This is a preliminary version of an article published by
S. Wilk, S. Kopf, S. Schulz, W. Effelsberg:
Social Video: A Collaborative Video Annotation Environment to Support E-Learning.
Proc. of World Conference on Educational Multimedia, Hypermedia and Telecommunications (EdMedia), pp. 1228-1237, June 2013.

Social Video: A Collaborative Video Annotation Environment to Support E-Learning

Stefan Wilk
Distributed Multimedia Systems
Technical University of Darmstadt
Darmstadt, Germany
stefan.wilk@cs.tu-darmstadt.de

Stephan Kopf
Department of Computer Science IV
University of Mannheim
Mannheim, Germany
kopf@informatik.uni-mannheim.de

Wolfgang Effelsberg
Distributed Multimedia Systems
Technical University of Darmstadt
Darmstadt, Germany
wolfgang.effelsberg@cs.tu-darmstadt.de

Abstract: Our social video system allows users to enrich video by additional information like external websites, hypertext, images, other videos, or communication channels. Users are able to annotate whole videos, scenes, and objects in the video. We do not focus on a single user accessing the system but on multiple users watching the video and accessing the annotations others have created. Our web-based prototype differs from classical hypervideo systems because it allows annotation (authoring) and navigation in videos by focusing on collaboration and communication between the users. The prototype is integrated into the online social network Facebook and was evaluated with more than 300 users. The evaluation analyzes the usage of the system with a learning scenario in mind and indicates a learning success of users.

Introduction

Despite the huge success of social media, videos do not fit very well into their context. Platforms such as *YouTube* or *Yahoo! Video* are the state of the art of combining social media and videos. They allow users to share, rate, and comment videos.

The typical functionality which is usually available in case of social media is limited in case of videos:

- Video players do only support a very limited user interaction, e.g., pressing buttons like play, pause, stop, or using a slider to move to a certain point in time.
- The annotation of web pages and even images is supported by social media today, e.g., by defining image regions with the mouse and adding labels or tags. A real annotation of videos is not possible.
- Another important aspect of social media is the collaboration and joint work of users. Besides sharing of content a real collaboration is not supported for videos.
- Social media uses web-based technologies for navigation and annotation. Video streaming is supported today, but web based annotation or editing of videos is not possible.

In analogy to the term social media we define *social video* to allow the creation and interaction of user-generated content, and to support social interaction with other users while watching or annotating videos. We believe that such a system supports knowledge construction if it allows the communication between users and provides functionality to annotate the content. Social video extends a special form of interactive video called hypervideos. A *hypervideo* maps the idea of hypertext – using hyperlinks to refer to external websites – to videos by allowing users to interact with objects in a video. We enhance the classical idea of hypervideo (the combination of hypermedia and video) by embedding fundamental parts of social interaction like communication and collaborative access of enriched videos.

A second aspect we consider is to use social videos in the context of e-learning. General problems are the linear structure of videos and the lack of possibilities to adapt to the speed of learners. These reasons limit the application of videos in knowledge construction despite their advantages. We believe that learning can be improved when using social videos and providing a system that is intuitive to use.

The major contributions of our approach are:

- Our social video system offers a single user interface that supports users to collaboratively annotate and access information by using a standard web browser.
- E-learning is supported by the immediate sharing of annotations between active users and additional tools for collaboration.
- Due to the integration into the social network Facebook, a large number of users tested and evaluated the system. A learning scenario shows that the user interface supports users in accessing relevant information and allows them to embed their own ideas.

The paper is structured as follows: The next section describes existing systems and ideas concerning videos and user interfaces for videos. We present the system architecture of our *interactive video learning system* afterwards. The processes of enriching video with additional information (annotation) and accessing the desired information are explained. The evaluation analyzes in detail how users accept and use our system. Finally, we conclude the paper and give a brief outlook on future work.

Related Work

By mapping the idea of hypertext to videos, the concept of hypervideo was born. The early systems did only respond to the need for accessing information. VideoBook (Ogawa, Harada, & Kaneko, 1992), for example, uses time-based scenario-dependent links to support video-to-video linking. HyperCafe (Sawhney, Balcom, & Smith, 1996) lets users decide the plot line of a story. In a cafe setting, users of the system choose which conversations they want to follow. The DIVER project (Pea et al., 2004) uses a panoramic camera for the recording of lectures in a classroom setting. The system facilitates the annotation of single regions in frames by addition of textual comments.

Early hypervideo systems, that were especially designed for knowledge construction like HTIMEL (HTML with Time Extensions) (Chambel, Zahn, & Finke, 2006), focused on distance learning using a web-based interface. The GUI design of HTIMEL clearly separates the video from additional texts or images. Interactivity is thus limited to follow static links. Hyper-Hitchcock (Girgensohn, Shipman, & Wilcox, 2003) implements a concept called detail-on-demand, which is a specialized version of a hypervideo. A major goal is to reduce the negative effects of hypermedia on users like disorientation. It restricts the amount of information by only allowing one outgoing link of a video at a time. The Interactive Shared Educational Environment (Marchionini & Pattee, 2003) allows users to communicate ideas via an internal chat system and to combine active reading and video. The idea to collaboratively watch videos was investigated by Zahn et al. (Zahn, Barquero, & Schwan, 2004). The authors further investigated how the success of learning is influenced by hypervideos, multimedia, and hypermedia. A user study shows that hypervideos are an effective tool to support learning. In addition, knowledge construction based on hypervideos was intensively researched for university courses (Stahl, Zahn, & Finke, 2005) and for learning processes in veterinary surgeries (Tiellet, Pereira, Reategui, Lima, & Chambel, 2010).

ToolClips (Grossman & Fitzmaurice, 2010) implements and validates the concept that, compared to text, videos support the learning processes better and that they help users to understand program functionalities much easier. The SmartPlayer project (Cheng, Luo, Chen, & Chu, 2009) adjusts the playback speed in response to the level of the user interest of a scene or a shot. SmartPlayer thus can adapt to the user's need in regard of the learning speed. In contrast to those video learning environments our systems allows users to pause videos and access additional information or even connect videos explaining details as well as discussing topics with friends.

System Overview

Our interactive video system implements the idea of hypervideos by combining videos with additional information. Considering the scenario of a collaborative video learning environment, the system provides the basic functionalities for users and their friends to access (navigate) and extend (annotate) the video collaboratively. It is implemented as a distributed web-based system, which makes it accessible via standard web-browsers and facilitates the easy integration into arbitrary web pages or social network sites. The system extends classical client-server models by an adaptive cloud-based server allocation approach. This adaptive approach is required due to the computational load of the system mainly caused by the automatic tracking algorithms that position annotated information according to the movement of objects in the video. Multiple servers responsible for the social video delivery are supported by one central registration server that clients connect on startup.

The ***client software*** implements navigation and annotation in one single user interface. The social video clients are responsible for bringing annotated videos to the users, allowing them to access information and to embed their own ideas. Figure 1 visualizes the graphical user interface of our system. A main goal of the system is to support users in finding the information they need. Besides the interactive video area, additional information can be accessed in the structural area, the dynamic information area, and the options menu (Figure 1). The interactive video area is required for user interaction and the starting point for learning processes. It shows the current social video and furthermore allows access to additional information via interactive hotspots. Each hotspot highlights a specific object and therefore has to move and resize according to the changes of the object. The structural area is comparable to a table of contents that facilitates a direct access to the in- and outgoing hyperlinks of the current social video. The structural area includes the video object tree, which shows all video objects and information nodes of the current social video in a tree view. The dynamic information area provides different tools to facilitate the access of information. It embeds the central chat interface, which allows the communication with friends from Facebook and active users of the system. In addition, it allows using and editing annotations of scenes and video clips.

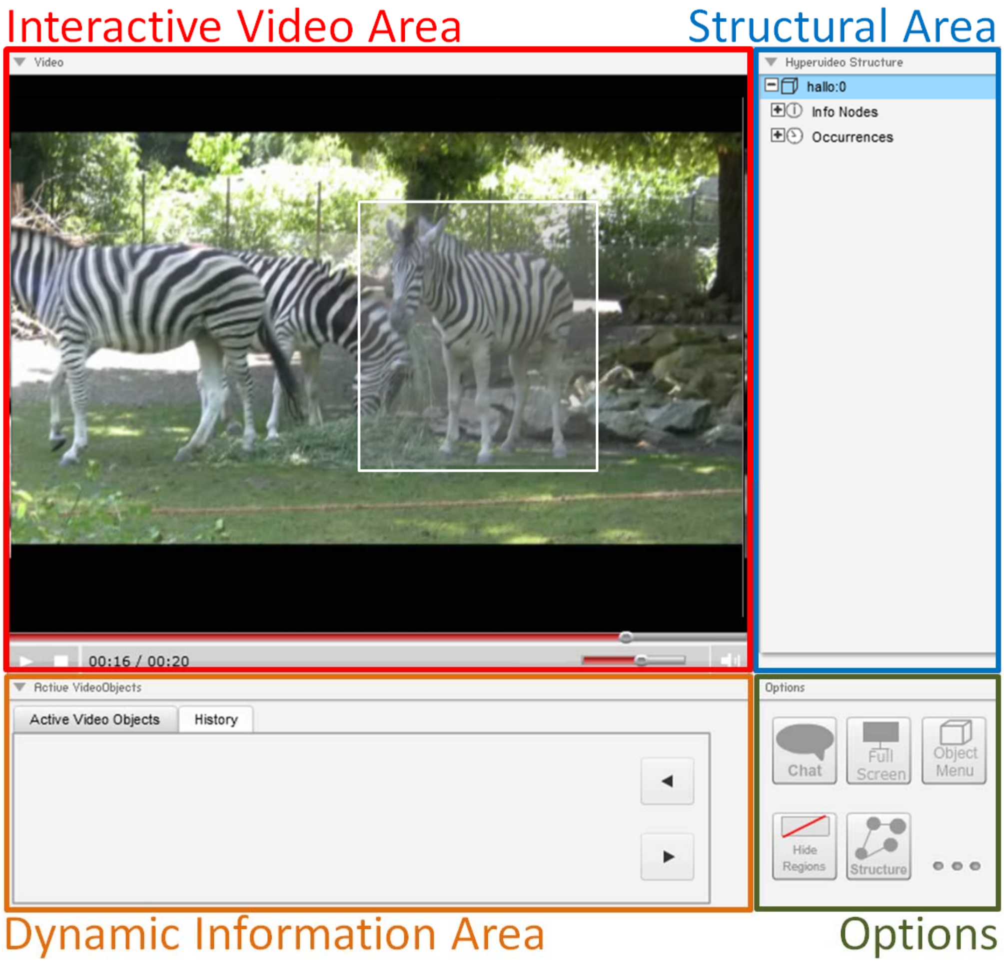

Figure 1: Overview of the graphical user interface of the interactive social video system

Options give access to a range of customization properties and additional tools. The client interface combines the two main concepts of annotation and navigation. During the playback of a video, a user can access any information visualized in hotspots in the video area or edit the content by directly clicking on the video. Annotating an object is triggered by drawing a rectangular region around the object. The initial marking of an object invokes the object tracking functionality on the support server. In contrast to many previous systems, objects are tracked automatically and a user does not have to define video objects in every frame.
The **central registration server** is the initial point of communication and refers clients to their designated **support server**. Annotations are saved independent of the video file in a social video document. A personalized copy of every social video document is sent to the client at the initial contact with the system. The social video document is managed by the assigned support server individually for each client. Consistency between the different clients is eventually guaranteed. The automatic tracking of objects in videos is the main task of the support servers. An adaptive algorithm selects the most suitable tracking technique from a diverse set of algorithms, including MeanShift, template-based matching, optical flow, or matching based on SURF-descriptors. Details about the implemented robust tracking approach responsible for tracking objects in the video can be found in Wilk et al. (Wilk, Kopf, & Effelsberg, 2012).

Annotation of Videos

Annotations made to a video are collected in data containers. We call such a container video object if it describes one concrete object within the video, e.g., a person or an animal. The visualization for the user of a video object is an interactive hotspot. The interactive hotspot is a rectangular overlay of an object in the video that is positioned and scaled according to the movements and deformations of the object (Figure 1).

The annotation – also called authoring – is a central concept of our social video system that allows every user to bring own ideas and knowledge into the video. A novel concept is to understand annotation not as a centralized but as a distributed and, at the same time, social process. Multiple users can work on the same objects at the same time. This turns our social video system into a platform for exchanging knowledge in a social way. To motivate users not only to consume information but also to participate in creating new content, the annotation process should be intuitive and very easy to use. Thus, to add information, the following three step process was implemented: create a new data container, add additional information to the container, and let the system automatically visualize it.

Creating a new data container is easily done by using the mouse and by clicking on the interactive video area. By drawing a rectangular region, the object in the current frame is marked and the first step of the annotation process is done. The initial marking of the object automatically invokes the server-based object tracking.

In a second step, users are able to **add additional information** to a data container. Multiple links may be added to one container. The prototype implements four types of links which are called information nodes. The system allows adding hypertext rendered within the system (**image-text node**), links to external webpages (**hyperlink node**), to refer to other social videos (**video node**), and the creation of a new chat topic (**communication node**). The combination of videos is essential in the context of the social video system. Administrators are able to preselect adequate video clips and offer them to users. However, the system makes it also possible for users to upload their own videos and to add links to existing ones. Every new video can automatically be used as a social video, and users can immediately annotate it after the upload and the automatic preprocessing by the server. Communication is essential for collaboration. **Communication nodes** generate new topics in the integrated chat system. The chat allows users to exchange questions, ideas, or comments. Each communication node supports a specific communication about one distinct video object and helps to understand of the content of the video clip. The content of a communication is persisted in the system, so that any user can access and continue the communication at any time. The chat system integrates other social media and supports the communication with Facebook users via Facebook Chat [1].

[1] https://developers.facebook.com/blog/post/2010/02/10/integrating-facebook-chat-everywhere/

Another central component of the system is the **automatic visualization**. Its task is to transform the annotated video object into a hotspot that automatically follows its movement. Please refer to Kopf et al. (Kopf, Wilk, & Effelsberg, 2012) for more information about the different nodes.

Navigation in Interactive Video Environments

Information in hypermedia is loosely structured, which allows learners to explore and discover knowledge in a nonlinear way. This makes the search of specific information much more difficult. Thus, a typical problem in hypermedia learning environments is the fact that users get disoriented. This is especially true for videos due to the additional complexity caused by the dynamics of a video. Our educational hypermedia system reduces the effect of disorientation by offering multiple possibilities for navigation. To support the navigation of users, we especially focus on the following questions a user might have:

- Which possibilities for navigation are available?
- In which video and point in time is the user currently located?
- Which parts of the video did the user visit before?

To make information accessible, a hotspot is visualized as an overlay of the annotated video object. Users can interact with the object by clicking on its hotspot. The system pauses the video and displays the information nodes in a radial menu. This menu includes an arbitrary number of associated information nodes. The video object carousel (Figure 2, left) maps all navigation elements that can be currently accessed in the interactive Video Area. With this component we want to give access to the active video objects independent of their position and size in the current frame. The carousel is integrated into the dynamic information area. Another possibility to access information is to use the video object tree (Figure 2, center) that shows all video objects of the current video. It differs from hotspots, because – independent of the motion, size, and visibility of the annotated object – the tree makes it possible to access object information at any time.

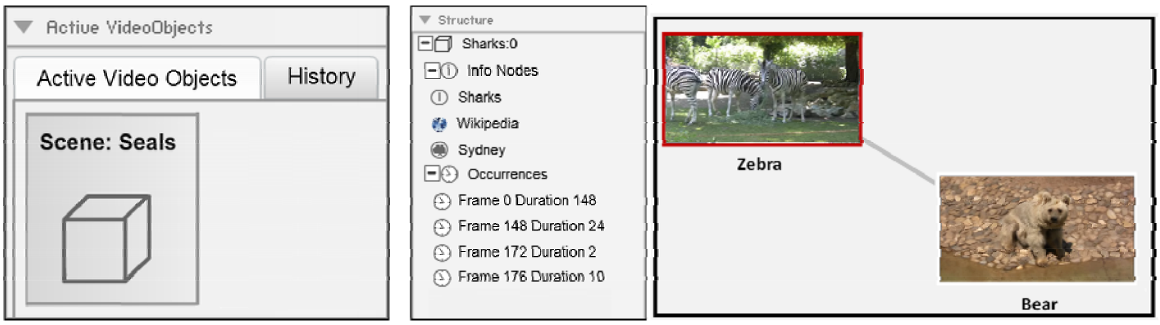

Figure 2: Video object carousel / history (left), video object tree (center) and structural view (right)

The concept of social videos is so powerful because they represent a rich space of connected information. On the other hand, this might also cause disorientation of users. To avoid disorientation, a user should always know where he is located in the hypermedia space. The structural view was designed to support users to determine their position in the overall video space (Figure 2, right). It visualizes all linked videos as a graph. A user can retrieve his current position in the hypermedia space and see possible navigation paths. Whenever the user moves the mouse to a node, a preview of the video is started that provides additional information about its content. By clicking on a node, the system jumps directly to the social video and starts playing.

EVALUATION

The evaluation of the system consists of a detailed analysis of the system in the context of usability and support of the user, as well as a small experiment about knowledge construction based on social videos in comparison with classical lecture recordings. For our initial prototype we have designed the system in a way to be easily embedded into the Online Social Network Facebook [1]. We have developed a survey module that is fully integrated into the social video system as part of the interactive video area. Upon using a feature of the social video system for the first time a web-based dialog is displayed that gathers information on strengths as well as weaknesses of the feature using a questionnaire. The questions have to be answered using a 5-step Likert-scale (McDaniel & Gates, 1998 p.247). For the rest of this work 1 is the worst and 5 the best score to achieve. In addition, to collect data about the usage of the system, interactions are logged when and how often a user activates a navigation tool. The evaluation with the survey tool was stopped after 300 fully completed evaluation data sets. The second experiment described in this work analyzes how well the system may be used for e-learning. The learning success of two user groups – the first using the social video system and the other group watching classical recordings – are compared.

Navigation and Annotation in a Social Video System

Accessing embedded information and linked videos is one of the most important features of the system. During the first evaluation experiment, 300 users evaluated different features of the social video system. Videos included two non-related topics. Users could use the system with videos and predefined information about 'zoology' and 'how to mix cocktails'. The video clips include eleven predefined video objects in total. Since interactive hotspots are designed as mean to access additional information, it was interesting to see, that users accepted with a majority (88%) that multiple hotspots are shown at a time. As long as none of the interactive zones overlap, users welcome the possibility to access as much information as possible. This is only the case as the interactive hotspots do not deter attention from the story being told. From the feedback of the users we, e.g., retrieved that choosing the background color is essential to not destroy the authenticity of the video. The best colors that we have chosen consisted of slight intensity variations of the predominant color in a video shot. Analyzing the logs, we observe that hotspots are intensively used. Around 69.96 % of all mouse clicks that were intended to retrieve information from video objects were done using interactive hotspots.

The structural view was intensively used in combination with the video object tree. Together, both tools resulted in 22.13 % that were intended to retrieve additional information from video objects. From the feedback of the users and inspection of the logs a typical usage pattern can be retrieved. The structural view was used to navigate between the different videos that are relevant for the current learning topic. When an interesting topic was found the video object tree was used to jump to the interesting part of the video clip. Thus users accessed both tools to search for interesting scenes, e.g., when they were interested in a special animal (topic 'zoology') or a special step in a recipe for a cocktail (topic: 'mixing a cocktail'). Support for this implication is given by the answers made in regard of the important features of the video object tree. We have designed the tool to easily access information and jump to different interesting video scenes but in addition the hierarchical layout and the possibility to use it as a table of content about what is happening are important for users. Both structural view (average: 4.24; variance: 0.31) as well as the video object tree (average: 3.82; variance: 0.92) gained high user satisfaction (1: very low, 5 very high). The combination was rated by two thirds of all users as good or very good (ratings 4 or 5). Especially the transparency about possible content a user may consume was seen as very important. But in contrast to the interactive hotspots (average: 4.07; variance: 1.01) and the structural view (average: 4.2; variance: 0.77) the usage of the video object tree (average: 3.44; variance: 0.94) seems to be more complicated.

In contrast to the previous described navigation tools the active video objects carousel (6.24 % of all mouse clicks), intended as an addition to the interactive hotspot, as well as the history component (1.66 % of all mouse clicks) were less frequently used. In addition, less benefit was seen by the users to access the active video carousel (benefit from using: average: 2.6; variance: 1.16) or the history (benefit from using: average: 3.09;

[1] https://www.facebook.com

variance: 0.64) of the system.

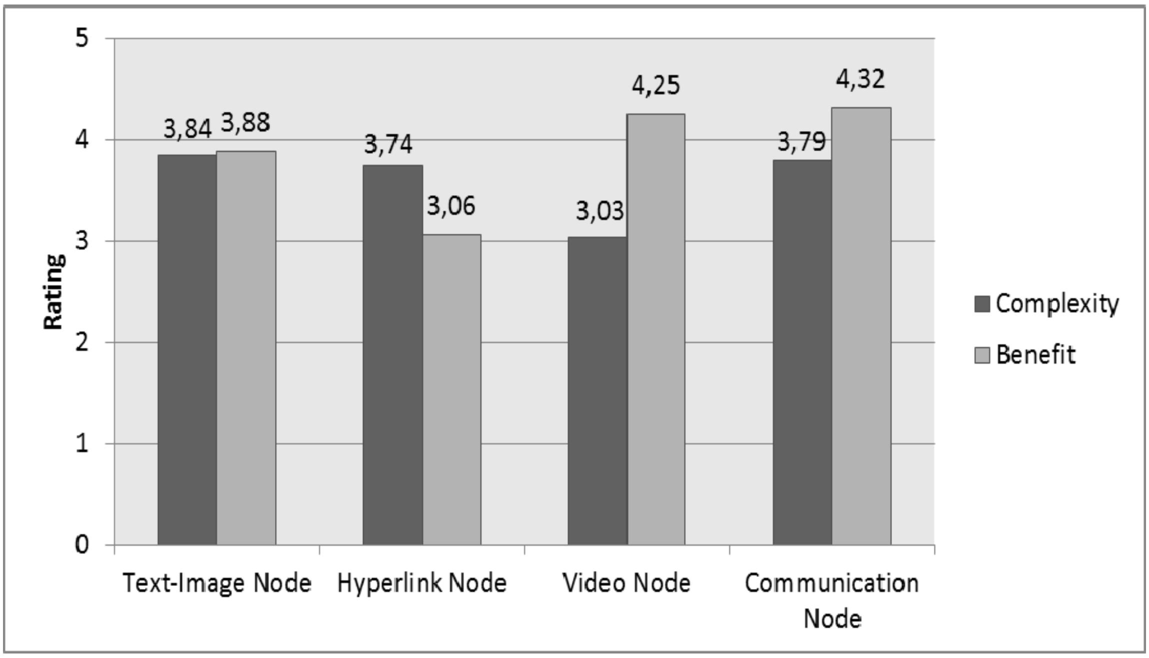

Figure 3: Information Nodes - Ratings for benefit of using each type of node (1: very little; 5: very high) and complexity of creating a new node (1: very complex; 5: very easy)

Besides different ways to access video object the types of information that can be associated with the social video document were of interest. Figure 3 shows the results of the evaluation for the benefit the users gained from using the individual node as well as the complexity to create them. With an average value of 4.32 (variance: 0.8), the benefit of communication nodes is higher compared to all other nodes. 96 communication nodes were initiated in the system that included 56 distinct questions or discussion. All questions were answered or discussed in the evaluation period. Users are highly motivated to help each other. The usage pattern for collaborative work on the system resulted in most of the cases that users added additional information nodes to an existing video object. Thus, a rich set of information was gathered once someone started to annotate an object in the video. Furthermore, 136 changes were made to existing information nodes. In most of the cases (107) the changes concerned text-image nodes. Users did enhance texts and added images to existing content. This observation is strongly supported by an additional benefit of 3.87 (variance: 0.81) when users are asked whether connecting own ideas with existing content is supported by the system. Another collaborative aspect is the fact that users of social media invite friends to participate. With a mean opinion score of 3.64 (variance: 1.2) most of the users tend to agree that self-controlled learning with videos can be significantly enhanced using additional information presented in a video upon request. Interesting to know that annotating videos was seen as even more important (average: 4.04; variance: 1.2). Embedding own ideas and the feature to ask questions using communication nodes support learners in adjusting videos to their own needs. Thus, collaboration and group based interaction in videos seem to be essential for the success of learning with social video. The ease of usage of the social video system was rated with a mean opinion score of 4.1 (variance: 0.79)

Impact on Knowledge Construction

The preceding sections show that the social video system is easy to use and that it is highly accepted by users. This section analyzes whether or not the social video system can be used for learning and knowledge construction. The following evaluation uses several video clips that describe how a Wordpress Blog [1] is created. The experiment compares the knowledge gain of an experimental group and a reference group. In contrast to the evaluation using a prototype integrated into Facebook this experiment has been held in a controlled environment – classroom setup with 15 students (10 users evaluating the social video system; 5 users watching pure video – reference group). The experiment lasted 1 hour but every user could choose when to finish the video session. In the initial setup 4 video clips with 32 information nodes were created. To eliminate advantages of the social video group all information added upfront was also part of the narration in the video clips. A pretest was constructed to determine the knowledge of users. After using the social video system a retest was performed that build upon the questionnaire of the pretest and added 10 more questions per topic.

Video Clip	Duration (mm:ss)	# Information nodes predefined	Average percentage of used information nodes
Wordpress – Main	12:47	12	32 %
Wordpress – Create Sections	0:31	5	53 %
Wordpress – Headers	0:52	6	62 %
Wordpress – Widgets	1:02	8	67 %

Table 1: Video clips of the topic 'Wordpress Blog'

After the pretest, the social video group uses the system whereas the video group does only access the videos using a normal video player. Users of both groups can access the same content. In comparison to the results of the pre-test, the learning effect in both groups is very high. The quota of correct answers of the social video group is 75.3% (variance: 8.5%; absolute increase in comparison to pretest: 58.7 %) compared to 54.7% of the video group (variance: 11.1%; absolute increase in comparison to pretest: 34.17 %). On average, the users of the social video group accessed the system 39.1% longer (video group - average usage time 20 minutes 40 seconds; social video group - average usage time: 28 minutes 45 seconds). We also identified a negative correlation between the length of a single social video sequence and the learning effect. Accessing longer videos decreases the learning effect at the same time. The percentage of correct answers after watching the main sequence of the learning scenario is around 48% whereas the effect in all other – much shorter – sequences is about 86%. Another observation is the usage of information nodes in social videos. 32% of the existing nodes are accessed in long sequences in comparison to 61.7% in short video sequences. The positive learning effects by offering information in videos seem to decrease as the duration of a single video clip increases. In correlation, the degree of information nodes accessed by a user is lower (Table 1). In this experiment the average duration of users watching the video clips differed from 26 minutes and 23 seconds in the social video group to 16 minutes 53 seconds in the reference group.

After reassignment of the groups a second scenario was constructed using the 'mixing a cocktail'-topic. While analyzing the usage patterns, it became obvious that especially the image-text nodes helped to gain an overview of ingredients and steps necessary for creating a cocktail. During this run the reassigned social video group reached 73 % (absolute increase in comparison to pretest: 55.9 %) of correct answers whereas 48.4 % (absolute increase in comparison to pretest: 39.2 %) was reached by the video group.

[1] http://www.wordpress.org

Discussion

The social video concept aims at combining videos with social media concept. We believe that social video is a natural evolution of Web 2.0 by collaboratively enriching video documents with links as well as communication topics. Nowadays, social media focuses on the web-based creation and exchange of user-generated data like text, web pages, or images. Although many websites allow users to view, upload, and share videos, the editing of the video content similar to the concept of the Web 2.0 is not supported yet. Our social video system makes the first step towards true social video. Our collaborative system allows users to access the common knowledge base and to enrich social video with their own ideas by adding hyperlinks, videos, text, and images. The user interface makes it possible to access all functionalities in one screen. This shifts the fundamental separation of presentation and authoring environment of existing video systems to a more collaborative, social media approach.
Offering an intuitive user interface and unified processes for navigation and authoring, the social video system was accepted by users. They especially liked the support for navigation between information nodes. Still, some interface components like the carousel component were more difficult to use. Due to the modular concept of the overall system, these components could easily be removed.

We believe that the presented system introduces social aspects in interactive video systems which may improve learning effects. Our social video group performed better than the comparison group. Especially when a group of short video clips was used, the possibility to retrieve additional information provided by video objects was used regularly. The results in the knowledge construction section can only be the start of a more detailed analysis on how social video supports collaborative learning.

Conclusion and Outlook

Our interactive video system is especially designed to help users to navigate in hypermedia space and to enrich it with own ideas. The evaluation results make it clear, that the system supports both concepts and is highly accepted as a multimedia learning environment. The navigation functionalities help to quickly and precisely access the information a user is looking for. Especially the combination of the structural view, that represents a graph of associated social videos, and the video object tree, that allows accessing detailed information in social videos, is most useful. We found some indications that such a system could improve learning scenarios, if concentration and focus of the scholars can be ensured. Especially long-running video sequences include the danger of mitigating the positive effect of social videos.

Still, the current implementation does not support a direct search of information nodes by entering keywords. A future version of the system should integrate a search engine to support a domain wide search in a rich set of social videos. Furthermore, new types of information nodes should be integrated to enhance the possibilities for collaboration.

Acknowledgements

This work has been co-funded by the German Research Foundation (DFG) in the Collaborative Research Center (SFB) 1053 „MAKI – Multi-Mechanism Adaptation for the Future Internet“.

References

Chambel, T., Zahn, C., & Finke, M. (2006). Hypervideo and Cognition: Designing Video-Based Hypermedia for Individual Learning and Collaborative Knowledge Building. In E. Alkalifa (Ed.), *Cognitively informed systems: utilizing practical approaches to enrich information presentation and transfer* (1. ed., pp. 26–49). Idea Group Pub.

Cheng, K.-Y., Luo, S.-J., Chen, B.-Y., & Chu, H.-H. (2009). SmartPlayer: User-centric video fast-forwarding. *Proceedings of the 27th international conference on Human factors in computing systems - CHI '09* (pp. 789–798). New York, New York, USA: ACM Press.

Girgensohn, A., Shipman, F., & Wilcox, L. (2003). Hyper-Hitchcock: Towards the Easy Authoring of Interactive Video. *Human-Computer Interaction INTERACT 2003* (pp. 33–40). IOS Press.

Grossman, T., & Fitzmaurice, G. (2010). ToolClips: an investigation of contextual video assistance for functionality understanding. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (pp. 1515–1524). New York, New York, USA: ACM Press.

Kopf, S., Wilk, S., & Effelsberg, W. (2012). Bringing Videos to Social Media. *2012 IEEE International Conference on Multimedia and Expo* (pp. 681–686). IEEE.

Marchionini, G., & Pattee, A. (2003). The interactive shared educational environment: user interface, system architecture and field study. *Joint Conference on Digital Libraries 2003* (pp. 291–300). Houston, USA: IEEE Comput. Soc.

McDaniel, C. D., & Gates, R. H. (1998). *Marketing research essentials* (1. ed., p. 462). Cincinnati: Taylor & Francis.

Ogawa, R., Harada, H., & Kaneko, A. (1992). Scenario-based hypermedia: a model and a system. In N. Streitz, A. Rizk, & J. André (Eds.), *Hypertext: concepts, systems and applications* (1st ed., pp. 38–51). New York: Cambridge University Press.

Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. (2004). The diver project: interactive digital video repurposing. *IEEE Multimedia*, *11*(1), 54–61.

Sawhney, N., Balcom, D., & Smith, I. (1996). HyperCafe: Narratic and Aesthetic Properties of Hypervideo. *Proceedings of the the seventh ACM conference on Hypertext - HYPERTEXT '96* (pp. 1–10). New York, USA: ACM Press.

Stahl, E., Zahn, C., & Finke, M. (2005). How can we use hypervideo design projects to construct knowledge in university courses? (pp. 641–646). Taipei, Taiwan: International Society of the Learning Sciences.

Tiellet, C. a. B., Pereira, A. G., Reategui, E. B., Lima, J. V., & Chambel, T. (2010). Design and evaluation of a hypervideo environment to support veterinary surgery learning. *Proceedings of the 21st ACM conference on Hypertext and hypermedia - HT '10*, 213.

Wilk, S., Kopf, S., & Effelsberg, W. (2012). Robust tracking for interactive social video. *2012 IEEE Workshop on the Applications of Computer Vision (WACV)* (pp. 105–110). IEEE.

Zahn, C., Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos - design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction*, *14*(3), 275–291.