MMVIS: A MultiMedia Visual Information Seeking Environment for Video Analysis

Stacie Hibino and Elke A. Rundensteiner
EECS Department, Software Systems Research Laboratory
The University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122 USA
E-mail: hibino@eecs.umich.edu, rundenst@eecs.umich.edu

ABSTRACT

Our MultiMedia Visual Information Seeking (MMVIS) environment is designed to support an *exploratory* approach to video analysis. Specialized subset, temporal, spatial, and motion dynamic query filters are tightly coupled with dynamic, user-customizable *relationship* visualizations to aid users in the discovery of data trends. Users can select two subsets (e.g., a subset of person P1 talking events) and then *browse* various relationships between them (e.g., browsing for temporal relationships such as whether events of type A frequently start at the same time as events of type B). The visualization highlights the frequencies of both the subsets and the relationships between them. This allows users to discover various relationships and trends without having to explicitly pre-code them. In this demonstration, we will focus on temporal analysis aspects of the system, presenting our temporal visual query language, temporal visualization, and an application to real CSCW data.

Keywords

Video analysis, dynamic queries, temporal query filters, interactive visualizations, trend discovery.

INTRODUCTION

Visual Information Seeking (VIS) is a framework for information exploration where users filter data through direct manipulation of dynamic query filters [2]. A visualization of the results is dynamically updated as users adjust a query filter, thus allowing them to incrementally specify and refine their queries. In this way, users also see the direct correlation between adjusting parameter values and the corresponding changes in the visualization of results. This approach has been shown to aid users in locating information, as well as for searching for trends and exceptions to trends—and to accomplish such tasks more efficiently than through traditional forms-based methods [1]. We thus extend the VIS framework to handle multimedia data sets, more specifically to perform video analysis [5]. Our extensions give users the power to explore various relationships between different types of video events, in a way that was not previously possible through other traditional means (e.g., timelines for temporal analysis, statistically based approaches, etc.) [6].

THE MMVIS ENVIRONMENT

Our MultiMedia Visual Information Seeking (MMVIS) environment is a system designed to study such an

application of VIS to video analysis. Several extensions to the original VIS framework were made to accomplish this:

- *subset query palettes* with multi-selection list filters for specifying multiple subsets of different types of events (e.g., all person P1 talking events)
- *specialized temporal, spatial,* and *motion* query filters for exploring the corresponding types of relationships between the subsets formed,
- user-customizable *spatio-temporal* visualizations for highlighting the occurrence of the selected subsets as well as the frequency of the specified relationships.

Sample Scenario

In our demo, we will use a sample scenario from a real CSCW case study to provide some context. Consider the case where researchers collect CSCW video data to analyze and characterize the process flow of a planning meeting between three subjects ("Carol," "Richard," and "Gary") collaborating from remote sites. The data is coded to indicate when each person speaks as well as to characterize the design rationale of what is being said (e.g., to indicate when criteria, alternatives, etc., take place in the meeting).

Selecting and Visualizing Subsets

In MMVIS, users first select two subsets (A and B) via subset query palettes (see Figure 1, Subset A query palette). We designed multi-selection filters so that users can select one or more items from a list of alpha-numeric data. Vertical bars along the side of the lists indicate the last action taken and its impact on the values of other parameters. In Figure 1, the Subset A query palette selects all Activity (Talking & NonVerbal) types of events while Subset B selects all design rationales [6].

Yellow transparent circles are displayed in the visualization to highlight the corresponding A events, as the user de/selects values from each parameter list. Similarly, blue transparent squares indicate B events. The radius of these transparent overlays represent either relative frequency (Figure 1), average duration, or total duration, customized according to the user's preference. Display options are available in the lower right corner of the main MMVIS window. By switching back and forth between display options, the user can gain additional information about the data (e.g., such as whether or not events with low frequency have relatively high average duration).

Exploring Relationships Between Event Subsets

Once users have selected subsets, they can then *explore* various relationships between members of these subsets

using the specialized relationship query filters. Our temporal query filters, forming a temporal visual query language (TVQL) [5], are presented to the user on a single palette (see Figure 1, Temporal Query palette). TVQL can be used to specify any one of thirteen temporal primitives (e.g., before, meets, equals) as well as combinations of such primitives. In Figure 1, TVQL specifies the relationship where events of type A start at the same time as events of type B, but A events can end before, at the same time as, or after B events end. This represents a combination of the starts, equals and started-by temporal primitives. The temporal diagram at the bottom of the palette visually confirms this, and is dynamically updated as users adjust any one of the temporal query filters.

As users manipulate the temporal query filters, they can also review the visualization of results (and changes in it) for trends and exceptions. The existence of a relationship between A and B events is visually indicated as a connector drawn between their centers. The width of the connector indicates the relative frequency of the temporal relationship. For example, Figure 1 indicates that Gary never starts talking at the same time as a Digression; and NonVerbal events frequently start at the same time as a Pause. TVQL can be used to easily *browse* variations on the temporal relationship specified. For example, users could adjust the second temporal query filter (endA-endB filter) to see how the visualization changes when Activities (Talking and NonVerbals) end before or at the same time as (but not after) Rationales end. This could be done simply by moving the right thumb to zero.

COMPARISON TO SIMILAR SYSTEMS
Although several video annotation and analysis systems have been developed, they have focused on novel approaches to video annotation, timeline-based formats for video analysis, or pre-coding relationships rather than searching for them [7, 4]. The novel approach to video analysis presented in MMVIS empowers users to *explore* the data in search of trends and exceptions to trends.

Other extensions to VIS have been done [3], but they do not address the spatio-temporal and *relative* exploratory needs of video analysis. MMVIS introduces some new extensions to VIS—the use of specialized temporal query

filters and spatio-temporal visualizations, tailored to highlight the strengths of *relationships* between different types of subsets.

STATUS AND FUTURE WORK
MMVIS has been implemented on a multimedia PC (MPC) platform using a ToolBook® interface to a database library. All temporal analysis components are fully integrated and functional. In the future, we plan to continue work on several aspects of MMVIS, including new visualizations and integration of spatial and motion query filters.

REFERENCES
1. Ahlberg, C., Williamson, C., & Shneiderman, B. (1992). Dynamic Queries for Information Exploration: An Implementation and Evaluation. *CHI'92 Conference Proceedings*. NY:ACM Press, pp. 619-626).
2. Ahlberg, C., & Shneiderman, B. (1994). Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *CHI'94 Conference Proceedings*. NY:ACM Press, pp. 619-626.
3. Fishkin, K. and Stone, M.C. (1995). Enhanced Dynamic Queries via Movable Filters. *CHI'95 Conference Proceedings*, 415-420. ACM Press.
4. Harrison, B.L., Owen, R., & Baecker, R.M. (1994). Timelines: An Interactive System for the Collection of Visualization of Temporal Data. *Proc. of Graphics Interface '94*. Canadian Information Processing Society.
5. Hibino, S. & Rundensteiner, E. (in press). A Visual Query Language for Temporal Analysis of Video Data. *The Design and Implementation of Multimedia Database Systems* (K. Nwosu, Ed.), NY: Kluwer Books.
6. Hibino, S. & Rundensteiner, E. (1995). Interactive Visualizations for Temporal Analysis: Application to CSCW Multimedia Data. UM Tech. Rep.#CSE-272-95.
7. Roschelle, J., Pea, R., & Trigg, R. (1990). VIDEONOTER: A tool for exploratory analysis (Research Rep. No. IRL90-0021). Palo Alto, CA: Institute for Research on Learning.
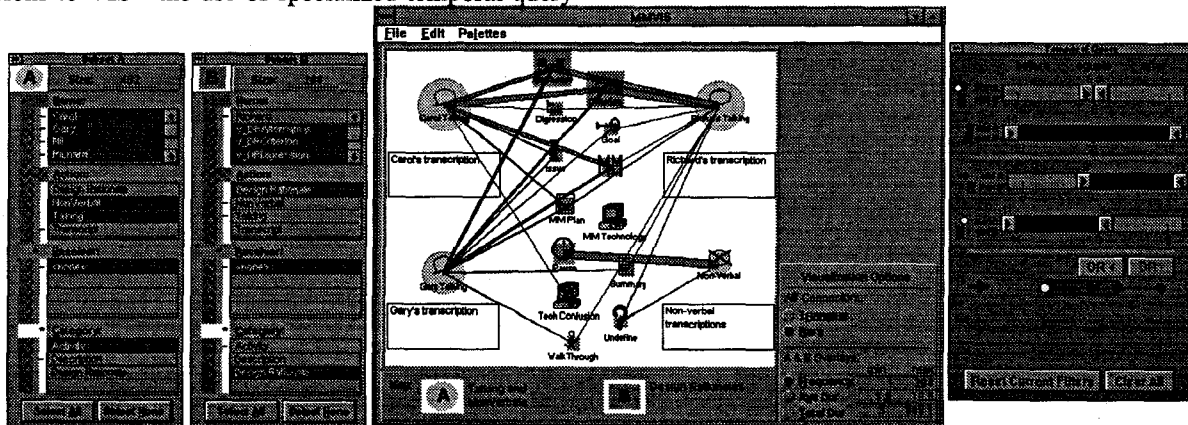
Figure 1. MMVIS Environment. Sample temporal analysis of CSCW video data on planning meetings.