CThru: Exploration in a Video-Centered Information Space for Educational Purposes

Hao Jiang[1,3] **Alain Viel**[2] **Meekal Bajaj**[1,4] **Robert A. Lue**[2] **Chia Shen**[1]

[1]Initiative in Innovative Computing, Harvard University
[2]Department of Molecular and Cellular Biology, Harvard University
[3]Department of Computer Science, Tsinghua University
[4]College of Computing, Georgia Institute of Technology

h-jiang@mails.thu.edu.cn {aviel@fas.|robert_lue@|chia_shen@}harvard.edu meekalbajaj@gatech.edu

ABSTRACT
We present CThru, a self-guided video-based educational environment in a large multi-display setting. We employ a video-centered approach, creating and combining multimedia contents of different formats with a story-telling education video. With the support of new display form factors in the environment, viewing a sequential educational video thread is replaced by the immersive learning experience of hands-on exploration and manipulation in a multi-dimensional information space. We demonstrate CThru with an animation clip in cellular biology, supplementing visible objects in the video with rich domain-specific multimedia information and interactive 3D models. We describe CThru's design rationale and implementation. We also discuss a pilot study and what it revealed with respect to CThru's interface and the usage pattern of the tabletop and the associated large wall display.

Author Keywords
Video-centered information space, self-exploration, multi-display environment.

ACM Classification Keywords
H5.1. Information interfaces and presentation (e.g., HCI): Multimedia Information Systems.

INTRODUCTION
Video is an efficient and engaging visual method to convey concrete educational concepts from both the macro and micro perspectives. Used appropriately, videos can be motivating and effective as a springboard to start interactive learning, and can take students on impossible field trips, such as travelling through the inside of our cells. "Information-rich realistic contexts, often referred to as 'macro-contexts', encourage the active construction of knowledge by learners, ..." [6]. Learners often have questions while watching an educational video.

Conventional video players do not provide interactivity within the video content, though allowing the viewers to stop, and to go back and forth to re-view the clips. It will be useful and could enhance the learning experience if other types of information, possibly at different details, levels and modalities, are also directly accessible within the video. Hypervideo [8], though useful in commercial applications, has severe drawbacks in the learning domain: simply linking to external information source which is decoupled from the video sequence, users are easily lost from the main track (video), therefore it is not suitable to be applied to educational processes.

Educational instruments can benefit from emerging technologies. Display form factors, in particular, have changed, and improved to accommodate various requirements: large, pixel-rich displays prevail in their visualization power, while touch sensitive surfaces provide under-the-finger interaction tangible possibility. Although showing potentials, they have yet been fully exploited in or applied to the education domain.

In recognition of the above issues, we have developed CThru, a video-centered information space for self-guided learning. In CThru, core multimedia education contents of different formats are used to construct a cross-referenced information space running along with the main video, conveying a variety of domain specific knowledge and information. Users watching the video can freely jump to or out of the information space for in-depth knowledge, thus a sequential video playback is replaced by the experience of immersing and exploring in a multi-dimensional space. CThru runs in a multi-display environment containing a large display wall and an interactive tabletop (Figure 1).

Figure 1. CThru running in a multi-display setting.

BACKGROUND AND RELATED WORK

We demonstrate CThru in a cellular biology education scenario, based on the animation *"Inner Life of the Cell"* (http://multimedia.mcb.harvard.edu, referred to as *Inner Life* below). *Inner Life* is a 3D computer graphics animation describing biological mechanisms occur within a human leukocyte. The 8.5 minutes' animation covers various aspects of leukocyte activities and interactions, making it an ideal main thread to organize an introductory cellular biology course.

We divided the entire animation into 7 self-contained chapters, and made one of them, the 45 seconds' chapter *"Membrane and Associated Proteins"*, into an interactive CThru application. We identified 9 visual molecular objects in the video, for which we constructed and embedded custom-made text explanations and multimedia modules. Currently the boundaries of those objects are marked manually. Our future plan is to experiment with automatic object identification and tracking approaches such as those presented in [3].

Various forms of hypermedia have been extensively explored and commercialized. In particular, hypervideo embeds clickable anchors into a video stream, allowing navigation from the video to external sources. Sawhney et al presented an early concept of hypervideo, HyperCafe [7], which provides temporal and spatial link opportunities in a video to allow users follow different conversation threads. Girgensohn et al [3] designed an authoring interface to embed structured videos in a central video segment as hyperlinks. Hypervideo has been used in commercial websites recently. VideoClix [8], an online interactive video library, links visual products in the video to sponsor websites. The key difference of CThru from existing hypervideo is that, while hypervideo systems aim at introducing external information and allowing users to link to them, users can easily lose the original video context in jumping to a new link; CThru, as an educational environment, maintains a user's learning video context when she freely explores and examines supplementary knowledge and information. To achieve this, external information is compiled into an organized space entirely embedded within the video context. CThru users are able to directly interact with this structured knowledge base, and freely switch between video playback and information space.

Recent research has reported approaches in interacting with videos. Dragicevic et al's DimP [1] allows a visual object in the video to be dragged along its trajectory. Goldman et al [3] use a different visual tracking method. These previous work leverages the spatial information of a video to reorganize the temporal sequence of its playback. A few research projects have also built domain-specific educational and scientific tools in a shared display environment. LivOlay [5] allows multiple user applications to be visually overlapped and compared on a shared data wall, which can be used in areas requiring visual analysis of datasets, such as astrophysics. In molecular and cellular biology domain in particular, Forlines et al [2] adapted a commonly used protein molecule viewer application for use in a multi-user, multi-display setting.

DESIGN GUIDELINES

Our objective was to create a self-guided educational space to enhance the learning experience and to accommodate different learning style for students. The development of CThru underwent several iterations incorporating the following guidelines:

Serving as a video player: The system should contain all the necessary functionalities provided in a regular video player, such as pausing/resuming current playback, and skipping forward/backwards along the time line.

An organized and interactive information space: As a video-centered system for educational purposes, external information sources embedded into the video are designed to convey supplementary knowledge and take-away messages. Their existence and logical relationship should be clear to the users, thus it is necessary to organize all external material and generate an embedded information space for them. Furthermore, this information space should be linked with the main video sequence, accessible and interactive, so that users watching the video can freely pause the video, dive into the information space, browse for educational material, and resume the original video.

Having distinctive multi-surface interfaces and adopting table-centric input: The data wall has large real estate and is pixel-rich, while the tabletop provides users with under-the-finger interactivity. A distinctive strategy shall be followed in designing the interface for each surface: The wall provides both fine-grained visualization display region and current learning context information; the tabletop has synchronized visualization region to maintain visual correspondence with the wall, as well as a control panel to give input to the system. Tabletop input controls both surfaces. As CThru runs in a public space, the interactive tabletop also supports multiple users to manipulate the system in either a simultaneous or turns-taking way.

A SELF-GUIDED VIDEO SPACE FOR EDUCATION

CThru starts by a user touching the tabletop to select a chapter of *"Inner Life"*, and watching it. Both the tabletop and the display wall surfaces render the video.

During the playback, tapping on the tabletop switches between pausing and resuming the video. Scrubbing table surface from left to right fast-forwards the video sequence, and from right to left rewinds the video. In the current CThru prototype, multi-finger and multi-hand gestures are supported on a multi-touch tabletop. Multiple users around the table take turns to give input using social protocols, As a collaborative tool, the interactive tabletop affords group awareness and mediates open style fact-to-face discussion

We have created an organized *information space*, to manage supplementary information of an educational video. In this information space, external material associated with the same visual object are grouped together as an education *topic*, and all topics covered in the video are indexed according to their temporal position.

When a topic appears in the video (currently the *"Membrane and Associated Proteins"* chapter), a flashing transparent layer appears on the table over its associated visual object. Touching this region will pause the video and trigger a pop-up interface of information space on both surfaces as shown in Figure 2. The paused frame is displayed in the background to maintain context (See 2F). Touching the background video closes the information space and resumes the video.

We manually identified representative shots in the video, and formed a time-ordered *key frame chain*, displayed on the left side of the information space (2A). Objects with supplemental information in the key frame chain are marked and labeled. The current frame in the chain is highlighted (2B); up to 3 topics in that frame can be displayed as separate columns to the right of the display (2C). Each topic has a name, a short text description, and portals to external multimedia material (2D).

A user can browse all the learning topics in the video by directly scrolling the key frame chain upwards or downwards. As the highlighted frame changes, so will the topics shown on the right; a fading effect is rendered during transition. A star icon (2E) is used as a visual cue in linking visual objects under the finger with its external information topic displayed on the right.

The user accesses external material by touching its portal thumbnail in the topic region. Here we followed the design guideline to display distinctive multi-surface interfaces (Figure 3): Central area of both the wall and the table is reserved to display the multimedia information. The current topic is displayed as a tab on the left side of the wall (3.Left). Portal thumbnails to other material of the same topic, as well as a return button back to the information

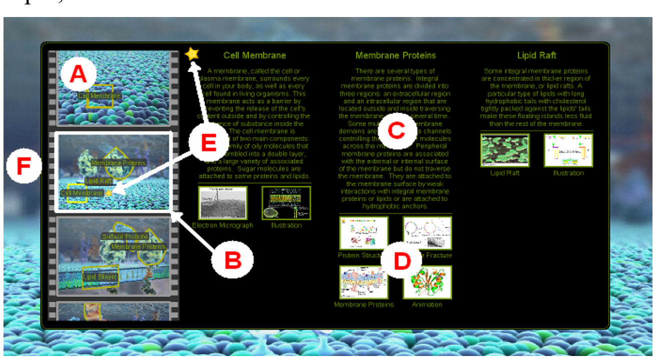

Figure 2. The interface of information space. (A) Key frame chain with marked and labeled objects. (B) Current frame is highlighted. (C) Text explanation. (D) Portals to educational modules. (E) Star icons linking visual object with its associated topic column. (F) Paused video frame as background.

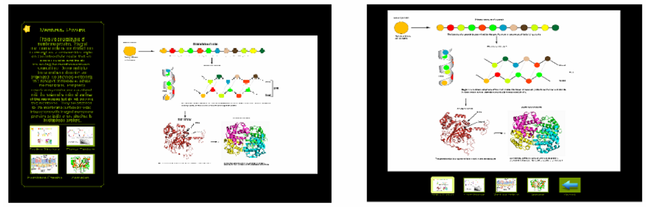

Figure 3. Interfaces on shared displays when presenting an educational module. (Left) Data wall. (Right) Tabletop.

space, are displayed at the bottom of the tabletop (3.Right). Transitions to and from the information space are animated to maintain visual fluidity.

External educational modules may have different modalities. CThru allows extensible education modules to be installed. These modules are linked to the video using an author-defined script file. So far, we have built the following types of interactive education modules:

Illustration slideshow: Illustrations of the same topic can be loaded and displayed in a slideshow. Touching the image flips one slide.

Movie: Besides the main video thread, external movies can also be played in the module. The same set of input methods described to control video playback before is also used here to control the movie.

3D protein viewer: We built upon the JyMOL (http://delsci.com/jymol) plug-in to create a 3D protein viewer module, which loads a protein by its access number in the protein data bank (http://www.rcsb.org), and renders its 3D structure on both surfaces. Gestural input (two-handed zooming and one-hand panning) on the tabletop changes the rendered structure's size and position.

Reference link: Topics in the video are sometimes cross-referenced. A reference link provides a shortcut to another topic related to the current one.

CThru is written in Java using Java 2D and Java Media Framework toolkits. It currently runs on a large display wall of six 50-inch cubes with a total resolution of 3072x1536. CThru has been ported onto two types of multi-touch tabletops: a top-projected DiamondTouch tabletop with a 1280x1024 resolution, and a rear-projected Microsoft Surface of 1024X768 pixels in resolution.

INFORMAL PILOT STUDY AND INITIAL FEEDBACK

We performed a two-stage pilot study and asked the participants their level of comfort with the presented content after each stage. In the first stage, each participant was instructed to view a 45 second video chapter. They were allowed to scrub and replay the video content. In the second stage, the users were presented the same video in the multi-display environment of CThru. The users were encouraged to independently explore the capabilities of the system. The investigators were available to answer questions about the system. This two-stage design was due to the early phase of our system development as we had limited supplemental material available.

Five users (two females) participated in the study: two undergraduates, two graduate students, and one computer science researcher. All participants had taken general biology classes in high-school; two in particular had worked with biologists before. None of them claimed previous knowledge about the cellular biology concepts described in the video. Each spent approximately 40 minutes totally in the two stages.

Feedback of Stage One: Non-interactive video. The participants generally felt that understanding content through the non-interactive video was challenging. The participants reported that while the content was visually engaging, the information presented in the video was too "dense". At the end of the session, Participant P2 said: *"... I wanted to open up Wikipedia while watching the video to understand the content better. ..."* Another participant P5 responded that *"There's so much going on that, correlating elements in the video with the narration was difficult"*.

Feedback of Stage Two: CThru. We received positive feedback on the ability to access supplementary information. The ability to pause and view tagged content was felt to ease the learning curve for understanding the video's content. The participants only spent a small portion of their time on the actual video, choosing instead to flip through the key frame chain and browse the other supplementary topics. While this may have been a result of the participants having viewed the clip in stage one, it raises interesting design questions for education tools such as whether scrubbing forward is desirable.

Observations showed that users were able to discover most interface features. Using the table surface to navigate through the video was not obvious to two users. Once this feature was pointed out, both felt comfortable scrubbing through the video directly. All users reported finding the interface of information space clear, and were able to correctly express the relationship between the highlighted boundaries in the key frame chain and topic columns displayed aside. Four users mentioned the importance of the current-topic star in helping them link the visual object under their finger to the exact column that has information on that topic. The participants expressed comfort using multi-touch gestures to manipulate 3D protein models, citing familiarity with other multi-touch interfaces.

One interesting issue arose during the interaction - the unequal distribution of user attention between the wall and tabletop. Our observation showed that the users' visual foci stayed on the tabletop most of time, even though the data wall provides three times as much resolution as the tabletop and is much larger. Most users only shifted their foci to the wall when reading text. Participant P5 said she felt more efficient when focusing on the surface she gave input to unless encountered text legibility problems. In fact she was willing to spend more time on the wall reading text and viewing high-resolution illustrations. As CThru is intended to be used in a public space as opposed to a lab setting, we believe the wall to play a more important role in less-stressful learning tasks due to the ease of visualization.

After the initial exploration period we asked our participants to identify certain elements in the video that were referenced in the narration. We also asked them to tell us where we could find specific details that were mentioned in the supplemental information. In both cases, the participants were unanimously able to connect the narration to the visual element and its supplemental information.

CONCLUSION AND FUTURE WORK

CThru is an ongoing effort in constructing a video-centered educational space. We experimented with organized supplementary information within sequential videos and introduced different display form-factors. Even with limited material used in the initial study, the feedback was encouraging. Future work includes exploring the efficacy of collaborative learning enabled by the shared surfaces. We will also consider conducting a formal user study when more interactive contents are generated. Another interesting issue to be examined is the interplay of interface and information presentation between the data wall and the tabletop.

ACKNOWLEDGEMENT

We thank Matthew Bohan and Tom Torello for kindly generating the supplemental multimedia material. The first author Hao Jiang is partially sponsored by the 863 Project 2008AA01Z132 and 2009AA01Z336.

REFERENCES

1. Dragicevic, P., Ramos, G., Bibliowitcz, J., Nowrouzezahrai, D., Balakrishnan, R., Singh, K. Video browsing by direct manipulation. *Proc. CHI '08*.
2. Forlines, C., Lilien, R. Adapting a single-user, single-display molecular visualization application for use in a multi-user, multi-display environment. *Proc. AVI '08*, 367-371.
3. Girgensohn, A., Wilcox, L., Shipman, F., Bly, S. Designing Affordances for the Navigation of Detail-on-Demand Hypervideo. *Proc. AVI '04*, 290-297.
4. Goldman, D.B., Gonterman, C., Curless, B., Salesin, D., Seitz, S.M. Video Annotation, Navigation, and Composition. *Proc. UIST '08*, 3-12.
5. Jiang, H., Wigdor, D., Forlines, C., Borkin, M., Kauffmann, J., Shen, C. LivOlay: interactive ad-hoc registration and overlapping of applications for collaborative visual exploration. *Proc. CHI '08*.
6. Pasquali, M. Video in science. Protocol videos: the implications for research and society. 2007. *EMBO reports*. v8, n8, 712–716.
7. Sawhney, N., Balcom, D., Smith, I. HyperCafe: narrative and aesthetic properties of hypervideo. *Proc. Hypertext '96*, 1-10.
8. VideoClix. http://www.videoclix.tv/watchdemo.