrated in Fig. 3.

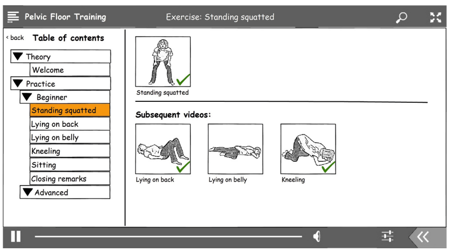

Fig. 2. Structured version of the prototype.

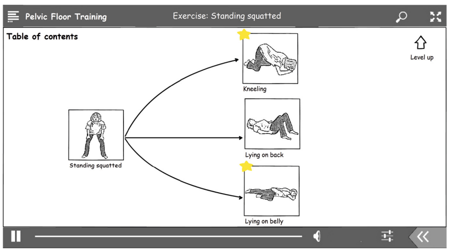

Fig. 3. Graphical (menu-like) version of the prototype.

[8] https://balsamiq.com/ (accessed January 25, 2015).

To answer the previously described questions, the test users got different tasks in a scenario with a medical training. The tasks included:

- Use of the table of contents as a help for navigation.
- Navigate through the video to the end scene.
- Recognize already visited video scenes.
- Recognize follow-up scenes in the representation.

The user test was conducted with a group of eight people ($N = 8$) all aged above 45. To make sure every tester had the same preconditions for the test, it was ensured that the participants did not meet between the tests. An introductory and explanatory text was read to the participants from a paper, then some basic questions about age and educational background were asked. After that, the tasks enlisted above were presented to the test users one after the other after a brief introduction to the GUI of the prototype. The users had to perform each task for both versions. While completing their tasks they were asked to comment what they were doing and why they were doing it with the think-aloud technique [16, p.256].

Result and Discussion. The user test revealed the following results: Both, volume and play/pause button were recognized correctly by all of the test users. The buttons for the table of contents, the search, and the full screen were recognized by half of the participants. The settings button was not recognized by anyone. The presentation of the follow-up scenes was problematic, connections were assumed that did not exist. A more clear separation may solve this problem. The currently selected scene was not recognized and should be highlighted. The video controls (play/pause, timeline, and volume control) were visible when the table of contents was open. This irritated the test users, because no connection between the table of contents and the video control existed. Accordingly, the video control panel should be hidden when the table of contents is displayed. Furthermore, it was suggested that the "<back" button should be removed from the top left side and an "X" should be added to the right upper corner of the view, which is well known from other applications. Already watched scenes should be

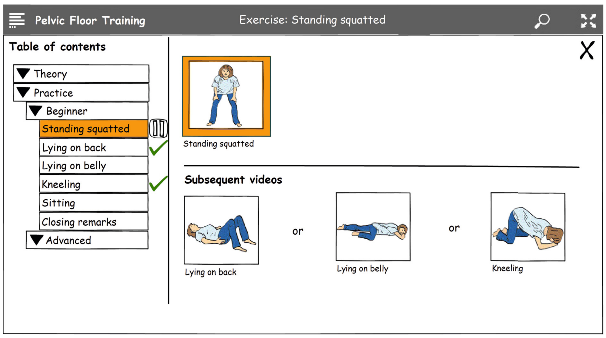

Fig. 4. Refined version of the paper prototype.

marked in the structure at the left side, instead on the scene thumbnails in the preview area. The structured version was preferred to the graphic version by the test users. This leads to the refined prototype in Fig. 4. It has less elements than the structured version in Fig. 2 which leads to a better overview.

4 Implementation

The implementation in the player software afforded extensions of the meta data format and the authoring tool. The existing meta data format described by Meixner and Kosch in [11] had to be extended. Besides a video for a scene, a thumbnail image for that video has to be referenced in the scene element in the XML file. Furthermore the authoring tool described by Meixner et al. in [12] needs to export a thumbnail for each scene. This was accomplished with an editor that allows the selection of the thumbnail.

The implementation of the table of contents in the player GUI was realized according to the results of the user test as described in the previous section and illustrated in Fig. 4. Thereby, the design was matched with the design of the existing player. A screen-shot of the implementation of the table of contents can be found in Fig. 5. The drop down fields of the prototype are implemented with the commonly known triangles. The current scene is highlighted in reversed colors. Instead of marking the thumbnail of the current video with a colored frame, it is displayed enlarged.

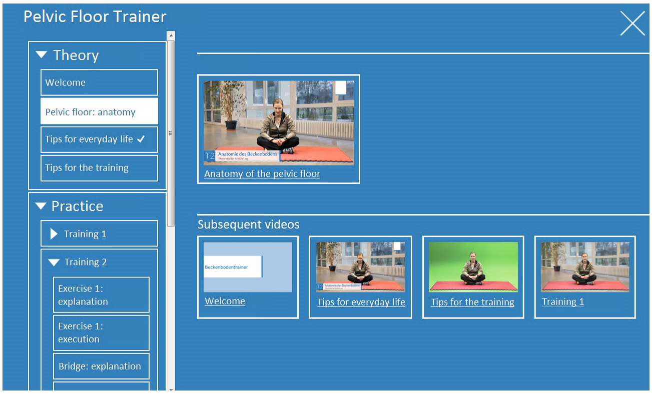

Fig. 5. Implemented version of the table of contents in the hypervideo player.

5 Final Evaluation

The implemented software was tested in a second user test.

Study Method and Participants. The test users had to perform similar tasks like those from the first test, but with the player and videos from a real scenario. The tasks included:

- Use of the table of contents as a help for navigation.
- Navigate trough the video to the end scene.
- Recognize already visited video scenes.

The second user test was also conducted with a group of eight people ($N = 8$) all aged above 45. None of the participants was a member of the first test user group. The same arrangements were made as in the first test (participants did not meet, introductory and explanatory text read from paper, basic questions). After that, the tasks enlisted above were presented to the test users one after the other after a brief introduction to the GUI of the prototype. The users had to perform each task on the tablet using the implemented software. In combination with that, the think-aloud technique [16, p.256] was used. Furthermore, the test users were asked to fill out the UEQ [9].

Result and Discussion. With the small number of answers, only a qualitative statement can be made, but the results indicate, that the implementation was evaluated positive in all categories of the UEQ [16, p.256], namely attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty.

Only few adaptations were made in the software after the user test. The basic structure was not changed. The representation of some elements was further improved: the button areas were enlarged and the distance between the thumbnails of the follow-up scenes was increased. The test users furthermore preferred a fully unfolded table of contents when it was first opened, which speaks against a menu structure which unfolds step-by-step and confirms the result of the survey.

6 Conclusion

In this work, we presented an extended presentation of a table of contents for hypervideos. The design was tested in a real world medical training scenario with the target group of people older than 45. This user group is a particular challenge since they sometimes have limited experience in the use of mobile devices. Furthermore, age-related physical deficiencies have to be taken into consideration for this user group.

We designed the user interface for an extended table of contents in three steps. An expert group and a survey were used to get hints on how a prototypical implementation should look like. The findings thereof were implemented in two different paper prototyped versions of the display. These were then tested in a user test with an in-group design, where each user had to test and evaluate both versions. As a result it can be noted that a divided view is desired. The table of contents should be presented in an easy-to-touch version on the left side of the view. A preview on the selected scene and follow-up scenes extracted from the underlying scene graph should be presented on the right side of the view. Unnecessary buttons from the underlying video player should be hidden to limit confusion. These findings were implemented in the existing open source SIVA

HTML5 player and tested with a second group of users. This test only revealed minor problems which were implemented in the final GUI.

However, this work has some limitations. The GUI was only tested on 7" and 10" tablets in landscape mode. Accordingly, the design should also be usable on monitors, but the usage in portrait mode needs a revision of the design. This is part of our future work.

References

1. Apostu, S., Al-Nuaimi, A., Steinbach, E., Fahrmair, M., Song, X., Möller, A.: Towards the design of an intuitive multi-view video navigation interface based on spatial information. In: Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2013, pp. 103–112. ACM, New York, NY, USA (2013)
2. Cunha, B., Pedrosa, D., Goularte, R., Pimentel, M.: Video annotation and navigation on mobile devices. In: Proceedings of the 18th Brazilian Symposium on Multimedia and the Web, WebMedia 2012, pp. 261–264. ACM, New York, NY, USA (2012)
3. Gerjets, P., Kirschner, P.: Learning from multimedia and hypermedia. Technology-Enhanced Learning, pp. 251–272. Springer, The Netherlands (2009)
4. Girgensohn, A., Wilcox, L., Shipman, F., Bly, S.: Designing affordances for the navigation of detail-on-demand hypervideo. In: Proceedings of the Working Conference on Advanced Visual Interfaces, AVI 2004, pp. 290–297. ACM, New York, NY, USA(2004)
5. Grigoras, R., Charvillat, V., Douze, M.: Optimizing hypervideo navigation using a Markov decision process approach. In: Proceedings of the Tenth ACM International Conference on Multimedia, MULTIMEDIA 2002, pp. 39–48. ACM, New York, NY, USA (2002)
6. Höffler, T., Leutner, D.: Instructional animation versus static pictures: a meta-analysis. Learn. Instr. **17**(6), 722–738 (2007)
7. Hürst, W., Götz, G., Welte, M.: Interactive video browsing on mobile devices. In: Proceedings of the 15th International Conference on Multimedia, MULTIMEDIA 2007, pp. 247–256. ACM, New York, NY, USA (2007)
8. Karrer, T., Wittenhagen, M., Borchers, J.: Pocketdragon: a direct manipulation video navigation interface for mobile devices. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2009, pp. 47:1–47:3. ACM, New York, NY, USA (2009)
9. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) USAB 2008. LNCS, vol. 5298, pp. 63–76. Springer, Heidelberg (2008)
10. Leitner, M., Wolkerstorfer, P., Sefelin, R., Tscheligi, M.: Mobile multimedia: identifying user values using the means-end theory. In: Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI 2008, pp. 167–175. ACM, New York, NY, USA (2008)
11. Meixner, B., Kosch, H.: Interactive non-linear video: definition and XML structure. In: Proceedings of the 2012 ACM Symposium on Document Engineering, DocEng 2012, pp. 49–58. ACM, New York, NY, USA (2012)
12. Meixner, B., Matusik, K., Grill, C., Kosch, H.: Towards an easy to Use authoring tool for interactive non-linear video. Multimedia Tools and Applications, pp. 1–26. Springer, US (2012)

13. Meixner, B.: Annotated interactive non-linear video - software suite, download and cache management. PhD thesis, Universität Passau (2014)
14. Miller, G., Fels, S., Ilich, M., Finke, M.M., Bauer, T., Wong, K., Mueller, S.: An end-to-end framework for multi-view video content: creating multiple-perspective hypervideo to view on mobile platforms. In: Anacleto, J.C., Fels, S., Graham, N., Kapralos, B., Saif El-Nasr, M., Stanley, K. (eds.) ICEC 2011. LNCS, vol. 6972, pp. 337–342. Springer, Heidelberg (2011)
15. Nguyen, C., Niu, Y., Liu, F.: Direct manipulation video navigation on touch screens. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2014, pp. 273–282. ACM, New York, NY, USA (2014)
16. Rogers, Y., Sharp, H., Preece, J.: Interaction Design: Beyond Human - Computer Interaction. Wiley, New York (2011)
17. Shipman, F., Girgensohn, A., Wilcox, L.: Combining spatial and navigational structure in the hyper-hitchcock hypervideo editor. In: Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia, HYPERTEXT 2003, pp. 124–125. ACM, New York, NY, USA (2003)
18. Tonndorf, K., Handschigl, C., Windscheid, J., Kosch, H., Granitzer, M.: The effect of non-linear structures on the usage of hypervideo for physical training. In: Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME 2015), IEEE, pp. 1–6 (2015)
19. Zahn, C., Barquero, B., Schwan, S.: Learning with hyperlinked videosdesign criteria and efficient strategies for using audiovisual hypermedia. Learn. Instr. **14**(3), 275–291 (2004)