

Extending and Evaluating Visual Information Seeking for Video Data

Stacie Hibino

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

ABSTRACT

Extending and adapting the visual information seeking paradigm for video analysis would empower casual users to *explore* temporal, spatial, and motion *relationships* between video objects and events. Several extensions are required to accomplish this: extensions to dynamic queries to specify *multiple* subsets, customized *temporal*, *spatial*, and *motion* query filters, and the design of new *spatio-temporal* visualizations to highlight these relationships. In my thesis research, I am working on these extensions by combining a new multimedia visual query language with spatio-temporal visualizations into an integrated MultiMedia Visual Information Seeking (MMVIS) environment. This research summary describes my overall approach, research goals, and evaluation plan.

Keywords

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

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]. If the VIS paradigm was extended and applied to video analysis, users would be empowered to explore various relationships (e.g., temporal relationships such as how often different types of events start or end at the same time) in a way that was not previously possible through other traditional means (e.g., timelines for temporal analysis) or other video analysis approaches (e.g., [3, 5]).

EXTENDING VIS FOR VIDEO DATA

I have identified several extensions to the original VIS framework that are necessary to adapt VIS for the analysis of video data.

These extensions include:

- mechanisms for selecting *multiple* subsets of different types of events,
- specialized *temporal*, *spatial*, and *motion query filters* for exploring the corresponding types of relationships between the subsets formed, and
- user-customizable *spatio-temporal* visualizations for highlighting, for example, the occurrence or frequency of the specified relationships.

In MMVIS, we provide subset query palettes (i.e., duplicate sets of query filters placed on palettes) for selecting multiple subsets. We have designed specialized temporal query filters [4], and have done some preliminary work on spatial and motion query filters. We have focused initial visualization work on temporal visualizations that cluster temporal relationships. The integrated MultiMedia Visual Information Seeking (MMVIS) environment currently supports the features listed above.

Scenario Applying MMVIS to CSCW Data

In order to better understand how MMVIS would work, consider the following scenario: HCI 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 (DR) of what is being said (e.g., to indicate when alternatives, digressions, etc., take place in the meeting). Researchers can use subset query palettes to select two subsets: A) talking and non-verbal events and B) DRs. They can then *explore* various relationships between members of these subsets using the specialized relationship query filters. Our temporal query filters form a temporal visual query language (TVQL) [4] and are presented to the user on a single palette (see Figure 1, Temporal Query palette). Keeping within the VIS paradigm, the visualization of results are dynamically updated as users specify the subsets as well as the temporal and/or spatial relationships. In Figure 1, TVQL specifies the relationship where A and B events start at the same time, but A’s end before or at the same time as B’s.

RESEARCH GOALS

The research goals of this thesis work are:

- to design a *multimedia visual query language* for temporal, spatial, and motion queries (via making significant extensions to dynamic query filters),
- to enhance the visual query environment with user-customizable *spatio-temporal* visualizations dynamically

© Copyright on this material is held by the author.

