INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
DEVELOPMENT AND PROOF-OF-CONCEPT OF AN INTERACTIVE VISUALIZATION SYSTEM FOR THE SPATIO-TEMPORAL ANALYSIS OF LINEAR POINT DATA

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Sang-Ki Hong, B.A., M.A.

*****

The Ohio State University
1997

Dissertation Committee:

Professor Morton E. O'Kelly, Adviser

Professor Lawrence A. Brown

Professor Joan M. Nichols

Approved by

Morton E. O'Kelly
Department of Geography
ABSTRACT

Point patterns, the locations of particular events mapped as points in a study area, are one of the most common types of spatial patterns. The term 'linear points' refers to points that occur exclusively on a line or a set of lines. The location of traffic accidents in a road network is a good example of linear point data. This study develops an interactive visualization system suitable for the analysis of linear point data to identify spatial, temporal, and spatio-temporal patterns. It also demonstrates the process of developing and applying statistical procedures to confirm the findings from visual analysis.

The visualization system consists of modules which allow the user to interactively query and display the various attributes of the data. The central module in this system, the Point Visualization module, allows for easy query construction through point-and-click methods. The results of these queries are instantaneously posted to the visual display. Visualization of the temporal components of the data is facilitated by using the vertical axis of the three-dimensional display. Further exploration of the temporal dimension is made possible through the temporal slider and slide show tools.
Traffic accident data collected for Pitt County, North Carolina for the period of 1986 through 1988 were used to demonstrate the capabilities of the visualization system. The visual analysis of the data indicated a possible space-time clustering in the case of alcohol-related traffic accidents. Subsequent statistical analyses confirmed the existence of space-time clusters for some combinations of critical interpoint distances and inter-event intervals.

The visualization system developed in this study is shown to be an efficient tool for the analysis of complex spatio-temporal data. The further expansion of the system will allow the development of a more generalized system capable of processing a wider variety of data.
Dedicated to my parents
ACKNOWLEDGMENTS

During my stay at Ohio State, so many people have given me their help, support, encouragement, and love. I would like to let them know how much I appreciate their help and support.

Many thanks go to the Department of Geography for supporting me throughout all my stay at Ohio State. Without the department’s funding and facilities, this study would have not been possible. I also thank my dissertation committee members: my adviser Morton O’Kelly, Larry Brown, and Joan Nichols. They all have been extremely supportive and kind to me.

I cannot thank enough my adviser Morton. Without his guidance and encouragement, as well as the computer equipment he provided to me, this study would never have been materialized. He picked me up when I was down and patted my back when I got something right. I will not forget his support.

I have had an extreme fortune to work with some of the most wonderful people. I cannot hope for a better work environment than the one I have had at the Department Technician’s office. For that, I like to thank my dear friend Jim DeGrand. I do not know how I can adequately express my gratitude to him. We spent countless hours working, talking, laughing, suffering together and throughout the whole time, he never let me
down. From Jim and his wife Cynthia, I have learned a great deal about being a really
decent human being. I will miss him and his family dearly.

I also like to thank Jens Blegvad for being such a good friend. Jens never
hesitated to lend a helping hand. The newest member of our Technician team, John
Larsen deserves my thanks for helping me understand difficult statistics. I send many
thanks to Jeff Arnold who painstakingly went through my manuscripts and made so many
valuable suggestions. This document reads a lot easier because of his efforts.

The biggest thanks of all should go to my family. I would not be here without the
unconditional support and love I got from my father, my mother, my wife and my
daughter. I am so glad that I could dedicate this work to my parents and make them
proud of me. I will be forever grateful to my wife Eun Jung and my daughter Sukyoon. I
do not believe that this work could have been completed without their love. I do not
think I would be complete without them, either.
VITA

December 18, 1961 ........................................ Born - Kwangju, Republic of Korea

1985 ........................................................ B.A. Geography
                                 Seoul National University, Seoul, Korea

1987 ........................................................ M.A. Geography
                                 Seoul National University, Seoul, Korea

1989 – present ............................................. Graduate Teaching and Research
                                 Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Geography
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 RESEARCH CONTEXT</td>
<td>5</td>
</tr>
<tr>
<td>1.1.1 The complexities of spatio-temporal data</td>
<td>5</td>
</tr>
<tr>
<td>1.1.2 Current status of linear point pattern analysis in geography</td>
<td>7</td>
</tr>
<tr>
<td>1.1.3 Current methods for linear point pattern analysis</td>
<td>9</td>
</tr>
<tr>
<td>1.2 IMPORTANCE OF THE RESEARCH</td>
<td>14</td>
</tr>
<tr>
<td>1.3 CONTRIBUTIONS OF THE RESEARCH</td>
<td>15</td>
</tr>
<tr>
<td>1.4 LIMITATIONS OF THE RESEARCH</td>
<td>17</td>
</tr>
<tr>
<td>CHAPTER 2 LITERATURE REVIEW</td>
<td>19</td>
</tr>
<tr>
<td>2.1 SCIENTIFIC VISUALIZATION AND VISUAL INFORMATION PROCESSING</td>
<td>21</td>
</tr>
<tr>
<td>2.1.1 Defining Scientific Visualization</td>
<td>21</td>
</tr>
</tbody>
</table>
2.1.2 Visual information processing

2.2 EXPLORATORY DATA ANALYSIS

2.3 DYNAMIC GRAPHICS AND SPATIAL DATA ANALYSIS

2.3.1 Dynamic Graphics and Data Analysis

2.3.2 Review of Selected Dynamic Graphics Systems

2.3.3 Dynamic Graphics for Spatial Data Analysis

2.4 GEOGRAPHIC VISUALIZATION

2.4.1 Analysis of complex spatial data through visual methods

CHAPTER 3 A FRAMEWORK FOR ANALYZING LINEAR POINT DATA

3.1 BUILDING A CONCEPTUAL FRAMEWORK FOR AN INTERACTIVE VISUALIZATION SYSTEM

3.1.1 Visual representation of point patterns

3.1.2 Treatment of the temporal dimension

3.1.3 Conceptual Framework

3.2 DESIGNING AN INTERACTIVE VISUALIZATION SYSTEM

3.2.1 Functional components of the system

3.3 VISUALIZATION TOOLS FOR TRAFFIC ACCIDENT DATA

3.3.1 Structural characteristics of the underlying network

3.3.2 Point Data Visualization

3.3.3 Detecting changes in the temporal component of data

CHAPTER 4 IMPLEMENTATION OF THE VISUALIZATION SYSTEM

4.1 DATA

4.2 STRUCTURE AND ORGANIZATION OF THE PROTOTYPE VISUALIZATION SYSTEM

4.2.1 Data Organization and File Structure

4.2.2 Module Organization
4.2.3 User Interface ................................................................................................................. 105

4.3 Modules and Functions in the Visualization System ......................................................... 105

4.3.1 IRIS Explorer map of the visualization system ................................................................. 107

4.3.2 Data Input .......................................................................................................................... 107

4.3.3 Point Visualization Module ............................................................................................. 109

4.3.4 Treatment of the Temporal Component in the Data ......................................................... 113

4.3.5 Network Visualization Module ....................................................................................... 117

4.4 Examples of the Graphic Display ....................................................................................... 119

CHAPTER 5 AN ASSESSMENT OF THE IMPLEMENTATION AND DISCUSSION ........ 128

5.1 Visualizing the Traffic Accident Patterns .......................................................................... 129

5.2 Testing the Presence of Space-Time Clusters ................................................................ 150

5.2.1 Methods and Test Results ............................................................................................ 150

CHAPTER 6 CONCLUSION ...................................................................................................... 167

BIBLIOGRAPHY .......................................................................................................................... 175

APPENDIX ................................................................................................................................. 188
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Spatial-Visual Operations</td>
<td>33</td>
</tr>
<tr>
<td>3.1</td>
<td>Visual variables to represent traffic accident data</td>
<td>89</td>
</tr>
<tr>
<td>5.1</td>
<td>Result of Knox’s Contingency Table Analysis</td>
<td>155</td>
</tr>
<tr>
<td>5.2</td>
<td>Result of Knox’s Contingency Table Analysis (with revised data)</td>
<td>158</td>
</tr>
<tr>
<td>5.3</td>
<td>Verification of Knox’s approach using Mantel’s method</td>
<td>162</td>
</tr>
<tr>
<td>5.4</td>
<td>Verification of Knox’s approach using Mantel’s method (with revised data)</td>
<td>163</td>
</tr>
<tr>
<td>5.5</td>
<td>Result of Mantel’s method with reciprocal transformation (with revised data)</td>
<td>165</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Dr. John Snow's map of the spread of cholera in central London in September 1854</td>
<td>26</td>
</tr>
<tr>
<td>2.2</td>
<td>The basic brushing method with highlighting</td>
<td>41</td>
</tr>
<tr>
<td>2.3</td>
<td>Conditioning on one variable</td>
<td>43</td>
</tr>
<tr>
<td>2.4</td>
<td>Shadow highlight operation</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>MacSpin display</td>
<td>52</td>
</tr>
<tr>
<td>2.6</td>
<td>Polygon Explorer display</td>
<td>56</td>
</tr>
<tr>
<td>2.7</td>
<td>Dynamic link display</td>
<td>58</td>
</tr>
<tr>
<td>2.8</td>
<td>Dynamic network node display</td>
<td>58</td>
</tr>
<tr>
<td>2.9</td>
<td>A SEFLOW display showing the Cartographic Representation Module</td>
<td>62</td>
</tr>
<tr>
<td>2.10</td>
<td>Muehrcke's Cartographic Process</td>
<td>66</td>
</tr>
<tr>
<td>3.1</td>
<td>Conceptual framework for the prototype visualization system</td>
<td>82</td>
</tr>
<tr>
<td>3.2</td>
<td>Visual variables applied to cartographic data types</td>
<td>88</td>
</tr>
<tr>
<td>4.1</td>
<td>Relative location of Pitt County in the state of North Carolina</td>
<td>97</td>
</tr>
<tr>
<td>4.2</td>
<td>Municipalities and major roads in Pitt County, North Carolina</td>
<td>98</td>
</tr>
</tbody>
</table>
4.3 Context level data flow diagram of the prototype system .....................100
4.4 Level 1 data flow diagram of the prototype system ............................101
4.5 Expanded DFD (level 2) of the “process user data” process ................102
4.6 User interface of the prototype visualization system ............................106
4.7 Overall structure of the prototype visualization system .......................108
4.8 Point Visualization Module .................................................................110
4.9 Slide Show Module (Timer) .................................................................116
4.10 Network Visualization Module ............................................................118
4.11 Example of point attribute query .........................................................121
4.12 Example of vertical displacement at side angle .................................123
4.13 Example of zoomed in display .............................................................124
4.14 Example of displaying the temporal spread ........................................125
4.15 Example of the use of color variable ...................................................127
5.1 Overall pattern of traffic accidents .....................................................130
5.2 Accident locations in and around Greenville, NC .............................131
5.3 Network display showing functional class as a color variable ..............133
5.4 Locations of alcohol-related traffic accidents .....................................134
5.5 Alcohol-related accidents in and around Greenville, NC (viewed from the south) .....................................................136
5.6 Alcohol-related accidents that occurred on Monday ..........................137
5.7 Alcohol-related accidents that occurred on Saturday ........................138
| 5.8 | Alcohol-related accident locations in and around Greenville, NC | 140 |
| 5.9 | Location of alcohol-related injury accidents at low speed | 142 |
| 5.10 | Location of alcohol-related injury accidents at medium speed | 143 |
| 5.11 | Location of alcohol-related injury accidents at high speed | 144 |
| 5.12 | Location of alcohol-related fatal accidents at medium speed | 146 |
| 5.13 | Location of alcohol-related fatal accidents at high speed | 147 |
| 5.14 | Location of traffic accidents that occurred on Jan 9, 1986 | 149 |
| 5.15 | Frequency distribution of interpoint distances | 152 |
| 5.16 | Frequency distribution of inter-event intervals | 153 |
Geographers have always been interested in studying the spatial patterns of particular phenomena. In general, these spatial patterns are represented by maps of points, lines, and polygons, based upon the types of data that are being examined. Of the various spatial patterns resulting from such mapping, this research specifically focuses on the analysis of point patterns.

Point patterns, the locations of particular events mapped as points in a study area, are one of the most common types of spatial patterns. Examples of point patterns are abundant at various scales. Tree positions in a forest, the location of birds' nests, imperfections in metals or rocks, galaxies, towns, and earthquakes (Ripley, 1981) can all be mapped as points, thereby creating spatial point patterns. Therefore, it is not hard to see that the study of spatial point patterns has long been a major research topic in many disciplines as diverse as ecology, epidemiology, biology, astronomy, geology, and geography.
A map of a point pattern has two major components: the points representing the objects being studied, and the geographical area in which they are located (Boots and Getis, 1988). Based upon the nature of the data, the study area can be represented in one, two, or three dimensions. When we are concerned with the distribution of highway exits, for example, the study area can be represented as a line (one-dimensional study area). On the other hand, the study of settlement locations in a region would require a two-dimensional study area. Furthermore, there may be a case, such as an earthquake study, where a three-dimensional study area is more appropriate.

There exists a large body of literature on the analysis of spatial point patterns, especially in the branch of statistics called spatial statistics. However, most spatial point pattern analyses in the statistics literature have focused on analyzing point patterns occurring in two-dimensional study areas. The study of the point pattern of childhood cancer occurrences is a typical example of such analyses. What is notably missing is the study of point phenomenon occurring exclusively on a line or a set of lines. We will call this subset of point patterns as 'linear point patterns,' following Roder's (1974) terminology. Some examples of linear point patterns are: the location of traffic accidents in a road network, the location of power outages in an electrical network, the level of pollution measured at various spots along a river, or the distribution of business activities along major roads.

Within the realm of linear point patterns, the complexity of point pattern analysis varies depending upon the characteristics of the data involved and the purpose of the study. For example, if we are concerned only with the 'location' of the points on a line
without any consideration of the attributes associated with each point, the point pattern becomes a purely geometric one and can be treated in a relatively straightforward manner. However, the analysis would be much more complex if the data under investigation includes both spatial and aspatial attributes, and if we wish to analyze the point pattern not only in its geometrical arrangement, but also in terms of the relationships between spatial and aspatial attributes. Analysis gets even more complicated when the data contain a temporal element. In this case, the analysis would involve not only examining the patterns in the spatial domain, but also searching for clues on a possible space-time interaction.

One good example of such a highly complex spatio-temporal linear point pattern is the pattern of traffic accidents. Since the accidents can only occur in a road network, the point pattern of traffic accidents is clearly linear. At the same time, each accident has spatial and aspatial attributes associated with it. Furthermore, typical accident ‘data’ includes the time of the accident that adds another dimension to the analysis. Although there is a considerable number of studies on traffic accidents in the field of accident analysis, most research has focused on identifying the relationship between the number of accidents and the factors (mostly aspatial, such as driver characteristics or the involvement of alcohol) contributing to the accident. In other words, some of the fundamental aspects of the accident data, namely the location of each accident and the underlying spatial characteristics of that location in the network, have largely been ignored.
Given the complex nature of traffic accident data, how can we efficiently analyze the patterns? It is well known that the techniques developed in the field of scientific visualization (SV), coupled with exploratory data analysis (EDA) methods, can be effectively used to detect potentially ‘meaningful’ patterns in complex multidimensional data such as traffic accident data. When these techniques are integrated into a dynamic graphical environment that facilitates rapid visualization and interactive exploration of the data, an efficient analytical system for spatial point patterns can be constructed.

Once an interesting or abnormal pattern is found, it must be put through a formal test to verify that the pattern is not generated by chance. In other words, the analysis needs to be moved to a confirmatory stage. It is also clear that there is a need to demonstrate the application of statistical procedures to confirm the validity of the patterns that are found through the use of the visualization system.

As discussed above, in analyzing complex linear point patterns such as traffic accident patterns in the spatial and temporal domain, a large void exists in both point pattern analysis and traffic accident research. The intention of this research is to fill in this void by developing a new method for linear point pattern analysis that combines the concepts and techniques developed in the field of SV, EDA, dynamic graphics, and spatial statistics. Specifically, this research has the following goals:

1) to develop an interactive visualization system suitable for the exploration of linear point data to identify spatial, temporal, and spatio-temporal patterns, and
2) to demonstrate the process of developing and applying statistical procedures to confirm the findings from visual analysis.

1.1 Research Context

1.1.1 The complexities of spatio-temporal data

Due to the enormous improvement in data collection and management technology, more and more spatial data are being collected every day. Not only is the volume of the collected data growing exponentially, but the information included in the data collection process is getting more and more detailed and diverse. Take, for example, the 911 calls generated every day in the U.S. Throughout the United States, each call that comes in is time-stamped, the location mapped, and the nature of the emergency recorded. For researchers interested in examining the spatial and temporal patterns and the underlying causes for creating the patterns in these emergency calls, this kind of detailed information can provide a tremendous research opportunity. However, another direct consequence of this data explosion is that much of the information being collected is simply wasted because, in most cases, we do not have adequate tools to undertake the challenge of analyzing and extracting useful information. Without a proper way of analyzing it, all we can do with the enormity of the data is simply 'warehouse' it.

Geographic Information Systems (GIS), defined as complex, integrated computer systems for the input, storage, retrieval, manipulation, analysis, and display of spatially-referenced data (Marble, 1984), is a perfect tool for handling spatial data including spatial
point patterns. Ideally, GIS should not only be able to help researchers organize and manipulate the data but also provide users with a powerful set of tools for analysis. Unfortunately, the analytical subsystems of current GIS packages do not provide the adequate means to fulfill this requirement. The lack of analytical functions in GIS, both in theoretical perspective and implementations, has been widely regarded as a fundamental weakness of the current state of GIS (Fotheringham and Rogerson, 1994). For GIS to meet the demand for analytical tools, more and more analytical functions must be incorporated into GIS. This can only be accomplished through the rigorous application and implementation of available techniques—many of them developed in the field of spatial statistics—as well as the development of new analytical methods within the context of spatial data analysis.

There have been significant efforts made to alleviate this 'data-rich, analysis-poor' situation by adopting new developments in the fields of scientific visualization (SV) and exploratory data analysis (EDA) to the realm of GIS. Most notable of these efforts is the series of studies conducted under the guidance of Dr. Duane Marble at The Ohio State University. For instance, Sandhu (1990) tackled the visualization problem of very large space-time data by combining EDA and SV techniques in the analysis of global earthquake data. A prototype visualization system for the analysis of large spatial flow data has also been created by Gou (1993). This study intends to continue the track of spatial data visualization efforts that have previously been made in GIS and to expand the analytical capabilities of GIS by developing new methodologies for the analysis of spatio-temporal point patterns.
1.1.2 Current status of linear point pattern analysis in geography

The first study on linear point patterns in geographic literature is found in Dacey's (1960) study of the spacing of river towns in the central lowlands of the United States. In his study, Dacey used the technique called 'reflexive nearest neighbor method' (Clark, 1956) to test the earlier observation on the spacing of the river towns made by Burghardt (1959), who indicated that "The larger river towns ... reveal an interesting uniformity of spacing along the rivers (p. 322)." The result of Dacey's analysis, however, indicated that the spacing of river towns on the Mississippi is more grouped than random, rejecting Burghardt's assertion. Another example of using reflexive nearest neighbor method was given in Boots and Getis (1988) for the study of "fast food" outlet locations in southern Ontario.

Nearest neighbor analysis seems to be the most popular method for analyzing linear point patterns. While all nearest neighbor analyses generate summary statistics that can be used to test the point patterns against a random pattern, methods differ in terms of the calculation of the distances between nearest neighbor points and the level of information used to generate the summary statistics (for example, Selkirk and Neave, 1984; Durbin, 1965). The use of single summary statistics, however, has been criticized by many researchers because the information on the distances between points is lost in conversion to a single statistic. The most common test that retains the original information on interpoint distances is the technique called "refined nearest neighbor analysis (Diggle, 1979)." This involves comparing the complete distribution function of the observed nearest neighbor distances with the distribution function of expected nearest
distances for a random pattern. This method was used by Roder (1974) to revisit the river town spacing analyzed by Dacey (1960). Interestingly enough, Roder’s result suggested that Burghardt’s original assertion of even spacing was correct.

Another notable technique developed for the analysis of linear point patterns is the K-cluster analysis suggested in Hsu (1973). Pointing out that the conventional approaches are limited to the analysis of point distributions in the total space, he argued for the use of K-cluster analysis, which is a derived Poisson model of point distributions in space. His study is also noteworthy because it is one of the few studies that recognized the importance of spatial point pattern analysis of traffic accidents.

As we have seen above, there are very few studies on linear point patterns in the geography literature. Although linear point patterns are quite common and represent a significant subset of point patterns in general, the relative lack of studies on such patterns has resulted in the neglect of this important set of spatial phenomena. It is highly unlikely that this lack of research is due to a simple lack of interest on the geographers’ part. It seems that the complexity of data and the lack of proper tools to analyze it have hindered the advancement of research in this area. Therefore, it is very important to construct a sound conceptual basis for examining the linear point patterns and to develop proper analytical tools to tackle this problem.
1.1.3 Current methods for linear point pattern analysis

Confirmatory vs. Exploratory Approach

The usual process of analyzing spatial point patterns is to first create a dot map of the occurrences of the phenomena. The dot map is then examined to see if the distribution of points is due to stochastic process. In other words, the pattern is compared to the pattern generated by complete spatial randomness, (CSR, after Diggle, 1983), by means of a statistical test. A typical null hypothesis is that the pattern is the result of a CSR process. If the test reveals that the null hypothesis can be rejected (i.e., the pattern is significantly different from that of a random process), then the next task is to find the underlying process that generated the pattern. Some spatial statistics (such as Moran's $I$-statistic (Moran, 1950) that measures the degree of spatial autocorrelation) can provide some clues about the underlying process. Finally, with improved understanding of the data, a new hypothesis is developed and the statistical testing process is repeated.

The sequence of the scientific inquiry described above closely follows what Tukey (1977) called 'confirmatory data analysis.' The problem with this traditional approach in the face of highly complex, multidimensional spatio-temporal data is that, frequently, the researcher has no prior notions about the structure and the relationships that might exist among the variables. The data go unexamined, either because the data set is too large, or because it is too complex to be properly analyzed with existing techniques. When there is no theory to construct a model of the data, the traditional 'confirmatory approach' is hardly useful. The task now is no longer simply testing the
mapped data patterns against a known distribution pattern generated from a carefully constructed model. Rather, the primary task is finding meaningful patterns, either spatially, temporally, or both, in the data to generate new hypothesis. In other words, at this stage, the exploration of data becomes the primary objective of the analysis.

Scientific Visualization, Exploratory Data Analysis and Dynamic Graphics

Visual representation of various characteristics of the data has been regarded as one of the most efficient ways to explain and explore the data in a large data set. The famous adage that "a picture is worth a thousand words" also applies to scientific research. Visualizing data can help reduce the amount of time and effort required to understand the data. Furthermore, it can also help researchers find significant patterns that may not be detected through non-graphic methods.

The need to visualize data leads us to the area of scientific visualization. Scientific visualization is a rapidly developing field aided by the explosive growth in computer technology and the recognition of human ability to process visual information. In the context of scientific visualization, "to visualize" refers specifically to using visual tools (usually computer graphics) to help scientists explore data and develop insights (MacEachren and Monmonier, 1992). Scientific visualization provides an exciting new way to perform scientific research by equipping scientists with a new set of powerful, interactive, and highly visual tools to analyze data.

Data visualization is most effective in an exploratory sense. One critical difference between exploratory data analysis (EDA) and the traditional confirmatory
approach is the absence of *a priori* hypotheses on the data. In EDA, the researcher first wants to know what is going on in the data in order to improve his/her understanding of the data. Hypotheses can be generated afterward through observing the data. Many visualization techniques have been developed for EDA, including identification, brushing, and projection pursuit, just to name a few. EDA techniques are particularly suitable when the characteristics of the data are largely unknown or the data are multidimensional.

Dynamic graphics has been developed largely in parallel with EDA. The term "dynamic" in this case means that a user can interactively change almost all the elements of the graphic display to suit his/her needs. With a simple mouse click, a user can change the layout of graphic windows, open up another window, select different variables to display, and dynamically adjust the threshold level to find meaningful patterns. Dynamic graphics display, therefore, opens up a whole new way to visualize data.

**Point Pattern Analysis in Spatial Statistics**

Spatial statistics is a subfield of statistics that deals with spatial patterns. Spatial patterns have long been studied in diverse disciplines such as ecology, forestry, archeology, cosmology, geography, and seismology. Many of the early techniques associated with analyzing spatial patterns have been developed outside statistics. The significant turning point for the modern development in spatial statistics was the publication of an article by Ripley (1977) in which the *K-function* was introduced. *K-function* is a powerful descriptive and modeling tool that was originally suggested by
Bartlett (1964) (Cressie, 1993). Although there have been techniques in distinguishing complete spatial randomness from spatially regular and clustered patterns in the field (for example, quadrat and distance methods), $K$-function plots give a picture of such behavior for mapped data at a multitude of scales (Cressie 1993, p. 579).

A spatial point pattern is the simplest form of spatial data. We are interested in analyzing point patterns to find out whether observed point patterns show a systematic tendency in their distributions, either in the form of clustering or regularity. In spatial statistics, the analysis of a spatial point pattern usually involves describing the first and second order properties of the point pattern (Bailey and Gatrell, 1995). First order properties are described in terms of intensity of the process, which represents the global trend in the data. Second order properties represent the spatial dependence in the spatial process resulting from the spatial correlation structure. In other words, it represents the local or small scale effects in the data. The aforementioned Ripley’s $K$-function, which is also called the ‘reduced second moment measure,’ is closely related to the second order properties and also measures the local effects.

There are several well-established techniques for point pattern analysis in spatial statistics: from the simple quadrat methods to the sophisticated modeling of spatial point processes. Details on these techniques can be found elsewhere (see, for example, Ripley, 1977; Cressie, 1993; Bailey and Gatrell, 1955). However, it is worth mentioning again that most real world applications of these techniques found in the spatial statistics literature focus on point patterns in two-dimensional study areas. The theoretical and
practical ramifications of applying these statistical methods to linear point patterns have not been sufficiently examined.

It is possible to extend spatial statistical techniques to incorporate the possible effect of “space-time interaction” (Knox, 1964) in point patterns. One approach is to extend the notion of complete spatial randomness to include space-time point processes, which then can be called complete spatio-temporal randomness (CSTR; Cressie, 1993). CSTR means that there is an absence of structure in time as well as in space. It is the natural null hypothesis against which observed space-time point patterns could be tested. Another frequently used alternative is the space-time clustering. Space-time clustering is said to exist if, among those events that are close in time, there are events that are closer in space than would be expected due to chance alone (McAuliffe and Afifi, 1984). Williams (1984) reviewed various methods to measure space-time clustering. The following is a brief summary of available techniques.

- Knox’s (1964) developed a method for testing the existence of space-time clustering by considering all possible pairs of points both in time and at a distance interval. His method has been utilized in the examination of time-space clustering of several diseases, most notable of which is the “space-time clusters” of Burkitt’s lymphoma (Doll, 1978).

- Mantel (1967) generalized Knox’s procedure by eliminating Knox’s assumption about the distribution of points (Knox’s assumed a Poisson distribution). He also suggested using a different ‘closeness measure,’ which can incorporate more information about the distance between points. His
statistic is $Z = \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij} T_{ij}$ where $S_{ij} = 1/d_{ij}$ and $T_{ij} = 1/t_{ij}$ are measures of spatial and temporal closeness.

- McAuliffe and Afifi (1984) proposed another approach using the distance between nearest neighbor occurrences as a measure of closeness. Their method has an advantage over Knox's (1964) or Mantel's (1967) method in that it does not depend on the choice of arbitrary constants required to construct a contingency table.

1.2 Importance of the research

This study explores the new area of dynamic visualization of linear point data, thereby broadening the scope of GIS research. By implementing SV and EDA in a spatial analytical context, the study provides a sound basis to test the usefulness of these techniques in spatial analysis, especially in hypothesis generation and testing. As mentioned earlier, the lack of analytical capabilities in GIS has been a major roadblock to the wider adoption of GIS in scientific endeavors. This research demonstrates how a powerful visualization system can be used to examine a complex spatio-temporal data set. The method used in this study for the implementation of the visualization system deliberately follows the standard system design process developed in the area of software engineering. There is, therefore, an ample opportunity for further development or modifications for other applications.
Dynamic visualization is an important issue in the continuing efforts in cartography to make visual representation of spatial data more dynamic and intuitive (for examples, see Monmonier and MacEachren (eds.), 1992). The prototype visualization system implemented in this study allows users to rapidly create multiple representations of selected subsets of the data. Therefore, the techniques used in this study can be easily adopted to test the effectiveness of a cartographic visualization scheme.

1.3 Contributions of the research

By developing a theoretical framework for the implementation of SV and EDA in a spatial context and by identifying the benefits and limitations in implementing these methods in GIS, this study makes a fundamental contribution to the development of GIS theory. Even with the abundance of available point data, efforts to develop a comprehensive visualization system have been scarce. The notable development made in SV and EDA in recent years and an ever-increasing demand for adequate tools for spatial data analysis call for significant efforts on the part of GIS researchers to incorporate more dynamic visualization tools in GIS. Although this research deals with only a single type of point data, it is hoped that the conceptual framework used in this study would serve as a good starting point for further research.

Furthermore, this research makes a contribution to the theory of scientific visualization by applying the concepts and techniques developed in SV to the analysis of spatio-temporal data. SV concepts and techniques have been applied to many areas of research. Strangely enough, however, it is hard to find applications in the context of
spatial data analysis. Consequently, it is difficult to tell which of the concepts and techniques in SV are suitable for analyzing spatial data and how useful they could be. Implementation, especially with empirical data in real world situations, provides both an effective testing ground and an opportunity for further refinement. Likewise, this research contributes to the development of EDA by identifying which EDA techniques are more suitable for dealing with spatio-temporal data and by demonstrating the use of EDA techniques in spatial data analysis.

In addition to bringing state-of-the-art techniques to GIS research, this study supplies a new environment for analyzing point data for researchers in related fields. This will enable researchers to easily discover significant patterns which otherwise could not have been detected, rapidly test their hypotheses, or simply experiment with the data to get intuitions on the data which can later be used for hypothesis generation. This new way of approaching data analysis will certainly benefit the quality of spatial analysis and also facilitate new discoveries based on new intuitions acquired from exploring the data.

The interactive environment and the graphical user interface used in this research reduces the time required for an analysis and also makes system usage more intuitive, thereby helping researchers do more in less time. The integrated environment allows the iterative process of data input, analysis, and display of the results to be performed in one environment. This means that researchers do not have to move their data and results between different machines, operating systems, or programs (such as from PC to mainframe and back to PC or from statistical package to graphics program). This will
help cut down the time and effort required for analysis and also eliminate the
incompatibility problem between various programs.

1.4 Limitations of the research

As a proof of concept, this research has some limitations. First of all, the
functionality of the system is limited to certain visualization and exploratory data analysis
techniques. Certainly, one of the goals of the research is to identify which techniques are
valid in the context of spatial data analysis. Nevertheless, this research could not cover
all the available techniques, and even with relevant techniques, some had to be eliminated
for the sake of implementation simplicity.

In a broader perspective, this research has limitations in the sense that it deals
only with a particular set of point data. Notwithstanding its large volume and
complexity, traffic accident data represents a very small portion of possible linear point
data. Since the system is developed with a specific input data structure, it cannot be
directly used for other types of data. This limitation stems from the recognition of the
fact that at this point implementing the conceptual framework in a working visualization
system is more important, as well as more feasible, than building a generalized system
that can interface with various external programs. Therefore, the data input system is
treated as fixed. As a result, significant changes would have to be made if the system is
to be used as a general purpose analytical system.

Finally, this research deals only with traffic accident data that are relatively fixed
in the temporal dimension. In other words, this system is not suitable for real-time data
analysis and visualization. Real-time treatment of linear point data, especially traffic accident data, is certainly of critical value in transportation management contexts. The implemented system can be extended to handle real-time input data by modifying the data input function in the system. However, given the time constraints in the study, this task must be relegated to future research.
CHAPTER 2

LITERATURE REVIEW

As mentioned in the first chapter, the aim of this research is to establish a theoretical framework for developing an effective visualization system for the analysis of linear point data. To base our framework on sound conceptual and theoretical grounds, we need to examine the cumulative contributions of research in relevant fields. Contributions from the fields of scientific visualization, exploratory data analysis, cartographic visualization, dynamic graphics, and spatial statistics are all valuable resources upon which we can build the proposed system. This chapter reviews these contributions.

We start our discussion with the rapidly emerging field of scientific visualization (SV) and examine how it fundamentally changed our way of dealing with data. This section defines scientific visualization in various contexts, especially in the context of spatial data analysis and also examines the way in which the method of scientific visualization is tied to geographic data analysis.
We then go back to the history of scientific methods to discuss exploratory data analysis (EDA)—an alternative to the more traditional confirmatory method of data analysis. What is it and how did it begin? What impact has it had upon the fundamental shift in scientific data analysis methodology? What is the relationship between EDA and scientific visualization and why is it relevant to the proper construction of a framework for visual data analysis? These are the questions that must be answered to fully understand the fundamental philosophical changes occurring in data analysis.

EDA relies heavily on visual presentation of data in the search for meaningful patterns. As such, it is closely tied to scientific visualization. In fact, many SV systems incorporate many EDA methods in their core functions. The nature of EDA as 'detective work' (Tukey, 1977) requires viewing data in many possible combinations of angles, dimensions, and projections. Therefore, EDA is invariably dynamic and requires a dynamic environment. While early EDA systems were implemented in very expensive, proprietary computer display systems, the rapid development in computer graphics hardware and the lowering of costs in building sophisticated graphics systems on more widely available platforms has brought many innovative dynamic graphics systems within easy reach of a researcher interested in exploring data visually and dynamically. It is such a system that this research focuses on and a careful examination of the advances and the innovations found in many available dynamic graphics systems can undoubtedly provide valuable insights.

Cartographic data representation is inherently visual. Throughout the long history of cartographic production and research, cartographers have developed highly
sophisticated methods of visual representation for an enormous amount of spatial and aspatial information. Recently, much effort has been made to free cartography from the limitation imposed by its traditional medium of choice, namely, printed maps and their ‘static’ nature. While the need for and the great potential of more ‘lively’ maps was recognized early on (see Thrower 1959), it was not until recently that much advancement in making maps more dynamic has been made. At the root of these newly invigorated activities, of course, is the explosion of computer technology and its inevitable incorporation into cartographic methods and production. We need to look closely at the contributions of cartographic research in this area to incorporate the findings and innovative techniques into the proposed system.

Finally, this chapter discusses the specific techniques that have been developed in the field of geography and spatial statistics in analyzing linear point data.

2.1 Scientific Visualization and Visual Information Processing

2.1.1 Defining Scientific Visualization

Improvements in data-gathering technology in recent years have effectively put us in the situation where the amount of the data gathered from various sources far exceeds our capability to analyze them. Likewise, the amount of geographic information that will become available in the future is expected to grow exponentially. For example, image data from EOS remote sensing satellites is expected to accumulate at a rate of one terabit
(10^{12} \text{ bits}) \text{ per day (Soffen 1990, quoted from DiBiase et al. 1992). All we can do, when met with the enormity of the data without proper ways of analyzing them, is to simply 'warehouse' them. Clearly, scientists and engineers needed a new method of data analysis, and from this need, the field of scientific visualization emerged.}

Visualization has been called the second computer revolution (Friedhoff 1989). The National Science Foundation report on Visualization in Scientific Computing (ViSC) (McCormick et al. 1987, p. 3) defines visualization as:

"... a method of computing. It transforms the symbolic into the geometric, enabling researchers to observe their simulations and computations. Visualization offers a method for seeing the unseen."

While visualization encompasses a wide range of research areas and topics such as computer graphics, image processing, computer vision, computer-aided design, signal processing, and user interface studies (McCormick et al. 1987), a consistent and coherent emphasis has always been on its role as an enabling technology that facilitates the perception, use, and communication of visual information.

Scientific Visualization, Cartography, and GIS

In dealing with spatial and even non-spatial relationships, geographers are most comfortable with a depiction that allows them to visualize relationships and connections (MacEachren 1992). In the realm of geographic information systems, visualization encompasses the historical discipline of cartography as well as the emerging fields of scientific visualization and information science (Clapham 1992). The definition offered by Buttenfield and Mackaness (1991, p. 432) identifies three different aspects of
visualization: computational, cognitive, and graphic design. Their definition of visualization is:

"...the process of representing information synoptically for the purpose of recognizing, communicating, and interpreting pattern and structure. Its domain encompasses the computational, cognitive, and mechanical aspects of generating, organizing, manipulating, and comprehending such representations."

Recently, visualization has been drawing more and more interest from geographers and cartographers, which has culminated in a special issue of Cartography and Geographic Information Systems on Geographic Visualization (Monmonier and MacEachren (eds.) 1992). Indeed, the works of geography and cartography have a great deal of traditional and natural interest in visualization. After all, maps are one of the oldest forms of visual tools to abstract, analyze and communicate spatial information. However, visualization, particularly in the form of maps in GIS, has not taken full advantage of developments in cartographic representation and communication research.

Computer-assisted geographic visualization is fundamentally different from traditional analog cartography that is bounded by what Goodchild (1988) called "pen and paper technology." Digital technology provides us a whole new way of dealing with spatial information by freeing the visualization process from the limitations of pen and paper technology. Some of the possible venues this enabling digital technology opens up include: the direct depiction of movement and change, multiple views of the same data, user interaction with maps, realism (through three-dimensional stereo views and other
techniques), false realism (through fractal generation of landscapes), and the integration of maps with other graphics, text and sound (MacEachren and Monmonier 1992).

2.1.2 Visual information processing

The renewed recognition of the power of human visual information processing

Human vision, despite its extremely efficient capability to detect subtle patterns and changes in complex subjects, has not been widely recognized as an ‘analytical tool’ for data analysis. However, as the amount and the dimension of information that needs to be processed and analyzed increased, and as the inadequacy of the numerical approach to these complex data became more and more apparent, visual analysis utilizing natural human visual information processing capability emerged as one of the powerful tools to be used in the analysis. Graphics can store enormous amounts of data in a very compact form and be presented in a manner in which many hidden relationships among variables can be easily recognized. As Tufte (1983) succinctly and aptly put, ‘graphics reveals data.’ Therefore, when complex data are presented in graphical forms, in many instances it is surprisingly easy to recognize underlying patterns by simply looking at it. As we can see in the following quote from Tufte (1990, p. 50), we use our visual information processing ability all the time in our everyday life:

“We thrive in information-thick worlds because of our marvelous and everyday capacities to select, edit, single out, structure, highlight, group, pair, merge, harmonize, synthesize, focus, organize, condense, reduce, boil down, choose, categorize, catalog, classify, list, abstract, scan, look into, idealize, isolate, discriminate, distinguish, screen, pigeonhole, pick over,
sort, integrate, blend, inspect, filter, lump, skip, smooth, chunk, average, approximate, cluster, aggregate, outline, summarize, itemize, review, dip into, flip through, browse, glance into, leaf through, skim, refine, enumerate, glean, synopsize, winnow the wheat from the chaff, and separate the sheep from the goats.

Visual inspection of data can not only provide a means to quickly detect patterns, but also provide a possible line of further investigation to reach a cause and effect relationship among depicted data. Probably the most well known example of simple graphic presentation of data leading to the discovery of an ultimate cause and effect relationship is the map created by Dr. John Snow regarding the spread of the cholera in central London in September, 1854 (Tufte 1983). He plotted the cholera deaths with dots and also marked the area’s 11 water pumps (Figure 2.1). Upon inspection of the plotted data, Snow found that cholera occurred almost entirely among those who lived near and drank from one of the water pumps (specifically, the Broad Street water pump, which is located in the center of the map). By having the pump’s handle removed, he could stop the epidemic, which cost more than 500 lives*.

* Recently Tobler (1994) wrote a BASIC program showing Dr. Snow’s map and Thiessen polygon boundaries around the pumps. It was developed for a student exercise in a course "Teaching Introductory Geographical Data Analysis with GIS" and shows how effective Dr. Snow’s map is in introducing the power of graphical analysis.
Figure 2.1 Dr. John Snow's map of the spread of cholera in central London in September 1854. Deaths are marked by dots and water pumps by crosses. The contaminated pump is located on Broad Street. (Source: Tufte 1983).
The role of graphics in scientific analysis

Quantitative graphics have been used for thousands of years in scientific endeavors (for a brief history of quantitative graphics, see Beniger and Robyn 1978). Graphics communicate quantitative and categorical information and are highly information intensive. Graphics, or graphical language, as opposed to written language, is used extensively to convey information because it does so effectively by using the enormous power of our eye-brain system to perceive geometrical patterns (Cleveland 1984). In his influential book on graphics and graphical principles, *Semiology of Graphics*, Bertin (1983, p. 12) recognized three fundamental functions of graphic representation:

- Recording information: creating a storage mechanism, which avoids the effort of memorization.
- Communicating information: creating a memorizable image, which will inscribe the information in the viewer’s mind.
- Processing information: furnishing the drawings, which permit a simplification and its justification.

From the earliest known map of Northern Mesopotamia on a clay tablet at about 3800 B.C. to modern computer graphic displays of the molecular structures of certain chemicals, graphical representation of information has served these fundamental functions of recording, communicating, and processing information. When these functions are facilitated by a rational and efficient tool, graphics become one of the major "languages" (Bertin 1983, p. 2) applicable to information processing.
The fact that graphics play such a vital role in information processing makes the job of maintaining graphical integrity in displaying data all the more important. Since "graphics offers unlimited choice of constructions for any given information (Bertin 1983, p. 100)," the decision as to what type of graphics should be used needs to be made carefully based on an evaluation of the properties of the data. Examples are abundant (for many good and bad examples, see Tufte 1983) in which the graphical integrity is compromised, intentionally or unintentionally, thereby giving a false impression of the facts or misleading the viewers into the wrong conclusion. Monmonier (1991) also showed how spatial information can be easily distorted with maps. This is especially true these days when there are many easily-accessible desktop mapping programs that do not require extensive cartographic background to create maps. Therefore, for graphics to be useful and truthful, it is worth remembering Tufte's (1983) principles to assure graphical integrity:

- the representation of numbers should be directly proportional to the numerical quantities represented
- labeling should be used to defeat graphical distortion and ambiguity
- show data variation, not design variation
- the number of information-carrying dimensions depicted should not exceed the number of dimensions in the data
- graphics must not quote data out of context

Traditionally, graphic tools were often used to begin a line of investigation and served a communication purpose once conclusions had been reached. But graphic means
were rarely considered acceptable ways to reach those conclusions and the real work of science was left to more “rigorous numerical analysis” (MacEachren and Monmonier 1992). However, as scientists struggle with the inundation of quantitative information that is becoming available, it is evident that graphics can play a much more critical role in data analysis, especially for exploratory purposes. Simply put, patterns and relationships are easier to detect in graphical displays. In this case graphics function as a research instrument. “When one can superimpose, juxtapose, transpose, and permute graphic images in ways that lead to groupings and classings, the graphic image passes from the dead image, the “illustration,” to the living image, the widely accessible research instrument it is now becoming (Bertin 1983, p. 4).” The resurgence of interest in statistical graphics and the plethora of studies on the visualization techniques and systems in various fields attest to the new role of graphics as a research instrument in data analysis.

**Abstract analytical thinking vs. visual thinking**

Vision is efficient because it produces abstractions from the complex input to the system. This ability to produce abstractions through vision comes before conscious (i.e., logical) processing of the information (Arnheim 1969; Friedhoff and Benson 1989). The primary advantage of visualization is that envisioning is a process of abstraction. As MacEachren and Ganter (1990) put it, “The mind’s strengths are in simplification, approximation and abstraction” (p.67). There are two distinct but related activities
involved in visualization: visual thinking and visual communication. According to DiBiase et al. (1992, p. 201):

"Scientists are engaged in visual thinking when their intent is to produce new knowledge and their method involves creating and interpreting graphic representations. When their intent turns to distributing existing knowledge in an unambiguous graphic form, they are engaged in visual communication."

It is well known that each side of the human cerebral hemispheres (two sides of the brain) are related to different mental functions. The monitoring of brain waves in the right and left hemispheres while the subject performed different tasks (Ornstein 1973) provides evidence that not only do the functions of the left and right sides of the brain differ markedly, but that both sides of the brain are also involved in highly complex cognitive functioning. The left hemisphere mode is verbal, objective, analytic, linear, symbolic, logical, rational, abstract, temporal, and digital; the right hemisphere mode is visual, holistic, spatial, analogic, concrete, synthetic, intuitive, nonrational, and nontemporal (Muehrcke 1980).

Unfortunately, scientists have largely ignored the mental functions of the right hemisphere of the brain in favor of the more verbal and numerical left side in their approach to analyses. For centuries, analytical thinking, which involves the mental faculties closely related to the brain's left hemisphere, has been encouraged and emphasized. At the same time, the faculties related to the right side of the brain, the visual and intuitive side, have been regarded as vague and subjective, and therefore unreliable for rigorous scientific activities. However, researchers have shown that although each half of the brain deals with reality in its own way, both appear to use high-
level cognitive modes which involve thinking, reasoning, and complex mental functioning (Muehrcke 1980). For creative thinking to occur, both sides of the brain need to work in concert. In other words, visual thinking complements abstract thinking.

Thinking visually means developing the cognitive skills related to the right side of the brain. It means realizing that "thinking can occur in other than verbal and mathematical modes. Sensory modes of thought, especially the visual mode, are at the very heart of thinking" (Muehrcke 1980, p.10). Visual images are fundamental in information processing regardless of the nature of the reality we are dealing with. Whether it be a physical reality, or the relationship among multiple variables, or a complex mathematical concept, if we can visualize it, it is much easier to understand and can possibly lead us to an insight.

Muehrcke (1980) classified visual images into the following three types: seeing, imagining, and graphic ideation (idea sketching). Seeing is not mere sensory information gathering. It is a creative and active cognitive process. We think as we see things. Muehrcke gives us Watson’s account of the discovery of the double helical structure for DNA as an example of thinking in the context of seeing. The second type of visual image, imagining, has to do with what we construct in our mind’s eye. Dreams are a typical example. Like thinking in the context of seeing, the "flashes of insight" (Muehrcke 1980, p.12), generated from the subconscious can also lead to great discoveries. The third type, graphic ideation or idea sketching, is what we generally do when we want to organize ideas. Rough sketching of ideas help organize and clarify our thoughts. Drawing graphic images itself is part of the mental process. In Muehrcke’s
words, “idea sketching acts as a mirror to reflect the visual mind.” The following quote from MacEachren et al. 1992, p. 99) sums up the fundamental role of visual thinking in information processing within the context of geography:

“Even when dealing with nonspatial relationships, geographers are most comfortable with a depiction that allows them to visualize the relationships and connections that in turn lead to hypotheses about underlying causes for the patterns that become apparent when data are presented in a spatial format.”

With both spatial and nonspatial relationships, the envisioning process therefore requires us to perform certain mental operations to form abstractions. Table 2.1 shows the mental operations specifically related to this visual mode.

The proper exercise of these visual-spatial operations can be facilitated through the use of tools that help us display the visual images of the things that we are interested in. At the simplest level, pen and paper can be used to plot the image. Paper maps are a perfect example of a visualization tool that helps to identify abstract spatial relationships existing among physical entities. Imagine how we would deal with geographic space if there were no maps. As the nature of the data and the relationships among objects get more complex, the tools that assist mental operations need to be more diverse and dynamic. It is from this necessity that we try to develop new visualization techniques and tools that are tailored to the specific nature of the data in which we are concerned.
2.2 Exploratory Data Analysis

The course of scientific investigation can be classified into two broad stages: the exploratory stage and the confirmatory stage (Young et. al. 1988). In the exploratory stage, scientists examine data without a priori notion to form hypotheses which are then tested in the confirmatory stage. Until recently, most of the attention in analysis was given to the confirmatory stage while the importance of the exploratory stage was largely ignored. Hypotheses were mostly generated based on ‘hunches’—although in many cases highly plausible hunches stemming from the analyst’s thorough knowledge of the subject. This lack of emphasis on the hypothesis-forming stage of scientific investigation was exacerbated by the lack of proper methods and tools to support such activities.

1. **Pattern Seeking**
   (a) **Closure**
   * filling-in an incomplete pattern
   * finding a target figure imbedded in a more complex image
   (b) **Matching** one pattern with another as wholes, or a detail-by-detail comparison of two or more patterns
   (c) **Categorizing**—distinguishing objects by recognizing common features. The way we literally invent our world.
   (d) **Pattern completion** by interpolation/extrapolation

2. **Visual Memory**
   Ability to retain visual imagery through a combination of vigorous perception and faithful remembering. Key to cognitive mapping.

3. **Rotating Images**
   Mentally rotating an object in its environment or rotating the point-of-view (viewpoint) in relation to the object.

4. **Dynamic Structures**
   Moving a single object in space or moving several objects in relation to each other. Examples: folding/unfolding operations; tie/untie knots.

5. **Visual Reasoning**
   Moving from concrete to abstract images (visual induction) or from abstract to concrete images (visual deduction)

6. **Visual Synthesis**
   A creative putting together of parts to form a greater whole. Involves manifold operations.

| Table 2.1. Spatial-Visual Operations (after Muehrcke 1980). | 33 |
This situation was fundamentally changed by the development of a new analytical method called exploratory data analysis. Tukey (1977, p. V), in his landmark book Exploratory Data Analysis, states that EDA:

"is about looking at data to see what it seems to say. It concentrates on simple arithmetic and easy-to-draw pictures. It regards whatever appearances we have recognized as partial descriptions, and tries to look beneath them for new insights. Its concern is with appearance, not with confirmation."

In his effort to define EDA more precisely, Good (1983) emphasized the importance of hypotheses formulation with the use of EDA. Basically EDA is a matter of looking for patterns that "we never expected to see (Tukey 1977, p. vi)." According to Good (1983), EDA contributes to two aspects of scientific discovery. One aspect is what he calls "successive deepening," a process in which a hypothesis is formulated and, if it explains enough, is judged approximately correct. Then the results are used to improve it. Examining residuals or treating them as if they were original data in EDA is a good example of this approach.

The other aspect of scientific discovery to which EDA contributes is the use of "divergent thinking", or the avoidance of mental ruts. EDA emphasizes looking at data from several points of view, thereby exposing hidden or unexpected patterns in data. This helps us to ask new questions, leading to the development of new hypotheses. As Tukey (1980, p. 24) aptly put it, "Finding the question is often more important than finding the answer." In light of this, Good's (1983) list of five aims of EDA is worth noting. They are: i) the presentation of data; ii) pattern recognition; iii) hypothesis formulation; iv) to look for hypotheses of greater explicativities; and v) to maximize
expected utility allowing for the guessed costs and delays of computation and thinking. What is consistent throughout these five aims is the emphasis on the use of graphics, pattern finding and hypotheses formulation. They are the essence of EDA.

Why is EDA important to GIS? While there is a notion that Geographic Information Systems are distinguished from automated cartography by their “emphasis on analysis (Goodchild 1987, p. 333),” the lack of analytical capabilities in GIS has been one of the main concerns of GIS researchers and users (for a comprehensive treatment of this topic, see Fotheringham and Rogerson 1994). The problem of incorporating spatial analytical tools into GIS to allow more rigorous and sophisticated analyses has been approached in various ways. Goodchild et al. (1991) classified four types of links that can exist between GIS and spatial analysis: i) stand-alone spatial analysis software; ii) loose coupling of existing GIS software with statistical software (either standard packages or purposely written); iii) close coupling of GIS software with spatial statistical software; and iv) full integration of spatial analysis with GIS. Although not included in the above classification, expanding the analytical functions in existing commercial GIS (e.g., Network module in ARC/INFO, see ESRI 1990) is also a viable option. Efforts made in the ‘coupling’ approach include linking GIS to existing statistical packages (e.g., Ding and Fotheringham 1991) and embedding spatial modeling into existing GIS (e.g., Batty and Xie 1994). In contrast to these more or less ‘loosely coupled’ approaches, some researchers have developed more ‘tightly coupled’ packages allowing dynamic linkages between GIS and the mapping of spatial relationships (e.g., Haslett et al. 1990). On the other hand, others have suggested that a whole new line of spatial
methods which are suitable for GIS environment be developed (Openshaw et al. 1987; Openshaw et al. 1994).

Differences in approaches aside, however, it is generally agreed in the discussions on developing spatial analysis tools in GIS that general data exploratory methods would be of great value within GIS. Several years ago, Walker and Moore (1988, p. 348) suggested that the state-of-the-art GIS is “superbly useful for extracting and describing...[spatial data]” but GIS “sorely lacks... routine capabilities...for developing hypotheses about spatially varying feature attributes.” Their statement is still valid as far as the analytical capabilities of GIS are concerned. To remedy the situation, we need to develop more and more diverse and innovative analytical methods and integrate them with GIS. As Openshaw (1994, p. 83) remarked, “Techniques are wanted that are able to hunt out what might be considered to be localised pattern or ‘database anomalies’ in geographically referenced data but without being told either ‘where’ to look or ‘what’ to look for, or ‘when’ to look.” In his discussion on spatial analysis and GIS, O’Kelly (1994) also emphasized the importance of EDA as the tool of choice of spatial analysts especially when they are confronted with large data sets that are generated in large models and GIS. He stated that “The key technical advance will be in pattern recognition, which intelligently allow the user to sift through the data, reduce dimensionality, find patterns of interest, and then order the GIS to find other instances or similar occurrences” (O’Kelly 1994, p. 74).

Some argue that entirely new and different analytical methods for these types of pattern searching and hypotheses generation be developed (most notably, Openshaw 36
1987). Others contend that we can learn a lot from applying exploratory data analysis methods developed in mainstream statistics (for example, see Haining 1990). Most likely, the problem at hand (the size and dimensionality of the data, and the intent of the research) will dictate which approach is most appropriate for each specific case. Whatever approach we may adopt, however, the more tools we have, the better off we will be.

2.3 Dynamic Graphics and Spatial Data Analysis

2.3.1 Dynamic Graphics and Data Analysis

We can substantially increase the efficiency of the graphical methods described above by creating a dynamic system in which the analysts can interact with the graphics. Dynamic environments are important for the analysis of complex, multidimensional data since the addition of dynamic capabilities to traditional static display greatly increase the power of graphical method. As Huber (1983) aptly noted: “We see more when we interact with the picture—especially if it reacts instantaneously—that when we merely watch.” An excellent overview of dynamic graphics and available methods is given by Becker et. al. (1987). They recognized two important properties of dynamic graphics: direct manipulation of graphical elements on a computer graphics screen; and a virtually instantaneous change of elements in graphics. In a dynamic graphical environment, the data analyst takes an action through manual manipulation of an input device and
something happens, virtually instantaneously, on a computer graphics screen. The analyst can therefore instantly see the result of his action: be it to identify data points with extreme values (labeling) or to change certain parameters of the system. The instant graphic feedback of direct data manipulation helps the analyst easily test various hypotheses or 'explore' other possible lines of investigation which otherwise would require multiple steps and would be clumsy to say the least.

In recent years, the combination of rapidly declining computer hardware costs and increased computing power of personal computers has resulted in many affordable dynamic data visualization packages that can be utilized in virtually every aspect of scientific research. There is no doubt that this trend will continue, and we can expect that dynamic graphics will be used more and more frequently by scientists dealing with complex multidimensional data.

Dynamic Graphics Techniques

In this section, we review some common dynamic graphics techniques. These techniques provide a useful set of methods that can be implemented in the proposed dynamic visualization system.

Identification

In analyzing data with labels, there is frequently a need to identify the label of a data point or a set of points. Identification is an interactive technique to rapidly turn the labels of data points on and off. It is bi-directional in nature. That is, in one direction,
one wants to know the label of a particular point. The other direction is to find the location of the elements in graphics when a label is selected. The former is called \textit{labeling} and the latter \textit{locating}. Identification is a basic but extremely useful operation, since the dynamic environment eliminates the need to look up additional sources of information while minimizing display cluttering.

**Deletion**

The basic purpose of deletion is to eliminate certain graphical elements so that we can better study the remaining elements. In a dynamic graphics system deletion can be performed by pointing to the element to be deleted with a mouse. After the point’s removal, the graph is automatically rescaled and redrawn on the screen. Deletion can be useful for removing outliers for a closer look at the remaining points or for filtering out subsets.

**Linking**

Linking is a method to link corresponding data points on different scatterplots. In conventional static displays, linking is carried out by either drawing lines between corresponding points on the two scatterplots or by using a unique plotting symbol for each point in one plot and the same symbol again in another plot. A third method is using the scatterplot matrix—all pairwise scatterplots arranged in a rectangular array. It is possible to perform a partial linking of points on different scatterplots in the matrix.
because it provides a highly integrated view of the data. In dynamic graphics, linking can be done by selecting a point or a group of points in one scatterplot, which dynamically highlights the corresponding points in other plots.

**Brushing**

Brushing is a dynamic method in which the data analyst moves a rectangle around the screen by moving a mouse (Becker and Cleveland 1987). The rectangle is called the *brush*. The data are displayed in a scatterplot matrix. Simply put, as the analyst moves the brush, the data points that fall inside the brush are highlighted. At the same time, the corresponding data points in the scatterplot matrix also get highlighted.

The power of brushing as an analytical tool comes from the fact that brushing is a collection of dynamic graphical methods for analyzing data in multiple dimensions. There are three main features of brushing: i) an ability to change the shape of the brush; ii) an ability to employ different brushing operations; and iii) an ability to use different paint methods (Becker and Cleveland 1987). Figure 2.2 demonstrates the basic brushing method. The scatterplot matrix shows three variables (hardness, tensile strength, and abrasion loss) as well as the rectangular brush in the (2,1) panel in the scatterplot matrix.
Figure 2.2 The basic brushing method with highlighting. All the points inside the brush in the current panel, as well as the corresponding points in other panels are highlighted (Source: Becker et al. 1987b).
The panel in which the brush is located is called the current (or active) panel. The brushing operation being done is ‘highlight’, that is all points in the current panel that are inside the brush, as well as the corresponding points in the other panels are highlighted.

By combining various brushing methods, we can create several brushing tools that are suitable for different analytical tasks. For example, the ability to change the brush shape provides a means to perform conditioning on one or two variables. Making the brush long and narrow has the effect of fixing the values of one variable to a certain range (Figure 2.3). We can move this brush in the horizontal direction to examine the behavior of the other variables. On the other hand, making the brush a square is the same as conditioning on two variables. Furthermore, this brush shape can be combined with different brushing operations in different paint modes. In Figure 2.4, the square brush is used with a shadow highlight operation. The shadow highlight operation is useful when highlighted points cannot be easily visually detected because of the large number of points. In the shadow highlight, only the points that are highlighted on the active panel appear on the other panels. The two paint modes (lasting and undone) are mainly used to select irregular clusters of points.

Brushing can be extended to other types of graphic displays besides the scatterplot matrix. For example, the combination of scatterplot matrix and rotation was used to display trivariate data by Becker et al. (1988) and Stuetzle (1987). Scatterplot matrix brushing has also been combined with maps to investigate geographic correlation (Monmonier 1989, 1990; Tang 1993).
Figure 2.3 Conditioning on one variable. The long and narrow shape of the brush fixes the values of one variable to a certain range (Source: Becker et al. 1987a).
Figure 2.4 Shadow highlight operation. Only the points that are highlighted in the active panel are displayed in the other panels. (Source: Becker et al. 1987b).
Scaling

Dynamic scaling allows the data analyst to control the aspect ratio (the length of the vertical axis divided by the length of the horizontal axis) of the graph dynamically. Determining the suitable aspect ratio is important because it influences the overall shape of the graph, hiding or revealing important trends in the data. For example, when the slopes of line segments are used to determine the rate of change of $y$ as a function of $x$ in a two-variable graph, the judgment of slope is affected by the aspect ratio. With dynamic scaling, the analyst can easily experiment with different aspect ratios and find the most visually revealing one for the given data.

Rotation

Rotation is used to convey the three-dimensional structure of point clouds through a rapidly changing sequence of two-dimensional views. We can rotate the cloud around one of three axes, define an arbitrary axis, or grab the point cloud and interactively rotate it. Rotation is particularly effective in searching for trivariate structure or examining the local dependence of one variable on the other two.

2.3.2 Review of Selected Dynamic Graphics Systems

This section reviews a few selected dynamic graphics systems that contributed much to the current state of visual data analysis. Most of the early systems were built on custom-made, highly specialized hardwares which greatly limited access to the system.
The advent of affordable, high-resolution color graphics workstations significantly changed the way in which these dynamic tools are created and used.

**An early system**

One early dynamic graphics system was developed by Fowlkes (1971) to make probability plots. In that system, the user could continuously change a shape parameter for the reference distribution on a probability plot displayed on the screen by turning a knob. The system showed the potential of dynamic graphical methods for data analysis (Becker et. al. 1987).

**PRIM-9**

Another very influential early system called PRIM-9 (Fisherkeller et. al. 1975) appeared in the mid 1970s. As the system acronym indicates, PRIM-9 was a set of dynamic tools for Picturing, Rotating, Isolating, and Masking multidimensional data in up to nine dimensions. Implemented on dedicated computer graphics hardware in the Graphics Interpretation Facility of the Stanford Linear Accelerator Center at Stanford University, the system introduced a whole new set of dynamic graphical methods that were adopted in many of the systems built throughout the 1970s and 1980s. Two direct descendants are PRIM-S at the Swiss Federal Institute of Technology, and PRIM-H at Harvard (McDonald 1988). PRIM-9 implemented several important dynamic graphics methods. The following section briefly discusses each of the dynamic graphics methods in PRIM-9, as described by Fisherkeller et al. (1988).
Picturing

Picturing is the ability to look at data from several different directions in multidimensional space. Any two of the current coordinates can be selected as horizontal and vertical axes. For each direction a button is provided to cycle through the coordinates by increments of one.

Rotation

Since we usually do not know which projection we want, continuously controlled rotation is a natural way to search through the multidimensional data space. Rotation is the ability to turn the data so that it can be viewed from any direction. Since specifying a general rotation axis in multidimensional space is complicated, rotation is limited to those associated with pairs of the current coordinates. The rotation process is made easy by two features: rotation reversal and rotation by increasing (accelerating) steps. In rotation reversal, a single button acts as a toggle button to change the direction of the rotation. Rotation in one direction occurs as long as the button is held down. The next time the button is pushed, rotation occurs in the reverse direction. Through rotation reversal, one can easily correct overshoot of the desired projection. Getting the desired projection is aided by varying the speed of rotation in increasing steps. Rotation occurs slowly at first and then accelerates. Combined with the rotation reversal feature, it minimizes the delay time for a large rotation while allowing fine adjustment.
Masking

Masking is the ability to select any subregion of the multidimensional space and display only those data points in that subregion. Masking is tied to the coordinates, so under rotation the data points will enter and leave the masked region. As the mask (specified in the current coordinate as front edge, back edge, and joint) moves back and forth along one of the current projection axes, the data points will disappear and reappear as they leave and reenter the masking region. This forms the basis of the operation called brushing (Becker and Cleveland 1987).

Isolation

Often, there is a need to select a subset of the data for further scrutiny. Isolation is the ability to select an arbitrary subset of the data and limit the operations to this subset only. The data points to be isolated are defined by interactively constructing a mask that includes only those points. Isolation is tied to the data points. In other words, the isolated set of data will remain the same as the rotation occurs, thereby allowing a full range of operations for the subset.

Automatic Projection Pursuit

One of the most interesting functions of PRIM-9 is a sort of "automatic pilot" for rotation, which is called automatic projection pursuit (Friedman and Tukey 1974).
The algorithm assigns to each projection a numerical index (I) that corresponds to the degree of data structuring present in the projection. Starting with the current projection (or any of the two possible principal axes if defined), the algorithm finds the projection corresponding to the first local maximum of the projection index by fixing the horizontal axis while the vertical axis is varied. Next, the search continues with the vertical axis fixed to the value found in the previous step and by varying the horizontal axis. This process continues until either an interesting view appears or until convergence is reached (no change in projection with subsequent searches).

One of the studies in which PRIM-9 was successfully used was the study of diabetes by Reaven and Miller (1979). In their attempt to define the nature of chemical diabetes, they analyzed multivariate medical data from 145 people using PRIM-9 to directly visualize the 3-D shape of the data set. What they found was a shape that resembles a boomerang with two wings and a fat middle. From this picture and from subsequent analyses, they concluded that there is an apparent separation between patients with chemical and overt diabetes and this separation may explain why patients with chemical diabetes rarely develop overt diabetes.

**Orion I**

Orion I is the direct descendant of PRIM-9. Orion I is different from earlier PRIM systems in that it uses relatively inexpensive raster graphics and microprocessor
technology to deliver more computing power (Friedman et al. 1982). While PRIM-9 used a large mainframe computer (IBM 360/91) to perform its calculations and cost millions of dollars to construct, Orion I used a much less expensive hardware platform (a SUN microcomputer with a special purpose arithmetic processor and high-resolution graphics display) (McDonald 1988).

In addition to incorporating the operations provided by PRIM systems, Orion I introduced some new EDA methods as well. The developers of Orion I tried to go beyond three-dimensional scatter plots (available in PRIM systems through the use of real-time rotation) to experiment with ways of looking at more than three variables at a time. To include higher dimensions colors were added to represent the values of additional variables. Projection pursuit models were expanded to allow interactive search for a projection using perception, judgment, and a knowledge of the context of the data. Also, to aid the interpretation of projection pursuit models, multiple windows could be concurrently displayed showing different variables. These windows were linked so that highlighting a data point in one window causes the corresponding data point in another window to be highlighted (MacDonald 1988). This was a significant improvement over the single white dots in a single graphic screen used in earlier PRIM systems.

MacSpin

With the rapid development of computer technology, especially the advent of low cost machines that support graphical interfaces, a new generation of dynamic graphics systems began to appear. Since Apple Macintosh was the first widely available machine
using a graphical interface, many statistical programs were designed for Macintosh. For a review and comparison of some of those programs (Data Desk, Exstatix, Fastat, JMP, StatView, and SuperANOVA), see Best and Morgenstein (1991). Insofar as dynamic graphics are concerned, the most influential of this new generation of software was a program called MacSpin (Donoho et al. 1988). MacSpin has been praised by Tufte (1988, p. 392) as "a genuine masterpiece, a marvelous interactive program." A thorough description of MacSpin has been given by Donoho et al. (1988). The following is a brief summary of MacSpin’s features.

MacSpin uses rotation to display 3-dimensional scatter plots (Figure 2.5) and offers dynamic graphics primitives such as animation, identification, and highlighting. Among the improvements over the previous “one-of-a-kind installations” (Donoho et al. 1988, p.339) are a wide variety of display options, a better user interface and tighter integration between user action and display. The developers of MacSpin placed a great deal of emphasis on user interface and supporting environment such as facilities for data input/output, data editing, graphical hardcopy, data transformation, and variations in the display options. It uses pop-up windows extensively to provide additional information on the data and to minimize screen cluttering. Basic screen layout is maintained throughout most analyses to avoid what is called “constant context switches” (p. 350)—one display (such as a scatterplot) is replaced by another completely different graphic element (such as a menu), then another (back to the scatterplot), thereby occupying user’s short-term memory with incidentals. It is an important point to consider in designing graphical
Figure 2.5 MacSpin display. The top figure shows a typical MacSpin display showing a point clouds. The bottom figure demonstrates the use of rotation. (Source: Donoho et al. 1988).
environments because forcing users to focus on the data is a critical element in graphical data analysis. The single most important contribution of MacSpin is that it freed dynamic graphical analysis from the expensive, dedicated machines in laboratory environments and brought it within easy reach of the public. Now more people can explore their data through graphical means and generate new and fresh ideas.

2.3.3. Dynamic Graphics for Spatial Data Analysis

Dynamic graphical methods has amassed significant interest not only in the aspatial domain but also in the spatial domain. What makes spatial data distinct from other data is that spatial data has an alternative key to access the data. In other words, the data can be accessed either by attribute or by location. Location is based on two continuous dimensions (x, y). There can be a multiplicity of possible conceptual data models for spatial data. In addition, spatial data has a distinctive feature of spatial dependence, the propensity for nearby locations to influence each other and to possess similar attributes (Goodchild 1992). These characteristics of spatial data add much complication to the development of dynamic graphics methods for it. By the same token, the characteristics, especially the inherent multidimensionality of spatial data and its spatial dependence, make dynamic graphics methods intuitively natural for spatial data analysis. Given the complexity of designing a system, it is not surprising to observe that there have not been many systems developed in this area. The following section reviews some selected systems.
SPIDER

One of the first attempts to apply dynamic graphics in the context of spatial data analysis was SPIDER (SPatial Interactive Data ExploreR), a system developed by a group in the Department of Statistics at Trinity College in Dublin, Ireland (Haslett et al. 1990; Haslett et al. 1991; Wills et al. 1990). The developers of SPIDER focused mainly on 'multiple spatial views of the data' and 'dynamic linking of statistical and geographical views'. As they put it, "The key to the tool is the dynamic linking of alternative views of data" (Haslett et al. 1991, p. 240). They made a variety of views available to the user: a histogram view, a map view, a scatter plot matrix view, a moving average view, and a trace view. These multiple views are linked to help identify relationships between variables (relationships between data points and their spatial variations and/or relationships among statistical variables). For example, by linking the histogram view and the map view, a user can select a portion of the histogram to identify the locations of the corresponding data points in the map view. Linking is the key in this environment. "Separately the statistical views are of some value, but that value is enormously enriched when the views are linked" (Haslett et al. 1991, p. 235).

SPIDER also possesses an ability to overlay multiple spatial views of the data. For example, satellite image, line coverage, and point data can be overlaid to give the user the orientation of the data. The images in this case, however, are used as a backdrop for the analysis and cannot be operated on. In this aspect SPIDER is more a statistical tool than a GIS. With its well integrated user interface constructed on a personal computer (Mac II), and its innovative way of linking a spatial view of the data to various
statistical views, SPIDER provides a good example of how dynamic graphics methods and tools benefit exploratory spatial data analysis.

**Polygon Explorer**

The idea of applying dynamic graphics to geographic data was also studied by McDougall (1991, 1992), who examined the possibility of applying commercially available EDA programs to spatial contexts. He also constructed a polygon-based data exploration system called Polygon Explorer. In his study, he first demonstrated the use of commercial software (JMP) to display irregularly spaced point data and raster data for exploratory analysis. These data can be easily imported to EDA programs through JMP's data transformation features. After imported, the data can be displayed as a form of scatterplot (for irregularly spaced point data) or spin plot (for raster data). Multiple linked views are possible for exploration.

For polygon data, McDougall developed a prototype interface, called Polygon Explorer for the Macintosh computer (McDougall 1992). This program includes a map display, some basic statistical graphs (bar charts, histograms, and a scatterplot) and a capability for cluster analysis (Figure 2.6).

These elements in the display are linked together. To further assist exploration, two features are added to the basic display: an ability to transform or re-express a variable to normalize the distribution; and a feature called the “Tukey yard” (following Tukey) or schematic box plot in EDA, which provides a rectangle in the scatterplot to give a visual cue to outliers. Finally, the prototype was used to assess how one widely
Figure 2.6 Polygon Explorer display. The display shows towns in Massachusetts with high unemployment. The rectangle in the scatterplot box is the “Tukey Yard.” (Source: MacDougall, 1992).
used statistical analysis (cluster analysis) can be implemented and used in a dynamic graphics environment. As a prototype, this program is quite limited in its data handling capability (one categorical and two continuous variables) and does not provide extensive analytical functions. However, as the author points out, the initial objective of the prototype was to determine the extent to which dynamic visualization could be achieved for hundreds of polygon data sets. As such, McDougall's study shows yet another possible research area that can benefit from a dynamic graphics environment.

**Dynamic Graphics for Network Visualization**

A series of research on creating dynamic graphical tools for network visualization has been done by Becker et. al. (1990a, 1990b). Unlike the research previously reviewed, Becker et. al.'s research focuses on the more "dynamic" nature of display by utilizing an advanced computer graphics environment. The term "dynamic" in this case means that a user can directly adjust the screen display with a pointing device to get instant visual feedback. In their research, Becker and his colleagues developed two complementary tools (dynamic link display and dynamic node display) that use dynamic graphics to display and manipulate network data on a high-performance graphical workstation (Figure 2.7 and 2.8). In the development process, they introduced several interesting techniques and one important concept, parameter focusing. This concept is explained below.

Static maps (as opposed to dynamic maps) are easily overloaded with too much information. By adding dynamic graphics capabilities to these static map displays,
Figure 2.7 Dynamic link display. The map shows only the routes on which airline delays are greater than 100 minutes. The threshold is set by the slider on the lower scale. (Source: Becker et al. 1990).

Figure 2.8 Dynamic network node display. The rectangles encode the average air traffic delays at each node. The horizontal extent of the rectangle shows delays for inbound traffic, the vertical extent for outbound traffic. The time is dynamically controlled by the slide bar below the map which is set to 14:30. (Source: Becker et al. 1990).
Becker and his colleagues produced a tool that manages information by focusing the display in several different ways, thereby helping the user identify geographical patterns. The key point is determining dynamically the correct set of parameters to produce a meaningful map. Dynamic graphics are particularly useful when: i) the space of all possible parameter values is large; ii) most combinations of the parameters do not lead to understandable displays; and iii) the maps are sensitive to particular values of parameters.

Becker et al. identified six classes of parameters to be set in the focusing process. They are:

a) statistics: the statistics to be displayed. Alternate statistics can be displayed in a back and forth fashion to allow comparison. Also, transformation of data values may be needed.

b) levels: the range of data values to be displayed. Any subset of data can be displayed. A generalization of this technique is brushing.

c) geography/topology: this parameter allows display of an appropriate subset of the data based on network geography or topology. Geographic zoom operation and control of network topology by deactivating (or reactivating) any nodes and associated links are used.

d) time: displayed time can be varied at will, providing snapshots at particular times or producing a movie-like effect.

e) aggregation: aggregation of statistics over geographical regions or a logical subset of a network.
f) size: the parameter controlling the overall size of the symbols drawn on the map. Line shortening (the line connecting two locations is drawn only part-way) for a link map and symbol sizing for a node map are used.

Becker et. al.'s (1990a, 1990b) research is very important because it is the first serious and systematic approach for network visualization where the advantage of a dynamic graphics environment is recognized and demonstrated through implementation. It also brought out important techniques and concepts to carry out dynamic visualization.

However, there are some limitations in Becker et al.'s study (1990a, 1990b) as well. First of all, some of the techniques are not actually implemented (e.g., geographic aggregation and brushing in particular). Secondly, the display of statistical information, particularly link-related statistics are not fully incorporated. Related to this is the lack of statistical graphs that can show the relationships between variables found to be interesting on the map. If the visualization system is to be an exploratory tool, then the researcher should have the means to investigate possible relationships thoroughly. Thirdly, the difference between point-to-point data and a true network is not considered. Despite the fact that the term *network* was used, and of course, node/link structure was discussed, the link map display shows only direct point-to-point connections, such as telephone calls. However, in many networks, the characteristics of the network itself plays a significant role in determining the magnitude of flow on it. We need to treat the network itself as an entity, not merely as a connection between nodes. Consequently, the link map display implemented may be suitable for the same kind of point-to-point data (e.g., migration,
airline traffic), but may not be appropriate for other types of networks. There is great room to further develop graphic displays so that they can handle more realistic networks.

Another very important study with regard to the current research was done by Gou (1993), who developed a prototype system called SEFLOW (Scientific visualization and Exploratory data analysis of large spatial FLOW data; see Figure 2.9). Interstate migration data among U.S. states for three time periods (1955-60; 1965-70; and 1975-80) were used to demonstrate the system's capabilities. Also, for the 1975-80 period, socioeconomic variables such as population density, unemployment, and per capita income were used to provide the necessary basis for statistical analysis. Implemented in the X windows system, the prototype consisted of five modules, each handling specific tasks. For example, the Forward Brushing Module allows brushing of data points in a scatterplot matrix. This matrix is linked to a map display, which shows the migration flows among states that are inside the brush. It is what Monmonier (1989) called 'geographic brushing.' Gou extended the idea of geographic brushing to include what he calls 'backward brushing.' In backward brushing, the analyst brushes a specific geographic space to reveal patterns of the highlighted data points in the scatterplot matrix. This makes the brushing a more complete and more versatile tool for data analysis.

Gou's (1993) study shows the potential benefit of developing a dynamic system suitable for flow data analysis. Dynamic graphics techniques, such as brushing and linking, are seamlessly integrated with cartographic display tailored to suit the need for the cartographic representation of flow data. Dynamic threshold setting of flow volume
Figure 2.9 A SEFLOW display showing the Cartographic Representation Module. SEFLOW uses various flow representation methods. The left map represents the flows with lines and numbers while the right map uses varying arrow widths proportional to the flow magnitudes. (Source: Gou 1993)
with two slide bars makes it easy to interactively create thought-provoking flow maps. In addition, the system's dynamic nature and various representational schemes for flow data allow the interactive examination of migration patterns while minimizing screen cluttering. In traditional paper- and table-based analysis, this kind of interactive pattern-searching would have been almost impossible.

Although quite successful in terms of implementing a new breed of dynamic tools for flow data analysis, the prototype has some drawbacks. First and foremost, Gou's study limits itself to a single type of flow data, namely point-to-point flow data. Therefore, the various cartographic representation schemes for flow data, such as using lines with numbers and varying arrow width proportional to the flow magnitudes, must be modified if they are to be used in network-based flows. Secondly, as Gou (1993) states in the study, the prototype does not have the capability to handle diverse input formats that are available for analysis. Thirdly, the volume of data that can be handled by the prototype is relatively small—only state level migration could be processed in a reasonable response time. The ability to deal with much larger volumes of data at finer spatial levels is desired. Finally, there are some minor drawbacks: the modules are structured in such a way that there is an abrupt change in display layout between the cartographic representation module and the other modules (forcing context switch); there are no subset operations for aspatial and spatial data; and no spatial zoom function is available. Even with these drawbacks, however, Gou's study provides us with an important reference point from which we can design a more flexible and dynamic data analysis system.
2.4 Geographic Visualization

The dynamic graphics method is not the only way of visualizing spatial data. In fact, the graphic portrayal of spatial information has been done for centuries by cartographers and geographers. MacEachren (1982) defined geographic visualization as "the use of concrete visual representations—whether on paper or through computer displays or other media—to make spatial contexts and problems visible, so as to engage the most powerful human information-processing abilities, those associated with vision." In this sense, geographic visualization is nothing new.

However, what sets the new geographic (or cartographic) visualization apart from this long tradition of the graphic portrayal of spatial information is the availability of new tools that can ease the creation of graphics as well as open up new means of portraying the data. With the rapid adoption of GIS in our dealings with spatial data, we are experiencing a rapid transition from the old static medium to the new dynamic medium. As MacEachren and Ganter (1990, p. 67) put it, "More recent development in geographic information systems are providing a testing ground for a variety of new cartographic and graphic tools for scientific visualization. One of the strongest links between cartography and geographic information systems is through cartographic visualization tools and their potential to increase the data synthesis and analysis capabilities of GIS." This section reviews some of the many visualization efforts made in the fields of cartography and geography.
2.4.1 Analysis of complex spatial data through visual methods

Cartographic animation

The traditional emphasis on visual communication in cartography, especially in the search for the 'optimal map,' has prompted some attempts to present the complex nature of spatial data in the form of animation and interactive visualization. Animation, in particular, has been regarded as a way of overcoming the static nature of printed maps and of addressing the dynamic element of change over time. As early as 1959, Thrower (1959) emphasized the potential role of animation in cartography in terms of visualizing the temporal process and believed that popular animation was just around the corner. A few years later, Cornwell and Robinson (1966, p. 82) investigated the possibilities for computer animated films in cartography, stating that "the prospect for the use of this versatile technique in accomplishing creative dynamic cartography is limited only by the imagination." However, as Campbell and Egbert (1990) pointed out, animated cartography is not much more widespread now than 20 some years ago when Tobler (1970) published the result of the first computer animated map sequence.

Tobler's (1970) well-known computer movie of the urban growth of the Detroit region was the first attempt to utilize the animation technique. In that study, the main purpose of the movie representations of the simulated population distribution in the Detroit region was "to provide insights, mostly of an intuitive rather than a formal nature, into the dynamics of urban growth (Tobler 1970, p. 238)."
The significance of the animation technique in presenting the result of complex spatial data was noted in a series of studies by Moellering (Moellering 1976, 1980a, 1980b). In his study of traffic crashes, he used computer animated film to display the objects of analysis in a dynamic temporal setting (Moellering 1976). He recognized two fundamental uses of animation in geographical settings: one as a cognitive device to facilitate the perception of spatio-temporal dynamics of the process; the other as a heuristic device to aid in suggesting hypotheses. These ideas have been further developed into the ‘direct control of animation sequence’ (Moellering 1980a) and the concept of the ‘real-time cartographic system’ (Moellering 1980b). Moellering suggested that the rapid development of cartographic information processing and manipulation techniques made Muehrcke’s (1972) exposition of the cartographic processing system (Figure 2.10) no longer adequate to describe the situation.

![Figure 2.10: Muehrcke's Cartographic Process.](image)

*Figure 2.10: Muehrcke's Cartographic Process. (after Muehrcke 1972)*
To solve the problem, he developed the concept of the ‘real’ and ‘virtual map’ (Moellering 1980b). Maps are classified into four categories: real map, virtual map type I, type II, and type III. The classification criteria are i) “Is the map permanent with tangible reality?” and ii) “Is it directly viewable?” He recognized the importance of transformations between virtual map type I (a map with no permanent tangible reality but directly viewable as a cartographic image, e.g., CRT map) and virtual map type III (underlying data which has neither tangible reality nor a directly viewable cartographic image, but readily convertible to other classes of virtual maps) in interactive systems. By developing very efficient algorithms for this type of transformation in a system, Moellering argued that we can dynamically ‘explore’ the cartographic surface (Moellering 1980b).

It was not until the end of the 1980s, however, that cartographic animation was rediscovered when cartographers recognized the potential of animation for the depiction and exploration of spatial and statistical relationships and patterns (Karl 1992). The subject of the potential benefit of cartographic animation was again brought up by Campbell and Egbert (1990). New animation techniques for spatio-temporal data (Monmonier 1989, 1990) and geoscientific processes (DiBiase et. al. 1991) were suggested. In addition, many programs that can be run from personal computers to create cartographic animation appeared on the market (for discussions of some of these programs and design issues, see Gersmehl 1990).

The combination of the renewed attention to and easy accessibility of animation softwares has brought exciting opportunities for cartographic animation. However, many
fundamental problems in cartographic animation remain to be solved. For example, there is no consensus on the exact definition of cartographic animation. Many design issues such as symbolization in animated maps, animation speed, legend design, and control of the animation sequence have not been properly addressed. While the studies aforementioned have certainly contributed to broaden the possible use of animation in cartography, they deal only with specific animation techniques for specific cartographic information. More comprehensive studies on the design, perception and production of animated map sequences is needed but not available (Karl 1992).

Interactive visual analysis of spatial data

The growth in interest in cartographic animation has further expanded into the area of interactive visualization. While animation has been largely regarded as more of an illustrative tool, the interactive visualization system emphasizes the investigative aspect of cartographic visualization. In cartographic animation, we can do one of the following three things: animate space—the process of panning and zooming around and into a large two-dimensional static image; animate time—the map is held still and the action played out upon it; or animate a combination of both (Dorling 1992). Traditional cartographic animation focused primarily on the second method of animation, that is, to record the changes in data through time. Changes are usually represented by either moving the symbols around the map or changing their color. These kinds of representations are suitable to illustrate the points that the producer of the animation wishes to make, but hardly serve to find unknown patterns in the data.
In dealing with visualization of spatio-temporal data, however, this method of animating time continues to be dominant (Gould 1989; MacEachren and DiBiase 1991; Okazaki 1993). Improving upon the technique of sequencing maps to focus the viewer's attention on a meaningfully ordered set of cartographic subpatterns (Slocum et al. 1990), a new animation technique was suggested by Monmonier (1990). Called 'atlas touring,' this technique involves creating a succession of views through the use of a 'graphic script,' composed using basic sequences called 'graphic phrase' (Monmonier 1989b, 1992). A graphic phrase is a programmed graphic sequence that acts as a building block to construct a graphic script. Using these graphic building blocks, the script developer can put together highly sophisticated animation sequences that highlight trends or anomalies in the data. As a customized guided tour of complex spatio-temporal data, atlas touring can effectively focus the viewer's attention on the main points of the presentation. The graphic script or atlas touring, however, serves mainly as a narrative tool to communicate complex information in a limited time. Only a limited control of animation (such as pausing to understand a complex scene) is allowed, and the user cannot deviate from the pre-programmed sequences of animation. A truly interactive tool would allow the user to freely roam inside the tour, going back and forth in time, and branching out as he/she finds interesting phenomena in the data even though the existence of such patterns were not foreseen by the script developers.

Animating space, on the other hand, can be much more useful in exploring data. Unfortunately, however, it has been given inadequate attention despite its value and adaptability as an exploratory tool (Dorling 1992). To explore, we need the ability to
freely search through the data with the full control of its display characteristics: the variables to be examined, the position, color, scale, and aspect of them, and the sequence with which they are chained together. Replacing static maps with a dynamic map means we eliminate hundreds of static maps with various insets and replace them with a single coherent spatial image that can be manipulated at will to suit our needs. In their comprehensive review of animation in scientific visualization, DiBiase et al. (1992) also noted that control over visual variables (position, size, value, texture, hue, orientation, and shape; see Bertin 1983), and dynamic variables (duration, rate of change, and order) are critical to fully exploit the potential of cartographic animation.

In developing the interactive visualization systems, many researchers recognized the potential of applying EDA methods and visualization techniques to geographic data (for example, Monmonier 1989, 1990; MacEachren and Ganter 1990). As a part of ongoing research to develop a methodology for the exploratory analysis of spatio-temporal data, Sandhu (1990) has created a prototype system called the Planetary Data Visualization System (PDVS). This system is special because it deals with a very large volume of global data (earthquake data). Using a supercomputer, the system allows the interactive visualization of global earthquake data with the use of a special temporal symbolization scheme and a basic set of EDA methods.

Noticing that most traditional statistical techniques do not adequately address 'geographic correlation'—the extent to which two variables are similar in spatial pattern—Monmonier (1989) introduced a technique that he called 'geographic brushing.' Based on a dynamic graphics method called 'scatterplot brushing' (Becker and Cleveland
geographic brushing links a map display to a scatterplot matrix. With the link, the otherwise aspatial display of statistical relationships among variables becomes an efficient means of geographic exploration, raising questions and suggesting hypotheses. This geographic brushing concept can be expanded to incorporate time-series data, making it a temporal brush (Monmonier 1990), where the user can control the time period for which the data are displayed with a temporal scrollbar. Therefore, in this high-interaction graphics environment (Becker et al. 1987), a user is equipped with a multitude of tools to conduct the search for meaningful patterns in the data. These tools include statistical tools such as a scatterplot matrix with brushing capability, a tool for spatial relationships such as a geographic brush, and a tool for temporal analysis such as a temporal brush.

Another way of improving the search process is to use human ability to deal with multiple input modes (text, graphics, audio, and motion video). Multimedia presentation is a good example of reducing the complexity of the overall message by spreading information among several sensory modalities. The explosive growth in multimedia titles in the personal computer software industry certainly demonstrates that multimedia presentation has now become a feasible (and preferred) means of conveying information. Market analysts expect that annual sales of CD-ROM titles will exceed 37 million units by 1995 (Colligan 1994, quoted from DiBiase 1994). One such use of multimedia in cartographic presentation has been reported by DiBiase (1994). As part of the Multimedia Encyclopedia CD-ROM title (New Grolier Multimedia Encyclopedia), 15 animated maps with sound on such topics as the American Revolutionary War, World
War II, and Magellan's circumnavigation of the world were produced. Despite various technical roadblocks in designing and producing multimedia maps, DiBiase reports that the process was satisfactory. With apologies to Campbell and Egbert (1990) who stated that cartographic animation is merely scratching the surface, he declared that "animated cartography is not just scratching the surface anymore (DiBiase 1994, p. 7)."

Simple multimedia, however, does not allow viewers to control the information flow. What they see is predetermined by the designers of the presentation. To overcome this lack of navigational control, the use of hypermedia was proposed (Buttenfield and Weber 1994). Hypermedia has an advantage over simple multimedia in that it links the multiple modes transparently, permitting associative browsing of data. In other words, users can jump directly to the relevant portion of the data for a closer look, and either come back to the previous stage of the data search or follow through the new thread. This kind of capability requires that the tool be 'proactive' (Buttenfield 1993) rather than 'interactive.' While interactivity provides the capabilities to respond to system actions that are anticipated by system designers (through system dialogue boxes and menus), proactive computing simulates a system responsive to commands and queries that may not have been anticipated by system designers (Buttenfield 1993). "When visual tools are proactive, users initiate queries and steer data presentation in a manner consistent with the associative power of the human intellect (Buttenfield and Weber 1994, p. 8)."

The ability to steer data presentation has been recognized as "the most exciting potential [use] of visualization tools" in modeling (McCormick et al. 1987). As an example, Buttenfield and Weber (1994) developed a prototype hypermedia system using
a biogeographical database (radial growth in trembling aspens). The system implements cartographic display, as well as statistical graphics or numeric tabulation. Additional information including text, range maps for North America, and photographic images are available at all times. Animation of aspen growth runs as the default and can be paused. Zooming and panning is possible. With zoom, automatic scale change is performed. Users can follow their own lead to view maps, statistics, text or metadata. In that sense, the system is proactive. However, as the authors point out, it is not fully proactive because it lacks a scripting language for viewers to develop their own links and command structure (Buttenfield and Weber 1994). Nevertheless, their prototype demonstrates the possibility of a whole new breed of geographic visualization systems where the user dictates what the system shows and how the system shows it rather than the system designer dictating what the user sees.

Since MacEachren and Ganter’s (1990) powerful arguments for cartographic visualization, many geographic visualization studies have appeared in the literature. A partial list of studies found in the literature covers a wide range of topics, such as using visualization techniques for environmental modeling (Mitasona et al. 1993), using visualization tools for regionalization (Hancock 1993), and visualization of the spatial structure of the social geography of a nation (Dorling 1993).

Another research area where the visualization techniques could be proven useful is visualization of spatial data quality. The quality of spatial data and databases has been a major concern for developers and users of GIS (Chrisman 1983). The volume and variety of available spatial data are increasing rapidly. At the same time GISs are
becoming more clever with an increasing number of automated functions available at the users’ fingertips. The ease of use and increasing level of sophistication, however, also increases the possibility of reaching wrong conclusions and poor decisions when the quality of the data and the fitness of the applied model are not adequately understood. Therefore, for a GIS to be used as a true ‘spatial understanding support system’ (SUSS, see Couclelis 1991), it should be able to represent the quality of the spatial data and the database that the decision is based on. In this regard, visualization can be used as a method for capturing, interpreting, and communicating quality information.

It is hard to define the exact meaning of the ‘quality’ of spatial data. The exact definition of quality has to be determined in the context of each application. It is, however, still useful to define it in general terms. When presenting NCGIA’s position on its initiative on spatial data quality visualization, Buttenfield and Beard (1991) indicated that the quality of spatial information “relates to accuracy, error, consistency, and reliability.” Accuracy refers to the discrepancy between measurement and a model. It is worth noting that the Proposed Standard for Digital Cartographic Data Quality (Moellering et al. 1988) adopts three accuracy measures (positional accuracy, attribute accuracy, and consistency) as part of the quality specification. Error is defined as the discrepancy between measurement and true value: errors can come from data collection (source error), data processing (process error), or the appropriateness of the application (use error) (Beard 1989).

On the other hand, MacEachren (1992) advocates using the term ‘uncertainty’ instead of ‘quality’ because analysts never know the precise amount of error in any
particular data. They are rather uncertain about the characteristics of data. The uncertainty, according to MacEachren (1992), comes from three sources: variability due to spatial aggregation, attribute aggregation, and uncertainty about temporal information. Regardless of which definition is used, however, the main issues are the following: i) how to capture and store quality information in a digital database; and ii) once the quality information is stored, how we can make it available to the user. Some possible research themes on these issues are found in Buttenfield and Beard (1991).

In an effort to develop a framework for spatial data quality visualization, application of the formalization method (through algebraic specification) has been suggested (Clapham and Beard 1991; Clapham 1992). In this case, the strategy is to construct a formal specification of two problem domains: data quality elements and the visual variables that encode quality information. The goal is to develop a clear, rigorous, implementation-independent definition of the problem components, establish a basis for the evaluation of graphic products, and provide a framework for integrating visualization tools through automatic synthesis or selection of the graphic representations of quality information within a GIS (Clapham and Beard 1991). If rigorously defined and well implemented, the formalization method will certainly help to define the problem and assist in selecting an appropriate visual representation. However, at present, the level of formalization is far from what is needed for successfully implementing the mapping operations from one domain (data quality) to the other (visualization).

Developing a systematic scheme for visual representation requires us to examine available visual variables and to determine which visual variables are relevant for the
given task. The possible use of Bertin’s visual variable in the context of spatial data quality visualization was examined by MacEachren (1992). Although most of Bertin’s graphic variables are found to have some degree of logical match with particular categories of uncertainty information, MacEachren indicates that the most logical graphic variable for depicting uncertainty is color saturation. Saturation can be varied from pure hues for very certain information to unsaturated hues for uncertain information. Another variable that can be useful is what he calls “focus.” “Presenting data ‘out of focus’ (as you would see it with an out-of-focus camera), or simply at lower spatial resolution, might be an ideal way to depict uncertainty” (MacEachren 1992, p. 14). Furthermore, a dynamic graphics environment provides an additional opportunity to make use of dynamic variables (duration, order, and rate of change; see DiBiase et al. 1992).

Some of these ideas have been implemented in a system called R-VIS (Reliability-VISualization) by MacEachren et al. (1993). R-VIS uses a variable-magnitude graduate symbol (triangle) to represent metadata. By making it dynamic, users can bring the quality information into the foreground as it is needed and put it back into the background if necessary. Another dynamic tool, a slider widget, dynamically sets the threshold value, thus making it easy to highlight the specific subsection of data to examine the possible presence of uncertainty. In addition, temporal uncertainty is depicted by animation (the duration of map displays is proportional to temporal certainty).

Veregin et al. (1993) also took advantage of a dynamic environment in their effort to design a data quality visualization system for choropleth maps. To circumvent the
difficulty of combining map and data quality information in a single map, they used multiple windows linked dynamically to display map and data quality information. One noble feature used in their system is the use of auditory variables ("Geiger counter" metaphor) to complement the visual portrayal of data quality information. Mimicking a Geiger counter, which indicates radiation levels with the frequency of "clicks," the number of clicks emitted when a mouse button is pressed is proportional to the data quality value at the current mouse location.

As discussed earlier, the term 'quality' is application dependent. Each application should define the quality of spatial data based on available uncertainty information, a proposed data model, and analysis methods. Studies tend to focus on very specific types of quality information and its representation issues. For example, Mackaness and Beard (1993) examined various methods to display the reliability of interpolated values and suggested several techniques: displaying sample point locations, values, and accuracy; combining kriged values with error estimates; and overlaying the estimated variance map generated as a function of interpolation itself (kriging was used as an example) onto the interpolated map. In dealing with soil map boundary uncertainties, on the other hand, Hassen (1993) reports that the changes in hue and saturation levels can be used effectively to convey uncertainties to soil map users.
CHAPTER 3

A FRAMEWORK FOR ANALYZING LINEAR POINT DATA

Traffic accident data contains two major components. The first component is the underlying street network system against which accident location is recorded. The second component is the characteristics of the traffic accident itself. This includes the location and time of each accident, as well as all of the attribute data associated with each accident. Therefore, developing a dynamic visualization system for such data requires that both components be properly handled and displayed.

This chapter provides a conceptual framework for implementing an interactive visualization system for analyzing linear point data. The framework is based upon an examination of the characteristics of the underlying street network and the types of traffic accident data that occurred on it. The examination of network and accident data characteristics provides us with the possible candidates that are suitable for visualization.

From these possible candidates, a design strategy for the visualization system can be developed. Each of the elements in the visualization system is structured to perform
specific visualization tasks while maintaining a consistent user interface to form a synergetic environment. Following the conceptual design of the system, user requirements for the visualization system are considered. Then the analysis of the accident data and the possible visual representation of them based upon Bertin’s (1983) visual variables is presented. Finally, this chapter introduces visualization tools that are specifically designed for this type of data and shows how the tools are integrated into the system to achieve a high degree of modularity as well as flexibility and expandability.

3.1 Building a Conceptual Framework for an Interactive Visualization System

The goal of this research is to develop an interactive visualization system in which researchers can explore the data to: i) find hidden patterns that are not easily detectable without the help of an intuitive, fast-feedback spatial and statistical display; ii) detect temporal anomalies (patterns, or trends) through visual examination so that subsequent research can be focused on the time-periods identified; and iii) perform statistical tests to detect the existence of space-time clustering in the data.

3.1.1 Visual representation of point patterns.

It is necessary to develop better symbolic representation of points on a graphics display. One of the interesting methods for representing linear point patterns is the strip map available in Arc/Info (ESRI, 1992). It shows the accident location on the map with some of the attribute data symbolized in various ways. At the same time, it shows the
temporal dimension of the data (e.g., day of the week). However, to be able to generate a display, the user of the system has to deal with multiple obstacles that can significantly limit the system’s use as an interactive tool. Not only must the user have in-depth knowledge of the system to create a display, but the complexities of symbol selection and the command-driven nature of the system practically eliminates the possibility of the system being used other than as part of a customized application. These shortcomings notwithstanding, however, the strip map in the Arc/Info system provides us a useful visual representation method. Therefore, we need to examine the available representation schemes and evaluate how efficient they are in communicating the complex information embedded in the data. New and efficient symbolization schemes for this kind of data need to be developed.

3.1.2 Treatment of the temporal dimension.

In addition to developing proper symbolization schemes for point data, the problem of how to treat the temporal component in the data needs to be addressed. The temporal dimension can be dealt with in two different ways. First, to detect any unusual clustering that occurs in the temporal domain—temporal anomalies—we can utilize the techniques developed in cartographic animation. One such use of the cartographic animation technique has been demonstrated by Moellering (1976) in his study of fatal traffic accidents. Another approach is to use a ‘temporal slide bar’ that allows the user to interactively control the sequence of display. This technique has been successfully used with area-based spatio-temporal data by Tang (1993).
Second, to examine the space-time interaction, the statistical tests for space-time clustering such as Knox’s contingency table approach (Knox, 1964) or Mantel’s generalized regression approach (Mantel, 1967) can be used to test the randomness of point patterns in the space-time domain. This is different from the detection of pure temporal anomalies because the techniques used to test the existence of space-time interaction are not sensitive to clustering in either of the dimensions alone. Therefore, with the space-time clustering methods, the question is whether the points close to each other in space are also close in time.

3.1.3 Conceptual Framework

To develop a theoretically sound basis for the developed analytical tool, the concepts and techniques in each of the aforementioned fields must be integrated into a seamless analytical system. Figure 3.1 shows the conceptual framework for the prototype system. This system takes the spatio-temporal linear point data as input. The visualization system analyzes the input data through two visualization modules: temporal visualization and spatial data visualization.

The core of temporal visualization is a tool for interactive control of the temporal display of the data. Based upon the methods developed in cartographic animation, the interactive control tool can provide a flexible method to search through the temporal dimension of the data.
Figure 3.1 Conceptual framework for the prototype visualization system
The spatial data visualization tool consists of two distinct but interrelated submodules. The network visualization module can display various network-related spatial and aspatial information. The physical characteristics of each link (for example, lane width) or categorical attributes (the functional class of the link), as well as node-related information can be displayed. The point visualization module handles the display of the point pattern through various representation schemes. These two modules are tightly linked so that an operation on one module can be directly reflected in the other display if needed.

The outcome of this visualization system is the detection of possible temporal, spatial, or space-time clustering in the mapped pattern. Based on the evidence of possible clustering, the researcher can proceed to perform various spatial statistical tests. A set of spatial statistical tools tailored for linear point data can help the researcher carry out these tests. At the same time, a new hypothesis can be formed either directly from the evidence of possible clustering or from the rejection/non-rejection result of a spatial statistical test. In conclusion, the visualization system successfully integrates the concepts and methods developed in SV, EDA, dynamic graphics and spatial statistics.

3.2 Designing an interactive visualization system

This section identifies the building blocks for designing a linear point data visualization system. In general, this identification process starts with an assessment of user requirements. Although a more formal assessment of user requirements is preferable
when designing a system, the limited resources available for this research prohibits the use of the formal method. However, we can identify a generic set of user requirements by asking questions such as: what kind of tasks would analysts want to do with traffic accident data?; what kind of tools should be available to assist them to perform such tasks?; and what type of interface would be suitable to provide a simple and easy-to-use launching pad for those tasks? Through the examination of the data, the following requirements can be identified.

The system should be able to:

1) Efficiently display various characteristics of the road network.

2) Efficiently display the multiple attributes in the point data in the spatial display to maximize the understanding of the data characteristics.

3) Provide an easy way of selecting and displaying spatial and attribute information in the data (zoom and pan, setting a specific time period to display).

4) Provide an easy way to control the spatial extent and scale of the display.

5) Provide an easy way of selecting one of the available symbolization schemes.

6) Provide a way to represent the temporal component of the data.

7) Make graphical techniques developed in EDA available.

8) Provide an entry point to access external statistical packages.

9) Provide an easy-to-use interface utilizing a menu-based, point-and-click type of navigation.
10) Allow the dynamic selection of display variables and a display method, i.e., the selection can be changed while maintaining all other aspects of display.

11) Respond to the user action in an interactive fashion.

3.2.1 Functional components of the system

Based on the user requirements above, we can identify three major functional components of the system. They are: i) filtering the data; ii) visual representation methods; and iii) handling the temporal component in the data.

Data Filtering

The user should be able to filter the data based on its characteristics. An ability to isolate subgroups and to deal with each individual subgroup is an important technique in exploratory data analysis, especially in a dynamic environment. Subsets can be isolated based on the following three characteristics: spatial, temporal and aspatial attributes.

Spatially, we may want to limit our analysis to a specific area for an in-depth look. The size of the display window will usually make the details in small areas impossible to decipher. For example, a downtown street network is usually considerably more dense than the rest of the network. We need to be able to zoom in and out to see details and overall patterns in the data. In addition, if the display window cannot cover the whole area of interest, we need a panning ability. Panning eliminates the need to repeatedly zoom in and out to see the desired area. Zoom and pan functions could be considered spatial filtering.
*Temporal filtering* creates a subset based on temporal constraints, such as specific year, month, or time of day. We frequently want to see what the rush-hour traffic accident pattern looks like and how it differs from non-rush-hour patterns. We may also want to see how the pattern changed after a major improvement of network structure, such as building a new bridge. The ability to take an instant snapshot of accident locations at a specific time and compare the snapshots at different times will enhance the understanding of overall accident patterns through time, as well as aid in detecting temporal anomalies.

The data can also be filtered based on attributes. For example, we may want to confine our attention to traffic accidents that occurred only on major roads such as highways and main arteries. This type of filtering is based on one of the attributes associated with links. Or we may want to see only the pattern of accidents that involved fatalities. By dividing up the data, we will be able to see the differences in accident patterns among different types of data. This capability—*attribute filtering*—will help us narrow our focus on certain attributes of the accident data that exhibit interesting patterns.

**Visualization representation: visual variables and point data**

The effectiveness of a visualization relies heavily on the selection of the appropriate graphic representation for the data that need to be displayed. Since Bertin (1983) devised a systematic approach to graphic symbolization in his *Semiology of Graphics*, his list of visual variables has been used in cartographic design research. The seven variables he identified include position, size, color value, texture, color hue,
orientation, and shape. McCleary (1983) applied Bertin’s visual variables to cartographic data and its three major data types—point, line, and area data—and examined the strengths and weaknesses of each symbolization method for particular data types. Similarly, DiBiase et al. (1992) evaluated the effectiveness of Bertin’s visual variables with respect to measurement levels (nominal, ordinal, and interval/ratio). Figure 3.2 shows Bertin’s seven visual variables applied to cartographic data types. Based on Bertin’s list of visual variables and some additional symbols that are suitable for traffic accident data, Table 3.1 shows the possible list of visual variables that can be used to represent traffic accident data.

Handling of the temporal component of the data

How data changes through time can be one of the most basic concerns for spatial analysts. The rush hour traffic accident pattern, for example, can be significantly different from that of the non-rush hour. We can also easily assume that the spatial patterns of night time traffic accidents would show some dissimilarity from day time patterns. In the geographic visualization arena, animation techniques are frequently used to represent temporal changes.
### Figure 3.2 Visual variables applied to cartographic data types. (Source: DiBiase et al. 1992.)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Point features</th>
<th>Line features</th>
<th>Area features</th>
<th>Nominal data</th>
<th>Ordinal data</th>
<th>Interval/Ratio data</th>
</tr>
</thead>
<tbody>
<tr>
<td>POSITION</td>
<td><img src="image" alt="Position" /></td>
<td><img src="image" alt="Position" /></td>
<td><img src="image" alt="Position" /></td>
<td>Effective</td>
<td>Effective</td>
<td>Effective</td>
</tr>
<tr>
<td>SIZE</td>
<td><img src="image" alt="Size" /></td>
<td><img src="image" alt="Size" /></td>
<td><img src="image" alt="Size" /></td>
<td>Not Effective</td>
<td>Effective</td>
<td>Effective</td>
</tr>
<tr>
<td>VALUE</td>
<td><img src="image" alt="Value" /></td>
<td><img src="image" alt="Value" /></td>
<td><img src="image" alt="Value" /></td>
<td>Not Effective</td>
<td>Effective</td>
<td>Effective</td>
</tr>
<tr>
<td>TEXTURE</td>
<td><img src="image" alt="Texture" /></td>
<td><img src="image" alt="Texture" /></td>
<td><img src="image" alt="Texture" /></td>
<td>Effective</td>
<td>Effective</td>
<td>Not Effective</td>
</tr>
<tr>
<td>HUE</td>
<td><img src="image" alt="Hue" /></td>
<td><img src="image" alt="Hue" /></td>
<td><img src="image" alt="Hue" /></td>
<td>Effective</td>
<td>Not Effective</td>
<td>Not Effective</td>
</tr>
<tr>
<td>ORIENTATION</td>
<td><img src="image" alt="Orientation" /></td>
<td><img src="image" alt="Orientation" /></td>
<td><img src="image" alt="Orientation" /></td>
<td>Not Effective</td>
<td>Not Effective</td>
<td>Not Effective</td>
</tr>
<tr>
<td>SHAPE</td>
<td><img src="image" alt="Shape" /></td>
<td><img src="image" alt="Shape" /></td>
<td><img src="image" alt="Shape" /></td>
<td>Not Effective</td>
<td>Not Effective</td>
<td>Not Effective</td>
</tr>
</tbody>
</table>

---

88
<table>
<thead>
<tr>
<th>data variable</th>
<th>visual variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>position</td>
</tr>
<tr>
<td>direction</td>
<td>relative position, arrow head</td>
</tr>
<tr>
<td>volume</td>
<td>width, color intensity</td>
</tr>
<tr>
<td>attributes</td>
<td>color</td>
</tr>
<tr>
<td>attributes/temporal variables</td>
<td>vertical displacement</td>
</tr>
<tr>
<td>time</td>
<td>temporal slider (interactive),</td>
</tr>
<tr>
<td></td>
<td>frames (animation)</td>
</tr>
</tbody>
</table>

Table 3.1 Visual variables to represent traffic accident data

However, animation is seldom suitable for complex information graphics because in many cases the information content of the graphic is too complex to understand in a short period of time. To make the motion smooth, animation usually requires at least 24 frames per second. Unless the pattern we are looking for is very distinct or understood beforehand through some other means, it is very difficult to find it through complex graphics at such a high speed. Thus, it can be said that animation is more suitable for communicative purposes in which we want to present changes over time in a clear and understandable manner.
As an alternative to animation, the slide show technique allows the user to control the pace of the display sequence. Since the user controls how long a specific image is displayed on the screen, he/she can make sure that there is enough time for him/her to understand the image. This can also allow the user to stop the sequence at any time and move back and forth between frames to verify the changes that are noticed during the sequence. This technique, however, still requires putting together some sort of time-line along which the graphics are presented. To make a slide-show, each frame must be constructed in advance in a pre-planned sequence. Therefore, the slide-show method would not be very useful in exploring unknown data where a pre-determined sequence cannot be constructed.

What we need in exploratory situations is an interactive tool that dynamically changes the content of the graphic display in response to the user's control of time, i.e., a temporal slide bar. By dragging the knob in the slide bar users can move back and forth along the time line at a pace that allows the user to understand the graphics. A slow, continuous drag of the knob would create an animation effect from which the analyst could detect abrupt temporal changes. The user can stop at that point and repeatedly move back and forth around that point in time to inspect more closely what has happened. Tang (1993) successfully implemented this kind of temporal slide bar with his dynamic choropleth map display.
3.3 Visualization Tools for Traffic Accident Data

3.3.1 Structural characteristics of the underlying network

The network visualization tool deals with the physical or structural elements in the underlying road network. This tool differs from others in that it deals with the variables that are relatively static. For example, the functional classification of a road in a network is relatively stable over long periods of time. Likewise, the overall topological structure of a network does not change overnight. Of course, with historical data, we can think of a evolving network from a small, poorly connected network to a large, highly connected one. However, historical examination can be easily achieved by treating each time period separately with multiple instances of the same display module.

This network visualization tool is further divided into a node-related display and a link-related display. The variables that can be displayed with this network visualization tool include location, link impedance, capacity, and functional class of the link. Of course, other relevant variables concerning physical elements of the link can also be included.

The layout of the display should be simple, since its main purpose is to aid the interpretation of traffic accident patterns on the road network. This module draws the boundary of the study area as a background and displays the position and topological structure of the network. The screen should not be cluttered with other details that could be accessed in a separate pop-up window with a click of a pointing device. This display is intended to give a broad sense of spatial relationships among link characteristics and
accident locations. However, the spatial zoom function should be available in case a closer look at a specific subset of a network, such as downtown streets, is required.

3.3.2 Point Data Visualization

The Point visualization tool deals with the actual point data in a network. Traffic accident points can be represented as simple dots on the display. However, the two-dimensional dot cannot be effectively used to represent a point in a three-dimensional space. Therefore, the system employs three-dimensional objects to represent points. A sphere was considered first, but to facilitate the speed of the display, an octahedron—a simplified form of a sphere—was chosen over a sphere. Descriptions of the specific functional modules follow.

Dynamic control of the level of point size

The large number of points that need to be displayed create a fundamental problem of 'crowding the graphic,' i.e., the display area is too small to adequately display all the data. To alleviate the problem of display overcrowding, dynamic control of point size needs to be implemented. This control adjusts the point size according to the spatial extent currently displayed. For example, when we want to look at a small scale (e.g., county-wide) spatial pattern, we need to increase the size of the points so the points representing accident locations are clearly visible. On the other hand, when we zoom into a specific area to examine detailed accident patterns, point size needs to be adjusted so as to minimize the point overlap.
Dynamic selection of displayed variables

The fundamental benefit of building a dynamic visualization system is that the system can provide users with an instantaneous feedback. The ability to select a variable to display among a list of available variables is therefore the core of the dynamic system. The system response time should be reasonably fast (a few seconds at the longest) so that users can see the changes in the display immediately after their action. Also, a predefined list of available variables greatly reduces the need for the user to go back to some external documentation to find the variables or to type in the variable name. Selecting a variable involves just one mouse-click action. The system updates the display without any further action on the user’s part.

Dynamic selection of subgroups

There are many instances in data analysis when the focus of investigation centers on a small set of data points that satisfies certain conditions. As described earlier in this chapter, this function can be called ‘data filtering.’ The selection of a subset of data points can be performed in three ways by choosing the filter variable from: i) attribute variables alone; ii) temporal components alone; or iii) a combination of the two. Therefore, it is possible to show, say, only the accidents that occurred on Friday and involved drunken driving. By providing many possible ways to combine variables, the subgroup selection tool can be a quite useful exploratory tool.
Zoom, pan, tilting and rotation of the display

The geographic zoom function should be available throughout the display module to reduce the clutter and to focus the user’s attention on specific geographic areas. Where there is a great competition for limited display space among various symbols, a zoom-in function is required. On the other hand, when network-wide patterns are under examination, a zoom-out function is required. When zoomed in, there is a frequent need to move around outside the zoomed area while maintaining the same scale. Panning eliminates the need to zoom out and zoom back in to view an adjacent area.

In addition to the zoom and pan functions, tilting and rotation of the display is needed to maximize the advantage of the full three-dimensional display. Tilting allows the easy exaggeration of the vertical axis, which helps to find the patterns that may exist in the temporal spread. Rotation, on the other hand, allows users to look at the data from different sides to check for patterns that may be hidden behind other points closer to the viewing port.

3.3.3 Detecting changes in the temporal component of data

Differences in spatial patterns at different time periods often provide important clues on which time periods should be examined in detail. One of the most important tools that this prototype visualization system provides is in its handling of the temporal components in the data in an intuitive way. As mentioned earlier in this chapter, temporal variables can be handled in the following three ways: i) through the use of a
temporal slide bar; ii) through slide show or animation; and iii) by means of vertically displacing temporal variables.

The temporal slide bar can be understood as slicing the data along the temporal axis. The system provides a variable selection widget that can be used to select the scale of the temporal slice. At the same time, a slide bar widget is also provided so that the users can drag it across to examine different temporal slices.

Slide show or animation is an extension of the temporal slide bar. Depending on the speed of dragging the slide bar widget, the module can be effectively used as a slide show display. In addition, a special module was developed to allow users to create a slide show display.

Vertical displacement of a temporal variable takes advantage of the three-dimensional capability of this visualization system. It is an application of the familiar space-time cube where the vertical axis is used to represent time. Since the system's graphic display module is fully three-dimensional and capable of displaying data at any angle, this visual representation method has proven to be truly innovative and useful. Moreover, the selection of the displayed variables is handled in dynamic and interactive fashion. Therefore, the system allows the user to maintain the same angle of the display while displaying different variables.
IMPLEMENTATION OF THE VISUALIZATION SYSTEM

This chapter focuses on the implementation of the visualization system whose functional requirements are specified in the previous chapter. The structure of the prototype system is described using data flow diagrams (DFD). It then describes the design structure of the prototype system and the interconnections among its various modules. A description of the user interface for the system and the individual modules with their functions follows.

4.1 Data

The study area for this research is Pitt County in the state of North Carolina. The data comes from the North Carolina Department of Transportation (NCDOT). NCDOT collected and compiled traffic accident into Arc/Info Dynamic Segmentation coverages. There were 3860 accidents recorded over a three-year period from 1987-1989. The author converted these coverages into point and line coverages and ultimately into flat ASCII files that can be fed into the visualization system. Figure 4.1 shows the location of the study area in the state of North Carolina.
Pitt County is a major regional center of over 100,000 people located about halfway between Raleigh and the Atlantic Ocean on North Carolina's coastal plain. The county seat is the city of Greenville, a community with a population of about 50,000 (Fig. 4.2). Historically dependent upon agriculture, Pitt County is the nation's number one producer of flue-cured tobacco. Pitt County has emerged as an urban center of education, medicine, and industry. The county is home to East Carolina University, the University Medical Center of Eastern Carolina–Pitt County, Proctor & Gamble, Glaxo-Wellcome Pharmaceutical Corporation, and other major industrial employers. The municipalities and major roads of Pitt County are depicted in Figure 4.2.
Figure 4.2 Municipalities and major roads in Pitt County, North Carolina
4.2 Structure and Organization of the Prototype Visualization System

The prototype visualization system is implemented as a set of modules in the IRIS Explorer (Numerical Algorithms Group 1995) data visualization software. The IRIS Explorer is a well-known scientific visualization software that has been employed in a variety of fields (e.g., earth science, physics, engineering, chemistry) for visualizing large data sets. It is a system designed to create powerful visualization maps, each of which comprises a series of small software tools called modules. A map is a collection of modules that carries out a series of related operations on a dataset and produces a visual representation of the results (NAG 1995).

The data flow of the prototype visualization system is shown in Figures 4.3 through 4.5. The data flow is described through the use of data flow diagram (DFD). The data flow diagram is a graphical modeling tool of structured analysis in software engineering that allows us to picture a system as a network of functional processes connected to one another by “pipelines” and “holding tanks” of data (Yourdon 1989). A rectangle in a DFD represents an ‘external entity,’ while a bubble (circle) represents a ‘process.’ The flow of data is shown as an arrow with the arrowhead indicating the direction of data flow. The symbols in DFD that show two parallel lines with text in-between represent repositories of data that are to be stored for use by one or more processes.

Figure 4.3 shows the context level data flow diagram of the prototype system. The context level DFD is a special DFD in which a single bubble represents the entire system. It shows how the system interacts with the external environment. As the figure
shows, the system interacts with the user through a user-interface (called a visualization control panel) to get command and data. After processing the user input, the system produces the necessary geometry data and displays them in the graphics window.

Figure 4.3 Context level data flow diagram of the prototype system

In Figure 4.4 the single bubble in Figure 4.3 (visualization system) is expanded to show two major processes involved in the system. The process that controls the user interface is referred to as “interact with user.” This process handles the opening of source data files, as well as the processing of all the choices that a user makes with regard to visual representation of accident and network location, attributes, and temporal ranges. The data gathered in this process is then fed into a process called “process user input,”
which manipulates the data based on the user's choices. This process is also responsible for creating geometry data that can be displayed in the graphics display.

The expanded diagram of the "process user input" process is shown in Figure 4.5. Each of the bubbles in this level 2 DFD shows specific system functions. For example, the "read input data" process handles data input, while the "select subgroup" process
Fig 4.5 Expanded DFD (level 2) of the "process user data" process
deals with the selection of a subset of points based on the result of an interactive user query. The output of each of these processes is fed into a process called "create geometry," which is responsible for generating a geometry data structure suitable for the graphics display.

4.2.1 Data Organization and File Structure

The traffic accident data sets used in this dissertation contain four types of data. The first type of data is accident location. The coordinate information of the traffic accident locations was generated from Arc/Info coverages by using the UNGENERATE command in Arc/Info. The coordinates were saved in flat ASCII text files. The second data type is the accident attribute information. These data were also generated from Arc/Info coverages (which in turn are created from event databases included in the Arc/Info dynamic segmentation). A list of the attribute data with an accompanying data dictionary are included in Appendix. The third type of data is the underlying road network. Since the format of ASCII output of the Arc/Info UNGENERATE command for line coverages differs from that of points, we need to treat the network coordinate data as a separate data type. Finally, we have the attribute information of the road network, which was saved in an ASCII file.

4.2.2 Module Organization

There are three major components in the system: i) the Data Input module; ii) the Point Visualization module; and iii) the Network Visualization module. The Data Input
module reads two different types of ASCII data files (coordinate data and attribute data) that are generated from Arc/Info coverages and prepares them into a suitable format that can be used in the visualization module. Since the format of the ASCII files that are generated from Arc/Info are different for point and line data, each visualization module employs a different data input module. The Point Visualization module creates a dynamic graphic display of the point data. It consists of several control submodules: the Point Size Control submodule; the Subset Selection Control submodule; the Color Variable Selection submodule; the Vertical Displacement Control submodule; the Temporal Slider Control submodule; and the Slideshow Control submodule. The Network Visualization module consists of the Network Display Control submodule and the Visual Representation Selection submodule.

There are some additional modules in the system that make it possible to generate dynamic graphics output. The data processed in the Point Visualization module and the Network Visualization module are displayed in a graphics window called the Render module. The Render module has a built-in capability to zoom-in and zoom-out, pan, and rotate the display in 3D space. The GenerateColorMap module is used to control the color-space of the display, while the Legend module is used to generate the legend display. Finally, the Timer module is used to create the slide show of selected data points. The following section describes the system’s user interface, as well as the individual modules that make up the system in greater detail.
4.2.3 User Interface

Figure 4.6 shows the user interface of the prototype visualization system. The computer display screen is divided into four different sections. The main graphics display window is located in the upper-left hand corner. The title of this window is ‘Render.’ This is the window in which all the graphic output is displayed. To the right of the ‘Render’ window is the ‘Point Visualization’ module. All the operations related to the visualization of point data is handled by this module. Almost all of the operations that a user performs are carried out by making appropriate choices from either a menu or a list of choices using a pointing device (mouse). The ‘Network Visualization’ module is located in the lower-left hand corner. To its right is a window entitled ‘Timer.’ This is the user interface of the ‘Slide Show’ module. Next to the ‘Timer’ window is a window called ‘Switch.’ It is used for turning the network display on and off.

4.3 Modules and Functions in the Visualization System

This section describes the modules implemented in the system along with their functions. First, we describe the overall organization of the modules that are implemented as IRIS Explorer modules. This is done by showing the IRIS Explorer map of the visualization system (Figure 4.7). Each rectangle on the map represents a module in the system. The rectangles are connected with lines representing the flow of data. The output of the previous module is input for the following module. The modules in Figure 4.7 are shown in minimized form to avoid cluttering the diagram and to bring attention to
Figure 4.6 User interface of the prototype visualization system
the data flow and the relationships between modules. After the description of the overall organization of modules, each individual module is described in detail.

4.3.1 IRIS Explorer map of the visualization system

The IRIS Explorer map in Figure 4.7 shows the overall structure of the modules implemented in the system. The visualization system is implemented with 11 modules. The main functionality of the system is found in three major modules: i) the Point Visualization module; ii) the Network Visualization module; and iii) the Render module. These modules are interconnected to various other support modules that perform specific functions. For example, the coloring of the point and network symbols is made possible through the use of the GenerateColormap module, which creates a standard colormap and feeds its output to visualization modules. The Timer module, which is responsible for the slide show display, controls the parameters for the temporal selection variable so that appropriate values for the variable can be fed as input into the selection routine of the Point Visualization module.

4.3.2 Data Input

Data input is handled within the Point Visualization and Network Visualization modules. Both the ‘Point Visualization’ module and ‘Network Visualization’ module have a pull-down menu called “File”. When this menu is selected, a file selection dialogue box pops up. From this selection box, users can select appropriate data files for display. Although two separate types of files (one for coordinates and one for attributes)
Figure 4.7 Overall structure of the prototype visualization system
are required for each module, users only need to select a common portion of the file names without the extension to select the data set. In other words, if the data set consists of `pitt_county.crd` (coordinates) and `pitt_county.atr` (attributes), users can select just "pitt_county" from the file list. The system automatically loads the necessary files.

### 4.3.3 Point Visualization Module

The 'Point Visualization' module is one of the key components of the visualization system (Figure 4.8). It has various types of user interface widgets. A menu item called 'File' is used for data set selection. The text input box under the menu can also be used to directly type in the data set name. Below the text input box are two on/off buttons. The 'Subgroup Selection' button is used to activate/deactivate the query buttons (a group of eight buttons in the middle). The 'Temporal Slidebar' button is used to turn on/off the temporal slider. To the right of the two on/off buttons is a dial for 'Sphere Size.' Users can use this dial to interactively control point symbol size. This kind of interactive control of the size of the visual representation is essential for generating appropriate displays at various scales.

#### Interactive Query Operation

When the 'Subgroup Selection' button is set to 'On', a group of eight buttons appear. These buttons are called 'option buttons.' Beneath each button, is a list of available options. By selecting relevant items from the list, users can perform a query on the point attributes. The module selects all the points that meet the condition as soon as
Figure 4.8 Point Visualization Module
the selection is made and displays these points in the graphics display. The buttons act just like 'AND' operator in a Boolean search. In other words, selecting an item from the list of items under one button, say the 'Alcohol related' item from the 'Alcohol Involvement' button, and selecting another item under another button, say 'Friday' from the 'Day of the Week' button, would display only the points that satisfy both conditions, that is the accidents that are alcohol-related and occurred on Friday. There are many items to choose from under each option button, and the combinations from the eight different buttons with different choices give an incredible amount of flexibility for query construction. In addition, since the subsequent choice does not release the subset of points already selected from the previous choice, the user can sequentially narrow the search to determine exactly what is happening with the points selected. This powerful query capability coupled with the almost instantaneous system response to the user query provides users with truly interactive visual feedback.

**Visual Representation of Attribute Data**

In visualizing complex multi-dimensional data, it is of critical importance to utilize as many visual representation methods as possible to maximize the visual feedback. The use of color is one of the most important visual cues to differentiate attributes embedded in the data. The use of color makes a huge difference when the visual symbols are displayed. Colors can reinforce the visual feedback when they are combined with the symbols already in the display. For example, if a circle represents one attribute in the data (e.g., a minor accident) and a square represents another (a fatal
accident), it would be hard to see the fatal accident patterns when many minor accidents dominate the display. If the fatal accidents are colored red, however, we would be able to easily see the patterns.

Color can also be used to display additional information. We could, for instance, color the symbols with another attribute (e.g., the speed at which the accident occurred) while maintaining the same symbolization scheme of circles and squares as in the above example. In this case, colors are being used to display an attribute that is different from the attribute represented by the point shapes.

In the lower-right hand corner of the 'Point Visualization' module is a selection box labeled 'Color Variable.' This box is used to select an attribute from a list of available attributes to color the points. We can use this selection box to show an attribute in the data. For example, when a user selects the 'Reportable Status' attribute for the color variable, the display will show points with colors representing the severity of the accident. In this case, red represents fatal accidents, green represents injury accidents, and so forth.

When colors are used with the vertical displacement method (a new way of displaying the temporal component in the data that will be described later), they reinforce the visual feedback from the display. For example, we can choose the 'Day of Week' as the variable to be used for the vertical displacement. We will see in the display seven layers of points with each layer representing a day of the week. When dealing with many points, differentiating the points that belong to a specific layer can be difficult. However, if we choose 'Day of Week' as the color variable so as to display the layers with different
colors, we can easily tell whether a point belongs to a certain day, such as Friday or Saturday.

4.3.4 Treatment of the Temporal Component in the Data

Traffic accident data are extremely complex. Not only do the data include spatial and aspatial information, they also include temporal information. Effectively presenting the temporal information is of critical importance for the efficacy of the system. The prototype system employs three representation methods to display the temporal data: vertical displacement; the temporal slider; and the slide show.

Vertical Displacement

An easy way of understanding the vertical displacement method is to think of a box. We use the bottom surface to represent the spatial dimension and the vertical axis to represent the temporal dimension. Each data point would then possess three coordinates, $x$, $y$, and $z$, where $x$ and $y$ are spatial coordinates and $z$ is the vertical coordinate representing time. When we plot the data points inside this box we would get a point cloud display. When the data used for $z$ is categorical or ordinal, we would see the points stacking up as layers based on their value. If the data are some values that can be treated as a ratio, we would see the points scattered all over the inside of this box. Either way, what is shown in the display would be a representation of both the spatial and temporal data. This method of display is called the ‘Vertical Displacement’ method.
The selection box for vertical displacement is located in the lower-left hand corner of the ‘Point Visualization’ module. The available choices are: none, day of week, month of the accident, year of the accident, report time, and chronological order. The last item, chronological order, requires additional explanation. This term is used to identify the method that converts the accident date into something that can be used as ratio data. Strictly speaking, the absence of absolute zero in the calendar date means it is not in ratio scale. However, for the given time range, it is possible to convert the calendar dates to another measure so that we can use them as if they were in a ratio scale. A frequently used method for this kind of conversion is to use Julian dates. We can convert any calendar date into a Julian date which can be used as ratio data. The prototype system employs an internal Julian date conversion routine so that all the dates used for temporal display are Julian dates. The use of Julian dates makes it possible to plot the data points along the z dimension based on the calendar date on which the accidents occurred.

Temporal Slidebar

One of the frequently used methods for displaying temporal data is to create an output that changes its display content over time. Animation is the most well-known technique of this kind. However, creating an animated sequence involves stringing many snapshots of the display in a predefined sequence. The variable and query cannot be changed in the middle of the animation. In other words, animation takes away the interactive nature of the visualization system. A temporal slidebar or slider, on the other hand, not only allows direct control of the temporal data that needs to be displayed, but
keeps the interactive control of the attribute display intact. The ‘Point Visualization’
module provides a temporal slider for the direct control of temporal data display.

The button labeled ‘Temporal Slidebar’ controls the activation of the slider. When it is set to ‘Off,’ the slider disappears from the interface to minimize confusion. When the slider is activated, users can move the bar using the mouse to display the accident locations for a particular time period. The subgroup operation mentioned earlier can be set to ‘On’ or ‘Off’ depending on the user’s intent. The button underneath the slider controls the year to which the date belongs. The current date appears in the upper-left hand corner of the slider. The system does not show the calendar date on the display, but rather displays the date in the command window to which users can easily refer.

Slide show

Sometimes we desire to just see how the pattern looks for a fixed period, such as for the entire month of January, 1988. The temporal slider can accomplish this by moving the bar repeatedly, but it is quite awkward. The module entitled ‘Timer’ is used for this situation (see Figure 4.9). This module allows a fine level of control over the starting and ending values of the time range, as well as control over the delay between display updates. Users can also specify a step value to be used for increment. Usually the step value is set to one. By using a different step value, we can explore some interesting hypotheses. For example, if we specify a step value of 7, the ‘Timer’ will display the accident locations on the same day of the week.
Figure 4.9 Slide Show Module (Timer)
It is fun to watch the slide show. Points pop up at various locations, then vanish as quickly as they appeared. During a slide show, a sequence of displays may catch our eye for whatever reason. There may be a sequence of days during which a particular area shows repeated occurrences of accidents or the overall accident pattern might show a hint of cyclical pattern. Whatever the case may be, if and when we detect an interesting phenomenon, we can certainly replay the sequence to examine it more closely. We can also use the ‘Temporal Slider’ to assist in the investigation.

4.3.5 Network Visualization Module

Displaying the underlying network is another major function of this prototype visualization system. The network display can simply show the shape and pattern of the road network to provide the context for the point display. However, the intricate relationships between traffic accidents and underlying road characteristics require a more flexible display scheme. There is a need to show the appropriate network attributes that may reveal hidden relationships between not only accident locations and road attributes, but also between accident attributes and road characteristics. The ‘Network Visualization’ module provides a flexible display method that makes possible interactive selection of network characteristics.

The control interface of the ‘Network Visualization’ module is shown in Figure 4.10. In the top portion of the module window is a menu item called ‘File’. This menu, along with the text input box under it, is used to select the input dataset. The use of the ‘File Selection’ dialogue box is identical to that in the ‘Point Visualization’ module,
Figure 4.10 Network Visualization Module
although the file format for the input data for network coordinates is different from that for point data. Below the text input box is a selection box labeled 'color variable.' This selection box is used to display the desired network attribute. The 'Network Visualization' module displays the roads in the network with different colors representing different values in their network attributes. It can also use the combination of color and line width to maximize the visual feedback. For example, the 'Functional Class' of the network is displayed with varying colors depending on the categorical value of the variable, but the line width does not change. On the other hand, the 'Average Daily Traffic' variable display uses both color and line width.

The ability to change the displayed network attribute instantly gives a great deal of freedom to investigate the various interrelationships between point data and network data. Not only can users observe what types of accidents are occurring on what types of roads, but they can also work on a possible hypothesis that might explain the observed phenomenon. When this capability is combined with interactive query with regard to point attributes, the visualization system provides an excellent analysis environment for generating new hypotheses and rapidly testing outcomes by visual inspection.

4.4 Examples of the Graphic Display

In this section, some examples of graphic display are provided to show how the display looks when some of the controls in the modules are used. There are five examples (Figure 4.11 through Figure 4.15) showing various combinations of spatial,
temporal, and attribute variables selected at different zoom levels and viewing angles. It needs to be mentioned, however, that the printed figures have severe shortcomings in depicting what really is displayed on the computer screen. There are two problems with printed pictures. First, the resolution (200 dots per inch) of the color printer used to print the graphic is not adequate. To be able to depict the sharp contrasts among point symbols (represented by colored octahedra), the resolution should be much higher. The printing process also took out a lot of tonal difference (especially in red), so the printed display looks dull. The second problem is that the printed graphics cannot convey the interactive nature of the system. The instant feedback that users receive when they change the variables, the scale, the size of point symbols, and the viewing angle is the most important factor in successful system implementation. In actual use, the system usually responds to a user action almost instantaneously and the users have almost unlimited freedom to look at, zoom in, zoom out, pan, tilt, and rotate the whole display. Keeping the limitations of the printed pictures in mind, some examples of the visualization system follows.

**Example of point attribute query and vertical displacement**

Figure 4.11 shows the traffic accident data displayed with the *alcohol involvement* attribute variable selected and *the month of the accident* used as the vertical displacement variable. The size of the octahedra have been increased significantly so as to show the overall county-wide pattern. The size is, of course, interactively controlled by the user.
Figure 4.11 Example of point attribute query
using the ‘Point Size’ dial. The legend indicates that the color value 0.5 indicates the locations of accidents that involved alcohol.

**Example of vertical displacement at side angle.**

The effect of the vertical displacement of the temporal component in the data is shown in Figure 4.12. The display is also tilted so that the vertical axis can be more clearly represented. Note that the size of the point symbols has been reduced to fit the display scale.

**Example of zoomed in display**

Figure 4.13 shows a zoomed display of Figure 4.12. Now we can clearly see the stacking of accident points based on the month of occurrence. This display clearly shows the limitations of a simple two-dimensional display. The overlapping of the accidents at the same location or in close proximity of each other cannot be represented with a 2-D display.

**Example of displaying the temporal spread**

Figure 4.14 shows what happens when you set the display angle in such a way that the eye level of the user is almost at the horizon. At this angle the temporal spread is much more obvious. When the user’s interest is discovering some temporal patterns in the data, this is a useful perspective from which to look at the data. Also note that this
Figure 4.12 Example of vertical displacement at side angle
Figure 4.13 Example of zoomed in display
Figure 4.14 Example of displaying the temporal spread
view shows the data from different compass directions. Here we are looking at the data from the South end of the county toward downtown Greenville.

**Example of using the color variable**

Figure 4.15 shows the display with a different vertical displacement variable (Day of the Week) and different attribute variable (Reportable status). The red point symbols represent fatal accidents, while the green represent injury accidents.
Figure 4.15 Example of the use of color variable
CHAPTER 5

AN ASSESSMENT OF THE IMPLEMENTATION AND DISCUSSION

This chapter demonstrates the operation of the prototype visualization system and discusses its efficiency in visualizing complex spatio-temporal point data. The demonstration focuses on what kinds of visual representation can be made possible through the use of the system, what are some of the interesting patterns found in the data, and how the temporal dimension is treated within the system. This chapter also presents a discussion of the way in which the presence of a space-time cluster can be statistically tested and shows the results of the analyses.

The data used in this chapter is the time and location of 3,680 cases of traffic accidents and their related attributes for the underlying road network of Pitt County in the state of North Carolina. For space-time clustering analysis, 253 cases of alcohol-related traffic accidents were used.
5.1 Visualizing the traffic accident patterns

The previous chapter described the layout and major modules in the system, as well as the system's user interface. In this section, we will focus on the resultant graphic displays generated from the system when the user made various choices through the control modules.

Display of overall accident patterns

One of the most basic graphic representations dealing with spatial data is displaying where the data items are. By looking at the locations of all relevant objects, a researcher can form a broad sense of the overall magnitude of the phenomenon and begin finding specific patterns that might be of interest. Figure 5.1 shows the output of the graphic display showing all the traffic accident locations in the study area. As evident in Figure 5.1, the occurrences of traffic accidents can be found for almost all of the major roads. Although the user can select a particular attribute to use as a coloring variable, the overlap of points (represented as octahedra) at the county-level scale severely limits the usefulness of the color. The dynamic nature of the interactive display can be very advantageous in that when it is necessary, the user can rapidly zoom in to focus on specific portions of the study area (for example, the city of Greenville in the center). An example of a zoomed-in display of all traffic accident locations in and around Greenville area is shown in Figure 5.2. Panning can minimize repeated zooming-in and zooming-out, while maintaining the same scale of the graphics.
Figure 5.1 Overall pattern of traffic accidents
Figure 5.2 Accident locations in and around Greenville, NC
The size of the accident point symbols (octahedra) can be interactively controlled to allow the user to see the overall pattern at a small scale by increasing the size or by decreasing the size to minimize the overlap of symbols when zoomed in. The network display can be activated or deactivated depending on the scale and on user needs. For example, by decreasing the point symbol size to zero and activating the network display, users can see what the overall structure of the network looks like. At the same time, users can select from available network attributes to display an attribute that meets their needs. For instance, Figure 5.3 shows the functional class of each road in the study area using different colors.

Display of alcohol-related traffic accident patterns

There could be a multitude of combinations of queries users can select. It would be both impossible and not very useful to demonstrate every possible point pattern in this section. In fact, being able to generate an unlimited number of graphic displays at a rapid pace is precisely the main benefit of the exploratory system. Therefore, this and subsequent sections show some of the many accident point patterns generated from the system, and describe the means through which the display is generated. By comparing the outputs from different user queries, this and subsequent sections also show how some interesting hypotheses concerning the traffic accident patterns in the study area can be generated.

In this section, the focus of the examination is the spatial pattern of the alcohol-related traffic accidents. Figure 5.4 shows the locations of alcohol-related traffic accidents.
Figure 5.3 Network display showing functional class as a color variable
Figure 5.4 Locations of alcohol-related traffic accidents
accidents for the period 1986 through 1988. We can observe from the figure that the locations of alcohol-related accidents largely coincide with the locations of municipalities in the study area. The main concentration of accidents is found in and around the city of Greenville. To the north, we can see some string of accidents that occurred in Bethel, and to the west, Fountain and Falkland. On the south side, Winterville, Ayden, and Grifton all show some accidents occurring. Simpson and Grimesland to the east also shows several alcohol-related accidents. It is also interesting to find that there are a number of accidents on the roads between municipalities. Notably, accidents that are south and south east of Greenville fall into this category.

Figure 5.5 shows the alcohol-related accidents in and around Greenville, viewed from the south. The vertical position and color of the points indicate the ‘day of the week’ variable, in which the lowest position represents Monday accidents and highest position Sunday accidents. We can clearly see that there are more points in higher vertical positions than lower positions. This indicates that more alcohol-related accidents occurred on weekends than weekdays, which is quite probable. To get the contrasting spatial patterns, we can plot the accident locations based on the day of week variable. Figure 5.6 shows the accident patterns that happened on Monday. Compare this to Figure 5.7, which shows the Saturday accident pattern. Not only did the number of accidents dramatically increase on Saturday, but the accidents in and around outlying towns are much more pronounced. There may be a reason for this discrepancy between the two patterns and an investigation of the factors that might have caused the discrepancy would certainly be warranted.
Figure 5.5 Alcohol-related accidents in and around Greenville, NC (viewed from the south)
Figure 5.6 Alcohol-related accidents that occurred on Monday
Figure 5.7 Alcohol-related accidents that occurred on Saturday
We can examine the spatial patterns of accident locations on a more local scale. Figure 5.8 shows accident locations in and around the city of Greenville. The colors for points represent the ‘reportable status’ variable, where red represents fatal accidents, green injury/non-fatal accidents, and yellow property damage only accidents. Due to both the limited space available on a printed page and the limits of reproduction, the output does not show the necessary details that can be observed on the computer display. In fact, where there are stretches of roads on which many accident locations are extremely close, they only show up as one point or something that looks very close to one point. However, when we zoom in to those areas and use the vertical displacement technique, we can find that many accidents occurred there. A good example of this overlap can be found in the stretch of NC11 near the intersection of NC33 (NW quadrangle in Figure 5.8). These limitations notwithstanding, we can still identify the locations of alcohol-related fatal accidents, as well as some stretches of road where alcohol-related accidents are more concentrated. The curved section of NC 33 east of Greenville shows a concentration of accidents. A closer look at this section of the road would be worthwhile.

The pattern of fatal and injury traffic accidents

Among traffic accidents, the accidents that involve fatality or bodily injury are much more serious than property damage only accidents. Accordingly, it would be quite appropriate to see what kind of spatial patterns the fatal and injury accidents create and how the patterns relate to various attributes of the accidents. One of the most studied attributes with respect to fatal traffic accidents is the speed at which the vehicle was
Figure 5.8 Alcohol-related accident locations in and around Greenville, NC
moving. Therefore, this section describes the graphic display of the patterns of fatal and injury accidents with regard to the speed of the accidents.

The locations of injury accidents seem to be distributed all over the major road network in the study area when they are viewed in the aggregate. However, when separate displays are created based on speed when the accidents occurred, we notice some differences in spatial patterns among the displays. Figure 5.9 shows the locations of injury accidents that occurred at low speeds (under 29 mph). To make the point pattern clearer, the road network is not shown in the graphics. We find a big concentration of accidents in Greenville, especially on the N-S stretch of NC11 and again on northbound NC11 from Greenville to Bethel. These stretches of NC11 are obviously quite injury accident prone.

Let us take a look at the pattern of medium speed (30-49 mph) injury accidents (Figure 5.10). The overall pattern is similar to that of low speed accidents, although the significance of NC 11 is significantly diminished, especially on the northern part of the county. Instead, there are more accidents near townships and also on the major east-west thoroughfare of US 264 between Farmville and Greenville.

Finally, Figure 5.11 shows the pattern of high speed accidents. The difference in the spatial patterns between high speed accidents and low or medium speed accidents is striking. The high speed injury accident locations are quite evenly distributed along all major roads with the exception of the center of Greenville. It is evident that more high speed accidents occurred on the segments of roads that connect the towns in the county.
Figure 5.9 Location of alcohol–related injury accidents at low speed
Figure 5.10 Location of alcohol-related injury accidents at medium speed
Figure 5.11 Location of alcohol–related injury accidents at high speed
Many of these roads are two-lane rural highways and they seem to be the most injury accident prone.

Would the same pattern hold for the fatal accidents? We examine this question using the output in Figures 5.12 and 5.13. Because there were only two low speed fatal accidents reported in the area, only medium speed (Figure 5.12) and high speed (Figure 5.13) fatal accidents are shown. It is interesting to note that there is again a marked difference in the distribution of accident locations. Medium speed accidents are concentrated in and around Greenville, while high speed accident locations show a wide dispersion. The pattern is very similar to what we just observed for injury accidents. Some studies on fatal traffic accidents have reported that the fatality rate for rural two-lane highways is higher than that for multi-lane divided highways. The dispersion of high speed fatal accident locations in Pitt County seem to support these findings.

Detecting possible temporal anomalies in traffic accident locations.

The use of the vertical displacement technique to display temporal patterns in traffic accidents has been described in the previous chapter. The vertical displacement technique is useful to see whether there are obvious clusters in the temporal dimension. However, there are instances where more traditional techniques of treating time are necessary. Conventionally, time dimension is displayed through the use of a slide show or animation.

The prototype visualization system provides two means of dealing with time, while maintaining user freedom to construct any kind of query that he/she is interested in.
Figure 5.12 Location of alcohol-related fatal accidents at medium speed
Figure 5.13 Location of alcohol-related fatal accidents at high speed
and with the color variable of his/her choice. The first is using the ‘Time slider.’ When
the user moves the time slider, the accident locations that satisfy the user’s query at the
particular point in time is displayed. Figure 5.14 shows an example of the graphic
display when the time slider is used. When the date is set to January 9, 1986 using the
time slider, four accident locations with their color representing the value of ‘Reportable
status’ variable are identified. In this instance the subset operation was turned off to
show all the accidents that occurred on that date. Since accidents are scattered throughout
the whole time dimension, it would be more appropriate to start with the subset operation
turned off when using the time slider. If a temporal anomaly is found for a certain date,
then a subsequent investigation can be conducted using interactive query.

The second method for displaying the time dimension in a more conventional way
is to use the ‘Slideshow’ module. As discussed in the module description section of the
previous chapter, the ‘Slideshow’ module provides the means to go through a chunk of
time without the interactive manipulation of time slider. Users can set both the initial
value, and ending value as well as the delay factor (in seconds) between the display
updates. By observing the accident location points appearing and disappearing on the
display, users can get a general sense of the spatial patterns over a given period of time.
If any interesting sequence of display is detected, we can readily go back and investigate
that particular time period with the time slider. It is quite unfortunate that the static
nature of the print medium does not allow the demonstration of the efficiency and
usefulness of this capability.
Figure 5.14 Location of traffic accidents that occurred on Jan 9, 1986
5.2 Testing the presence of space-time clusters

In this section, we shift our focus to tackling the following question: Is there statistically meaningful evidence of the presence of space-time clusters in the data? Some of the graphic displays described in the previous section seem to indicate that there may be a possible space-time cluster in the data. As seen in the display of accident locations, the sheer number of data points and the excessive overlap of points at some spots make it extremely difficult to conduct a statistical test for all the data points. Nor is the all-inclusive test desirable since the objective of the statistical test is to determine if the location and the timing of particular types of accidents show any clustering in both space and time. Therefore, the alcohol-related traffic accident subset of the data (253 cases) was selected for examining the possibility of space-time clustering.

The methods that are applied are Knox’s contingency table approach and Mantel’s generalized regression approach. These methods along with their test results are described in the following section.

5.2.1 Methods and Test Results

Knox’s Contingency Table Approach

Knox’s test is based on a consideration of all \( n(n-1) \) ordered pairs that can be formed from an observed set of \( n \) events. To identify space-time clusters, we need to identify the pairs that are close both in space and in time. The ‘closeness’ in this case is binomial, i.e., a pair is classified to be ‘close’ if they are within a certain critical distance
(or time interval) set by the researcher. These pairs of events are then placed into one of four cells in a 2x2 contingency table. If we have \( X \) pairs that are close in space and \( Y \) pairs that are close in time, and if we assume the independence of space and time, the number of pairs \( Z \) that are simultaneously 'close' in both space and time will have an approximate Poisson distribution with mean \( \frac{XY}{n(n-1)} \). If our observed value of \( Z \) is 'unusually large' when compared with the expected value, we can reject the null hypothesis of a space-time independence in favor of space-time clustering.

The problem with Knox's approach is the use of arbitrary critical distances in the analysis. Depending on the choice of the distance value used, we can get quite different results from the same data. For the selection of the critical distance, Knox suggested using a value that is meaningful to the nature of the problem or derived from an inspection of the data (Knox 1964). However, as in the case of alcohol-related traffic accidents, there are cases where the researcher does not have any \textit{a priori} notion of the appropriate value for the critical distance. In this situation, the frequency distribution of data can suggest some clues for suitable values. We can also compensate for the weakness of the analysis due to the use of an arbitrary critical distance by repeating the analysis using different sets of critical distances.

Figures 5.15 and 5.16 show the frequency distributions of inter-event distances for space and time, respectively. In Figure 5.15, we notice that there is a natural break point at distance 2000 m as well as at 7000 m. Whether to classify the events within 7000 m of each other as being 'close' in space depends on the size of the study area. Given the size of Pitt County and the size of the municipalities in the county, a separation
Figure 5.15 Frequency distribution of interpoint distances
Figure 5.16 Frequency distribution of inter-event intervals
of 7000 m could be treated as being spatially close. Unlike the distance data, the frequency distribution of the time interval data (Figure 5.16) does not suggest any obvious breakpoints. Therefore, it seems reasonable to use several time intervals to see if any of the intervals give rise to any interesting pattern. The values chosen are 0 (the accidents that occurred on the same day), some multiples of 7, which may explain the weekly trend in the data, as well as 30 days to represent a month. The value of 365 is used purely for comparison.

One important point that needs to be made regarding the distance measure is that the distance used in this analysis is the distance over the network, not the Euclidian distance used in most area-based point pattern analyses. As mentioned earlier, linear points can only occur in a line, which is frequently a part of the network. Using straight Euclidian distance can severely misrepresent the actual distance between two points. Therefore, all of the inter-point distances used in the analysis are the shortest-path distance between events. This use of network distance provides a much more realistic measure against which to test the evidence of the space-time clustering in the data.

Table 5.1 shows the result of the Knox's contingency table analysis for space-time clustering.
Distane(m) | Interval (days) | 0  | 7  | 14 | 21 | 28 | 30 | 90 | 365
---|---|---|---|---|---|---|---|---|---
500 | Exp | 0.6068 | 6.6438 | 12.899 | 18.92 | 25.253 | 27.027 | 76.894 | 276.86
| Obs | 16  | 24  | 28  | 36  | 46  | 46  | 96  | 268 |
1000 | Exp | 1.1109 | 12.162 | 23.613 | 34.636 | 46.229 | 49.476 | 140.77 | 506.84
| Obs | 16  | 26  | 38  | 48  | 60  | 60  | 152 | 468 |
1500 | Exp | 1.7275 | 18.913 | 36.72 | 53.861 | 71.889 | 76.938 | 218.9 | 788.17
| Obs | 16  | 30  | 42  | 64  | 80  | 80  | 220 | 760 |
2000 | Exp | 2.5031 | 27.406 | 53.207 | 78.046 | 104.17 | 111.48 | 317.19 | 1142.1
| Obs | 18  | 36  | 62  | 92  | 112 | 114 | 324 | 1108 |
7000 | Exp | 15.038 | 164.65 | 319.66 | 468.88 | 625.82 | 669.78 | 1905.6 | 6861.3
| Obs | 30  | 186 | 350 | 504 | 632 | 676 | 1918 | 6860 |

Numbers in bold italic font represent the cases where the observed numbers are greater than the expected numbers.

Table 5.1 Result of Knox’s Contingency Table Analysis

The result of Knox’s analysis shows some unusually large numbers of observed cases, when compared with the expected number of cases from a theoretical distribution. Therefore, we can tentatively reject the null hypothesis of no clustering in favor of clustering for some distance-interval combinations. The reason for the tentativeness is explained below. Particularly when the distance of 500 m is used, most of the number of observed cases far exceeds the expected number up to the time interval of 30 days. This strong tendency for clustering weakens as the critical distance used in the analysis increases, but we could still find some significant differences between the observed number and expected number up until the distance of 2000 m and the time interval of 7
days. Along the temporal dimension, the greatest difference in the two numbers is found when the time interval of 0 days (the same day event) is used. At this time interval, the result shows significant evidence of clustering up to a distance of 7000 m. As the time interval increases, the evidence of clustering weakens throughout the all distance ranges, but the smaller critical distances show much more persistent evidence of clustering than the larger distances.

Although some of the space-time combinations suggest the presence of clustering, the pattern of differences between expected value and observed value suggests that something else is going on in the data. The greatest discrepancies between the expected value and the observed value were consistently found when the same day accident was considered. This pattern prompted a closer examination of the same day accidents.

By limiting the inter-event distance to the same day, 16 accident pairs were identified. When the attributes of these 16 accident pairs were examined, a plausible explanation for the pattern, that is, the stronger space-time clustering for same day events, was found. It turned out that the 16 pairs that were extremely close in space and time were actually 8 cases of alcohol-related accidents that were reported twice. The reason why these accidents were reported twice may be because both of the drivers involved were found to be driving under the influence and a separate accident report was generated for each driver. Therefore, the two data points are in the same location and at the same time. It is a matter of judgment whether to treat a case such as this as separate accidents or as a single event.
From an exploratory data analysis perspective, the fact that the results of the Knox’s test led the author to a closer examination of a particular set of points, which resulted in finding the double-counted cases in the data is significant. The author did not know these cases existed in the data. It is true to the meaning of ‘exploratory,’ therefore, when the analysis result suggested the anomalies in the data.

Identifying the location and the time of the cases in which both of the drivers involved were under the influence is valuable in and of itself. However, the influence of these cases on the outcome of the analysis seems so heavy that there is a need to reexamine the data with these cases being treated as single events.

Table 5.2 shows the result of Knox’s analysis when the above mentioned cases were treated as single events. With the new data set, the result again shows some evidence of space-time clustering in some distance-interval combinations. At the critical distance value of 500 m the observed number exceeded the expected numbers at the time interval of 7, 28, 30, and 90 days. At the interpoint distance of 2000 m, there is one instance of distance-interval combinations that produced a larger observed number than the expected (the combination of 2000 m and 0 day). At 7000 m interpoint distance, all the observed numbers for all the inter-event intervals, with the exception of 365 days, exceeded the expected numbers.
Numbers in bold italic font represent the cases where the observed numbers are greater than the expected numbers.

Table 5.2 Result of Knox’s Contingency Table Analysis (with revised data)

<table>
<thead>
<tr>
<th>Distance(m)</th>
<th>0</th>
<th>7</th>
<th>14</th>
<th>21</th>
<th>28</th>
<th>30</th>
<th>90</th>
<th>365</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 Exp</td>
<td>0.4376</td>
<td>5.8639</td>
<td>11.407</td>
<td>16.658</td>
<td>22.274</td>
<td>23.806</td>
<td>67.712</td>
<td>243.76</td>
</tr>
<tr>
<td>500 Obs</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>16</td>
<td>26</td>
<td>26</td>
<td>72</td>
<td>230</td>
</tr>
<tr>
<td>1000 Exp</td>
<td>0.827</td>
<td>11.082</td>
<td>21.558</td>
<td>31.482</td>
<td>42.096</td>
<td>44.991</td>
<td>127.97</td>
<td>460.68</td>
</tr>
<tr>
<td>1000 Obs</td>
<td>0</td>
<td>8</td>
<td>18</td>
<td>26</td>
<td>38</td>
<td>38</td>
<td>124</td>
<td>416</td>
</tr>
<tr>
<td>1500 Exp</td>
<td>1.2968</td>
<td>17.377</td>
<td>33.802</td>
<td>49.363</td>
<td>66.005</td>
<td>70.543</td>
<td>200.65</td>
<td>722.34</td>
</tr>
<tr>
<td>1500 Obs</td>
<td>0</td>
<td>12</td>
<td>22</td>
<td>42</td>
<td>58</td>
<td>58</td>
<td>192</td>
<td>690</td>
</tr>
<tr>
<td>2000 Exp</td>
<td>1.899</td>
<td>25.446</td>
<td>49.5</td>
<td>72.287</td>
<td>96.657</td>
<td>103.3</td>
<td>293.83</td>
<td>1057.8</td>
</tr>
<tr>
<td>2000 Obs</td>
<td>2</td>
<td>18</td>
<td>40</td>
<td>68</td>
<td>88</td>
<td>90</td>
<td>288</td>
<td>1012</td>
</tr>
<tr>
<td>7000 Exp</td>
<td>11.324</td>
<td>151.74</td>
<td>295.17</td>
<td>431.05</td>
<td>576.37</td>
<td>616</td>
<td>1752.1</td>
<td>6307.6</td>
</tr>
<tr>
<td>7000 Obs</td>
<td>14</td>
<td>166</td>
<td>324</td>
<td>468</td>
<td>590</td>
<td>630</td>
<td>1758</td>
<td>6280</td>
</tr>
</tbody>
</table>

The result of Knox’s test with the revised data set is different from the test result using the original data set. First of all, the significance of the same day events has decreased significantly. In the test with the revised data, only two critical interpoint distances (2000 m and 7000 m) show the observed number exceeding the expected number when the same day is used for the critical inter-event interval. This is quite different from the first test result with the original data shown in Table 5.1 in which the importance of the same day events in producing the evidence of space-time clusters was much pronounced. This difference is expected because the double-counted pairs were
treated as single events in the second test. Secondly, the evidence of the space-time clustering is weaker in the second test result than that from the first. Many distance-interval combinations which produced the evidence of clustering with the original data no longer showed the evidence in the second test. In addition, the revised data showed that the space-time clustering was the strongest at the 7000 m interpoint distance. While some minor evidences of clustering were found at 500 m distances, the differences are too small to declare the existence of space-time clustering. Note that 7000 m was one of the breakpoints in the frequency distribution of interpoint distances (see Figure 5.15). This, coupled with the Knox's test result, suggests that if we consider the events within 7 km of each other as being 'close' in space, there is a strong evidence of space-time clustering.

One other thing we discover from Table 5.2 is that the assumption of a Poisson distribution in Knox's test seems to be holding up well in almost all space-time ranges. The close approximation of observed numbers to the theoretical expectation illustrates the fact quite well.

**Mantel's Generalized Regression Approach**

Mantel (1967) extended Knox's test by eliminating the assumption of a Poisson distribution from the Knox test. His test statistic is defined as \( Z = \sum \sum X_{ij}Y_{ij} \), where \( X_{ij} \) represents the 'distance' between events \( i \) and \( j \), and \( Y_{ij} \) is the corresponding 'time interval.' The null distribution of \( Z \) can be obtained through a finite population approach. If we have \( n \) locations of cases in space and \( n \) locations in time, the hypothesis of no
clustering is equivalent to one that the locations in space are matched at random with the
locations in time, there being a total of \( n! \) equiprobable sets of matchings. We can obtain
the null distribution of \( Z \) by listing the \( n! \) possible permutations of data and computing \( Z \)
for each permutation. Mantel (1967) devised a procedure through which the mean and
variance of the sampling distribution can be calculated without using \( n! \) possible
permutations of the data. Mantel’s procedure is summarized below.

Consider two \( n \times n \) matrices, one of \( X_{ij}’s \), the other of \( Y_{ij}’s \), and both with zero
diagonals. The quantities of \( X_{ij} \) and \( Y_{ij} \) are two different measures relating the \( i \)th element
of a sample to the \( j \)th element. The mean of \( Z \) is expressed as:

\[
\text{Exp } Z = \text{Exp} \sum \sum X_{ij} Y_{ij} \\
= \sum \sum X_{ij} \text{Exp } Y_{ij} \\
= \sum \sum X_{ij} \sum \sum Y_{ij} / n(n-1)
\]

We can write for the variance of \( Z \), the shorthand

\[
\text{Var } Z = \sum \sum X_{ij} X_{kl} \text{ Cov } (Y_{ij}, Y_{kl}) \quad \text{[eq. 5.2]}
\]

where the summation is over all permutations of the subscripts, although the cases \( i = j \) or
\( k = l \) can be ignored. We can write that, in general,

\[
\text{Cov } (Y_{ij}, Y_{kl}) = \text{Exp } Y_{ij} Y_{kl} - \text{Exp } Y_{ij} \text{ Exp } Y_{kl} \quad \text{[eq. 5.3]}
\]

where

\[
\text{Exp } Y_{ij} = \text{Exp } Y_{ij} = \sum \sum Y_{ij} / n(n-1) \quad \text{[eq. 5.4]}
\]

Using Equations [5.3] and [5.4], we can rewrite Equation [5.2] as

160
The value of $\text{Exp } Y_{ij}Y_{kl}$ still needs to be resolved. The exact procedure for deriving this value and the subsequent mean, variance and test statistic can be found in Appendix 1 and Appendix 2 of Mantel (1967). Ultimately, we get the $t$ statistic from the observed value ($Z$), expected value ($\text{Exp } Z$), and the variance ($\text{Var } Z$) with:

$$t = \frac{Z - \text{Exp } Z}{\sqrt{\text{Var } Z}}$$

If we let $X_{ij}$ and $Y_{ij}$ each take the value 1, if the events are close in both time and space, and 0, if they are not, then $Z = \sum \sum X_{ij}Y_{ij}$ would equal twice the number of close pairs both in time and space in Knox's contingency table. Therefore, we can verify the test result of Knox's approach using Mantel's statistic. Tables 5.3 and 5.4 show the results of Mantel's methods using symmetrical 1's and 0's with the original data and revised data, respectively.
The results of Mantel’s methods are fairly consistent with those of Knox’s when the original data set (253 cases) was used. The strongest evidence of space-time clustering is again found in predominantly short space distances and extremely short time intervals. The weakening of evidence for clustering also shows very similar patterns as Knox’s. Although Mantel’s method does not make any assumptions about the underlying distribution, the similarities of the resultant patterns in the statistic is remarkable. This again confirms that the assumption of the Poisson distribution in Knox’s approach was valid.

When Mantel’s method was applied to the revised data (245 cases), the result (Table 5.4) seems slightly different from that of Knox’s. Mantel’s method shows evidence of significant space-time clustering when the critical space distance of 7000 m...
and critical time intervals of 7 and 14 days are used. For space distance of 7000 m, Knox’s result was a bit more pronounced.

<table>
<thead>
<tr>
<th>Distance (m)</th>
<th>0</th>
<th>7</th>
<th>14</th>
<th>21</th>
<th>28</th>
<th>30</th>
<th>90</th>
<th>365</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>-0.4695</td>
<td>0.04014</td>
<td>-0.2997</td>
<td>-0.1168</td>
<td>0.57581</td>
<td>0.32862</td>
<td>0.40264</td>
<td>-0.8667</td>
</tr>
<tr>
<td>1000</td>
<td>-0.6474</td>
<td>-0.6634</td>
<td>-0.5533</td>
<td>-0.7104</td>
<td>-0.4622</td>
<td>-0.7644</td>
<td>-0.2721</td>
<td>-1.993</td>
</tr>
<tr>
<td>1500</td>
<td>-0.8136</td>
<td>-0.9275</td>
<td>-1.4722</td>
<td>-0.7658</td>
<td>-0.7248</td>
<td>-1.1008</td>
<td>-0.4758</td>
<td>-1.084</td>
</tr>
<tr>
<td>2000</td>
<td>0.05263</td>
<td>-1.0667</td>
<td>-0.9847</td>
<td>-0.3707</td>
<td>-0.652</td>
<td>-0.9708</td>
<td>-0.2866</td>
<td>-1.2035</td>
</tr>
<tr>
<td>7000</td>
<td>0.61923</td>
<td>0.90946</td>
<td><strong>1.34726</strong></td>
<td><strong>1.4537</strong></td>
<td>0.46765</td>
<td>0.46514</td>
<td>0.12136</td>
<td>-0.1949</td>
</tr>
</tbody>
</table>

Numbers in bold italic font represent significance at $\alpha = 0.1$ ($\alpha = 0.2$ for two-tailed test)

Table 5.4 Verification of Knox’s approach using Mantel’s method (with revised data)

Mantel’s Method using Reciprocal Transformation

The value and justification of using Knox’s procedure is that it allows the avoidance of the extreme variability in space distances and time distances by considering all space distances and time intervals in excess of a critical value as equivalent. Even if he does not differentiate between time and space distances less than his critical values, Knox has gone a long way towards adducing evidence for clustering from his data (Mantel 1967).
However, the very fact that Knox’s procedure does not differentiate the distances within the critical value means that not only is the choice of the critical distance absolutely essential in finding the clustering, but the valuable information about the interpoint and inter-event distances are lost in the analysis. To avoid the use of critical distance, Mantel suggested a data transformation to spread out the short distances while collapsing the range of long distances. With the following transformation, the numerical rather than the categorical values can be employed as spatial and temporal ‘closeness’ values:

\[ X_{ij} = \frac{1}{d_{ij} + k_s} \quad \text{and} \quad Y_{ij} = \frac{1}{t_{ij} + k_t} \]

where \( d_{ij} \) and \( t_{ij} \) are, respectively, distance and time interval between events \( i \) and \( j \), and \( k_s \) and \( k_t \) are arbitrary constants chosen to allow for events that have identical space or time coordinates.

While Mantel’s method with reciprocal transformation avoids the use of critical distance, the choice of the additive constants can affect the outcome of the analysis. If the constants chosen are too small, the region near zero will be unduly expanded with resulting loss of power for detecting clustering. No optimum criteria are known for the choice of the critical constants for either Knox’s or Mantel’s methods (McAuliffe and Affi 1984). In the absence of reasonable intuitive bases for selecting appropriate constants, Mantel suggested a trial-and-error approach using a spectrum of constants (Mantel 1967).
Table 5.5 shows the result of Mantel’s method with reciprocal data transformation. The data used here are the revised alcohol-related traffic accidents (245 cases). The constants $k_s$ and $k_t$ in Table 5.5 represent the additive constants added for space distance and time interval, respectively.

<table>
<thead>
<tr>
<th>$k_s$</th>
<th>$k_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>0.1</td>
<td>0.509754</td>
</tr>
<tr>
<td>500</td>
<td>-0.14357</td>
</tr>
<tr>
<td>1000</td>
<td>-0.03352</td>
</tr>
<tr>
<td>2000</td>
<td>0.0796</td>
</tr>
</tbody>
</table>

Numbers in bold italic font represent significance at $\alpha = 0.05$ ($\alpha = 0.1$ for two-tailed test)

Table 5.5 Result of Mantel’s method with reciprocal transformation (with revised data)

With the revised data, we find three combinations of the additive constants that are statistically significant at a significance level $\alpha = 0.05$ ($\alpha = 0.1$ for two-tailed test). Statistical evidence of space-time clustering occurs when the additive distance constant is $k_s = 0.1$ and when the additive time constant is $k_t = 1$, $k_t = 14$, and $k_t = 21$. It is
interesting to find that the power of the test shows a similar pattern with regard to the combinations of the additive constants to the results from Knox's and Mantel's symmetrical zero-one approaches. Recall that both Knox's and Mantel's analysis results showed at least some meaningful space-time clustering at the distance of 7000 m and the time intervals of up to 21 days. Using a small additive constant has the effect of reducing the influence of the event pairs in close proximity. It is because a small additive constant used in the reciprocal transformation produces a big number, thereby expanding the distance range used in the analysis. Therefore, using a small number for the additive constant has the effect of stretching the distance between close event pairs, i.e., reducing the influence of the close pairs. The fact that the evidence of the space-time clustering occurs exclusively with a low additive distance constant \( k_s = 0.1 \) tells us that when the influence of the events that are very close in space is somewhat reduced by using a very small additive constant, the reciprocal analysis detects the evidence of space-time clustering. In other words, when we classify the events that are within 7 km from each other as being 'close' in space, we can find space-time clustering in the data. The most suitable inter-event intervals that are effective in detecting the clustering are 14 days and 21 days.
CHAPTER 6

CONCLUSION

This research began with a rather simple notion: How do we develop a system with which point data can be effectively visualized and patterns be detected? Further elaboration on this notion brought out the objectives of the study, namely: i) to create an efficient analytical system for the visualization of complex spatio-temporal linear point data; and ii) to statistically test for the existence of space-time clustering in the data.

This chapter summarizes the outcome of the implementation of the prototype system and the statistical analysis. What has been accomplished and what are areas for further improvement? What suggestions can be made concerning the direction of future research? What is the implication of this study for the development of geographic analysis? This chapter attempts to answer these questions.
Results of the research

The most obvious result of this research is the creation of the prototype visualization system. Given the volume and complexity of the data that was used in the study, the system response turned out to be quite impressive. The point visualization portion of the system is well suited for revealing the complex interactions between location and attributes in the data. The use of color for reinforcing the visual representation of certain attributes and also for providing additional information in the display proved to be an effective visualization technique. Interactive query operation coupled with rapid visual feedback provided a great amount of flexibility to examine the data in every possible attribute combination. The network visualization module worked well with the point visualization module by providing different views of the underlying network. The system provides a useful tool for investigating the intricate interrelationships between point location, point attributes and network characteristics.

The prototype visualization system also made significant headway into the means for displaying and analyzing the temporal component of data. The vertical displacement technique provided an intuitive way to detect temporal as well as space-time clusters. The temporal slider provided a direct control over displaying only the relevant data points for a given temporal range. The slide show allows the user to observe the sequence of events in both the space and time dimensions and to detect any anomalies that would have otherwise gone unnoticed.

From the perspective of exploratory data analysis, a tool that can lead the user to a pattern that is not easily detectable is most valuable. The strength of a visualization
system should be evaluated by examining how the system encourages users to try and find unknown, hidden patterns that may exist in the data. By providing an effective interactive visualization system with many easy-to-use tools, the prototype system seems successful in this regard.

When a pattern that interests the researcher is found or when there is suspicion about the possible existence of space-time interaction, we need to employ appropriate statistical analyses to verify the suspicion. The latter part of this dissertation dealt with the statistical methods that are suitable for linear point data. Probably the biggest difference between this research and other studies on space-time clustering is the use of network distance as the spatial closeness measure. The nature of the linear point data makes the results of statistical tests based on Euclidian distance highly susceptible to distortion. Calculating all possible interpoint distances over the network for 253 data points, which results in 63,256 distance measures, was a computation-intensive task. Without the assistance of a high-end computer and the power of GIS, the task would have been nearly impossible to complete. The use of the distance over a network made the outcome of the statistical analyses presented in this study more meaningful.

The findings from the statistical analyses on space-time clustering using alcohol-related traffic accident data are also interesting. Using both Knox’s and Mantel’s method, it was verified that the result of the analyses are influenced by the choice of critical value (the critical distance in space and in time for Knox’s method; the additive constants in Mantel’s reciprocal method). Repeated tests for space-time clustering found that for at least certain combinations of critical values, there is enough evidence to reject
the null hypothesis of no space-time clustering. The critical interpoint distance that produced the evidence of space-time clustering was identified as 7 km, while the critical time intervals were 14 days and 21 days. It is also worth noting that during the first phase of statistical analysis, a hidden fact in the data was revealed. The author did not notice the existence of the same accident that was reported twice when both of the drivers were under the influence. Close examination of the data found 16 cases of this nature. The fact that those cases were found fits very well the spirit of exploratory data analysis. These cases were hidden deep inside the data. If it were not for the statistical test and the subsequent visual check on their locations, they would not have been discovered.

Finally, it is a pleasure to report that the capability of the hardware platform (a Silicon Graphics Indigo 2 Impact workstation) was excellent and that the overall response from the system was very good. On the software side, the IRIS Explorer visualization software within which the visualization system was implemented as custom-built modules proved to be an excellent development environment. The layout and structure of the IRIS Explorer's programming environment was conceptually clean and easy to understand. It is certain that IRIS Explorer visualization software is a good platform on which many sophisticated visualization modules can be built.

Limitations of the research

As a proof-of-concept prototype visualization system, the implemented system is capable of demonstrating the power of visualization for use in spatio-temporal analysis. However, the system has some limitations as well. The first and foremost improvement
that needs to be made to the system is incorporating the statistical procedure within the system. Currently, the statistical analysis has to be carried out outside the system. This creates an interruption in the flow of analysis. With a tight coupling between the visualization module and the statistical procedures, the investigation of statistically meaningful patterns can be much more easily accomplished.

A second limitation of the system is the inability to perform queries for the combined point and network data. For example, there may be a need to select a subset of points that satisfies conditions from both of the data sets, for example, displaying fatal accidents that occurred on two-lane roads only. We can bypass this limitation by including the network information as part of the point attributes. However, that requires manipulation of the data at the source level. For long term solutions, therefore, it would be useful to provide a query capability for combined data attributes.

There is a need to improve the way the output is displayed as well. Currently, the graphics windows does not provide picking capability (choosing a point using a pointing device and getting a response from the system). In other words, when the user require information on certain points, or on a section of network, it would be desirable to allow the user to point directly to it and get the necessary information. To achieve this, using a different data type (for example, ‘Lattice’ data type in IRIS Explorer), rather than geometry data type, needs to be considered. In general, bringing out a pop up window is the most frequently used technique to display additional text information and its implementation should be serious considered.
Finally, there needs to be a more efficient way of calculating interpoint distances. The author could not find any currently available GIS packages that can handle this task efficiently. For example, although ESRI’s Arc/Info GIS does provide powerful network and dynamic segmentation functions, it is extremely difficult to calculate the interpoint distances over the network without extensive modification of underlying network structure. The essence of the problem lies in the fact that the network distance calculation can only be performed between nodes in Arc/Info, and there is no easy way to generate a set of pseudo nodes from a given point data coverage. The fact that many points are in very close proximity also creates a problem in converting point coverage to pseudo nodes on a network. Therefore, there is a critical need to provide a function that can take a point coverage and a network coverage and to generate interpoint network distances from the combination of the two.

Future Research Directions

We are living in a data-rich society. It is certain that our data gathering and warehousing capability will increase exponentially in the future. Rather than being overwhelmed by a mountain of data, a well-designed visualization system can function as an extremely effective tool for extracting valuable information from the data. This study demonstrated how some of the advantages of scientific visualization can be garnered to further our knowledge on spatio-temporal analysis of point patterns. Future research should focus on applying various leading-edge technologies toward the improvement of the state of spatial analysis.
One thing that naturally comes to mind is expanding the current visualization system to a general point visualization system that can handle not only linear point data but area-based point data as well. Developing a generic point visualization system that is capable of handling a multitude of data sources without the conversion and manipulation of data at the source level would be the next immediate challenge. There are many hurdles to overcome to achieve this goal. Data input functions should be vastly improved to allow direct access to data stored in digital form. More suitable data formats for visualization system should be developed and the means through which the data inside the system is manipulated and used to generate the necessary geometry should be improved.

On the theoretical side, the effectiveness of various visual representation schemes for point and network display need to be studied. Formal evaluations on the effectiveness of visual representation schemes would greatly help to refine available methods. At the same time, the search for new and innovative visual representation methods should continue.

The need for integrating statistical procedures within the system was already mentioned. In addition, there is a need to more actively incorporate EDA techniques in the system and to develop a link between EDA methods and existing visualization methods. As for statistical testing to detect clusters, more elaborate procedures that can test the existence of spatial clusters in the linear data are needed. The modification of Ripley’s *K-function* analysis for linear data would certainly be a possibility.
As more powerful computing capabilities and more high-end graphic workstations become available on spatial analysts' desktops, it is certain that visual analysis and visualization systems in analyzing complex spatio-temporal data will play a more and more useful role in diverse research areas. It is the author's belief that there will soon be a day when using a fully-integrated, powerful visualization system in spatial analysis is no longer a novelty but the norm.

ACSM, 1992, Geographic Visualization, *Cartography and Geographic Information Systems*, 19(4),


177


McAuliffe, T. L., and Affifi, A. A., 1984, Comparison of a nearest neighbor and other approaches to the detection of space-time clustering, *Computational Statistics & Data Analysis* 2, 125-142.


183


Tobler, W.R., 1994, *Cholera: a BASIC program showing the Dr. Snow’s map and Thiessen polygon boundaries around the pump.*, In Santa Barbara, CA: National Center for Geographic Information and Analysis.

185


Warner, J., 1990, Visual Data Analysis into the '90s, *PIXEL*, 1(January/February), 40-44.


APPENDIX

Data dictionary for traffic accident data used in the study.

Table Name: ACCIDENT

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPLN</td>
<td>1</td>
<td>F-Fatal Accident</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Injury Acc./Non-fatal</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Property Damage Only</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Non-Reportable</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Private Property</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>Not stated</td>
</tr>
<tr>
<td>DAY</td>
<td>0</td>
<td>Not stated</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Monday</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Tuesday</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Wednesday</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Thursday</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Friday</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Saturday</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Sunday</td>
</tr>
<tr>
<td>ALCOHOL</td>
<td>0</td>
<td>No alchol involvement</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Alchol or drugs detected</td>
</tr>
<tr>
<td>HITAC</td>
<td>0</td>
<td>No hit and run</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Hit and run</td>
</tr>
<tr>
<td>HUMID</td>
<td>0</td>
<td>Not stated</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Dry</td>
</tr>
<tr>
<td>Wet</td>
<td>Muddy</td>
<td>Snowy</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not stated</th>
<th>Clear</th>
<th>Cloudy</th>
<th>Raining</th>
<th>Snowing</th>
<th>Fog, smog, smoke, dust</th>
<th>Sleet or hail</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Single vehicle accident</th>
<th>Ran off road</th>
<th>Hit fixed object</th>
<th>Hit non-fixed object</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi-vehicle accident</th>
<th>Car vs car</th>
<th>Car vs truck or bus</th>
<th>More than two vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not stated</th>
<th>00 - 29 mph</th>
<th>30 - 49 mph</th>
<th>50 - 79 mph</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Table Name: ROADCLASS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Rural principal arterial -</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Rural principal arterial - Other</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Rural minor arterial</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Rural major collector</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Rural minor collector</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Rural local</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Urban principal arterial -</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Urban principal arterial - Other</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Urban principal arterial - Other</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Urban minor arterial</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Urban collector</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Urban local</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>1-8</td>
<td>Actual number of lanes</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Nine or more lanes</td>
</tr>
</tbody>
</table>