

L.IVE: An Integrated Interactive Video-based Learning Environment

Toni-Jan Keith Monserrat¹
tjmonsi@gmail.com

Yawen Li²
iamliyawen@gmail.com

Shengdong Zhao¹
zhaosd@comp.nus.edu.sg

Xiang Cao³
xiangcao@acm.org

¹NUS-HCI Lab, National University of Singapore, Singapore

²NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore

³Lenovo, Beijing, China

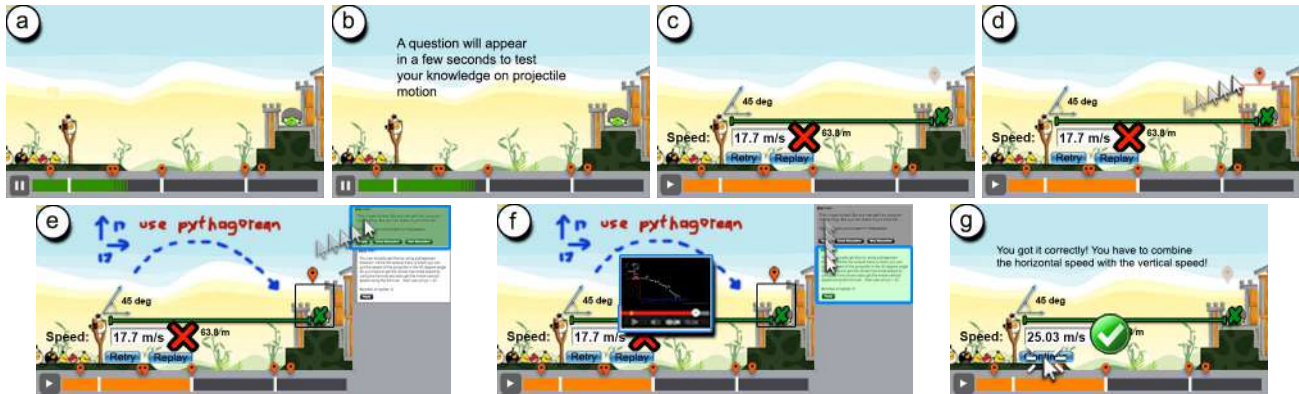


Figure 1: A physics lecture video based on Angry Birds is played on the L.IVE system (a). A lecturer uses an embedded text annotation to indicate a quiz is coming (b). The video is automatically paused with an in-place quiz that asks the speed required to hit the green target. Users' answers will be assessed and feedback given based on their individual responses (c). Users see a comment tag above the green target, and move the cursor over as the tag fades in (d). Clicking on the comment tag displays the comments on the side. Hovering over the comments reveals an associated annotation, which can be hand-drawn (e) or shown as a linked video (f). Correctly answering the quiz question advances the video (g).

ABSTRACT

In this paper, we introduce L.IVE: an online interactive video-based learning environment with an alternative design and architecture that integrates three major interface components: video, comment threads, and assessments. This is in contrast with the approach of existing interfaces which visually separate these components. Our study, which compares L.IVE with existing popular video-based learning environments, suggests advantages in this integrated approach as compared to the separated approach in learning.

Author Keywords

L.IVE; interactive video; video-based online learning

ACM Classification Keywords

H.5.1. Information interfaces and presentation: Multimedia Information Systems - Video

INTRODUCTION

Recently, video-based online learning environments have emerged as a hot topic in education. Some popular

examples include Khan Academy, Udacity, edX, Coursera, and other websites that post their videos in YouTube educational channels. Together, they offer a great variety of courses and attract millions of users each month [2].

Although the content, design, and layout differ among these environments, they all have three interface elements to support key educational activities: videos to deliver the lecture content (learning), comment threads for sharing (discussion), and assessments to check learning progress (assessment) [5].

Despite their successes, current interface designs seem to have one drawback: the three components are presented separately, either by page layout or page separation (Fig. 2 left). This potentially creates additional burdens for learners to cognitively link different sources of information together, thereby impairing the effectiveness of their learning [1]. For example, a comment placed below the video may contain useful information for learners to understand the key lecture concepts; however, it can be easily overlooked, as most people tend to focus on the video. Furthermore, because the duration of each video typically spans at least a few minutes, it can be difficult for learners to identify which part of the video the comment is associated with.

Similarly, while it may seem natural to assess users' learning after watching a video segment, visually separating assessment questions from video content can make it harder for learners to recall the relevant information for answering questions. This may be less ideal as Constructivism

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2473-1/14/04 \$15.00.
<http://dx.doi.org/10.1145/2556288.2557368>

suggests that it is beneficial to allow learners to explore and access relevant information “just-in-time,” so they can better construct mental images for enhanced learning [4].



Figure 2: Current video-based learning environments visually separate the video, comments, and assessments (left). L.I.V.E tightly integrates the three components (right).

Therefore we propose an alternative design to visually integrate comment threads and assessment questions tightly with the video. We created the L.I.V.E (Integrated Interactive Video-based Environment for Learning) prototype to test the potential benefit of the proposed integrated approach as compared with the current separated approach. Our evaluation of 18 participants indicated that they learned more efficiently with the L.I.V.E integrated approach than with the baseline approach: participants showed 20% greater scores, measured as post-test performance compared with pre-test performance, when they experienced content with L.I.V.E. Our post-experimental questionnaire also indicated that most participants preferred the integrated approach as compared with the separated approach.

The concept of adding either comment thread or assessment to videos is not new. However, based on our analysis, users’ benefit may be maximized when both comment thread and assessment are tightly integrated with the video. We developed L.I.V.E, the first system and design to integrate all three components and apply them in the context of online video-based learning. Further, we also contribute an architecture to facilitate the implementation of this integrated design on the web, where video-based learning systems are becoming increasingly popular. Finally, we conducted the first empirical evaluation to reveal the potential benefits of this integrated design for online video-based learning.

RELATED WORK

Numerous works have added various types of interactive elements in videos. Due to space constraints, we can only list a few representative ones. The Hvet Design [11] used links in videos to external information to improve user learning. NoteVideo reverse-engineered the video to convert visual glyphs of lecture notes into clickable links to make it easier for learners to access the different parts of video and has shown the new interface improves various learning tasks [9]. CWaCTool implemented in-place annotations and comments in video for sharing and discussing annotated information [10]. In-place annotation and in-context comment threads have also been tried in documents [12]. In addition, learning assessments have been included as part of games [3], and most recently be experimented in video [6].

While the previous work has provided inspirations, there has not been an attempt to integrate the three components of online video-based learning environment together. Furthermore, there hasn’t been a formal evaluation to reveal the potential benefits/drawbacks of such integrated approach for learning and education, hence providing motivation for our work.

L.I.V.E SYSTEM

Our vision for the L.I.V.E system is for it to be an interactive video player that allows users to view lecture video, discuss with both the teaching staff and peers, and assess their learning within or in close proximity to the video itself. The goal is to tightly connect information presented from the different sources to provide an enriched, holistic learning experience.

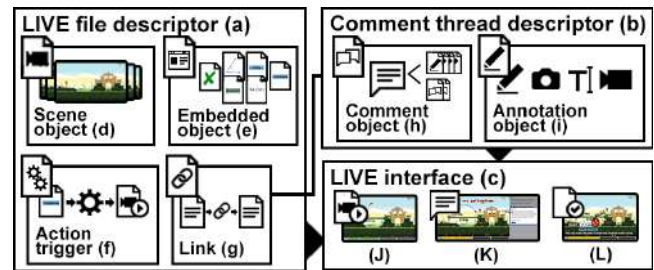


Figure 3. An overview of the Live system

Figure 1 illustrates an example of using the L.I.V.E system in a physics lecture on projectile motion. As in the example in the figure, educational information is often presented in three main formats: video, comment threads, and assessments. Though used somewhat differently, all three are equally important in working together to enhance the user’s ability to learn efficiently. Existing systems separates these elements and may hamper with the user’s ease of learning of abstract and complex information. Thus, there is a need for an interface design and underlying software architecture to support integrated organization and presentation of information that connects the three source components and allow users to easily interact with them.

System Implementation and Architecture

We constructed a data structure to organize the different interactive elements based on their spatial and temporal relationships to the video as well as their logical relevance to the lecture content. This hierarchy of information is described in two files called the *Interactive Video File Descriptor* (Fig. 3a) and the *Comment Thread Descriptor* (Fig. 3b). Both are JSON documents that are based on SIVA’s interactive video XML document design but are modified to allow in-place objects in video rather than at pre-defined places (top, bottom, left, right) [8].

Interactive video file descriptor

The file descriptor defines four types of entities: a list of scene objects, a list of embedded objects, a separate list for action triggers, and a link to a comment thread descriptor (Fig. 3). *Scene Objects* (Fig. 3d) define the main videos (Fig. 1a) and their starting and ending times and can be

linked to each other to provide a seamless flow to create one integrated timeline. *Embedded Objects* (Fig. 3e) define integrated interactive elements such as assessment forms, buttons for navigation, and additional in-place information (i.e., texts, sub-videos, and images, see Fig. 1b). They have a starting time of appearance and an ending time of appearance. The location of the element in the video space is defined by its x and y coordinates. *Action Triggers* (Fig. 3f) links to events, like showing or hiding interactive elements, are encapsulated in an action trigger's data definition. User interactions on elements will usually call these action triggers to issue a change in the interface. Link (Fig. 3g) is a vector that points to a separate comment thread descriptor for the L.IVE interface to load.

Comment thread file descriptor

All user annotations and comment threads are encapsulated in a separate file called a *Comment Thread Descriptor* (Fig. 3b). The separation of the comment thread descriptor and the interactive video file descriptor allows users to share the interactive video with their peers without copying the comment threads of the original interactive video. A comment thread descriptor contains a list of comment objects and annotation objects. Comment object is either a comment thread starter or a reply to a comment (Fig. 3h). A comment thread starter type has data connecting it to a scene object and holds x and y coordinate and a timestamp of the video. This visually integrates it to the video. A comment reply is connected to another comment data object, creating a comment thread. Annotation object is a user-created in-video annotation encapsulated in an annotation object (Fig. 3i) and is connected to comment data. Each annotation object can be any of these types depending on their encapsulated object data: free-hand drawing, text, image, or video. The annotation object's location is defined by its x and y data coordinates.

The system is implemented using three main web technologies: HTML5, Cascading Style Sheets (CSS), and Javascript (with JQuery 1.8 Library and JSON), which are responsible for the structure, style, rendering, and interactivity of all objects and action triggers in the L.IVE interface. The current mechanism of authoring a L.IVE video is through manually editing the markup language in a text editor. The WYSIWYG interface for authoring the L.IVE video is in current development.

USER STUDY

To evaluate our design, we performed an experiment to test the L.IVE interface and compare it with a current video-based online learning interface (baseline) (Fig. 2) to investigate any potential differences in their abilities to facilitate learning.

Participants: Eighteen participants (6 females), ranging in age from 20 to 29 years were recruited from within the university community to take part in the experiment. 15 of them have previous experience with existing online video-based learning environments.

Apparatus: The study was conducted using a desktop computer running on Windows 7 OS with the usual mouse-keyboard input. The L.IVE system's implementation was as previously described. The baseline system was implemented using HTML5, CSS, and JavaScript.

Task and stimuli: Two 10-minute biology videos from Khan Academy were selected as the lecture videos¹. Each video has 10 comments and 3 assessment quizzes. All comments were selected from the existing comment threads on these videos from Khan Academy. We made some modifications to the comments by removing unrelated posts and answered some questions with text explanations and/or links to external information and videos. We also added x and y coordinates to fix appearance of linked videos on the L.IVE interface. The structure and appearance of comments for both conditions were the same. We also developed the assessment quizzes. Because assessments were not embedded in the baseline, they were implemented in the usual way: after every 2-3 minutes of a video segment, the users were taken to another page to complete the assessment. The timing of the assessment in the baseline is the same as in L.IVE.

Design and procedure: A within-participant design was used. Each participant watched both videos (v1, v2) using both interfaces (L.IVE, baseline). The order of the interfaces was counter-balanced while the order of the videos remained the same (e.g., participant 1 watched v1 using L.IVE, then watched v2 using baseline; participant 2 watched v1 using baseline, then watched v2 using L.IVE).

To measure the knowledge gain, for each video, a 10-question pre-test (before watching) and a 10-question post-test (after watching) were administered to each participant. The two sets of questions tests the same type of knowledge but were asked in slightly different ways (i.e. one question would be asked in identification form with a why question after, and the other question would be in essay form. The difference in the number of questions answered correctly in the pre- and post- tests was recorded. After they finished watching both videos using the two interfaces, participants shared their preferences and experience in an interview.

Before the experiment, participants were asked to spend about 2 minutes each to familiarize with the two interfaces. Each participant performed the entire experiment in one sitting, including breaks, in approximately 1 hour.

RESULTS

Two tailed t-Tests with 5% alpha-level revealed that there was a significant effect of interface type on difference of pre-test and post-tests scores. The percentage scores of participants using L.IVE interface (54.21%, pre-test score=41, post-test score=8.3, gain=7.89) was significantly higher than when using the baseline interface (44.36%, pre-

¹ v1) <http://bit.ly/14iSTFS>
v2) <http://bit.ly/15CQ3wB>

test score=3.7, post-test score=6.8, gain=6.43) ($t_{17} = 2.98, p < 0.01$). The L.IVE interface resulted in an additional 22% score as compared with the baseline interface. However, the overall learning time participants spent using the L.IVE interface (16 min 38 sec) were not significantly different from time spent using the baseline interface (15 min 44 sec) ($p = 0.62$). For embedded assessments, all participants eventually completed the assessments correctly after doing the experiment. Participants answered L.IVE and baseline assessments incorrectly, on average, 1.11 and 1.83 times, respectively. This indicates an average of 0.72 or 65% more unsuccessful attempts when using baseline.

In addition to the results above, the majority of participants preferred the design of the L.IVE system (13/18) over the baseline design (5/18). Participants expressed that the in-context annotations, comment threads and assessments were helpful in getting to know the bigger picture of the information. The ease of access to information in comments while watching video helped them understand and absorb more information. The in-context assessments also helped them to recall information.

DISCUSSION

The feedback of the participants also provided additional insights that can guide the design of future video-based learning environments. Although most participants preferred the L.IVE interface, a few participants still preferred the baseline interface as it allowed them to focus on the video first. This suggests that the L.IVE system may not be suitable for everyone; it is recommended that future video-learning environments provide an option for users to switch back to the traditional interface if they prefer.

In addition, while 10 comments are easy to manage, participants pointed out that an excessive number of comments may clutter the video interface and distract their attention. It is suggested that comments and annotations should be monitored and managed so that irrelevant ones can be filtered out and potential visual clutter can be minimized or avoided [7].

Lastly, participants suggested that it would be useful to allow personalized comments targeted to specific audiences instead of the entire viewing group. For example, a student may want to raise a question only for the teaching staff to look at; alternatively, one may want to start a discussion only with several of her close friends.

CONCLUSION

In this paper, we proposed an alternative design to visually integrate comment threads and assessment questions tightly within video. We implemented this alternative design in a prototyping L.IVE system and contributed a system architecture for organizing and presenting information from the three components on the web. Our evaluation of 18 participants indicated that they learned more efficiently with the L.IVE integrated approach than with baseline, showing

20% score gain with the former. Our post-experimental questionnaire and interviews revealed that most participants preferred the integrated approach as compared with the separated approach. Their input also highlighted potential challenges to watch for when deploying the system. In the future, we would like to evaluate L.IVE in a real online course setting and explore other values or experiences our integrated approach can bring to areas other than learning.

ACKNOWLEDGEMENT

This project is funded by the National Research Foundation (NRF) and managed through the multi-agency Interactive & Digital Media Programme Office (IDMPO) hosted by the Media Development Authority of Singapore (MDA) under Centre of Social Media Innovations for Communities (COSMIC).

REFERENCES

1. Ainsworth, S. and Van Labeke, N. Multiple Forms of Dynamic Representation. *Learning and Instruction* 14, 3 (2004), 241–255.
2. Faviero, B.B.F. Major players in online education market - The Tech. 2012. <http://tech.mit.edu/V132/N34/education.html>.
3. Harpstead, E., Myers, B.A., and Alevan, V. In-search of Learning: Facilitating Data Analysis in Educational Games. *In CHI '13*, (2013), 79.
4. Jonassen, D. Designing Constructivist Learning Environments. *In Instructional design theories and models: A new paradigm of instructional theory 2*. 1999, 215–239.
5. Kellogg, S. Online learning: How to make a MOOC. *Nature* 499, 7458 (2013), 369–371.
6. Macwilliam, T., Aquino, R.J., and Malan, D.J. Engaging Students Through Video: Integrating Assessment and Instrumentation. *Journal of Computing Sciences in Colleges* 28, 6 (2013), 169–178.
7. Mayer, R. Cognitive Theory of Multimedia Learning. *In The Cambridge handbook of multimedia learning*. 2005.
8. Meixner, B. and Kosch, H. Interactive Non-linear Video: Definition and XML Structure. *In DocEng '12*, (2012), 49.
9. Monserrat, T.-J.K.P., Zhao, S., McGee, K., and Pandey, A.V. NoteVideo: Facilitating Navigation of Blackboard-style Lecture Videos. *In CHI '13*, (2013), 1139.
10. Motti, V.G., Fagá, R., Catellan, R.G., Pimentel, M.D.G.C., and Teixeira, C. a. C. Collaborative synchronous video annotation via the watch-and-comment paradigm. *In EuroITV '09*, (2009), 67.
11. Tiellet, C.A.B., Pereira, A.G., Reategui, E.B., Lima, J.V., and Chambel, T. Design and evaluation of a hypervideo environment to support veterinary surgery learning. *In HT '10*, (2010), 213.
12. Zyto, S., Karger, D., Ackerman, M., and Mahajan, S. Successful classroom deployment of a social document annotation system. *In CHI '12*, (2012), 1883.