

ST-TrajVis: Interacting with Trajectory Data

Tiago Gonçalves¹, Ana Paula Afonso¹, Bruno Martins², and Daniel Gonçalves²

¹Departamento de Informática – Faculdade de Ciências da Universidade de Lisboa

²Instituto Superior Técnico, INESC-ID

tgoncalves@lasige.di.fc.ul.pt, apa@di.fc.ul.pt, bruno.g.martins@ist.utl.pt, daniel.goncalves@inesc-id.pt

Huge amounts of movement data are nowadays being collected, as a consequence of the prevalence of mobile computing systems and location based services. While the research interest on the analysis of spatio-temporal data has also significantly increased, there are still several open challenges in areas such as interaction and information visualization. In this paper, we present the first steps of a research project that aims to study the usability of visualization techniques of mobility data. We present ST-TrajVis, an application for the visualization of movement data, based on the innovative combination of two popular techniques, namely a 2D map and a space-time cube, augmented with data processing techniques supporting the interaction with interesting subsets of the data. We conducted a user study to assess the usefulness of ST-TrajVis, and to obtain feedback regarding the users interaction with the different techniques. The results suggest the adequacy of the combination of 2D maps with space-time cubes, the existence of some features of interest to users, and the need to conduct further comparative studies between the different techniques.

Visualization, Interaction, Spatio-Temporal data, Trajectories, Space-Time Cube, 2D Map

1. INTRODUCTION

Several challenges regarding the visual analysis of movement data have been reported in the literature on trajectory visualization, which has been significantly increasing due to the popularity of location based services. The recent advances in mobile position tracking and data management technologies fostered the community's interest, and enhanced our capability of collecting large amounts of data, representing the mobility of people, animals, and natural phenomena (Lee and Krumm 2011).

By definition, a trajectory consists of the evolution of a moving object's spatial properties over time (Dodge et al. 2008). Such changes are often represented as a time-stamped ordered series of location points $P = \langle x_i, y_i, z_i, t_i \rangle$, that compose a trajectory $T = \{P_1, P_2, \dots, P_n\}$, where each x_i , y_i , and z_i , associated to a point P_i represent the geographic coordinates of latitude, longitude, and altitude, at time t_i . Given the characteristics of the data, it is crucial to take into consideration both the spatial and the temporal properties (Andrienko et al. 2010). Each point, or group of points, may also contain one or more thematic attributes associated or derived from the data, representing, for instance, the transportation mode associated to the object. In order to extract useful and relevant information from these large datasets, and to allow the identification of movement patterns and behaviours of individuals

or groups of moving objects, it is crucial to study and develop adequate methods (Dodge et al. 2008). Among these, data visualization and interaction techniques have become increasingly relevant.

Throughout the years, the visual analysis of the spatio-temporal properties of trajectory data, on geographic information systems (GIS), has been possible with static maps, space-time cubes, animated maps, and/or small multiples (Andrienko et al. 2011). Each of these categories may rely on different types of thematic mapping techniques, for the definition of the visual properties of the symbology used. These visualizations may also rely on different pre-processing techniques for the simplification, filtering, and/or enhancement of the data (Lee and Krumm 2011). Despite the several studies, there are still open issues in the fields of visualization and user interaction. Several questions addressing the usefulness of these techniques, the differences between them, and their adequacy for different tasks, emphasize the current lack of empirical data regarding these techniques (Kjellin et al. 2010). In fact, these issues are likely to become more relevant, as more people get involved with spatio-temporal data issues (Andrienko et al. 2010).

In this paper, we present the initial steps of a research project that aims to study the usability of visualization and interaction techniques for mobility

data. We present ST-TrajVis, an application that uses a set of techniques for the visual analysis of trajectory data. We also present a user study with 5 participants, aiming to assess the usefulness of the visualization techniques used, and to obtain some feedback for an initial analysis of the users' interaction with techniques that should be used for trajectory visualization. First, the paper outlines important concepts and related work. Then, it follows with a presentation of ST-TrajVis and its validation through a small user study. The paper finishes with some conclusions and with a description of the challenges for future work.

2. VISUALIZATION OF TRAJECTORY DATA

Andrienko et al. (2010) argue that most visualizations of trajectory data emphasize the spatial properties of the data over the temporal ones. These should not be neglected, as that may undermine the usefulness of the visualization. In addition, they also emphasize that *everyone is a spatio-temporal analyst* and thus a possible user of these data.

In addition to spatial and temporal properties, trajectory data may also include thematic/attributive properties. Based on these, several queries can be made. Peuquet (1994) identified three types, namely: (i) *when+where* \rightarrow *what*, to state the properties of an object at a given time, (ii) *when+what* \rightarrow *where*, to state the location(s) of object(s) at given time(s), and (iii) *where+what* \rightarrow *when*, to state the time or set of times when an object or more was at a certain spatial area. Andrienko et al. (2003) adapted these queries by considering two main components ("when" and "what+where") and two search levels, (elementary and general), resulting into four types of queries: (i) *elementary when and what+where*, to describe characteristics of an object at given time, (ii) *elementary when and general what+where*, to describe the situation at a given time moment, (iii) *general when and elementary what+where*, to describe the dynamics of the characteristics of one or more objects at a certain location over time, and (iv) *general when and what+where* to describe the evolution of the overall situation over time.

Among the methods used to explore and answer to these questions there are several visualization techniques. These should provide useful representations of feature changes/dynamics, which can be achieved through the manipulation of the visual properties of the symbols used to represent the data. Bertin (1967) defined marks as the basic units of visualization. These can be points, lines, areas, and volumes, and they can be modified through the use of certain methods, called visual variables, namely, location, size, shape, value, colour, orientation and

texture. Through the manipulation of these variables, the spatio-temporal and the thematic properties of the data can be represented and/or emphasized in the visualization.

Several categories of techniques have been proposed for the visual exploration and analysis of trajectory data. Among the most common and relevant approaches are the two-dimensional maps and space-time cubes. In the 2D static map, the multiple possibilities for combining the different marks and visual variables have led to the proposal of different approaches for the visualization of spatio-temporal data. For instance, arrows and lines can represent the evolution of an object's position from a starting to a destination point. Certain symbols with pre-defined shapes may represent the locations or events visited by an object. Also, the use of unique colours, transparency levels, or text labels may help on the representation of an object's category, an event's recency, or its position in time (Kraak and Ormeling 2010). Thus, despite their common features, any two visualizations, based on a 2D map, may be significantly different from each other.

The space-time cube technique combines the temporal and spatial properties of the data on a three-dimensional space. Time and space are represented within a cube, where two of its axes, the $x - y$ axis, represent spatial information, while the third, the z -axis, represents time. Typically, the *higher* the information is in cube, the *most recent* it is. Contrarily to other techniques, time is displayed as a spatial position on the cube, allowing other visual variables to be used for the representation of other attributes (Kjellin et al. 2010). Also, interactive techniques that change the point of view within the space-time cube should be used, since the three-dimensional properties of the technique may induce the occlusion of information (Kjellin et al. 2010).

However, when the dataset's dimension and/or complexity is too large, purely visual approaches may be insufficient (Andrienko et al. 2011). Large amounts of data often lead to an over-plot and over-cluttering of the information. As such, when analysing large amounts of data, it is necessary to combine visualization methods with data pre-processing techniques, including data filtering, aggregation or transformation (Deng et al. 2011). Filtering methods consist in the selection of a subset of the entire dataset. This can be achieved based on the temporal properties of the data (e.g., finding all trajectories between a start and an end date), the spatial properties (e.g., finding all trajectories within a geographical area), or the thematic ones (e.g., finding all trajectories of a specific user). Aggregation methods reduce the amount of data to be displayed

by merging subsets of data, based on any of the dataset's components. Finally, data transformation consist in the extraction or reorganization of the dataset. For example, extracting an object's speed based on the its spatial and temporal positions, converting the dataset into another format, or simply reducing its size (e.g. with algorithms such as Douglas-Pecker, to reduce the number of points in a trajectory). When properly used, these methods may reduce the size of the data, and support the identification of general features out of a *noisy* dataset (Andrienko et al. 2011).

However, despite the existence of several techniques for the visualization and interaction with spatio-temporal data, several open challenges remain unsolved due to the lack of empirical knowledge regarding the usability of these techniques, the differences between them, and possible factors that may have an effect over the users' performance.

3. ST-TRAJVIS DESIGN

We developed ST-TrajVis, a web application for the visual exploration of trajectory data, to obtain feedback for an initial analysis of how the users' interact with some of the existing techniques. An example screenshot of the application can be seen in Figure 1. ST-TrajVis is divided into four main interactive components, namely a 2D map, a space-time cube, a data querying area, and a data enhancement area. In the following sections, these components will be described with more detail.

3.1. Data Querying and Enhancement

The data querying area allows filtering the dataset according to spatial and/or temporal properties before visualizing it. ST-TrajVis allows the definition of a geographical area, the start and end dates of search, and the period of hours to be visualized (e.g., visualize the trajectories between 12h and 17h). Additionally, ST-TrajVis provides a data enhancement section that allows the de/activation of the representations of speed, and of the trajectory's recency. It also supports the smoothing of the trajectories, by the temporal and/or the spatial distances between points (e.g. display points at a minimum distance of 20m and/or at a minimum temporal distance of 5 minutes). Finally, this section also controls the display of the possible stop locations, based on the amount of time an object was around a certain geographical location.

3.2. 2D Map and Space-Time Cube

The visualization of the spatial and the temporal properties of the data is made by the combination of

two visualization techniques, not often seen together, namely a 2D map and a space-time cube.

In the last years, several studies have presented these techniques, attempting to compare them, or proving their usefulness in a given context. Although more studies addressing these techniques are needed, some results suggest that both techniques are adequate for different types of data and tasks. While some 2D maps seem more adequate to help answering *elementary* questions, and when metric properties are involved, space-time cubes seem to be beneficial for answering to *general* questions of for ordinal information needs (Kristensson et al. 2009; Kjellin et al. 2010).

As such, ST-TrajVis combines these techniques into the same visualization, exploring the 2D map's focus on spatial information and adequacy in answering to *elementary* questions, and the space-time cube's focus on temporal information and adequacy in answering to *general* questions. In both visualizations, trajectories are displayed as time-ordered points connected with lines. We use colours to represent a time period, defined by the hour of the day. At maximum, four colours are used, representing different time periods, depending on the selected time interval. If the interval is over 24 hours, each colour will represent six hours. If the time interval is less than six hours, each colour will represent one hour and a half. In addition, the lines' width and transparency are, optionally, used to represent the object's approximated speed, and the data's recency, respectively.

Both visualizations provide interaction techniques based on panning and zooming. The space-time cube, in particular, allows the rotation of the cube, changing its point of view (Kjellin et al. 2010), and the rescaling of the cube's height, changing the scale of its temporal representation.

For the representation of the stop locations, in the 2D map, ST-TrajVis places circles at the estimated locations, sized according to the amount of time spent in that area. In the space-time cube, that information is displayed with vertical lines, whose length is proportional to the amount of time spent at that estimated location.

Finally, both visualizations are linked with highlighting techniques. When one point is selected, both visualizations highlight that point, providing additional information about it. If the selected feature is a trajectory point, ST-TrajVis describes its absolute position, date, and the approximated speed of the object. If the feature is a stop position, it will present a list of the time intervals of when the object was detected at that location.

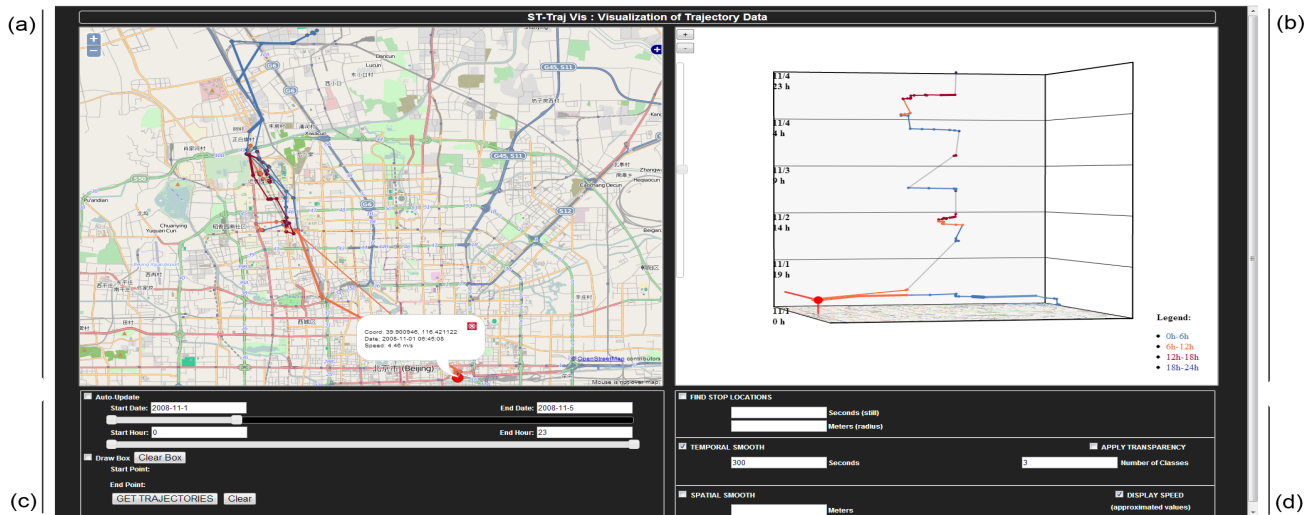


Figure 1: ST-TrajVis. (a) 2D Map (b) Space-Time Cube (c) Data Query Component (d) Data Enhancement Component.

4. USER STUDY

In this section, we describe a usability study conducted with ST-TrajVis. The main goals of this study consisted in assessing the usability of the visualization and interaction techniques used. We also aimed to obtain some data regarding the users' experience with the techniques, which will allow the improvement of the visualizations' features, and the identification of factors of interest to take into consideration when conducting future studies.

A total of 5 participants volunteered to the study, aged between 23 and 28 (Av: 25.6, SD: 2.5). Although all users were experts with computer applications and GIS, all of them were novice regarding applications for analysing spatio-temporal data.

4.1. The Data

The data used in this experiment consists of a subset of the trajectory dataset that was released as part of the GeoLife project (Zheng et al. 2008). Each entry in the dataset is represented as a sequence of time-stamped points, containing the information of latitude, longitude and altitude. Given the early stage of the project and the users' expected unfamiliarity with the tasks proposed, in the scope of this work, the dataset was reduced to the trajectories of one user, in the period of one month.

4.2. Procedure

At the beginning of the experiment, all of ST-TrajVis' features were demonstrated and explained to the participants. Then, each participant had some time to experiment the features of the application, get used to them, and clarify any doubts.

After that, each participant performed six tasks. The first two aimed to assess the participants' ability to interact with the main temporal search mechanisms. These consisted in obtaining all trajectories between a start and end days (T1), and all trajectories from a certain hour of the day to another, within a start and end days (T2). The other tasks aimed to assess the participants' ability to explore and interpret the different types of information provided by the different visual methods. In the third task, the participant had to compare and identify the days when the represented user had moved the most (T3). In the fourth task, the participant had to identify the regions where the represented user spent the most time on (T4). The fifth task consisted in comparing and identifying the days when there was a faster movement (T5). Finally, in the last task, the participant had to determine which hours of the day had the most movement registered (T6). For all tasks, in addition to the users' answers, the amount of time required to complete each task was recorded.

Besides being invited to share their opinions during the execution of the tasks, at the end of the experiment, the participants were asked to answer a questionnaire that aimed to collect their opinions. On a 0 to 10 scale, they had to classify their agreement with certain sentences, namely: (Q1) how well the application's features helped on the tasks; how good were the representations of (Q2) hour (colour), (Q3) data recency (transparency) and (Q4) speed (line thickness); and how useful were the (Q5) 2D map and the (Q6) space-time cube for the completion of the tasks. Finally, the participants were asked about: (Q7) which features they most appreciated; (Q8) which additional features they would suggest; and (Q9) whether they were able to describe any movement pattern(s) from the dataset.

4.3. Results

Figures 2 and 3 represent the participants' mean completion task times, and the mean scores given at the end of the tasks.

All participants, except one, completed the tasks successfully and accurately. Given the feedback of that participant, it was hypothesized that he was not able to finish Tasks T3 and T5 due to performance and perspective issues with the space-time cube.

From the analysis of Figure 2 and the results of Tasks T1 and T2, it can be hypothesized that the application is intuitive enough for the configuration and request of trajectory data. In fact, three participants commented that the learning curve to interact with the application was small, and that the interface's features were intuitive. On the other hand, the results also suggest an overall increasing performance from Task T3 to T6. It can be hypothesized that, despite the different tasks, there was a learning effect, that alongside the knowledge obtained in the familiarization phase, helped the participants performing more easily the tasks.

The analysis of the participants' scores for the first six questions (Figure 3) suggests an overall positive feedback towards the application's features, even if less expressive on question Q3 ("How useful was transparency for the representation of the trajectories' recency?"). Overall, these results suggest that the use of a colour classification code for the representation of temporal periods was the most useful (Q2-Q4). On the other hand, despite the small difference in scores, three participants have shown a higher preference for the space-time cube over the 2D map (Q5-Q6) and commented that the space-time cube was their "main visualization" to perform the tasks, while the 2D map, although good, was mostly used as a complementary tool.

Overall, the most appreciated feature was the display of stop locations, followed by the optional representation of speed, and the existence of filtering and smoothing methods (Q7). Even so, the participants commented their interest on having more filtering and interaction techniques (Q8). Finally, regardless of their performance, all participants identified similar movement patterns of the represented user (Q9).

4.4. Discussion

Despite the small number of participants, the results of this study highlight some interesting issues regarding the application and its features.

Considering the participants' feedback, the results suggest the adequacy of the combination of a 2D

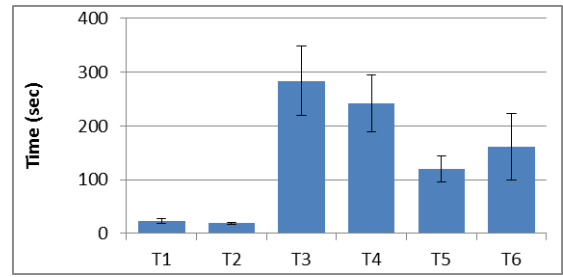


Figure 2: Mean completion times for each task, with standard error bars

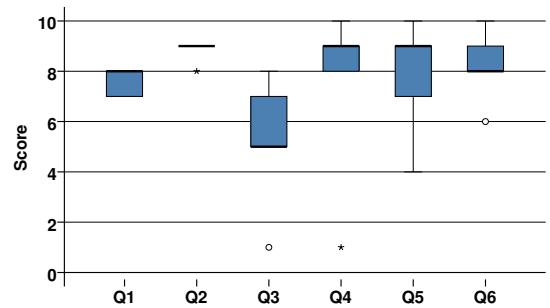


Figure 3: Mean participant scores, for the first six questions of the questionnaire

map with a space-time cube. Despite the problems of one participant with the space-time cube, this technique received a very positive feedback; even being preferred, by some, over a two-dimensional visualization. This is somewhat unexpected, given the participants' unfamiliarity with the technique, and the two-dimensional approach's popularity for representing georeferenced data. In addition to the small learning curve associated with the application, these results suggest the usefulness of the space-time cube for the visualization of spatio-temporal data and encourage the conduction of further comparative studies, to explore the differences between these two techniques, and their adequacy in different tasks.

Another interesting result is related to the features used and/or suggested. Overall, participants have shown interest in dynamically changing the visualization, namely through the de/activation of the representation of some attributes, and through the suggestion of having a wider range of interaction techniques (e.g., filter the data directly in the visualization). However, the spatial and the temporal smoothing features, and the use of transparency to represent the data's recency, were barely used. This could suggest that the participants did not feel the need to use them or these features must be presented in a more intuitive way.

In addition, the dataset used in this study described the trajectories of one user. One of the reasons presented by the participants regarding their preference for the space-time cube was the un-cluttered visualization of the trajectories, in comparison to the 2D map. If the visualization displayed the mobility of more users, there would be, most likely, some cluttering. As such, this feedback emphasizes the need to compare these techniques with larger amounts of trajectories and with different complexities, to assess their scalability.

5. CONCLUSIONS

Recent advances in tracking technologies and location-based services have given rise to the interest and collection of trajectory data. However, the spatio-temporal properties of this type of information still raise several challenges that cross different research areas, such as visualization and interaction. Assuming that *everyone is a spatial-temporal analyst* (Andrienko et al. 2010), and therefore, a potential user of these data and techniques, this issue is likely to become increasingly relevant.

We presented ST-TrajVis, a web application for the visual exploration and analysis of trajectory data. This is the first step of our ongoing research work aiming to address the lack of empirical knowledge regarding the usefulness and the differences between visualization techniques for trajectory data. We also presented a user study for the usability assessment of ST-TrajVis and to obtain some feedback from users regarding their experience with the techniques provided. Despite its small-scale, the study suggests that ST-TrajVis provides a set of features easy to learn and intuitive to use, even for less experienced users. More importantly, it highlights the need to conduct further studies, including: the comparative assessment of the effect of the number of trajectories and their complexities over 2D maps and space-time cubes; or the exploration of usable methods to better interact with the visualizations.

Ongoing work regarding this project passes through: the improvement of ST-TrajVis namely, by adding new and/or more complete filtering mechanisms; the development of taxonomies of visualization tasks and techniques, for the exploration of trajectory data; the development of new prototypes for the comparative assessment of the visualization techniques based on those taxonomies and possible factors of interest, as number of trajectories or trajectory complexity; and, taking into consideration the feedback given by the users, study possible new interaction methods with the visualizations.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of FCT (Portuguese Science and Technology Foundation) through project SInteliGIS (PTDC/EIA-EIA/109840/2009) and the scholarship with reference SFRH/BD/78837/2011.

REFERENCES

- Andrienko, G., Andrienko, N., Bak, P., Keim, D., Kisilevich, S., and Wrobel, S. (2011). A conceptual framework and taxonomy of techniques for analyzing movement. *Journal of Visual Languages and Computing*, 22(3):213–232.
- Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M.-J., Schumann, H., and Tominski, C. (2010). Space, time and visual analytics. volume 24, pages 1577–1600.
- Andrienko, N., Andrienko, G., and Gatalisky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. volume 14, pages 503–541.
- Bertin, J. (1967). *Semiology of Graphics: Diagrams, Networks, Maps*.
- Deng, K., Xie, K., Zheng, K., and Zhou, X. (2011). Trajectory indexing and retrieval. In *Computing with Spatial Trajectories*, chapter 2, pages 35–60. Springer.
- Dodge, S., Weibel, R., and Lautenschitz, A.-K. (2008). Towards a taxonomy of movement patterns. *IV 2008*, 7(3-4):240–252.
- Kjellin, A., Pettersson, L. W., Seipel, S., and Lind, M. (2010). Evaluating 2d and 3d visualizations of spatiotemporal information. volume 7, pages 19:1–19:23.
- Kraak, M.-J. and Ormeling, F. (2010). *Cartography Visualization of Spatial Data*. Pearson Education, 3rd edition.
- Kristensson, P. O., Dahlbäck, N., Anundi, D., Björnstad, M., Gillberg, H., Haraldsson, J., Mårtensson, I., Nordvall, M., and Ståhl, J. (2009). An evaluation of space time cube representation of spatiotemporal patterns. *IEEE Transactions on Visualization and Computer Graphics*, 15(4):696–702.
- Lee, W.-C. and Krumm, J. (2011). Trajectory preprocessing. In *Computing with Spatial Trajectories*, chapter 1, pages 3–33. Springer.
- Peuquet, D. J. (1994). It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. volume 84, pages 441–461.
- Zheng, Y., Li, Q., Chen, Y., Xie, X., and Ma, W.-Y. (2008). Understanding mobility based on gps data. In *UbiComp 2008*, pages 312–321. ACM.