Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Int. J. Human-Computer Studies 62 (2005) 127–157
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Evaluating event visualization: a usability study of COPLINK spatio-temporal visualizer

Wingyan Chung[a,*], Hsinchun Chen[b], Luis G. Chaboya[b], Christopher D. O'Toole[b], Homa Atabakhsh[b]

[a]*Department of Information and Decision Sciences, College of Business Administration, The University of Texas at El Paso, 500 W. University Avenue, El Paso, TX 79968, USA*

[b]*Artificial Intelligence Lab, Department of Management Information Systems, The University of Arizona, 1130 East Helen Street, McClelland Hall 430, Tucson, AZ 85721, USA*

Received 14 January 2004; received in revised form 18 June 2004; accepted 31 August 2004
Available online 4 November 2004

Abstract

Event visualization holds the promise of alleviating information overload in human analysis and numerous tools and techniques have been developed and evaluated. However, previous work does not specifically address either the coordination of event dimensions with the types of tasks involved or the way that visualizing different event dimensions can benefit human analysis. In this paper, we propose a taxonomy of event visualization and present a methodology for evaluating a coordinated event visualization tool called COPLINK Spatio-Temporal Visualizer (STV). The taxonomy encompasses various event dimensions, application domains, visualization metaphors, evaluation methods and performance measures. The evaluation methodology examines different event dimensions and different task types, thus juxtaposing two important elements of evaluating a tool. To achieve both internal and external validity, a laboratory experiment with students and a field study with crime analysis experts were conducted. Findings of our usability study show that STV could support crime analysis involving multiple, coordinated event dimensions as effectively as it could analyze individual, uncoordinated event dimensions. STV was significantly more effective and efficient than Microsoft Excel in performing coordinated tasks and was significantly more efficient in doing uncoordinated tasks related to temporal, spatial and aggregated information. Also, STV had compared favorably with Excel in completing uncoordinated tasks related to temporal

*Corresponding author. Tel.: +1 915 747 5496; fax: +1 915 747 5126.
E-mail address: wchung@utep.edu (Wingyan Chung).

doi:10.1016/j.ijhcs.2004.08.005

and spatial information, respectively. Subjects' comments showed STV to be intuitive, useful and preferable to existing crime analysis methods.

Keywords: Event visualization; Information visualization; Usability; Evaluation methodology; Taxonomy; COPLINK; Spatio-Temporal Visualizer; Crime analysis; Intelligence analysis; National security; Law enforcement

1. Introduction

In many real-world applications, information is primarily represented as events having time and space as their primary attributes (Jain, 2003a). Oftentimes, the number and complexity of events increase quickly, resulting in voluminous collections of data. For example crime incidents recorded in a police department's databases multiply quickly as they can happen at any time of a day, as do patients' visits to a hospital, customer orders in a restaurant, traffic accidents occurring on highways and scheduled tasks of a project. Consequently, it becomes increasingly difficult to analyze such a large amount of event data. Information overload and failure to indicate overall trends of events thus hinder further discovery of the potentially useful data. These problems elicit further inconvenience in certain domains such as law enforcement and intelligence analysis, where ability to gain insights from information and timely aggregating of information from multiple events are critical.

Information visualization holds the promise of amplifying human cognition in processing voluminous information (Card et al., 1999). In particular, event visualization uses techniques that facilitate understanding a large number of events. It has been applied to visualization of patients' medical records (Plaisant et al., 1999), highway accident reports (Fredrikson et al., 1999) and criminal incident records (Buetow et al., 2003). The benefits of event visualization are three-fold. The parallel processing capability of human vision reduces difficulties in understanding many dimensions of an event. Spatial and temporal data, representing the two most important event dimensions, can be portrayed vividly by event visualization. In addition, aggregation of event details can enhance discovery of trends and patterns. Event visualization assists human interaction with complex event data by leveraging human vision's power to handle multiple event dimensions concurrently. However, from a human–computer interaction (HCI) perspective, little work has been done to address the coordination of event dimensions with the types of tasks involved, thus leaving the question of how the two issues affect performance unanswered (Hewett et al., 1996). Moreover, how visualizing different event dimensions can benefit human analysis rarely has been examined in previous work.

In this paper, we propose a taxonomy of event visualization and present a methodology for evaluating event visualization tools and techniques. We report the results of evaluating the COPLINK Spatio-Temporal Visualizer (STV), an event visualization tool that integrates spatial, temporal and aggregated data to support coordinated visualization of crime events. At the local law enforcement level, STV can be used to summarize crime data, identify crime trends and reveal criminals' behavioral patterns. At the national security level, STV can provide valuable intelligence by helping to trace terrorists' activities. We believe that STV and results of its evaluation contribute to the field of intelligence and security informatics.

The rest of the paper is structured as follows. Section 2 surveys the current status of event visualization, with a focus on evaluation of event visualization tools. Based on this survey, a taxonomy of event visualization is presented. Section 3 presents research questions and their background. Section 4 presents the functionality of STV. Section 5 describes the methodology for evaluating STV. Section 6 reports and discusses the findings. Section 7 concludes the paper and discusses future directions.

2. Literature review

As a basic representation of human activities, an event has been defined as a significant occurrence or happening at a single point in space-time (Jain, 2003b). Event visualization can be defined as the visual representation of spatial, temporal and other dimensions of events. These events are usually part of human daily life and involve time and location. The following sections survey event visualization techniques and tools applied in different domains and the evaluation of these techniques and tools. To restrict the boundary of the survey, we only review those tools and techniques applied to visualizing events.

2.1. Event visualization techniques and tools

Temporal and spatial dimensions are the primary focus of many event visualization techniques and tools. We have organized our review according to event dimensions and the applications involved. Table 1a provides detailed descriptions of the tools and techniques.

2.1.1. Temporal visualization support

Visualizing temporal information is a common feature of many event visualization tools and techniques. Time is commonly represented by a horizontal line, which has been used by many earlier software systems and in paper-based event visualizations. We herein focus our review on computer-based event visualization systems developed in the recent decade. LifeLines summarizes medical records as a set of lines displayed on a zoomable TimeLine (Plaisant et al., 1999). Similar diagnoses are grouped on horizontal lines, each representing evolution of a group of events over time. AsbruView extends the work of LifeLines by incorporating temporal annotations and constraints (Kosara and Miksch, 2001b). Other event visualization tools for visualizing medical records include KNAVE (Shahar and Cheng, 2000), KHOSPAD (Combi et al., 1999) and timeline Browser (Cousins and Kahn, 1991). A related information architecture for visualizing personal histories has been proposed in Plaisant et al. (1998).

Table 1
A taxonomy of event visualization

(a) Event visualization tools and techniques

Dimension	Explanation/metrics	Code
Event dimensions		
Time-related	Day, date, minute, second, week, month, year, century, light-year	D1
Space-related	Location, place name, geographic information	D2
Content-related	Document details, event details	D3
Person-related	Persons' name, date of birth, address of residence, contact information, etc.	D4
Aggregated	Sum of event statistics, mean, standard deviation, periodic statistics (e.g. total occurrences on all Mondays)	D5

Domain	Representation of events	Code
Applications of event visualization		
Medicine	Patient records, clinical histories	A1
Law Enforcement	Highway accidents, crime incidents, criminal-justice records	A2
Business/news	Marketing opportunities, purchasing orders, distribution channels, news events	A3
Computer-mediated communication	Emails, chat room messages, newsgroup messages	A4
Environmental study	Complicated environmental phenomena	A5
Entertainment	Sports games, video shows	A6

Object	Meaning	Code
Metaphors used in event visualization		
Horizontal line/bar	Time, history	V1
Box	Event, incident	V2
Circle	Incident, milestone of time	V3
Map	Geographic region, location,	V4
Tree	Organizational chart, hierarchy	V5
Spike	Aggregate of information	V6
River	Time, history, thematic strength	V7
Spiral	Serial periodic data	V8
Glyph	Person, place, incident, message	V9
Flower	Person, message	V10

Techniques and tools	Descriptions	Applications	Event dimensions	Metaphors
Examples of tools and techniques				
LifeLines (Plaisant et al., 1999)	Provides a general visualization environment for personal histories, such as patients' clinical records. Allows users to access details of a patient's past records, zoom in or out of the timelines to view trends or details, high relationships by	A1	D1, D3	V1, V2

Table 1 (*Continued*)

Techniques and tools	Descriptions	Applications	Event dimensions	Metaphors
	query search, change the coding of display attributes, outline interested facets and summarize events			
AsbruView (Kosara and Miksch, 2001b)	Designed to fulfill the following requirements: representing relationships between temporal intervals, representing temporal uncertainty and unknown parts, specifying precision in time, hierarchical decomposition of event details, showing different facets of events	A1	D1, D3	V1, V2
KNAVE (Shahar and Cheng, 2000)	Enables dynamic examination of temporal information at multiple levels of abstraction and change in level of granularity	A1	D1	V1, V2
KHOSPAD (Combi et al., 1999)	Displays clinical information with history-oriented and relation-oriented views	A1	D1	V1
Time line browser [Cousins and Kahn, 1991]	Visualizes instant events (such as instantaneous clinical measurements) and intervals with duration on a timeline	A1	D1	V1
MMVIS (Hibino and Rudensteiner, 1998)	Answers the query of how often a set of events happen after another set of events	A6	D1, D3, D5	V1, V2, V9
ThemeRiver (Havre et al., 2002)	The horizontal distance between two points indicates the length of the time interval. The vertical distance, or width, of the river indicates the collective strength of the selected themes. It can show the history of an event collection	A3	D1, D5	V1, V7
TimeStore (Yiu et al., 1997)	Represents email messages and their arrival time as the vertical and horizontal dimensions of a task-date chart	A4	D1	V1, V3

Table 1 (*Continued*)

Techniques and tools	Descriptions	Applications	Event dimensions	Metaphors
TimeStore-Taskview (Gwizdka, 2002)	Has the same graphical representation as TimeStore but focuses on pending task management	A4	D1	V1, V3
Spiral graph (Carlis and Konstan, 1998; Weber et al., 2001)	Uses thickness of lines to represent the amount of data and different colors to represent different data types	A5, A6	D1, D5	V1, V6, V8
ESRI's ArcWeb (http://www.esri.com/)	Uses a GIS map view to display data, including street data, high-resolution imagery, weather data, topographic data and flood data	A3, A5	D2, D3	V3, V4
ViNeu (Kreuseler, 2000)	Supports both two-dimensional maps and three-dimensional landscape representations. Different methods for rendering multidimensional data within virtual 3D scenes are provided, in combination with several navigation techniques	A5	D1, D2	V1, V4
PANGAEA (Diepenbroek et al., 2002)	Analysis and visualization of metainformation and analytical data is supported by a number of software tools distributed as freeware from its Web site. Also includes mini-GIS PanMap, the plotting tool PanPlot and Ocean Data View (ODV) for the exploration of oceanographic data	A5	D2, D3	V3, V4
Perseus DL's map of top events (Smith, 2002)	Involves detecting and visualizing documents with date and place information. Collocations of date-place were displayed on time-lines and maps. More detailed displays could highlight key names and phrases associated with a selected event	A3	D1, D2, D3	V3, V4

Table 1 (*Continued*)

Techniques and tools	Descriptions	Applications	Event dimensions	Metaphors
Time manager (Andrienko et al., 2000)	A set of Java Applet enabled tools that provide animation of dynamic changes of events over geographical locations and time	A3, A5	D1, D2, D3	V1, V4
ReCAP (Brown, 1998)	Map-oriented searches are supported to provide a GIS display of the area along with plotted criminal activity. Patterns developed over time are plotted cyclically	A2	D1, D2	V1, V3, V4
CrimeStat (Levine, 2002)	Describes spatial distribution, provides distance and "hot spot" analysis and supports modeling of a variety of crime analyses (e.g., estimating the location of a serial offender, analyzing clustering in time and space)	A2	D2, D3, D5	V1, V3, V4
COPLINK STV (Buetow et al., 2003)	Integrates spatial, temporal and aggregated information on a single interface to facilitate the analysis and discovery of crime trends and patterns. Provides GIS, timeline and periodic pattern views that are synchronized	A2	D1, D2, D3, D5	V1, V2, V3, V4, V5, V8
Snap (Fredrikson et al., 1999; North and Shneiderman, 2000)	Based on the relational data model (Codd, 1970), Snap interconnects the visualization tools selected by users and constructs coordinated visualization interfaces for exploration of data and relationships	A2	D2, D3, D5	V1, V2, V3, V4, V5
Chat circles (Cherny, 1995)	Represents each chat room participant by a colored circle on the screen in which his or her words appear. The circles grow and brighten with each message and they fade and diminish in periods of silence and disappear when the participants disconnect from the chat room	A4	D1, D3, D4	V3

Table 1 (*Continued*)

Techniques and tools	Descriptions	Applications	Event dimensions	Metaphors
Loom (Donath and Robertson, 1994)	Supports visualizations of Usenet participants. The patterns and texture of the events within the group are reflected in the digital fabric of Loom	A4	D1, D3, D4, D5	V1, V3
PeopleGarden (Xiong and Donath, 1999)	Uses floral representation to visualize message archives. The number of petals of a thread-based flower equals the number of messages posted for that thread, while the number of leaves represents the number of participants in this thread. In addition, the height of a flower indicates how long the thread lasts	A4	D1, D4	V9, V10
CommunicationGarden (Zhu and Chen, 2002)	Has a similar representation to PeopleGarden. Uses a self-organizing map algorithm to categorize messages. Also represents each participant as a flower with the number of petals representing the degree of participation	A4	D1, D4, D5	V9, V10
EventViewer (Jain, 2003)	Displays events on a map, a timeline and an event table in a coordinated manner	A6	D1, D2, D3	**V1, V4**

(b) Evaluation methodology of event visualization

Methodology	Explanation	Code
Direct comparison	Involves comparing the tool or technique with a similar benchmark tool or technique	M1
Qualitative user study	Users are observed during the study and interviewed after using the tool or technique	M2
De-featuring approach	Tests individual visual elements by using different types of tasks	M3
Lab experiment	Involves systematic testing with subjects in a controlled environment	M4

Performance measure	Explanation	Code
Efficiency	Time used to perform a task	S1
Effectiveness	Proportion of correctly answered questions over all answered questions	S2

Table 1 (*Continued*)

Performance measure	Explanation	Code
Accuracy	Number of errors found	S3
User satisfaction rating	Users provide subjective scores (often a Likert scale is used) along various dimensions	S4
Verbal and written comments	Users provide verbal feedback after using the tool or technique	S5
Learning time	Time spent on learning the tool or technique	S6

Previous work	Descriptions	Methodologies	Measures
Examples of event visualization evaluation			
LifeLines (Alonso et al., 1998)	The study found that LifeLines could give a better overall summary of the records than the tabular representation. LifeLines achieved a higher efficiency while fewer errors were found in the tabular format. There was no significant difference in user satisfaction ratings	M1, M4	S1, S3, S4
AsbruView (Kosara and Miksch, 2001a, b)	Each participant used AsbruView to author a plan for their everyday work and had to fill out questionnaires before and after using the tool. The participants said that the AsbruView's metaphors and the way temporal uncertainty was handled were easy to understand and use. They also found the use of color helpful and liked the fact that they could change a plan's type at any time	M2	S5
ThemeRiver (Havre et al., 2002)	Two users participated in the qualitative study and answered questions about their understanding of the metaphor, their ability to identify themes, whether visualization helped raise new questions and differences between interpretations. Verbal protocol was captured and users were asked to complete a short questionnaire, eliciting feedback and possible enhancements to the system	M2	S5
Snap (North and Shneiderman, 2000)	In the first phase of the evaluation, six users were asked to construct coordinated visualization interface using Snap for S6 tasks on browsing population statistics of various States of the United States. The following were recorded: subjects' background information, learning time, whether they were successful or not in the task and time to completion of the task. In the second phase, eighteen subjects were asked to use three interfaces—detail	M4	S2, S4, S6

Table 1 (*Continued*)

Previous work	Descriptions	Methodologies	Measures
	only, no coordination, coordination—to perform a variety of tasks. Subjects' performance time and satisfaction were recorded		
TaskView (Gwizdka, 2002)	21 students participated in the study. Microsoft Outlook was used as a benchmark for comparison. Efficiency and effectiveness were measured	M1, M4	S1, S2, S4
CommunicationGarden (Zhu, 2002)	Five types of tasks (identify, cluster, compare, rank, correlate) were selected in the evaluation study. Effectiveness and efficiency were measured across different task types when compared against Netscape Messenger	M1, M3, M4	S1, S2, S4, S5

Temporal relationships often provide insights into the events involved. A MultiMedia Visual Information Seeking environment was developed to allow users to select two subsets of events and then browse and query to identify temporal relationships (Hibino and Rudensteiner, 1998).

Visualizing documents published over time can facilitate understanding of event trends. ThemeRiver depicts thematic variations within a large collection of documents over time (Havre et al., 2002). A river metaphor was used to visualize a document collection's timeline and respectively select thematic content and thematic strength as indicated by the river's directed flow, composition and changing width. However, ThemeRiver cannot visualize such important details as event locations, aggregate statistics and persons involved. This problem may be due to the choice of the river metaphor that limits the representations of some dimensions such as space.

Electronic mail (email) has become a major communication channel through which temporal events are recorded. TimeStore is an interface that uses time of arrival as the primary dimension for the display of emails (Yiu et al., 1997). Time is represented along the x-axis of a two-dimensional chart and message senders are listed along the y-axis and sorted in various ways. Similarly, TimeStore-TaskView (Gwizdka, 2002) uses the same graphical representation as TimeStore but was designed to manage pending tasks. Although events are sorted according to their occurrence time, the task-date table displays only a few sparsely distributed event points, consuming much screen space. Thus, the use table view may not be suitable to visualizing sparsely occurring events.

A spiral metaphor has been used to visualize various serial periodic event data. The Spiral Graph uses thickness of lines to represent data volume and different colors to represent different data types (Weber et al., 2001). It has been applied to visualizing sunshine intensity, in which the spiral effect enables easy periodic comparison of clouding periods or detection of sunrise and sunset. Both two- and three-dimensional displays can be used in spiral visualization, as demonstrated in applications to visualizing year–month chimpanzee food consumption, displaying periodicity of sound and showing seasonality of movie releases over time (Carlis and Konstan, 1998).

2.1.2. Spatial visualization support

Visualizing spatial data are another major feature of event visualization. Mapping events on a two-dimensional space is most commonly found because human vision often can process information on such a space efficiently. Established in 1969, ESRI (http://www.esri.com/) is one of the leading companies providing software solutions on geographic information systems (GIS) technology. ArcWeb USA, a major product of ESRI, is a comprehensive offering of nationwide data and services that include street data, high-resolution imagery, weather data, topographic data and flood data. ESRI's software helps organizations understand customer needs, analyze site locations, visualize and map demographic data and identify market trends. For example ArcView is a software used by law enforcement agencies to map crime incidents. Similar to ESRI, MapInfo (http://www.mapinfo.com/) helps businesses and governments analyze and derive insights from location-based information.

Mapping events is also useful for analysing environmental data. ViNeu is a system for visual analysis of complicated environmental phenomena in which contexts are given as spatial and temporal dependencies (Kreuseler, 2000). PANGAEA is an information system for processing, long-term storage and publication of georeferenced data related to earth science fields (Diepenbroek et al., 2002). A map view of oceanographic data is provided. In addition, GIS has been applied to modeling highway development (Jha et al., 2001), supporting a facility location decision (Noon and Hankins, 2001) and creating simulated images for understanding forest landscapes (McDonald and Stokes, 1998).

2.1.3. Multidimensional visualization support

Events involving more than one dimension frequently occur because humans often handle multiple event dimensions concurrently with their various senses (e.g., sight, sound, touch). These events require powerful visualization capabilities to display various dimensions (e.g., space, time, person, event aggregation) and to reduce information overload. A number of techniques and tools have been developed to support multidimensional event visualization.

2.1.3.1. Generic event visualizers. In the Perseus Digital Library Project, historical events with date and place information extracted from unstructured text have been detected and visualized (Smith, 2002). Because probabilistic techniques were used in identifying events, the approach is more applicable to visualizing historical events rather than emerging events, which are seldom well covered in published documents.

Time Manager supports interactive exploration of spatial data on the Internet over time (Andrienko et al., 2000). It has been integrated into several Web applications, such as viewing dynamic changes in forest areas, demonstrating movement of white storks in Africa, visualizing earthquake occurrences over time and showing time variations of thematic data (e.g., gross domestic product at market prices).

Snap-together visualization (Snap) enables users to dynamically mix and match visualizations and coordinations (North and Shneiderman, 2000). It has been applied to visualization of highway incident data in which temporal, geographical and categorical aggregations have been viewed (Fredrikson et al., 1999). Data points are aggregated to provide the benefits of summarization. Because Snap provides many visualization options for users to choose from (e.g., list, page, table, spotfire, outliner, treemap, etc.), it may overwhelm novice users and hinder effective event exploration.

EventViewer displays events on a map, a timeline and an event table and allows for searching (Jain, 2003b). It has been applied to visualizing inventory demand in different locations over time and to showing events happening during a football game. Users who want more details about a particular event can double-click any three display areas (what, where and when). The advantage is to integrate different dimensions of event details.

2.1.3.2. Domain-specific event visualizers. Several application domains often have demanded multi-dimensional visualization support. In the law enforcement and intelligence community, spatio-temporal visualization can help crime analysts and intelligence experts identify crime trends and related criminal activities effectively and efficiently. The Regional Crime Analysis Program is an operational environment for crime analysis (Brown, 1998) that provides a GIS display and plots data to show cyclical patterns. Also for crime analysis, CrimeStat is a spatial statistics program for the analysis of crime incident locations (Levine, 2002). It supports "hot spot" analysis and modeling of crime analysis. The COPLINK STV provides an integrated visualization environment that combines periodic, timeline and GIS views to allow simultaneous examination of the same data in three different views (Buetow et al., 2003).

In computer-mediated communication (CMC), knowledge sharing can be facilitated through the use of powerful event visualization systems. Chat Circles (Cherny, 1995) represents each chat room participant by a colored circle that changes with degree of user participation. Loom supports visualizations of Usenet participants and their interactions in a threaded newsgroup (Donath and Robertson, 1994). Both PeopleGarden (Xiong and Donath, 1999) and CommunicationGarden (Zhu and Chen, 2002) use floral representation to visualize archives of a CMC system. Most CMC tools enable visualization of large volumes of message archives created over time, thereby helping to reduce information overload. Aggregated information of the archives is vividly displayed in appealing glyphs (e.g., flowers in PeopleGarden and CommunicationGarden). However, detailed knowledge of the events (e.g., persons involved, geographic differences) cannot be gleaned from the visualization.

2.2. Evaluation of event visualization

Evaluation is an important step towards a better understanding of the usability of event visualization tools. This section describes previous work in evaluating these tools and techniques. Particular emphasis is placed on the evaluation methodologies and performance measures for coordinated visualization.

2.2.1. Evaluation methodologies

A commonly used methodology is to compare a tool or technique against a benchmark in a controlled laboratory environment. In an evaluation of LifeLines, a tabular format was used as a benchmark for comparison on viewing personal history records based on the speed, accuracy and user satisfaction ratings and recall data (Alonso et al., 1998). The results showed that LifeLines representation led to much faster response times, mainly for questions that involved interval comparisons and making inter-categorical connections. Since only a timeline metaphor was tested in the evaluation, it is unknown whether spatial and aggregated information of events would be useful. Also, a typical event involves many more dimensions than a personal record, thereby making evaluation of these dimensions more challenging.

In a TimeStore-TaskView (TaskView) email interface evaluation, TaskView was compared with Microsoft Outlook and was found to be more efficient for overview tasks but less efficient in tasks relating to message details (Gwizdka, 2002). The study mainly concerned time as the dimension, thus ignoring such possible dimensions in the email content as persons, places, event aggregates, etc.

A two-phase evaluation of Snap-together visualization (Snap) studied the aspects of coordinated visualization that caused improved performance and examined whether users could construct appropriate coordinated visualization (North and Shneiderman, 2000). In the first phase (construction), six subjects were asked to construct coordinated visualization interface using snap for tasks on browsing US states' population statistics. In the second phase (operation), eighteen subjects were asked to use three interfaces—detail only, no coordination, coordination—to perform a variety of browsing tasks (e.g., visual lookup, compare, search for target value). The study concluded that coordinated visualization was critical when access to details was needed, but was not necessary if only overview information was needed. When disjointed views were presented, subjects desired and expected coordinated visualization. The research points out the general benefits of coordinated visualization, but did not address whether coordinated visualization improves event visualization. It was not clear whether the tasks used in the second phase were actually those typically performed in analyzing event data and how different dimensions of an event affect performance was not studied.

A qualitative user study was often used to obtain user feedback. To assess the usefulness of AsbruView (Kosara and Miksch, 2001b), a qualitative user study was conducted with six physicians (Kosara and Miksch, 2001a). The participants found the visualization metaphor, including timelines and three-dimensional plan-level-time view, easy to understand. They also found the use of color helpful because it helped explain the functions involved (e.g., the use of gray for undefined components was understood by all subjects). However, the validity of conclusions could be enhanced with quantitative data (in addition to qualitative data) and more subjects. Moreover, as admitted by the authors, the experimental environment was not consistent for all participants, thereby further undermining the validity of conclusions.

In an evaluation of ThemeRiver (Havre et al., 2002), the results showed that two users had no difficulty in understanding the metaphor and could identify strongly represented themes. Questionnaire responses showed that ThemeRiver was useful for identifying macro trends but less useful for identifying minor trends. However, the validity of the results is highly questionable due to the small number of participants. It is not clear whether the tasks and questions used could cover all the tool's capabilities. Also, the evaluation mainly covered the temporal dimensions of document content but not other possible dimensions such as spatial and event details found in documents.

To study what role visualizations perform, a de-featuring approach was proposed and used to evaluate four information retrieval interfaces (Morse and Lewis, 2000). The approach uses a visual task taxonomy (Zhou and Feiner, 1998) that contains a large number of tasks performed by visualization tools. Examples of these tasks include Associate, Background, Categorize, Cluster, Compare, Correlate, Distinguish, Generalize, Identify, Locate, Rank, and Reveal. The de-featuring approach can be used exhaustively to test the capabilities of a visualization by mapping from the visual task taxonomy to a specific domain (information retrieval in Morse and Lewis, 2000). It has been used to create a social visualization tool known as CommunicationGarden (Zhu, 2002) that was found to outperform Netscape Messenger in terms of efficiency in all task types and in terms of effectiveness in "identify" tasks. The two had comparable effectiveness in "compare" and "correlate" tasks. The study points out the importance of distinguishing different task types using the visualization task taxonomy (Wehrend and Lewis, 1990; Zhou and Feiner, 1998; Morse and Lewis, 2000), especially for analysis purposes.

2.2.2. Performance measures

Various measures have been employed in evaluating event visualization tools and techniques. Efficiency (measured by time spent) has been a widely accepted measure (as used in Alonso et al., 1998; Gwizdka, 2002; Zhu, 2002) because visualization can help reduce the time needed to understand vast amount of information which typically places a high cognitive load on humans. Another commonly used measure is users' subjective comments or ratings (e.g. Kosara and Miksch, 2001b; Havre et al., 2002) that can directly reflect users' reactions. However, conclusions from subject comments may not be generalized to other situations. Effectiveness (often measured by precision and recall) and accuracy (measured by the correctness of task performance) were also used to study how the techniques or tools assist human work (e.g., Alonso et al., 1998; North and Shneiderman, 2000; Zhu, 2002).

2.3. A taxonomy of event visualization

Based on the above review, we have developed a domain-independent taxonomy of event visualization. Various event visualization tools and techniques can be classified in terms of event dimensions, application domains, and visualization metaphors used, as shown in Table 1a. Event dimensions are important because they serve to represent key information of the events (e.g., time, space). Application domains are related to human use of the tools or techniques. The domains listed in Table 1a only represent those appeared in our literature review as it is not possible to exhaust all possible domains. Visualization metaphors help users to understand the complicated event data by using visual clues. Evaluation methodologies and measures for evaluation visualization are summarized in Table 1b. It can be seen that much more work has previously been done on developing event visualization tools and techniques than on evaluating them.

The proposed taxonomy contributes to event visualization research in several aspects. While previous research has proposed taxonomies for diagram research (e.g., papers and taxonomies reviewed in Blackwell and Engelhardt, 2002), they generally deal with the representational aspects of visualization and have been criticized for lacking consideration of the context involved (e.g. social context, task and interaction) (Blackwell and Engelhardt, 2002). We attempt to mitigate this concern by considering in our taxonomy the event dimensions and application domains, which constitute the context of applying event visualization tools and techniques. Moreover, a taxonomy designed for classifying and studying event visualization research has not been found in previous research. Although there were efforts to classify visualization research by different means (e.g., a taxonomy of software visualization; Price et al., 1993), event visualization is not widely explored. Our proposed taxonomy tries to fill in this gap. In addition, our taxonomy has considered different types of evaluation methodologies used to study the usefulness of event visualization tools and techniques. Such consideration was not found in previous taxonomies.

3. Research questions

Previous work in evaluating event visualization tools and techniques has not specifically addressed the coordination of spatial, temporal, and aggregated information. Many studies tested with only a small number of participants, thus making the validity of their findings questionable. Comparing a tool with a selected benchmark has commonly been done but failed to reveal how coordination of different event dimensions affected performance in human analysis. Furthermore, as shown in Section 2.2.1, although the taxonomy of visualization evaluation tasks was shown to be valuable in evaluating general graphical interface (Wehrend and Lewis, 1990; Zhou and Feiner, 1998; Morse and Lewis, 2000), it surprisingly has not been widely applied to evaluating event visualization tools. In this research, we try to answer the following questions:

1. How can an event visualization tool be evaluated to study the coordination of spatial, temporal, and aggregated information?
2. How does coordination among the three dimensions (i.e., spatial, temporal and aggregated information) affect the performance of the event visualization tool?
3. What is the usability (measured by effectiveness and efficiency of task performance) of the event visualization tool for facilitating human analysis?

4. Research testbed

Our research testbed, the COPLINK STV, is a coordinated event visualization tool for crime analysis. A previous version of STV has been presented in Buetow et al. (2003). In this section, we briefly highlight the major functions and describe changes made to STV.

4.1. The COPLINK spatio-temporal visualizer (STV)

STV supports visualization of crime incident data provided by COPLINK, a software system that runs on top of the Tucson Police Department's (TPD) current databases and records management system. The award-winning COPLINK project (http://ai.bpa.arizona.edu/COPLINK) has been conducted by researchers at the University of Arizona, in collaboration with the TPD and the Phoenix Police Department since 1997.

STV provides three views to support effective and efficient discovery of spatial, temporal and periodic patterns (see Fig. 1). The GIS view displays a map of Tucson to help locate geographic clusters of crime incidents. The timeline view on a chart shows crime incidents as square boxes arranged in chronological order, with groups of incidents displayed at a hierarchy on the left of the chart. The periodic pattern view provides aggregated information of a collection of incidents. A circular chart is used to display how many incidents occurred over a specified period in a selected time unit.

Located at the lower part of Fig. 1, the control panel allows users to maintain central control over the three views. It was modified from the previous version to display clearly the start and end times of the selected global and local times and to include an entity information box. Global time bounds of the displayed incidents are controlled through a series of drop-down menus. Zoom-in timeframes can be adjusted by moving the zoom time slider. Detailed information (such as crime type, address, date and time) of an incident selected in the timeline view is displayed in the entity information box located on the right of the zoom time slider.

4.2. A crime analysis example

To illustrate STV functionality, we explored a scenario in which a crime analyst had been assigned to the task of examining bank robbery data between October and December of 2001 (Fig. 1). The scenario is hypothetical because we used data

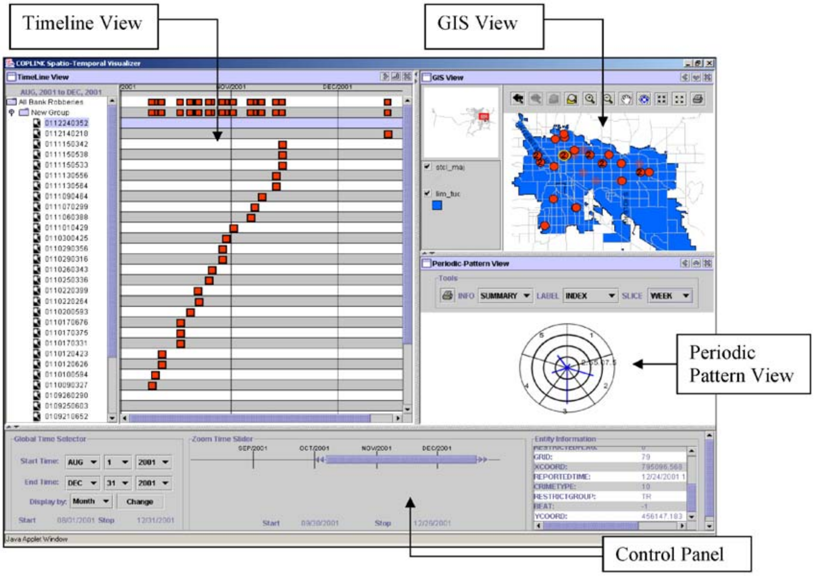

Fig. 1. Visualizing bank robbery incidents (September–December, 2001).

scrubbed from real TPD data, which we were not allowed to disclose to the public. Nevertheless, these scrubbed data had the same properties of real data for analysis purpose. By looking at the timeline view, the analyst immediately saw a significant drop-off in bank robbery incidents after the Thanksgiving holiday (November 27). On the GIS view, he noticed a cluster of robberies in the North-west side of the city. Inspecting the periodic pattern view revealed that most of the incidents took place around the second or third week of the month. Further analysis showed that an area around the intersection of Euclid Avenue and Grant Road (north-west side of Tucson) appeared to be the center of a concentration of activity.

For the crime analyst, STV raised many questions: Why was there the sudden disappearance of robberies after Thanksgiving? Why was the first week of each month almost devoid of robberies? Why were so many banks in the same area hit at the same time? Such provocative questions might have been hidden if other methods such as an SQL search had been used, because the results might not reveal any easily observed patterns. The spatial and temporal crime trends suggested a need for an increase in patrols in areas where many incidents of bank robbery had occurred, particularly within time periods of apparently increased activity.

5. Evaluation methodology

This section describes our methodology in evaluating STV, thereby answering Research Question 1 stated in Section 3. We explain below the experimental design, hypotheses and experimental tasks and procedure.

5.1. Experimental design

We employed a de-featuring approach (Morse and Lewis, 2000) in our experimental design because it can be tailored to a specific domain (crime analysis in law enforcement). From our literature review, we do not find previous work that studies coordination among different event dimensions in evaluating event visualization tools. Thus, in our experimental design, we considered the spatial, temporal, and event aggregate dimensions, as well as combinations of them when applying the visual task taxonomy (Wehrend and Lewis, 1990; Zhou and Feiner, 1998; Morse and Lewis, 2000) to designing tasks.

Based on our interview with a TPD crime analyst and a criminal intelligence officer (CIO) about their daily work and frequently used tools, we adopted the following types of tasks in our task design: (1) Identify—find an event with certain features; (2) Compare—tell the differences between events/sets of events; (3) Rank—arrange events in a certain order to show the best or the worst cases; and (4) Cluster—group events according to certain attributes. Other types of tasks, although performed occasionally, were not selected because the above-mentioned four are most common and more than four task types would likely lengthen the experiment undesirably.

The target users of STV are TPD crime analysts and CIOs. CIOs work in TPD's divisions (e.g., north-west Tucson, downtown, etc.). Their daily work includes analyzing crime incident reports, mapping crime hot spots, producing information flyers for police officers to provide them with criminal mug shots and crime details. The timeframe of analysis ranges from a few days to four weeks. Specializing in crime types (e.g., aggravated assaults, gang crime, narcotics crime, sex crime), crime analysts work in the TPD's main station, where they analyze overall crime trends in Tucson, produce crime-specific periodic (monthly, annual) reports, maintain close communication with detectives and senior TPD officials and suggest strategies for combating crimes. Their timeframe of analysis ranges from 1 to 6 months.

To achieve both internal and external validity, we designed a controlled laboratory experiment and a field study to evaluate STV.

5.1.1. Laboratory experiment

A laboratory experiment helped us to control the experimental factors easily and to ensure that results would be more generalizable and have high internal validity. Based on an interview of the daily work of crime analysts and CIOs, we chose Microsoft Excel (a tabular format) as a benchmark for comparison; it also had been used for browsing event information in a previous study (Alonso et al., 1998). Crime analysts and CIOs typically use Excel to generate periodic reports, calculate statistics, and identify trends.

Thirty University of Arizona students, who did not have any relationships with the authors prior to the experiment, participated as subjects to evaluate and compare STV and Excel. During the one-hour experiment, each subject was trained to use each of the two systems and was required to know basic computer operations (e.g. using a mouse, typing with a PC keyboard). The order of using the two systems was randomly assigned to avoid bias due to sequence of use. In addition to providing a spreadsheet of crime incident data, we provided Excel users with a physical map of Tucson to help them easily map the X and Y coordinates from the tables to the map. Excel users were also taught to use simple analysis functions such as filtering and sorting and were allowed to use all other Excel features. After using each system, a subject was asked to fill in a post-section questionnaire (designed based on Davis, 1989; Lewis, 1995) and to compare the two systems.

5.1.2. Field study

A comprehensive field study involving task performance, questionnaire survey and interview was conducted at the TPD. Four crime analysts and two CIOs participated as subjects. Each of them used STV to perform the tasks, filled in a questionnaire comparing STV and existing crime analysis methods and then was asked to provide detailed comments.

5.2. Experimental tasks

Scenario-based crime analysis tasks having definite answers were used in both laboratory experiment and field study. For example a scenario was "All crime incidents of residential robbery from May 1, 2001 to September 30, 2001" and a task was "Over the entire period, rank the seven weekdays in descending order of the total number of incidents occurring (from the most to the fewest)." The suitability of all the tasks was validated by a veteran TPD detective who has served in law enforcement for more than 30 years.

To assess the effect of coordination of event dimensions (temporal, spatial, and aggregated information), we divided the tasks into two sets. Each set contained all four selected types (identify, compare, rank, cluster). In the first set (Set A, uncoordinated tasks), subjects could find answers using only one of the three views of STV. For example the task "Identify the name of a road that is the closest to the incident that, among all displayed incidents, happened at the most southeast side of Tucson." is an example of "identify" task and is uncoordinated because it only asks for spatial information. In the second set (Set B, coordinated tasks), subjects were expected to use a combination (and coordination) of more than one view of STV to find the answers. For example the task "Focus on the time period between August 1 and August 31. Identify incidents occurring north of Speedway. Among all the incidents selected, which two incidents occurred closest in time to each other?" is a "cluster" task and is coordinated because it asks for both spatial and temporal information.

Eight tasks were included in each scenario and their order as well as their types and natures are summarized in Table 2. Three scenarios were designed and randomly assigned to different systems in the laboratory experiment to avoid bias due to sequence of use.

No definite time limit was imposed on each task but we timed the performance of each task performed by each subject. Typically a subject could finish an uncoordinated task in about one minute and a coordinated task in three minutes.

Table 2
Nature and types of experimental tasks

Task types	Uncoordinated task (set A)	Uncoordinated task (set B)
Identify	Task 1	Task 6
Rank	Task 2	Task 7
Compare	Task 3	Task 5
Cluster	Task 4	Task 8

The performance was measured by effectiveness (accuracy of answer) and efficiency (time to completion). As each subject was asked to perform similar tasks using the two systems, a one-factor repeated-measures design was used, because it gives greater precision than designs that employ only between-subjects factors (Myers and Well, 1995). All verbal comments were analysed using protocol analysis (Ericsson and Simon, 1993).

5.3. Hypothesis testing

In developing our hypotheses, we tried to demonstrate how to evaluate an event visualization tool considering the different dimensions and task types involved, which has not been studied in previous work. We therefore compared the following four settings to answer Research Question 2 stated in Section 3:

1. A tabular format (Excel) used for uncoordinated tasks (baseline measure).
2. A tabular format (Excel) used for coordinated tasks.
3. STV used for uncoordinated tasks (Set A).
4. STV used for coordinated tasks (Set B).

Because the tasks in Set B are more complicated than those in Set A, we expected subjects to achieve different performance levels using different systems. Three groups of hypotheses were tested as follows.

H1: The effectiveness of performing coordinated tasks using STV is not significantly different from that of performing uncoordinated tasks using STV.

H2: The efficiency of performing coordinated tasks using STV is lower than that of performing uncoordinated tasks using STV.

H3: For coordinated tasks, STV is more efficient and effective than a tabular format.

H4: For uncoordinated tasks, STV is more efficient and effective than a tabular format.

H5: STV achieves better user ratings than a tabular format in the lab experiment.

H6: STV is more effective and efficient than a tabular format for performing different task types.

H7: STV is rated by crime analysts and CIOs as more effective and efficient than other methods.

The first group includes H1 and H2, which relate to STV's performance on uncoordinated and coordinated tasks. The rationale behind H1 is that, because STV's coordinated visualization of the three dimensions could help users understand a large number of crime events in a short time, it would accomplish the more complicated coordinated tasks as effectively as the less complicated uncoordinated tasks. The rationale behind H2 is that users needed more time to understand the questions and find answers for coordinated tasks than for uncoordinated tasks.

The second group of hypotheses includes H3–H6, which relate to comparison between STV and a tabular format (MS Excel). In H3, we believed that STV's coordinated views would enable it to perform more effectively in the more complicated coordinated tasks that tabular format did less well and STV would save users' time in gaining understanding of events because of the benefits of visualization. Hypothesis H4 was related to the effectiveness and efficiency of using the two systems to perform uncoordinated tasks involving different event dimensions such as time, space and aggregation of information. We believed that STV's visualization capabilities would enable it to outperform a tabular format in uncoordinated tasks.

Owing to the above-mentioned advantages of STV, we believed that STV would achieve better user ratings in usefulness, ease of use and information display and interface design (H5). Also, across different task types we selected, STV would perform better because the various functions would enable users to identify, rank, compare and cluster event information more easily (H6).

The third group includes only H7, which relates to experts' ratings on STV as compared with their existing methods for crime analysis. Because STV provides an appealing interface and integrates different event dimensions, we believed that it would be rated better than the existing methods.

To test each of the hypotheses H1–H6, we used a pairwise *t*-test with a degree of freedom equal to 58. The test compared the mean performance values of using the two systems to study whether their performances were statistically different. To test H7, we asked the experts to tell which systems they preferred to use to perform the tasks. The test values are stated in Section 6.

6. Experimental results and discussions

This section reports and discusses the findings of our study. Table 3 details the results of hypothesis testing. Table 4 lists subjects' profiles. Table 5 summarizes results of preferences of students and TPD crime analysts and CIOs. Tables 6 and 7 provide summaries of student and expert subjects' verbal comments.

6.1. STV's performance on uncoordinated and coordinated tasks

Hypothesis H1 was confirmed. With the help of STV, subjects could easily understand how different event dimensions were coordinated. They could visualize the dimensions involved in coordinated tasks, thereby achieving effectiveness

Table 3
Results of hypothesis testing

Hypothesis	Uncoordinated tasks		Coordinated tasks		p-value	Result[a]
	Mean	S.D.	Mean	S.D.		
H1 (effectiveness[b])	0.82	0.15	0.9	0.17	0.065	Confirmed
H2 (efficiency[c])	51.47	14.5	111.15	24.92	0	Confirmed
Hypothesis	STV		MS Excel		p-value	Result
	Mean	S.D.	Mean	S.D.		
H3						Confirmed
Effectiveness	0.86	0.12	0.74	0.14	0.001	STV is better
Efficiency	81.31	17.13	131.68	23.20	0.000	STV is better
H4						Partially confirmed
Effectiveness-temporal information	0.72	0.36	0.73	0.31	0.861	No difference
Efficiency-temporal information	55.73	23.72	83.50	37.22	0.001	STV is better
Effectiveness-aggregated information	0.88	0.22	0.65	0.42	0.008	STV is better
Efficiency-aggregated information	40.97	23.99	143.17	49.25	0.000	STV is better
Effectiveness-spatial information	0.84	0.21	0.79	0.21	0.326	No difference
Efficiency-spatial information	54.36	13.99	80.92	31.87	0.000	STV is better
H5						Confirmed
Usefulness[d]	1.71	1.01	3.43	1.72	0.000	STV is better
Ease of use[d]	1.85	0.82	3.03	1.46	0.001	STV is better
Interface[d]	1.84	0.95	3.63	1.53	0.000	STV is better
H6						Partially confirmed
Effectiveness-identify	0.88	0.14	0.78	0.20	0.031	STV is better
Effectiveness-rank	0.87	0.22	0.67	0.27	0.003	STV is better
Effectiveness-compare	0.84	0.19	0.84	0.17	1.000	No difference
Effectiveness-cluster	0.86	0.20	0.68	0.22	0.000	STV is better
Efficiency-identify	53.28	18.92	108.63	36.75	0.000	STV is better
Efficiency-rank	95.05	36.71	179.78	42.58	0.000	STV is better
Efficiency-compare	70.22	15.44	114.77	32.68	0.000	STV is better
Efficiency-cluster	106.68	25.64	123.53	33.53	0.016	STV is better
H7	Number of tasks[e] in which experts prefer					Confirmed
	STV		Others	No preference		
Effectiveness-uncoordinated tasks	16		4	4		STV is better
Effectiveness-coordinated tasks	23		0	1		STV is better
Efficiency-uncoordinated tasks	22		1	1		STV is better
Efficiency-coordinated tasks	22		1	1		STV is better

[a]An alpha error of 5% was used.
[b]Effectiveness ranges from 0 to 1.
[c]Efficiency was measured by the time used (in seconds).
[d]Usefulness, ease of use, and interface design were rated by subjects on a 7-point Likert Scale, with 1 being the best.
[e]Total number of tasks considered in each hypothesis = 6 (experts) $\times$ 4 (tasks) = 24.

Table 4
A summary of subjects' profile

Subjects	Profile
Experts	Four crime analysts, 2 criminal intelligence officers (3 males, 3 females); Average experience in TPD: 12.8 years; Average experience in crime analysis work: 4.83 years; Specializations: gang crimes, auto theft, burglary, aggravated assaults, armed robberies, gun tracing and other crimes
Students	Gender: 15 males, 15 females; Education: 21 undergraduate students, 3 with associate degrees, 4 with bachelor's degrees, 2 with master's degrees

Table 5
Subjects' preferences

Dimension	Number of students who preferred[a]		Number of crime analysts and CIOs[b] whose first preference was		
	STV	Excel	STV	ArcView	Others[c]
To identify a crime incident with geographical information only, I would use	29	1	5	1	0
To identify a crime incident with information about time only, I would use	18	12	4	2	1
To find crime patterns and trends over time, I would use	28	2	6	1	0
To find crime patterns and trends over a geographic area, I would use	28	2	4	3	0
To do statistical computation on crime information, I would use	15	15	4	1	1
To generate crime summary reports, I would use	23	7	4	1	0
To rank a set of crime incidents according to a given crime attribute (e.g., incident reported time), I would use	21	9	4	2	0
To compare two crime incidents based on a given crime attribute (e.g., incident location), I would use	24	6	4	1	1
To analyze criminal behaviors over time, I would use	29	1	3	1	1
To analyze criminal behaviors over geographic locations, I would use	28	2	4	2	0
To analyze criminal behaviors over geographical locations and a period of time, I would use	29	1	6	0	0

[a]The total number of student subjects is 30.
[b]Four crime analysts and two CIOs participated in the study.
[c]Other tools for crime analysis included RMS, MS Access, MS Excel, SQL, case reports, and COPLINK.

Table 6
A summary of students' comments

Category	Comments and number of subjects having the comments			
	STV	No.	Excel	No.
Positive comments	User-friendly, easy to use	12	Familiarity	4
	Visually appealing, visualization helps view data easily	17	Efficiently sort, filter and organize data	9
	Very impressive, great system	4	Easy to use the system	9
Negative comments	Difficult to use time slider	4	Difficult to find locations on map	7
	Hard to see details in entity information window	3	Straining on eyes to view a lot of data	5
	No double-click function to view event details	3	Easy to make mistakes	5
	Numbers on periodic view are too small	2	Interface is very unpleasant	4
	GIS highlighting is hard to see	2		
Improvements	Allow typing in streets location to find incidents	3	Some way to install the map and coordinates	5
	Clearer numbering on periodic pattern chart	2	Make the system more user-friendly	4
	Help function	4	Enhance the graphical user interface	5
	Statistical analysis functions	3		

comparable to that in uncoordinated tasks. H2 was also confirmed because uncoordinated tasks required less time and coordinated tasks often were broken up into several subtasks that took up more time. We conclude from the results that STV could support crime analysis involving multiple, coordinated event dimensions as effectively as it could in analysis involving individual, uncoordinated event dimensions, although the tasks in the former category took more time.

6.2. Comparing between STV and a tabular format

6.2.1. Performance in coordinated and uncoordinated tasks

Hypothesis H3 was confirmed, showing that STV's powerful coordinated visualization enables more effective and efficient performance than tabular format. Subjects made fewer mistakes in STV than in Excel and noticed their mistakes more easily in STV. Moreover, subjects had trouble deciding which filters to use in Excel where the only textual information presented was in table cells with minimal visual cues.

H4 was partially confirmed. While both systems displayed temporal information to the same degree of detail and accuracy, STV's timeline visualization actually

Table 7
A summary of crime analysts' and CIOs' comments

Category	Comments and number of subjects having the comments			
	STV	No.	Excel	No.
Positive comments	Very useful, very easy to learn	3	ArcView has nice-looking map	2
	Integrate different event dimensions	2	Familiarity with existing methods	1
	Great to see immediate results	3		
Negative comments	The entity window should be wider and be able to show details without having to scroll and some attributes should be removed	3	Having to flip through different systems (RMS, SQL, Access, ArcView) to get one result	2
	Map should have labeling functions	1	ArcView is time-consuming, not user-friendly and not reliable	2
	Zoom time slider is too sensitive to move back and forth on time line	1		
Improvements	Tell hot spots	1	Training is needed	1
	Click on an incidents and bring the information of incidents	1		
	Clicking on a point and displaying details of the incident	1		

enabled faster performance than Excel's functions (e.g., filter, sort) that required many keystrokes. Thus STV was found to have outperformed Excel in efficiency in uncoordinated tasks relating to temporal information. STV's periodic pattern view allowed subjects to obtain aggregated information very quickly and easily, while Excel had a more complex procedure for aggregating information and required more setup and computing times.

A closer look into subjects' actions revealed that when subjects used Excel to perform tasks relating to spatial information, they made use of the physical map we provided that had important details of street names and regional information. Such details were not provided in STV's GIS view due to limitation in screen size and a lack of detailed data. Also, due to unfamiliarity with STV, few subjects realized that they could resize the GIS's map to obtain greater accuracy, thus further reducing the effectiveness. On the other hand, STV allowed subjects to easily select a rectangle containing the desired incidents and filter out undesired ones. In contrast, subjects spent more time filtering and sorting the X and Y coordinates in Excel and in mapping them onto a physical map than they used in STV. Thus STV was significantly more efficient than Excel in performing tasks relating to spatial information.

Based on the above findings, we conclude that STV was significantly more effective and efficient than Excel in coordinated tasks and was significantly more efficient in uncoordinated tasks relating to temporal, spatial and aggregated information. Also, STV had effectiveness comparable to that of Excel in uncoordinated tasks relating to temporal and spatial information.

6.2.2. System usability

H5 was confirmed, demonstrating STV's superior usability. Even though we allowed subjects to use all advanced functions of Excel, subjects still rated STV to be more useful because Excel did not support certain important crime analysis functions (e.g., crime mapping, trend identification) as well as STV. STV also provided a cleaner interface than Excel, which offered many confusing options. In addition, subjects preferred the visual display of STV to Excel's tabular format. Therefore, we concluded that STV achieved significantly better usability ratings than Excel in terms of usefulness, ease of use and information display and interface design.

6.2.3. Performance on different types of tasks

Hypothesis H6 was partially confirmed, showing that STV enabled effective and efficient performance in almost all types of tasks. The tabular format provided by Excel actually hampered subjects' identification of correct items because of information overload. The need to use filter and sort functions in Excel to cluster items also added an extra burden. In contrast, STV almost automated the ranking of event information and hence facilitated "rank" tasks. Visual effects provided by STV greatly enhanced the efficiency of performing all the four types of tasks. However, STV did not allow for highlighting multiple incidents happening at one time—an important function to support "compare" tasks; Excel did. This explains why we found no significant difference in performing "compare" tasks. From the results, we concluded that in comparison with Excel, STV was more efficient in all four task types, more effective in "identify," "rank," and "cluster" tasks and had comparable effectiveness in "compare" tasks.

6.3. Comparing between STV and existing crime analysis methods

In the field study, the experts were asked to compare STV with their existing methods for crime analysis and indicate their preferences in terms of effectiveness and efficiency. The results of evaluating H7 showed an overwhelming favor towards STV, as shown in Table 3. In these promising results, relatively fewer (though still over 75%) experts agreed that STV was more effective in uncoordinated tasks than their existing methods.

6.4. Subjects' preferences and comments

6.4.1. Effects of coordinated visualization

The students found STV to be much more useful to complete their tasks than Excel. A main reason was the power of STV's coordinated visualization; a large number (17) of subjects with this comment as shown on Table 6. For example student #8 said: "The visual effects really helped me understand and search the information easily," student #19 said: "This system is amazing. You can see when and what time the incident happened exactly." In contrast, students mentioned that Excel strained their eyes. For example student #6 said: "It's very straining on the eyes to look up data (that are) so compact and similar." On the contrary, STV allowed students simply to look at points on an interactive map. This might explain why many students using Excel complained about the difficulty of finding map locations on the physical map.

The TPD crime analysts and CIOs were very excited about being able to coordinate such different event dimensions as time and space at the same time (see Table 7). For example expert #1 said: "Zoom time slider function is great and fast! I love the map and corresponding timeline view. Very helpful to know what you're looking at." As STV allowed users to immediately see the results of their actions (such as moving the time slider), it gave a powerful advantage over other systems. Despite these encouraging responses, there might be biases arising from different backgrounds of experts. We believe that this problem can be alleviated when more experts participate in a future study in which STV has been deployed in the TPD.

6.4.2. User friendliness

The students found STV to be easier to use than Excel because STV did not hide any of its functionality in hard-to-remember locations while Excel provided many confusing options (see Table 6). For example student #18 noted that STV was "very visual, everything seems tangible, better performance, map is already included, graphical capabilities are easily accessible and visual," while student #17 said that in Excel, it was "hard to follow the steps to use it." Although some found Excel also to be easy to use, there were complaints about Excel's unpleasant and misleading interface. Even after training, students were confused as to which function would be best used for a task at hand.

The experts found STV to be more intuitive and useful than their existing methods. As expert #3 said: "It (STV) is very useful because you don't need to use different tools to come up with the results." Expert #4 deemed STV to be "very useful…gives you everything in one step." The consolidated functionality of STV allowed subjects to complete tasks in a much more efficient manner than before. STV's superior performance was remarkable considering that the experts had had much more training in using their standard set of tools (see Table 7). They found having to flip between standard systems to be tedious and time consuming.

6.4.3. Improvements needed

We received valuable input from students on improvements needed for STV. To speed up the search, they wanted to be able to find a street on the map simply by typing the street name instead of searching the map manually. This feature would be great for finding an area with which the user might be unfamiliar. It also was suggested that the numbers in the periodic pattern view should be enlarged.

The experts suggested that STV should identify hot spots on the map, thereby allowing them to easily find areas of abnormal activity. One expert mentioned that it would be useful for the GIS view to open up entity information when the user double-clicked on an incident. This would allow the user to quickly retrieve case details.

From the experts' and students' comments, we have concluded that STV's coordinated visualization capability assisted human analysis and that STV was very user-friendly and performed better than existing crime analysis methods.

7. Conclusions and future directions

In this paper, we have proposed a domain-independent taxonomy of event visualization. To close gaps in previous research, we have developed a methodology for evaluating event visualization tools and techniques and have applied it to evaluating the COPLINK STV, a coordinated event visualization tool for visualizing crime incidents. Findings from our study are very encouraging, as shown in the superior performance of STV in many different types of tasks and in coordination of different event dimensions. STV was found to be intuitive and useful and was perceived more favorably than existing crime analysis methods. Through the study, we believe a better understanding in HCI issues (such as how coordinating event dimensions affects performance and how different task types can be analysed) has been achieved.

Given the encouraging findings, a promising future direction is to extend STV into a fully functional crime analysis system. Immediately after the field study, we received queries from TPD's crime analysts and CIOs about plans to deploy the tool. With more experts using the tool, we would be able to conduct a larger scale field study in order to overcome the disadvantages of a limited number of expert participants and restricted generalizability of findings.

Another interesting future direction would be to extend the applicability of STV to other domains requiring analysis of spatio-temporal events. Based on the current STV system, we are creating a generic STV tool that can be applied to any domain. One potential application is the analysis of terrorism incidents as the concern of national security greatly increases following the terrorist attacks. Tools and techniques enabling deeper analysis and visualization are needed (Chen et al., 2004). The distributed and unusual nature of terrorism incidents provides a good domain for applying STV to visualizing terrorist activities. Another interesting domain is business intelligence (BI), where distributed business stakeholders in a multilingual world interact at different times (Chung et al., 2004; Chung, 2005). Visualization of BI using STV is expected to improve analysis capability over previous efforts (e.g., Chung et al., 2003).

In terms of HCI issues, more visualization metaphors can be developed to illustrate event dimensions more effectively. A challenging future direction is to develop a model that can be used to accurately predict performance of event visualization tools, based on limited information or before the tools are at full function. Achieving such a goal would greatly facilitate design and development of tools and is likely to benefit other areas of HCI.

Acknowledgments

This project has been partly supported by NSF Digital Government Program, "COPLINK Center: Information and Knowledge Management for Law Enforcement," #9983304, July 2000–June 2003 and NSF Information Technology Research, "Developing A Collaborative Information and Knowledge Management Infrastructure," NSF/IIS #0114011, September 2001–August 2004. We thank Detective Tim Petersen, Ty Buetow, Tom Cushna, Damien Daspit, Lt. Jenny Schroeder, Dan Casey, Karin D. Quiñones, all the student subjects and TPD crime analysts and criminal intelligence officers who participated in the user study for their support and assistance in this research.

References

Alonso, D.L., Rose, A., Plaisant, C., Norman, K.L., 1998. Viewing personal history records: a comparison of tabular format and graphical presentation using LifeLines. Behaviour & Information Technology 17, 249–262.

Andrienko, N., Andrienko, G., Gatalsky, P., 2000. Visualization of spatio-temporal information in the Internet. In: Proceedings of the 11th International Workshop on Database and Expert Systems Applications. IEEE Computer Society, Greenwich, London, UK, pp. 577–585.

Blackwell, A.F., Engelhardt, Y., 2002. A meta-taxonomy for diagram research. In: Anderson, M., Meyer, B., Olivier, P. (Eds.), Diagrammatic Representation and Reasoning. Springer, Berlin, pp. 47–64.

Brown, D.E., 1998. The regional crime analysis program (RECAP): a framework for mining data to catch criminals. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, San Diego, CA, USA, pp. 2848–2853.

Buetow, T., Chaboya, L., O'Toole, C., Cushna, T., Daspit, D., Petersen, T., Atabakhsh, H., Chen, H., 2003. A spatio-temporal visualizer for law enforcement. In: Proceedings of the First NSF/NIJ Symposium on Intelligence and Security Informatics (ISI 2003). Springer, Heidelberg, Tucson, AZ, USA, pp. 181–194.

Card, S., Mackinlay, J., Shneiderman, B. (Eds.), 1999, Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann Publishers, San Francisco, CA.

Carlis, J.V., Konstan, J.A., 1998. Interactive visualization of serial periodic data. In: Proceedings of the Eleventh ACM Annual Symposium on User Interface Software and Technology. ACM Press, San Francisco, CA, USA, pp. 29–38.

Chen, H., Chung, W., Xu, J.J., Wang, G., Chau, M., Qin, Y., 2004. Crime data mining: a general framework and some examples. IEEE Computer 37 (4), 50–56.

Cherny, L., 1995. The MUD register: conversational modes of action in a text-based virtual reality. Ph.D. Dissertation, Department of Linguistics, Stanford University.

Chung, W., Chen, H., Nunamaker, J.F., 2003. Business intelligence explorer: a knowledge map framework for discovering business intelligence on the Web. In: Proceedings of the 36th Hawaii International Conference on System Sciences. IEEE Computer Society, HI, USA.

Chung, W., Zhang, Y., Huang, Z., Wang, G., Ong, T.H., Chen, H., 2004. Internet searching and browsing in a multilingual world: an experiment on the Chinese Business Intelligence Portal (CBizPort). Journal of the American Society for Information Science and Technology 55 (9), 818–831.

Combi, C., Portoni, L., Pinciroli, F., 1999. Visualizing temporal clinical data on the WWW. In: Proceedings of the Joint Conference on Artificial Intelligence in Medicine and Medical Decision Making. Springer, Berlin, pp. 301–311.

Chung, W., 2005. Business Stakeholder Analyzer: An automatic classification approach to facilitating collaborative commerce on the Web. In: Hawaii International Conference on System Sciences, Island of Hawaii: IEEE Computer Society.

Codd, E.F., 1970. A Relational Model of Data for Large Shared Data Banks. Communications of the ACM 13 (6), 377–387.

Cousins, S.B., Kahn, G., 1991. The visual display of temporal information. Artificial Intelligence in Medicine 3, 341–357.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13 (3), 319–340.

Diepenbroek, M., Grobe, H., Reinke, M., Schindler, U., Schlitzer, R., Sieger, R., Wefer, G., 2002. PANGAEA-an information system for environmental sciences. Computers & Geosciences 28 (10), 1201–1210.

Donath, J., Robertson, N., 1994. The sociable Web. In: Proceedings of the Second International World Wide Web Conference. Elsevier, Chicago, IL.

Ericsson, K.A., Simon, H.A., 1993. Protocol Analysis: Verbal Reports as Data. MIT Press, Cambridge, MA.

Fredrikson, A., North, C., Plaisant, C., Shneiderman, B., 1999. Temporal, geographical and categorical aggregations viewed through coordinated displays: a case study with highway incident data. In: Proceedings of the ACM Conference on Information and Knowledge Management Workshop on New Paradigms in Information Visualization and Manipulation, pp. 26–34.

Gwizdka, J., 2002. Future time in email—design and evaluation of a task-based email interface. Technical Report IML-2002-04,. Interactive Media Laboratory, Department of Mechanical and Industrial Engineering, University of Toronto.

Havre, S., Hetzler, E., Whitney, P., Nowell, L., 2002. ThemeRiver: visualizing thematic changes in large document collections. IEEE Transactions on Visualization and Computer Graphics 8 (1), 9–20.

Hewett, T.T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., Verplank, W., 1996. ACM SIGCHI Curricula for Human–Computer Interaction. The Association for Computing Machinery, Inc., New York, NY, USA.

Hibino, S., Rudensteiner, E.A., 1998. Comparing MMVIS to a timeline for temporal trend analysis of video data. In: Proceedings of the Advanced Visual Interfaces (AVI 98).

Jain, R., 2003a. Events in heterogeneous environments (Keynote speech). In: Proceedings of the International Conference on Data Engineering. IEEE Computer Society Press, Bangalore, India, pp. 8–21.

Jain, R., 2003b. Experiential computing. Communications of the ACM 46 (7), 48–54.

Jha, M.K., McCall, C., Schonfeld, P., 2001. Using GIS, genetic algorithms, and visualization in highway development. Computer Aided Civil and Infrastructure Engineering 16 (6), 399–414.

Kosara, R., Miksch, S., 2001a. Metaphors of movement: a visualization and user interface for time-oriented, skeletal plans. Artificial Intelligence in Medicine 22 (2), 111–131.

Kosara, R., Miksch, S., 2001b. Visualizing complex notions of time. In: Proceedings of the 10th World Congress on Medical Informatics. IOS Press, Amsterdam, Netherlands, pp. 211–215.

Kreuseler, M., 2000. Visualization of geographically related multidimensional data in virtual 3D scenes. Computers & Geosciences 26 (1), 101–108.

Levine, N., 2002. CrimeStat: a spatial statistics program for the analysis of crime incident locations (v 2.0), Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC.

Lewis, J.R., 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. International Journal of Human–Computer Interaction 7 (1), 57–78.

McDonald, T., Stokes, B., 1998. A forest landscape visualization system. In: Proceedings of the Seventh International Conference on Computers in Agriculture. American Society of Agricultural Engineers, Orlando, FL, USA.

Morse, E., Lewis, M., 2000. Evaluating visualizations: using a taxonomic guide. International Journal of Human–Computer Studies 53 (5), 637–662.

Myers, J., Well, A., 1995. Research Design and Statistical Analysis. Lawrence Erlbaum Associates, Hillsdale, NJ, USA.

Noon, C.E., Hankins, C.T., 2001. Spatial data visualization in healthcare: supporting a facility location decision via GIS-based market analysis. In: Proceedings of the 34th Annual Hawaii International Conference on Systems Sciences. IEEE Computer Society, HI, USA, p. 10.

North, C., Shneiderman, B., 2000. Snap-together visualization: can users construct and operate coordinated visualizations? International Journal of Human–Computer Studies 53 (5), 715–739.

Plaisant, C., Shneiderman, B., Mushlin, R., 1998. An information architecture to support the visualization of personal histories. Information Processing and Management 34 (5), 581–597.

Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., Shneiderman, B., 1999. LifeLines: using visualization to enhance navigation and analysis of patient records. In: Proceedings of the American Medical Informatics Association Annual Fall Symposium. Hanley & Belfus, Inc., Philadelphia, PA, USA, pp. 76–80.

Price, B.A., Baecker, R.M., Small, I.S., 1993. A principled taxonomy of software visualization. Journal of Visual Languages and Computing 4 (3), 211–266.

Shahar, Y., Cheng, C., 2000. Model-based visualization of temporal abstraction. Computational Intelligence 16 (2), 279–306.

Smith, D.A., 2002. Detecting events with date and place information in unstructured text. In: Proceedings of the Second ACM/IEEE CS Joint Conference on Digital Libraries. ACM Press, Portland, OR, USA, pp. 191–196.

Weber, M., Alexa, M., Müller, W., 2001. Visualizing time-series on spirals. In: Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'01), San Diego, CA, USA.

Wehrend, S., Lewis, C., 1990. A problem-oriented classification of visualization techniques. In: Proceedings of the IEEE Visualization, pp. 139–143.

Xiong, R., Donath, J., 1999. PeopleGarden: creating data portraits for users. In: Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology, Asheville, NC, USA, pp. 37–44.

Yiu, K.S., Baecher, R., Silver, N., Long, B., 1997. A time-based interface for electronic mail and task management. In: Design of Computing Systems: Social and Ergonomic Considerations, Proceedings of the Seventh International Conference on Human–Computer Interaction (HCI International '97). Elsevier, San Francisco, CA, USA, pp. 24–29.

Zhou, M.X., Feiner, S.K., 1998. Visual task characterization for automated visual discourse synthesis. In: Proceedings of the ACM Conference on Computer-Human Interaction, pp. 392–399.

Zhu, B., 2002. Information visualization for knowledge repositories: applications and impacts. Ph.D. Dissertation, Department of Management Information Systems, The University of Arizona, Tucson.

Zhu, B., Chen, H., 2002. Visualizing the archive of a computer mediated communication process. In: Proceedings of the Second ACM/IEEE-CS joint conference on Digital libraries. Portland, OR, USA, p. 385.