H. Reiterer & O. Deussen (Hrsg.): Mensch & Computer 2012
München: Oldenbourg Verlag, 2012, S. 283-292

A Zooming Concept for an Interactive Non-linear Video Authoring Software

Britta Meixner, Christoph Grill, Harald Kosch

Chair of Distributed Information Systems, University of Passau

Abstract

Authoring tools for interactive non-linear videos or hypervideos often use the graph-based pattern for defining the structure between single elements of the multimedia-presentation. Visualization improvements like zooming and focus+context are rarely used. Elements of one category (for example videos, images, etc.) are represented by one uniform-looking icon, which makes it difficult to find elements in larger projects. This work illustrates that semantic and geometric zooming in combination with overview+detail views and a fisheye zoom are able to increase the attractiveness of the GUI and decrease the task completion time for several tasks.

1 Introduction

Interactive non-linear videos or hypervideos provide a non-linear structure and offer additional information with a main video. They can be used in e-learning, for tours through cities or houses and more. The underlying structure of such videos is an (extended) scene graph which links single scene with each other. Additional information (annotations) is linked with the scenes. Thereby a start and an end time for display is provided as well as the position where the annotation is displayed in the player. For example, a tour through the ground floor of an ordinary one family house can have about twenty scenes and thirty annotations. A well designed and clearly structured graphical user interface makes it possible for non-professionals to create their own interactive non-linear video. Bulterman and Hardman (2005) describe four different patterns for the authoring of multimedia presentations, namely the “structure-based pattern”, the “timeline-based pattern”, the “graph-based pattern” and the “script-based pattern”. Tools like the HyperProp system (Soares et al. 2000), Riva Producer Enterprise[1] and the hypervideo authoring system described by Chang et al. (2004) use the graph-based pattern for the creation of multimedia presentations. The scenes in these tools are hard to differentiate, because they are all represented by the same uniform icon. No hint

[1] http://www.matchware.com/en/ (accessed January 15, 2012)

on the content of the video is given. Annotations are either not visible in the graph representation or are shown on the same layer as the videos, which can be confusing for authors creating interactive non-linear videos, where the annotations form a unified whole with the main videos. Other authoring paradigms can for example be found in (Chambel et al. 2004), (Stahl et al. 2005) and (Stahl et al. 2006). These works focus on learning with hypervideo and describe different tools and paradigms for the creation of hypervideos. The Diver Project (Pea et al. 2004) allows creating new videos from existing ones, by cutting segments out of videos and by providing zoomed presentations of parts of a video. Zooming can be found in these works, too, but it is used to present a region of the video with more detail. A more detailed description of related work can be found in (Meixner et al. 2010).

2 Research Contributions

This work shows several possibilities to extend the graph-based paradigm in order to provide more information without switching to another paradigm. The main contributions of this work are:

- a concept for semantic zooming in the scene graph (Section 4),
- a concept for a fish-eye zoom in the scene graph (Section 4) and
- a usability evaluation of the implementation in comparison to the previous version of the scene graph (Section 5).

We also give an overview over related work dealing with semantic or fish-eye zooming (Section 3).

3 Related Work

According to Cockburn et al. (2009), four different schemes exist for the manipulation of datasets displayed in an interface. The approaches can be categorized as “overview+detail”, “zooming”, “focus+context” and “cue-based”, whereby only the first three are relevant for this work. **Overview+detail** provides different presentations of one information, whereby one presentation is a rough overview and the other one shows a smaller part with more details. Examples for overview+detail can be found at Google maps with street view or Microsoft PowerPoint. **Zooming** is a technique using “temporal separation”, whereby only one zoom step is visible at a time. Geometric zooming "enlarges objects while zooming in and shrinks them when zooming back out" (Buering et al. (2006)). Semantic zooming shows different types of information about an object when the magnification of an object changes (Perlin & Fox 1993). Approaches for semantic zooming can be found in Frisch et al.(2008), Stengel et al. (2011) and Mulloni et al. (2010). **Focus+context** combines overview+detail and zooming to one presentation. Elements are displayed with more details while the context around them is shown with less details (fisheye) (Furnas (1986)). Fisheyes can be found in Rao & Card (1994) or the Mac OS X Dock icon-panel. Semantic Fisheyes combining the lens-effect with a semantically more detailed presentation can be found in Zhang et al.

(2010) and Yoo et al. (2008). Combinations of fisheye and semantic zooming are presented by Paternò & Zini (2004) and Buering et al. (2006) whereas the latter uses a geometric-semantic zooming.

4 Zooming Concepts

Different zooming concepts are needed for a better navigation and usability of the scene graph of the SIVA Producer (see Meixner (2010)). Overview+detail is used to show which area of the graph is currently displayed, if the size of the graph does not fit the editor area. Using the Eclipse Rich Client Platform[2], the overview area is freely positionable and can be closed if not needed by the user. The overview (see figure 1, (1)) is integrated from very early versions of the software on and will not be examined in detail in this work. Geometric zooming with four different levels was also integrated during the implementation of the scene graph and is not object of this work. The scene graph provides only uniform icons (see green rectangles in figure 1, (2)) for all scenes which only differ in the name-labels of the scenes. We decided to use the concepts of semantic zooming and focus+context to provide more information to the user.

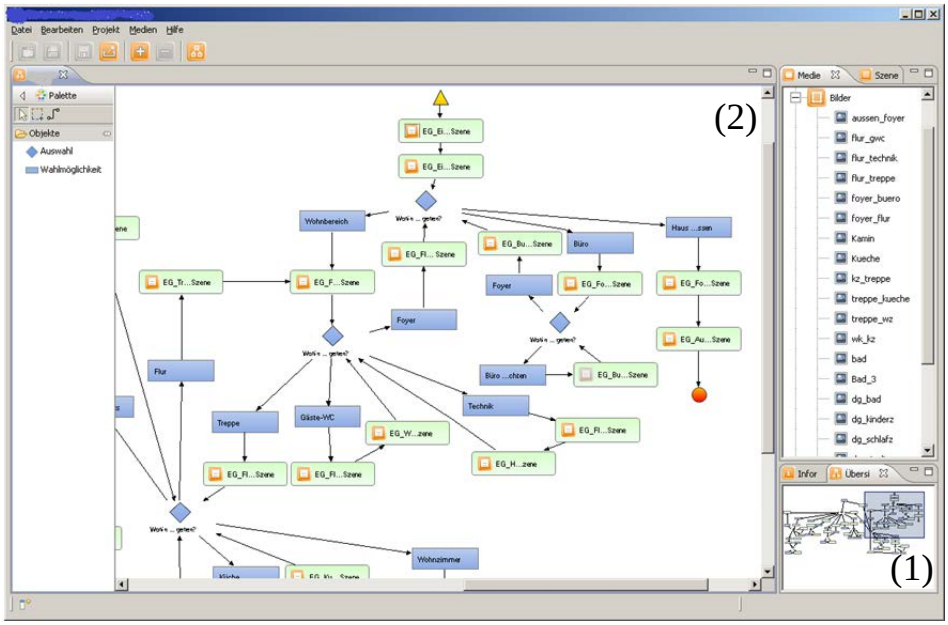

Figure 1: Previous version of the scene graph with overview (1) and detail (2) views

2 http://wiki.eclipse.org/index.php/Rich_Client_Platform (accessed March 16, 2012)

An expert group (see Choi (2009)) with three experts (*N = 3*) in the area of interactive non-linear video authoring was conducted to decide how many zoom levels make sense and which information should be displayed at which zoom level. Possible elements were determined in a brainstorming. The display of each element was discussed for each zoom level. First ideas for the layout of the graph nodes were discussed. It has been decided that three zoom levels and a semantic fisheye are useful which provide the following information to the user:

- Zoom level 1: name of the scene and annotation-icon
- Zoom level 2: name and preview of a scene, quantity of each kind of annotation
- Zoom level 3: name, preview, duration and source video of a scene; name, type and position of each annotation
- Semantic fisheye: more detailed scene preview, thumbnails of annotations

Based on the outcome and discussions during the expert group meeting, up to four meaningful paper prototypes (Snyder (2003)) were created and reviewed by members of the expert group. These were integrated into a survey (Salant & Dillman (1994)) to decide the following questions:

- Which icons are the most meaningful for indicating the absence/presence of an annotation in a scene?
- How to abbreviate long scene names?
- How to arrange the elements in zoom level 2?
- How to arrange the elements in zoom level 3?
- How to arrange the elements for a semantic fisheye?

The survey contained twelve questions and was conducted with Lime-Service[3]. Images and possible answers of subsequent questions were influenced by previously answered questions. If a certain icon was selected in one of the first questions, this icon was used for the wireframes in the following questions, too, for example. The questionnaire was carried out with 102 students/university employees (*N = 102*) between the ages of 20 and 35. The results are given in percent; the p-value is calculated with a binomial test. As a result, the following findings can be noted:

1. The most meaningful icon for an annotation (creation) is a pen (60.8 %) (others: 9,8 % - 26,5 % - 2,9 %) ($p < 0.001$).
2. The most meaningful icon for the presence of an annotation is a star-like symbol in the upper left corner of the icon (47 %) (others: 37 % - 16 %) ($p < 0.05$).

[3] http://www.limeservice.com/ (accessed March 9, 2012)

3. Long names should be abbreviated at the end with "…" (64 %) (others: 19 % - 17 %) ($p < 0.001$).
4. The video preview in zoom level 2 should be positioned above the annotation information but below the scene name (57 %) (others: 28 % - 15 %) ($p < 0.001$).
5. The position information of an annotation in zoom level 3 should either be positioned before or after the annotation name (before: 54 %, after: 46 %) ($p = 0.11$).
6. Scene information like scene name and duration should be placed below the scene preview but above the annotation information (70 %) ($p < 0.001$).
7. No clear result emerged on the question how the fisheye should be realized. One possible solution is the expansion of a scene node (49 %), the other solution is a tooltip-behavior (51 %) ($p = 0.69$).

The results of the survey were summarized to final paper prototypes due to their statistical significance. The result can be found in figure 2. We decided to put the position information (result 5, ($p = 0.11$)) before the annotation name for a better overview. This leads to a grouping of the icons for the type and the position of the annotation. Furthermore, we decided to implement the fisheye (result 7, ($p = 0.69$)) with a tooltip-behavior, because it is easier to handle in the graph and does not lead to confusion with zoom level 2. The appearance of a scene node in zoom level 1 is quite similar to the original appearance except the new annotation icon and another abbreviation of long scene names. Zoom level 2 shows two preview images of a scene as well as a compact list of annotations and their quantity. The most detailed level is zoom level 3, where duration and source video of a scene are shown below the scene preview. Each annotation is enlisted with icon, position in the player and name. The fisheye shows more preview images of the scene and only thumbnails of the annotations as well as the scene name in full text.

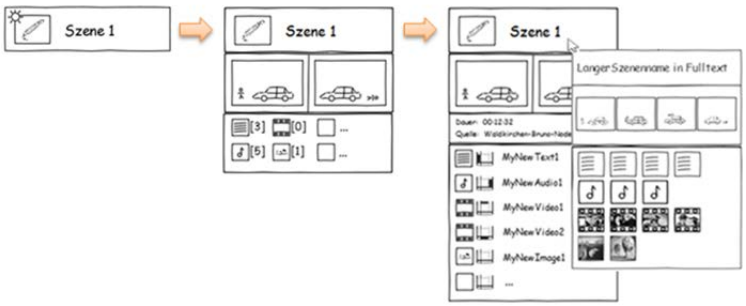

Figure 2: Paper prototypes of zoom levels 1 to 3 and the semantic fisheye as a tooltip

Screenshots of the implementation can be found in figure 3 and figure 4. The color design was fit to the rest of the GUI of SIVA Producer with mainly grey and orange color shades. A zigzag line was inserted to indicate that the preview images of the scene are part of a video.

Due to the geometric magnification of the scene nodes from zoom level 1 to 3, overlapping of scene nodes can occur. To avoid overlapping, a graph layout algorithm was implemented. It is called Fast-Node-Overlap-Removal and can be found in Dwyer et al. (2006) and Dwyer et al. (2007).

Figure 3: Screenshot of the semantic fisheye showing two frames of the scene, two image annotations and one text annotation

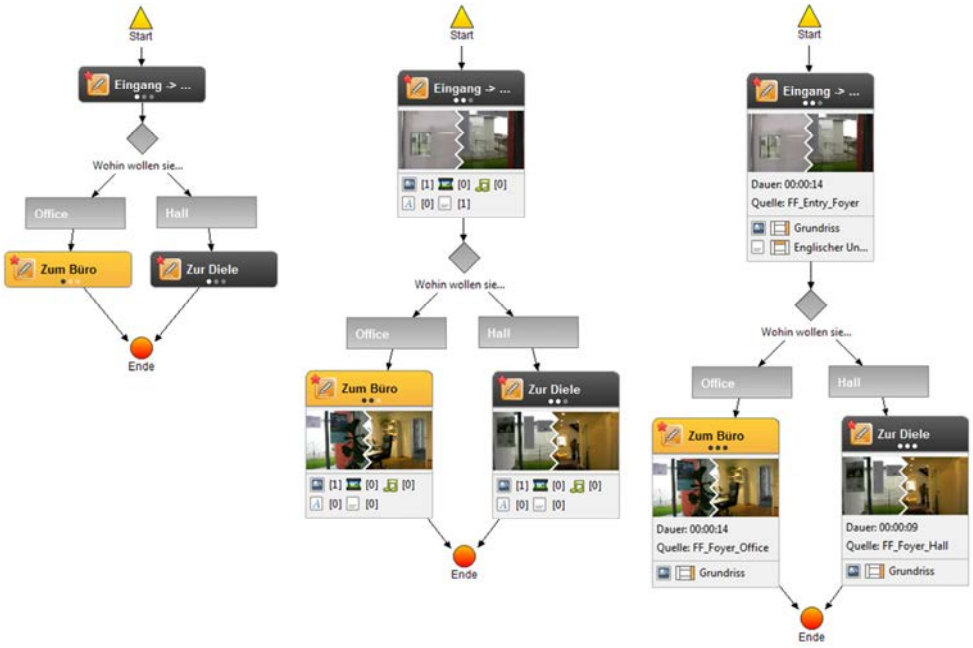

Figure 4: Screenshots of all zoom levels: Zoom level 1 (left) shows only scene names and indicates that each of the three scenes contains annotations. Zoom level 2 (centered) shows a preview of the scene and how many annotations are added to a scene. Zoom level 3 (right) is the most detailed one; it shows two frames of a scene, the duration of the scene and the source video. Furthermore the title of an annotation is displayed as well as its type and position.

5 User Evaluation

First user tests were performed to show that the proposed test methods are suitable for more extensive testing. The goal of the tests is to show that the new graph layout is more intuitive and allows shorter task completion times. User observation and the User Experience Questionnaire (Laugwitz et al. (2008)) were conducted with twenty users (*N = 20*). These were divided into two groups of ten users each. The test users were aged between 20 and 28 (group 1: $\overline{x_1} = 24.5, \sigma_{x_1} = 2.16$, group 2: $\overline{x_2} = 23.7, \sigma_{x_2} = 1.9$) and distributed randomly in two disjoint groups; a between subject design was used. Group 1 tested the old and group 2 the new graph layout. The users had to perform the ten tasks enlisted in table 1 after a short introduction to the basic functions of the software. The tasks are typical for working with the GUI of the scene graph. Time for completing the different tasks was taken for each task.

ID	Question	Average task completion time		*p*
		old layout	new layout	
1	Describe the content of the graph in about 5-7 sentences.	30.8 s ($\sigma = 5.06$)	33.4 s ($\sigma = 7.29$)	0.261
2	What shows the “entrance”-video?	14.4 s ($\sigma = 2.65$)	9.4 s ($\sigma = 2.69$)	< 0.001
3	Name all scenes with a “layout”-annotation.	49.7 s ($\sigma = 7.21$)	25.9 s ($\sigma = 8.68$)	< 0.001
4	Which scene has the highest quantity of annotations?	39.9 s ($\sigma = 8.46$)	23.4 s ($\sigma = 5.55$)	< 0.001
5	Does scene “to the exit” contain annotations?	3.1 s ($\sigma = 0.70$)	3.2 s ($\sigma = 1.54$)	0.847
6	How many annotations contains scene “entrance→foyer”?	6.7 s ($\sigma = 2.00$)	7.8 s ($\sigma = 2.71$)	0.293
7	How many scenes contain subtitles?	46.4 s ($\sigma = 7.32$)	17.2 s ($\sigma = 3.09$)	< 0.001
8	Are video-annotations used in this interactive video?	39.8 s ($\sigma = 7.67$)	14.1 s ($\sigma = 3.21$)	< 0.001
9	How long is scene “hall→foyer”?	8.2 s ($\sigma = 2.04$)	8.3 s ($\sigma = 2.61$)	0.921
10	Which is the source video of scene “entrance→foyer”?	15.5 s ($\sigma = 4.98$)	8.6 s ($\sigma = 2.24$)	< 0.001

Table 1: Task completion times for different tasks and corresponding p-values (calculated by Welch's t-test)

The results of both tests can only be seen as a rough trend because of the small user groups. Summing up table 1, we can say that the new layout shows a better task completion time for eight out of ten tasks. It can be noted that tasks which affect only one scene are completed only slightly faster (5s at the maximum) with the new graph layout. Tasks concerning more than one scene or a comparison of scenes reveal substantial better task completion times (and better significance values) using the new scene graph. Both groups were asked to complete the UEQ, which can only show a tendency because of the small test group. Further statistical evaluations do not make sense because of the small test group. The results (see figure 5) reveal tendentially better values in all categories for the new layout of the scene graph than

for the old one. Especially the values for attractiveness, efficiency and novelty show a rise. As a result of the pre-test, it can be noted that both tests are suitable for testing with larger user groups.

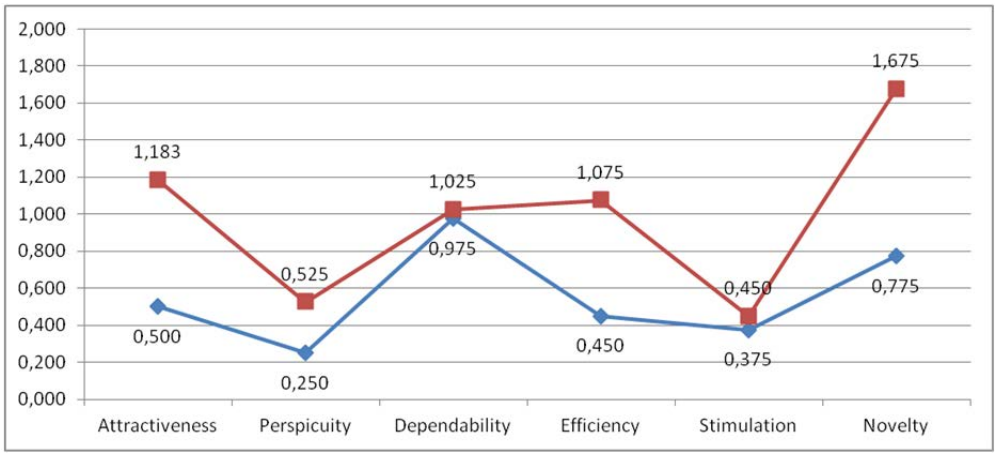

Figure 5: Results of the UEQ

6 Conclusion

We have introduced layouts for three different semantic zoom levels as well as a fisheye tooltip in this work. The content of each zoom level and the fisheye was determined by an expert group. The findings of this expert group were transferred into paper prototypes which were used in a survey to find the most intuitive solution for each question. The implemented solution was tested with ten users and compared to the old version of the scene graph, which was also tested with ten users in a pre-test. This pre-test was used to ensure the suitability of the user test (with measurement of the task completion time) and the UEQ for testing with larger user groups. Qualitative findings showed that the task completion time could be improved in eight out of ten tasks, especially in tasks concerning more than one scene or a comparison of scenes. Furthermore a UEQ indicated that especially the attractiveness, the efficiency and the novelty were improved in the new version of the scene graph. Tests with more users are necessary to be able to provide quantitative and statistically significant values. While some results in the survey had a clear tendency to one of the possible solutions, no clear decision was made according to the arrangement of the position information of an annotation in zoom level 3. Furthermore, the realization of the fisheye had no clear tendency. Further tests are necessary to decide which solution provides a better user experience.

Acknowledgements

This work was partially 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 2.0”.

References

Buering, T., Gerken, J. & Reiterer, H. (2006). *User Interaction with Scatterplots on Small Screens - A Comparative Evaluation of Geometric-Semantic Zoom and Fisheye Distortion*, IEEE Transactions on Visualization and Computer Graphics, Vol.12, No.5, pp. 829-836

Bulterman, D. C. A. & Hardman, L. (2005). *Structured multimedia authoring. ACM Trans. Multimedia Comput. Commun. Appl.* 1, 1, ACM, New York, NY, USA, pp. 89-109

Chambel, T., Zahn, C. & Finke, M. (2004). *Hypervideo design and support for contextualized learning.* IEEE International Conference on Advanced Learning Technologies, 2004. Proceedings. pp. 345-349

Chang, H. B., Huang Hsu, H., Liao, Y. C., Shih, T. & Tang, C. T. (2004). *An object-based hypervideo authoring system.* In: IEEE International Conference on Multimedia and Expo, 2004. ICME '04. vol. 3, pp. 2219 – 2222

Choi, Y.J. (2009). *Providing Novel and Useful Data for Game Development Using Usability Expert Evaluation and Testing*. In *Proceedings of the 2009 Sixth International Conference on Computer Graphics, Imaging and Visualization* (CGIV '09), Washington, DC, USA: IEEE Computer Society, pp. 129 –132

Cockburn, A., Karlson, A. & Bederson, B. B. (2009). *A review of overview+detail, zooming, and focus+context interfaces.* ACM Comput. Surv. 41, 1, Article 2 (January 2009), New York, NY, USA: ACM, pp. 2:1-2:31

Dwyer, T., Marriott, K. & Stuckey, P. J. (2006). *Fast node overlap removal*. In Proceedings of the 13th International Conference on Graph Drawing, GD'05, Berlin, Heidelberg: Springer-Verlag, pp. 153–164

Dwyer, T., Marriott, K. & Stuckey, P. J (2007). *Fast node overlap removal: correction*. In Proceedings of the 14th International Conference on Graph Drawing, GD'06, Berlin, Heidelberg: Springer-Verlag, pp. 446–447

Ericsson, K. A. & Simon. H. A. (1998). *How to study thinking in everyday life: Contrasting Think-Aloud protocols with descriptions and explanations of thinking*. Mind, Culture, and Activity, 5(3):178–186

Frisch, M., Dachselt, R. & Brückmann, T. (2008). *Towards seamless semantic zooming techniques for UML diagrams.* In *Proceedings of the 4th ACM Symposium on Software Visualization* (SoftVis '08). New York, NY, USA: ACM, pp. 207-208

Furnas, G. W. (1986). *Generalized fisheye views*. In Mantei, M. & Orbeton, P. (Ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '86), New York, NY, USA: ACM, pp. 16-23

Laugwitz, B., Held, T. & Schrepp, M. (2008). *Construction and Evaluation of a User Experience Questionnaire*. In *Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work* (USAB '08), Holzinger, A. (Ed.), Berlin, Heidelberg: Springer-Verlag, pp. 63-76

Meixner, B., Siegel, B., Hölbling, G., Lehner, F. & Kosch, H. (2010). *SIVA suite: authoring system and player for interactive non-linear videos*. In Proceedings of the International Conference on Multimedia (MM '10). ACM, New York, NY, USA, pp. 1563-1566

Mulloni, A., Dünser, A. & Schmalstieg, D. (2010). *Zooming interfaces for augmented reality browsers*. In Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10). New York, NY, USA: ACM, pp. 161-170

Paternò, F. & Zini, E. (2004). *Applying information visualization techniques to visual representations of task models*. In *Proceedings of the 3rd Annual Conference on Task Models and Diagrams* (TAMODIA '04). New York, NY, USA: ACM, pp. 105-111

Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W. & Hoffert, E. (2004). *The Diver Project: Interactive Digital Video Repurposing*. IEEE MultiMedia 11, 1 (January 2004), pp. 54-61

Perlin, K. & Fox, D. (1993). *Pad: an alternative approach to the computer interface*. In Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '93). New York, NY, USA: ACM, pp. 57-64

Rao, R. & Card, S. K. (1994). *The table lens: merging graphical and symbolic representations in an interactive focus+context visualization for tabular information*. In Plaisant, C. (Ed.): *Conference Companion on Human Factors in Computing Systems* (CHI '94), New York, NY, USA: ACM, pp. 222-230

Salant, P. & Dillman, D. (1994). *How to conduct your own survey*. Business Reference. Wiley

Snyder, C. (2003). *Paper prototyping: the fast and easy way to design and refine user interfaces*. TheMorgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann (2003)

Soares, L. F. G., Rodrigues, R. F. & Saade, D. C. M. (2000). *Modeling, authoring and formatting hypermedia documents in the hyperprop system.* Multimedia Syst. 8, pp. 118–134

Stahl, E., Zahn, C. & Finke, M. (2005). *How can we use hypervideo design projects to construct knowledge in university courses?*. In Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The next 10 Years! (CSCL '05). International Society of the Learning Sciences, pp. 641-646

Stahl, E., Finke, M. & Zahn, C. (2006). *Knowledge Acquisition by Hypervideo Design: An Instructional Program for University Courses*. Journal of Educational Multimedia and Hypermedia. 15 (3), pp. 285-302. Chesapeake, VA: AACE

Stengel, M., Frisch, M., Apel, S., Feigenspan, J., Kästner, C. & Dachselt, R. (2011). *View infinity: a zoomable interface for feature-oriented software development*. In *Proceedings of the 33rd International Conference on Software Engineering* (ICSE '11). New York, NY, USA: ACM, pp. 1031-1033

Yoo, B., Lea, J. & Kim, Y. (2008). *The seamless browser: enhancing the speed of web browsing by zooming and preview thumbnails.* In Proceedings of the 17th International Conference on World Wide Web (WWW '08). New York, NY, USA: ACM, pp. 1019-1020

Zhang, K., Wang, H., Tran, D. T. & Yu, Y. (2010). *ZoomRDF: semantic fisheye zooming on RDF data*. In Proceedings of the 19th International Conference on World Wide Web (WWW '10). New York, NY, USA: ACM, pp. 1329-1332

Contact Information

Britta Meixner
Chair of Distributed Information Systems, University of Passau
Innstraße 43, 94032 Passau