VISUAL ANALYSIS

OF

MULTIPLE ROUTE DEVIATIONS

by

DEMET YILMAZ

Submitted to the Graduate School of Engineering and Natural Sciences

in partial fulfillment of the requirements for the degree of

Master of Science

SABANCI UNIVERSITY

Spring 2010

VISUAL ANALYSIS OF MULTIPLE ROUTE DEVIATIONS

APPROVED BY:

Assist. Prof. Selim Balcısoy …………………………

(Dissertation Supervisor)

Assoc. Prof. Albert Levi …………………………

Assoc. Prof. Burçin Bozkaya …………………………

Assist. Prof. Hüsnü Yenigün …………………………..

Assist. Prof. Nihat Kasap …………………………

DATE OF APPROVAL: …………………………

VISUAL ANALYSIS OF MULTIPLE ROUTE DEVIATIONS

Demet Yılmaz

EECS, M.Sc. Thesis, 2010

Thesis Supervisor: Assist. Prof. Selim Balcısoy

Keywords: Spatio-temporal Data, Visualization of Multiple Routes, Route Deviation, Glyph, Level of Detail

ABSTRACT

With the collection of excessive amounts of movement data, the analysis of routes has become very important in the exploration of these datasets. To reveal the hidden information in the data, many visualization approaches are proposed according to analysis type. The analysis in this thesis involves visualization of route deviations which occur due to several reasons such as traffic congestion, accidents, weather conditions, or operational issues. The proposed system integrates two visualization techniques for multiple routes visualization: circle-glyph representation and route coloring. While route coloring emphasizes deviations on general trends, circle-glyph visualization concentrates on the detailed analysis of deviations. Level of detail is applied in some visualization parts. A usability study with statistical analysis of its results is performed to evaluate the success of the proposed approaches.

ÇOKLU ROTA SAPMALARININ GÖRSEL ANALİZİ

Demet Yılmaz

EECS, Yüksek Lisans Tezi, 2010

Tez Danışmanı : Asist. Prof. Selim Balcısoy

Anahtar Kelimeler: Yer-Zaman verileri, Çoklu Rotaların Görşelleştirilmesi, Rota Sapması, Glif, Detay Seviyeleri

ÖZET

Rota analizi, büyük miktarlarda hareket verilerinin toplanmasıyla önem kazanmaktadır. Bu verilerde saklı bilgileri ortaya çıkarmak için farklı analiz amaçlarına yönelik çok farklı görselleştirme metodları önerilmiştir. Bu tezdeki analiz, rota sapmalarının görselleştirmesini konu alır. Rota sapmaları; trafik sıkışıklığı, trafik kazaları, hava koşulları ya da işletme ile ilgili bir çok nedenden ötürü oluşabilir. Bu tezde önerilen sistem, çoklu rota sapmaların görselleştirilmesinde iki teknik sunar: daire-glif gösterimi ve rota boyama. Rota boyama, sapmaları genel şekilde ifade ederken, daire-glif gösterimi sapmaları detaylı bir şekilde gösteren bir tekniktir. Bazı görselleştirme kısımlarında detay seviyeleri tekniği kullanılmıştır. Önerilen yaklaşımların başarılarını ölçmek için istatistiksel yöntemler kullanarak kullanılabilirlik testleri yapılmıştır.

To my parents...

ACKNOWLEDGEMENT

I wish to express my gratitude to my supervisor, Selim Balcısoy, whose expertise, understanding, and patience, added considerably to my graduate experience. I am grateful to him not only for the completion of this thesis, but also for his unconditional support from the beginning.

I am greatly indebted to Burçin Bozkaya who has given his valuable ideas throughout this thesis and made suggestions for the editing of my thesis.

I am grateful to my thesis committee members Albert Levi, Burçin Bozkaya, Hüsnü Yenigün and Nihat Kasap for their valuable reviews and comments on the dissertation.

I would like to express my special thanks to my family for their all valuable supports and belief in me. I would like to thank all my friends and colleagues, particularly to Süheyla Çetin, Ali Demir and Sabahattin Öge for their friendship and assistance.

Finally, I would like to thank TÜBİTAK-BİDEB for their financial supports during my M.S. studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE OF ABBREVIATIONS

GIScience	Geographic Information Science
POI	Point of Interest
LOD	Level of Detail

1. INTRODUCTION

1.1. Visualization : An Overview

Why should we be interested in visualization?

Because, the human visual system is a pattern seeker of enormous power and subtlety [1].

Visual system has its own rules formed between the eye and the visual cortex of the brain. If data is presented according to these rules, important and informative patterns can easily be perceived; if not, our data becomes incomprehensible or misleading. Hence the main point is understanding how perception works and reflecting this to the representation of data in a proper way [1].

With the aid of computers, visual representations are not limited to statistic illustrations of data but can be made dynamic and interactive. There is a variety of methods derived by these ways. Especially, in the last years, visual representations are widely found on the Web. Many visual analysis tools of data are available, for instance ManyEyes [2] is a platform where users can upload their data, create, and share visualizations, and discuss them.

Despite the existence of many visual analysis tools, there is a huge diversity in terms of dataset contents and the same dataset can be analyzed with different perspectives. One type of dataset vastly dealt with is movement data which has both spatial and temporal components. By means of current tracking technologies, movement data sets have been recently collected in huge amounts. Many researches are performed on visualization of movement data, such as pattern extraction from the behaviour of moving objects [3], finding interactions between moving entities [4], spatio-temporal aggregation of trajectories [5] etc. Each one considers movement data from a different angle and proposes different solution techniques. This thesis handles movement data yet from another perspective by proposing a novel visualization system to visualize deviations existing on movement data.

The proposed visualization system is developed through the study of multiple routes visualization and applying appropriate techniques to see the route deviations which take place due to several reasons, such as traffic congestion, accidents, weather conditions, or operational issues.

1.2. Problem Definition and Solution Approach

Interpreting the movement of a mobile object such as a person, an animal, a vehicle, or a particle requires analyzing the datasets in detail. Since, the data is available in huge amounts, proper visualization techniques need to be devised. Route visualization becomes important at this point. Up to now, many techniques have been applied in route analysis area, but the deviation visualization of routes is not formally defined and studied. This study looks at the issue from a different perspective and aims to form a basis in the deviation visualization frame.

Visualizing movement data is not a trivial task if you have multiple routes which are formed by the entities in the movement data set. Displaying them just as they are, (using their spatial information) would give a cluttered scene which will probably make the understanding and exploration of the dataset impossible. Movement data used in this thesis includes multiple routes and many points to be served on these routes. Each point should be served by one vehicle travelling on its route within a time constraint. If the vehicle does not visit order points on time and by following the route assigned to it, *route deviations* occur.

Defining a visual representation to explain route deviations is the main problem handled in this thesis. So, some methods are developed to visualize this scenario. Displaying routes, vehicle paths, and order points are arranged by the controls on the interface. For order points, levels of detail are defined and according to the change in levels of detail due to the position of camera, order points are clustered to prevent visual complexity. For representation of routes and vehicle paths, different colors and sizes are employed. In order to comprehend deviations, two methods are developed: circle-glyphs and route coloring. In each one, different visualizations are created to reflect deviations in particular ways.

1.3. Summary of Contributions

The proposed system integrates several concepts and visualization techniques in visualizing route deviations. The main contributions of this research are:

- Introducing route deviation concept in analyzing movement data.
- Developing definitions in deviations, and their realizations on a visual platform in real time.
- Providing users the opportunity to choose between different approaches when displaying multiple route deviations.
- Offering a highly interactive environment which visualizes (some part of) data at different detail levels according to the degree of visual cluttering.

1.4. Thesis Outline

This chapter briefly describes why visualization is important and draws attention to the need for visualization of movement data. It also gives an abstract about the visualization system developed and the main aspects of approaches used. Finally, it offers a list of contributions. Subsequent chapters contribute to the thesis as follows:

Motivation and Related Work: The second chapter gives a brief historical background of information visualization and route visualization in geographical visualization context. Some examples of visualization systems are presented.

Visualization System: The third chapter describes details of the proposed system and explains the approaches used in the system.

Case Study: The fourth chapter demonstrates the usability study performed to test users' perception on our two approaches developed in displaying route deviations on multiple routes.

Results and Discussions: The fifth chapter denotes the accomplishments and limitation of the work and presents the results. It also refers to further works and possible improvements for the method.

2. MOTIVATION AND RELATED WORK

2.1. Historical Background

"A picture is worth ten thousand words."

As the old Chinese proverb says, pictures have an important role in people's life beside words. Whatever the era, visual elements have been used as a medium in expression of thoughts reflecting its own times. Some wall paintings were discovered in Çatalhöyük which was the centre of advanced culture in the Neolithic period in Konya, Turkey. Paintings discovered during excavations in Çatal Höyük illustrate animals and hunting scenes [6].

Figure 2.1 A Wall painting of a bull, deer, and man, Çatalhöyük, Turkey

As the needs in human life change, visual stories told by pictures undergo various changes. With the evolution of technology, hand drawings are not as needed as before; instead, computers have been used. Although visualization is not strictly bound to computers, it progresses as technology evolves. Also, it was not until the 18th century that we see the beginning of the charts and graphics used today. [7] Techniques and instruments for precise observation and measurement of physical quantities were being developed by the 16th century. It was the 17th century when growth was seen in the theory and in practice of analytic geometry, theories of errors of measurement, and the birth of probability theory. Over the 18th and 19th centuries, information about people i.e. social, moral, medical, and economic statistics began to be gathered in large amounts. These progresses gave rise to innovations in visual thinking; diagrams were used in illustration of mathematical proofs and functions. The close relation of the numbers of the state (the origin of the word "statistics") and its geography made the visual representation of such data possible on maps (thematic cartography) [9].

Data collection in huge amounts and wide varieties yielded different techniques to emerge. The need in obtaining understandable information from this data resulted in the appearance of different research areas in visualization.

2.2. Information Visualization

Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information [8].

Although the definition relates information visualization with technology, considering the place of visualization in historical developments, it can be considered to cover all of the previously described techniques in a broader concept. In this sense, information visualization takes us back to the earliest scratches of forms on rocks and the earliest use of diagrams in the history of science and mathematics [9].

Today, information visualization has a great range of application areas covering every industry and all tasks where understanding of the intrinsic structure in data is crucial. Some noticeable examples are in the field of economical/financial analysis, representation of large hierarchies, medical training/assistance, and engineering/physics [10].

Information visualization techniques differ in great deal such as X-Y plots, line plots, and histograms. A classification technique explained in [11] handles classification by basing it on three criteria: the data to be visualized, the visualization technique, and the interaction technique used.

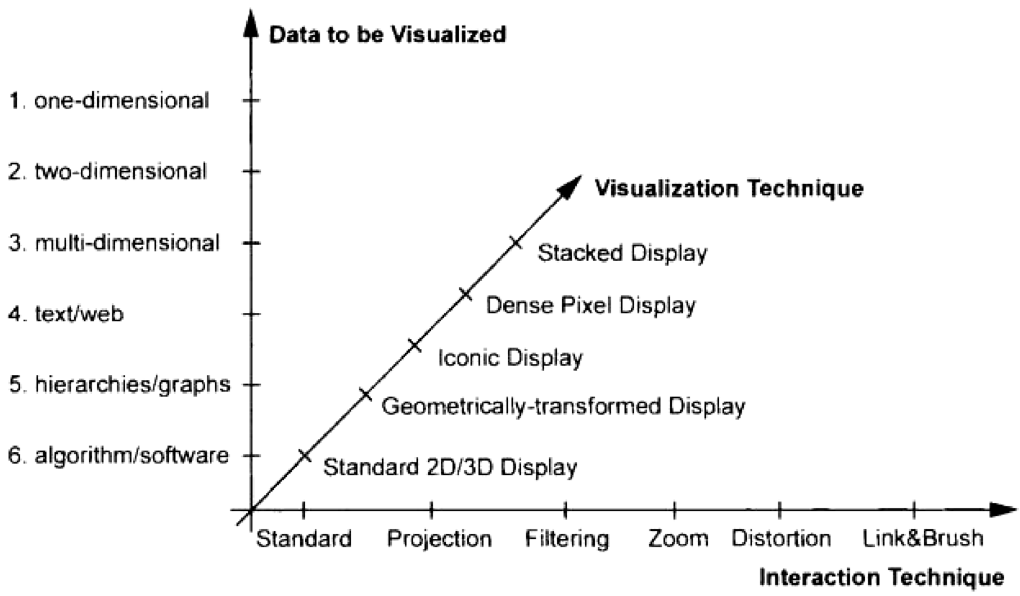

Figure 2.2 Classification of Information Visualization Techniques [11]

The data type to be visualized (Shneiderman, 1996) may be:

- One-dimensional data such as temporal (time-series) data,
- Two-dimensional data such as geographical maps,
- Multidimensional data such as relational tables,
- Text and hypertext such as news articles and web documents,
- Hierarchies and graphs such as telephone calls and web documents, and
- Algorithms and software such as debugging operations.

The visualization technique used may be classified as:

- Standard 2D/3D displays such as bar charts and X-Y plots,
- Geometrically transformed displays such as landscapes and parallel coordinates,
- Icon-based displays, such as needle icons and star icons,
- Dense pixel displays, such as the recursive pattern and circle segments, and
- Stacked displays, such as treemaps and dimensional stacking.

The third dimension, interaction techniques, allows users to directly navigate and modify visualizations, as well as select subsets of the data for further operations. Some of the examples are:

- Dynamic Projection,
- Interactive Filtering,
- Interactive Zooming,
- Interactive Distortion,
- Interactive Linking,
- Brushing.

These techniques are useful for data exploration but limited to relatively small and low dimensional data sets. In the last decade, a large number of novel information visualization techniques have been developed, allowing visualizations of large volumes of multidimensional data sets.

2.3. Human Visual Perception and Visual Mapping

Vision is a complicated system that requires numerous components of the human eye and brain to work in accordance. When the photoreceptor neurons (called photoreceptors) in the retina collect light, they send signals to a network of neurons which then generate electrical impulses that go to the brain. After the brain processes impulses, we are given information about what we see. [12]

What we see may not contain so meaningful information all the time because data may not be so simple to understand at a glance. Also the human vision system is not capable of understanding complex scenes. Hence these complex scenes are required to be made perceptible by applying proper representation methods. The significance of visualization techniques come into the picture at this stage. These techniques involve a mapping which is specifically called as *visual mapping*. It is a mapping process between data aspects and visual variables, i.e., assigning specific visual characteristics to data attributes in order to facilitate visual sense-making. [13] Visual characteristics should be assigned so that it should reflect its own attribute clearly. Visual variables are used in this context. Graphic designer Jacques Bertin's vocabulary of the graphical techniques is commonly used to encode information in presentation graphics [14]. In graphical presentations, graphical marks are used such as points, lines, areas, surfaces, and volumes to encode information. They have modifiable attributes such as position, size, shape, value, color, orientation, and texture which are called as Bertin's visual variables.

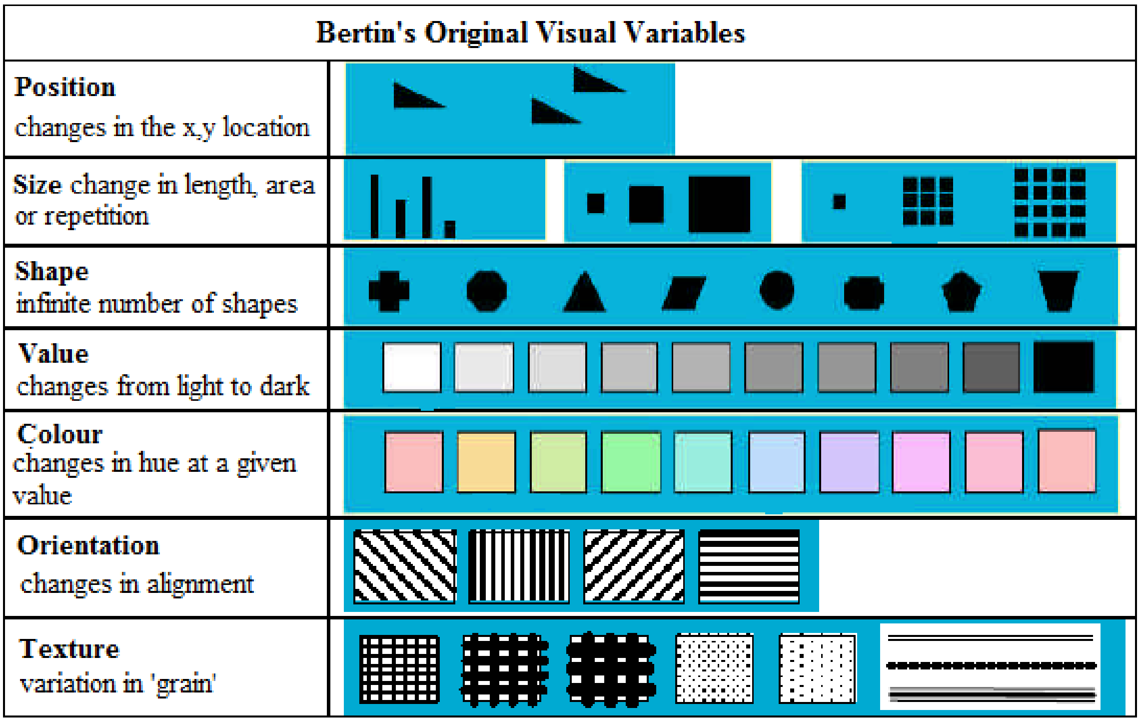

Figure 2.3 Bertin's Original Visual Variables [14]

Mackinlay expanded Bertin's variables [15]. He also investigated the effectiveness of visual variables and suggested different orderings based on the task. Before Mackinlay, Cleveland and McGill provided an observational study which stated that people accomplish perceptual tasks associated with interpretation of graphical representations with different accuracy levels. But they did it only for quantitative cases; their ordering does not handle the encoding of qualitative information. Mackinlay extended their ranking using existing psychophysical results and various analyses of different perceptual tasks.

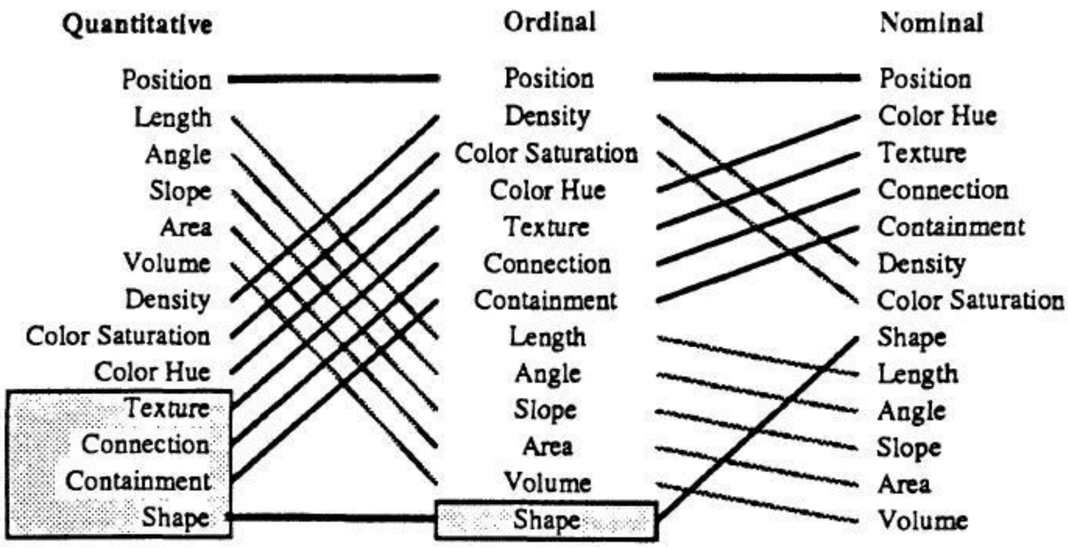

Figure 2.4 Mackinlay's ranking of perceptual tasks [15]

Same information can be represented by different visual variables but it has a substantial influence in the perception. Hence it is very important to use appropriate variables in representations.

2.4. Maps and Geographical Visualization

The physical world is a commonly used subject in visualization. When geospatial data is collected in progressively larger size, geographical visualization comes into prominence. Geographical visualization, geovisualization, in short, addresses the visual exploration, analysis, synthesis, and presentation of geospatial data by integrating approaches from cartography with those from other information representation and analysis disciplines, including scientific visualization, image analysis, information visualization, exploratory data analysis, and GIScience [16].

Cartography, the study of map making, is substantially related with geovisualization. Geographical information is kept beneath maps and can be made noticeable by certain techniques. Since static maps have a limited exploratory capability, geovisualization techniques, on the other hand, enable users of interactive maps to explore different layers of the map, zoom in or out, and change the visual appearance of the map, usually on a computer display [17].

2.5. Route Visualization

When the location information of entities became more important by widespread use of Location-based services (LBS), a variety of visualization techniques appeared on display devices such as computers, mobile phones, or other such devices. In the last decade, digital maps such as those provided by Microsoft Live (www.live.com), Google Maps (maps.google.com) and Yahoo Maps (maps.yahoo.com) have become increasingly popular [19]. Due to the complex nature of movements, providing insight to user is not very trivial. Hence, a variety of techniques are proposed in this area. In this sense, route visualization has been one of the most popular fields of interest in geovisualization.

One contemporary application of geovisualization is related with route maps. Route map is a depiction of a path from one location to another [18]. They have been widely preferred in forming driving directions through the Web recently. This paper aims to improve the usability of route maps by automatically designing and rendering them by putting maps into a clearer form. The designed system, LineDrive, is found preferable by user feedbacks than standard computer generated route maps.

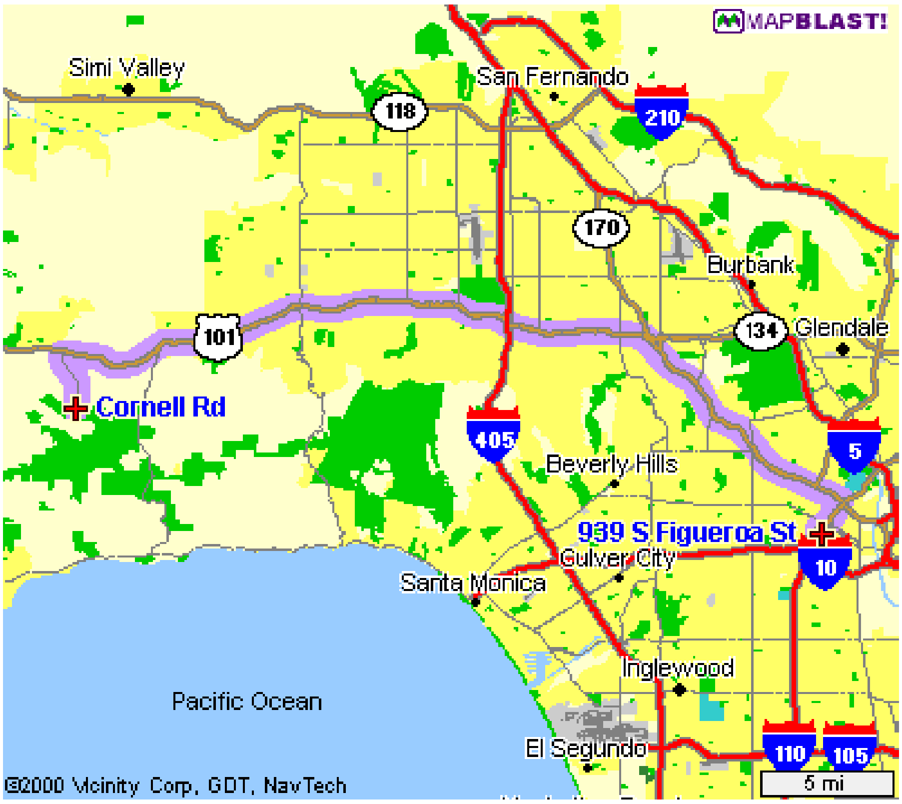

Figure 2.5 Standard computer-generated route map [18]

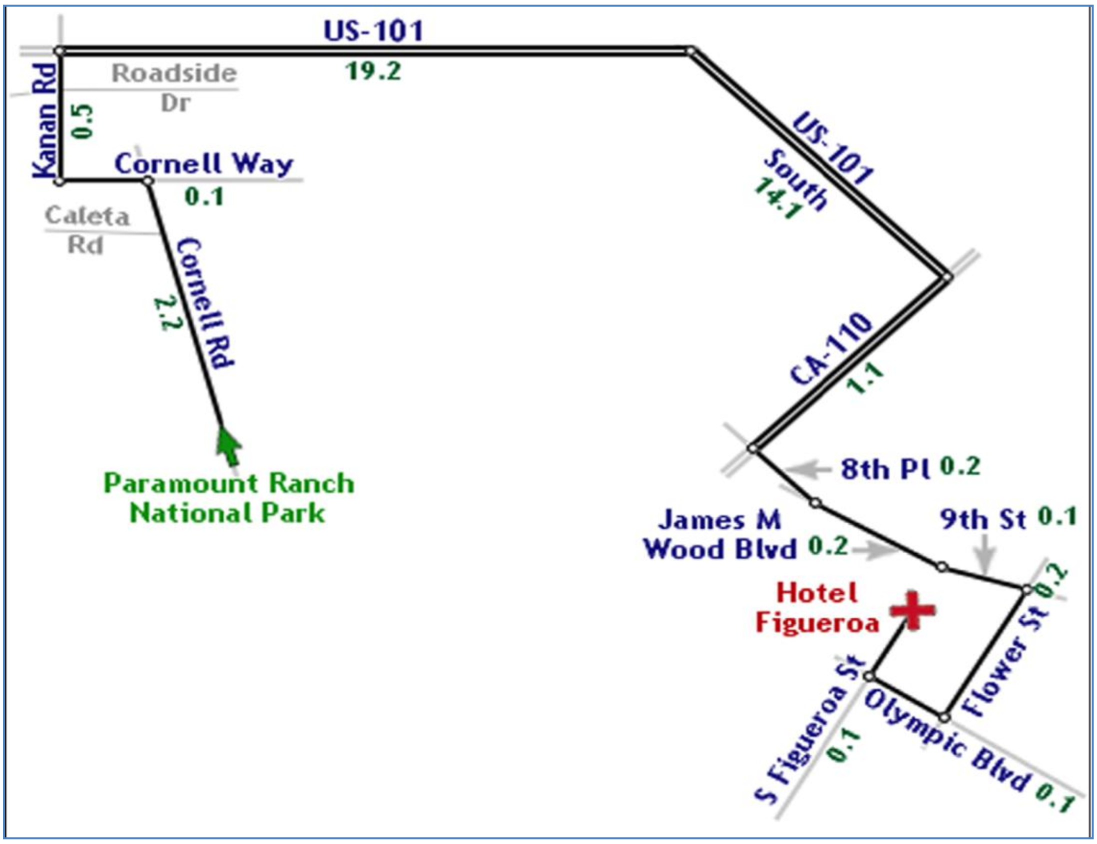

Figure 2.6 LineDrive map [18]

Another application is developed for tourist map creation [19]. An automated system has been developed for designing tourist maps that selects and highlights the information most important to tourists. By using image analysis and web-based information extraction techniques, a map that emphasizes the most important landmarks is generated. Cartographic generalization techniques such as simplification, deformation, and displacement are also used.

Routes may lack of geographical location information. But still they have spatial information. Another similar term used in this concept is *trajectory*, the path a moving object follows through space as a function of time [20]. This definition also corresponds to the term *spatio-temporal* since a trajectory has both space and time component. As shown in Figure 2.7, X-Y plane corresponds to space and Z axis to time component.

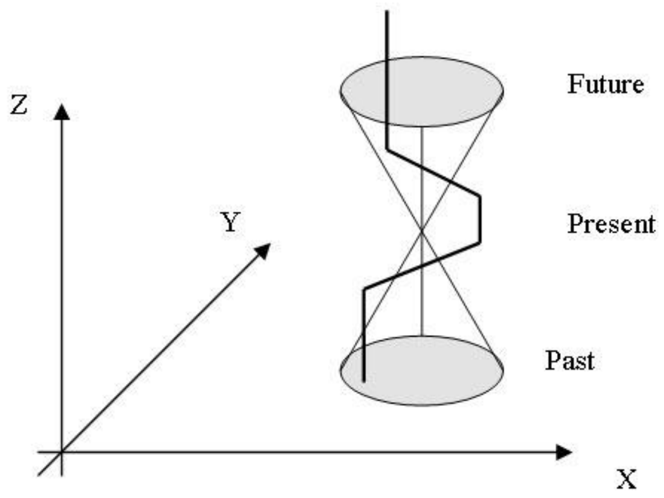

Figure 2.7 Space and Time Path [21]

Trajectories are widely studied elements in this field. One such study [22] investigates visualization of trajectories existing in large amounts. Since traditional visualization techniques such as space-time cube cannot provide effective visualizations, this study proposes visualizing them in a summarised form. The aim is to find a way for automated generation of graphical models of groups of similar or related trajectories. By using a partitioning method of the territory, aggregate moves are obtained.

Figure 2.8 On the left: about 1500 trajectories of cars moving in the city of Milan, on the right: Resulting graphical model for the cluster of trajectories shown on the left [22]

Another related application is about storytelling in geotime [23]. This study aims to analyze behaviour of mobile objects in time and geography within a single highly interactive 3D view. Seeing events, connections, and movements in a combined temporal and spatial view allows analysts to detect, examine, and understand patterns much more quickly than with traditional analysis tools.

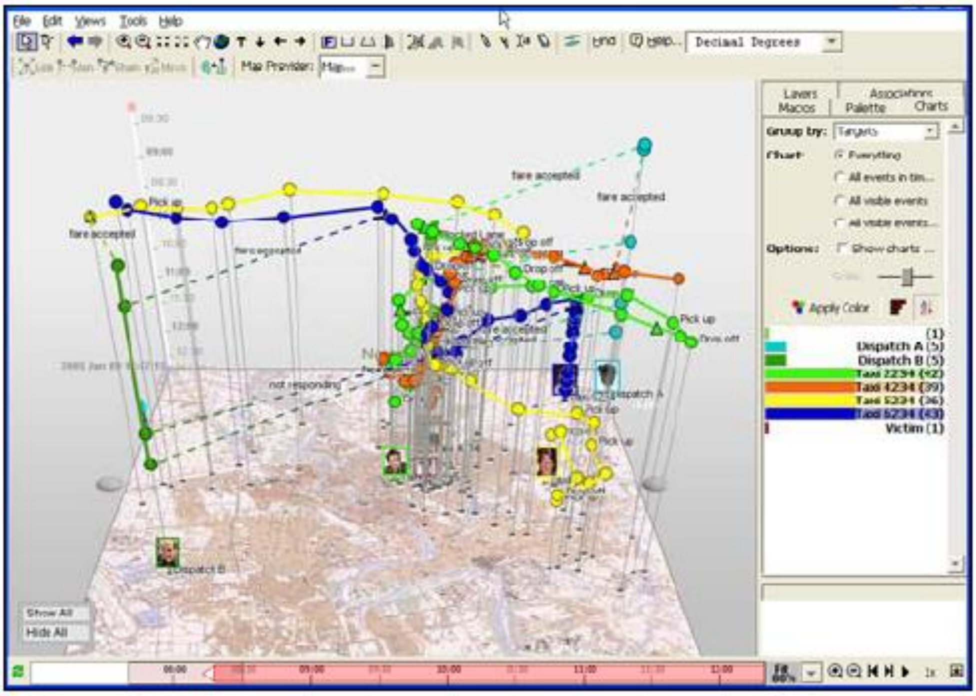

Figure 2.9 GeoTime for analysis of events in time and space [23]

2.6. Glyphs in Information Visualization

Glyphs (also referred to as icons) are graphical entities that convey one or more data values via attributes such as shape, size, color, and position [24].

An object or symbol for representing data values. Glyphs are generally a way of representing many data values and are sometimes called icons. A common glyph is the arrow, often chosen to represent vector fields. The arrow depicts both speed and direction at a point. (Keller and Keller, 1993)

The most prominent example for glyphs is the Chernoff Faces [Chernoff, 1973]. Chernoff used the different parts of human face such as mouth, nose, eyes, eyebrows, head etc. in order to encode different attributes of multidimensional data set. Glyphs are graphical objects that can be derived according to multidimensional data structure to be visualized. They are well suited for displaying complex, multivariate data sets since they make representations easy to understand. Since multiple attributes are combined in a single graphical unit, data set is more comprehensible.

In his study [24], Ward presents an overview of multivariate glyphs and provides taxonomy on glyph placement strategies, namely, about how to position glyphs on the screen. Once a glyph has been designed and generated, it must be placed at a location in display space (2-D or 3-D). The position of glyphs can be very effective in improving the detection of similarities, differences, clustering, outliers, or relations in the data. Methods also vary as to whether they allow overlapping glyphs, whether they are space-filling, or whether they employ empty space to highlight differences. Some glyph examples are given below.

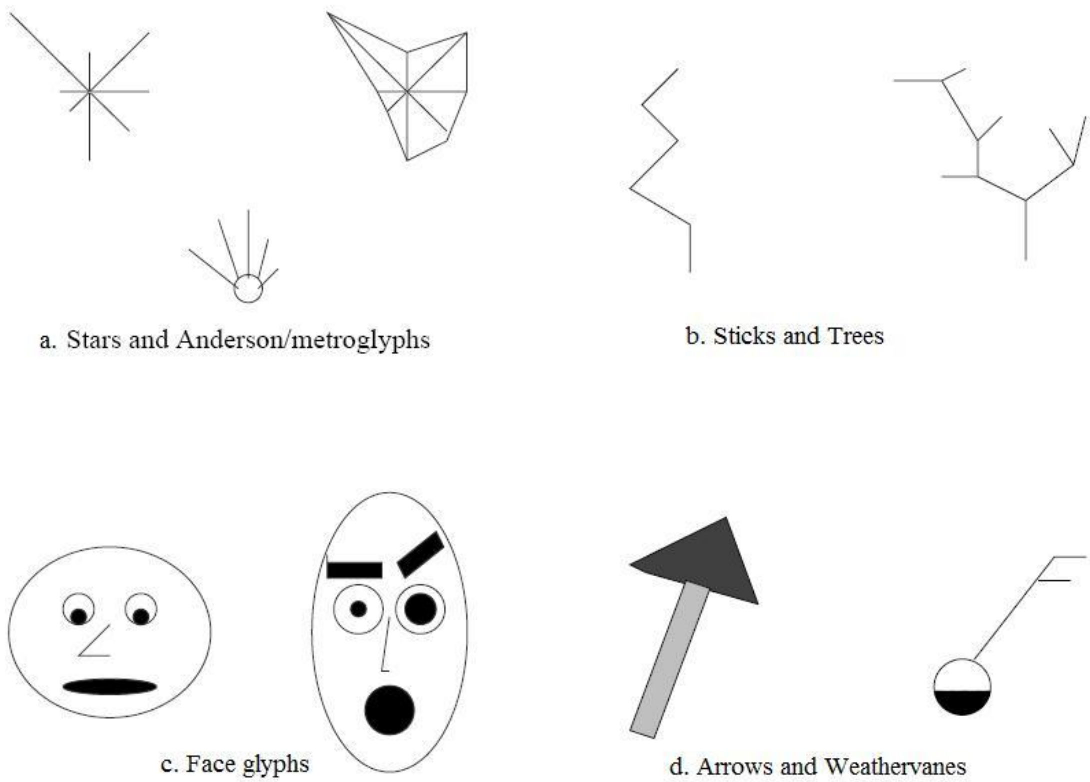

Figure 2.10 Some glyph examples from Ward [24]

3. VISUALIZATION SYSTEM

The goal of this study is to develop a comprehensive system to visualize route deviations related with vehicles. Today, various organizations use a fleet of vehicles to service a set of orders. For example, a big bread company might use several trucks to deliver breads to their customers' shops or markets. A cargo company might schedule daily routes to deliver items to their customers' homes. To carry out these transportation services efficiently, it is necessary for organizations to understand the benefits of vehicle management. Vehicle management can include a range of functions, such as vehicle financing, vehicle maintenance, vehicle telematics (tracking and diagnostics), driver management, fuel management and health & safety management [25]. With rising fuel costs, vehicle tracking has become more of an issue. Hence, businesses look for increased efficiency and cost savings. Due to the several reasons such as traffic congestion, accidents, weather conditions or operational issues; route deviations occur. Although, preventing the first three reasons may not be possible, problems arose from driver behaviours may be evitable.

Once the vehicle tracking data has been collected, managers want to be informed of vehicle behaviours. In this thesis, we considered that if one wants to keep track of vehicular deviations, visualizing this event will simplify understanding the scenario. In this sense, we provided a single united platform for obtaining all imperfect routes. So, we developed some methods to reflect route deviations on visual platform.

Some of the definitions used in this context are given below:

- Route : Scheduled path that a vehicle has to travel.
- Vehicle path : The path that this vehicle has actually travelled.
  Expected case: vehicle path should be exactly same as its scheduled route.
- Route deviation: It is the difference measure of vehicle path from its scheduled route according to some criteria which will be explained.

- Route deviation visualization: aims to visualize the deviation of routes using proper visualization techniques.

3.1. Event Description

Typically, a route has one or more order or customer points to visit. These points are also called as POIs (point of interest), which can be found useful in a specific context. On each route, one vehicle travels and delivers its items to these order points. Hence, the vehicle makes mandatory stops at these fixed order stops and will adhere to the constraint of arriving on schedule at these stops. Again, the rule is that the vehicle should;

- make its deliveries to all order points on time
- follow its scheduled route exactly.

If the above rules are not respected, then a deviation occurs. Possible deviations considered in this study are:

- Distance deviation (or km deviation): How many kilometers the vehicle deviates from its scheduled route. Simply, it is the distance difference between scheduled route and actual vehicle path in kilometers.

- Time deviation: How many minutes the vehicle deviates from its agreed time. The vehicle can be late or early at its arrival to order points. But, one is typically interested in latency to see problematic ones, and calculations are made based on late ones. There are three kinds of time deviations defined for a route:

    - **Simple Time Deviation** : Vehicle travels from an initial point to a final point on its route. Usually, initial and final points are depots. Simple time deviation is about arriving at the final point on time. That is, it is the minute difference between (pre-defined) arrival time at the final point and the actual time vehicle arrives at this point. If the vehicle finishes its journey on time, then simple time deviation is zero.

    - **Individual Time Deviation** : It is the sum of time deviations for order points arrived late.

- **Average Time Deviation** : This is another measure in time deviations which is individual time deviation over number of points having time deviations (points arrived late). It aims to find an average time deviation value for those points having time deviations.

  (individual_time_deviation / number_of_time_deviated_points)

3.2. Input Data

There are four types of input data used in the evaluation of route deviations:

- Route data:

  Route includes location information about all points forming the route. Each point has location information as latitude and longitude values.

- Vehicle trajectory data:

  Vehicle data includes location and time information. Each data record indicates that vehicle passes from a particular location at a particular time.

- Order point data:

  Order points have location and specific time value so that the order point should be served at this time value.

- Route Information data:

  A file containing general information about routes such as total route length, number of order points on this route etc.

Since, the aim is to handle multiple routes, 20 routes is considered enough and 20 routes' data is generated and visualized in this study.

3.3. Visualization

Visualization process of our study is simply generating a scene with routes, vehicles, and order points with the underlying map. These components can be set visible/invisible from the interface. According to what is being examined, one can turn off the visibility of any of them. The background map can also be set as visible/invisible. A general view from the system is given below:

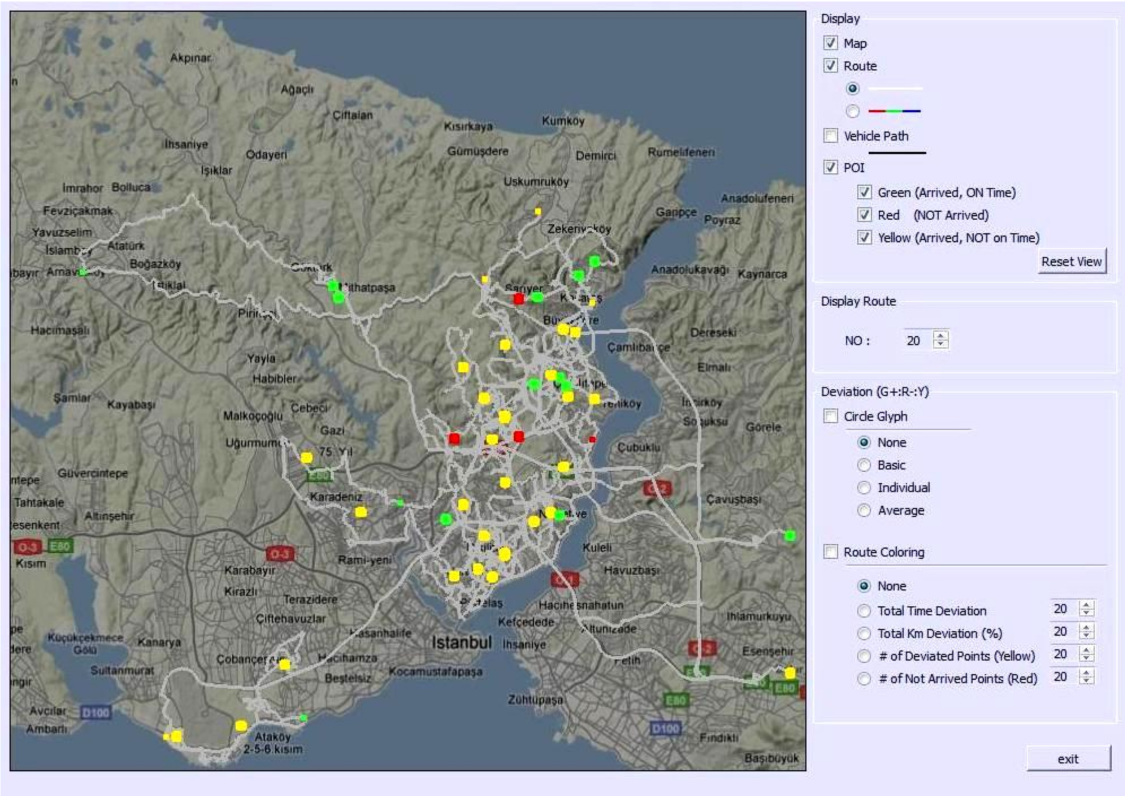

Figure 3.1 General view from the system

Routes can be displayed either in same color (white) or in multiple colors. For vehicle paths we use the same color (blue) for all vehicles not to confuse their colors with those of their routes (if routes are in multiple color mode).

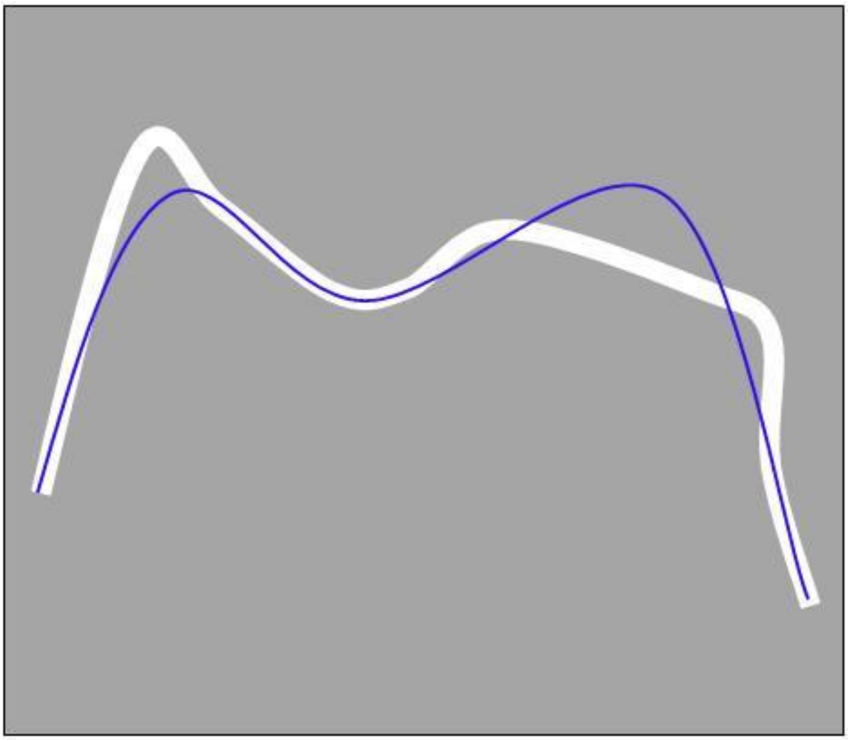

Figure 3.2 Representations of proposed route (white) and travelled vehicle path (blue)

For order points there are three cases in terms of their arrival information. Basically, an order point is arrived on time, not arrived, or arrived but not on time. So we assign colors to each one: *Red* for not arrived points, *Green* for points arrived on time or early, *Yellow* for points arrived late. We use this approach, since this scenario resembles the case in stoplights where red indicates an unsafe situation (which is the stop signal), green a safe situation (which is the go signal) and yellow a caution signal. Likewise, assigning meaningful daily-life colors to order points eases understanding instead of using different shapes for three kinds of points (POIs) or using another distinguishing method.

One challenge in POI visualization is that POIs can be so close to each other which can cause visual cluttering. In order to prevent this, we cluster the ones which fall close enough to each other within the level of detail framework (LOD). LOD is a tool which involves decreasing the complexity of a scene as it moves away from the viewer [26]. At each LOD, distances between each point are calculated and points falling into same cluster are merged to form a single point. Distances between each point are calculated in terms of pixel distance at each LOD since pixel distance between two points is directly related with LOD. When you zoom in, pixel distance between two points increases, and zoom out, it decreases under perspective. The size of the clustered point is bigger relative to the single point for distinguishing because it includes more than one point. The route in Figure 3.3 includes two small POIs (from top to bottom: green, yellow), and four clustered POIs (from top to bottom: red, red, yellow, yellow ones). The location of clustered POI is set by points' average coordinate value which constitutes this clustered POI. That is why some POIs in the figure are not drawn upon the route. Their locations have a meaning such that there are some points around these locations not exceeding the limit distance related with pixel closeness.

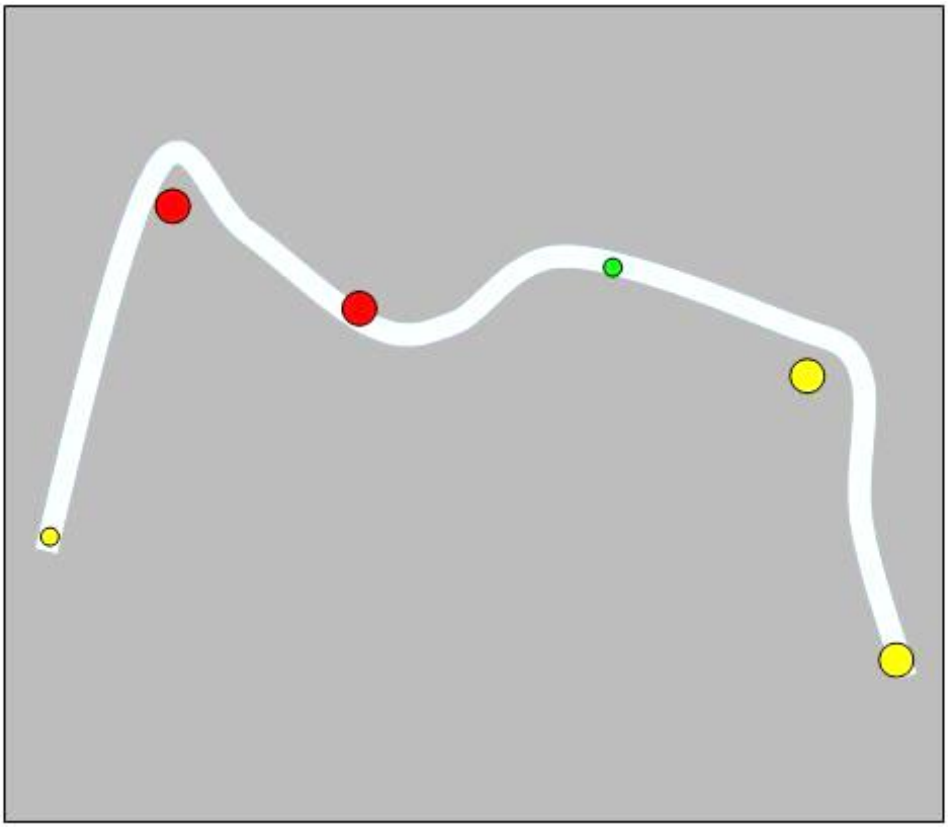

Figure 3.3 Representations of clustered and single POI(s) of a route

Colors for these clustered POIs are assigned according to the dominant point in each cluster. By dominant, we mean the most problematic ones, that is, red points dominate yellow points and yellow ones dominate green ones. If the viewer wants to see the points that are not arrived, s/he zooms in to a small/big red POI. If this is a clustered POI, this tells that there is at least one red POI in that cluster. If the viewer wants to see points which are arrived late, s/he zooms in to a small/big yellow POI. If this is again a clustered POI, this tells that there is at least one yellow POI and no red POI in that cluster. Finally, for big green POIs, it can be derived that, all points in this cluster are green meaning that they are all reached on time (or early).

3.3.1. Deviation Visualization on Single Route

On a single route, there are points where deliveries should be made as explained in the previous sections. Since, there are two kinds of deviations defined, km deviation and time deviation; both should be observed on a single route.

Before explaining how the deviation problem is addressed, it should be emphasized that the analysis in this study does not try to cover the deviation at every point (every single point which form the route or vehicle path) but only the order points' deviation. The focus in time deviation is whether the vehicle arrives at POIs on time or not: if a POI is not arrived on time, how many minutes the vehicle is late at that location. For km deviation the focus is whether the vehicle exactly follows its pre-defined route: if not, where the deviations occur.

In order to visualize km deviation for a single route, different approaches can be used. Since km deviation relates with vehicle path and its actual route, one should wonder overlapping and non overlapping parts. Ideally, vehicle path and its route should be identical and overlap everywhere. If the case is not like this, detecting non overlapping parts should be easy. In detection of these parts, we make use of *thickness* visual variable which is a *size* visual variable. In the background, the route is rendered thicker; in the foreground, the vehicle path is rendered thinner. In Figure 3.2, it can be seen that route and vehicle path are distinguishable enough. The route is in white and thicker than the blue vehicle path. Non overlapping parts can be easily detected. The distance deviation percent gives information about this deviation type but it will be explained in Circle-Glyph section in detail.

At the beginning of this study, we tried another approach to visualize km deviation. It was giving meanings to road segments. A segment is defined as a pair of two consecutive fixed stops (POIs) on a route. But then we saw that it may not be applicable for some routes. The approach was coloring route segments in accordance with the degree of km deviation, i.e. if the vehicle is deviating gradually at a particular segment, then this segment should be emphasized gradually. To realize this, applying gradients of a color tone to these segments has been found reasonable. Tones of yellow may be used, just to show that at darker yellow parts, the vehicle is increasingly deviating more and more and lighter tones indicate the opposite. This method could work in most cases, but in a case where the vehicle uses same section of a road on a round trip destination, this scenario would mix colourings since one road part is drawn upon another. Thus, this approach is abandoned.

In visualizing time deviation, color assignment to POIs is as explained before. Additionally, since obtaining time deviation is the aim, we assigned time deviation minutes for these points. To do this, an arc is generated around the point having time deviation and this arc is set visible at maximum level of detail; i.e., zooming as much as the interface allows. Figure 3.4 is an example of yellow points with arcs. The *Clock analogy* is utilized here where minutes are read from the yellow arc around points just as in analog clocks. In Figure 3.4, from the upper to lower one, you can observe a few (3~), 5~ and 1 hour 20~ minutes time deviations, respectively. Each black circle around the time arc, as in the third one, indicates 1 more hour in addition to minute values. If there are two black circles, then there are 2 hours more, in addition to minute arc value, etc.

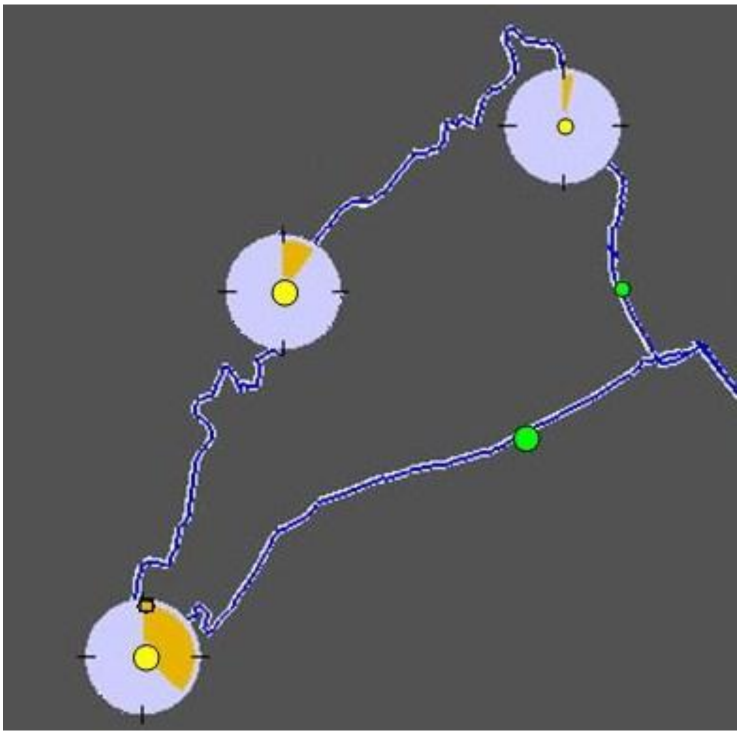

Figure 3.4 Representations of time arcs around deviated (yellow) POIs

3.3.2. Deviation Visualization for Multiple Routes

In order to visualize deviations on a single route, we find it enough to show minute arcs around POIs, and for road parts where vehicle does not follow its scheduled route, to draw the route path thicker than its vehicle. But what is the case if we have multiple routes and want to see the deviations for each of them. Such a scenario is visualized in Figure 3.5: 10 randomly created routes and their related vehicle paths and POIs. The result generates a visual clutter. Pink points are order points, they are only drawn where they are located, and namely, they are single points, not clustered. It is almost impossible to make deductions about movement of vehicles under this scenario.

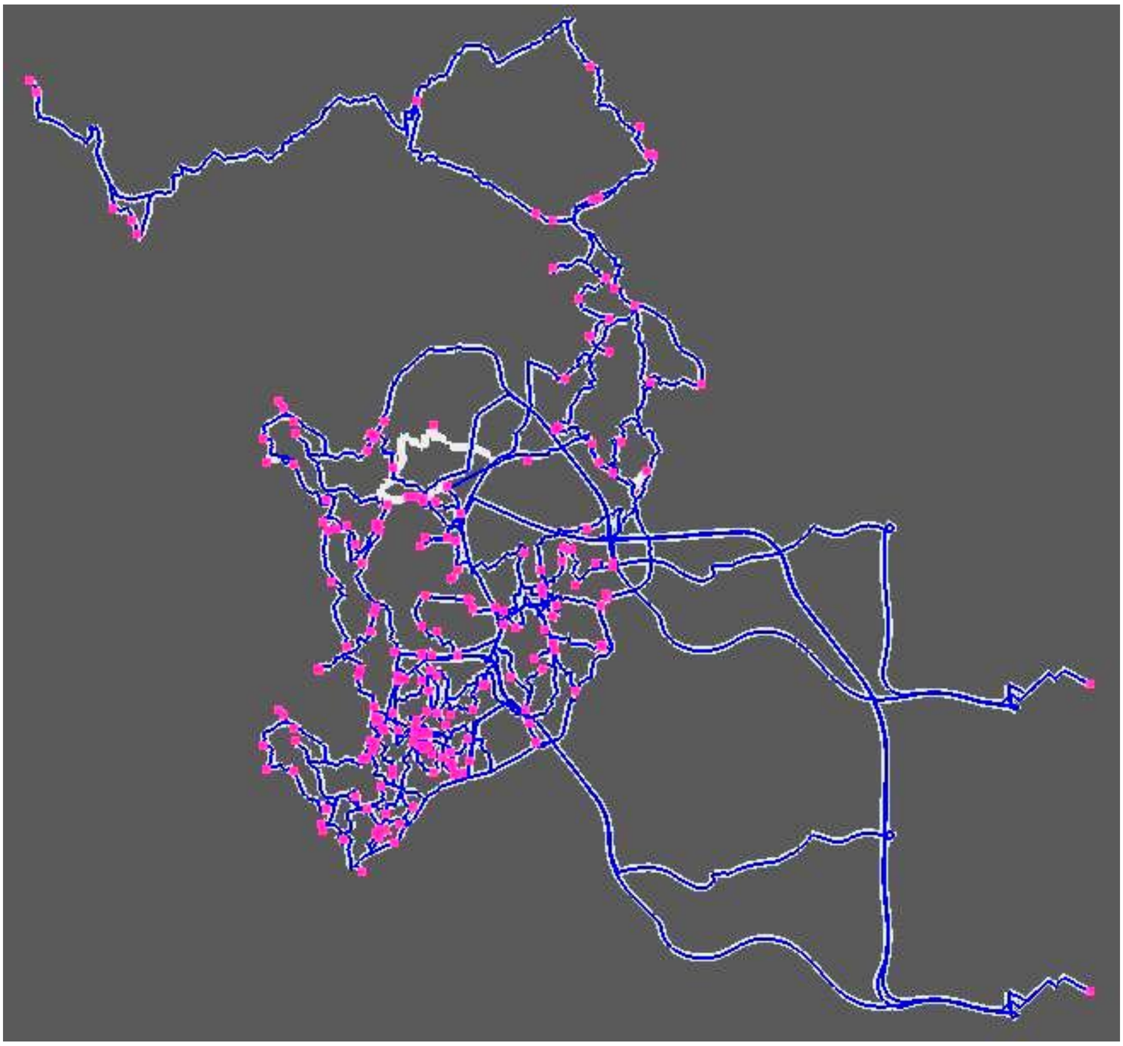

Figure 3.5 Representation of 10 routes (white) and their related vehicle paths (blue) and POIs (pink)

There are multiple routes to handle and all deviations should be addressed. So, we propose two different methods: circle-glyph representation and route coloring.

3.3.2.1. Circle-Glyph Representation

In order to handle multiple route deviations, the first method proposed is circle-glyph representation. Each route has one circle-glyph to reveal its deviation. Circle-glyphs are used to show the two kinds of deviations in its structure.

There are 3 types of information kept in a glyph. An example of circle-glyph is given in Figure 3.6.

- **Textual part:** rendered in the middle and shows how many green, red, and yellow POIs exist in this route. The format is “green:red:yellow”. In the example figure, the circle-glyph has 13 green points and 4 yellow points.
- **Time deviation arc:** designed in accordance with clock analogy just as are POIs at maximum LOD in showing their time deviation. The arc is rendered in yellow and clockwise if the vehicle arrives later than expected; or green and counter clockwise if the vehicle arrives earlier than expected. Black circles around the arc also show these many hours as deviation in addition to arc minute. The base circle is divided into 12 parts by small lines as in an analog clock solely to ease readability. In the example figure, 2 hours and 8-9 minutes time deviation can be approximately observed.
- **Km deviation bar:** (at the bottom) shows a percent value related with the total distance of a route and the distance value that its vehicle travels. The bar is divided into four equal parts to ease readability. If the bar is yellow, then the vehicle has travelled more than it should; if green, then less than it should. In the figure, one can observe 30%-40% km deviation, approximately.

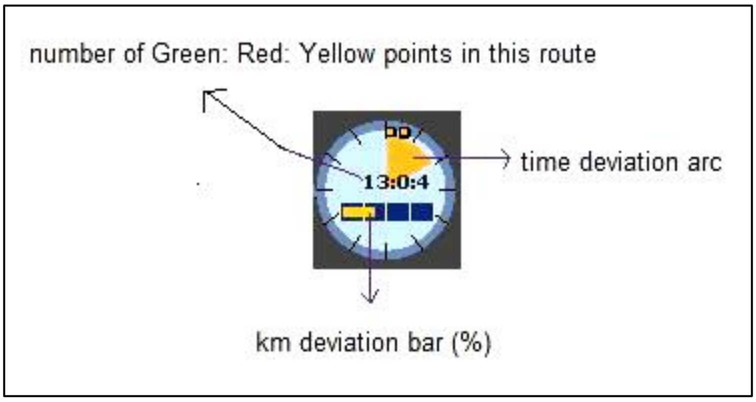

Figure 3.6 Representation of circle-glyph

There are three kinds of circle glyphs in which only time deviation arc is calculated differently, according to: simple, individual, and average time deviations (explained in Event Description part). One can select any from the interface.

After glyphs are generated according to each routes' deviation information, the next issue is how to place them on the screen. We tried two approaches here and tested them in a usability test to determine which is more perceptible. The first approach in the placement of circle-glyphs on the screen was placing each glyph to the location, which is found by averaging all points forming its route. We call it *center placement*. This very simple approach aims to place the glyph somewhere in the center of its route. Examples of this placement are given in Figure 3.7. As seen from the figure, this placement does not handle overlapping between route and the glyphs. (Instead, the method tries to prevent overlapping among glyphs, explained next). Since glyphs are used to embody routes and their related information, if one wants to learn about routes, just looking at its glyphs will be sufficient.

Figure 3.7 Center Placements of Circle-Glyphs

When attempting to put glyphs at these average coordinate locations, they may overlap with other glyphs. To get rid of this problem, glyphs are shifted. The algorithm works like this: the first glyph is placed at its center location, while placing n^{th} glyph, a control is performed from the first glyph to $(n-1)^{th}$. If the current glyph overlaps with these previous glyphs, than the current glyph is shifted right (the right was merely our choice, left can also be used or up or down) until the overlapping disappears. At the end of the algorithm, all overlaps among circle-glyphs are eliminated. This was the first approach in placing glyphs. The result of the placement for our input data which has 20 routes is given in Figure 3.8.

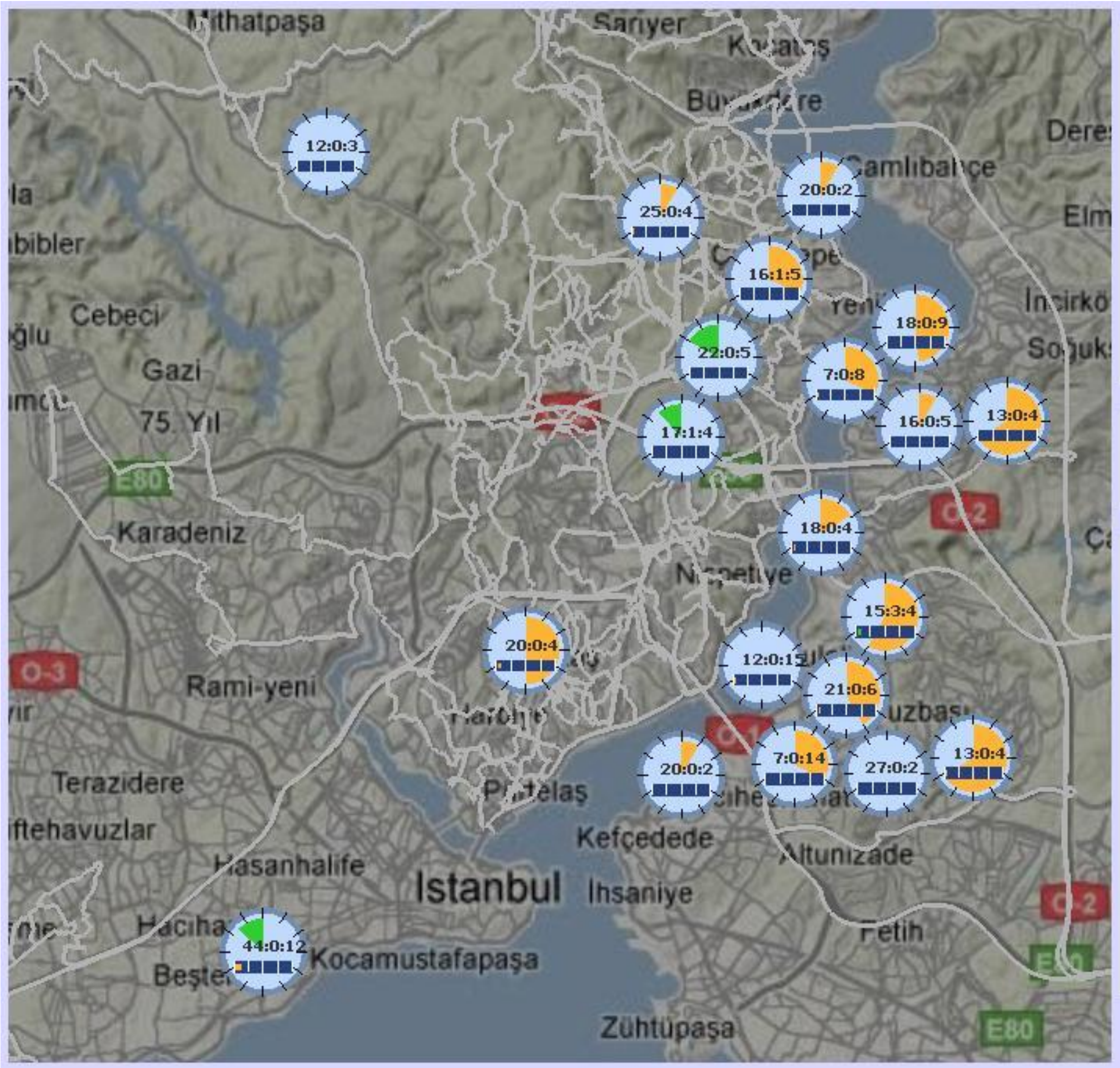

Figure 3.8 Circle-glyphs for 20 routes under the Center placement

The second method in placing circle-glyphs is positioning them on the screen boundary. The study in [27] has been an inspiration to us in this placement. All glyphs are placed along the screen boundary. They are just placed one after another from the first one to 20^{th}. Resulting scene of this placement is given in Figure 3.9.

In a usability study, we test these placements' effects in human perception.

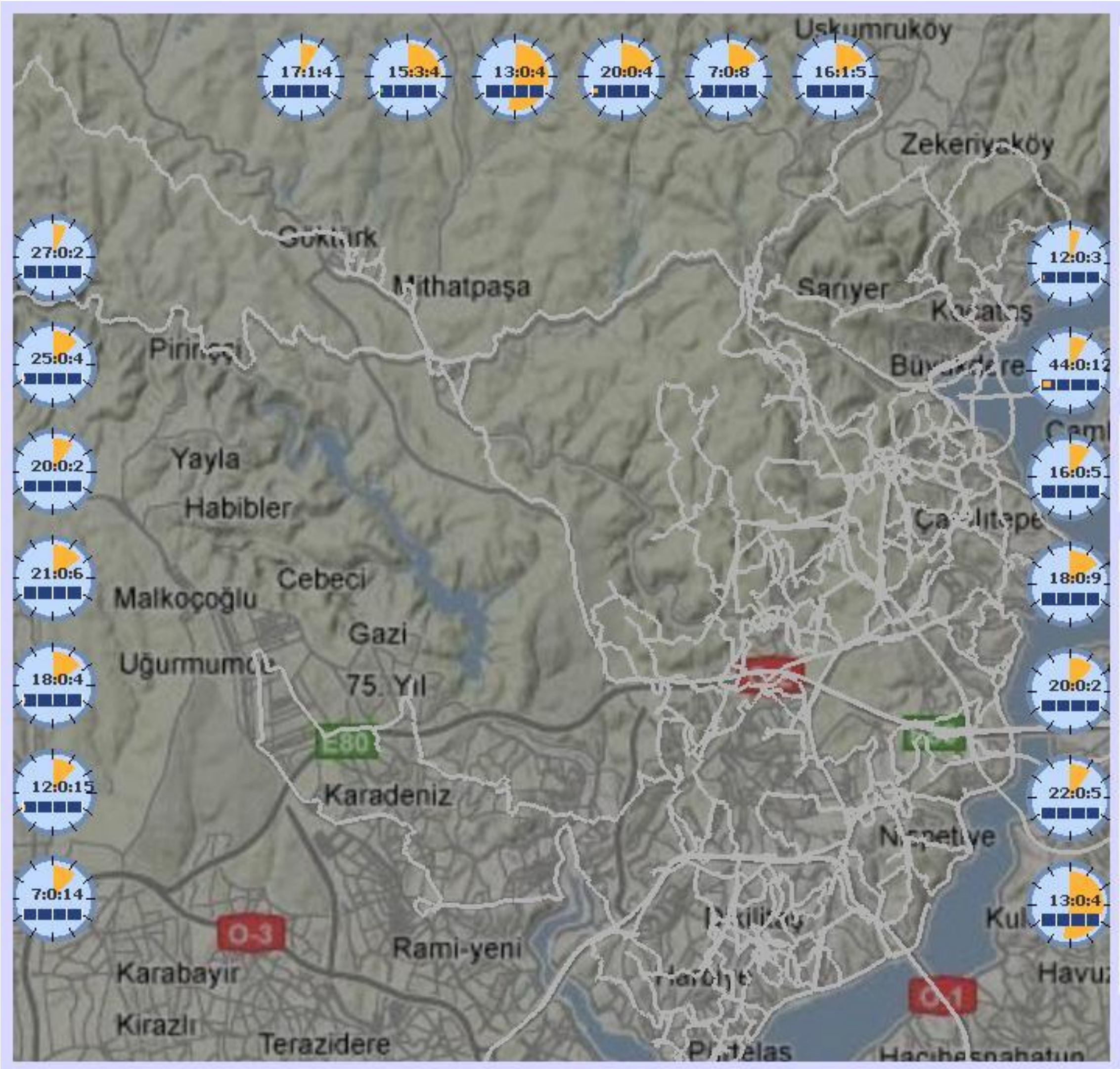

Figure 3.9 Circle-glyphs for 20 routes under the Border placement

In both placement algorithms, user can select a route number from the relevant spinbox (including numbers from 0 to N, where N is the number of routes) on the interface to see the information of the selected route: the route path, the vehicle path, the POIs and also the circle-glyph. So, for each individual route, one can understand which glyph belongs to which route easily. Additionally, as previously mentioned, routes can be displayed either in same color (white) or in multiple colors. If routes are in multiple colors mode, each route and the surrounding circle of its glyph are assigned the same color. This provides a visual cue for the users to understand which glyph corresponds to which route, even there are multiple routes. But, the routes are drawn on top of each other if they share same road part. Due to these overlappings, some routes can be invisible. In these cases, this approach (same color assigning) may not be adequate; and the user may not find the corresponding route of a glyph. Single route example of this scenario is given in Figure 3.10.

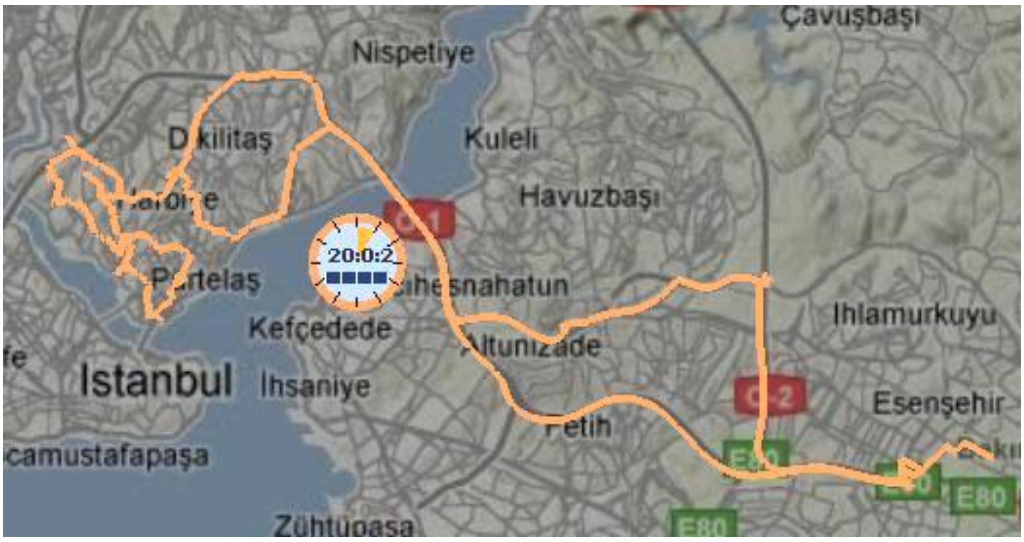

Figure 3.10 The route and the surrounding circle of its glyph are rendered in same color, when routes are in multiple color mode

3.3.2.2. Route Coloring

The route coloring approach aims to highlight routes set as prominent according to a criterion. This criterion can be any of the four below:

- Total time deviation in hours and minutes
- Total distance deviation in kilometers
- Number of Deviated Points (yellow POIs)
- Number of Not Arrived Points (red POIs)

Under this criterion, the most prominent route is colored in red; the least prominent, white; and in between routes, in tones between red and white according to the degree of the criterion. In Figure 3.10, route coloring is performed according to total time deviation. The maximum red route is the most prominent one which means that total time deviation is maximum here. According to the degree of time deviation, other routes are colored. Following tones of red, pink tones can be seen. Finally, when time deviation is very low, routes are colored in white.

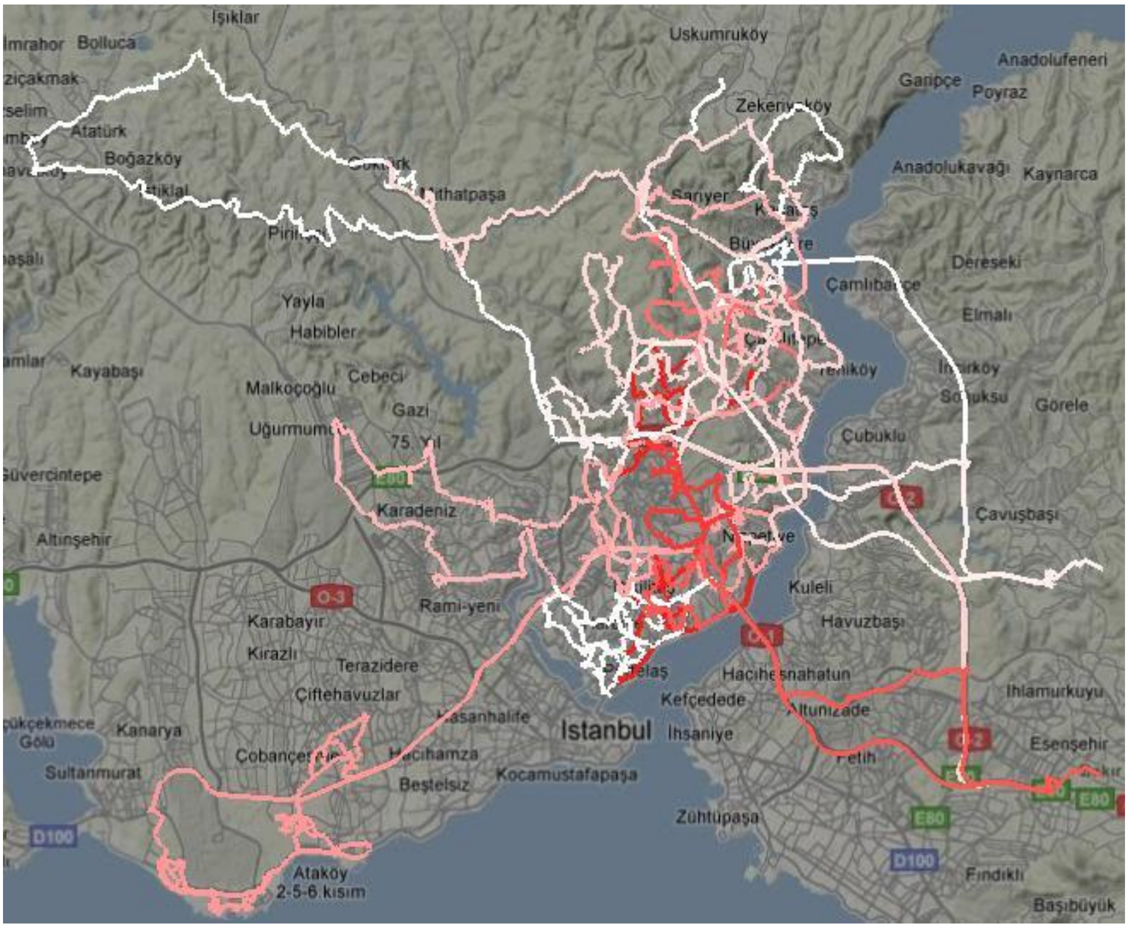

Figure 3.11 Route Coloring according to time deviation

Although colored routes are seen all together, from the interface, single routes can be selected from the maximum deviated one to the less deviated. Another example is in Figure 3.11 showing route coloring according to km deviation.

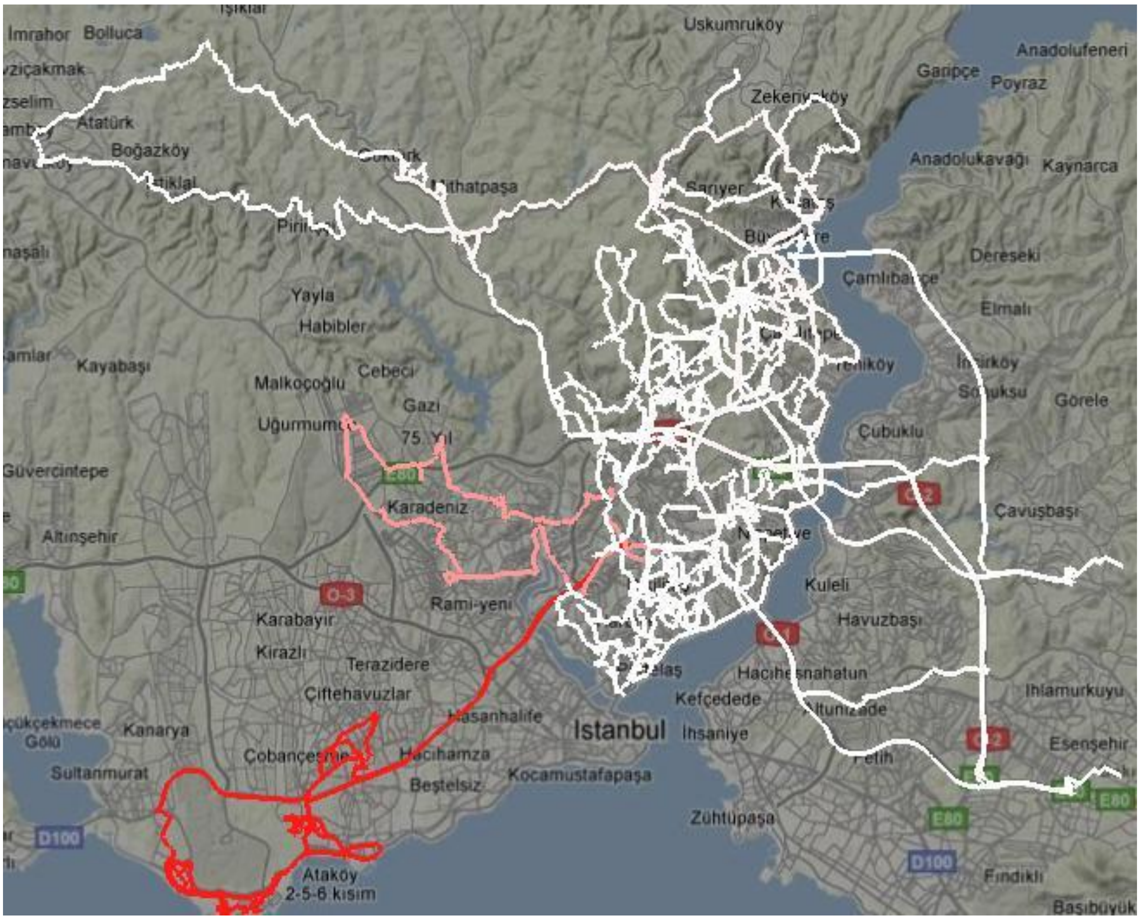

Figure 3.12 Route Coloring according to km deviation

Single routes denoted from maximum important to less important are given in Figure 3.12 based on route coloring according to km deviation scene in Figure 3.11. So, the top route has maximum km deviation; the middle one, second maximum km deviation; and the bottom, third maximum km deviation. Since this third one is white, it is expected that the remaining routes are white, too. Because, route coloring denotes all routes from maximum to minimum in terms of the criterion chosen.

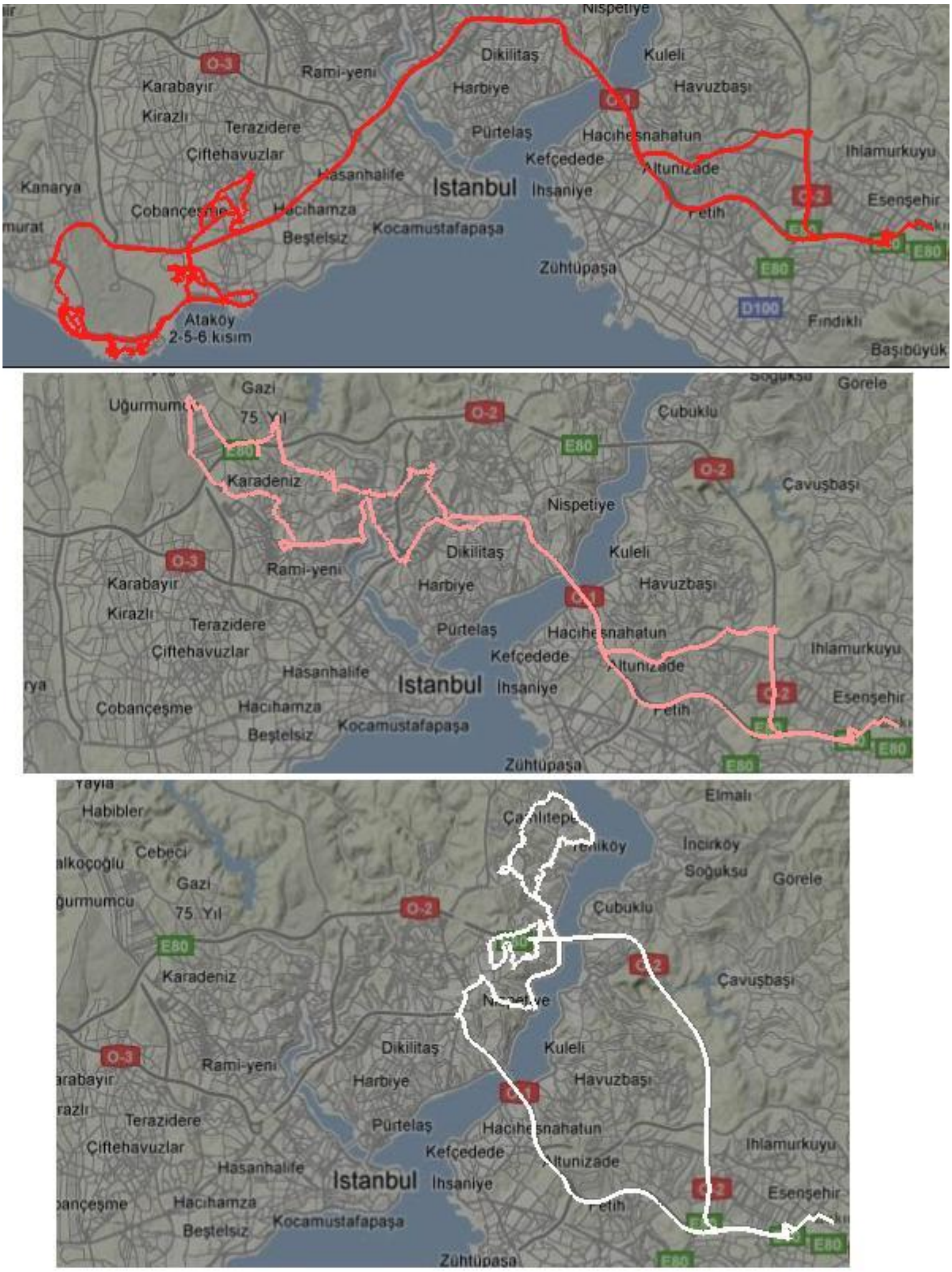

Figure 3.13 Route Coloring according to km deviation: Top 3 routes are shown in terms of having maximum km deviation

4. CASE STUDY

The experiment was conducted with 14 people, where 5 of them were female. The participants were college students with their age ranging from 20 to 28. All participants had either normal or corrected-to-normal vision. No participant was color-blind. Only fundamental computer using skills were required.

The experimental apparatus was a Toshiba Satellite A200-1BP notebook computer with a Core 2 Due 2.0 GHz processor and 2 GBs of memory, an additional attached USB mouse. The computer had an NVIDIA GeForce 7300 video card with 256 MB of video memory and with a resolution of 1280x800 pixels.

4.1. Experiment Design

The dataset comprises 20 routes (explained in Input Data section) and their associated vehicle path and POIs. The tasks were designed to comprehend multiple route deviations in such datasets clearly.

The experiment has 3 sessions;

- **a warm-up session:** the objective of the experiment and the nature of the visualization system are explained to the participants,
- **a demonstration session:** the participant is shown some example tasks and in what way the tasks can be completed,
- **evaluation session:** the tasks are given and response times of each participant are recorded.

The evaluation tasks are designed in two groups: Reading and Interaction. Reading tasks aim to test two methods on circle-glyph placements by using the system's readability capability. Since there is a lot of information displayed in the system, we also wanted to see if they are shown in a readable way and the system is serving our purpose. On the other hand, Interaction tasks aim to test the two methods on visualization of multiple route deviations by using the system's interaction capability. We also wanted to see whether the user is able to interact with the system in a proper way.

In the experiment, each participant is asked to answer some questions from the table given in Table 4.1. In Reading tasks session, the first 7 of 14 participants answered the question in T1, and the second half answered the question in T2. Likewise, in the Interaction tasks session, the first 7 of 14 participants answered the question in T3, and the second half answered the question in T4. The division of participants is needed to prevent familiarity bias in the two groups: between T1 and T2; and between T3 and T4. A common problem encountered in similar usability studies is that users learn the dataset in the first tasks and complete the subsequent tasks much faster by remembering the previous tasks. By this remembering effect, the time spent to complete a task does not accurately reflect the visualization method's actual contribution. To eliminate this source of error, if a participant is asked the question in T1, the response of the same participant to T2 is not used as this will cause familiarity bias. Same is valid for the Interaction tasks between T3 and T4.

Response times of each participant are recorded. The participants are forced to answer questions correctly after they are told to reconsider their answers if a wrong answer is given. If a wrong answer is given, then the participant spends more time trying to find the correct answer, which results in an increased task completion time. Accuracy was not included as an extra measure because of two reasons: first, the aim of the experiment was to find the time it takes to reach the correct answer. Secondly, most of the participants gave the correct answer in their first attempts. Thus accuracy was not counted as an informative measure.

Task Type	Task No	Task Definition
Reading	T1	Center Placement: How many routes have time deviation (latency)?
Reading	T2	Border Placement: How many vehicles travel less than they are expected (km deviation)?
Interaction	T3	Use Circle-glyphs: Find the route with max km deviation What is the number of yellow points in this route?
Interaction	T4	Use Route Coloring: Find the route with max km deviation What is the number of POIs in this route having time arc around?

Table 4.1 Tasks used in the evaluation of our methods

- **Reading tasks**:

In reading tasks, T1 and T2, the participants are given two static images of the scene and are asked to answer two questions related with these images. In these questions, we tried to test glyph placement strategies' effects in perception which are center and border placement. Although the questions seem different, both of them include a similar counting operation on glyphs. The difference only results from the part of the glyph at which the participant is looking while answering the question. In T1, the participant counts time related values on the *time deviation arc*; and in T2, distance related values on the *km deviation bar*.

In the first reading task, T1, circle-glyphs are placed according to the center placement, which places each circle-glyph on the center of its route (which was explained in Circle-Glyph Representation section in detail). The test image we used in this task is given in Figure 4.1. In this task, we asked the participants to find the number of routes having time deviation (latency) in this image. So that the participant should have to look at all glyphs' *time deviation arc* parts and count the ones having a yellow time deviation arc.

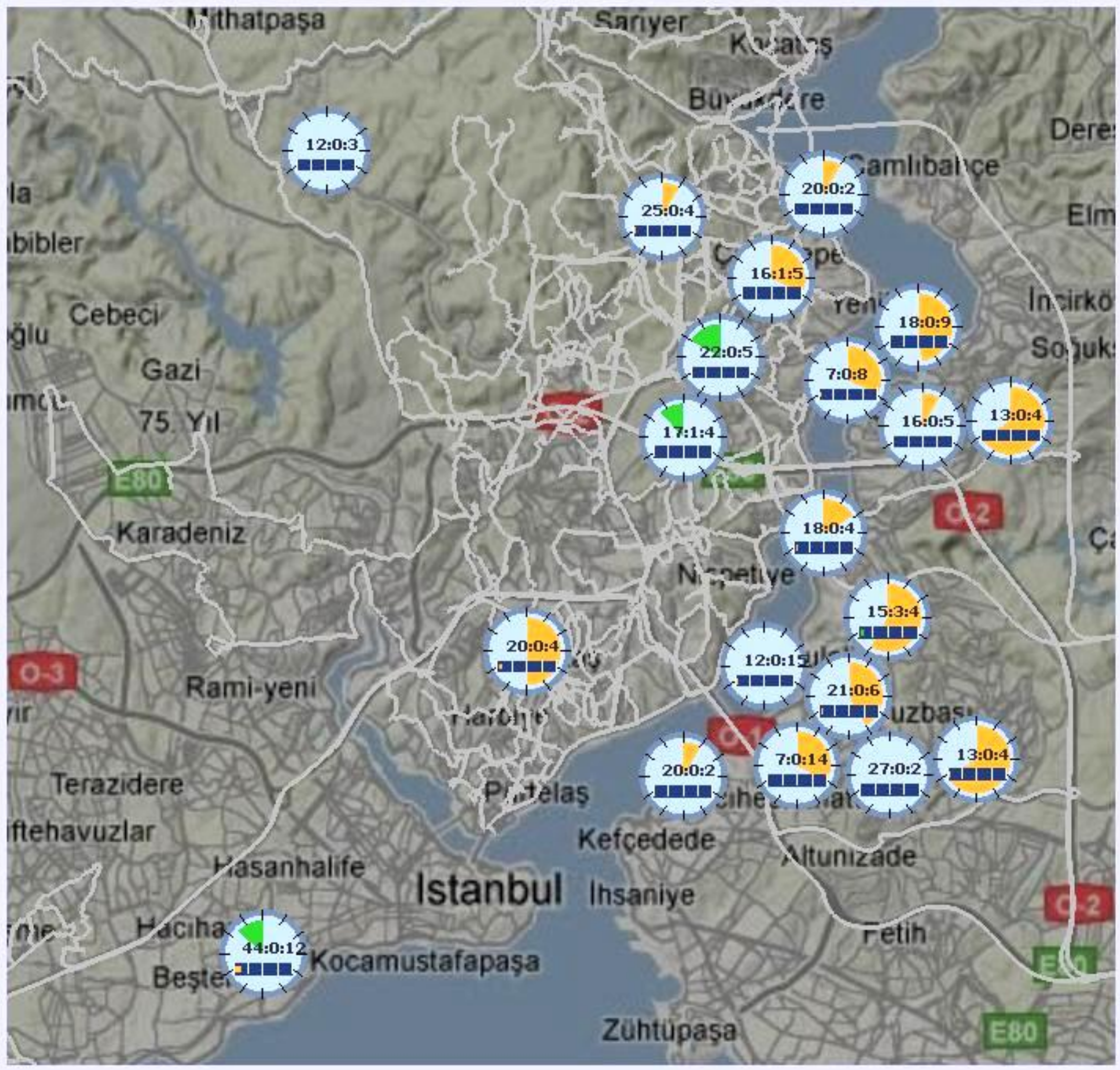

Figure 4.1 First reading task, T1, circle-glyphs are in center placement mode

In second reading task, T2, circle-glyphs are placed according to the border placement which one by one places each circle-glyph on the screen boundary (which was explained in Circle-Glyph Representation section in detail). The test image used in this task is given in Figure 4.2. In this task, we asked the participants to find the number of vehicles travelling less than their scheduled route path. So that the participant should have to look at all glyphs' *distance deviation bar* parts and count the ones with green segments in it.

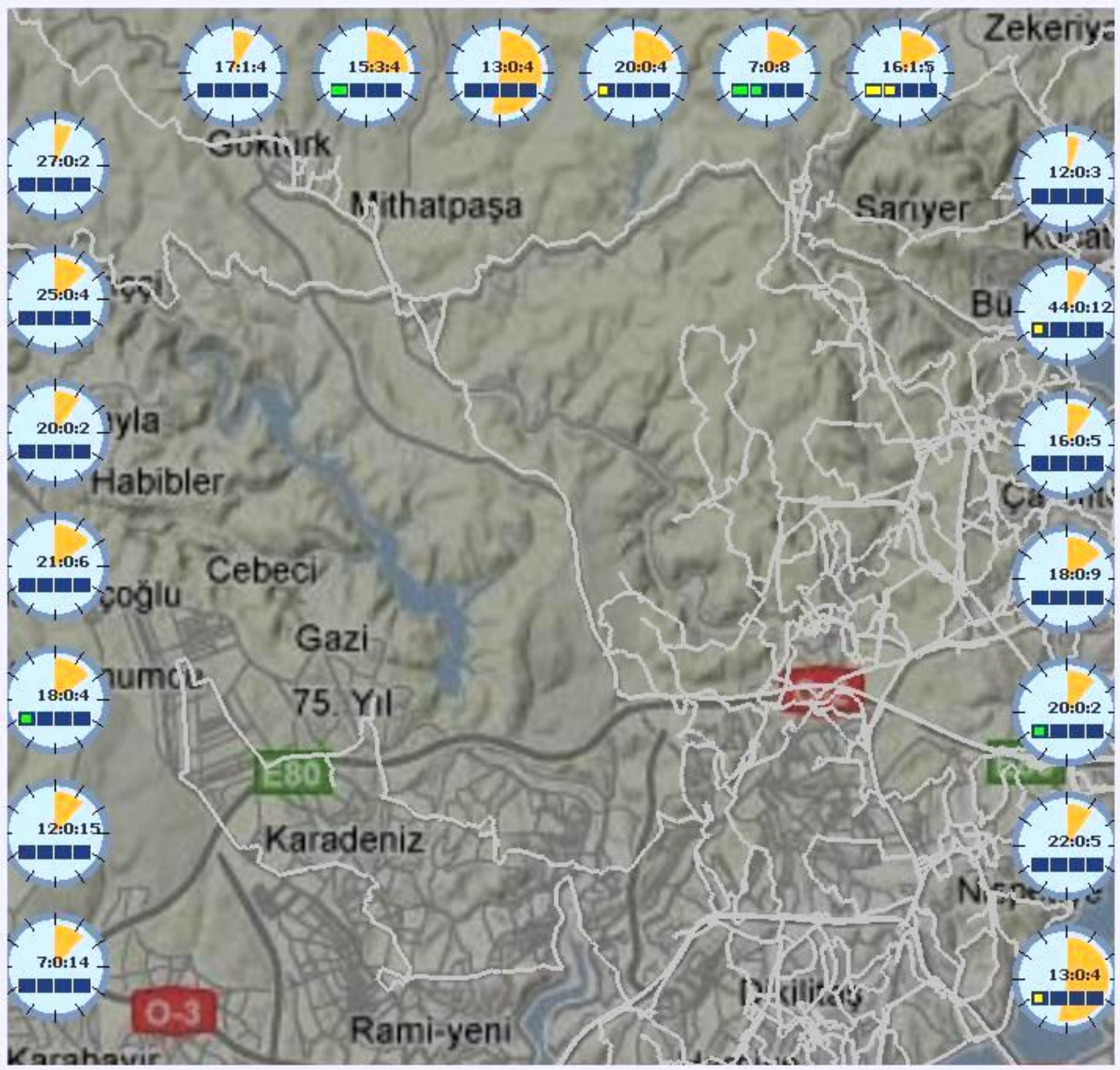

Figure 4.2 Second reading task, T2, circle-glyphs are in border placement mode

Both tasks include a glyph search and counting task, but the first task is performed when circle-glyphs are in center placement, and the second task is performed when circle-glyphs are in border placement modes.

Our hypothesis for the reading tasks, T1 and T2, is that border placement which places the circle-glyphs on the screen boundary helps easier and quicker understanding of deviations than center placement. Because in the border placement mode, glyphs are around the screen, and routes are in the center, and this image decreases the complexity of the scene. But in center placement, both glyphs and routes are drawn together in the center of the screen; lots of overlaps and intersections occur which result in a more complex scene. Also, in border placement, because the glyphs are lined up along the border of the scene, searching over them should be much easier compared to searching in the center placement mode.

- **Interaction tasks**:

In interaction tasks, T3 and T4, the participants are allowed to interact with the user interface. In interaction questions, we tried to test which multiple deviations handling method is more comprehensible and obtains quicker response from users. To test the usability of these methods: circle-glyphs or route coloring; at first, we asked the participants to the find the route with maximum distance deviation by using these methods. At this point, the user should search and find this route. The first 7 of 14 participants find it by using circle-glyphs, and the remaining half by route coloring.

In first interaction task, T3, when the participant finds the route with maximum distance deviation by using circle-glyphs, s/he should come up with a screen as shown in Figure 4.3 (The scene is partially shown in this figure). The user should have identified the route with maximum distance deviation as the bottom glyph in the figure, since the segment in *distance deviation bar* is the biggest among all other glyphs. Hence, the user should have up to now made a search over all glyphs. In the second part of the question, we asked for a readable value for this route: "the number of yellow points in this route". Then from the textual part in the glyph, s/he should read it as 12. In the figure, the textual part in the bottom glyph is 44:0:12, which shows the number of Green:Red:Yellow points, respectively.

Figure 4.3 First interaction task, T3, scene from circle-glyphs usage task

In second interaction task, T4, when the participant finds the route with maximum distance deviation by using route coloring, s/he should come up with a similar screen in Figure 4.4. At the beginning of this task, we set the visibility of POIs to true, since the second part of the question is asking a readable value related with POIs. So, setting the visibility of POIs will not be an extra work for the participants. The scene is again partially shown in this figure. In the second part of the question, we asked again a readable value for this route: “the number of POIs having time deviation arc around”. Then by just looking at POIs, s/he should say it as 6 as it is shown in the figure.

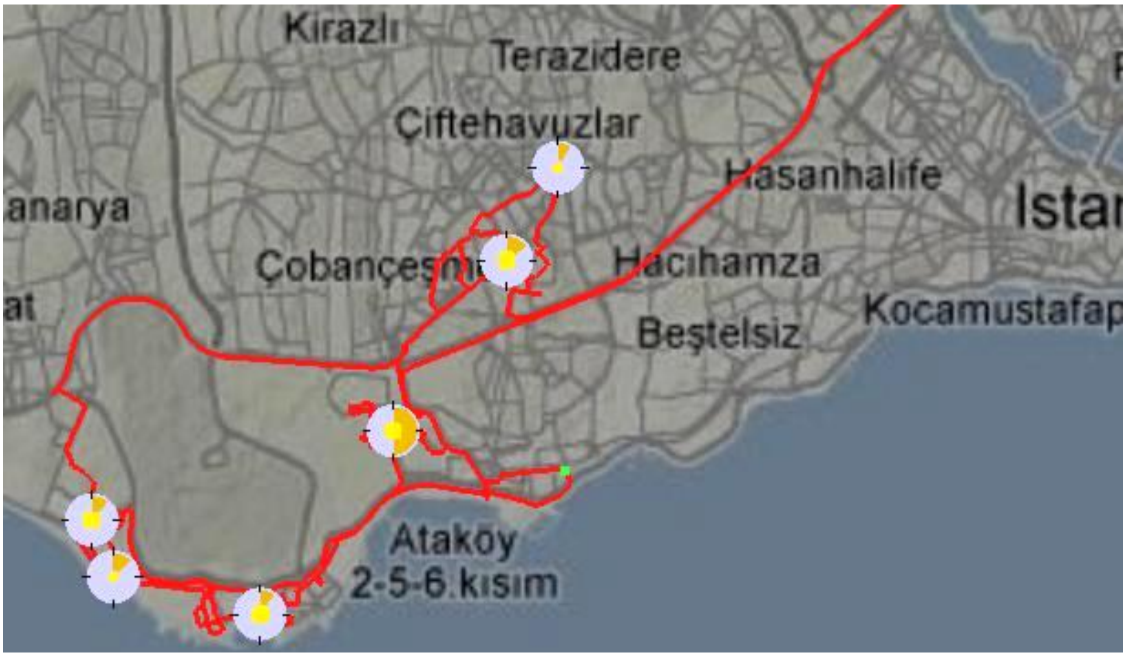

Figure 4.4 Second interaction task, T4, scene from route coloring task

Our hypothesis in these interaction tasks, T3 and T4, is that by using route coloring method, users should give quicker response compared to circle-glyphs. Because route coloring sorts the routes according to the criterion chosen (criterion is *km deviation* in T4) and this should be done in an easier way in route coloring.

4.2. Test Results

Evaluation tasks (T1, T2, T3, and T4) are performed on 14 participants overall. Response times to the tasks are measured. The results of this experiment (means and standard deviations) are given in Table 4.2.

	T1	T2	T3	T4
Mean	15.41	13.25	35	24.54
Std Dev	4.50	3.45	7.04	5.37

Table 4.2 The mean completion time (in seconds) and the standard deviation of the mean values for each task.

In order to compare our previously mentioned methods, we examined statistical significance of the difference between the task completion times in the Reading tasks group, between T1 and T2 and the Interaction tasks group, between T3 and T4. Our goal was to determine whether placing circle-glyphs on boundaries works better in the Reading tasks group and whether route coloring does better in the Interactions tasks group. To draw such a conclusion, we must detect a significant difference in the means of two sets of values in each group. Hence, we performed analysis of variance (ANOVA) on mean values which is specifically two-sample t-test.

Since every test of significance [28] begins with a *null hypothesis* followed by an *alternative hypothesis***,** first we formed them for our tests. Null hypothesis forms a basis for the argument, yet unproven. Alternative hypothesis is a statement of what a statistical hypothesis test is set up to establish. Hypotheses are always stated in terms of population parameter, such as the mean. So, we make calculations according to the mean values given in Table 4.2. Also, there is the property of alternative hypotheses as being *one-sided or two-sided.* A one-sided hypothesis claims that a parameter, mean in our case, is either larger or smaller than the value given by the null hypothesis. A two-sided hypothesis claims that a parameter is simply not equal to the value given by the null hypothesis, that is, the direction does not matter. Since, we claim that one of the techniques in both reading and interaction groups should give larger means, and the others smaller, our test is one-sided.

- **Test Result on Reading tasks:**

In order to compare the methods used in reading tasks, T1 and T2, we apply a t-test on them. At first, we form null hypothesis and alternative hypothesis for glyph placement strategies: center placement (which is applied in T1) and border placement (which is applied in T2).

Null hypothesis: Border placement is no better than center placement in terms of task completion times, on average.

Alternative hypothesis: Border placement is better than center placement, on average (this is the hypothesis that we claimed at the beginning).

Our aim in this test is to see whether we can reject the null hypothesis and conclude that the alternative hypothesis is true. Calculations are done [29] and t-value is calculated as 1.47. After finding t-value, we examined the result of this statistical analysis at different significance levels from t-distribution table [30] [31]. The outcomes revealed that the difference between the mean task completion times were statistically significant at the 10% significance level for the Reading tasks group. This finding means that we can reject null hypothesis only under this significance level. And this value is not statistically small enough for us to say that the border placement performs better than the center placement as much as we considered.

- **Test Result on Interaction tasks:**

In order to compare the methods used in interaction tasks, T3 and T4, we apply a t-test between them. At first, we form null hypothesis and alternative hypothesis for visualization of multiple route deviations methods: circle-glyph usage applied in T3 and route coloring applied in T4.

Null hypothesis: Route coloring is no better than circle-glyphs, on average.
Alternative hypothesis: Route coloring is better than circle-glyphs, on average. This is the hypothesis that we claimed at the beginning.

Our aim in this test is to see whether we can reject the null hypothesis. Calculations are done [29] and t-value is calculated as 4.57. Having calculated the t-value, we examined the result of this statistical analysis at different significance levels from t-distribution table [30]. The outcomes revealed that the difference between the mean task completion times were statistically significant at the 0.05% significance level for these two methods. This means that we can reject null hypothesis under this significance level. Since this value can be statistically counted as small, we can say that the route coloring method created an experimentally verified simplicity when compared to circle-glyphs, as we claimed in our hypothesis. This made statistically significant reductions of task completion times by using route coloring method. Because in order to complete a route finding task by the using circle-glyph method, user needs to perform low-level tasks such as searching all glyphs and comparing them with others to obtain the desired route. These low-level tasks are eliminated in route coloring method. This facility is reflected on the statistically significant reduction of task completion times.

An important remark about Interaction tasks should be made here. Since route coloring does not give details about route deviation such as deviation time, km etc. as in circle-glyphs, one can consider that the two methods are not doing the same job and cannot be considered as comparable tasks. But related tasks, in T3 and T4, are formed intended for finding problematic routes under the criterion asked. In the route coloring task (T4), the participant is not asked about the details that a circle-glyph includes in its structure. Both tasks, T3 and T4, require a route searching and a simple reading part about the route found and we compared the methods' effects in route searching part. Box plots [32] for the two test groups are given in Figure 4.5 and Figure 4.6.

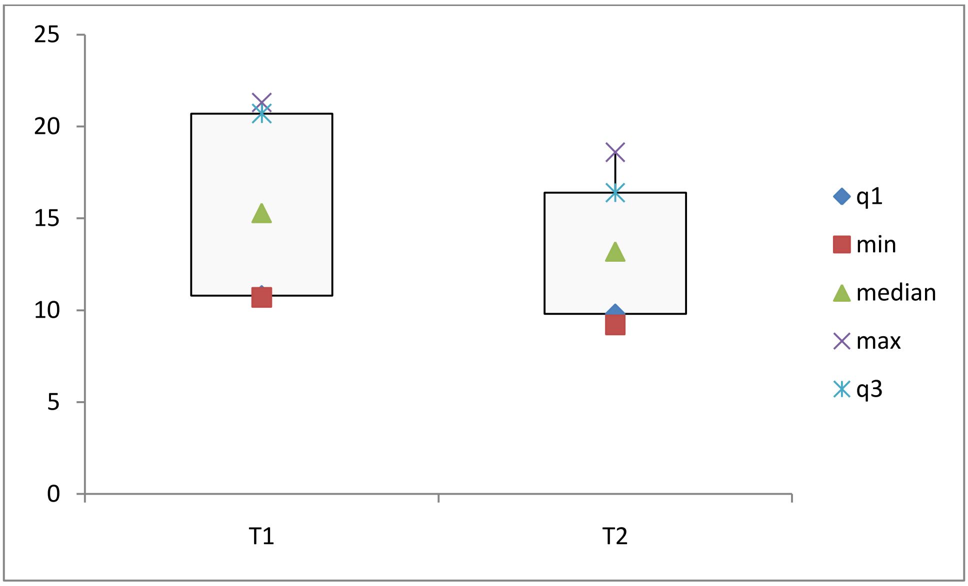

Figure 4.5 The boxplot of the completion times of Reading tasks: T1 & T2

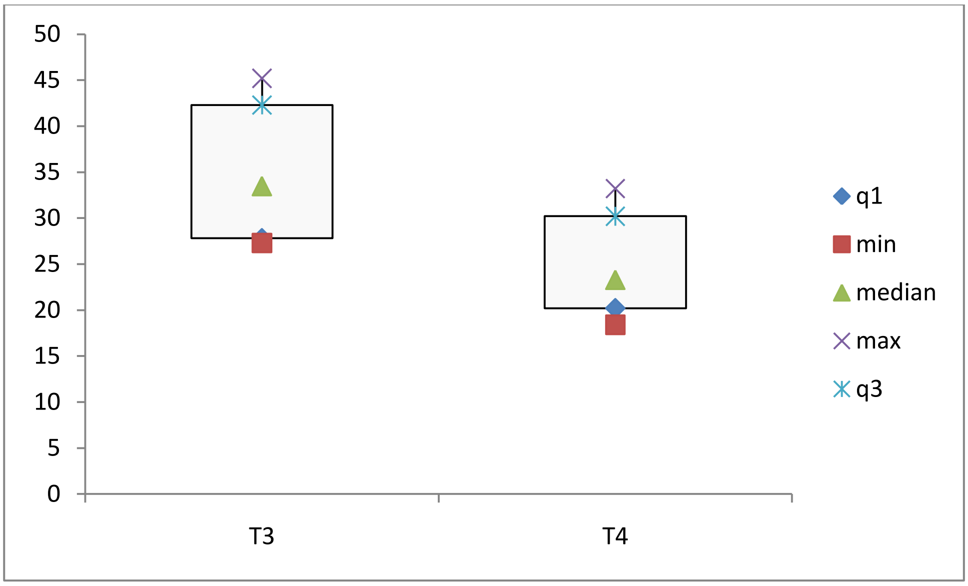

Figure 4.6 The boxplot of the completion times of Interaction tasks: T3 & T4

5. RESULTS and DISCUSSIONS

5.1. Results

With the recent rise of mobile devices equipped with GPS receivers, movement data can be collected in excessive amounts. The movement of a mobile object needs to be interpreted since valuable information is kept in these datasets. By just looking at the numbers related with locations, it will not be possible to comprehend the data and make deductions. In order to retrieve valuable information from movement data, multifarious techniques are developed in geovisualization frame.

In this thesis, we analyze some movement data we have created and propose several methods in visualizing deviations on routes. As far as we know, visualization of route deviations is not yet defined nor studied in the literature. Many researchers focus on trajectory clustering techniques and discovering patterns from movement data. But this study looks at the issue from a different perspective and aims to form a basis in the deviation visualization frame. The problem of deviation is composed of two parts, spatial and temporal. In the deviation scenario, there is one vehicle which travels on a route assigned to it. The vehicle should exactly follow its route and make deliveries to some pre-defined order points on time. If the vehicle fails either one of these two tasks, it is said to have a *deviation*. If the vehicle does not follow its route, this is termed *distance deviation*, which is the spatial part of the deviation problem. If the deliveries to order points are not made on time, this case is termed *time deviation*, which is the temporal part of the deviation problem. Additionally, we have multiple routes and their vehicles and POIs in this scenario. After deriving these concepts for the deviation problem, we created several methods to visualize this scenario in a clear way. Methods we developed in this thesis aim to respond to multiple routes case, as well. We tried to address this problem step by step. First, we considered how to visualize deviations on a single route. Then we dealt with visualizing deviations on multiple routes.

In the proposed system, level of detail is used in the visualization of order points. Because due to the close locations of some order points, overlaps and intersections between them may occur. Since we give information using both color of points and time deviation arc around them, we wanted to present the information given on points in a legible way. According to the degree of the closeness between the viewpoint and the points, clustering is applied to these points to decrease visual complexity.

In visualization of multiple routes deviations, we propose two methods: circle-glyphs and route coloring. In the circle-glyphs method, the first consideration was designing a glyph to cover all aspects of deviations. In designing circle-glyphs, we considered what to contain in the glyph structure about deviations and designed according to this consideration. Another issue was placing the glyphs on the screen and here we proposed two methods, being center and border placement. We tested their effects in human perception in a usability study. In route coloring, we provide users to sort the routes according to the level of deviation so that the users are able to obtain the most deviated routes on the fly. The criterion of deviation is set as an optional value so that users are able to sort the routes according to time deviation, distance deviation etc. To test the usability of circle-glyphs representation versus route coloring, we formed evaluation tasks related with them. Discussions of our usability test are presented in the next section.

5.2. Discussions and Further Study

There are some issues in this system which need further study. Our dataset includes 20 routes, we also tried the system with 30 routes which was quite successful as well. But what if we have hundreds of routes, and want to display deviations by circle-glyphs? Glyph placement strategies may not be adequate here, because the screen is a limited place and representing hundreds of circle-glyphs at once will not be a solution since the glyphs will not fit in the screen. Thereby, to make the system more flexible according to data size, specifically, number of routes; different approaches need to be devised. On the other hand, route coloring is more successful at this point because it does not suffer from data size. Since it sorts routes according to the deviation criterion chosen, the most prominent routes are always on the foreground and if users want to obtain the deviated routes, from maximum deviated to minimum deviated ones, they can find them easily.

The causes of deviations vary widely. Traffic congestion, accidents, weather conditions, and operational issues are among the prominent ones. Exploring the reason behind how an actual behaviour deviates from a desired or typical behaviour is another issue which is not handled in this study. Since characterization of vehicle behaviours is needed in this context, it requires different analysis techniques to apply on the movement data. Also, it requires analysis of behaviours of mobile objects in detail; and weekly or monthly data of mobile objects may be needed to understand the vehicle activity under the deviation causes.

The visual environment we propose in this study tries to present information about deviations in a clear way. The aim is to provide users both the overview (by the summary glyphs and route coloring approaches) and the details of the deviations (by zooming to POIs, showing time deviation amount etc). The methods developed provide viewing the problematic routes on the fly. Also, widely used color assignments are done in accordance with the colors in daily life to ease understanding (such as in stoplight context). Instead of visualizing deviations by visual elements, one can consider to put all the numbers related with the deviation details to a table. This naive approach can be informative enough, but it cannot reflect overall view of deviations for multiple routes. Suppose you look at a table filled with 20 routes' deviation details as it is in our input data, (details include coordinate values of points forming route path and vehicle path, the unreached order points, the time deviation amount for all deviated points and all other deviation details) and you want to obtain the most problematic routes. You eventually have to make a search over the table and make the all comparisons from mind. Additionally, the visual elements we deal with in this study consist of routes which can only make sense when drawn on maps. A table cannot explain the route and its vehicle's deviation (distance deviation) visually; instead it can only include coordinate values of points forming route path and vehicle path from which it is impossible to deduce anything. The table can only include the total length of paths and the difference of them in representing distance deviations. Consequently, our system provides the viewers to understand the structure of deviations in a coherent way by visual elements, when the naive approach becomes inadequate in explaining the deviations in most cases.

In the usability study, we tried to test most aspects of the deviation visualization methods. The first test was performed on glyph placement strategies: center placement and border placement. Our hypothesis is that border placement is easier to understand and simpler than center placement. Each method is tested on 7 participants. In order to claim our hypothesis, we have to obtain statistical difference between the results in two groups. To understand the statistical difference, we applied t-test on the outcomes obtained from the participants. The result of the test revealed that the difference between the mean task completion times was statistically significant at the 10% significance level for these two methods. This value is not so small statistically which means that the difference of two methods is not as great as we expected. But it can still be said that border placement is found more successful than center placement with a confidence interval of 90%.

The second test was performed to test the usability of deviation visualization methods for multiple routes: circle-glyph and route coloring. Our hypothesis is that route coloring provides a simpler approach than circle-glyphs in finding most deviated routes. Each method is again tested on 7 participants. To understand the statistical difference between the two groups, we applied t-test on the outcomes obtained from the participants. The result of the test revealed that the difference between the mean task completion times was statistically significant at the 0.05% significance level for these two methods. This value is small enough statistically to say that the difference of two methods is as great as we expected. Hence, it can be said that route coloring method is found more successful compared to circle-glyph with a confidence level of 99.95%. That is, route coloring method created an experimentally verified simplicity in visualization of multiple route deviations, which manifests itself in significant reduction of task completion times.

In this thesis, we tried to cover all aspects of route deviations and reflect them in our visualization approaches. In general, the system is found to be easy to understand and informative enough by the participants' feedbacks and the results of the usability test proved it.

REFERENCES

[1] Colin Ware, “Information Visualization: Perception for Design”, Second Edition, 2004.

[2] F. B. Viégas, M. Wattenberg, F. Ham, J. Kriss, M. McKeon, "Many Eyes: A Site for Visualization at Internet Scale", IEEE Transactions on Visualization and Computer Graphics, vol. 13, No. 6, 2007.

[3] Andrienko, G. and Andrienko, N. “Extracting Patterns of Individual Movement Behaviour from a Massive Collection of Tracked Positions” ,
Workshop on Behaviour Modelling and Interpretation , Technical Report 42, TZI Technologie-Zentrum Informatik, Universitaet Bremen, pp.1-16, 2007

[4] Natalia Andrienko, Gennady Andrienko, Monica Wachowicz, Daniel Orellana, “Uncovering Interactions between Moving Objects”, GIScience, 5th international conference, Proceedings, pp.16-26, 2008

[5] Gennady Andrienko, Natalia Andrienko, “Spatio-temporal aggregation for visual analysis of movements” IEEE Visual Analytics Science and Technology (VAST 2008) Proceedings, IEEE Computer Society Press, pp.51-58, 2008

[6] Wall painting in Turkey, Available from:
http://en.wikipedia.org/wiki/Wall_painting_in_Turkey, retrieved 20.06.2010

[7] Juan C. Dürsteler, “The History of Visualisation”, The digital magazine of InfoVis.net, 2002

[8] Stuart Card, Information visualization, in A. Sears and J.A. Jacko (eds.), The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Lawrence Erlbaum Assoc Inc, 2007.

[9] Michael Friendly, “Milestones in the history of thematic cartography, statistical graphics, and data visualization”, August 24, 2009

[10] Information Visualization, Available from:
http://www.infovis-wiki.net/index.php?title=Information_Visualization, retrieved 20.05.2010

[11] Daniel A. Keim, Christian Panse, Mike Sips, "Information Visualization: Scope, Techniques and Opportunities for Geovisualization", 2003

[12] Diane M. Szaflarski, "How We See: The First Steps of Human Vision", Access Excellence Classic Collection:
http://www.accessexcellence.org/AE/AEC/CC/vision_background.php, retrieved June, 2010.

[13] Visual Mapping, Available from:
http://www.infovis-wiki.net/index.php?title=Visual_Mapping, retrieved 20.05.2010

[14] J. Bertin, "Semiology of Graphics: Diagrams, Networks, Maps", University of Wisconsin Press, 1983.

[15] J. Mackinlay, "Automating the design of graphical presentations of relational information", ACM Transactions on Graphics, vol. 5, no. 2, pp. 110-141, 1986.

[16] J. Dykes, A.M. MacEachren, M.-J. Kraak, "Exploring Geovisualization", International Cartographic Association, pp 2005

[17] Jiang, B., Huang, B., and Vasek, V., "Geovisualisation for Planning Support Systems", In Planning Support Systems in Practice, Geertman, S., and Stillwell, J. (Eds.), pp. 177-191, Springer, 2003

[18] Maneesh Agrawala, Chris Stolte, "Rendering Effective Route Maps: Improving Usability Through Generalization", Proceedings of ACM SIGGRAPH 2001, 2001

[19] Floraine Grabler, Maneesh Agrawala, Robert W. Sumner, Mark Pauly, "Automatic Generation of Tourist Maps**"**, ACM Transactions on Graphics, vol. 27, num. 3, 2008

[20] Trajectory, Available from:
http://en.wikipedia.org/wiki/Trajectory, retrieved 10.05.2010

[21] Ali A. Alesheikh, "Developing a Spatio-Temporal Database For Moming Objects", 2006

[22] Natalia Andrienko, Gennady Andrienko , "Building a Visual Summary of Multiple Trajectories ", (Fraunhofer Institute IAIS), 2009

[23] Ryan Eccles, Thomas Kapler, Robert Harper, William Wright, "Stories in GeoTime", 2007

[24] Matthew O. Ward, "A taxonomy of glyph placement strategies for multidimensional data visualization", 2002

[25] Fleet Management, Available from: http://en.wikipedia.org/wiki/Fleet_management, retrieved 27.07.2010

[26] Level of Detail, Available from: http://en.wikipedia.org/wiki/Level_of_detail, retrieved 25.05.2010

[27] Pushpak Kartnik, David Cline, Stefan Jeschke, Anshuman Razdan, Peter Wonka, "Route Visualization using Detail Lenses", IEEE Transactions on Visualization and Computer Graphics, 16(2):235-247, 2010.

[28] Tests of Significance, Available from: http://www.stat.yale.edu/Courses/1997-98/101/sigtest.htm, retrieved 01.07.2010

[29] Two-Sample T-Test, Available from: http://ccnmtl.columbia.edu/projects/qmss/the_ttest/twosample_ttest.html, retrieved 01.07.2010

[30] Distribution Tables, Available from: http://www.statsoft.com/textbook/distribution-tables/#t, retrieved 28.06.2010

[31] Douglas C. Montgomery, George C. Runger, "Applied Statistics and Probability for Engineers", Third Edition, 2003

[32] Box Plots and Five-Number Summaries, Available from: http://www.sinclair.edu/centers/mathlab/pub/findyourcourse/worksheets/122,220/ConstructingBoxPlots.pdf, retrieved 01.07.2010