To appear in *Proceedings of Advanced Visual Interfaces 1998 (AVI'98)*, May 1998.

Comparing MMVIS to a Timeline for Temporal Trend Analysis of Video Data

Stacie Hibino[†]
Bell Laboratories, Lucent Technologies
1000 E. Warrenville Road
Naperville, IL 60566 USA
hibino@research.bell-labs.com

Elke A. Rundensteiner
Computer Science Department
Worcester Polytechnic Institute, 100 Institute Rd.
Worcester, MA 01609-2280 USA
rundenst@cs.wpi.edu

ABSTRACT
Our MultiMedia Visual Information Seeking (MMVIS) environment provides an exploratory visual paradigm for temporal trend analysis. In this paper, we present the results of a user interface study evaluating the utility of MMVIS. We compare MMVIS to a timeline-based approach for analyzing temporal trends in real video data. We evaluate the quantity, complexity and accuracy of temporal trend observations made within each interface, compare the number of positive versus negative trends found, and collect feedback on user satisfaction. Our results show that subjects made interesting and complex observations of temporal trends using either interface. The results also indicate some advantages and biases of each interface, such as 1) timeline subjects make more errors during analysis and 2) timeline subjects are biased against identifying negative trends such as exceptions. At the same time, however, subjects appreciate the familiarity of timelines. Because we designed the MMVIS architecture to provide users with a library of visualizations, we thus include a discussion on enhancing the utility of MMVIS through incorporating a timeline into it in the future.

KEYWORDS: User interface evaluation, dynamic queries, video analysis, multimedia visual information seeking, temporal analysis.

INTRODUCTION

Our MultiMedia Visual Information Seeking (MMVIS) environment provides users with a novel interactive visualization approach to analyzing temporal relationship trends in temporal data such as video [11, 8]. In MMVIS, users can interactively select two subsets of events and then dynamically browse and query for temporal relationships between the selected subsets (e.g., to determine how often subset A video events occur 0, 1, or 2 seconds after subset B video events). We support this temporal browsing within MMVIS by tightly coupling specialized temporal query filters (referred to as TVQL, our temporal visual query language [12, 8]) with a dynamically updated temporal visualization (TViz) of results.

We have evaluated the utility and usability of MMVIS through one case study and two user interface studies. In our case study, we applied MMVIS to the temporal analysis of real CSCW video data of a design meeting [10]. This case study illustrated how our approach *can* be used to examine and identify temporal trends and how *different* types of temporal relationships (e.g., temporal sequences *or* overlaps) can be easily explored within MMVIS.

In our first user study, we evaluated the TVQL interface outside of the context of MMVIS [9]. We compared the users' ability to specify and interpret various types of temporal queries using TVQL versus a forms-based temporal query language (TForms). The study showed that while users spent more time learning TVQL than TForms, they were also able to specify temporal queries more efficiently and accurately with TVQL than with TForms.

In this paper, we now describe our second user study evaluating the fully integrated MMVIS environment. In this study, we evaluate the utility of MMVIS by comparing and contrasting its usefulness for temporal analysis to an alternative means of doing analysis. We chose to use a basic timeline for this comparison since 1) timeline-based formats are commonly used in video annotation and analysis systems [7, 5] and 2) we plan to incorporate a timeline format into MMVIS in the future and this comparison enables us to study a timeline for temporal analysis and characterize its utility in contrast to TViz (i.e., to validate the usefulness of having more than one type of temporal visualization available for temporal analysis).

Rather than using the same video data evaluated in our CSCW case study [10], we applied MMVIS to a second real video data set—video of the men's beach volleyball gold medal final game of the 1996 Summer Olympics. Applying MMVIS to this second video data illustrates the flexibility of the system for handling video from different types of domains (i.e., sociological versus sports video data) as well as the power of TVQL and the integrated MMVIS framework for exploring temporal relationships across these different domains. We present details on how we coded the volleyball video data in the next section.

In our MMVIS user study, subjects analyzed the volleyball data using MMVIS or the timeline format. The subjects looked for *temporal frequency trends* such as "which team made the most number of points?" as well as *temporal relationship trends* such as "do players from one team serve to a specific player on the other team more often?" Their

[†]This work was conducted while the author was at the University of Michigan and was supported in part by a University of Michigan Rackham Thesis Grant.

task was to answer a series of true/false, multiple choice, fill-in, and free form questions—first on temporal frequency trends and then on temporal relationship trends. In this paper, we compare the utility of the MMVIS and timeline interfaces for temporal analysis by evaluating the subjects' answers to the free form questions.

We compare the utility of the two interfaces for temporal analysis based on the quantity, complexity and accuracy of observations that subjects made, a comparison of the number of positive versus negative trends found in each interface, and feedback on user satisfaction for each interface. We also include an example-based comparison of the efficiency of using the interfaces for finding more complex trends. Our results show that while subjects could make interesting and complex observations of temporal trends using either interface, there were differences between the interfaces in terms of accuracy (timeline subjects made more errors) and the number of positive and negative trends found (timeline subjects did not identify any negative trends). In our discussion, we also highlight the advantages and biases of each interface and present suggestions for incorporating a timeline into the MMVIS framework.

EXPERIMENTAL METHOD

Design

Subjects were divided into two groups according to interface used—one group used MMVIS and the other used a timeline for temporal analysis. A between subjects design was then used to compare the interfaces. Subjects in each group performed all tasks for their given user interface.

Participants

Ten undergraduate and graduate students (six males and four females) participated in the study. All subjects had participated in the temporal visual query language (TVQL) user interface study [9] or were familiar with the TVQL interface. None of the subjects had used MMVIS nor seen the system applied to the volleyball video. Subjects had expertise and experience in either video analysis (VA) or databases (DB). Each subject was paid ten dollars an hour for a maximum of thirty dollars. All subjects had at least five years of computer experience and were familiar with the Macintosh and/or Windows operating systems.

All subjects were asked to rate their knowledge of volleyball by selecting one of four choices representing whether or not they had: very little, if any, knowledge about the game, a vague recollection about the basic rules of the game, definite knowledge about the basic rules of the game, or extensive knowledge about volleyball and volleyball strategies. All subjects had at least a vague knowledge about the rules of the game. Slightly more than half of the subjects had basic or extensive knowledge of the game. Because the basic rules of volleyball are relatively simple, this domain knowledge did not appear to affect the subjects' ability to analyze the video.

The distribution of volleyball domain knowledge within each group was exactly the same—with two subjects having a vague recollection of the game of volleyball, one subject being confident about knowledge of volleyball rules, and two subjects with extensive knowledge of rules and strategies of the game. In addition, each group had three male, and two female participants. The primary background differences between the groups was in database (DB) versus video analysis (VA) expertise. The DB to VA ratio of expertise was four to one for the MMVIS group and two to three for the timeline group. However, we did not examine differences between subjects based on expertise in this study, since no significant difference between VA and DB subjects was found during our TVQL user study [11].

Procedure and Materials

Video Coding. The sample video is from the two-man beach volleyball finals of the 1996 Summer Olympics. In this final game, two USA teams were competing for the gold medal: Team Red (Mike Dodd and Mike Whitmarsh) and Team Black (Karch Kiraly and Kent Steffes). The following types of events were abstracted as video annotations in order to capture and analyze the essence of the game: individual player actions, errors, plays, and rallies. Each individual player's action was coded with their name and one of the following actions: *block, dig, hit, kill, pass, serve*, or *set*. Any given action starts when the player first contacts the volleyball and ends when the next player contacts the ball or until the ball is considered "dead" (e.g., when it hits the ground). *Errors* were coded separately from players' actions. A separate error was coded for each player action that was an error so that 1) every error started and ended at the same time as an individual player action and 2) if one player's error was immediately followed by another player's error, the errors were coded as two separate errors rather than one longer error. *Plays* and *rallies* were also coded separately from players' actions. A *play* consists of one to three consecutive actions by players on the same team. A *block* is an exception to this rule in that it is not included in the overall action count during a play. Three types of plays were coded:

- *side-over play:* a play that is successfully sent over the net but doesn't immediately result in a point or side-out.
- *point play:* a play which results in a point scored. A *point play to Black* and a *point play to Red* indicate which team earned the point.
- *side-out play:* a play which results in a side-out (i.e., turnover). A *side-out play to Black* and a *side-out play to Red* indicate which team won the ball and will serve it next. No points are scored in a side-out.

A *rally* consists of one or more series of plays ending with a point or side-out play. A *point-rally to Black* and a *point rally to Red* indicate which team won the rally and scored a point. Similarly, a *side-out rally to Black* and a *side-out rally to Red* indicate which team won the rally and obtained serving possession of the ball without scoring a point. The types of rallies are mutually exclusive so that there is never any temporal overlap between rallies.

Figure 1 presents the sample timeline fragment shown and described to all subjects in a reference sheet. Note that a separate row in the timeline is provided for each type of event (e.g., a Dodd block, Kiraly serve, side-over play, etc.). In addition, events are grouped by one of five categories: Team Red, Team Black, Error, Plays, and Rallies. These categories are mutually exclusive and are indicated in the key along the far left-hand side of the timeline fragment depicted in Figure 1.

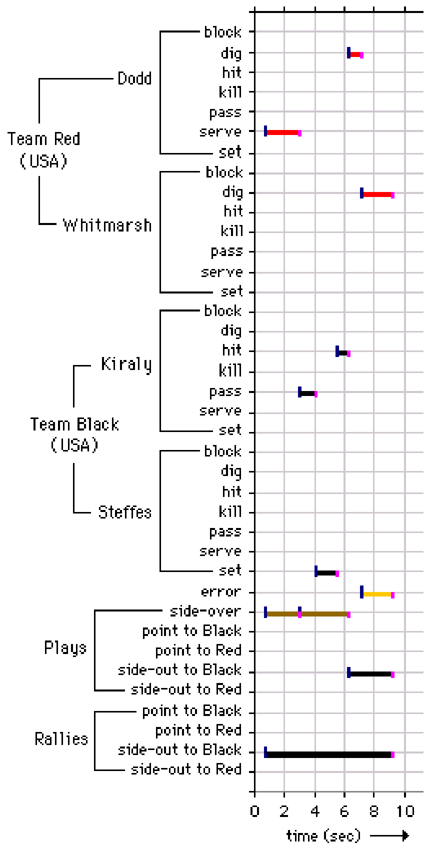

Figure 1. Timeline fragment of the first rally of the volleyball game.

The timeline fragment in Figure 1 shows the first rally of the game, including the following plays and actions:

Play	**Action(s)**
side-over	• Dodd starts the rally off with the first serve of the game
side-over	• Kiraly pass • Steffes set • Kiraly hit
side-out to Black	• Dodd dig • Whitmarsh dig (error)

Since the first rally ended in a *play side-out to Black*, the rally is coded as a *rally side-out to Black*. In the next rally (not shown on the timeline), team Black will serve. The action proceeds at a very quick pace and this first rally only took ten seconds. The full game lasted about thirty minutes. Each player action, error, type of play and type of rally of the video was coded, for a total of 660 video events.

MMVIS. While we have described MMVIS in more detail elsewhere [11, 8], we summarize its key features here. The *main MMVIS window* consists of a *visualization area* anchored in the upper left corner, a *key* for subset selection below the visualization, brief *instructions* in the upper right corner, and *visualization options* in the lower right corner.

Although MMVIS is designed to provide several alternative visualizations from which the user can select, the current prototype uses our abstract temporal visualization (TViz) of results. TViz initializes the visualization area with icons of the various types of events in the given data set.

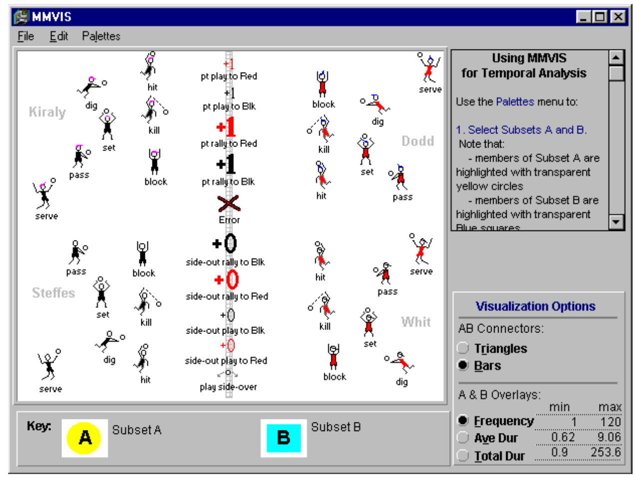

Figure 2. Main MMVIS window of the volleyball data.

In Figure 2, for example, the visualization area of the main MMVIS window indicates the various types of volleyball events taking place in the volleyball video data. Note that 1) each icon (i.e., type of event) in TViz corresponds to a separate row in the timeline and 2) the icons are arranged on the screen to provide an underlying context for temporal trend analysis (e.g., each player's actions are grouped by player, players are divided by teams, etc.)

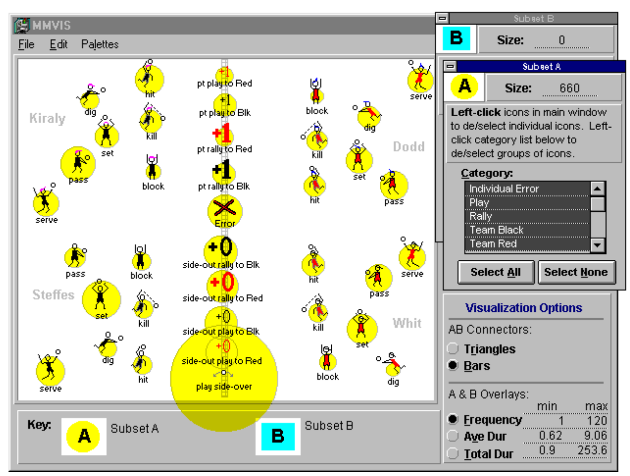

Figure 3.‡ Comparing temporal frequency of events.

In MMVIS, users conduct temporal analysis by comparing overall trends in temporal frequency and duration of events, selecting two subsets of events (Subset A and Subset B) and then using our specialized temporal dynamic query filters (our temporal visual query language, TVQL) to analyze temporal relationship trends such as "how often do A events start at the same time as B events?" Figure 3, for example, illustrates how users can compare temporal frequency of events. In this example, we have used the Subset A query

‡ Note: corresponding color plates are available for Figures 3-4 and 6-9.

palette to select all types of events and the Subset B palette to select none. The visualization options in the lower right of the main MMVIS window show that the A and B subset selection highlighters indicate the relative *frequency* of events. Thus, we see that on the Black team, Steffes sets more often then Kiraly during the volleyball game whereas on the Red team, Dodd and Whit seem to set fairly evenly.

If we set Subset A (indicated by circle overlays) to Dodd and Whit serves, and Subset B (square overlays) to Kiraly and Steffes passes and digs, we can then view the *AB sequence* relationship ( ) to compare how often each player on team Black receives serves from team Red. Figure 4 shows how we can use TVQL (the temporal query palette) to specify the AB sequence relationship and how the bars between A and B events in TViz are correspondingly updated to indicate the strength of the temporal relationship. In this example, the bar between Dodd serve and Kiraly pass is very similar in thickness to the bar between Dodd serve and Steffes pass, thus indicating that Kiraly and Steffes fairly evenly receive Dodd's serve. On the other hand, the bar between Whit serve and Kiraly pass is much thicker than the bar between Whit serve and Steffes pass, thereby indicating that Kiraly receives many more of Whit's serves than Steffes does. This trend provides some explanation as to why Steffes sets more often during the game than Kiraly sets. That is, Steffes sets every ball that Kiraly receives, since the same player cannot touch the ball twice in a row. If Kiraly receives a serve more often than Steffes, then Steffes is forced to set more often than Kiraly in plays involving serve reception.

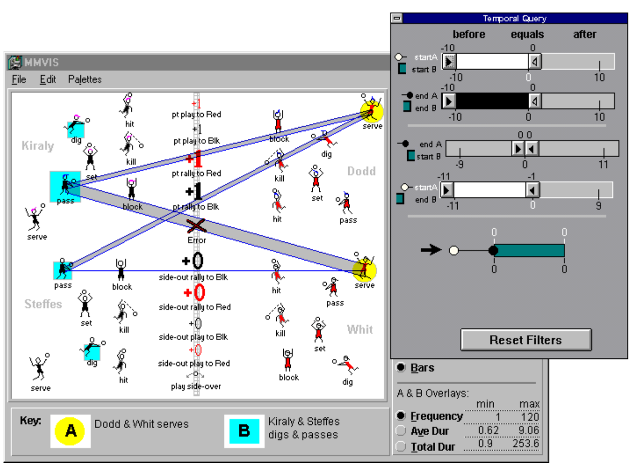

Figure 4. Using MMVIS to compare serve reception by specifying the AB sequence relationship ( ) in TVQL and reviewing the results in TViz.

Timeline. Timeline subjects were given a hard copy timeline of the full volleyball game data displayed over two 8.5" by 11" sheets of paper, with the first half of the game on one sheet and the second half of the game on the other sheet. A fragment from the first sheet of the timeline used is displayed in Figure 5. Timeline subjects also had access to the same timeline online, displayed in an *Apple Image Viewer*. This image viewer provides a simple interface for scrolling and zooming an image file.

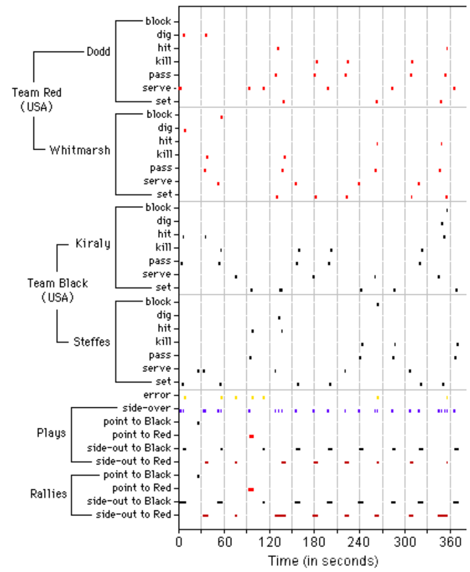

Figure 5. Fragment of the timeline used by subjects in the study (reduced 55%).

Procedure. At the start of each testing session, subjects completed consent forms and indicated their knowledge of the game of volleyball. They then read two reference sheets—one on the basic rules and terms used in two-person beach volleyball and one describing the sample volleyball video and how it was coded. This was followed by a time of clarification, to answer any questions subjects had on the terms described in the reference sheets or the way the video was coded. The remainder of the procedure for each interface was divided into three parts: *training, temporal analysis,* and *post-questionnaire on user satisfaction*.

Part I: Training. Timeline subjects were given simple verbal instructions on how to use the online image viewer. Since subjects were familiar with timeline formats and were given handouts describing the types of video events presented in the timeline, no other training was required.

MMVIS subjects were given online training materials including: a review of TVQL, description of changes to TVQL since the first study, description of MMVIS, and hands-on practice for MMVIS. The MMVIS hands-on practice included directions and practice time for 1) selecting subsets, 2) using TVQL and reviewing TViz for pre-selected subsets, and 3) using the full MMVIS environment for selecting subsets and exploring temporal relationships between events. The training materials followed the same order as the testing materials. During the first two hands-on practice tasks, only a limited set of functionality, corresponding to the practice tasks, was made available. For example, during subset selection practice, subjects could not access TVQL. Limiting the functionality of the system in this way, we used a "training wheels" approach [3] to teach subjects how to use the system.

After MMVIS subjects reviewed online training materials, they had time to ask clarification questions, and then were asked to demonstrate that they could specify two common queries—the *equals* ( ) and the *meets* ( )

temporal query relationships. All MMVIS subjects were able to specify these queries on their first try.

Part II: Temporal Analysis. Part II of each testing session was divided into two subparts—Part II.a regarding occurrence and duration (i.e., frequency and relative duration) of events and Part II.b on examining temporal relationships between various types of events. Table 1 summarizes the number and type of questions for each part. The MMVIS and timeline questions were isomorphic so that equivalent, though not identical, questions were given for each interface. A full listing of the questions used in Part II can be found in [8].

Table 1. Summary of the number and type of questions given in Part II of user testing (T/F=True/False, MC=Multiple Choice, FI=Fill-In, FF=Free Form).

	T/F	MC	FI	FF
Part II.a. Analysis of Occurrence and Duration				
MMVIS	3	3	2	1
Timeline	3	3	2	1
Part II.b. Analysis of Temporal Relationships				
MMVIS	6			2
Timeline	6			1

In Part II.a, subjects had to answer three true/false, three multiple choice, two fill-in and one free form question. In the case of the free form question, subjects were asked to identify at least two findings of their choice, explain how they came to their conclusions, and indicate if and why they found their results to be expected or surprising. MMVIS subjects were restricted to subset selection during Part II.a since TVQL was not required to answer these questions.

In Part II.b, all subjects first answered a series of six true/false questions. In both interfaces, one scenario was used to answer the first three true/false questions and a second scenario was used to answer the next three true/false questions. That is, for each scenario, the subjects were asked to answer questions about temporal relationships between specific types of events. For each scenario, MMVIS subjects used a version of the system where the A and B subsets were preselected for them and MMVIS functionality was limited to TVQL so that they only needed to adjust TVQL to answer the questions. In addition, subset selection was disabled so that users could not select alternative subsets in any of these pre-determined scenarios.

Figure 6 presents a sample screen shot of the third MMVIS scenario, indicating the types of events preselected for subset A and B (A events are highlighted with transparent circle overlays; B events are indicated with square overlays) and showing the results of a user-specified TVQL query (i.e., the *meets* temporal relationship), used to examine AB sequences of events. Limiting MMVIS functionality by preselecting A and B subsets allowed us to examine subjects' ability to use TVQL and interpret TViz to analyze particular temporal relationship trends.

Part II.b also included free form questions in addition to the multiple choice questions. The MMVIS subjects had two free form questions—one based on a third scenario (and hence preselected A and B subsets), and one where no scenario was specified and subjects were free to select any A and B subsets and examine temporal relationships between these new subsets. Since it was impossible to constrain the timeline interface for the free form question (e.g., we could suggest A and B subsets on which subjects could focus, but we could not enforce this scenario in the case of a free form question), we only gave timeline subjects one free form question to answer in Part II.b. While MMVIS subjects had more practice in answering free form questions in Part II.b, timeline subjects were given hints and suggestions during their final free form question on the types of relationships they could examine (some timeline subjects followed these hints, while others investigated their own questions).

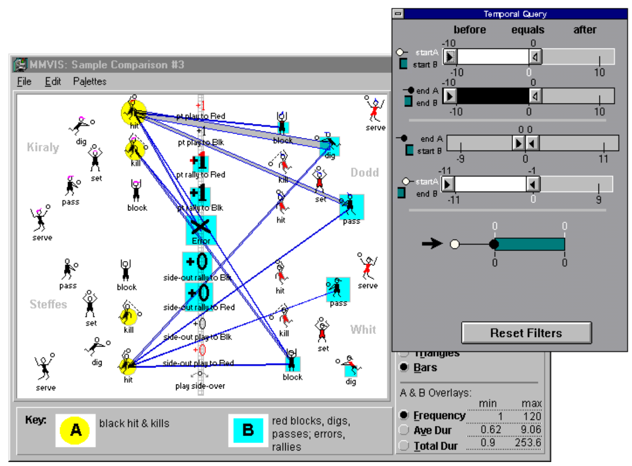

Figure 6. Sample screen shot of the third MMVIS scenario used during testing.

Part III: Post Questionnaire on User Satisfaction. At the end of the interface testing, subjects completed a user satisfaction questionnaire which included: a subset of rating scales from QUIS [4] on overall reaction, learning, features of the screen, system terminology, and system capabilities; open-ended questions on what subjects liked or disliked about the system, what they thought was easy or difficult about the interface, and any final comments they had about the system such as any specific suggestions for improving the interface. In addition, an informal post-interview was conducted to answer any of the subjects' remaining questions, to collect any additional feedback, and to discuss any comments included in the post-questionnaires.

Hardware and Software Setup. The training module, testing materials, MMVIS scenarios and MMVIS prototype were all computerized materials that were developed in *Asymetrix Multimedia ToolBook v3.0*. The online timeline was saved as a single image file and displayed using the *Apple Image Viewer*. The testing sessions were conducted on Dell Pentium 90 (90 MHz Pentium) desktop machines equipped with 17-inch SuperVGA monitors and running Microsoft Windows NT.

Types of Data Collected

We collected the following data for both interfaces: background information on subjects' knowledge of volleyball, logfile information on time taken and answers given during online testing, post-questionnaires on user satisfaction, observational data, and notes from informal post-interviews. In the case of the timeline, we also collected paper timelines used in order to further examine the marks and folds subjects made while using them. During MMVIS training, we also collected information on time spent reviewing online training materials.

RESULTS AND DISCUSSION

Data Analyzed

Since the multiple-choice, true/false, and fill-in questions were used to provide feedback and additional examples to subjects on using the corresponding interfaces for temporal analysis, we focus our evaluation on the observations subjects made during the free form questions. As described in the previous section, timeline subjects had *two* free form questions while MMVIS subjects had *three* free form questions. However, in the second free form MMVIS question, subjects could *not* choose or change the A and B subsets selected. Only in the final free form question could MMVIS subjects both specify A and B subsets and explore temporal relationships between them. We thus compare the observations of the two timeline free form questions to the first and last MMVIS free form questions.

Coding Scheme and Overall Results

Observations from each group were coded according to the following criteria: quantity and complexity, accuracy, and the presence of positive versus negative trends. The coding scheme for rating the type and complexity of observations is presented in Table 2 for temporal frequency and duration trends and in Table 3 for temporal relationship trends. The accuracy of each observation was rated as true (t), partially true (pt), false (f), or neutral (n—for observations that were merely comments or interpretations). Each trend was also rated as positive or negative. A *negative trend* indicates the *absence* of a relationship whereas a *positive trend* represents the *presence* of a relationship. A sample negative temporal trend includes an observation such as "serves are *never* directly met with blocking."

When comparing MMVIS to the basic timeline format for identifying temporal data trends, we see that, on average, subjects using MMVIS spent significantly more time than timeline subjects on the last free form question format and on all of Part II.b, the analysis of temporal relationship trends ($p < 0.05$; see Table 4). Note that the time to answer the free form questions is only an estimate in that it 1) includes the time subjects took to make and enter (i.e., type) their observations, 2) includes time used to explore any additional observations that were not recorded, and 3) is dependent on individual subject's intrinsic motivation to search for trends (subjects were asked to find at least two trends or exceptions to trends but were not required to do so; thus, some subjects made more than two observations, but not all subjects did so).

Table 2. Coding scheme characterizing type and complexity of subjects' observations of temporal *frequency and duration* trends.

Code	Meaning	Example
ac	average duration comparison: comparison between different types of events based on average duration of the events.	Members of the black team took less time to serve than members of the red team.
tc	total duration comparison: comparison between types of events based on total duration of the events.	Steffes spent much more time setting than anyone else.
co	count only, where the total count is less than 20: this type of observation could be determined by simply counting individual events in the timeline format.	Kiraly only had two blocks.
co+20	count only, where the total count is >= 20.	Steffes set the ball 33 times.
co+vt	count only using video time reference.	The red team only scored 3 points in the first half of the game.
dp	density preview: a glance at the overall picture of TViz or the timeline could be used to determine which type of event occurs the most or the least frequently.	The most common event was side-over.
dp4+	density preview with focus on four or more types of events: overall qualitative comparison between four or more types of events.	Kiraly hits and kills much more than Steffes.
cc	count and compare, where total count is < 20: this type of observation could be determined by counting two or more types of individual events and comparing their totals.	The black team had two more blocks than the red team (9 blocks vs. 7 blocks).
cc+20	count and compare, where total count >= 20.	The black team has more actions (196) than the red team (174).
cc+vt	count and compare using video time reference.	Each team scores a point within the first two minutes of the game, but after these initial points, there is a long time where neither team scores a point.

Overall, we also see that MMVIS subjects made a total of more temporal *frequency and duration* trend observations than timeline subjects (Table 5), but fewer temporal *relationship* trend observations than the timeline subjects (Table 6). However, timeline subjects made more errors in their observations. We provide a more detailed discussion of these results below.

Table 3. Coding scheme to characterize the type and complexity of subjects' observations of temporal *relationship* trends.

Code	Meaning	Example
trel1	temporal relationship comparison between *one* type of (A,B) pair types.	Steffes' hits are never blocked by Dodd. (Compares sequence relationship for (Steffes hit, Dodd block).)
trel2	temporal relationship comparison between *two* types of (A,B) pair combinations.	Both Kiraly and Steffes made hits that were errors. (Compares the *equals* temporal relationship for (Kiraly hit, Errors) and (Steffes hit, Errors).)
trel3+	temporal relationship comparison between *three or more* types of (A,B) pair combinations.	The black team makes more errors than the red team (comparison between all (A, Error) pairs, where A is any individual player's action).
spa	sequence pattern analysis: examine the sequence of three or more types of events. (This can only be done with the timeline.)	In examining the sequence of pass-set-hits between Kiraly and Steffes...
spa+	sequence pattern analysis along with additional complex temporal relationship analysis. (This can only be done with the timeline.)	Comparing the pass-set-kill sequences of the Black team to their points and turnovers, we see that...

Table 4. Summary of average time spent (in minutes) in the free form and each part of the user testing (*=statistically significant ($p < 0.05$) for Scheffe post-hoc analysis).

Average Time Spent (in minutes)	MMVIS	Timeline
Part II.a. Analysis of Occurrence and Duration		
1st free form question	17.8	14.7
average total time for Part II.a.	20.7	19.8
Part II.b. Analysis of Temporal Relationships		
2nd free form question	11.1	-
last free form question	* 19.6	11.8
average total time for Part II.b.	* 43.3	23.4

Table 5. Summary totals of accuracy of answers to free form questions on temporal *frequency and duration* trend observations (t=true, pt=partially true, f=false, n=neutral).

Relationship	MMVIS				Timeline			
(see Table 2)	t	pt	f	n	t	pt	f	n
ac	3				1			
tc	2							
co					1			
co+20								
co+vt								
dp	1				2		1	
dp4+	1				1			
cc								
cc+20	8				2			
cc+vt								
Totals	15				7		1	

Table 6. Summary totals of accuracy of answers to free form questions on temporal *relationship* trend observations (t=true, pt=partially true, f=false, n=neutral).

Relationship	MMVIS				Timeline			
(see Table 3)	t	pt	f	n	t	pt	f	n
trel1					3		2	
trel2	3							
trel3+	6		1		5	2	1	
spa					1			
spa+					1			
interpretation				1				
Totals	9	-	1	1	10	2	3	-

Analyzing Temporal Occurrence and Duration

In MMVIS, temporal occurrences and durations of events are examined via subset selection. By using "select all" from either of the subset selection palettes, users can gain a visual overview comparing the relative frequency, average duration, or total duration of various types of events. In addition, they can access quantitative totals for each type of event by moving the mouse cursor over its corresponding icon in TViz. In the timeline, users can compare the frequency of events by examining and comparing the density of events in any given row or by counting the number of events that occur in a particular row. They can also compare the duration of events by looking for events displayed as long lines in the timeline versus those drawn as short blips on the page or screen.

When examining the types of temporal occurrence and duration trends observed in each interface, we see that the MMVIS subjects took advantage of the extra built-in features of MMVIS, making average duration comparisons (*ac*), total duration comparisons (*tc*), and several large count and compare (*cc+20*) observations. MMVIS subjects also made interesting use of the subset selection palettes for doing large count and compare operations (Figure 7). For example, one subject observed that "the black team has more actions (196) than the red team (174)" while another subject compared the number of actions between team members, noting that "Kiraly is involved with more action than Steffes (102-94), whereas Dodd and Whitmarsh have equal number of actions (87-87)." In order to accomplish the first observation, the subject 1) set subset A to all actions of the black team, 2) set subset B to all actions of the red team, and then 3) compared the tallies at the top of the subset selection palettes (see Figure 7).

While it is feasible to make these high count and compare observations with the basic timeline, it would be much more time consuming to do so and the actual observations made by timeline subjects confirm that indeed none of them took the time to make these types of observations. We can improve an online timeline by providing mechanisms for accessing this information. In particular, we note that in addition to obtaining quantitative values of how often a *specific* type of event occurs, users may also want access to quantitative information on *groups* of different types of events. Furthermore, in both MMVIS and the timeline, it would be useful to access frequency information—of both individual and groups of various types of events—over user-specified time intervals. This would, for example, allow users to compare temporal frequencies over the first half of the video to that of the second half of the video, to frequencies for the whole video.

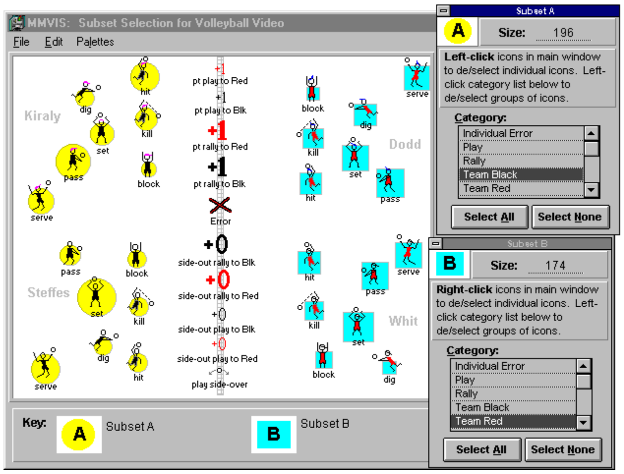

Figure 7. Sample interesting use of subset selection to compare frequency sums of groups of events.

Analyzing Temporal *Relationship* Trends

MMVIS subjects analyze trends in temporal relationships between various events through the following process: 1) selecting A and B subsets (if they are not already pre-selected for them), 2) using TVQL to specify a particular temporal relationship to analyze, and then 3) reviewing TViz to make a qualitative comparison between the presence or absence of AB connectors between A and B types of events. In contrast, users of the timeline interface must a) use a Gestalt approach for identifying gross patterns or b) locate individual events by hand and compare them to events in separate rows of the timeline to determine whether or not there is a trend in the temporal relationships between the types of temporal events examined.

Comparing the quantity and accuracy of temporal *relationship* trends observed using each interface, we found that while timeline subjects identified more relationship trends than the MMVIS subjects, they also made more errors in their observations. More specifically, timeline subjects made a total of ten true, three false and two partially correct observations while MMVIS subjects made nine true, only one false, and one neutral observations (Table 6). Examining timeline errors individually, we found that these errors appear to be related to cases where subjects inaccurately estimate trends based on glancing over or spot-checking the data rather than analyzing the data in a more systematic manner. Because data is widely distributed in the timeline format, subjects may also have focused on one portion of the timeline and then over-generalized over the full timeline. On the other hand, the timeline errors may be an indication that when trends are not obvious, humans are not very good at comparing distributed information. We plan to reduce these types of timeline errors by using an enhanced online timeline and/or by integrating a timeline into the MMVIS framework.

Another difference between the interfaces is in the complexity of trends identified. The timeline results contain five *trel1* observations, which are less complex than the *trel2* and *trel3+* observations made with MMVIS. In the case of the timeline, it is more work to evaluate the more complex relationships and the data shows that there was some aversion to this extra work by subjects, even for a relatively small data set. This also illustrates the potential benefit of enhancing the timeline by integrating it with TVQL into the MMVIS framework.

While timeline subjects made more of the less complex observations, some of these subjects did make some complex observations. One timeline subject, in particular—though he had only a vague recollection of the game of volleyball—spent a lot of time examining and annotating details of the timeline and identifying more complex trends such as the *spa* and *spa+* types of temporal trends. These types of trends, based on temporal sequences of events and comparing these sequences to other types of events, are not possible in the current version of MMVIS and thus illustrate the importance of making different types of visualizations available within MMVIS.

One of the biggest differences between the two groups is in identifying positive versus negative relationship trends. Subjects using the timeline identified *only* positive trends while the MMVIS subjects identified *both* positive and negative trends. In TViz, a negative trend is represented by the *absence* of a connector between an event type of Subset A and an event type of subset B. Sample negative temporal trends found by the MMVIS subjects include: "serves are *never* directly met with blocking" and "Whit is the only player to *not* commit an error when serving." The lack of negative trends in the timeline interface indicates that a timeline visualization might be biased towards looking only for positive trends rather than negative ones. This distinction provides support for including a visualization into MMVIS such as TViz that is *not* timeline-based.

We were surprised to see that some of the subjects actually took the time to identify the more complex type of temporal relationship trends. However, when we took a closer look at the observations in each group, we found additional differences between relationship trends involving high versus low frequencies of events. In fact, seven out of the fifteen timeline observations on temporal relationship trends were based on cases where one subset had a relatively low frequency (consisting of less than or equal to twelve individual events) while *none* of the MMVIS temporal relationship observations were based on such a low frequency. This indicates subjects' tendency towards looking for trends where search is reduced as much as possible, and raises the question of whether or not a timeline visualization will scale-up to handle longer video and/or a larger number of video events.

Although some timeline subjects did take the extra time to analyze temporal relationship trends where both subsets had a relatively *high* frequency, such analysis could be done more efficiently with MMVIS. For example, one observation made by both a timeline and an MMVIS subject was that the black team made more errors than the red team. This observation is interesting and somewhat surprising, given that the black team won the game. In order to make this observation, an MMVIS user could select all team black and all team red actions as subset A, errors as subset B, and then use TVQL to select the *equals* temporal relationship (see Figure 8). In contrast, a timeline subject would have to locate each occurrence of the 28 errors, identify which team made the given error, and tally the results by hand.

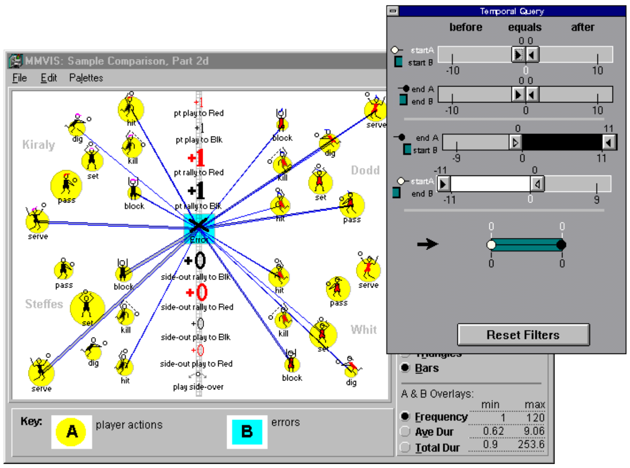

Figure 8. Using MMVIS to identify types and relative frequency of errors that occurred during the game.

In addition to improving efficiency in this situation, MMVIS also 1) summarizes which types of actions occur most frequently as errors, and 2) allows users to explore similar temporal relationships. That is, Figure 8 (an analysis conducted by one of the MMVIS subjects) shows how TViz can be used to easily compare results such as errors in serves between different players, identify where errors do not occur, etc. In this way, we see how results to our TVQL query potentially provides answers to several temporal queries at once. Using TVQL, we can examine what types of events occurred immediately before errors (Figure 9). In order to find the results to the same query in a timeline format, the user would have to go back to each individual error again and create a new tally of results corresponding to user actions before errors.

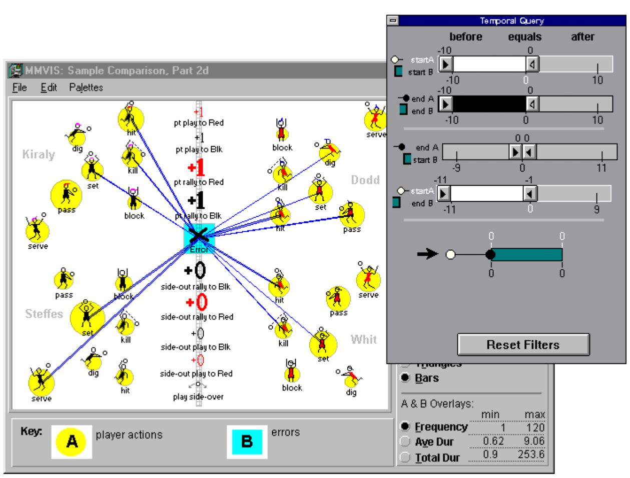

Figure 9. Updating the query in Figure 8 to examine what types of events occurred before errors.

Comparing User Satisfaction

Table 9 summarizes subjects' ratings of each interface based on their overall reaction to the system and their experience with learning the interfaces. While the post-questionnaires contain space for rating other aspects of the user interface, most of these were not as applicable to the paper timeline and hence many subjects did not answer them. A Scheffe post-hoc analysis indicated that these ratings were not significantly different.

Table 9. Comparison of user satisfaction ratings: average ratings for overall reaction and learning categories of QUIS, based on a five point scale.

QUIS Category	MMVIS	Timeline
Overall Reaction	3.7	3.1
Ease of Learning	3.6	3.9

INCORPORATING A TIMELINE INTO MMVIS

If a timeline were integrated into MMVIS and tightly coupled to TVQL, we could use vertical lines to connect individual events meeting the temporal relationship criteria specified via TVQL. If only a portion of the timeline were viewable at a time, these vertical lines could aid us in making local comparisons by simply reviewing the frequency of connectors between individual events. For example, if *Kiraly set* was followed by *Steffes hit* X times, then there would be a total of X lines between X pairs of *Kiraly set* and *Steffes hit* sequences.

These vertical lines, however, would be distributed across the timeline, potentially making it difficult for users to abstract an overall summary of the strength of temporal relationships between events. In TViz, this type of global visual summary is currently indicated by the thickness of AB connectors. To address this problem, we could incorporate a qualitative visual summary in a timeline along the key of the timeline (e.g., by drawing a bar between row labels of the timeline). This could then aid timeline users in making global comparisons over the full video in addition to local comparisons over only a portion of the video.

In general, TViz is more biased towards *global* rather than *local* trend finding, since it currently only provides results for the full video. While we can improve the use of TViz for identifying local trends by providing mechanisms for users to select a portion of the video to analyze, we are still missing some of the details available in a timeline format. Thus, another option for integrating a timeline into MMVIS would be to link appropriate timeline fragments to corresponding portions of TViz and making these timeline fragments available in a separate window at the users' request. Then, rather than restricting a timeline view to be used only as an *alternative* to TViz, we could provide timeline information *in addition* to TViz. TViz would represent a query preview and the timeline would present detailed information on the query results. This use of dynamic queries for preview and detail has been explored in other VIS applications [14]. In the future, we plan to explore and evaluate several alternatives for integrating a timeline format into MMVIS for temporal analysis.

RELATED WORK

In the past, video analysis[1] has focused on supporting users in accessing individual events or in reviewing timeline-based views of all or a time slice of events (e.g., [5, 7]). In these types of environments, users must either 1) *pre-code* temporal relationships rather than querying for relationships between events or 2) use a Gestalt approach to reviewing the timeline in search of temporal patterns.

While recent work in interactive visualizations for analyzing temporal data has emerged to empower users with more sophisticated tools, this work has also been timeline-based [15, 13, 6]. In addition, work by [15] is limited in that it has thus far focused on displaying relatively small data sets at a time (e.g., timeline data of a single patient's medical history) and work by [13] is limited to the analysis of temporal sequences. Eick and Lucas [6] present one of the more sophisticated environments available and have identified some classes of patterns which their system supports users in finding (e.g., periodic sequences). However, they still rely on a user's Gestalt interpretation of the data rather than providing direct support for posing temporal relationship queries. Although these systems may be susceptible to biases we found in timeline-based temporal analysis (e.g., bias against locating negative trends), they can also provide some insights as we work on incorporating a timeline into MMVIS in the future.

Our TVQL and MMVIS environment are unique extensions to dynamic query filters and visual information seeking (VIS [1]) for the purpose of analyzing temporal trends in video data. Previous user studies comparing VIS and dynamic query interfaces to alternative query mechanisms (e.g., forms-based query interfaces) have demonstrated the power of this direct manipulation approach for various query tasks ranging from finding a particular data item to searching for data trends and exceptions to trends [2, 16]. In this paper, we have further demonstrated the power of exploratory query paradigms and have, in particular, demonstrated that our approach is not only useful for identifying interesting data trends, but have also shown its utility in finding *exceptions* to trends.

CONCLUSION AND FUTURE WORK

In this paper, we compared the utility of MMVIS to a basic timeline for analyzing temporal trends in video data. Subjects used MMVIS or a timeline to analyze video of the men's beach volleyball final game of the 1996 Summer Olympics. Both MMVIS and timeline subjects were able to make interesting and complex temporal trend observations. However, timeline subjects made more errors in their observations and were biased towards identifying positive temporal trends. MMVIS subjects, on the other hand, identified both positive and negative trends (e.g., a negative trend such as "Whitmarsh is the only player who did *not* make a service error"). At the same time, some timeline subjects took advantage of the detailed information and layout available in the timeline format, observing trends in sequences or localized trends that changed over time. We thus discussed how we might integrate a timeline with TVQL as an alternative visualization in MMVIS. In the future, we plan to investigate the addition of a timeline into MMVIS, continue studying and improving our TVQL interface, apply our approach to new temporal data sets, and iterate on our user testing of the system.

[1]Note that we use the term *video analysis* to refer to an object-level analysis of relationships between events rather than a bit-level analysis for the purpose object extraction.

REFERENCES

1. Ahlberg, C., & Shneiderman, B. (1994). Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *CHI'94 Conf. Proc.* ACM Press, 313-317.
2. Ahlberg, C., Williamson, C., & Shneiderman, B. (1992). Dynamic Queries for Information Exploration: An Implementation and Evaluation. *CHI'92 Conference Proceedings*, NY:ACM Press, 619-626.
3. Catrambone, R. and Carroll, J. (1987). Learning a Word Processing System with Training Wheels and Guided Exploration. *CHI'87 Conf. Proc*. NY:ACM Press, 169-174.
4. Chin, J., Diehl, V. and Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *CHI'88 Conf. Proc*. ACM Press, 213-218.
5. Davis, M. (1993). Media Streams: An Iconic Lang. for Video Annotation. *Telektronikk 4.93:Cyberspace* 89(4), 59-71.
6. Eick, S. and Lucas, P. (1996). Displaying trace files. *Software—Practice and Experience*, 26(4), 399-409.
7. Harrison, B.L., Owen, R. and Baecker, R.M. (1994). Timelines: An Interactive System for the Collection of Visualization of Temporal Data. In *Proc. of Graphics Interface'94.* Toronto: Canadian Info. Proc. Society, 141-148.
8. Hibino, S. (1997). *MultiMedia Visual Information Seeking*. University of Michigan PhD dissertation.
9. Hibino, S. and Rundensteiner, E. (1997a). User Interface Evaluation of a Direct Manipulation Temporal Visual Query Language. *ACM Multimedia'97 Conf. Proc*., 99-107.
10. Hibino, S. and Rundensteiner, E. (1997b). Interactive Visualizations for Temporal Analysis: Application to CSCW Multimedia Data. In *Intelligent Multimedia Info. Retrieval* (M. Maybury, Ed.), 313-335. Cambridge, MA: MIT Press.
11. Hibino, S. and Rundensteiner, E. (1996a). MMVIS: Design and Implementation of a Multimedia Visual Information Seeking Environment. *ACM Multimedia'96 Conference Proceedings.* NY:ACM Press, 75-86.
12. Hibino, S. and Rundensteiner, E. (1996b). A Visual Multimedia Query Language for Temporal Analysis of Video Data. In *Multimedia Database Systems: Design and Implementation Strategies* (K. Nwosu, B. Thuraisingham and P.B. Berra, Eds.), 123-159. Norwell, MA: Kluwer Acad. Publ.
13. Jerding, D.F., Stasko, J.T. and Ball, T. (1997). Visualizing interactions in program executions. *1997 Int'l Conference on* Software *Eng. (ICSE'97) Proc*. NY:ACM Press, 360-370.
14. North, C., Shneiderman, B. and Plaisant, C. (1997). Visual Information Seeking in Digital Image Libraries: The Visible Human Explorer. In *Information in Images* (G. Becker, Ed.), Thomson Technology Labs (http://www.thomtech.com/ mmedia/tmr97/chap4.htm).
15. Plaisant, C., Milash, B., Rose, A., Widoff, S. and Shneiderman, B. (1996). Lifelines: Visualizing Personal Histories. *CHI'96 Conf. Proc.* NY:ACM Press, 221-227.
16. Williamson, C. and Shneiderman, B. (1992). The Dynamic HomeFinder: Evaluating Dynamic Queries in a Real-Estate Info. Exploration System. *SIGIR'92 Proc*. NY: ACM Press.