Visualization Techniques of Trajectory Data: Challenges and Limitations

Tiago Gonçalves [1], Ana Paula Afonso [1], and Bruno Martins [2]

[1] Faculdade de Ciências da Universidade de Lisboa,
[2] Instituto Superior Técnico, INESC-ID
tgoncalves@lasige.di.fc.ul.pt,apa@di.fc.ul.pt,bruno.g.martins@ist.utl.pt

Abstract. As a consequence of the prevalence of mobile computing and location based services, huge amounts of movement data are nowadays being collected. While the research interest on the analysis of trajectory data has also significantly increased, there are still several open challenges in areas related to geographic information systems. Despite the existence of several techniques for the visualization of movement data, it is still unclear how usable and useful these techniques are, how can they be improved, and in which tasks they should be used. In this paper, we highlight the current limitations in the visual exploration of trajectory data, and present the ongoing research aiming to address those issues. For that, we propose the development of taxonomies addressing visualization tasks, techniques, and data, based on empirical data, through systematic comparative usability studies, and present an overview of the current results.

1 Introduction

The intrinsic relation between what we do and where we are emphasizes the importance of the analysis of geo-referred data [1]. This fostered the community's interest and enabled the collection of huge amounts of spatio-temporal data representing the trajectories of people, animals, and natural phenomena [2].

Trajectory data is an important factor on many processes and activities in several areas of research, like traffic analysis, or the identification of behaviours and moving patterns. By definition, a trajectory consists in the evolution of a moving object's spatial properties over time [3]. Typically, this type of data is represented as a time-stamped oriented series of location points, $P = \langle x_n, y_n, t_n \rangle$ or $P = \langle x_n, y_n, z_n, t_n \rangle$, that compose a trajectory $T = \{P_1, P_2, ...P_n\}$, where x_i, y_i, and z_i represent, respectively, the geographic coordinates of latitude, longitude and altitude, at time t_i. Moreover, each point, or group of points, may contain additional attributes, derived or associated from the data, representing other types of information. These subsets of the data, often called thematic, may represent, for instance, the object's speed at a given instant, its current state, or its category. Therefore, time, space, and thematic attributes play an important role when analysing these data [4].

In order to support the extraction of useful and relevant information from these large datasets, it is crucial to develop and study adequate data visualization techniques. In the context of geographic information systems, several techniques have been proposed to help on the visualisation and exploration of the spatio-temporal properties of trajectory data [5] [6] [7]. Despite the various approaches, these techniques can be grouped into four main high-level categories, namely: (i) static maps, (ii) space-time cubes, (iii) animated maps, and (iv) small multiples [7].

However, despite the useful results obtained, several challenges and open issues can still be identified. In particular, in the field of visualization and human computer interaction, it is still unclear how usable these techniques really are, and in which tasks they should be used.

In this paper, we highlight the main limitations regarding the visual exploration of trajectory data, and describe the work in progress addressing these issues. The following sections address an overview of the existing work, followed by the description of the main limitations in the visualization of trajectory data, and finally, the description of the ongoing research's objectives and methods addressing these limitations.

2 Related Work

The following section presents some relevant concepts associated with the visual exploration of trajectory data, addressing the characterization of trajectory data, and the methods used to process and, later, visualize the data.

According to Peuquet [8], spatio-temporal data and, in particular, trajectory data can be characterized by three main components: space (*where*), time (*when*), and topics (*what*). In turn, several questions can be made by combining these components, namely: (i) *when+where* → *what*, to state the properties of an object at a given time; (ii) *when+what* → *where*, to state the location(s) of object(s) at given time(s); and (iii) *where+what* → *when*, to state the time or set of times when an object or more was at a certain spatial area.

Other authors, such as Andrienko et al. [9], argue that, asides from these factors, analysts should also take into consideration search levels, *elementary* and *general*, depending if the focus of the analysis is one or more objects, respectively, and cognitive operations, such as identify and compare. By combining these levels with two main search targets (*when* and *what+where*), the authors identified four types of questions: (i) *Elementary When and What+Where*, to describe characteristics of an object at given time (e.g. where is person A at 9 am?); (ii) *Elementary When and General What+Where*, to describe a situation at a given time moment (e.g. Are persons A and B close to each other at 9 am?); (iii) *General When and Elementary What+Where*, to describe the dynamics of the characteristics of an object over time at a certain location (e.g. which buildings did person A visited during the day?); and (iv) *General When and What+Where*, to describe the evolution of the overall situation over time (e.g. who visited the most locations between 9 and 10 am?).

To properly answer these questions, adequate visualization techniques are needed. This often means mapping temporal information representing feature changes or dynamics of an object [10]. These can be obtained through the adaptation of certain visual variables [11], including, among others, colour, size, shape and orientation, that can be used over different marks, namely, points, lines, areas, surfaces and volumes. Through the adaptation of these visual marks, it is possible to represent the characteristics and variations of spatial, temporal, and thematic attributes.

In addition, due to the sheer amount of data present in trajectory datasets, purely visual approaches may be insufficient [7]. This requires visualization methods to be preceded and/or combined with data processing techniques to reduce the amount of data, including methods such as data cleaning, filtering, and/or smoothing techniques.

Data cleaning consists of the deletion of erroneous dataset entries [2]. Data filtering methods consist of the selection of a subset of the dataset, based on the spatial, the temporal, and/or the thematic proprieties of the data. Data smoothing methods transform the dataset by reducing the level of detail of a trajectory, for instance, using algorithms such as Douglas-Pecker, to reduce the number of points in a trajectory [12]. An additional solution consists of aggregation methods to merge subsets of the data, based on their similarity (e.g. direction of movement, position in time/space).

After cleaned/filtered, these *raw trajectories* can be enriched with more data, becoming *semantic trajectories* [13]. This extra information is used to represent, among others, the mode of transportation of the moving object, or the activities performed (e.g. working, shopping) [14]. These can be obtained by inferring data from the original dataset (e.g. determine the dislocation mode of the object based on its speed) or by combining them with other datasets (e.g. determine that a person is eating, since it is close to a restaurant).

In order to properly visualize these transformed trajectory datasets several visualization methods were presented.The following sections present an overview of the most used techniques for spatio-temporal, and trajectory data visualization.

2.1 Static Maps

Two dimensional static maps are one of the most common approaches for representing any type of information of a geographical location.

The analysis of the existing literature reveals a wide range of options to represent trajectory data information on static maps. This happens due to the several possible combinations of visual variables into complex symbols [15] and/or the use of data processing techniques [16].

Typically, lines/arrows are used to represent the spatio-temporal properties of an object's movement from a starting to a destination point, while specific symbols with pre-defined shapes may represent different events occurring in space (see Figure 1). Similarly, other visual attributes like colour, line thickness, or transparency may be used to represent an event's category or recency [10].

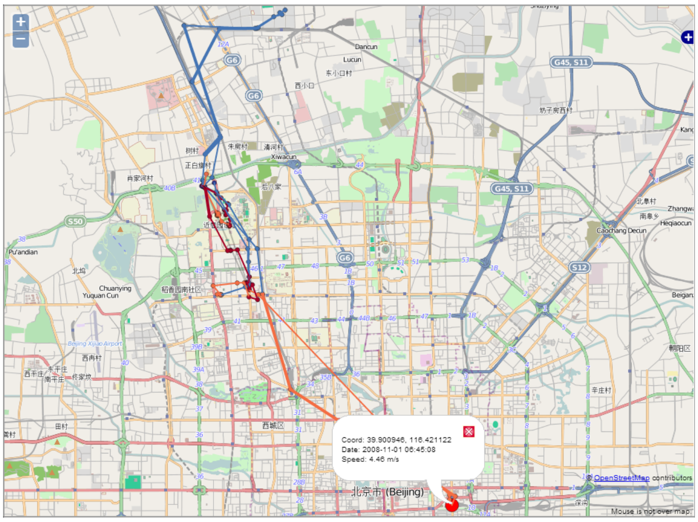

Fig. 1. Two-dimensional static map representation. Colours are used to represent the period of the day, while line width is used to represent the object's approximated speed.

2.2 Space-Time Cube Map

With the increasing evolution of computer graphics, three dimensional visualizations have become more common and are seen as a promising way to represent complex information [17].

The space-time cube map is a visualization technique that emphasizes the idea that time and space are inseparable, by using the three dimensions to represent both attributes [17]. Typically, the $x - y$ axis of the cube are used to represent spatial information, while the third dimension, the z-axis, is used to represent time. Usually, time increases along the z-axis, implying that the higher the information is within the cube, the most recent it is (see Figure 2) [18].

In this type of map, time is represented as a spatial position, therefore, it is not required other visual variables to convey that information [18]. However, this technique may be affected by the human perceptual issues of three dimensional environments [18]. These may induce the occlusion of information, and problems of interpretation, in particular, of metric properties [18], regarding three dimensional perspectives. Therefore, it is recommended the use of interactive techniques that allow the user to change the point of view within the space-time cube [18].

2.3 Animated Maps and Small-Multiples

Although different in presentation, animated and small multiple maps follow a somewhat similar approach in the way the data is handled. In both techniques, data is divided between several maps, each one representing the state of a phenomenon, at a different time period. However, while animated approaches display

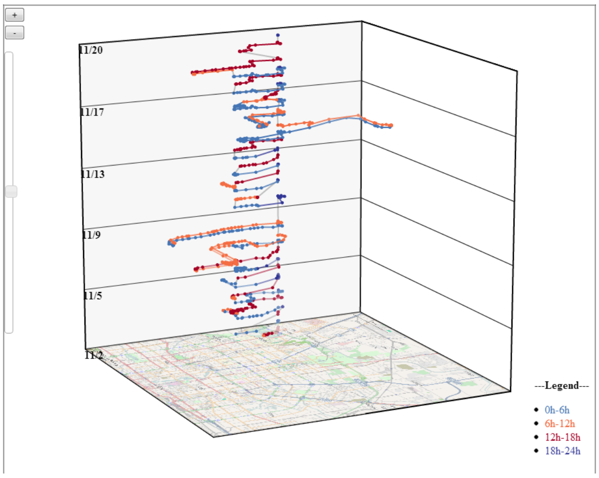

Fig. 2. Space-time cube representation. Colours are used to represent the period of the day, while line width is used to represent the object's approximated speed.

these maps (frames) as a sequence in a single view, small multiple maps present them juxtaposed to each other (see Figures 3 and 4).

Preliminary research argues that animated displays can help revealing spatio-temporal patterns that are not evident with common static representations [19]. Besides, unlike static approaches, these have an additional dimension that can be used to present information. However, the longer the animation, the higher the amount of data to be presented. Since one frame will not be always visible, this may raise some cognitive and perceptual limitations [20]. Controls over the animation can also be given to the user, allowing him to move towards or backwards on the animation, thus allowing frames to be visible again.

On the other hand, although small multiple maps minimize cluttering issues by splitting the information through separated maps, and allow the comparison of multiple time periods at the same time, these depend on the screen size, which determines the number of maps that can be presented simultaneously. Moreover, small multiples are discontinuous spatial presentations, which may require a larger effort from the viewer when a large number of maps is used, since it requires the user to mentally connect all maps into an ordered sequence.

3 Current Challenges and Limitations

Recently, several approaches addressing the visualization of trajectory data have been reported. However, despite the results obtained, we can still identify several challenges and open issues.

Andrienko et al. [9] stressed that existing visualizations often neglect the temporal properties of the data, which may undermine the visualizations. In addition, they emphasize that *everyone is a spatio-temporal analyst* and, thus,

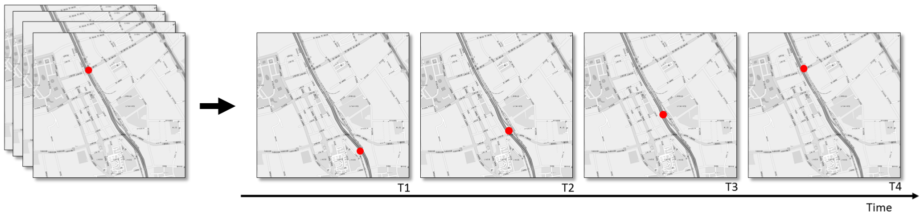

Fig. 3. Animated map representation. Each map represents an instant in time, and is displayed once.

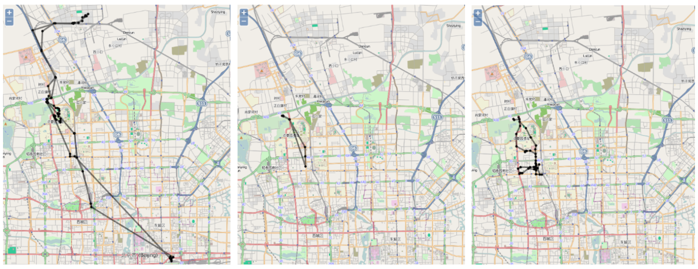

Fig. 4. Small multiple map representation. In this example, each map represents one day of recorded movement. All maps are visible.

it should not be assumed that the analysis of spatio-temporal data will always be conducted by a specialist, therefore, making the community of users potentially unlimited.

On the other hand, although the number of visualization studies has increased, usability has been, somewhat, neglected [21]. When evaluating a certain visualization, most usability studies tend to be either limited to case studies and/or focus on highly expert users of the system [21].

In addition, previous studies suggest an uncertainty regarding the techniques and the procedures used on comparative evaluations [22][23]. In fact, the results of some studies are, sometimes contradictory, in particular, when dealing with animated representations [22]. Tverky et al. [23], in particular, suggest that there is no evidence to conclude that animated representations are more helpful than others, since previous studies either fail to assure that informationally similar visualizations are used, or have unbalanced procedures, with different techniques. Moreover, some authors suggest possible factors that may have an impact on the comparison of these visualization techniques [5] [6].

On the other hand, given the diverse results obtained in terms of the adequacy with different techniques, some authors highlight the need for new taxonomies of tasks and data, based on empirical findings in different types of tasks [17].

Based on these limitations, it is important to study more deeply the visualization of trajectory data and explore the empirical data, based on the users' performance, to better support this knowledge.

4 Proposed Approach and Expected Results

Despite the existence of several studies addressing the importance of the visual analysis of trajectory data, several limitations have been identified, which emphasize that the knowledge associated with these techniques is still scarce.

The results of previous studies suggest some techniques' adequacy for certain types of tasks (and data) over others. However, while some studies present mixed conclusions, others highlight some uncertainty regarding the procedures, and possible interfering factors in some experiments. Although some taxonomies, partially, address some of these issues [7], these are rarely based on empirical data. As such, to minimize these limitations, we propose the development of taxonomies of tasks, data, and techniques for trajectory data, through the systematic empirical assessment of the existing main groups of techniques and possible impacting factors over them, and, consequently, the development of an evaluation framework.

Usually themes such as the description of the data, the visualization techniques, and the types of tasks/operations over spatio-temporal data tend to be addressed separately. However, we argue that these are actually related to each other. Similarly to the work of Andrienko et al. [9], cognitive tasks like the identification, or the comparison of certain data features can be done considering the spatio-temporal and thematic properties of the data. Moreover, these can also be associated with the properties of visual variables, that determine how adequate a variable is to represent a certain type of data [11].

Consequently, as one of the first steps in this research, we have defined a taxonomy addressing these various components. While the first connection between these components (tasks-data) provides a set of tasks that can be used in a comparative assessment between techniques, the second connection (tasks-variables) highlights the visual methods that can be used to better represent the required information for certain tasks. Nevertheless, despite these initial results, we argue that this taxonomy can still be expanded by addressing techniques for the visualization of trajectory data, based on the results of comparative studies.

When exploring trajectory data, several factors may have an effect over the users' performance. Based on the considerations of previous research, we expect useful results from the empirical assessment of some factors, such as: (i) density/complexity of information, (ii) type of dataset, (iii) the familiarity of the user with the spatial location, (iv) the size of the visualization.

In order to assess these factors, it is necessary to conduct comparative evaluations between the existing techniques, following the tasks identified in the previously mentioned taxonomy.

It is unlikely to find a visualization technique that outperforms every other [7], however, it is plausible to assume that these evaluations may reveal new problems with the techniques, new factors to be explored, and, by consequence, new variations to be tested [24]. It is expected an overall reduction in the user's performance with a higher information density and/or complexity, since the larger the dataset, the higher the probability of over-plotting/cluttering of information. Similarly, the more types of information to represent, the more visual variables may be needed to encode all of the information. Consequently, this may increase the cognitive workload needed to interact with the visualization.

On the other hand, with the advances in mobile technology, users have become more proficient in interacting with smaller screens. This raises several challenges, since the smaller the screen, the less information a visualization may contain. In turn, this may hinder small multiple map techniques, since the number of possible visible maps is reduced. However, whether and/or how these factors significantly affect the user's performance with other techniques is still unclear.

In order to obtain relevant results, it is necessary to have *some* users to participate on the experiments. Since *everyone is a spatio-temporal analyst* [9], and thus, the community of users of spatio-temporal visualizations is potentially unlimited, we argue that the conduction of comparative studies with (supposedly) inexperienced users may not only reveal important results, but may also be considered as necessary.

Previous studies argue that novice users can be particularly helpful, due to three main reasons [5]. First, it is difficult to find expert user in a reasonable quantity that allow to obtain reliable results. Second, proving a visualization's usefulness for novice users, may provide a useful empirical building block to help researchers recruiting less experienced users. Finally, if novice users are able to properly interact with a visualization, there should be no reason to assume that expert users would not be able to do the same (if not better).

In fact, our previous research goes in agreement with these considerations [25]. Following the previously mentioned taxonomy, and through the use of several processing data techniques, we developed and tested ST-TrajVis, an application for the visualization of trajectory data. The results suggest that the techniques used, such as the space time cube, are easy to learn and intuitive to use, even for less experienced users. More importantly, it highlighted the need to conduct further studies addressing the effects of the number of trajectories, the complexity of the data representation, and the exploration of usable and noticeable methods to better interact with the visualizations.

Ultimately, our aim is to develop an evaluation framework of trajectory data visualizations. In fact, the process of this research addresses several components expected on an evaluation framework, namely: a set of representative visualization tasks, associated with common (trajectory) data properties (through the definition of the taxonomies); the identification of possible users of these visualizations, of factors that may have an effect over the user's performance, and metrics to evaluate them (through the comparative assessment of visualization techniques); and also a set of relevant results, from previous experiences, that may contribute for future research/analysis.

5 Conclusions and Future Work

The visual exploration of trajectory data has become an important element on several activities, which raised the interest and the collection of several trajectory datasets. Despite the various studies addressing this type of data, there are still many open challenges that cross different research areas, such as visualization and human-computer interaction.

While several visualization approaches have been reported, the lack of concrete (and empirical) knowledge regarding the usability and usefulness of these techniques, alongside the potentially unlimited users are still relevant issues.

In this paper, we highlighted the main limitations of the visual exploration of trajectory data, and presented the ongoing research addressing those issues. We argue that, through systematic comparative user studies, it will be possible to: (i) develop taxonomies of tasks, data, and visualization techniques, that allow spatio-temporal analysts to understand, and evaluate whether a certain visualization technique is useful for a given task; (ii) empirically compare the different visualization techniques; (iii) effectively identify possible factors that may have an impact over the users' performance, when interacting with techniques for the visualization of trajectory data; (iv) improve existing visualization techniques.

As the first steps in our work, based on the analysis of previous studies, we developed a taxonomy addressing trajectory data components, cognitive visualization tasks, and visual variables. Although, we argue that this taxonomy needs to be expanded, it contributed to the creation and assessment of a small visualization application that allowed for a preliminary interaction with novice users, in terms of spatio-temporal data analysis, and highlighted several important challenges.

Current ongoing work consists on the expansion of the referred taxonomy to address visualization techniques for trajectory data, and its use on the comparative user-based evaluation of representative prototypes of the various visualization techniques.

References

1. Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W.: Geographical Information Systems and Science. 2nd edn. John Wiley & Sons, Ltd (2005)
2. Zheng, Y., Zhou, X., eds.: Computing with Spatial Trajectories. 1st edn. Springer (2011)
3. Dodge, S., Weibel, R., Lautenschtz, A.K.: Towards a taxonomy of movement patterns. IV 2008 **7**(3-4) (2008) 240–252
4. Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S.I., Jern, M., Kraak, M.J., Schumann, H., Tominski, C.: Space, time and visual analytics. Volume 24. (2010) 1577–1600
5. Kristensson, P.O., Dahlbäck, N., Anundi, D., Björnstad, M., Gillberg, H., Haraldsson, J., Mårtensson, I., Nordvall, M., Ståhl, J.: An evaluation of space time cube representation of spatiotemporal patterns. IEEE Transactions on Visualization and Computer Graphics **15**(4) (2009) 696–702
6. Willems, N., van de Wetering, H., van Wijk, J.J.: Evaluation of the visibility of vessel movement features in trajectory visualizations. Computer Graphics Forum (2011) 801–810
7. Andrienko, G., Andrienko, N., Bak, P., Keim, D., Kisilevich, S., Wrobel, S.: A conceptual framework and taxonomy of techniques for analyzing movement. Journal of Visual Languages and Computing **22**(3) (2011) 213–232

8. Peuquet, D.J.: It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. Association of American Geographers **84**(3) (1994) 441–461
9. Andrienko, N., Andrienko, G., Gatalsky, P.: Exploratory spatio-temporal visualization: an analytical review. Volume 14. (2003) 503–541
10. Kraak, M.J., Ormeling, F.: Cartography Visualization of Spatial Data. 3rd edn. Pearson Education (2010)
11. Bertin, J.: Semiology of Graphics: Diagrams, Networks, Maps. (1967)
12. Heckbert, P.S., Garland, M.: Survey of polygonal surface simplification algorithms (1995)
13. Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M.L., Gkoulalas-divanis, A., Macedo, J., Pelekis, N., Theodoridis, Y., Yan, Z.: Semantic trajectories modeling and analysis. ACM Computing Surveys (2012) 42:1–42:32
14. Xie, K., Deng, K., Zhou, X.: From trajectories to activities: a spatio-temporal join approach. In: Proceedings of the 2009 International Workshop on Location Based Social Networks. LBSN '09, ACM (2009) 25–32
15. Pu, J., Liu, S., Ding, Y., Qu, H., Ni, L.: T-watcher: A new visual analytic system for effective traf?c surveillance. In: Proceedings of the 14th International Conference on Mobile Data Management. MDM '13 (2013) 127–136
16. Andrienko, G., Andrienko, N.: Spatio-temporal aggregation for visual analysis of movements. In: Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on. (2008) 51–58
17. Kjellin, A., Pettersson, L.W., Seipel, S., Lind, M.: Evaluating 2d and 3d visualizations of spatiotemporal information. ACM Transactions on Applield Perception **7**(3) (2010) 19:1–19:23
18. Kjellin, A., Pettersson, L.W., Seipel, S., Lind, M.: Different levels of 3d: an evaluation of visualized discrete spatiotemporal data in space-time cubes. Information Visualization **9**(2) (jun 2010) 152–164
19. MachEachren, A.M., Coscoe, F.P., Haug, D., Pickle, L.: Geographic visualization: Designing manipulable maps for exploring temporally varying georeferenced statistics. In: Proceedings of the 1998 IEEE Symposium on Information Visualization. INFOVIS'98 (1998) 87–94
20. Harrower, M., Fabrikant, S.: 4. In: The Role of Map Animation for Geographic Visualization. Wiley (2008) 49–65
21. Chen, C.: Top 10 unsolved information visualization problems. Computer Graphics and Applications, IEEE **25**(4) (2005) 12 – 16
22. Lobben, A.: Influence of data properties on animated maps. Annals of the Association of American Geographers **98**(3) (2008) 583–603
23. Tversky, B., Morrison, J.B., Betancourt, M.: Animation: can it facilitate? International Journal of Human-Computer Studies **57**(4) (2002) 247 – 262
24. Tohidi, M., Buxton, W., Baecker, R., Sellen, A.: Getting the right design and the design right. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '06 (2006) 1243–1252
25. Gonçalves, T., Afonso, A.P., Martins, B., Gonçalves, D.: St-trajvis: Interacting with trajectory data. In: Proceedings of the 27th International BCS Human Computer Interaction Conference (HCI 2013). (2013) 1–6