Tangible Video Editor: Designing for Collaboration, Exploration, and Engagement

Jamie Zigelbaum, Michael Horn, Orit Shaer, Robert J. K. Jacob
Tufts University, Medford, Mass. USA
{jamie.zigelbaum, michael.horn, orit.shaer}@tufts.edu, jacob@cs.tufts.edu

ABSTRACT
This paper introduces the Tangible Video Editor (TVE), a tangible interface for editing sequences of digital video clips. The TVE features multiple handheld computers embedded in specialized cases. It can be used on any surface under normal lighting conditions without the need for an augmented surface, computer vision system, or video projector. We aim the Tangible Video Editor for use in educational and community settings. Hence, we designed it to engage users in a collaborative editing experience and to encourage the exploration of narrative ideas. We evaluated our design using a comparative study that involved 36 participants using the Tangible Video Editor and a standard GUI-based, non-linear video editor.

Author Keywords
Tangible user interface, distributed cognition, digital video editing, interface design, CSCW, physical interaction.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION

The TVE is a tangible interface for editing sequences of digital video clips. Inspired by traditional film editing systems such as Moviola [12] and Steenbeck [19], the Tangible Video Editor combines the benefits of physical editing with the power of current non-linear editing (NLE) techniques.

The TVE works without a computer vision system, video projector, or augmented surface. Instead, the interface incorporates multiple handheld computers embedded in physical *tokens* (objects that represent digital information). The system works on any surface under normal lighting conditions, and it avoids the latency problems associated with many object-tracking technologies. The components can be easily manipulated—users can pick them up, look at them from various angles, and arrange them in their work environment to best suit their needs.

Figure 1. Working with the Tangible Video Editor.

We designed the TVE for use by novice video editors in educational and community settings. Our goal was to develop an easy to use video editing system that implements the basic functionality of a video editor and is aimed at engaging users in the editing process, facilitating collaboration, and encouraging exploration. To achieve this goal we applied an iterative design cycle and evaluated multiple prototypes with users. We evaluated our final prototype against a similar graphical video editor to examine how the choice of editor facilitates *collaboration*—the extent to which partners mutually participate in a task, *exploration*—the extent to which the interface encourages users to interact with the content and try new things, and *engagement*—the extent to which users enjoy using the interface and the extent to which they were involved in their task. We also discuss the role of distributed cognition [4] as it applies to tangible interaction.

Case for Tangible Video Editing

In educational and community settings the process of editing content—the discussions, exchange of ideas, and development of mutual understanding—is often more important than the final product. Hence, video editing software aimed at these settings should support a collaborative editing process. Tangible user interfaces (TUIs) are widely cited as supporting collaboration [1, 5, 20]. Furthermore, when designing *synchronous groupware*, interfaces designed for multiple simultaneous users, designers must often weigh the merits of power vs. operability. The physicality of TUIs solves many of the problems laid out by Gutwin and Greenberg in their discussion of tradeoffs between power and workspace awareness [3]. As TUIs are more specialized than graphical user interfaces (GUIs).they often trade the power available from symbolic manipulation (e.g. menu commands, keyboard shortcuts) for visibility and continuous feedback that are particularly important when multiple users are interacting with a system simultaneously. We claim that a TUI is a good fit for our design goals: it enables multiple users to work together psychologically and physically, it provides them with a customizable workspace, and it embodies a clear syntax for completing video editing tasks that allows them to focus their mental energy on the content of their work.

In this paper we will first discuss related and influential work. We will then give an account of our design process and the components of the TVE. Next, we will present our evaluation, and summarize our findings. The paper will end with a conclusion and ideas for future work.

RELATED WORK

Related work is divided into three categories: tangible interfaces for digital media, empirical experiments, and conceptual frameworks for TUIs.

Tangible Interfaces for Digital Media

The design of the TVE was influenced by several previous projects which involve the manipulation of digital media through physical objects. An early example is Video Mosaic by Mackay and Pagani [10]. Essentially an enhanced storyboarding tool, Video Mosaic addresses the issue of representing time spatially by combining the benefits of real-world interaction (in this case with paper) and the power of digital manipulation. Another example is Myron Krueger's work on VIDEOPLACE [9]. This is an early example of a computer scientist promoting the advantages of interface design based on perceptual processes used in the real, physical world. This theme is echoed in the work of Snibbe et al. [18] who built devices that delivered haptic feedback for media control. *Big Wheel* and their other interfaces demonstrate the value of designing systems that mimic the physical laws of nature.

Recent examples of tangible video editing include the mediaBlocks [23] system, which provides a compelling interface for capturing, editing, and displaying multi-media using wooden blocks to represent content. Some of the most interesting aspects of the system include the media browser and the media sequencer. These components allow users to view and sequence digital video by arranging the blocks in a physical device. A related interface, also from the Tangible Media Group at MIT, is the Tangible Video Browser [21]. With this system, physical tokens represent video clips. When the tokens are placed on the browser interface, they become objects that can be manipulated to navigate through the clip iteself.

Other relevant examples include TellTale [1], EnhancedMovie [7], and DataTiles [15]. TellTale, a toy caterpillar whose body is made up of segments that can be arranged in any order, is an interface for creating sequences of audio clips. Each segment can contain a short audio recording, and by rearranging the segments children can create sequences of sounds to tell different stories. EnhancedMovie features an augmented desk, which allows users to make a movie by editing a sequence of pictures using hand gestures. Hand gestures (such as closing all fingers above an image) allow users grab, release, select, and browse clips displayed as images on the table. Finally, DataTiles, though not built specifically for editing video, is a powerful tool for tangibly manipulating digital media. Users of DataTiles arrange transparent tiles on a video screen that senses their position and displays information within each tile.

Empirical Experiments

Several studies have presented empirical data to attempt to quantify the advantages and disadvantages of tangible systems compared to their graphical counterparts. Here we describe some of the studies that most influenced our evaluation of the TVE. The Senseboard [8] is a tangible interface that allows users to manipulate physical tokens on a grid to interact with abstract data such as scheduling information. An experiment was conducted which compared completion times for a scheduling task with participants using the Senseboard, a digital whiteboard interface, and a paper sticky note interfaces. Subjects using the Senseboard were able to complete the task slightly faster than with the other systems and also preferred it slightly more.

At the University of Washington, Chen-Je Huang performed an experiment that demonstrated a statistically significant difference between the tangible and graphical user interfaces for spatial tasks involving accuracy and reproduction time [6].

Finally, Scott, Shoemaker, and Inkpen conducted a study that investigated children's behavior while collaborating on paper-based and computer-based puzzle-solving activities [15]. The study was notable for its use of video coding to compare the extent to which children collaborated with each interface.

Conceptual Frameworks for TUIs

In his dissertation [24], Ullmer introduces the Token + Constraint approach for tangible interaction. He emphasizes the use of mechanical constraints to perceptually encode digital syntax. While our design attempts to minimize the use of mechanical constraints such as docking stations and racks, it emphasizes the use of physical properties to encode digital syntax.

The TAC paradigm [17] uses Ullmer's Token + Constraint approach as its basis. It then extends the concept of constraints, stressing that tokens (physical objects that represent digital information) and constraints (physical objects that constrain the behavior of tokens) are relative definitions. Tokens can serve as constraints by mutually constraining each others behavior. In our system, tokens (clip holders) mutually constrain each other: two clip holders can only be connected in a certain way and while they are connected their manipulation is constrained.

Finally, our own work on Reality-Based Interaction [14] provides the foundations for a design framework based on users' pre-existing knowledge of interaction with the real world. We designed the TVE so that users would know how to use it based on their existing knowledge of interacting with objects in their 'real world'.

DESIGN AND IMPLEMENTATION

The Tangible Video Editor is the result of an iterative design process in which successive prototypes were created, evaluated with users, and refined.

Design Process

To gain insight on video editing, we visited the ESPN *SportsCenter* studios in Bristol, Connecticut where we observed expert video editors at work. We interviewed them about the art of video editing, the limitations of the technology, and the improvements that they would add to current GUI video editing systems. We also surveyed existing NLEs for professional as well as amateur use.

Based on our observations, we built a series of low fidelity prototypes out of paper and foam core. These early design explorations led us to the implementation of our first functional prototype, which employed computer vision techniques coupled with a ceiling-mounted LCD projector. This computer vision based implementation suffered from high latency; often the projected image would lag behind its target on the table's surface. We evaluated this prototype with six participants and used their suggestions and comments to design the successive implementations.

As the latency problem associated with computer vision crucially affected user experience, we decided to seek implementation methods where object tracking would not be necessary. To experiment with alternative designs, we built two additional low fidelity prototypes. The first was built rapidly using high-density foam and served for brainstorming and task simulation. The second was built from laser cut acrylic and served for evaluation with users. Both prototypes were built with a large board to constrain the movement of tokens and provided docking stations where users could load video clips and play them.

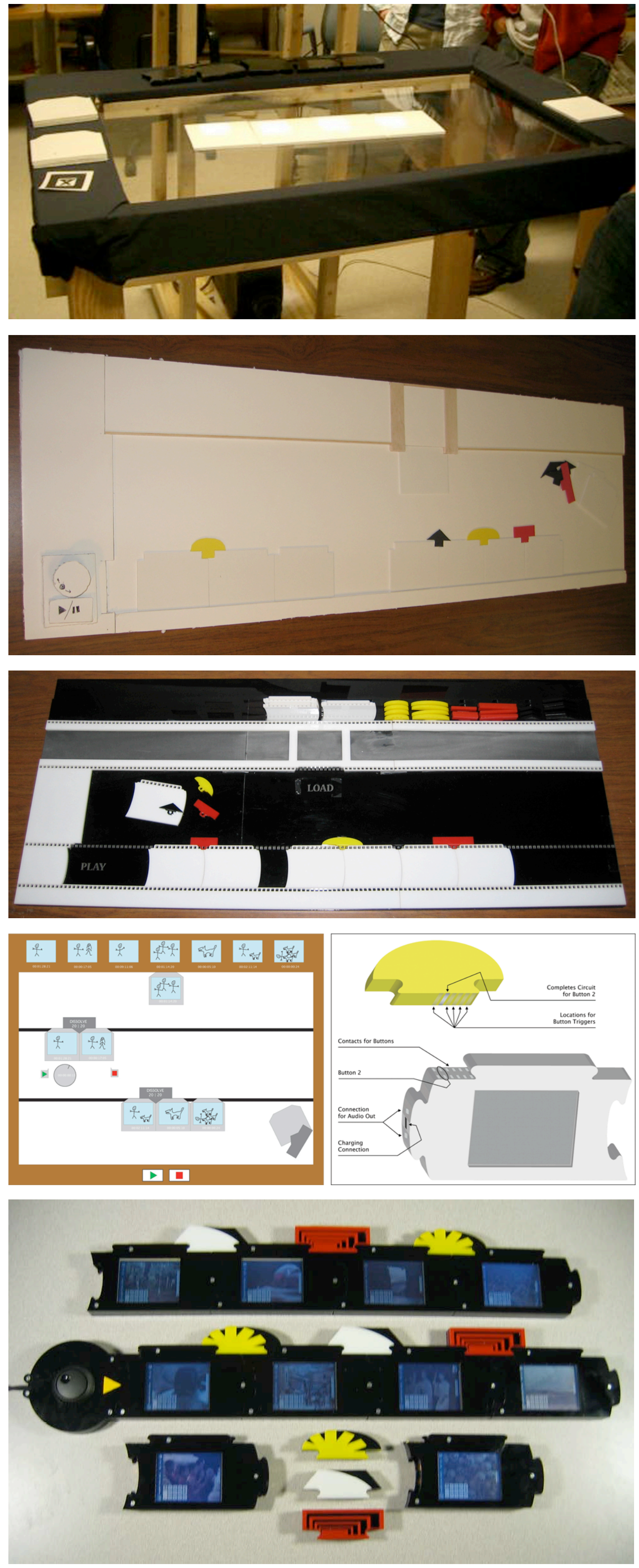

Figure 2. From top to bottom: a) the first functional prototype (computer vision implementation); b) the foam-core prototype; c) acrylic prototype; d) design mockups for the computer vision implementation (left) and final version (right); e) the final version of the TVE.

We evaluated the acrylic prototype with eight participants working in pairs. We used a wizard-of-oz technique (where the experimenters complete the role of the computer) in order to refine our metaphors for editing linear sequences of video clips. The results from this phase of user testing were mixed: users reported having a positive and engaging experience but complained about the way the physical design restricted them. We also observed that often users were unsure of how to interact with the system.

Following this user study we decided users would benefit from a less rigid structure that could enable them to organize the clips in any way they wanted. This decision led us to our final design, which we implemented using embedded handheld computers.

The current version of the Tangible Video Editor (see [22] for a video of this version) features a *play-controller*, 10 *clip-holders*, and 15 *transition connectors* (see Figure 2). These components operate in conjunction with a desktop PC and a monitor for viewing video content. Each clip-holder consists of a Compaq iPaq 3650 Pocket PC running the Microsoft PocketPC 2002 operating system, mounted inside a custom-built plastic case. After being assigned a particular video clip, a clip-holder displays the first frame from the clip on its LCD screen. The cases, which resemble large jigsaw puzzle pieces, connect together to form linear sequences of video clips. Users can also place various transition connectors between adjacent clip-holders to add transition effects such as fade, minimize, and rotate.

With our current implementation, users assign a video clip to a clip-holder by keying in an ID number on the clip-holder's touch screen display. In our experiment, we gave participants 28 paper cards representing the collection of video clips for them to browse through. Each card displayed a full color print of the first frame from the video clip on one side and the clip's ID number on the opposite side. These cards fill the role of the *clip pallet* found in many NLEs. A typical NLE clip pallet is a matrix of thumbnail images representing each available video clip.

Design Rationale

Traditional film editing systems provide editors with a rich sensory environment that allows them to utilize enactive (muscle) memory and haptic (force) feedback. The task of cutting and splicing film into sequences for playback involves the use of physical tools such as cutting arms, and taping stations. The affordances of these tools help convey both their purpose and their means of use. In contrast, current state-of-the-art, NLE software such as Final Cut Pro, Premiere, and AVID provide filmmakers with little of the physical richness employed by their predecessors. However, they provide editors with powerful features such as the ability to reuse shots without making a new print from a negative, to undo actions as quickly as they were done, and to initiate new projects with the click of a button. Our goal was to combine the benefits of traditional, physical film editing with the advantages of digital film editing.

Physical Syntax

TUIs draw upon users' existing knowledge of interacting with the 'real world' to convey to users what interactions an interface can (or cannot) support. They often use form and mechanical constraints [24] to physically express digital syntax.

Figure 3. Clip-holders are shaped like jigsaw puzzle pieces.

The physical design of the TVE clip-holders indicates their means of use through a jigsaw puzzle metaphor (see Figure 3). Thus, most users have a hint that the clip-holders are supposed to connect together, and they can build on prior knowledge to assemble them. Even users who have never seen a jigsaw puzzle before have some cues because the function of the pieces is inherent in their design [2]. This metaphor also helps reduce the possibility of syntax errors. The shapes of the parts simply prevent them from being assembled incorrectly.

The metaphors for physical items utilized by GUI applications (i.e. icons, widgets, etc.) do not always follow the same rules as the physical world, and users must learn their specific language of interaction. For example, we observed in our testing that users of a GUI video editor took time figuring out how to remove a clip from a sequence (e.g. key press, drag and drop, menu item). TVE users, on the other hand, simply picked up a clip-holder and moved it out of the way.

Case Design

The cases for the clip-holders, transitions, and play-controller were constructed from layers of 1/8 inch thick extruded acrylic sheets, cut with an industrial laser cutter. The iPaqs were removed from their original cases and mounted inside the clip-holders. A top layer of acrylic holds the iPaq in place and hides its internal components. Only the touch screen display is visible to the user. Copper connectors run along the outside edges of the cases where the clip-holders interlock with transitions and other clip-holders. The copper connectors are wired to the audio-out, audio-in, and application shortcut buttons of the iPaqs (see Figure 4). When two clip-holders are joined, an electrical connection is formed between the audio out channel of the right clip-holder and the audio in channel of the left clip-holder. Likewise, when a transition is placed between two clip-holders, it closes a circuit to press one of the iPaq's application shortcut buttons.

Clip-to-Clip Communication

A clip-holder must be able to pass information to adjacent clip-holders about its current clip ID number and transition. This information flows along a sequence from the rightmost clip-holder to the leftmost clip-holder, through the play-controller to the PC. Each clip-holder in the sequence receives a data stream from its neighbor to the right, appends information about its own clip and transition, and then passes the new data stream to its left. This data stream is transmitted in a continuous loop.

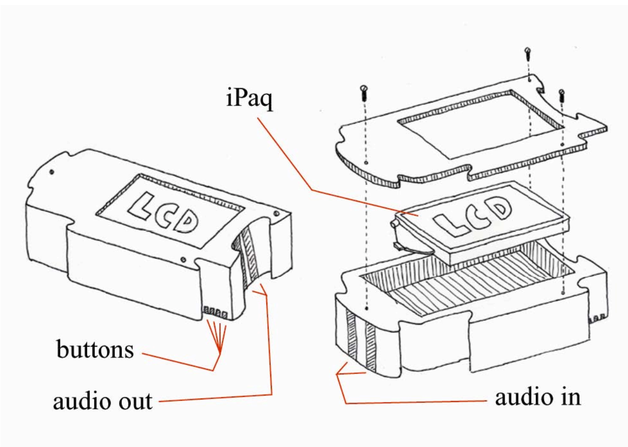

Figure 4. Diagram of clip-holder assembly and connections.

We use the iPaq's audio channels (microphone and speaker jacks) to encode ASCII data in a frequency modulation format. This format, similar to early telephone modems, transmits binary zeros as tones of a base frequency and binary ones as tones of a higher frequency. The data transmission rate with this technique is slow (roughly thirty bytes a second) but adequate for our purposes. We considered alternative means for transmission including sending data over infrared or through a serial port. However, unlike a serial port, which requires several pins, audio only requires soldering two connections. Audio transmissions also avoid potential infrared signal interference problems and can be easily interpreted by various platforms. Any device with both a microphone and a speaker could conceivably be wired to work with the TVE system.

Transitions

The TVE uses three types of transitions (minimize, rotate, and fade) to insert graphical effects between adjacent clips. When inserted, a transition closes a circuit that generates a button press event on the iPaq. Each transition type generates its own distinct button event. The software running on the iPaqs registers these events and inserts the corresponding transition data into the clip-holder data stream.

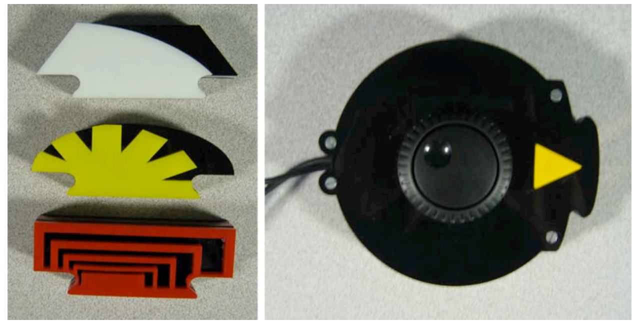

Figure 5. The three types of transitions (top: fade, middle: rotate, bottom: minimize) and the play-controller.

Play-Controller and PC

The play-controller is a circular device that includes a yellow play button and a jog wheel. Users connect the play-controller to a clip-holder or the beginning of a sequence of clips holders that they want to view. Pressing the play button forwards the data stream to the desktop PC and triggers an application to play the clip sequence with transition effects. The application, written using Macromedia Flash, dynamically composes the movie from pre-loaded video clips.

EVALUATION

We conducted a study to evaluate whether the TVE provides benefits, compared to a typical graphical editor, in supporting collaborative editing, exploration of alternative narratives, and users' engagement. Our study was also designed to explore whether the TVE facilitates distributed cognition.

Task

The participants' task was to create a short movie from pre-existing video clips. We provided 28 video clips for subjects to work with. Each clip ranged from 3 to 15 seconds in length. Three types of transitions were available for subjects to add where they saw fit. Each pair of participants was told to work together to create a movie that would tell a meaningful story. We informed them that we would ask them about their story when they were finished. To motivate the participants we told them that their movie would be judged by an impartial committee, and that the best movie would be displayed anonymously on the project website.

Experimental Design

We used a between-subjects design in our study: subjects worked on only one of the two conditions tested, the TUI condition or the GUI condition. We designed our experiment using the co-discovery method where users work collaboratively to perform a task. For us the advantage of this technique over a think-aloud protocol is that the subjects talk about their task naturally. We divided the subjects into pairs, and each pair was randomly assigned to perform the task under one condition.

Pairs editing in the TUI condition used the TVE, and pairs editing in the GUI condition used Microsoft Movie Maker (MMM) [11]. MMM is a typical non-linear editing application for home video that we chose for its ease of use and comparable functionality to the TVE. For the GUI condition subjects were instructed to limit their use of MMM to viewing clips, sequencing clips and adding one of three types of transitions that we identified for them by pointing to their location in the MMM interface. For the TUI condition subjects were given a few minutes to get comfortable using the TVE before their session began to offset the divergence in the experience subjects had with TUIs and GUIs.

For both condition subjects were asked to arrange up to a maximum number of clips from the 28 available clips; for the TUI condition the maximum was five clips, for the GUI condition the maximum was ten clips. The variation in maximum clips for the two conditions was necessitated by technology constraints discovered during testing. Although all ten clip-holders were functioning we found that the TVE was more stable when we used only five of the ten clip-holders in the task.

Subjects

36 subjects participated in our experiment, 12 males and six females participated in the GUI condition. Their ages ranged from 20 to 58 with a median age of 27. 12 males and six females participated in the TUI condition. Their ages ranged from 18 to 74 with a median age of 27. Participants in six pairs (three from each condition) indicated that they knew each other. Subjects were recruited by word of mouth and were not paid.

Procedure

For each condition participants filled out a pre-test questionnaire and a consent form. We then explained the task to the participants, demonstrated the operation of the equipment, and asked them to perform the task. Afterwards we asked each participant to fill out a post-test questionnaire about the task and the editing process. Each session was filmed for further analysis.

Metrics

We have defined several metrics to be used in our measures.

We define a *turn* as a transition in conversation from one partner to another (disregarding one word responses such as "*yeah*" or "*OK*").

We define *turns per minute* as the number of turns taken by a pair divided by the total session length.

We define *mutual discussion ratio* as the amount of time one participant spoke during a session divided by the amount of time his or her partner spoke. Single word responses were disregarded.

We define *total discussion ratio* as the sum of both partners' speaking time divided by the session length.

We define an *alternative timeline* as a variation in a sequence of clips that fundamentally changes the participants' movie

Measures

We have identified measures that are associated with each of the following components: collaboration, exploration, and engagement. For each measure we present a definition as well as a set of metrics. Participants filled out a post-test questionnaire in which they answered 16 questions about their experience on a 5-point Likert scale where 1 indicates the participant strongly disagrees, and 5 indicates the participant strongly agrees. We used these answers as well as data extracted from videotapes of the experimental sessions as evidences for analysis.

It is important to note that many of these measures are subjective. We attempted to minimize the error associated with different analyst's subjective interpretation of measures by using one member of our team to conduct all the video analysis.

Collaboration

We define *collaboration* as the extent to which partners mutually participate in the task. We measured collaboration with the following metrics:

C1. Mutual discussion ratio
C2. Turns per minute
C3. Questionnaire answers Q1-Q4 (see Table 1)

Exploration

We define *exploration* as the extent to which the interface encouraged participants to interact with the video clips and try alternative narratives. We measured exploration with the following metrics:

X1. Number of alternative timelines created by a pair
X2. Questionnaire answers Q5, Q7, and Q11 (see Table 1)

Engagement

We define *engagement* as the extent to which participants enjoyed the task and the extent to which participants were involved in the task of creating the movie. We measured engagement with the following metrics:

N1. Total discussion ratio
N2. Questionnaire answers Q12-Q16 (see Table 1)

Results

We used a two-tailed *t*-test to determine the statistical validity of both the questionnaire results and the video analysis. The results are shown in Table 1.

Collaboration

Our results indicate that partners took more turns per minute with the TUI condition and contributed more evenly to the discussion with the TUI condition (C2, C1). In addition, participants in the GUI condition indicated that

			Metric	TVE	MMM	*p<0.10, **p<0.05
Collaboration	C1		Mutual discussion ratio	1.42	1.77	
	C2		Turns per minute	3.98	3.56	
	C3	*Q1.*	I would have rather worked alone on this project.	3.17	2.44	*
		Q2.	I enjoyed working with my partner.	4.28	4.11	
		Q3.	Working with my partner made our editing process more creative than it would have been working alone.	3.94	3.94	
		Q4.	Either my teammate or I did most of the work in editing our film.	2.22	2.94	*
Exploration	X1		Alternative timelines	3	2.67	
	X2	*Q5.*	The interface that I used (TVE or MMM) was easy to use.	4.11	4.33	
		Q7.	I became very familiar with the content of every video clip available for use in the film.	1.61	2.82	**
		Q11.	I had trouble figuring out how to do what I wanted while using (TVE or MMM).	2.22	1.79	
Engagement	N1		Total discussion ratio	0.34	0.34	
	N2	*Q12.*	I spent more time figuring out how to use the interface (TVE or MMM) that thinking about our film and how I wanted it to be.	1.56	1.22	*
		Q13.	I was really excited to make our short film.	4.22	3.5	**
		Q14.	I was very enthusiastic about the editing process.	4.22	3.44	**
		Q15.	I enjoyed creating our story and figuring out which clips to use to tell it.	4.28	4.11	
		Q16.	On the whole the editing experience was very fun.	4.56	3.94	**

Table 1. Mean results form the video analysis and questions on the post-test questionnaire answered on a 5-Point Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree).

one person tended to do most of the editing work (Q4). However, none of these results were statistically significant. We observed during the evaluation that participants in the GUI condition would most often designate one person to operate the keyboard and mouse for the duration of the task. Participants in the TUI condition most often shared all of the work, including operation of the play button.

Results from the data show that users of the TVE were more likely to have preferred working alone on their task (Q1). This result was not significant, and we hypothesize that relationships between participants who knew each other prior to the study might have biased responses to this question.

Exploration

On average, participants in the TUI condition did explore more alternate timelines (3.0 vs. 2.67), but this result was not statistically significant.

Participants with the GUI condition reported that they became more familiar with the content of every video clip than participants with the TUI condition (2.82 vs. 1.61, p=0.001). In analyzing the session video tapes we noticed that pairs working on the GUI editor tended to remain silent and view all available clips prior to discussing what story to tell, whereas pairs working on the TUI editor began discussion of story within the first minute of the session. Our implementation of the TVE in the experimental condition called for participants to enter an ID number on the clip-holder's touch screen with a stylus to change clips, this took a few seconds to do, while users of MMM only had to double click on a thumbnail of a clip to watch it. Answers to Q7 indicate that participants were more likely to remember the contents of all available clips in the GUI condition. The anecdotal evidence of user behavior stated above may help explain this.

Engagement

Participants with the TUI condition reported that they were more excited about creating films (Q13) and more enthusiastic about the editing process (Q14). They also reported that they had more fun overall (Q16). These results were found to be statistically significant.

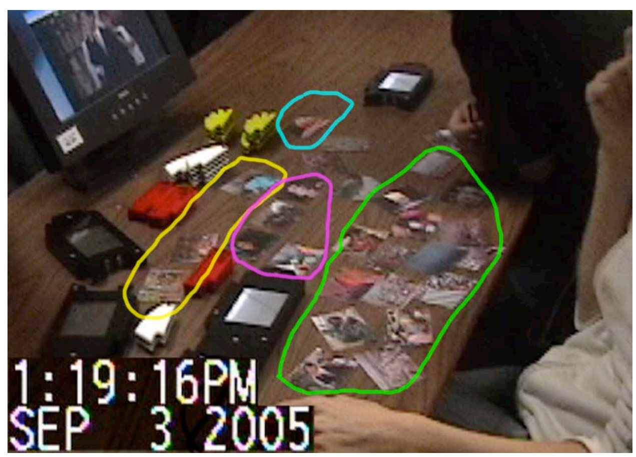

Figure 6. Clips arranged on table at the third minute of a TVE session. Clips circled in green (right) are being evaluated; purple (middle) are a possible timeline; yellow (left) have been tried and removed from the timeline; the blue clip (top) has been discarded.

Distributed Cognition

Pairs working on the TVE began to change their environment as soon as they started the task. This is notable as evidence of distributed cognition, which has been shown to aid memory by encoding information pertaining to the cognitive process in the user's workspace [13]. While using Movie Maker, or any other GUI non-linear editor, users are not given the same ability to manipulate and arrange information as they are with physical objects and TUIs. On GUIs users are restricted to organizing information on a two-dimensional screen within the confines of a display space. We found that pairs working on the TVE arranged the paper screen shots for each video clip in meaningful patterns in their work area (see Figure 6).

We observed that typically users on the TVE grouped clips that they were considering using in their films closer to their bodies and put clips they did not want to use farther away. They also arranged the clips into groups on the table e.g. religious clips in a group, nature clips in a group, etc. This use of their workspace to offload information is a powerful tool absent from the GUI environment.

Summary

The TVE interface worked well during our study, and participants were able to creatively arrange video clips with it into personal creations. Although it is a functional interface suitable for testing, the TVE is by no means a polished, production interface. By limiting the functionality of Movie Maker during testing, we were able to create a balanced comparison between the two conditions.

The statistically significant results from our study indicate that users of the TVE shared the work more evenly than users of Movie Maker as we had hypothesized, and that users enjoyed using the TVE more than Movie Maker. We also found evidences that the TVE facilitated distributed cognition.

CONCLUSIONS AND FUTURE WORK

In this paper we have introduced the Tangible Video Editor (TVE), a tangible interface for editing sequences of digital video clips. By embedding handheld devices in the interface components, the TVE can be used on any surface under normal lighting conditions and without the need for an augmented surface, computer vision system, or video projector. We presented results from a study comparing the use of the TVE and Microsoft Movie Maker. These results show that our TVE was successful as a video editing interface and provided benefits in supporting collaboration, exploration and users' engagement as well as facilitating distributed cognition.

The current TVE implementation was designed for simplicity and focused on the fundamental operations of digital video editing. In order to truly recapture some of the lost advantages of traditional film editing we are planning to implement advanced features of video editing such as video scrubbing, force-feedback temporal controls and trimming in future versions of our TVE.

ACKNOWLEDGEMENTS

We thank Lynda Bayas, Matt Chesler, Jeffrey Held, Jeffrey Livingston, Ann Norman, Wen-Zea Kuo, and Ajahne Naphtali Santa Anna for their work on the hardware and software of the TVE, and for helping with the evaluation; Chris Rogers and the Tufts University Future Technologies Lab for the use of their laser cutter and workspace; Angela Chang and Oren Zuckerman for their discussions and suggestions; and Gene Caputo and ESPN for their hospitality and help. This work was supported by the National Science Foundation (grant IIS-0414389).

REFERENCES

1. Ananny, M., Supporting children's collaborative authoring: practicing written literacy while composing oral texts. In *Proc. Computer Support for Collaborative Learning 2002*, Lawrence Erlbaum Assoc. (2002), 595–596.
2. Fishkin, K., Gujar, A., Harrison, B., Moran, T., and Want, R. Embodied user interfaces for really direct manipulation. In *Communications of the ACM, 43, 9,* (Sep. 2000), ACM Press (2000), 74–80.
3. Gutwin, C. and Greenberg, S., Design for individuals, design for groups: tradeoffs between power and workspace awareness, In *Proc. CSCW 1998*, ACM Press (1998), 207–216.
4. Hollan, J., Hutchins, E., and Kirsh, D., Distributed cognition: towards a new foundation for Human-Computer Interaction research, In *TOCHI, 7, 2,* (June 2000), 174–196.
5. Hornecker, E., Understanding the benefits of graspable interfaces for cooperative use. In *Proc. of Cooperative Systems Design 2002*, 71–87.
6. Huang, C., Not just intuitive: examining the basic manipulation of tangible user interfaces, In *Ext. Abstracts CHI 2004* , ACM Press (2004), 1387–1390.
7. Ishii, Y., Nakanishi, Y., Koike, H., Oka, K., and Sato, Y., EnhancedMovie: movie editing on an augmented desk. In *Adjunctive Proc. of Ubicomp 2003*.
8. Jacob, R. J., Ishii, H., Pangaro, G., and Patten, J., A tangible interface for organizing information using a grid, In *CHI 2002*, ACM Press (2002), 339–346.
9. Krueger, M. W., Gionfriddo, T., and Hinrichsen, K., VIDEOPLACE—an artificial reality. In *Proc. CHI 1985*, ACM Press (1985), 35–40.
10. Mackay, W. and Pagani, D., Video Mosaic: laying out time in a physical space. In *Proc. of the Second ACM International Conference on Multimedia*, ACM Press (1994), 165–172.
11. Microsoft Movie Maker. http://www.microsoft.com/windowsxp/downloads/updates/moviemaker2.mspx/.
12. Moviola History. http://www.moviola.com/Mov_history.html/.

13. Patten, J., Ishii, H., A comparison of spatial organization strategies in graphical and tangible user interfaces, In *DARE 2000*, 41–50.
14. Reality-Based Interaction. http://www.eecs.tufts.edu/~jacob/theory/.
15. Rekimoto, J., Ullmer, B., and Oba, H., DataTiles: a modular platform for mixed physical and graphical interactions. In *Proc. CHI 2001*, ACM Press (2001), 269–276.
16. Scott, S.D., Shoemaker, G.B.D., and Inkpen, K.M., Towards seamless support of natural collaborative interactions, In *Proc. of Graphics Interface 2000*, 103–110.
17. Shaer, O., Leland, N., Calvillo-Gamez, E. H., and Jacob, R. J., The TAC paradigm: specifying tangible user interfaces. In *Personal and Ubiquitous Computing 8, 5,* (Sep. 2004), 359–369.
18. Snibbe, S. S., MacLean, K. E., Shaw, R., Roderick, J., Verplank, W. L., and Scheeff, M., Haptic techniques for media control, In *Proc. UIST 2001*, ACM Press (2001), 199–208.
19. Steenbeck Flatbed Editor. http://www.steenbeck.com/.
20. Suzuki, H. and Kato, H., AlgoBlock: A tangible programming language, a tool for collaborative learning, In *Proc. of 4th European Logo Conference 1993*, 297–303.
21. Tangible Video Browser. http://tangible.media.mit.edu/projects./tvb/.
22. Tangible Video Editor. http://www.cs.tufts.edu/~jacob/TVE/.
23. Ullmer, B., Ishii, H., and Glas, D., mediaBlocks: physical containers, transports, and controls for online media, In *Proc. SIGGRAPH 1998*, ACM Press (1998), 379–386.
24. Ullmer, B., Ishii, H., and Jacob, R. J., Token+constraint systems for tangible interaction with digital information, In *TOCHI 2005*, ACM Press (2005), 81–118.