

The Landscape of Time-based Visual Presentation Primitives for Richer Video Experience

Yasuhiro Yamamoto¹, Kumiyo Nakakoji¹, Akio Takashima²

¹KID Laboratory, RCAST, University of Tokyo
4-6-1 Komaba, Meguro, Tokyo, 153-8904, Japan
{yxy, kumiyo}@kid.rcast.u-tokyo.ac.jp
<http://www.kid.rcast.u-tokyo.ac.jp/>

²Meme Media Laboratory, Hokkaido University
N-13, W-8, Sapporo, Hokkaido, 060-8628, Japan
akiota@meme.hokudai.ac.jp
<http://km.meme.hokudai.ac.jp/akiota/>

Abstract. As technology advances, we have increasingly more opportunities to use video for our knowledge work, such as monitoring events, reflecting on physical performances, learning subject matter, or analyzing scientific experimental phenomena. Existing video-related software tools are either for as-is viewing or editing and do not support such knowledge-intensive processes. We argue for a variety of interactive presentation tools for richer video experiences in active watching. Based on the Time-based Visual Presentation (TbVP) framework, which separates presentation from content and views interaction methods as transformations between temporal and visual media data properties and user experience properties, this paper presents twenty-seven TbVP primitives and provides the landscape of rich interaction methods for videos. The primitives are illustrated with five scenarios that use videos for knowledge work.

1 Introduction

As technology advances, we have increasingly more opportunities to use video for our knowledge work, such as monitoring events, reflecting on physical performances, learning subject matter, or analyzing scientific experimental phenomena. The focus on most existing video viewing tools for interacting with video data, such as Quick-Time Player, RealPlayer, or Windows Media Player, has not been on supporting such processes. Such tools presume that viewers passively watch videos from the beginning to the end, appreciating it as a complete work of art or entertainment on an as-is basis. Video editing tools, such as Final Cut Pro or Premier, allow users to produce a variety of ways to interact with video data and to explore the space of visual effects. However, their goal is to save the results of such interactions and produce another set

of videos. They do not help users to simply interact with video data without changing the original content.

We argue that we need a variety of interactive presentation tools for richer experiences in using videos for our knowledge work. Our tools are intended to support *active watching* in contrast with *passive watching*. By applying Adler's notion of active reading to the video viewing experience [1], we argue that richer interaction schemes are necessary for active watching (Fig. 1).



Fig. 1. Opportunities have been widened for people to be engaged not only in passively watching videos but also in more actively watching videos

Viewers need to be able to interact with videos in various presentations depending on the purpose of their tasks: whether they are (1) checking for anomalies in events recorded over a long period of time (e.g., a surveillance video); (2) trying to identify a trend in the entire event recorded on video (e.g., a video record of a soccer game); or (3) learning the process recorded in a video (e.g., a video of a cooking procedure).

In the first case, a user may want to quickly skim the entire movie to look for possible distinctively different frames. The user may then want to focus more on those frames by very slowly looking at them. In the second case, such as a video of a soccer game, the user may want to skim the entire video, develop an overview of how the game went, and then focus on particular frames that show a certain player doing something spectacular. Those focused frames need to be situated within the context of the game as a whole. With the third case, a user may want to repeat a certain part of the video that shows crucial steps in a process. The user may also want to view multiple parts of the frames simultaneously to compare “before” and “after” procedures.

The goal of our research is to enable a user to interact with a video in a variety of ways: by interactively changing temporal and visual properties of a video, to identify trends, to discover unexpected phenomena, or to deeply examine a particular segment of a video. Our approach is to separate the presentation of a video from its content. A video presentation has both temporal and visual properties. Temporal properties include the playing speed and which part of the video to play. Visual properties include the size, opacity, or rotation of a video frame. Interaction methods on how to control speed and the appearance of frames while the video is being played have not been studied much. Combinations of multiple controls over temporal and visual properties would open up a vast space of possible presentation styles.

Our Time-based Visual Presentation (TbVP) model uses the notion of *media data values* and *user experience values*. Interacting with a video and changing its temporal properties can be expressed as a transformation from the video's Media Data Time (MDT) to the User Experience Time (UET). Changing the visual properties is expressed as a transformation from the video's Media Data Visualization (MDV) to the

User Experience Visualization (UEV). We have designed interaction methods for various transformations from MDT to UET and from MDV to UEV to develop TbVP primitives. This paper shows how combinations of the TbVP primitives support rich-video viewing experiences.

2 Rich Video Experiences for Active Watching

2.1 Tools for Active Watching

Our research has explored how people interact with materials and external representations in creative knowledge work, and we have developed interaction design methods for such processes [14][18][21][22]. Understanding depends not only on the experiences or the knowledge that each viewing person already has, but also on the dynamically emerging relation between the person and the representations. How people interact with text and images in everyday life involves not mere naive information-receiving processes but complex knowledge-construction processes. Videos as knowledge materials are no exception.

Adler and Doren used the phrase *active reading* to refer to reading with critical thinking and learning [1]. Active reading involves not only interactively changing the appearance of the reading material through highlights and annotations, but also changing reading styles, such as riffling through pages to overview its content, layout, and atmosphere; or poring over the specific parts of a book in which the reader has an interest. Schilit et al. developed a touch panel device named XLibris that supports the former process [16]. Speed-dependent Automatic Zooming for a document browser [3][9] has proved to be an effective approach supporting the latter process. The view automatically zooms out when the user scrolls rapidly so that the perceptual scrolling speed in screen space remains constant.

The experience of reading a book differs among readers, depending on the representation or what kind of interaction is possible [8]. We argue that the same holds for watching videos. In the same way as with reading, watching video data, such as a movie or animated visualization [12], can be *active watching* as well as *passive watching*.

Some studies have supported active watching, even though they might not have used this particular phrase. Video summarization and annotation tools, many of which are developed as research tools [4][10][19][22], allow users to bookmark frames in video; to associate text, images, video, URLs or other materials (such as presentation slides) with some frames or parts of videos as annotations, to use 2D-spatial positioning to represent relationships among segmented frames, and to insert XML tags. Those approaches serve for video viewing as knowledge construction by adding more information and context for video data to enrich the video viewing experience.

In contrast, few studies have addressed changing the temporal and visual aspects of video data to help knowledge construction. Visual and temporal aspects are the two fundamental properties of video data. Manipulation of time has had much fewer studies compared to the field of visualization. Our research goal addresses this challenge in designing interaction methods for both temporal and visual properties.

2.2 Separation of Presentation from Content

Presentation and content have been separated in traditional media. By using a document browser, for instance, one can change the appearance of the text by enlarging the font size, increasing the line spacing, highlighting certain phrases, or converting text from a single column into multiple columns, all of which are possible without changing the content of the text. Or, by using an image viewer, one can view images in various sizes, as a thumbnail list, or in a slide show, depending on what kind of experiences the viewer desires through viewing the images.

Viewing thus does not necessarily assume that data are presented in an as-is manner. Compared to these kinds of text and image changes, however, many existing video viewing tools support almost only as-is viewing. The few variations include half-size viewing, full-screen viewing, or repeating a certain part. They do not support viewing in varieties of different presentations.

This paper presents our approach to provide richer video viewing experiences. Although such experiences could be enabled by using high-functionality video editing tools, we argue that when the goal is to view videos and not to produce different video contents, such editors would be too cumbersome in the same way as it is too cumbersome to use Photoshop for viewing a set of images and not for editing. A variety of visual effect systems have been developed [5][6]; however, many of these are for production of new forms of media, and not for understanding the video content.

One example of an approach that shares our goal involves Rapid Serial Visual Presentation (RSVP) techniques [17]. RSVP is a presentation method that serializes a number of images and stacks them into a pile that a user can interactively flip through. Wittenburg et al. applied the method for consumer video devices [20], in which a sequence of image frames are generated from a video into a 3D trail, and as the user plays forward or backward, the trail advances or recedes while the image in the foreground focus position is replaced.

This paper argues for the need for such types of viewing options for users. We do not provide a single viewing mechanism suitable for a specific task. Rather, this paper provides a framework and the landscape of possible interaction methods for richer video experiences.

2.3 TbVP Model

Our approach uses the properties of Media Data (MD) and User Experience (UE) to distinguish between what the data originally have and what a user actually experiences. Interacting with media is viewed as transforming MD values into UE values. For example, zooming into a frame is regarded as changing a given value of a visual property (e.g., size, resolution) into different values of the properties (e.g., larger size, higher resolution). Changing the visual appearances of a representation is explained as a transformation from MDV into UEV. In the same manner, fast-forwarding the video is regarded as changing a given value of a temporal property (e.g., playing speed) into different values of the properties (e.g., faster play), which is explained as a transformation from MDT into UET.

Numerous systems have been developed to summarize and visualize video data by automatically transforming the raw video data into alternative temporal, or visual representations. For example, Video Manga identifies key frames of a video, and then adjusts their sizes to pack them on the page in a style reminiscent of a comic book [19]. Stained-Glass visualizations identify important areas in the key frames and emphasize them in irregular shapes [2]. Nam and Tewfik introduced the Dynamic Video Summarization technique for summarizing video data, a system that modifies the local sampling rate to make it directly proportional to the amount of visual activity in localized sub-shot units of the video [15]. Girsensohn et al. described detail-on-demand hypervideo, which allows users to watch short video segments and to follow hyperlinks to see additional details [7].

The goal of these approaches, however, is not to enrich video experiences but rather to enable a faster, more effective grasp of the video content than a user can get by watching the original video.

As described in the previous section, we view knowledge construction as a dynamic process by which people interact with external representations. Because no system will be able to generate appropriate visualizations and interaction methods for every conceivable need and situation, our goal is to create a system by which users can actively watch video data.

3 TbVP Landscape and its Primitives

This paper presents twenty-seven TbVP primitives that we have developed by using Macromedia Director. Each TbVP primitive is a video viewer for a particular type of interactive presentation of a video with a basic set of video manipulations. In the course of the development of the TbVP primitives, we have developed the landscape of TbVPs (Fig. 2), which guides the design and development of further video viewing interaction methods based on the TbVP framework.

The primitives are described under the following six different areas of interaction methods for richer video experience:

- (1) direct temporal transformation
- (2) direct visual transformation
- (3) integrated views
- (4) instrumental temporal transformation
- (5) instrumental visual transformation
- (6) integrated presentations

Note that these primitives presume the video viewer archetype, TbVP000, consisting of (a) the video-file-selection pane; (b) the play, pause, and go-back-to-start buttons; (c) the frame-viewing pane; and (d) the horizontal tracker-bar. The length of the tracker-bar corresponds to the entire video, and the bar has an indicator that allows a user to select a particular frame of the video. All TbVP primitives are based on TbVP000, and have these same four fundamental interface elements. See Fig. 2 for the landscape of all primitives discussed in this section.

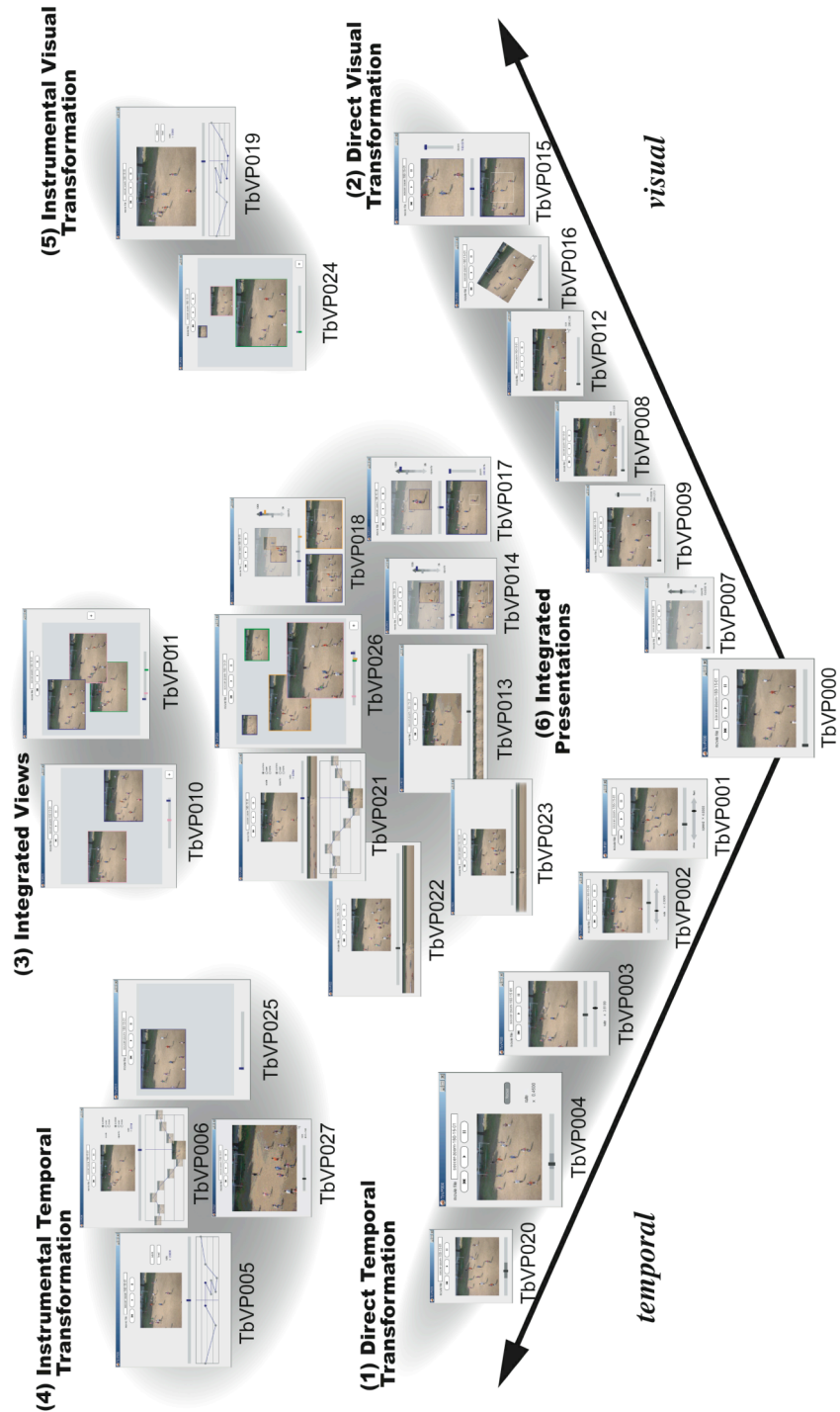


Fig. 2. The TbVP Landscape

(1) Direct temporal transformation.

In this area of the video interaction space, a user directly interacts with the temporal properties of a video. The horizontal bar in the bottom of TbVP001 allows a user to change the speed of playing the video in the range of 0.1 (i.e., 10 times slower) to 10 (i.e., 10 times faster). The bar of TbVP002 allows a user to reverse the speed in addition to TbVP001, in the range of -5 (5 times faster playing backward) to 5 (5 times faster playing forward).

The oval indicator of the second tracker bar in TbVP003 allows a user to specify the frame around which the viewer slows down when playing the video. When a user pushes the “focus now” button of TbVP004, the viewer quickly “rewinds” the video for 5 percent of the length of the entire movie, and then replays the rewound part at the slower speed, allowing a user to carefully examine frames that have just passed. TbVP020 allows a user to specify a certain period of time on the tracker bar for which the viewer slows down during the play.

(2) Direct visual transformation.

In this area of video interaction space, a user directly interacts with visual properties of a video. The vertical slider bar of the view allows a user to change the opacity (TbVP007) or the frame size (TbVP009) of the video in the viewer. With TbVP008, a user may directly grasp the corner of the movie viewer and resize it proportionally. TbVP012 is the same as TbVP008 except that the viewer does not keep the aspect ratio (proportions) of the original size. With TbVP016, a user can rotate the video image while playing it by dragging a corner of the frame window.

TbVP015 allows a user to specify and enlarge an area within the frame displayed in the lower frame pane and show it in the upper frame pane. The zooming factor can be changed by using its vertical slider.

(3) Integrated views.

Tools in the other four areas provide multiple frame views of a single video. In contrast, tools in this area can show multiple frame views of different videos. TbVP010 allows a user to freely position a frame in a 2D space. TbVP011 allows a user to freely position multiple frames in a 2D space. A set of trackers for each frame can be controlled in terms of the single tracker bar. Such free positioning of objects in a 2D space serves as a spatial hypertext representation allowing a user to gradually formalize emerging relationships among objects [11][13][22].

(4) Instrumental temporal transformation.

In (1) and (2), users manipulate the temporal and visual properties of a video by directly changing the objects. With tools in this area, a user would use an instrument to change temporal property values. The single-stroked line graph view of TbVP005 is a graphical interface to represent the transition of the playing-speed and the playing-direction (i.e., forward or backward) of the video while keeping the temporal continuity of video. The horizontal axis represents the original frames of the video, and the vertical axis represents the play forward/backward speed. Users drag line segments

through the graph to indicate which parts of the video should be played at what speed and in which direction. TbVP006 is an extension of TbVP005. TbVP006 shows the thumbnail images on the graph that are derived from the corresponding frames in the video.

With TbVP025, the vertical position of the viewer in the space determines the speed of play. As a user drags the viewer toward the right edge of the space, the playing speed gets faster; placing the viewer around the middle stops the play, and moving the viewer toward the left edge reverses the direction, playing faster and faster. The horizontal bar of TbVP027 allows a user to change the size of the frame, which also changes the speed of play—the larger, the slower; and the smaller, the faster.

(5) Instrumental visual transformation.

With TbVP019, a user does not directly change the size of the frame window, but instead, the size dynamically changes as the playing speed changes, which can be specified through the line-graph interface described above. With TbVP024, the horizontal position of the viewer in the space determines the frame size. As a user drags the viewer toward the top edge, the size gets smaller; toward the bottom, the viewer becomes larger.

(6) Integrated presentations.

The TbVP primitives in (1),(2),(4) and (5) presume that there is only one frame view for each moment. In contrast, the primitives in this area show two or more frame views simultaneously in their viewers.

TbVP013 shows a list of ten thumbnails extracted from the original video and placed along with the tracker bar as a sparse overview. A user can thus browse the rough overview of the entire video and decide which frame to look at.

TbVP022 and TbVP023 both show a list of two hundred frames (translucent frames in the case of TbVP023) extracted from the original video and overlaid and lined up along the tracker bar as a dense overview. Although a user is not able to distinguish each frame in the dense overview, slight color changes in the dense overview might indicate possible anomalies in the video.

Different from the above three primitive tools, TbVP014 provides two frame views. A user specifies the area from the bottom frame, which is overlaid on top of the frame view in the top. The two views can have different temporal manipulators; therefore, one can overlay a partial image of a forthcoming part of the video on top of the frame currently played. With TbVP018, a user can have two overlaying images. TbVP017 is similar to TbVP014 except that a user can resize the overlaid portion in the frame view in the top.

TbVP021 combines TbVP006 and TbVP025, where the speed is specified through the graph interface, and the change of the speed results in the change of the size of the frame. TbVP026 combines TbVP024 and TbVP025; moving a viewer toward the top right corner of the space makes the viewer get smaller and play faster.

4 Scenarios

This section presents five scenarios to illustrate how the TbVP primitives are combined and used for specific tasks. The five scenarios are: (1) a security person analyzes a video recorded by a surveillance camera, (2) a coach of a local soccer team analyzes the video of the team's last game, (3) an overseas tourist learns by video how to cook a Japanese pancake using a table-top BBQ, (4) a student learns how to juggle by watching the video of his expert buddy, and (5) a scientist designing a bio module system analyzes the behavior of paramecia in a charged petri dish. Those scenarios are chosen to emphasize different aspects toward richer video experiences.

(1) A Surveillance Video. When a security person analyzes a video recorded by a surveillance camera, the primary task is to quickly zap the entire video, and search for possible anomalies in long-term events recorded on the video. We use TbVP003, TbVP015, and TbVP022 to design a tool for such a scenario (Fig. 3).

The dense overview of TbVP022 gives the security person a rough idea of where to focus in the video. The security person can set the oval pointer of TbVP003 to view the part of the video around the frames of interest in slow speed. TbVP015 allows zooming-in a particular part of a frame in the video that seems suspicious.

Once the security person becomes aware of a suspicious part of the video, the rest of the video may not be of much interest and thereby viewers such as TbVP004 may not be necessary.



Fig. 3. A TbVP View for the Surveillance Video Scenario

(2) A Video of a Soccer Game. When the coach of a local soccer team analyzes the video of the team's last game, the primary task is to understand how the team's formation changed over time in terms of each player's position on the soccer field. We use TbVP001, TbVP003, TbVP004 and TbVP017 to provide such a macroscopic view for the coach (Fig. 4).

By using TbVP001, the coach may want to quickly browse the entire game on the video with a very fast playing speed. While browsing the game, whenever the coach

notices an interesting phenomenon in a game, he/she can press the “focus now” button of TbVP004 to rewind and replay at a slower speed to further examine what was going on. If the coach notes a particular part of the video (such as a shooting scene), he/she can set the indicator with TbVP003 so that the viewer plays the video in fast speed, except it plays the shooting scene at the much slower speed, which allows the coach to concentrate on the scene. TbVP017 allows the coach to zoom-in on a particular player’s play within a frame.



Fig. 4. A TbVP View for the Soccer Game Analysis Scenario

(3) A Video on a Cooking Procedure. When an overseas tourist learns by video how to cook a Japanese pancake using a table-top BBQ, the tourist is interested in knowing every detail of the procedure of how to mix ingredients, turn over a pancake, and top it with garnishes. At the same time, because the tourist is not familiar with the cuisine, he/she also wants to know the relationship between each step in the procedure and how the resulting cooked dish looks. We use TbVP011, TbVP013 and TbVP014 to design a viewer for such a situation (Fig. 5).

With TbVP011, the tourist can compare multiple phases of the cooking procedures, for instance, scenes after 1 minute, 3 minutes, and 10 minutes from the beginning of the cooking time. The sparse overview of TbVP013 gives the tourist a rough idea of what would happen in each phase of the procedure. TbVP014 allows the tourist to compare how the pancake would look when finished by overlaying a part of the pancake from different time on top of the currently displayed frame view.

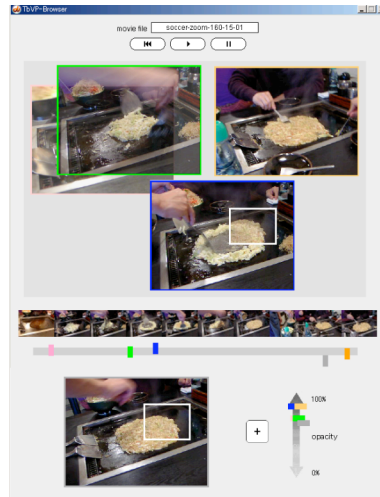


Fig. 5. A TbVP View for the Cooking Video Scenario

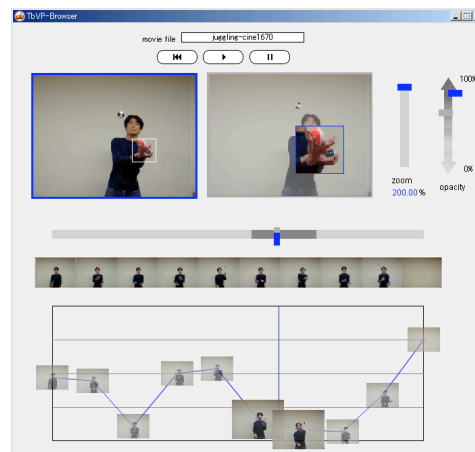


Fig. 6. A TbVP View for the Juggling Performance Video Scenario

(4) A Video on Juggling Performance. When a student wants to learn how to juggle by watching the video of his expert buddy, he/she wants to view particular parts of the video again and again to learn specific skills the expert plays in a video. We use TbVP006, TbVP013, TbVP017 and TbVP020 to design a video viewer for the student (Fig. 6).

With TbVP006, the student can program how to view the video by specifying the playing speed for the entire movie by using the line-graph interface. Thumbnails on the line-graph of TbVP006 together with the sparse overview of TbVP013 would help the student to have a rough idea of which times his expert buddy uses which techniques. TbVP017 allows the student to zoom-in on a particular part of the frame that

shows the expert's hand movement. The student can repeat a certain part of the video many times by using TbVP020.

(5) A Video on Paramecia. Because the size of the body of a paramecium is very small (about 0.1 mm) and the movement of paramecia in a charged petri dish is very fast with constantly changing directions, a scientist analyzing the behavior of paramecia in such an environment to design a biomodule system needs to view the video from multiple perspectives, compare a number of frames, and zoom-in on particular parts of the frame. This type of video viewing task is in some sense the most challenging kind to support.

We use TbVP002, TbVP003, TbVP005, TbVP011, TbVP016, and TbVP017 to design a viewer for the scientist (Fig. 7). TbVP002 allows the scientist to play forward and backward at various speeds. When the scientist gets interested in a particular scene, he/she can indicate the frame with the TbVP003 indicator to play the scene at a slower speed. TbVP005 allows the scientist to save a particular mode of viewing the video so that he/she can share this viewing experience with peers in a research group. Most of the frames look alike, so thumbnail displays on a line-graph interface are unnecessary, eliminating the need for TbVP006. TbVP011 allows the scientist to freely position multiple video frames in a space, and he/she can use the space to organize hypotheses as a spatial hypertext representation; for instance, video frames for a particular hypothesis can be placed toward the top edge, whereas those for another hypothesis can be placed toward the bottom edge. Because paramecia keep moving in many directions, the rotation provided by TbVP016 would be necessary to compare some aspects of the behavior. Zooming into a part of a frame in TbVP017 would help the scientist to deeply examine the behavior of a particular paramecium.

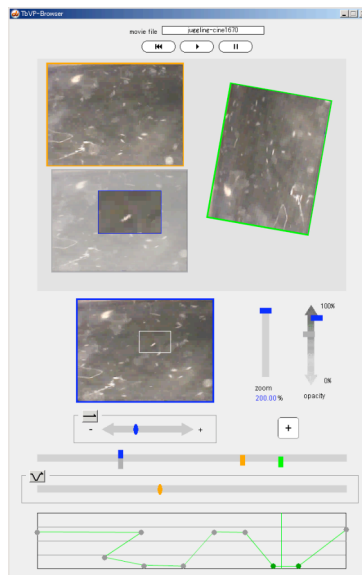


Fig. 7. A TbVP View for the Paramecia Analysis Scenario

5 Concluding Remarks

This paper has focused on interaction methods for users to actively be engaged in video viewing tasks to experience a variety of presentation styles of video data, rather than to edit and produce new video contents. Our approach is to develop the landscape of various types of interaction design primitives with which users could “talk about” their needs for a wide variety of rich video experiences through active watching. We have identified possible patterns of combining TbVP primitives for specific types of video viewing tasks, ranging from a monitoring task to a scientific investigation.

Our immediate future work includes developing a language and an environment to combine some of the primitives to design such a video viewing environment for a specific task. In doing so, we have to cope with design, technical, and psychological challenges, including how to handle complex user operations on multiple video views, how to solve conflicting operations (e.g., when visual properties are dynamically determined by more than one determinant), and how not to overwhelm a user’s cognitive process and yet enable a rich video experience. We will continue this work through user observations and exploration of user interface devices that might be better suited for this kind of task.

Acknowledgements

The authors would like to thank Elisa Giaccardi for her valuable comments for organizing this paper. This research is supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (A), 16200008, 2004-2007.

References

1. Adler, M. J. and Doren, C. V.: How to Read a Book, Simon and Schuster, New York, 1972.
2. Chiu, P., Girgensohn, A., Liu, Q.: Stained-Glass Visualization for Highly Condensed Video Summaries, Proceedings of 2004 IEEE International Conference on Multimedia and Expo (ICME 2004), June (2004)
3. Cockburn, A., Savage, J., Wallace, A.: Tuning and Testing Scrolling Interfaces that Automatically Zoom, Portland, OR., ACM Press, April (2005) 71-80
4. Correia, N. and Chambel, T.: Active video watching using annotation, in Multimedia’99 Proceedings (Part 2), ACM Press (1999) 151-154
5. Fels, S., Lee, E., Mase, K.: Techniques for Interactive Video Cubism. Proceedings of the Multimedia’00. Marina del Rey, CA. (2000) 368-370
6. Fukuchi, K., Mertens, S., Tannenbaum, E.: EffectTV: A Real-time Software Video Effect Processor for Entertainment. Lecture Notes in Computer Science, Vol. 3166. Springer-Verlag, Berlin Heidelberg New York (2004) 602-605
7. Girgensohn, A., Wilcox, L., Shipman, F., Bly, S.: Designing Affordances for the Navigation of Detail-on-Demand Hypervideo, Proceedings of the Working Conference on Advanced Visual Interfaces, AVI 2004, May (2004) 290-297

8. Hornbaek, K. and Fraekjaer, E.: Reading Patterns and Usability in Visualizations of Electronic Documents, *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 10, No. 2 (2003) 119-149
9. Igarashi, T. and Hinckley, K.: Speed-dependent Automatic Zooming for Browsing Large Documents, in *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology*, ACM Press (2000) 139-148
10. Li, F.C., Gupta, A., Sanocki, E., He, L.-W., Rui, Y.: Browsing Digital Video. *Proceedings of the CHI'00*. Hague, Netherlands. ACM Press (2000) 169-176
11. Marshall, C. C. and Shipman, F. M.: Spatial Hypertext: Designing for Change, in *Communications of the ACM*, Volume 38, Issue 8 (1995) 88-97
12. Nakakoji, K., Takashima, A., and Yamamoto, Y.: Cognitive Effects of Animated Visualization in Exploratory Visual Data Analysis, in *IEEE Proceedings of Information Visualisation*, (2001) 77-84
13. Nakakoji, K., Yamamoto, Y., Reeves, B.N., Takada, S.: Two-Dimensional Positioning as a Means for Reflection in Design of Interactive Systems (DIS 2000), *ACM Press*, New York (2000) 145-154
14. Nakakoji, K. Yamamoto, Y., Aoki, A.: Third Annual Special Issue on Interface Design, *Interactions*, ACM Press, Vol.IX.2, March+April (2002) 99-102
15. Nam, J. and Tewfik, A. H.: Dynamic Video Summarization and Visualization, in *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2)* (1999) 53-56
16. Schilit, B. N., Golovchinsky, G., Price, M. N.: Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM Press / Addison-Wesley Publishing Co. (1998) 249-256
17. Spence, R.: Rapid, Serial and Visual: A Presentation Technique with Potential. *Information Visualization*, Vol.1, No.1 (2002) 13-19
18. Takashima, A., Yamamoto, Y., Nakakoji, K.: A Model and a Tool for Active Watching: Knowledge Construction through Interacting with Video, *Proceedings of INTERACTION: Systems, Practice and Theory*, Sydney, Australia, November (2004) 331-358
19. Uchihashi, S., Foote, J., Girgensohn, A., Boreczky, J.: Video Manga: Generating Semantically Meaningful Video Summaries, in *Proceedings of the ACM Multimedia 99* (1999) 383-392
20. Wittenburg, K.B., Forlines, C., Lanning, T., Esenther, A.W., Harada, S., Miyachi, T.: Rapid Serial Visual Presentation Techniques for Consumer Digital Video Devices, *ACM Symposium on User Interface Software and Technology (UIST)*, November (2003) 115-124
21. Yamamoto, Y., Nakakoji, K.: Interaction Design of Tools for Fostering Creativity in the Early Stages of Information Design, *International Journal of Human Computer Studies (IJHCS)*, Special issue on Creativity, E.A. Edmonds, L. Candy (Eds.) (2005) (in print)
22. Yamamoto, Y., Nakakoji, K., Aoki, A.: Spatial Hypertext for Linear-Information Authoring: Interaction Design and System Development Based on the ART Design Principle, *Proceedings of Hypertext2002*, ACM Press, June (2002) 35-44