UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

MAP-BASED INTERACTION WITH TRAJECTORY DATA

Doutoramento em Informática
Especialidade em Engenharia Informática

Tiago José Lopes Gonçalves

Tese orientada por:
Professora Doutora Ana Paula Pereira Afonso e
Professor Doutor Bruno Emanuel da Graça Martins

Documento especialmente elaborado para a obtenção do grau de doutor

2017

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

MAP-BASED INTERACTION WITH TRAJECTORY DATA

Doutoramento em Informática
Especialidade em Engenharia Informática

Tiago José Lopes Gonçalves

Tese orientada por:
Professora Doutora Ana Paula Pereira Afonso e
Professor Doutor Bruno Emanuel da Graça Martins

Júri:
Presidente:
- Nuno Fuentecilla Maia Ferreira Neves, Professor Catedrático
  Faculdade de Ciências da Universidade de Lisboa

Vogais:
- Doutora Marlène Villanova-Oliver, *Assistant Professor*
  Université Grenoble Alpes, França;
- Doutora Maria Beatriz Alves de Sousa Santos, Professora Associada com Agregação
  Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro;
- Doutor Manuel João Caneira Monteiro da Fonseca, Professor Associado
  Faculdade de Ciências da Universidade de Lisboa;
- Doutora Ana Paula Pereira Afonso, Professora Auxiliar
  Faculdade de Ciências da Universidade de Lisboa, Orientadora;
- Doutor Paulo Jorge Cunha Vaz Dias Urbano, Professor Auxiliar
  Faculdade de Ciências da Universidade de Lisboa

Documento especialmente elaborado para a obtenção do grau de doutor

Este trabalho foi financiado pelas seguintes instituições:
FCT (Fundação para a Ciência e Tecnologia) - ref: SFRH/BD/78837/2011
LASIGE, Faculdade de Ciências, Universidade de Lisboa - ref: UID/CEC/00408/2013
INESC-ID, Instituto Superior Técnico, Universidade de Lisboa - ref: UID/CEC/50021/2013

2017

Acknowledgements

I would like to thank to everyone who made this endeavour possible.

First of all, thank you Mr. José and Mrs. Maria do Céu Gonçalves (which I more often address as dad and mom) for the never ending support and patience.

Then, I would like to thank to my supervisors, Professors Ana Paula Afonso and Bruno Martins, for their continued work, revisions, and advices, that in addition to the work here described, greatly contributed to my personal and professional growth.

In addition, thank you to all of my current and former colleagues who directly or indirectly contributed for this work, important names include: André, Diogo, José, Luis, Nádia ...

Naturally, I would also like to thank to the LASIGE and INESC-ID research units, for all the conditions and support provided.

As shown throughout nearly all chapters of this document, several volunteers contributed in various long-duration usability studies and, as such, I would also like to thank their participation.

Several names are missing from this list, nevertheless, these thanks extend to them as well.

Finally, in a more formal note, I would like to acknowledge that this work was supported by *Fundação para a Ciência e Tecnologia* (FCT) through the Individual Ph.D. Schoolarship with reference SFRH/BD/78837/2011, and through fundings to the LASIGE Research Unit, from the Faculdade de Ciências, Universidade de Lisboa and to INESC-ID, Instituto Superior Técnico Universidade de Lisboa, with references UID/CEC/00408/2013 and UID/CEC/50021/2013.

Resumo

Atualmente, várias decisões importantes estão diretamente relacionadas com o conhecimento e a análise de informação relativa a determinadas localizações ou ao movimento detetado sobre as mesmas. Estas decisões podem ter um impacto significativo quer em contextos fundamentalmente pessoais (e.g., que caminho seguir para determinado ponto de interesse?) como em contextos com impacto social (e.g., será necessário expandir o serviço de táxis para aumentar a cobertura de uma determinada região?). Por sua vez, este facto tem fomentado o interesse de várias comunidades de investigação, que procuram perceber os efeitos e potencialidades da utilização de informação georreferenciada de forma a melhorar a vida dos cidadãos/utilizadores em variados contextos.

Com a crescente popularidade associada a sistemas de computação móvel e tecnologias de geoposicionamento, também a recolha de grandes quantidades de dados espácio-temporais se tornou cada vez mais comum, eficiente e eficaz, sendo, consequentemente, cada vez mais comum a análise de informação georreferenciada relacionada com as trajetórias de pessoas ao longo do tempo.

Por definição, uma trajetória consiste na evolução da posição de um objeto ao longo do tempo. Tipicamente, este tipo de dados são representados como uma sequência de pontos ordenados temporalmente e complementados com informação temática, que descreve o objeto ou a localização em função de vários atributos não espaciais ou temporais (e.g., velocidade ou meio de transporte).

Devido às propriedades espaciais associadas com dados de trajetórias, as técnicas baseadas em mapas interativos são frequentemente consideradas como ferramentas essenciais para a visualização destes dados. Em particular, mapas estáticos bidimensionais (2D) destacam-se como uma das técnicas mais comuns para a representação de qualquer tipo de informação georreferenciada. Tirando partido da manipulação das variáveis visuais (e.g., cor ou opacidade) associadas a marcas, como pontos ou linhas, este tipo de técnicas suporta a visualização de vários eventos, padrões, e relações associadas às trajetórias sobre um plano geográfico. Contudo, apesar das suas vantagens, os mapas estáticos 2D tendem ainda a negligenciar a visualização dos dados num contexto temporal. Para minimizar este problema, estas técnicas necessitam, geralmente, de utilizar uma simbologia mais complexa ou visualizações adicionais.

Com o crescente desenvolvimento associado à computação gráfica, mapas tridimensionais (3D) e, em particular, cubos espácio-temporais (*space-time cube* - STCs) têm sido considerados como possíveis alternativas para a visualização de trajetórias. Os STCs combinam as representações de informação espacial e temporal na mesma vista, tipicamente, recorrendo à representação de

informação espacial através dos eixos x e y, enquanto que o eixo z é usado para distribuir a informação temporalmente. Apesar de se poder recorrer a estratégias de representação semelhantes à de mapas estáticos comuns (e.g., alteração de cores ou tamanhos de pontos e linhas), a associação da terceira dimensão ao contexto temporal dos dados suporta, para além da representação direta de informação temporal, a visualização do estado de um ou vários objetos em várias instâncias de tempo diferentes.

Ainda que pudessem ser considerados mais tipos de técnicas de visualização baseadas em mapas, a análise de mapas estáticos 2D e STCs engloba uma diversidade significativa em termos de possíveis representações para trajetórias. Consequentemente, a escolha de uma determinada técnica de visualização pode ser considerada como um problema importante, por vezes até controverso. Apesar do crescente número de aplicações presentes na literatura para a visualização deste tipo de informação, a validação da utilidade e usabilidade destas tende a ser baseada em exemplos/casos de uso, frequentemente negligenciado o papel fundamental do utilizador (humano) no processo de validação (e.g., será que o utilizador consegue perceber e utilizar com facilidade as funcionalidades disponibilizadas?).

No contexto deste trabalho, consideramos que estes problemas são particularmente importantes e prováveis de se tornarem cada vez mais relevantes. Apesar da existência de regras de *design* úteis para a criação de mapas, estas ainda são algo limitadas num contexto interativo. Por outro lado, os resultados existentes do ponto de vista empírico são ainda demasiado esparsos e difíceis de generalizar. Nos últimos anos, tem havido um maior interesse em apresentar informação espácio-temporal a mais utilizadores, refletido com a crescente popularidade de ferramentas de monitorização pessoal em contextos reais e virtuais, baseadas na análise de informação espácio-temporal. Por sua vez, ao contrário do que era comum, uma percentagem significativa dos utilizadores destas aplicações são parcial a completamente inexperientes em termos de análise e visualização de informação espácio-temporal, sendo, ainda assim, expostos ao papel de analista.

Neste trabalho pretendemos abordar estes problemas, pelo que foram delineados os seguintes objetivos principais:

- Obter um conjunto de resultados quantitativos e qualitativos que permitam caracterizar as vantagens e desvantagens de diferentes tipos de técnicas baseadas em mapas (mapas estáticos 2D e STCs), para a visualização de dados de trajetórias;

- Obter um conjunto de resultados quantitativos e qualitativos que permitam averiguar a viabilidade de diversos melhoramentos em técnicas de visualização existentes (mapas estáticos 2D e STCs);

- Identificar um conjunto de *guidelines* para a seleção e concepção de mapas estáticos 2D e STCs.

Nesta dissertação, apresentamos um estudo divido em quatro partes focado na comparação e validação de abordagens para a visualização de conjuntos de dados referentes a trajetórias através de mapas estáticos 2D e STCs.

Numa primeira fase focámo-nos na análise das contribuições existentes na literatura mais relevantes para esta investigação. Nesta etapa, foram analisadas as principais características geralmente associadas a dados de trajetórias e os tipos de técnicas baseadas em mapas, para a visualização deste tipo de dados, assim como os principais tipos de tarefas de análise visual de informação necessários aquando da utilização de uma dada aplicação. Estas tarefas incluem, entre outras, a localização de informação numa visualização, a identificação de informação baseada na sua representação, a comparação de dados ou análise de padrões e associações. Adicionalmente, propusemos um modelo hierárquico de tarefas, de modo a simplificar e organizar as várias propostas existentes na literatura, assim como simplificar a concepção de estudos de usabilidade baseado nas tarefas descritas.

Com base neste modelo de tarefas, na fase seguinte abordámos a caracterização de mapas estáticos 2D e STCs baseados na sua utilidade e usabilidade para a resolução de vários tipos de tarefas. Para isso, comparámos protótipos baseados nas duas técnicas em estudos de usabilidade, cujos resultados obtidos sugerem, globalmente, que apesar de ambas as técnicas serem usáveis por utilizadores inexperientes, os mapas estáticos 2D são mais adequados em tarefas de localização e/ou focadas nos componentes espaciais dos dados, sendo também a técnica de visualização preferida dos utilizadores. Por outro lado, STCs são mais adequados em tarefas de associação e tarefas de comparação focadas numa análise espácio-temporal dos dados, ainda que estejam dependentes da capacidade do utilizador em interagir com os mesmos.

Baseados nestes resultados, numa terceira fase procurámos minimizar as limitações detetadas nestas técnicas, recorrendo a um conjunto de alternativas que incluem:

- Utilização de gráficos temporais focados na representação de informação temporal;
- Uso de *pistas espaciais*, no STC;
- Uso de um plano espacial móvel, dentro do STC;
- Uso de uma vista geral em 2D, dentro do STC;
- Combinação de mapas 2D com STCs através da justaposição e/ou transição entre ambas as técnicas.

Através de um conjunto de estudos de usabilidade com estas abordagens, os resultados obtidos permitiram-nos chegar às seguintes conclusões:

- Os utilizadores preferem STCs melhorados através da combinação com outras técnicas (i.e., com plano móvel e/ou vista geral em 2D);
- O desempenho dos utilizadores é superior com STCs melhorados, comparativamente com STCs simples;
- A combinação de ambas as técnicas (por justaposição) é preferível à utilização de qualquer técnica individualmente;

- A possibilidade de transformar uma visualização noutra é uma funcionalidade aceite (e desejada) por utilizadores inexperientes;
- A combinação de ambas as técnicas minimiza o risco de comportamentos, por parte dos utilizadores, que resultem num menor desempenho.

Finalmente, através dos resultados obtidos nos vários estudos relativos às fases anteriores, identificámos um conjunto de *boas-práticas* relativas à escolha e concepção de mapas estáticos 2D e STCs. Globalmente, concluímos que a utilização de mapas estáticos 2D é sempre desejada, ainda que este não seja o foco principal da visualização. Por outro lado, a escolha relativa à utilização e combinação de STCs está dependente de três critérios principais:

- Relevância da informação temporal;
- Capacidade de interação;
- Espaço disponível no ecrã.

Para além disso, sugerimos também um conjunto de funcionalidades a serem incluídas aquando da utilização destas técnicas, que incluem, entre outras, a customização da representação da informação, controlos de granularidade temporal, alteração de projeções, e visibilidade constante de referências temporais.

Palavras-chave: Trajetórias, Visualização de Informação, Cubo espácio-temporal, Mapa 2D, Usabilidade, *Design Guidelines*

Abstract

With the increasing popularity of location based services and mobile tracking technologies, the collection of large amounts of spatio-temporal data became an increasingly common, easier, and more reliable task. In turn, this has emphasized the possibility of analysing georeferenced information, particularly associated with human trajectory data, to identify and understand movement patterns and activities, ultimately, supporting decision making in various contexts.

In order to properly analyse and understand the spatio-temporal and the thematic properties associated with these data, adequate visualization techniques are needed. Due to the spatial properties of trajectories, map-based techniques, such as 2D static maps or 3D space-time cubes (STCs) are considered as essential tools for their visualization. However, despite the increasing number of visualization systems, the study regarding their usability, alongside the role of the human user, sometimes with a limited background in data visualization and analysis, are often neglected. In addition to the somewhat disperse, and sometimes even contradictory, results in the literature, these factors, ultimately, emphasize the lack of knowledge to support the choice of particular visualizations, and their design, in different types of tasks.

This dissertation addresses these issues through three main sets of contributions, focusing on inexperienced users, in terms of data visualization and analysis: i) the characterization of the dis/advantages of existing map-based techniques (2D static maps and STCs), depending on the types of visual analysis tasks and the focus of the analysis; ii) the improvement of existing visualization techniques, either through the inclusion of additional spatial cues within the STC, or combining both types of techniques in various ways; and iii) the identification of design guidelines for trajectory data visualization, describing various considerations/criteria for the selection of different map-based visualization techniques and their possible interactive features.

Keywords: Trajectory data, Information Visualization, Space-time Cube, 2D Map, Usability, Design Guidelines

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the prevalence of mobile computing systems and location based services, large amounts of spatio-temporal data are nowadays being collected, representing the mobility of people performing various activities in different contexts. However, despite the increasing interest in the exploration of these data, there are still several open challenges, related with visualization and human-computer interaction.

To support the extraction of relevant information from the spatio-temporal and the thematic properties associated with human trajectories, it is crucial to develop and study adequate interactive visualization techniques. In addition to the properties of the visualizations themselves, it is important to take into consideration the types of information present within the data and, more importantly, the types of tasks that a user might need to consider to achieve a given goal. The understanding of these factors may, in turn, allow the identification of design guidelines, which should, consequently, simplify the development and assessment of a given interactive visualization.

The Ph.D. research described in this dissertation addressed these challenges. The following chapter presents the main motivations regarding this work and its main research goals, followed by the various contributions that were achieved and, finally, by an overview of this document's structure.

1.1 Motivation

Everything we do is intrinsically related to a location at any given time (Longley et al., 2005). As such, unsurprisingly, many important decisions are directly related with the proper understanding of information based on locations and the movement detected over those locations (Andrienko et al., 2013; Dodge et al., 2008). Some of these decisions may have a significant impact over society and business, like *should a taxi service invest more money to expand its coverage of a certain location*, or *how can a large building be evacuated safely and effectively*, while others may only have interest on a personal level, e.g., *should a person take a certain road to go home* or *which route should a person take to pass through several points of interest*. This has fostered the

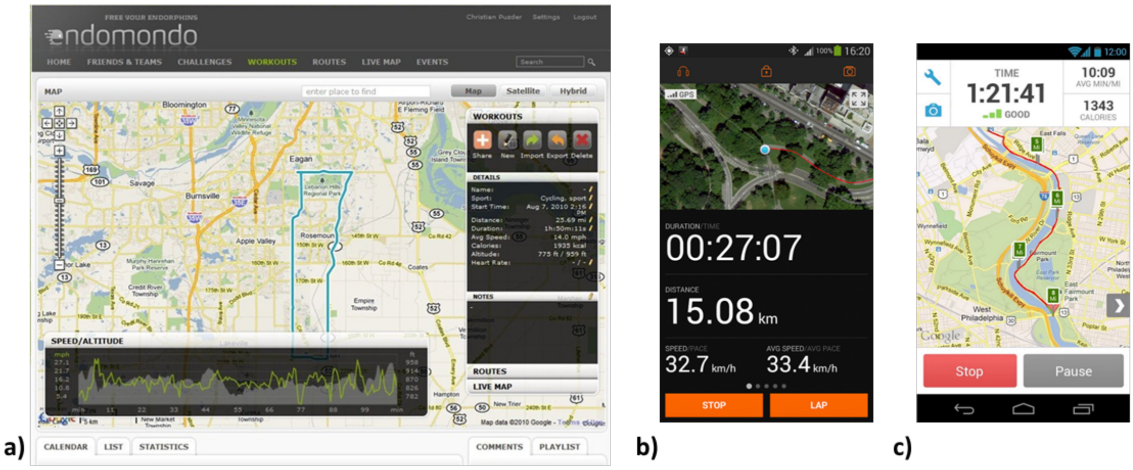

Figure 1.1: Sports applications with trajectory data visualization, namely a) Endomondo b) Sports-Traker c) RunKeeper.

interest of researchers, who have been trying to understand the effects of location-based information, particularly associated with human movement, as a means of improving the lives of users in various contexts (Andrienko et al., 2013). In addition, the increasing popularity and accuracy of mobile technologies has enabled a significantly higher quantity and quality of data. In particular, Global Positioning Systems (GPS), cellular mobile telephony, and Radio-frequency identification tags (RFID) allow people, animals, and objects to be located over long periods of time. This, in turn, enables the creation of spatio-temporal and trajectory datasets (Lee and Krumm, 2011) and, consequently, provide relevant data for researchers and non-researchers to analyse.

From a practical point of view, this interest is nowadays reflected by the increasing number of studies focusing on the usage of spatial trajectories, like GPS taxi data, to assess the state or improve the quality of various urban services (Ding et al., 2013; Gao et al., 2013; Bischoff et al., 2015; Pu et al., 2013). Similarly, social and cellphone network data can also be used to examine spatial trajectories, as these can record their users' locations whenever they perform relevant actions (e.g., log-in, post a message, or take a picture). Consequently, these can, for example, support the study of the effects of social events over the mobility in a city or the correlation of discussion topics with given locations (Naaman, 2011; Girardin et al., 2008; Lorenzo et al., 2013). An even more significant example consists in the increasing use of *mobile exertion* applications, such as *Endomondo* (Endomondo LLC, 2007), *Sports-Traker* (Sports Tracking Technologies Ltd., 2004), or *RunKeeper* (FitnessKeeper, 2009), that allow sports practitioners, often inexperienced in terms of data visualization and domain analysis, to record and visualize their movements in the form of spatial trajectories, as well as information extracted from that data (Figure 1.1).

By definition, a trajectory consists in the evolution of a moving object's spatial properties over time (Parent et al., 2012; Dodge et al., 2008). Typically, these data are represented as a temporally ordered sequence of location points, complemented with thematic information, either derived from the registered data (e.g., speed or transportation mode) or associated from other datasets (e.g., events happening at specific locations involved in the trajectories). Based on this information, it is possible to identify three main data components, namely: space (*where?*), time (*when?*), and attributes (*what?*) (Peuquet, 1994), from which their analysis should allow the extraction of valuable knowledge concerning the actions and events/activities that may occur in the observed world (Andrienko et al., 2010; Neumann, 2005).

Several recent studies have proposed data mining methods to automatically extract valuable information from raw spatio-temporal datasets (Keim et al., 2008; Zheng and Zhou, 2011). One relevant area of research concerning with these is *visual analytics*, which proposes the combination of the advantages of human cognitive and analytical abilities with computer data processing as a means of developing knowledge, methods, and technologies (Andrienko et al., 2013). This combination should give the user full control over the direction of the analysis along the various tasks needed, while reducing the cognitive workload associated with it. In addition, the system is expected to provide adequate tools to display and interact with the information (Andrienko et al., 2010; Keim et al., 2008). Consequently, visual analytics is an integrating discipline that addresses the issues with data processing and emphasizes the fundamental roles of visualization, as a medium to communicate information to the user and human-computer interaction, as the process through which the user communicates with the application and interprets the data (Andrienko et al., 2013; Keim et al., 2008). Furthermore, visual analytics also suggests that automatic methods may not be enough (Andrienko et al., 2013), due to the high complexity often associated with this data, i.e., large volumes of data, often composed by very different types of information, difficult to process and associate.

Considering the spatial characteristics associated with trajectory data, interactive maps are often regarded as important tools for their visualization. Although theoretically it is possible to work out conclusions without maps, it is often unadvised to do so since, without them, it would be extremely difficult to extract relevant geospatial information (Kraak and Ormeling, 2010).

Two-dimensional (2D) static maps are still one of the most used techniques to represent any type of georeferenced information (Figure 1.2 a). Typically, these take advantage of the manipulation of visual variables, like colour or size, associated to different visual marks (e.g, points, lines, and areas) to display various types of information over a geographical plane, such as those present in trajectories (Bertin, 1967). Even with a simple representation, it is still possible to use a map to answer various questions concerning the location being depicted, like the distance between points of interest, the position of locations in respect of each other, or the size of the areas occupied by said points. In turn, these factors have lead this type of technique to be considered fundamental in the representation of the spatial component associated with trajectories. Despite that, 2D static maps still tend to undermine the representation of the data's temporal component, as these usually limit themselves by displaying a single snapshot in time (Kraak and Ormeling, 2010). To solve this problem, there is often an increasing complexity of the visualization, such as when using more intricate symbols over a map to depict the evolution of attributes through time, or when combining the map with additional visualization techniques, such as time graphs (Andrienko and Andrienko, 2013).

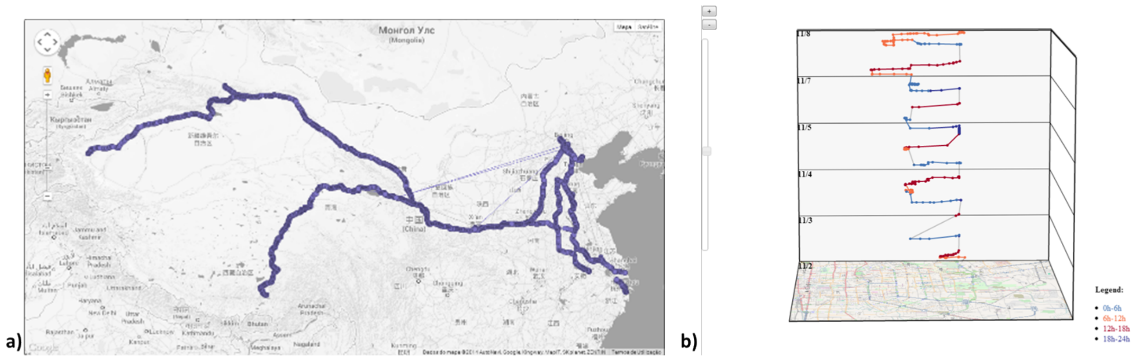

Figure 1.2: Map-based techniques for trajectory data visualization a) 2D static map b) 3D space-time cube.

With the increasing advancements of computer graphics, three dimensional (3D) maps and, in particular, space-time cubes (STCs) (Hägerstrand, 1970) (Figure 1.2 b) have been proposed as viable options for the visualization of trajectories (Kjellin et al., 2010b; Kraak, 2008). Space-time cubes represent both spatial and temporal information within a 3D view, where, similarly to common 2D approaches, the x and y axes are used to represent spatial information (e.g., in the form of latitude/longitude coordinates); however, unlike conventional 3D maps, which use the z-axis to represent information like altitude or depth, STCs use the third dimension to represent time (Hägerstrand, 1970). Similarly to 2D maps, trajectories can be displayed as a sequence of symbols, graphically encoded to represent variations in the thematic attributes. However, since time is represented as a spatial position, other visual variables become available to represent the thematic attributes (Kjellin et al., 2010b). In addition, space-time cubes also allow the representation of various layers of information, each one defined as a plane in the z-axis (Thakur and Hanson, 2010; Tominski et al., 2005, 2012) representing the state of an object in different moments in time. This implies that the higher (or lower, depending on the orientation of the temporal scale) the data is within the STC, the most recent that information is. This allows answering to various questions concerning the temporal aspects of the events detected over a location, like temporal distance between events, their order, frequency, or duration.

Although other types of map visualization techniques could be listed, like animated or small-multiple maps (Kraak and Ormeling, 2010), the analysis of these two categories (2D maps and 3D STCs) suggests already a significant diversity, in terms of data representation. Consequently, the choice for an adequate visualization technique can be seen as an important and, sometimes, a controversial challenge. Despite the increasing number of visualization applications, their empirical validation and the study of their usability have been somewhat neglected (Chen, 2005). In the recent years, this situation has been changing, with the publication of studies providing some type of validation regarding their proposed visualizations. Nevertheless, as suggested by Elmqvist and Yi (2012), some of these studies still follow the approach of *validation through awesome example*, by often limiting themselves on case studies, thus often overlooking the role of the human subject in the design and validation processes (Elmqvist and Yi, 2012; Roth, 2012).

In addition, despite the existence of some studies which address some of the aforementioned problems, several authors have already pointed out their uncertainty regarding the procedures and the contexts used, due to a biased or unbalanced evaluation that, ultimately, benefits certain visualization techniques over others (Tversky et al., 2002; Kristensson et al., 2009; Lobben, 2008; Robertson et al., 2008).

In this work, it is advocated that the current lack of user studies validating the usefulness of the proposed techniques, alongside the possible uncertainty associated to the already existing studies, are important issues. Despite the existence of useful design rules in the areas related to this work, such as cartography, there has been a limited research on the generation of geographic insight based on the interaction with map-based applications, also known as cartographic interaction (Roth, 2013). Moreover, the available empirical evidence that may help minimizing these issues is still considered too fragmentary and difficult to generalize (Andrienko et al., 2010; Roth, 2013). Consequently, alongside the existent uncertainty in terms of the results of previous studies, these factors further emphasize the current lack of actual knowledge regarding the adequacy and usability of existing map-based techniques for the acquisition of spatio-temporal insights, in the different types of tasks. These also suggest the need to empirically assess and compare existing types of (map) visualization techniques. However, more than just understanding the goals a user may have when interacting with a certain visualization, it is fundamental to take into consideration the types of visual analysis tasks that need to be addressed in order to achieve those goals (Valiati et al., 2006).

In recent years, there has been an increasing interest in showing complex spatio-temporal information to a wider population. This interest was first noticed with the growing popularity of mapping APIs (Application Programming Interfaces), like Google Maps, or geotagging photo features, that supported the use, and sometimes creation, of maps and geographical information by individuals not necessarily interested in geography as a science (Flanagin and Metzger, 2008). Nowadays, this interest is reflected and expanded by the increasing popularity of performance monitoring tools, both in the analysis of spatio-temporal *real life* and *virtual world* scenarios. Nevertheless, it is important to acknowledge the fact that the data presented to these (new) users, even if not necessarily related to their professional lives, may still be relevant on a personal context. More importantly, it should also be noticed that, given the popularity and availability of these data and applications, a significant portion of these users are inexperienced in terms of data analysis or visualization (Huang et al., 2015; Drachen and Schubert, 2013b). This fact is particularly important since it suggests that any user, regardless of their background and experience, may be put in the role of a *spatio-temporal analyst* (Andrienko et al., 2010) and, therefore, is a potential user of techniques for the visualization of spatio-temporal data, even if in less complex scenarios. This shows that, more than just conducting empirical evaluations of existing map techniques, it is also important to take inexperienced users into account when conducting those evaluations.

1.2 Research Goals

Existing literature in the relevant domains related to this research reveals some noticeable challenges associated to the areas of visualization and human-computer interaction. More specifically, despite the observed growth in the number of visualization techniques for spatio-temporal and trajectory data, there is still a significant lack of empirical support regarding the advantages and disadvantages of the existing techniques. This lack of support is further emphasized by some uncertainty associated to previous studies, which can be seen as a result of an unbalanced evaluation, often resulting in a visualization being favoured over other(s). Alternatively, this uncertainty comes also from a significant paradigm change in modern spatio-temporal data visualization and analysis, with a stronger focus on the role played by the computer/technology and an expansion of the community of potential users, usually lacking experience in terms of data visualization and analysis.

The Ph.D. research here described addressed these challenges. However, considering the wide range of possibilities provided by these issues, this work focused on the evaluation of the main types of map techniques for the visualization of trajectory data by inexperienced users, in the role of spatio-temporal data analysts, even if in a simplified context. The following sections describe the main research objectives in more detail and a summary of the conducted research process.

1.2.1 Obtain Empirical Data to Support the Dis/advantages of a Given Technique

As previously mentioned, there exists a significant lack of empirical data regarding the adequacy of existing visualization techniques. Therefore, this work proposes the conduction of usability studies for the comparison of the main groups of visualization techniques. Out of the four groups of techniques mentioned above (2D maps, 3D STCs, animated maps, and small-multiples), the focus of this research is on the comparison of 2D maps and 3D STCs. This decision will be further discussed in the next chapter. In addition, although these techniques can be used for the visualization of various types of data, this research focuses on the usage of these techniques for the visualization of movement data describing human trajectories.

It is important to notice that several approaches for the representation of thematic attributes are independent from the technique or dataset used. For instance, colours can be used to represent a certain set of categories, regardless if they are used on a 2D or 3D display; similarly, said categories can be associated with the types of transportation, the mover's identification, or any type of thematic data organizable into categories. Consequently, while acknowledging the importance of thematic information in the representation and analysis of trajectory data, this research focused mostly on the simultaneous representation of spatial and temporal information in 2D maps and 3D STCs.

Furthermore, to properly conduct the proposed usability studies, it was fundamental to also analyse the types of visual analysis tasks a user may need to perform to achieve a given goal.

Although, several authors have already proposed various taxonomies of tasks (Wehrend and Lewis, 1990; Vasiliev, 1997; Roth, 2012), there exists a noticeable redundancy, which must be clarified prior the conduction of any usability study.

1.2.2 Improve Trajectory Data Visualization Techniques

Based on the properties of the different groups of map visualization techniques, alongside the results of previous studies, it is expected that certain techniques outperform others in specific types of tasks. However, it is unlikely that the analysis of a given dataset will be always focused in just one type of task. Therefore, while the previous research goal aims to help identifying which scenarios are more adequate for certain visualization techniques, it is still important to understand how to improve the capabilities of a given visualization technique to be usable in several types of tasks. For that, following up the comparative evaluation of 2D maps and 3D prototypes, this work proposes the assessment of possible improvements over these techniques and the analysis of their adequacy and acceptance by an inexperienced population.

1.2.3 Identify Design Guidelines for Map-based Trajectory Data Visualization

Besides the problems that were already mentioned in the previous sections, it is noticeable a current lack of design guidelines regarding cartographic interaction, and thus, trajectory data visualization. Designers working with spatio-temporal data visualization have thus to face various questions with uncertain and unsolved answers, including the types of visualization techniques to be used and which features these should support. For that reason, another goal of this work consists in identifying a general set of design guidelines regarding interactive maps for analysing trajectory data.

Although it is important to acknowledge the importance of the results of previous related studies, alongside the knowledge obtained during the development of prototypes, the goal of this Ph.D is focused on the results obtained through the usability studies conducted. This implies the analysis not only of the users' performance, in terms of efficiency and effectiveness, and their comments during and after completing the tasks (i.e., subjective preferences), but also of an additional set of metrics that, alongside the direct observation of the users' actions, support the analysis of the strategies applied by them during the tasks, when interacting with the different components from the proposed map-based visualizations.

Based on the aforementioned objectives, five steps were identified describing the process of this work. These are depicted in Figure 1.3. This figure will be revisited at the end of each chapter, in which new information will be added, when appropriate, providing a schematic summary of the contributions presented in the chapter, and its role in the scope of this work.

The first step implies the collection and analysis of previous works existing in the literature, which, as described in the next chapter will also result in the identification of a simplified vi-

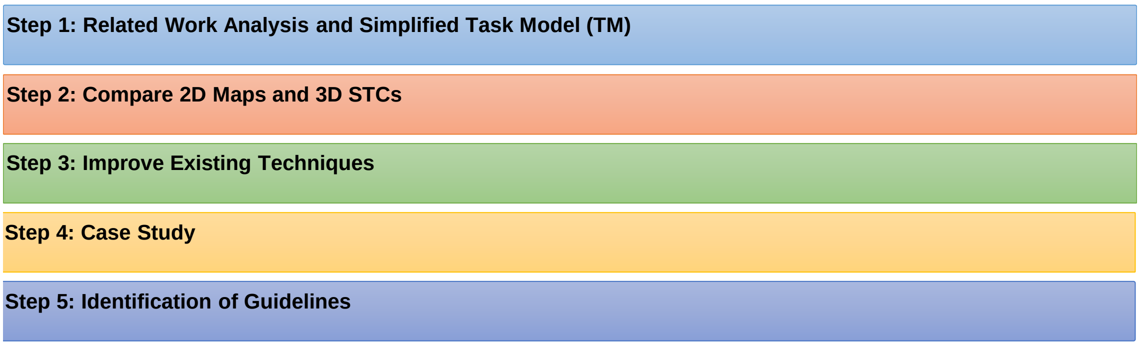

Figure 1.3: Summary of the various steps of this research.

sual analysis task model. Steps 2 and 3 focus on the comparison and improvement of existing techniques, particularly 2D maps and 3D STCs, and both will be composed by comparative user studies, further described in Chapters 3, 4, and 5. As listed in step 4 and discussed in Chapter 6, it was also considered necessary the analysis of a case study. In particular, the information obtained in the previous steps is applied into the area of Visual Game Analytics, where a prototype, VisLol, is described and compared to an already existing application. Finally, based on the information acquired throughout the various steps, the most noticeable patterns are identified, in Chapter 7, as a general set of design guidelines.

1.3 Contributions

The contributions of this Ph.D. research were divided into four main categories. Scientific contributions consist of the most significant findings obtained through the various user studies, alongside the direct outcomes from the analysis of the literature, the process and the results related with and obtained in the context of this research. Software contributions include all tools and prototypes developed to support the conducted experiments and achieve the research goals. Publications list all published documents in scientific conferences and journals, associated to this Ph.D. research. Finally, legacy refers to subsequent projects inspired by this research.

1.3.1 Scientific Contributions

Throughout the experiments conducted in the context of this Ph.D., it was possible to retrieve evidences showing and suggesting that:

- Inexperienced users, in terms of the application domain and visual analysis, are able to interact and work out conclusions with 2D maps and 3D STCs, despite their visual complexity and lack of familiarity. This is addressed with more detail throughout the discussion sections of the next chapters describing the usability studies.
- In their general form, 2D static maps are more adequate, in terms of efficiency and effectiveness, than STCs, in tasks focused on the data's spatial components. This is described

with more detail in the second usability study of Chapter 3, in which the users took less time to complete tasks and were generally more accurate when using the 2D static map prototype, in *location* tasks.

- On the other hand, STCs are more adequate, in terms of effectiveness, than 2D static maps, in tasks also dealing with the data's temporal components and tasks requiring the analysis of patters/relationships between data elements. This is primarily addressed in the same chapter as the previous point, where users were significantly more accurate in *associate* tasks when using the STC prototype.

- Despite acknowledging or praising the STC's role in the prototypes, users prefer 2D static maps as their main visualization component. In four of the experiments that were conducted (from Chapters 3 to 6), users have rated 2D static map prototypes higher than STC prototypes, according to their preferences the users have also attempted to convert the STC prototypes into 2D static maps, and they used mostly the 2D static map, when both techniques were available (particularly in tasks dealing with spatial information).

- Increasing the volumes of data to be displayed affects negatively the interaction with both techniques. While 2D maps are more affected in temporal-based tasks, STCs are more affected in location-based tasks. This factor is addressed in more detail in Chapters 3 to 5. In these experiments, increasing the amount of data displayed during the usability studies caused users to take more time and be less accurate with the prototypes. When just one trajectory was displayed, no significant differences were detected, in terms of performance.

- The use of temporally focused representations, such as timelines, is an effective compensation to the lack of temporal representation in the 2D map and is considered as essential by inexperienced users. It can also be useful when employed alongside an STC, although less significantly. This is addressed with more detail in Chapter 5. The analysis of the users' actions and comments related to the experiments that are there described revealed that they interacted significantly more with timelines in temporal tasks, when used alongside 2D static maps.

- Using additional spatial context cues over STCs, in the form of moveable map planes alongside the STC's height or in the form of a 2D data overview, significantly improves the usability of STCs in location-based tasks, closing the gap between STCs and 2D static maps in terms of effectiveness. This is addressed in the usability studies described in Chapters 4 and 5. In the first, the results have shown that both proposals for improving the STC provided better results than a baseline STC. In the second, the results have shown that users were not significantly more accurate when using the 2D static map comparatively to an improved STC.

- Despite the increasing visual and interactive complexity, combining 2D static maps and 3D STCs is a viable approach for inexperienced users to visualize trajectory data. This is addressed in Chapters 4, 5, and 6. The first discusses the attempts of users to convert one technique into the other to complete the tasks. The second shows how the use of a 2D map overview supports a better performance, comparatively to a baseline STC. The third shows how combining 2D static maps and 3D STCs is not significantly outperformed by other simpler/isolated techniques, and how this is seen as an overall preferable approach by the study participants.

Based on the analysis of previous works and also based on the combined results obtained in the conducted experiments, the following high-level theoretical contributions were produced:

- A systematic survey was developed with relevant concepts addressing: i) the main characteristics of trajectory data; ii) visualization/interaction tasks used to achieve a given goal; and iii) the main types of map techniques for trajectory data visualization. Through this analysis, a simplified hierarchical task model was developed, that once combined with the other addressed components, can be applied in the evaluation of visualization techniques.

- A systematic set of usability tests was also developed comparing different approaches for visualizing trajectory data with 2D static maps and 3D space-time cubes. The results were analysed based on the record of user action logs, common performance metrics (e.g., effectiveness, efficiency and preferences) and additional metrics including: number of actions, ratio of visualization component used or 3D mitigation actions/visualization type transitions.

- A set of general guidelines was produced regarding the design of 2D maps and 3D space-time cubes for trajectory data visualization, focused on an inexperienced user population.

1.3.2 Software Contributions

To conduct the usability studies referenced in the previous sections, some software development work was necessary. Consequently, the following tools were created:

- *STCJS* (Gonçalves, 2017a): API (Application Programming Interface) developed in HTML5, over the THREE.JS library (Cabello, 2010), that supports the creation and interaction with visualizations based on space-time cubes for the analysis of trajectory data.

- *TRAJMAP2DJS* (Gonçalves, 2017c): API developed in HTML5, over the THREE.JS library, that supports the creation and interaction with an information layer built with Google Maps, for the visualization of trajectory data.

- *TIMELINEJS* (Gonçalves, 2017b): API developed in HTML5, over the THREE.JS library, that supports the creation and interaction with timelines for the visualization of data from a temporal perspective.

- Various prototypes using the aforementioned APIs, including:

ST-TrajVis: Application displaying a 2D map and a STC for the visualization of a user's trajectories - this prototype is described in more detail in Chapter 3;

HTV-2DMap and *HTV-STC* (HTV - Human Trajectory Visualizer): Applications implementing a basic 2D map and STC, respectively, alongside a basic timeline, for the visualization of human (walking) trajectories from multiple subjects - these prototypes are described with greater detail in Chapter 3, under the names of 2D Map and STC, respectively;

ImpSTC-P1, *ImpSTC-P2*, *ImpSTC-P3*: Applications implementing three variations of the STC, the first with general improvements over *HTV-STC*, the other two implementing additional features to improve STCs in location-based tasks, namely additional moveable maps and map overviews - these prototypes are described in more detail in Chapter 4, under the names of P1, P2, and P3 respectively.

TTV-2DM, *TTV-STC*, *TTV-3VC* (Taxi Trajectory Visualizer): Applications implementing a 2D map, an STC, and a combination of both techniques, respectively, alongside timeline representations, for the visualization of taxi data - these prototypes are described with greater detail in Chapter 5, under the names of 2DM, STC, and 3VC respectively.

VisLoL: Application implementing a combination of 2D maps and STCs, alongside timeline representations, for the visualization of the trajectories and events associated to a set of players in a *League of Legends* match - this prototype is described in more detail in Chapter 6.

1.3.3 Publications

This Ph.D. research originated the following publications, grouped by publication year:

2017

Pedro Vieira, Tiago Gonçalves, Ana Paula Afonso, and Maria Beatriz Carmo. Mapas Animados para a Análise do Desempenho Pessoal em Jogos. Encontro Português de Computação Gráfica e Interação, EPCGI 2017, pgs. 12 (accepted for publication).

2016

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins, Why not both? - Combining 2D maps and 3D pace-time cubes for human trajectory data visualization, Proceedings of 30th British Computer Society Conference on Human-Computer Interaction Conference, HCI 2016, pgs. 10.

Tiago Gonçalves, Ana Paula Afonso, António Ferreira, Ana Rita Vieira. Trajectory Data Visualization on Mobile Devices with Animated Maps, Proceedings of the 37th Annual Conference of the European Association for Computer Graphics, EUROGRAPHICS 2016, pgs. 4.

2015

Tiago Gonçalves, Ana Rita Vieira, Ana Paula Afonso, António Ferreira. PATH - Visualização de Percursos Pessoais com Mapas Animados em Dispositivos Móveis. INFORUM - Simpósio de Informática 2015, pgs. 14.

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Cartographic Visualization of Human Trajectory Data: Overview and Analysis. Journal of Location Based Services, 9(2):138-166, 2015, pgs. 28.

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Combining and Assessing 2D Maps and Space-Time Cubes for Trajectory Data. British Computer Society Conference on Human-Computer Interaction, BCS HCI 2015, pgs. 2.

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Improving Spatial Awareness for Human Trajectory Visualization in Space-Time Cubes. Human-Computer Interaction - INTERACT 2015 - 15th IFIP TC 13 International Conference, pgs. 6.

2014

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Visualizing Human Trajectories: Comparing Space-Time Cubes and Static Maps. Proceedings of the 2014 British Computer Society Conference on Human-Computer Interaction, BCS HCI 2014, pgs. 6.

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Visualization Techniques of Trajectory Data: Challenges and Limitations. Proceedings of the 2nd AGILE Ph.D. School 2013 CEUR Workshop, 1136, pgs. 10.

2013

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins, Daniel Gonçalves. ST-TrajVis: Interacting with Trajectory Data. Proceedings of the 2013 British Computer Society Conference on Human-Computer Interaction, BCS HCI 2013, pgs. 6.

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins, Daniel Gonçalves, Visual Exploration of Human Mobility Datasets, Mobile Ghent 2013, pgs. 1 (extended abstract).

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins, Visual Analysis of Mobility Data. International Conference on Mobile Data Management 2013, MDM 2013, pgs 4.

In addition to the aforementioned list, two additional contributions were published during the course of this Ph.D. work, namely:

Tiago Gonçalves, Ana Paula Afonso, Maria Beatriz Carmo, Paulo Pombinho. Comparison of Off-screen Visualization Techniques with Representation of Relevance on Mobile Devices. Proceedings of the 2013 British Computer Society Conference on Human-Computer Interaction, pgs. 8.

Tiago Gonçalves, Ana Paula Afonso, Maria Beatriz Carmo, Paulo Pombinho, Overview vs Detail on mobile devices: a struggle for screen space. Proceedings of the 2012 British Computer Society Conference on Human-Computer Interaction, BCS HCI 2012, pgs. 6.

Furthermore, at the time of this document's writing, three additional papers, focused on the most recent results and future work, await revision for a conference and two different scientific journals, namely:

Tiago Gonçalves, Ana Paula Afonso, Bruno Martins. Cartographic Interaction with 2D Maps and 3D STCs for Trajectory Data Visualization, pgs. 20.

Tiago Gonçalves, Paulo Pombinho, Nuno Carreiro, Ana Paula Afonso, Maria Beatriz Carmo. VisLoL - The role of interactive maps in the performance assessment of MOBA players, pgs. 12.

1.3.4 Legacy

In addition to the aforementioned contributions, this work has also inspired the creation of related projects, also carried out at the University of Lisbon. In particular, three MSc thesis have benefited from the tools that were developed, and from the results obtained through this research. More specifically, these projects had the following contributions:

Visualização de Trajectórias Humanas em Dispositivos Móveis (Visualization of Human Trajectories on Mobile Devices (Vieira, 2016)) - this MSc thesis addressed the usefulness and usability of map-based techniques for trajectory visualization by inexperienced users, on a mobile environment (e.g., using smarthphones or tablets), and was inspired by the growing popularity of tracking applications. In particular, this work focused on the use of (simple) animated maps and compared two possible approaches in animated representations with a, more common, 2D static map (Gonçalves et al., 2016a, 2015d). This research was conducted by Ana Rita Vieira and supervised by Prof. Dr. Ana Paula Afonso, by Prof. Dr. António Ferreira, and by the author of this document, attaining a final grade of 16 out of 20.

Técnicas de Visualização para Melhorar o Desempenho em Jogos Online (Visualization Techniques for Performance Improvement in Online Games (Carreiro, 2016)) - this MSc thesis was inspired by the growing popularity of competitive online video games, such as MOBAs (Massive Online Battle Arena) games, which are often dependant of a proper understanding of the games' spatial and temporal related attributes, as well as the increasingly common collection of game-related spatio-temporal data, associated with the players. More specifically, this work addressed the adequacy and the effects that map-based spatio-temporal visualizations, such as 2D static maps and 3D space-time cubes, have over players' understanding of the events recorded in a game. This work also addressed the issue of inexperienced users, in terms of visualization and analysis, interacting with trajectory data. For that, it compared 2D static maps and 3D STCs with existing non map-based approaches, and it took into account the users' familiarity with the data and its effects over the analysis (Carreiro, 2016). This research was conducted by Nuno Carreiro and supervised by Prof. Dr. Ana Paula Afonso, Prof. Dr. Maria Beatriz Carmo, and by the author of this document, attaining a final grade of 17 out of 20.

Visualization of Spatio-temporal Information for Personal Performance Analysis in Games - this MSc thesis consists of a direct follow up of this Ph.D. research, aiming to address some of the challenges identified as future work in the previous MSc thesis, which include the study of the effects of domain knowledge and personal motivations over the data analysis, while still studying techniques for spatio-temporal and trajectory data visualization dedicated to inexperienced users. This research is currently being conducted by Pedro Vieira and is being supervised by Prof. Dr. Ana Paula Afonso and by the author of this document.

1.4 Document Structure

The remainder of this document is organized as follows:

Chapter 2 provides a review and detailed analysis of relevant existing work in the literature. In particular, this chapter focuses on the main characteristics associated with trajectory data, the main types of visual analysis tasks required to achieve a given goal, and it overviews the main types of map-based techniques available for the visualization of trajectory data. After that, it focuses on the most relevant comparative studies conducted with the various techniques. This chapter also presents the first contribution of this research, namely a simplified hierarchical organization of visual analysis tasks, which was used as the basis of the user studies described in the subsequent chapters.

Chapter 3 presents two usability studies with three prototypes for the visualization of human trajectory data, i.e., St-TrajVis, HTV-2DMap, and HTV-STC. The first study, conducted with ST-TrajVis, helped analysing the reception of inexperienced users towards map-based techniques for the visualization of trajectory data. The second study, conducted with the remaining prototypes, focuses on the comparison of the main characteristics of 2D static maps and 3D STCs.

Chapter 4 presents another comparative study, with three prototypes implementing variants of the STC technique, namely ImpSTC-P1, ImpSTC-P2, and ImpSTC-P3. These were developed based on the results presented in the previous chapter, and focused on the analysis and validation of possible improvements over the STC technique for *location*-based tasks.

Chapter 5 builds upon the results of the previous experiments and addresses the issue of combining 2D static maps and 3D STCs for trajectory visualization. For that, three prototypes were created, i.e., TTV-2DM, TTV-STC, and TTV-3VC, and evaluated in a comparative user study. In addition to the analysis of the usefulness of combining both techniques, this study helped further comparing 2D static maps and 3D STCs.

Chapter 6 presents VisLoL, a prototype developed in the context of a case study that takes advantage of the problems addressed in the three previous three chapters, applied to the area of video game analytics. More specifically, it presents a user study comparing the interaction and analysis made with a map-based visualization, combining 2D maps and 3D STCs, with an application not based on maps. This study also addresses the problem of the effects that the familiarity with the dataset may have over the analysis, by an inexperienced user.

Chapter 7 reviews and discusses the results and contributions addressed in the previous chapters. Based on the results of that complete analysis, this chapter identifies a set of design guidelines for trajectory data visualizations, with a focus on an inexperienced population.

Chapter 8 finalizes this document by describing the main conclusions drawn from the various steps of the Ph.D. research, presenting also the most promising and challenging research opportunities for future work.

Chapter 2

Interactive Visualization of Trajectory Data

This chapter fulfils two main roles. First the chapter presents the most relevant concepts associated to this Ph.D. research, alongside the related work in the literature responsible or dedicated to them. Then, to avoid unnecessary repetitions, the chapter presents the first contribution of this research, namely a simplified hierarchical visual analysis task model, based on the study and systematization of the aforementioned concepts. For that, this chapter is divided into six main sections. The first provides an overview regarding the concept of *cartographic interaction*, and establishes the context for the rest of the chapter. The next three sections present and analyse a relevant set of concepts and the related work based on three main questions (Aigner et al., 2007), respectively:

What is presented: focused on the *main components of trajectory data*.

Why is it presented: focused on the types of *user objectives* and tasks that can be conducted when using a given visualization, and followed by the proposal of a simplified view over these types.

How is it presented: focused on the characterization of the most general *types of map visualization techniques* to represent trajectory data, based on their properties and expected dis/advantages.

Considering this research's proposal for the conduction of comparative user studies of visualization techniques, the fifth section describes and discusses some of the most relevant comparative studies present in the literature. The chapter ends with a summary of the various points mentioned above.

2.1 Cartographic Interaction

To extract useful information from human trajectory datasets, adequate representations are needed. As a consequence of the inherent spatial properties of spatio-temporal data and trajectories, maps are often considered important tools for the representation and analysis of these data (Kraak and

Ormeling, 2010). However, due to the large volumes of data and the complexity generally associated with mobility datasets, it is often difficult to produce a single and sufficient representation of the data. Therefore, to enable a better understanding of this data, it is crucial that these representations support interactive features that allow users to visualize the data, partially or entirely, from different perspectives (Roth, 2013). This is directly associated to the concept of cartographic interaction.

Overall, cartographic interaction is defined as the dialogue between a human user and a map mediated through a computing device (Roth, 2013). This, in turn, implies the existence of three main interconnected components, namely:

The user: usually responsible for the beginning of the interaction. The user may select different map views depending on the visualization goals, usually implying some kind of change over the map;

The map: presents information to the user, based on his/her actions with the computing device. The current map may be changed depending on how effectively this helps the user achieving his/her goal;

The computing device: responsible for the translation of the user's actions/requests into a map view, and the display of the same view back to the user.

For a better understanding of these components, the following paragraph presents a scenario related to cartographic interaction with trajectory data:

A certain user recorded his/her trajectories, e.g., during jogging exercises, and wants to analyse them to estimate where to go next time, in order to improve his/her performance. The user visualizes the data in an application that displays the trajectories on a map, also allowing the visualization of that data according to several thematic attributes (e.g., speed). The user starts by loading the data into the application and ***locate*** *the trajectories in the map. Then, the user requests the application to display the recorded speed taken along the trajectory. The user may explore the map, e.g., through panning and zooming operations, and may be able to* ***identify*** *how fast he/she moved over certain locations. The user may then* ***compare*** *those same locations or the instants in time during the exercise, in which he/she moved faster or slower. Based on the analysis of the map and the path taken, the user may be able to* ***associate*** *his/her speed to the type of path/street taken, and conclude, for example, that he/she was slower in busier parts of the city, comparing to local parks or the outskirts of a city. Finally, based on the collected information, the user concludes which areas of the city he/she will want to avoid next time.*

Roth (2013) described the procedure associated with cartographic interaction based on the *stages of action model*, originally proposed by Norman (2002), and which can also be used to describe the aforementioned scenario. Overall, this consists of a cyclic model composed by seven stages, namely:

Forming the goal: requires developing a notion of what needs to be done. Typically, a *goal* is poorly defined and domain specific (Roth, 2012). In the presented scenario, this stage corresponds to the user deciding that he/she wants to analyse his/her data and the locations in which he/she may practice exercise with a higher performance.

Forming the intention: requires determining the tasks/objectives needed to achieve a certain goal. These objectives can be defined as the user's cognitive input for the cartographic interaction (Roth, 2012). In the presented scenario, this stage corresponds to the definition of the tasks *locate*, *identify*, *compare*, and *associate*, which helped the user in the various stages of the analysis.

Specifying the action: requires the determination of the actions that might be needed to perform a certain task. Typically, these actions involve using the interactive features provided by the computing device to request/change the state of the map view. In the presented scenario, this stage corresponds to the user's decision of interacting with the map, e.g., through panning or zooming the map.

Executing the action: consists in the user's input to the computing device, following the action(s) defined in the previous stage. In the presented scenario, this stage corresponds to the user's request for information about movement speed, or when panning or zooming the map;

Perceiving the state of the world: consists in the user's acknowledgement of a new state for the map view, as a result of the actions performed in the previous stage.

Interpreting the stage of the world: consists in the user's process to make sense of the data displayed. In the presented scenario, this stage corresponds to the user's translation of what is shown in the visualization into usable information, including the understanding of how fast was the movement at certain places, or determining in which locations a slower movement was recorded.

Evaluating the outcome: consists in the user's decision on whether the information acquired is enough to accomplish the goal created at the first step of the interaction. The outcome of this evaluation may result in the end of the interaction, in the beginning of a new cycle with the same goal (if the goal was not achieved), or in the beginning of a new cycle with a new goal. In the presented example, this stage corresponds to the user's conclusion that the map allowed him/her to decide which areas of the city should be avoided, to improve performance, thus ending the interaction with the map-application.

As previously mentioned, to properly develop and evaluate adequate visualization techniques for trajectory data, it is important to take into account the capabilities of the users, particularly when these are likely to be inexperienced, in terms of visualization and analysis. In addition, it is also important to consider the various constraints related to the aforementioned stages. In particular, although user goals may be poorly defined and domain specific (Roth, 2012), making it difficult to properly study them, the types of visual analysis tasks that might need to be completed to achieve a certain goal are relatively well known, and have been the focus of several studies.

Throughout the literature, several taxonomies of *interaction primitives* have been proposed, focused on the tasks, types of interactive actions, and data properties, associated with cartographic interaction (Roth, 2012). The following sections of this chapter describe and analyse the most relevant concepts and taxonomies associated to these topics.

2.2 Main Components of Trajectory Data

By definition, a trajectory consists in the evolution of a moving object's spatial properties over time (Dodge et al., 2008), often characterized by a set of *stop* and *move* features, depending whether the object is stopped or in movement (Spaccapietra et al., 2008).

Typically, trajectories are represented as a temporally ordered sequence of location points $P = \langle x_i, y_i, z_i, t_i \rangle$, that compose a trajectory $T = \{P_1, P_2, ...P_n\}$, where x_i, y_i, and z_i represent spatial coordinates (e.g., latitude, longitude, and altitude), and t_i represents the time instant when the spatial position was recorded. Furthermore, each trajectory point or the trajectory as a whole can be further enriched with *annotations* or *thematic attributes*. These may represent any type of information that adds knowledge to the trajectory and can provide a more detailed characterization of the moving object (Parent et al., 2012). More specifically, these attributes may include both quantitative and qualitative information, derived from computations over the original raw data (e.g., speed values or distance travelled), from the combination with other datasets (e.g., the weather at given location or time), or inferred from the existing data (e.g., transportation mode based on the speed during movement) (Schuessler and Axhausen, 2009; Xie et al., 2009; Yuan et al., 2012).

Usually, trajectories involve large volumes of data, therefore making it difficult to constantly keep and track records of movement data (Parent et al., 2012). In addition, several external factors, such as satellite triangulation issues or signal blocking, may create erroneous dataset entries. Consequently, prior to the visualization and analysis of these data, several pre-processing methods tend to be used to remove erroneous data entries (*data cleaning*), or reduce their size, either based on a set of criteria (*data filtering*) or by merging/transforming entries with similar properties (*data smoothing* and *aggregation/clustering*) (Deng et al., 2011; Heckbert and Garland, 1997; Wang et al., 2009).

According to various studies in the literature, the application of these methods can support the extraction of additional information, to better understand trajectories (Renso et al., 2012; Schuessler and Axhausen, 2009; Xie et al., 2009; Andrienko and Andrienko, 2013). Consequently, it is important to acknowledge the importance of pre-processing stages, prior to the visualization of the data. Nevertheless, it is also important to mention that these methods are beyond the scope and goals of this work, and thus this section will not enter in further details.

Ultimately, a trajectory consists in the path generated by an object over an area during a certain amount of time. Consequently, and regardless if the pre-process of trajectories requires the removal of erroneous locations or the inclusion of additional thematic attributes, it is possible to identify three fundamental sets of information directly associated to these data, namely space (*where?*), time (*when?*), and objects (*what?*) (Peuquet, 1994; Andrienko and Andrienko, 2005; Neumann, 2005; Qian et al., 1997).

2.2.1 Space

Although in the majority of the literature the term *space* is usually associated with the Earth's surface and its geographical features, it can refer to several non-geographical contexts (Longley et al., 2005; Drachen and Schubert, 2013b). In fact, space is a complex type of information that can be structured based on several different properties, like coordinate-systems or hierarchic relations between locations. Several types of geometric structures can be used to represent space, including points, lines, areas, and volumes (Bertin, 1967; Andrienko and Andrienko, 2013; Kraak and Ormeling, 2010). This, allows different spatial characteristics to be emphasized, such as distance between points/locations, altitude, slope of the terrain, or category of location, e.g., home, work, entertainment or restaurant (Andrienko and Andrienko, 2013).

Depending on the context, space can be treated as a one-, two-, or three-dimensional property of a given trajectory (Andrienko and Andrienko, 2013). Typically, when the understanding of the distances between location points or their identification is sufficient, representing space as a one-dimensional property may be enough. However, when it is important to understand how the position of an object changes over time, or understand the possible relations between locations, two- or three- dimensional representations of the data are more appropriate, which may influence the choice of a visualization technique for that data.

In the context of this research, both geographical and non-geographical representations of space were used throughout the various experiments. These include datasets containing (two-dimensional) trajectories recorded from the following locations:

China (geographical) - data made available through the GeoLife Project. This dataset will be further explained in the next chapter;

New York (geographical) - data made available by the Taxi and Limousine Commission of New York City. This dataset will be further explained in Chapter 4 and 5;

League of Legends (non-geographical) - data made available by Riot Games, through their Riot Games API. This dataset will be further explained in Chapter 6.

2.2.2 Time

Time is a complex type of data, which, similarly to spatial locations, can be described based on several criteria, including scale, scope, arrangement, and viewpoint (Andrienko et al., 2010; Aigner et al., 2007; Kaufman, 2004).

In terms of temporal scale, time can be considered as ordinal, discrete, or continuous. An ordinal scale implies relative order of relations between events (Aigner et al., 2007). A discrete scale is usually associated to a complex hierarchical system of granularities that include seconds, minutes, hours, or days (Andrienko et al., 2010). With a continuous scale, time can be mapped into numeric values (Aigner et al., 2007).

It is also important to analyse the duration or focus of temporal data. Typically, two dimensions are considered for this criterion, namely, time-points or moments and time intervals, depending if the temporal extent defining the data is greater than zero (Vasiliev, 1997; Aigner et al., 2007; Andrienko et al., 2010).

Temporal information also deals with cycles and re-occurrences that need to be taken into consideration, regardless of their frequency (Vasiliev, 1997; Frank, 1998; Neumann, 2005; Andrienko et al., 2010). As such, temporal information can be arranged into two categories, namely linear and cyclic time, depending if time is perceived as a continuous sequence or a set of re-occurring events, respectively.

Finally, the viewpoint of temporal information can be subdivided into ordered time, branching time, and multiple perspectives. Ordered time considers that events occur one after the other (Aigner et al., 2007). Branching time is described as a metaphor to assist in the analysis of alternative scenarios, and time with multiple perspectives considers the representation of more than one point of view of an observed fact (Andrienko et al., 2010).

In the context of this research, the datasets used in the experiments can be described by this criteria. Considering the focus on trajectory data recorded from real people (either in geographical or virtual locations), it is important to take into consideration the order in which the various actions took place (e.g., which location was visited, or which events happened, first?). As such, although time can be represented in a continuous scale (e.g., timestamps), its analysis is made through an ordered viewpoint. In addition, both time-points and periods can be considered when analysing the data. For example, while the arrival of a person to a specific location occurs in a time-point, the trip towards that location happens throughout a period of time. Moreover, although time related to a trajectory is generally seen as a (linear) continuous sequence of events, it is possible that some of those re-appear over time (e.g., a person visiting the same locations in various periods of time, such as going to work or home every day).

2.2.3 Objects

In addition to the properties associated with space and time, it is also important to categorize the objects/entities that create or are related to the trajectories. Objects are usually characterized by various properties, describing non-location and non-temporal information (Qian et al., 1997), like age, health condition, or mode of transportation. Naturally, the analysis of these attributes may have an effect over the movement itself and, consequently, over how the user may understand the data through a given visualization (Peuquet, 1994; Andrienko et al., 2008; Roth, 2012).

In the same way thematic properties can help understanding trajectories, objects can also be characterized in function of their relation with space and time. In particular, Andrienko et al. (2013) classified objects into five general categories, namely:

Spatial objects**:** that can be characterized as having a location in space;

Events**:** that can be characterized as having a location in time;

Spatial events**:** that can be defined according to spatial and temporal positions;

Mover**:** a type of spatial object where its position changes over time;

Moving events**:** a type of event where its position changes over time.

As previously mentioned, a trajectory is defined as the evolution of the position of an object over time. In addition, trajectories can be further characterized as a set of *stop* and *move* features, depending on whether the object is stopped or in movement (Spaccapietra et al., 2008). Comparing to the aforementioned categories, it is possible to see that these concepts are directly associated to those of *movers* and *spatial events*, which are, in fact, often considered to be the most relevant types of objects in trajectory data analysis (Andrienko et al., 2013).

In the context of this research, all datasets used are related with human trajectories. Therefore, considering the aforementioned categories, a vehicle, a human, or even a human's avatar in a virtual environment can be considered a *mover*. Meanwhile, the actions caused by, or experienced by, those *movers* can be considered as the *spatial events*, such as a person visiting a specific location, a taxi dropping a passenger, or a video game character defeating an enemy in the map. Despite their similarities, each type of *mover* and their trajectories can be characterized based on different sets of information. For example, if the *mover* is considered to be the human, he/she may be characterized in function of the mode of transportation. If the *mover* is a taxi, this may be characterized based on its driver, its number of passengers, the profit it made, or the type of payment received. If the *mover* is a virtual character controlled by a person, e.g., in a video game, this may be described in function of its number of enemies killed or game related currencies obtained.

2.3 User Objectives and Tasks

Each one of the three sets of information associated with spatio-temporal and trajectory data can be characterized in terms of the other two (Andrienko et al., 2013). In addition to the previously explained triad model, of *space*, *time*, and *objects*, Peuquet (1994) also describes three basic types of relations based on the combinations between the different sets. This supports three different types of questions over the data, namely:

When+where $\rightarrow$ ***what*****:** to state the properties of an object at a given time and location;

When+what $\rightarrow$ ***where*****:** to state the location(s) of object(s) at given time(s);

Where+what $\rightarrow$ ***when*****:** to state the time or set of times when one or more objects were at a certain spatial area.

Based on this framework, Andrienko et al. (2003) consider the relations between two *search targets* and two *search levels*. The *search targets* consist of the spatio-temporal components to be analysed, namely *when* and *what+where*, which can be known *a priori*. This categorization provides an additional emphasis to the temporal component and simplifies Peuquet (1994)'s approach. Furthermore, the *search levels* correspond to the amount of features under analysis, i.e., analysing one feature (*elementary*) or several (*general*). The combination between these two dimensions results in four different categories:

Elementary When and What+Where**:** to describe the characteristics of an object at a given time moment;

Elementary When and General What+Where**:** to describe the situation at a given time;

General When and Elementary What+Where**:** to describe the dynamics of the characteristics of an object at a certain location over a period of time;

General When and What+Where**:** to describe the evolution of the overall situation over a period of time.

Amini et al. (2015) propose a somewhat similar model based on the combination of three dimensions that include: the *search targets* described by Peuquet (1994) (*space*, *time*, and *objects*); two *search levels*, *singular* and *plural*, similar to those proposed by Andrienko et al. (2003); and two levels regarding the knowledge of the data component, namely *known* and *unknown*, depending on whether the information is explicitly given about a certain data component, or if the user has to discover it, respectively. Although implicit in the previous models, by clearly defining this third dimension (*known/unknown*), this framework has the potential to support the characterization of more complex scenarios, depending on the types of components that may need to be discovered, i.e., if there are unknown components, the more complex the scenario.

Roth (2012) emphasizes the importance of Andrienko et al. (2003)'s taxonomy for cartographic interaction since the spatial data component is always taken into consideration and associated to the object. More specifically, space is either already known, and acts as a possible constraint, or it is a main component under consideration, with the other two acting as constraints (Andrienko et al., 2003). Moreover, unlike other approaches, in addition to *search levels* and *search targets*, this taxonomy also supports the concept of *cognitive operations*, defined as the visual analytic processes applied to the representation. The authors propose two types of *cognitive operations*, namely *identify* and *compare*. These operations are equivalent to the tasks/objectives present in the *forming the intention* state, as described by Norman (2002) and Roth (2012), and presented in Section 2.1. The inclusion of these tasks is particularly important, since they significantly change the context of the aforementioned categorizations. For example, a *comparison* task in a *general when elementary what+where* context is different from an *identification* task in the same context. Even if *search levels* and *search targets* may be the same, the objective regarding them is different (e.g., finding if a driver did travel at an illegal speed, comparatively to stating the locations where the driver travelled the fastest). As such, the way in which the user would analyse, interact, and visualize the data will, necessarily be different if the user has to identify or compare something. Consequently, exploring these tasks can help understanding trajectory data, including the types of visualizations used to represent them.

Throughout the literature, several taxonomical approaches have been proposed to address the possible types of tasks, present in the *forming the intention* state, associated with several domains complementary to trajectory visualization, which include cartography, human-computer interaction, and visual analytics. Figure 2.1 depicts the most relevant studies, grouped by the proposed visualization objectives. Listed on the left side of the figure, each study is associated with a differently coloured label. On the right side of the figure, each objective is labelled according to the studies that refer it and is then scaled and ordered according to the number of studies referencing it, i.e., the higher the number of studies referencing that objective was, the larger its representation.

As mentioned in the previous chapter, one of the main objectives of this Ph.D. research implies the conduction of usability studies, which are dependent on the characterization of the tasks described in this section. The analysis of the Figure 2.1 reveals some redundancy, since the same objectives are being identified by several studies and similar or overlapping objectives are being described with different names. In addition, not only the total number of tasks proposed is considerably high for an usability study, but it is also relevant to notice that the number of tasks proposed by the various studies is extremely varied. Therefore, without a clear characterization, using these tasks as a baseline for evaluation can be extremely difficult, if not impossible, due to their redundancy and, most likely, overwhelming number.

To make sense of these contributions, this section presents a simplified view over the proposed visualization objectives (Gonçalves et al., 2015a). For that the following criteria were used:

1. Merge the objectives with similar definitions, or that have the same outcome, keeping those that differ the most among each other;

2. Distribute the remaining objectives into different levels. If an objective can be considered as a specific instance from another, while still being identifiable from others, then this objective is moved to a lower level. If an objective is not adequate in the context of visualization objectives, it is moved to the lowest level (orthogonal).

3. Repeat this process until all tasks are addressed.

Figure 2.2 depicts the result of this simplification. The application of the aforementioned criteria resulted in a four level model, in which the highest level, on top of the figure, represents the most general objectives. The two following levels list specific divisions of the objectives from the levels above. The last level, at the bottom of the figure, lists orthogonal objectives that do not

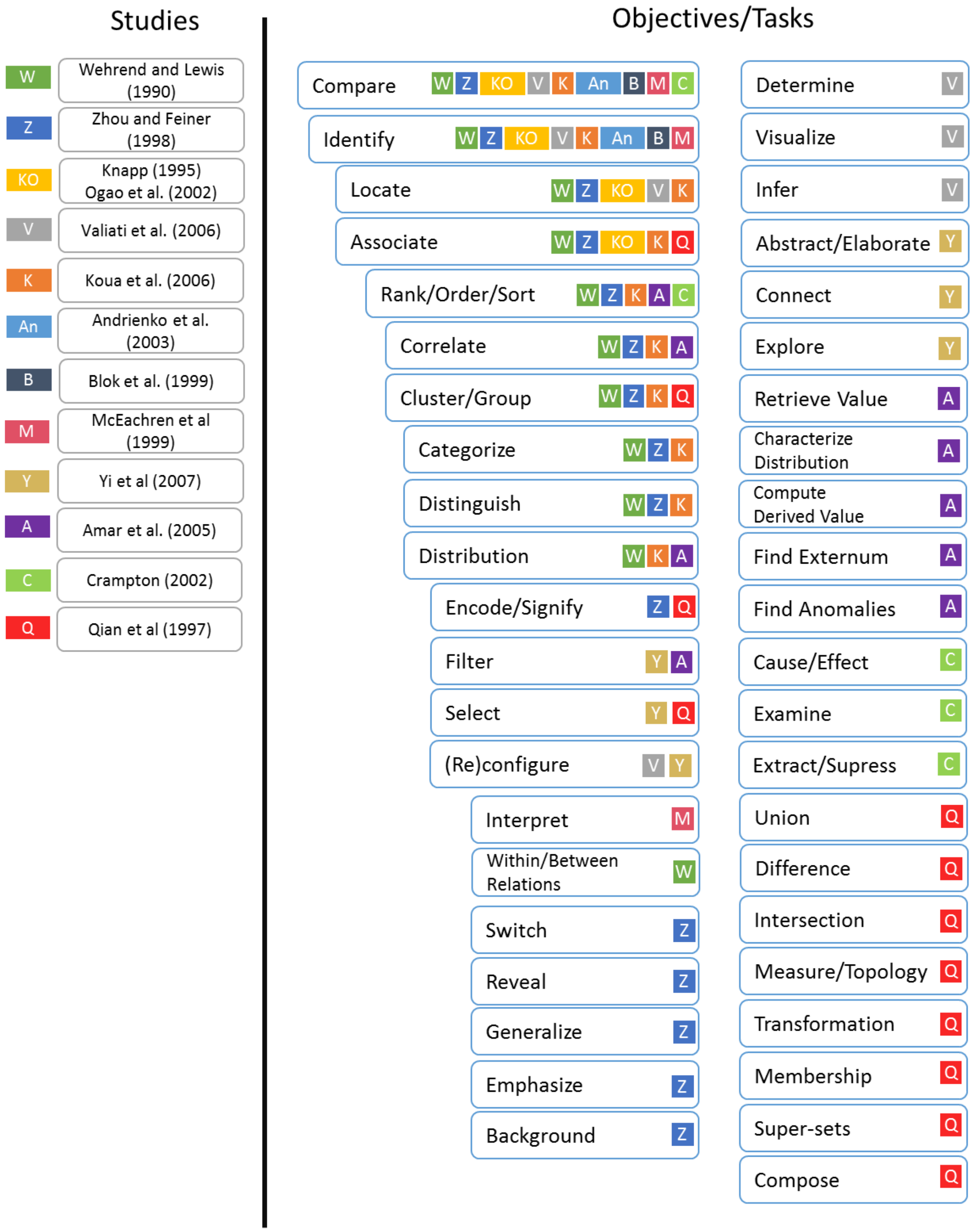

Figure 2.1: List of Task/Objective-based taxonomies, grouped by the proposed objectives.

overlap with the others. To make sense of this model, the following sections describe each level in more detail.

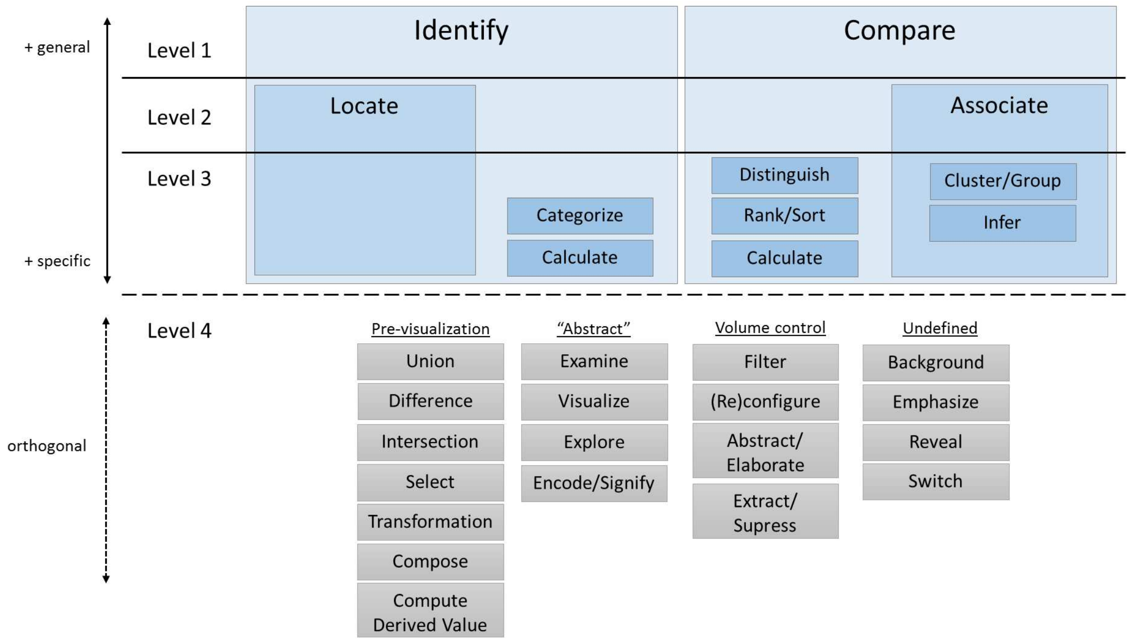

Figure 2.2: Simplified model of proposed visual analytics objectives.

2.3.1 First level - Identify and Compare

The first level of this model is composed by the two most mentioned visualization objectives, namely *identify* and *compare*. These consist of the basis for several taxonomies, including the ones presented by Andrienko et al. (2003), Blok (2000), and MachEachren et al. (1999). They also play an important role in more complex taxonomies, such as the ones presented by Wehrend and Lewis (1990), Zhou and Feiner (1998), or Koua et al. (2006). Despite the small variations between the existing taxonomies, *identify* is generally defined as the examination and description of the characteristics of objects in the visualization. *Compare* is often defined as the detection and measurement of similarity/difference relationships and changes between objects. These can, in turn, be considered as the most general types of objectives, due to the wide scope and applicability provided by their definitions.

2.3.2 Second level - Identify, Compare, Locate, and Associate

The second level of Figure 2.2 includes two additional objectives, namely *locate* and *associate*, originated from the tasks *identify* and *compare*, respectively. These four objectives are the basis of taxonomies such as the ones proposed by Knapp (1995) and Ogao and Kraak (2002). Typically, *locate* is defined as the determination of the existence or position of an object within the visualization, while *associate* is defined as the determination of relationships between two or more objects.

The separation of *identify* from *locate* implies a relevant emphasis on the representation of the data for the *locate* task (*where?* and *when?*, as further discussed in the next section), which in a map visualization tends to coincide with the data's spatial components. Comparatively, the *identify* task may be narrowed to other information (Ogao and Kraak, 2002), typically related with thematic attributes (*what?*). This information can be already known and act as a possible constraint in the location of an object(s) (Valiati et al., 2006; Ogao and Kraak, 2002), or it may need to be identified, after the object(s) being located (Roth, 2012).

Based on the type of visualization, as discussed in Section 2.4, and how the data is organized, the temporal component of the data can be addressed by *location* and *identify* objectives, depending on whether this information is used to position an object or event in time (e.g., an event *e* occurred at 9h30), or if it is used to identify an object or event (e.g., the end of the work-shift).

On the other hand, the specification of the *associate* objective from the *compare* objective implies an important distinction between detecting possible relationships or correlations and determining similarity or differences between sets of data. For instance, analysing the trajectories of a set of users and conclude that they are slower when they move together as a group (*associate*), in contrast to analysing them and concluding that they move slower comparatively to other periods of time (*comparison*).

2.3.3 Third level - Specific Cases

The third level of the figure depicts the least mentioned objectives in the literature, that usually consist of special or specific cases of the objectives presented in the level above (*identify*, *locate*, *compare*, and *associate*). This level highlights six additional objectives, namely: *categorize*, *distinguish*, *rank*, *calculate/determine*, *cluster*, and *infer*, present in taxonomies, such as the ones proposed by Wehrend and Lewis (1990), Koua et al. (2006), Amar et al. (2005), Zhou and Feiner (1998), and Valiati et al. (2006).

Typically, *categorize* is described as defining objects based on a given classification (Koua et al., 2006), thus being equivalent to describing an object's characteristic according to a specific set of rules (e.g., categorizing a mover as *slow* based on the speed values that were identified).

Distinguish consists of the recognition of an object as different from one or several others, while *rank* consists of the attribution of an order, according to some ordinal metric. These can be considered two specific types of *comparison* objectives, depending if the data under analysis are nominal or ordinal, respectively (Bertin, 1967).

Determine, proposed by Valiati et al. (2006), is defined as the calculation of precise values. For a simplification of language, this objective will be addressed as *Calculate*. Although this type of task may, actually, be seen better as an *operator* (Roth, 2012) (i.e., the actions executed to complete a task), it is important to reference it due to its relation with the analysis of numeric data, comparatively to the *distinguish* and *rank* objectives, that generally deal with nominal and ordinal data, respectively. As evidenced in Figure 2.2, two types of *calculate* tasks are considered, one related to the *identify* task, the other with *compare*. These depend on the type of values being calculated. *Calculate-identify* deals with values solely based on an object's attributes, which results in the identification of a new attribute (e.g., calculating a movers' average speed, based on the size of the representation). *Calculate-compare* deals with values obtained through the properties of different objects, e.g., calculating the difference in speed between two movers at a certain time, to compare how faster one was moving relatively to the other.

The objectives *cluster* and *infer* can be considered as specific cases of the *associate* objective. Typically, *cluster* is defined as the detection of objects with similar attributes, which implies the comparison of those objects and the consequent grouping due to their similarity. *Infer* is described as the determination of cause/effect relationships between data components, also implying the comparison of two or several objects, over certain conditions, and the consequent detection of a pattern (e.g., detecting high movement speeds at certain hours of the day, over a period of time).

2.3.4 Fourth level - Orthogonal Tasks

The final level of Figure 2.2 lists a group of objectives that offer minimal overlap with the ones that were previously described. Some of these tasks are present in a similar categorization described by Roth (2012), aimed at organizing existing taxonomies based on visualization and cartographic objectives. Roth (2012) justifies this similar disparity due to the fact that the majority of those cases are better understood as *operators*, rather than cartographic objectives/tasks. Consequently, even though these may not be considered as cognitive inputs over the visualization, they may still affect the visualization.

In this level, some of the objectives that were proposed may actually precede the displaying of the data itself, such as the operations that were identified by Qian et al. (1997), in a taxonomy of spatio-temporal visualization and database operations. These operations are focused on the manipulation of the data and include: *union*, *difference*, and *intersection*, to merge sets, and obtain unique or common elements between sets, respectively; *select*, to obtain information sets, based on certain constraints; *transformation*, to configure an element's properties; and *compose*, to merge several elements into a single complex one. In addition, *compute derived value* (Amar and Stasko, 2004) is described as the computation of an aggregate numeric representation of data sets.

Other objectives can be applied during the analysis of the displayed data. Some may be considered as abstract operations, like *explore/visualize* (Valiati et al., 2006; Yi et al., 2007) and *examine* (Crampton, 2002), that consist of changing the parameters and inspecting a visualization, respectively. Others are responsible for how to display the data on the visualization, like *encode/signify* (Zhou and Feiner, 1998; Qian et al., 1997). In comparison, objectives such as *filter/suppress*, *extract*, *(re)configure*, or *abstract/elaborate* are responsible for the control over which sets of information are displayed. *Filter* (Amar and Stasko, 2004) and *suppress* (Crampton, 2002) are defined as the acquisition of information based on conditions set over data attributes. Contrarily, *extract* (Crampton, 2002) is defined as the highlighting of specific features on a visualization. *Reconfigure* consists in the arrangement or change of the visual characteristics used to represent the data (Amar and Stasko, 2004; Yi et al., 2007). *Abstract/elaborate* is defined as the request for information in more or less detail (Yi et al., 2007).

Finally, Zhou and Feiner (1998) proposed a set of objectives for which no definitions were provided, namely *background*, *emphasize*, *reveal*, and *switch*.

2.3.5 Omitted Objectives

As stated above, one of the criteria for the conception of this simplified view was the omission of objectives considered to be overlapping with existing ones. It is important to emphasize that, by omitting these objectives, they are not being labelled as useless. Nevertheless, this simplification was a necessary step to prevent confusion and redundancy.

Find externum (Amar and Stasko, 2004) is defined as the identification of extreme attribute values, thus, being the consequence of a *ranking* task, that allows determining the order of attribute values. In addition, operations such as *retrieve value* (Amar and Stasko, 2004) and *(characterize) distribution* (Koua and Kraak, 2004; Wehrend and Lewis, 1990; Amar and Stasko, 2004) are defined as the acquisition of an attribute from a known object, or a set of objects, respectively. Thus, both operations are equivalent to the *identify* objective described in the previous sections. Finally, the remaining objectives were considered similar to *associate*, since their function is typically focused in the detection of relationships between objects (*within/between relations*, *cause/effect*, *correlate*), the search for objects that go against those relationships (*find anomalies*), and objects that share similar characteristics among themselves (*generalize, membership, super-sets*) or among real life correspondents (*interpret*).

2.4 Types of Map Visualization Techniques

This section describes some of the most relevant techniques for map-based visualization of trajectories. In particular, thematic maps are considered as an important method for the visualization of trajectories given the spatial characteristics associated with the data. More than just presenting the topology of a given location, these provide an additional set of layers with information, distributed over the mapped location, based on the temporal and thematic properties associated to the moving object (Reichenbacher, 2007). This provides an effective, and often crucial, means for summarizing and communicating complex georeferenced information (Kessler and Slocum, 2011), which helps answering to various spatio-temporal related questions (Vasiliev, 1997; Kraak and Ormeling, 2010).

Although other visualization techniques may be used instead to work out conclusions regarding trajectory data, such as timelines/time-graphs representing variations of attribute values in time (Andrienko and Andrienko, 2013), their isolated use is not recommended, since these techniques tend to neglect and not fully represent the spatial component of the data (Andrienko et al., 2011). This may significantly limit the ability to formulate and solve relevant spatial problems (Kraak and Ormeling, 2010). Nevertheless, these methods can still be used in conjunction with map visualization techniques to provide additional points of view over the data (Riveiro et al., 2008; Hurter et al., 2009; Pu et al., 2013).

Despite acknowledging the existence and importance of additional types of visual representations, this section is focused on map visualization techniques, particularly, in terms of visualizing both spatial and temporal data simultaneously. Moreover, it is also important to mention that these can either be the main feature of an application, or it can be one of many components of the application, in the sense that it can be viewed alongside other map and non-map visualization techniques (Andrienko and Andrienko, 2013; Hurter et al., 2009; Riveiro et al., 2008; Pu et al., 2013).

Throughout the literature, several map-based techniques have been proposed for the visualization and exploration of the spatio-temporal and thematic properties of trajectory data (Andrienko et al., 2011; Andrienko and Andrienko, 2013). For the most part, these are the result of the possible combinations of the different types of visual variables from visual marks (i.e., points, lines, areas, and volumes). Previous studies have already focused on the identification of the most relevant types of visual variables, like *position/location*, *shape*, *colour hue/value/saturation*, or *size*, among others (Bertin, 1967; Green, 1998; MacEachren, 1995). Each of these properties can be further characterized based on their adequacy to represent certain types of information or capture the users' attention. For example, attributes like *size* or *position* are considered to be better to represent numerical data, while *colour hue* or *shapes* are better to represent ordered and categorical data, respectively (Wolfe and Horowitz, 2004; Garlandini and Fabrikant, 2009; Mackinlay, 1986).

As previously mentioned, this research is focused on the differences between map-based visualization techniques and how these deal with the simultaneous representation of space and time. For that reason, throughout the following sections and the remainder of this document, there is a focus on a high-level categorization of four types techniques, which include, *static maps*, *space-time cubes*, *animated maps*, and *small-multiple maps* (Kraak and Ormeling, 2010; Andrienko et al., 2013; Gonçalves et al., 2014; Gonçalves et al., 2013a), and details the main characteristics associated to these high-level types of visualization techniques, based on the existing literature.

2.4.1 Static Maps

For many years, paper maps were the only way to represent georeferenced information and, in particular, trajectory data. These maps are still used today, which led users to become more familiarized with static maps, particularly with two dimensional maps.

It is important to mention that, in this work, the term *static* maps is used in contrast with *animated* maps. While the first type may use animated effects to emphasize static information, the second effectively requires the use of animation to represent the evolution of an object's movement. Therefore, considering this research's focus on interactive map visualization, *animated* maps (that require animated features) should not be mistaken with *interactive* maps (Roth, 2013) that provide features that may take advantage of animations to support the change in appearance or contents within the map.

Typically, the most simple approach for representing the evolution of one or several movers' trajectories consists in using lines or point symbols over the geographical representation. In turn, spatial events may be represented as icons with a different shape (Vasiliev, 1997; Phan et al., 2005). These symbols can be adapted to represent variations of the thematic attributes, through the change of colours, shapes, sizes/widths, or any other visual variable (Kraak and Ormeling,

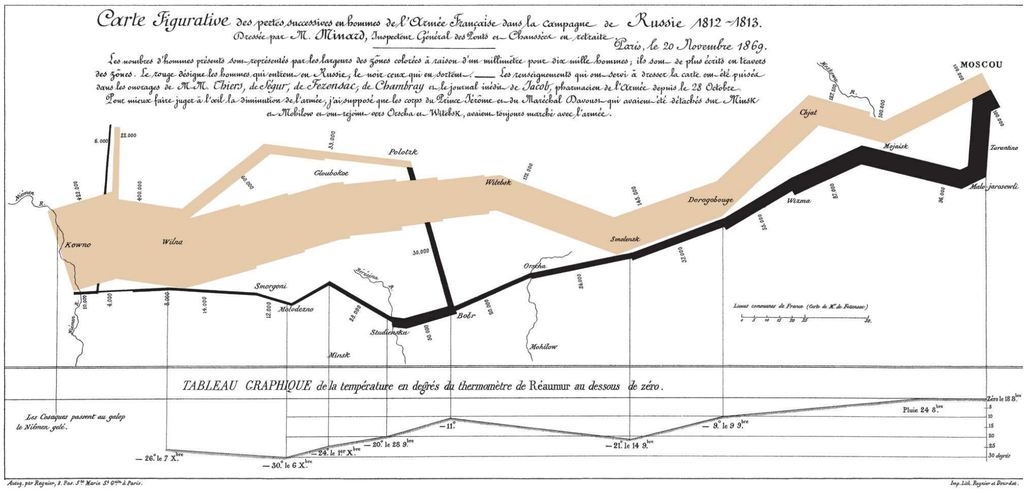

Figure 2.3: Minard's static map representing Napoleon's army campaign in Russia (source: https://cartographia.wordpress.com/category/charles-joseph-minard).

2010; Bertin, 1967; Perše et al., 2009; Perin et al., 2013; Lucey et al., 2014; Wei et al., 2013; Tani et al., 2014).

Figure 2.3 illustrates one of the most commonly referenced maps in the literature following this style of representation, namely Minard's 1869 map describing the campaign of Napoleon's army in Russia (Kraak and Ormeling, 2010). In this example, line marks are used, overlaying a simplified topographic view, to represent the trajectory followed by Napoleon's troops. Colour is used to categorize the troops' actions (*attacking* - cream; *retreating* - black). In addition, line thickness is used to represent the number of troops, while several text symbols are used to identify specific locations and to associate the width of the lines with the real number of troops. Finally, this map also provides a small diagram representing the variation of the temperature over the geographical area. Through this map, it is possible to easily understand the considerable decrease in the number of troops, associated with the low temperatures of each region.

Considering the often large dimensions of trajectory datasets, more complex representations might be needed to prevent issues, such as over-cluttering of information, or to support the visualization of more complex information. A common approach consists of grouping similar, and frequent, trajectories or events and showing them as one (Tobler, 1987; Andris and Hardisty, 2011; Pu et al., 2013; Shao et al., 2016). This approach reduces the cluttering of trajectories, and although these visualizations may not clearly reflect the temporal dimension of the data, they show cumulative movements that occurred during a certain period of time (Andrienko et al., 2013). Similarly to simple approaches, aggregated trajectories can be enhanced with thematic information, through the manipulation of visual variables, for instance, in Figure 2.4 movements are represented as arrows proportional to the number of objects that followed that trajectory (Tobler, 1987; Andrienko and Andrienko, 2008)).

Another common approach consists of using 2D histograms or *heatmaps*, to show the frequency of a certain type of spatial event. The colouring of these maps can, in an intuitive way,

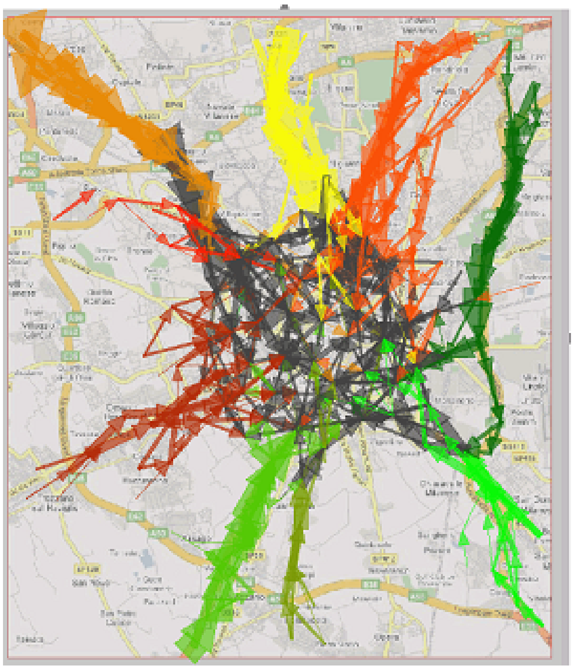

Figure 2.4: Trajectories represented by aggregated arrows (source: Andrienko and Andrienko (2008)).

transmit the notion of which areas are more often used (Pingali et al., 2001; Goldsberry, 2012; Pileggi et al., 2012; Andrienko et al., 2013). Similarly, density maps can also be considered as an alternative for aggregated trajectory visualization. These maps consist of an aggregate overview of the data, where wider density fields correspond to larger amounts of data, and their colour and saturation variations may represent the instant in time or the variation of an attribute, for instance, the average speed of the moving objects (Scheepens et al., 2011, 2012).

Although time is an inherent property of trajectory data, from the four aforementioned categories of techniques, static maps are the ones in which temporal information may be the most easily overlooked. To prevent that, time can be represented similarly to a thematic attribute, for instance, by using visual variables like transparency (Booker et al., 2007). Alternatively, additional symbols can also be used, like: textual timestamps, explicitly showing the actual time when the object passed through different landmarks; or cross-lines, orthogonal to the trajectory followed by the object, in which, the closer the cross-lines, the slower the object was moving (Kjellin et al., 2010b). Another alternative consists in the use of additional tables and graphs, inside or outside the map, to clearly represent the variation of attributes during time (Riveiro et al., 2008; Pu et al., 2013). Andrienko and Andrienko (2008), for instance, suggest the representation of aggregated movements with complex symbols, similar to tables (Figure 2.5), where the various cells are coloured according to different thematic attributes and distributed by days of the week (columns) and hours of the day (rows). On the other hand, works such as those of Ivanov et al. (2007); Andrienko and Andrienko (2007); de Oliveira and de Souza Baptista (2012) and Thudt et al. (2013) describe visualization systems with separated views for temporal information, to compensate the

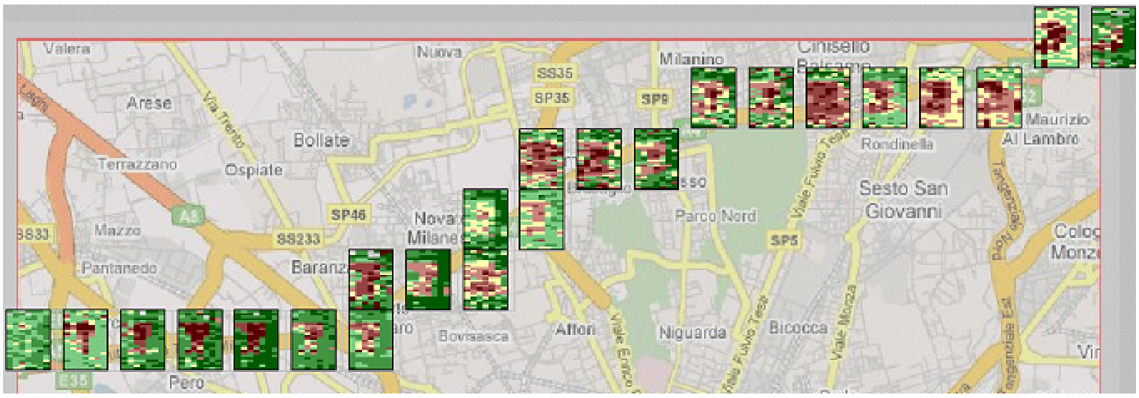

Figure 2.5: Mosaic diagrams showing the variation of the median speeds in spatial compartments by days of the week (columns) and hours of the day (rows) (source: Andrienko and Andrienko (2008)).

analysis of spatio-temporal and trajectory data on 2D static maps. Consequently, through these types of representations, time usually needs to be *identified* as a specific attribute in the map or in its supplementary visualizations.

Despite the focus of this section in the representation of trajectories in 2D static maps, several of the concepts described can also be applied over three-dimensional (3D) maps. However, given the similarities between three-dimensional maps and space-time cubes, their characteristics are presented in the next section.

2.4.2 Space-Time Cubes

With the advancements within the field of computer graphics, three dimensional visualizations have become more common, and are continuously being seen as a promising way to represent complex information (Kjellin et al., 2010b; Bleisch, 2012).

However, despite the similarity of the strategies used in 2D and 3D visualizations (e.g., similar use of visual variables), 3D representations are not always considered as useful, due to the limitations associated with human perception. With the existence of an additional spatial dimension, the point of view within the visualization constitutes a critical factor, since it may affect the perception of the measurements. Consequently, 3D visualizations are often regarded as more adequate in representing qualitative, rather than quantitative information (John et al., 2001; Kjellin et al., 2010a,b; Seipel, 2013). Nevertheless, and despite the lack of general guidelines for the design of efficient 3D visualizations (Kjellin et al., 2010b), several representations of trajectory data (and spatio-temporal) can be found in the literature (Bach et al., 2014, 2016).

Like in 2D static maps, the simplest way for representing trajectories or spatial events, consists in the use of points/lines graphically encoded to represent the variations of positions and thematic attributes over time (Pingali et al., 2001; Kapler and Wright, 2004; Eccles et al., 2008; Ware et al., 2006). Alternatively, more complex icons can also be used, ranging from simple cylinders, representing stop locations (Buard and Brasebin, 2011), to more complex symbols, depicting human

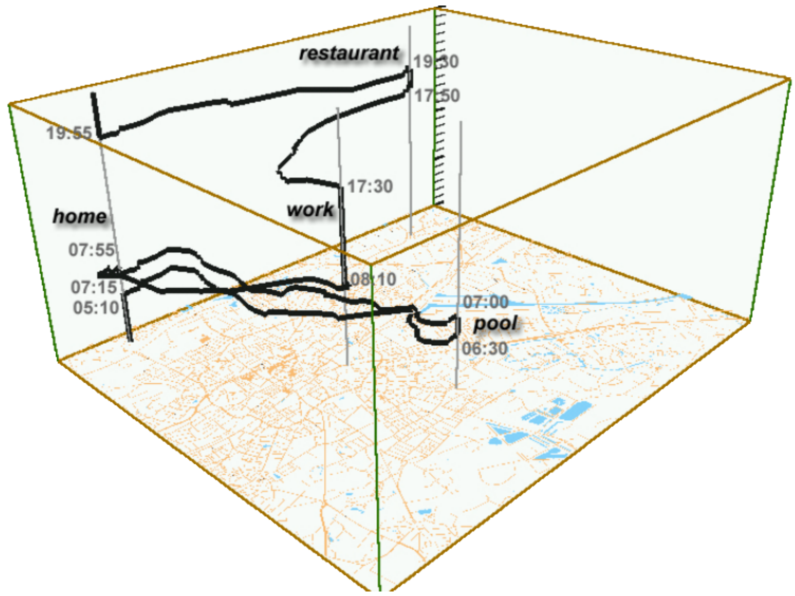

Figure 2.6: Trajectories represented on a space-time cube visualization (source: http://www.aviz.fr/ bbach/spacetimecubes/figures/Kraak2003spacetimecube.png).

dynamics, e.g., icons with different shapes to represent various actions (Liu and Lai, 2009).

Also similarly to 2D static map approaches, time can be represented with the methods discussed in the previous section, like textual labels, to specifically describe the time an object passed through a landmark, or the amount of time it stayed there (Liu and Lai, 2009). However, the existence of an additional spatial dimension, which is sometimes not fundamental for the understanding of the data, can be used instead to better emphasize the spatio-temporal properties of movement data. This type of approach is denominated space-time (Hägerstrand, 1970) and an example is depicted in Figure 2.6.

In the context of trajectory data visualization, the space-time cube technique (STC) has been considered a possible alternative to common 2D maps. This technique emphasizes the idea that time and space are inseparable, by representing both spatial and temporal information within a 3D cube (Demsar and Virrantaus, 2010). Typically, two of its coordinates/axes (e.g., the $x - y$ axes) are used to represent spatial information (e.g., latitude and longitude positions) and the third dimension (i.e., the $z-$axis) is used to represent time. Although with exceptions (Amini et al., 2015), time usually increases alongside the STC's height, implying that the higher the information is within the cube, the more recently the event happened (Kjellin et al., 2010a). Figure 2.7 represents these concepts with a simple example. Assuming that the points I and II correspond to events, it is possible to see that the x and y coordinates of the two events are the same, corresponding to the location A; however, event II is higher within the cube. This means that although event I and II occurred at the same location (A), event II is more recent than I.

Although somewhat similar to 2D maps for the representation of spatial and thematic data, using the third dimension to convey temporal information minimizes the 2D map's limitations in representing time, and implies that time needs to be *located* within the STC, rather than *identified*,

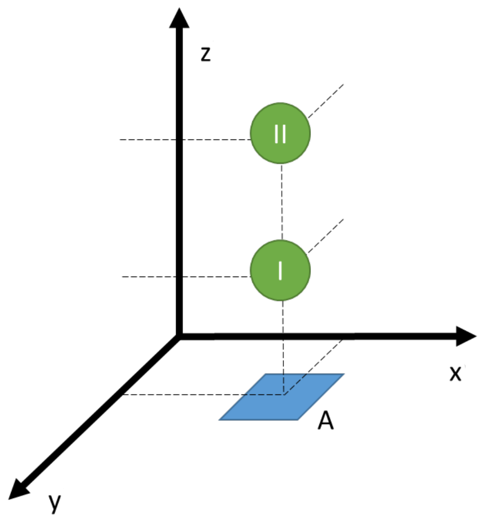

Figure 2.7: Representation of events at different time moments, but in the same location, on a space-time cube.

like in the 2D map (i.e., time is conveyed by a location, instead of a colour or a sequence of icons). This also allows the use of visual variables to represent other thematic attributes, instead of using them to depict temporal properties in a 2D map (Kjellin et al., 2010b). More importantly, representing time as a *location* supports the portrayal of several layers of information, each one defined as a plane along the $z-$axis, representing the state of a geographical area during a specific moment in time (Tominski et al., 2005; Tominski and Schulz, 2012; Thakur and Hanson, 2010).

Identifying movement features in one or several trajectories can be significantly facilitated by an STC visualization. For example, when using a line to represent a certain trajectory, stop locations will be represented by vertical lines, i.e., the location was the same during a period of time; on the other hand, diagonal lines will represent movement, i.e., different locations visited at specific times (Amini et al., 2015).

Since the third dimension is used to represent time, STCs may not be the most adequate technique to represent 3D trajectories, like those of planes. In these situations, either a 3D static map may be used, or the third spatial coordinate can be represented as a thematic attribute, e.g., by using different colours for different levels of altitude (Hurter et al., 2009).

Although it has been reported that users may show more certainty on their decisions when using 3D visualizations (Seipel, 2013; Kveladze et al., 2015), some studies have reported that users felt some discomfort after interacting with these visualizations (Seipel and Carvalho, 2012). Furthermore, it is also expected that STCs are affected by the limitations of human perception with three dimensional environments. More specifically, the adequacy of this technique depends on whereas the properties of the structures under analysis are invariant under affine transformations, i.e., if the scale and other visual properties of the displayed information are preserved or when, for instance, the STC's rotation changes (Kjellin et al., 2010a). Moreover, similarly to 2D maps, previous studies argue that STCs are also negatively affected when a large number of trajectories is on display, given the possibility of over-plotting and over-cluttering (Scheepens et al., 2012). Demsar and Virrantaus (2010), for instance, argue that 10 trajectories are enough to make the STC an unsatisfactory approach. Interestingly, previous studies have also suggested a possible effect of gender over the usage of 3D visualization techniques, which consequently is likely to affect STCs (Vecchi and Girelli, 1998). Particularly, male users have shown a better performance when dealing with tasks and visualizations related to a 3D view, comparatively to female users who may excel in using/analysing (2D) static approaches.

Several methods have been proposed to mitigate the aforementioned issues, including interaction, filtering, data transformation, and data aggregation techniques. The simplest approach consists in providing navigation methods that allow the user to change the point of view within the visualization. These allow the user to obtain more information from different points of view, some of which may be more favourable to the visualization of the data (Kjellin et al., 2010b). On the other hand, Kraak (2008) suggests the interactive adjustment of the STC's spatial plane to facilitate the location of objects in space and time. Demsar and Virrantaus (2010) propose the aggregation of trajectory data and their visualization on density maps over the STC technique. Similarly, previous studies have also proposed the use of stacking symbols to represent the evolution of certain thematic attributes from aggregated trajectories and spatial events over time, e.g., through methods like 3D pencil/helix icons (Tominski et al., 2005), data vases (Thakur and Hanson, 2010), or trajectory walls (Tominski et al., 2012). In addition, Drecki and Forer (2000) and Andrienko et al. (2013) suggest the use of time transformation techniques for the visualization of multiple trajectories in the STC. In this case, the transformation is achieved through the manipulation of the temporal references within the trajectories, to facilitate comparisons, e.g., by providing common start or end times for the different trajectories.

2.4.3 Animated and Small-Multiple Maps

Despite their differences in presentation, animated and small-multiple maps follow a similar approach in terms of how the data is handled. Both techniques support the visualization of various maps or various different states of the same map. However, animated maps display those maps in a sequence and in a single view (Harrower and Fabrikant, 2008). In comparison, small-multiple maps display those maps consecutively and juxtaposed to each other (Andrienko et al., 2003). Figures 2.8 and 2.9 depict examples of each technique. In the first, the visualization is composed by several maps, each one displayed once at a specific time moment. In the second, each juxtaposed map represents one day of registered movement.

In the context of this characterization for animated maps, this section focuses on *temporal animations* (Harrower and Fabrikant, 2008; Kraak and Ormeling, 2010). More specifically, these deal with the representation of events, dynamically, in a chronological order (e.g., a moving symbol representing the trajectory of an object). These should not be confused with *non-temporal animations*, which can represent changes of attributes of a dynamic phenomenon, with no direct relation with world time (e.g., the use of blinking symbols to attract attention to certain objects).

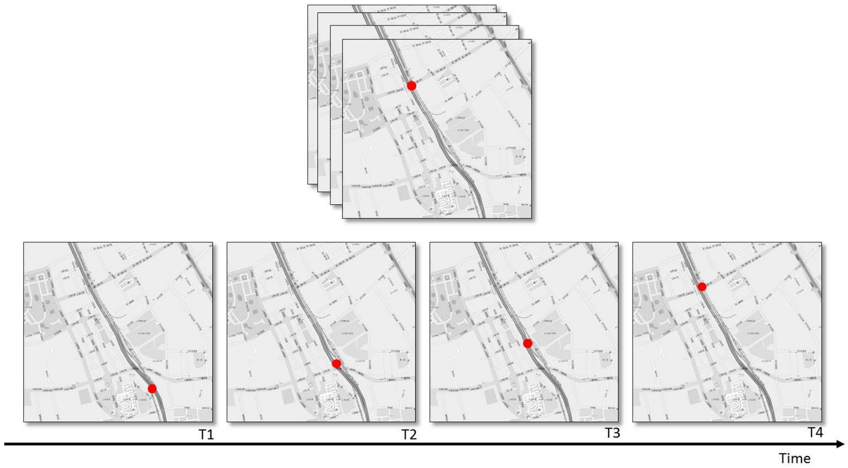

Figure 2.8: Representation of an animated map.

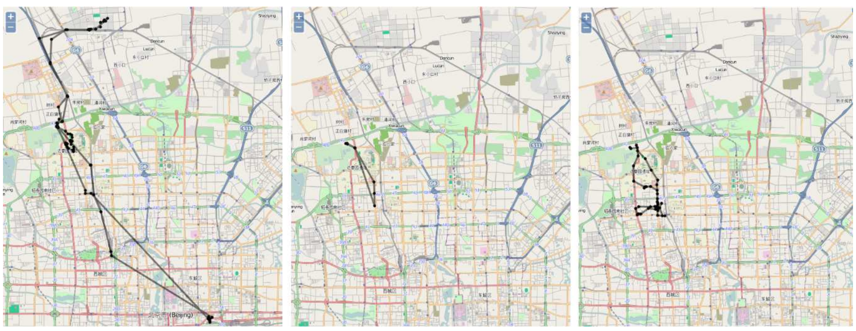

Figure 2.9: Representation of a small-multiples map visualization.

Unlike static approaches, animated maps have an additional dimension that can be used to present information and which, in theory, allows the use of other visual variables to represent thematic information, similarly to STCs. This also supports the existence of additional parameters, specific for animated displays, that can be used to help on the presentation and exploration of information (Robertson et al., 2002). These include, among others: animation speed, step (i.e., the interval between successive time moments), and smoothness (i.e., creation of intermediate frames through interpolation) (Andrienko et al., 2003).

Each map (or frame) within an animated map represents a time moment, and, as a result, time is perceived in function of the direct relation between display and real time (e.g., faster objects in the animation correspond to the faster ones in reality). On the other hand, each map in a small-multiples visualization represents the state of a phenomenon at a different time instant or period. Changes are perceived through the analysis of the succession of individual maps, and their visual characteristics. Consequently, time is represented as a spatial sequence that the user needs to follow to properly understand the temporal variation of the data (Kraak and Ormeling, 2010).

Preliminary research studies suggest that animated displays can be particularly useful for presentation and exploration purposes (Ogao and Kraak, 2002) and that they can help revealing spatio-temporal patterns that are not evident with common static representations (MachEachren et al., 1998). For presentation proposes, animations should be simple enough to properly convey the intended message. On the other hand, for exploration purposes, it is advised to provide tools that help manipulating or controlling the visualization of the spatio-temporal properties of the data, such as a time-slider, to support a faster navigation through the data and to detect relevant information (Andrienko and Andrienko, 2007; de Oliveira and de Souza Baptista, 2012). Furthermore, the amount of data to be presented increases with the duration of the animation. Even with the addition of interactive controls that allow users to control the animation, one frame will not always be visible on the display. These factors may, in turn, raise some cognitive and perceptual limitations, since the longer the animation, the less likely all relevant information will be memorized by the user (Harrower and Fabrikant, 2008).

In opposition, each small multiple map is always visible to the user, allowing the visualization of two (or more) time moments simultaneously (or periods, depending on the amount of time represented on each map) (Andrienko et al., 2003). Considering the ephemeral nature of features in animations, small multiple maps may have a smaller overload in the users' memory, in comparison to animated maps (Griffin et al., 2006). Consequently, small multiple approaches have been considered an adequate method to detect changes and repetitions, identify behaviours, and compare different phenomena at different periods of time (Tufte, 1997; Andrienko et al., 2003; Demissie, 2010). Nevertheless, although each map represents less information than a conventional static map or STC, considering the limited size of the visualization, the dimension of each map and the symbols used to present the information will be smaller as the number of maps increases, which may lead to a loss of detail (Ostermann, 2010). Another limitation is related to the fact that, although a movement along a trajectory represents a continuous scenario, small multiples are not continuous spatial representations. To properly perceive the continuity of the movement, users need to combine the information from each individual map into a single event. This task may require a significant effort when a considerable number of maps are used (Kraak and Ormeling, 2010; Tufte, 1997). Even so, some authors state that the effort required with this technique is not significant comparing with animated representations (Griffin et al., 2006).

2.5 Comparative Studies

Despite one of the motivations behind this research being the lack of user studies and empirical data regarding the different types of map representations for spatio-temporal and trajectory data visualization, it is important to acknowledge the already existing studies. Although limited, some studies comparing STCs, static, animated and small-multiple maps have already been reported. These consist of user studies based on the analysis of user performance and subjective preferences, and case studies analysing the characteristics of the techniques. Section 2.4 focused, for the most part, on the main conclusions regarding the second group of studies. As such, and considering that the goals of this work are dependent on the empirical comparison of the existing techniques, this section focuses on the results of the most relevant comparative user studies.

Table 2.1 summarizes some of the most important studies, identifying the types of techniques compared, and their most relevant conclusions. Although initially intended for it to also include the types of tasks used, several studies lacked a proper description of them. This, in turn would lead to several empty cells in the table, thus, potentially interfering with its analysis. It is important to mention that, despite this research being focused on the interactive visualization of human trajectory data, there is still significant knowledge to be acquired from studies applying these techniques in other spatio-temporal related contexts. Notable examples include authors like Cutler (1998) or Robertson et al. (2008) which compared animated with static and small multiple displays, respectively, regarding their capabilities in making users recall information, and in being used to present and analyse spatio-temporal data. Slocum et al. (2004) aimed at examining cognitive issues associated with the display of a dynamic geographic phenomenon, taking into account a population with varying levels of experience in geographic analysis and visualization. Griffin et al. (2006) studied user performance and the overall cognitive workload involved in reading animated maps to detect moving clusters. Kjellin et al. (2010a) and Seipel (2013) compared two different STC approaches, based on their levels of 3D, and contribute to this analysis by suggesting the similarities between weak 3D and 2D maps. These levels, *weak* and *strong* 3D, depend on the technique being a static monocular or a head-tracked stereoscopic visualization, respectively.

The analysis of Table 2.1 highlights some important characteristics of these techniques. Overall, STCs are considered to be an effective technique for the analysis of spatio-temporal patterns and to have a generally small learning curve (Kjellin et al., 2010a,b). Despite that, the usefulness of this technique appears to be related with the projection used. Kjellin et al. (2010a) did not find any significant differences between *weak* and *strong* 3D STC approaches. Contrarily,

Table 2.1: Comparison of static maps, space-time cubes, animations, and small multiples.

Study	**Techniques**	**Conclusions**
Koussoulakou and Kraak (1992)	Animation, Small multiples	Animated maps better than small multiples to detect moving clusters
Patton and Cammack (1996)	Animation, Static cloropleth map	Animation has better performance
Cutler (1998)	Static map, Animation	Animation has lower performance
Slocum et al. (2004)	Animation, Small multiples	Animation is adequate for identifying general trends; Small multiples are adequate for comparing time points
Griffin et al. (2006)	Animation, Small Multiples	Animation is better to detect moving clusters
Robertson et al. (2008)	Animation, Small multiples	Animation is confusing, but *fun* to use and more effective for presentation proposes; Small multiples have better performance in analysis tasks
Kristensson et al. (2007, 2009)	Static map, Space-time cube	Static Map is adequate for elementary questions; Space-time cube is efficient for general questions and has a small learning curve
Demissie (2010)	Static map, space-time cube, animation, small-multiples	Animation is the most adequate to count movement features (stops, returns, speed changes)
Kjellin et al. (2010a)	Strong and weak space-time cubes	Effective visualization technique; No performance differences between strong and weak 3D approaches
Kjellin et al. (2010b)	Static map, Space-time cube, Animation	Static map is beneficial when dealing with metric properties; Space-time cube is beneficial when dealing with ordinal information; Interactive techniques are important; Animation is, overall, ineffective
Willems et al. (2011)	Static Density map, Space-time cube, animation	Density maps are good for detecting stop features; Space-time cubes are good for detecting *busy* locations, however, are the least scalable; Animation is good for detecting fast moving objects; Least robust to clutter
Seipel (2013)	Static map, Strong and weak 3D space-time cubes	Static maps have a similar performance with weak 3D STC; Strong 3D approaches lead to more errors
Bisantz et al. (2014)	2D map and STC, with animated controls	No significant differences between 2D and 3D representations; Increased workload with 3D representations; Meta information regarding locations should be provided
Amini et al. (2015)	2D map and Space-time cube, with animation controls	STC is efficient in cluster-based question; STC was picked over 2D, even if considered *busy* and *messy*; STC is beneficial in examining sequences of events to identify complex behaviours
Kveladze et al. (2015)	Static map, Space-time cube	Domain experts are better than non-domain experts to solve tasks; Although, both techniques were highly valued, 2D views were always considered necessary

Seipel (2013) claim that *strong* 3D levels may lead the users to be less accurate, while *weak* 3D approaches may be considered similar to 2D maps (Seipel, 2013; Bisantz et al., 2014). Other studies present results that go somewhat against these conclusions. Kristensson et al. (2009), for example, claim that space-time cubes are efficient for complex spatio-temporal tasks. Similarly, Amini et al. (2015) state that this technique is useful for the examination of sequences of events to identify complex behaviours. Additionally, Kjellin et al. (2010b) praise the STC for its ability to help users in analysing some movement characteristics, namely the order or arrival of objects to certain places. Likewise, Willems et al. (2011) and Amini et al. (2015) state that this technique is good for detecting which locations are the most busy and answer to cluster-based questions (i.e., questions involving the meetings of multiple objects), respectively.

Small-multiples and single static map approaches are among the most used when comparing with other techniques, particularly animated representations. The results of previous studies suggest that static maps are beneficial for elementary tasks and when dealing with metric properties (Kristensson et al., 2009). In addition, the use of aggregation techniques prior to the visualization, such as in density maps, alongside the manipulation of their visual variables, such as colour, to represent temporal and thematic attributes, is considered to be adequate to help finding stop features (Willems et al., 2011). Moreover, the studies addressing the small-multiples map technique suggest its adequacy in analysis tasks, comparatively to animated representations, and when comparing information on different points in time (Slocum et al., 2004; Robertson et al., 2008).

Although the same argument can be applied to the STC technique, the analysis of the previous studies addressing animated representations suggests some mixed, if not contradictory results. Robertson et al. (2008) state that although animated techniques may be considered confusing, they still seem to be well accepted by the users and to be adequate for presentation purposes. Similarly, other studies concluded that animated maps are better than other representations for the identification of moving trends and the detection of moving clusters (Koussoulakou and Kraak, 1992; Slocum et al., 2004; Griffin et al., 2006; Demissie, 2010; Willems et al., 2011). These conclusions make sense considering that movement/animation is considered a pre-attentive visualization attribute (Wolfe and Horowitz, 2004), thus making information represented by this technique more noticeable to the users. However, authors like Cutler (1998) and Kjellin et al. (2010b) argue that these techniques are generally ineffective, with lower performance, and that static approaches may support an acquisition of knowledge as good, if not better, than animated approaches. This fact was already noted by some authors, like Tversky et al. (2002), Lobben (2008) and Robertson et al. (2008), who suggest the possibility of biased or unbalanced evaluations that, ultimately, benefit certain visualization techniques over others.

It is important to acknowledge the fact that these studies cover a wide period of time and that the technology available for the creation of interactive displays was, comparatively to the date of this document's writing, more limited. Nevertheless, both the contradictory nature of some of the results presented above, alongside the uncertainty associated with them, emphasizes one of the current limitations within this research subject - the lack of empirical knowledge regarding the existing types of techniques.

Another important detail that must be taken into account is related with the possibility of combining the different types of visualization techniques. This is a very important topic in this research and it will be addressed with more detail in future chapters. Assuming the results of the previous studies to be valid, they ultimately suggest that different techniques are more useful in specific types of tasks or applications. Therefore, it is plausible to assume that combining the different types of techniques may help minimizing the expected limitations from their individual components. Interestingly, some studies published during this research somewhat address this idea, thus emphasizing the motivations discussed in Chapter 1. Kveladze et al. (2015) present a visualization application that allows the user to switch between a 2D static map and a 3D space-time cube, by using an interactive menu option, and applied it on a user study, comparing the capabilities of domain and non-domain experts in spatio-temporal data visualization. Based on the analysis of interviews conducted with the participants, the results suggest that both groups were receptive towards this feature. This idea is also supported by Amini et al. (2015), who propose the ability to switch between 2D and 3D map visualization techniques to allow the exploration of the data from different points of view. In addition, Amini et al. (2015) and Bisantz et al. (2014) combine the basic concepts of animated maps with 2D static map and the 3D STC techniques.

Despite not reporting any user study, Bach et al. (2014) generalize the concept of STC into a *conceptual representation that helps to think about temporal data visualization techniques*. From this generalization, the authors identify various operations that support the extraction of information in the form of other visualization techniques, including 2D maps, namely through time cutting and time flattening operations. These operations will be addressed in the next chapters, when discussing the effects of using interactive map planes and time granularity controls in the STC technique, respectively. Despite the valuable contributions of these works, particularly in the context of this research, these studies are considerably more focused in distinguishing the different map-based visualization components rather than exploring the actual consequences of their combination.

Finally, despite the fact that the majority of the existing works use animated or small-multiple maps in a 2D context, as suggested by the aforementioned studies, each frame from an animated and a small-multiple map can be represented in the form of a 2D static map or in the form of a 3D space-time cube. However, by applying animated features into these techniques, users will necessarily have to spend more time looking at the data, to visualize the animation, at least in their first interaction. On the other hand, as previously mentioned, the existence of animation supports the use of additional controls, such as pause or speed controls, to interrupt, and to in- or decrease the animation's speed. Therefore, the addition of animated features or the separation of a single map into various representations can significantly change how the data is represented temporally, and thus how the data is explored.

2.6 Summary

Trajectory data have become the focus of several studies, given their importance in decision making by various groups of users. As a consequence, it is fundamental to study and develop adequate techniques for the visualization of this type of data.

This chapter addressed the main concepts associated to this issue, based on five main sections. The first provided an overview regarding the concept of *cartographic interaction*, and established the context over which the following sections are related. The second looked into the definition and the main components related with trajectory data, namely the concepts of *space*, *time*, and *objects*. The third section focused on the main types of visual analysis tasks a user might need to perform to achieve a given goal. The results of this analysis highlighted some agreement between existing studies, since they identified similar tasks, but also some redundancy, since various identical tasks were presented with different names. As such, this chapter also presented a simplified and hierarchical model of tasks, based on their level of specificity and adequacy as visual analysis tasks. This simplified task model will be used in the following chapters as one of the basis for comparative user studies. The fourth section described the main types of techniques used to visualize trajectory data. Due to the spatial properties associated with this type of data, the target of this analysis was in high-level types of map-based techniques, particularly: *static maps*, *space-time cubes*, *animated maps*, and *small-multiple maps*. Finally, the fifth section described and discussed some of the most relevant comparative studies related to these techniques present in the literature.

Despite the various studies addressing the visualization of trajectory and spatio-temporal data, there are still several important challenges and open issues that need to be addressed. The analysis of the literature provides several examples of applications to visualize spatio-temporal and trajectory data; however, there is still a significant lack of empirical data to evaluate the usability of various approaches and to support the benefits of certain visualization and interaction features over others. Furthermore, although some studies have already been conducted, there is still some uncertainty regarding their results, due to somewhat contradictory findings and, sometimes, uneven procedures followed, which can give an advantage to certain visualization approaches over others. In addition, when evaluating a certain visualization, several usability studies tend to be addressed in *ad hoc* manner, not fully taking into consideration the different types of visual analysis tasks, and being either limited to case studies or focused on the most experienced users of the system under analysis. However, as previously stated, this type of data and visualizations are becoming increasingly available to a wider population of users, often inexperienced, yet interested, in visualizing and exploring this data in personal contexts (e.g., improving health, or sports/gaming performance).

This Ph.D. research aims to help minimizing these issues, by focusing on three different aspects: i) obtain empirical data to support the dis/advantages of existing techniques; ii) identify possible improvements; and iii) identify design guidelines for trajectory data visualization techniques. In particular, out of the four aforementioned types of map techniques, this work is focused on the usage of 2D static maps and 3D space-time cubes by inexperienced users in terms of data

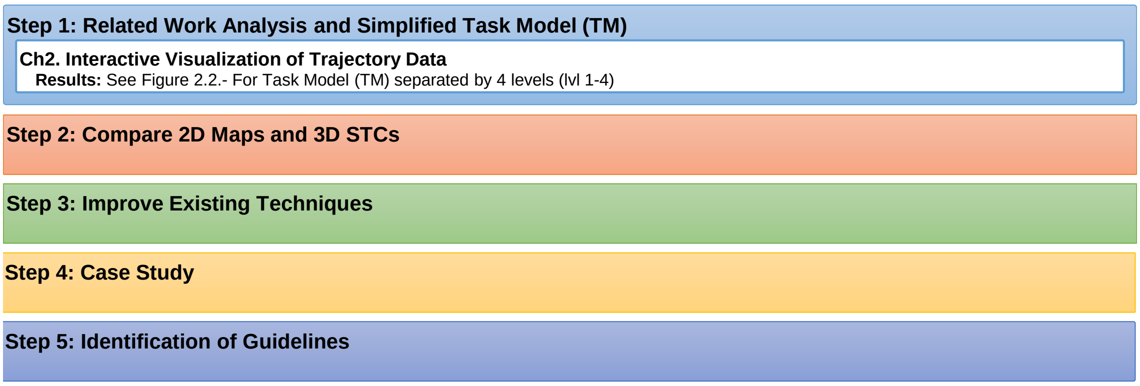

Figure 2.10: Summary of this research and the contributions of Chapter 2.

visualization and analysis. While 2D static maps are among the most well known approaches, the STC is a still lesser known approach that, despite its antiquity (Hägerstrand, 1970), raises several questions, generally associated to its 3D properties. Furthermore, animated and small-multiple maps can be applied over 2D or 3D/STC maps. Therefore, it is important to conduct more usability studies with these techniques, before considering the inclusion of animated or small-multiple features.

Figure 2.10 depicts the evolution of the scheme presented in Chapter 1 (Figure 1.3) alongside the role of the contributions of this chapter. These contributions address the first step of this research. The next chapter describes the first set of comparative studies between 2D static map and 3D space-time cube approaches with inexperienced users, which were used to obtain some preliminary feedback regarding the users' capabilities in using these techniques and to compare their adequacy in the aforementioned tasks for the visualization of trajectory data.

Chapter 3

Comparing 2D Static Maps and 3D Space-Time Cubes

It is important to evaluate existing map-based techniques for the visualization of trajectory data, taking into account the types of visual analysis tasks a user might need to address and the capabilities of inexperienced users in terms of data visualization and analysis. In the previous chapter, several map-based approaches to visualize trajectories have also been described alongside the interest towards two-dimensional (2D) static maps and three-dimensional (3D) space-time cubes (STCs).

This chapter describes the first two approaches addressing these challenges. The next section presents and evaluates ST-TrajVis, a prototype for the visualization of trajectory data, through the use of a 2D static map and a 3D STC. Based on the results obtained with ST-TrajVis, showing that inexperienced users may indeed be able to extract information from these map-based visualization techniques, the section after it presents a follow up study comparing two different prototypes based on 2D static map and 3D STC visualization techniques, to identify the types of tasks in which each technique is more adequate.

3.1 ST-TrajVis: Interacting with Trajectory Data

This section focuses on the design and evaluation of the ST-TrajVis prototype (Gonçalves et al., 2013b). The main goals of this study consisted in the preliminary assessment of the ability of inexperienced users to explore and analyse human trajectory data, alongside the identification of possible factors of interest to take into the following steps of this research. These include the different approaches users may adopt when interacting with these techniques, and the possible combination of these different types of maps. The following sections describe the trajectory dataset used by the ST-TrajVis and its main features in terms of visualization and interaction, followed by the procedure adopted for the user study, and ends with the analysis and discussion of the results.

3.1.1 The Data

The data used for the development and evaluation of ST-TrajVis consists of a subset of the trajectory dataset released as part of the GeoLife project (Zheng et al., 2008, 2009, 2010). The complete dataset contains, approximately, 17.621 trajectories of 183 users, recorded between 2007 and 2012. The trajectories were recorded by different GPS loggers and GPS-phones, with each point having a spatial or temporal distance of, approximately, 5 to 10 meters and 1 to 5 seconds, respectively. The various points recorded represent the locations associated to the users' outdoor movements during their daily routines. The majority of these trajectories were recorded in China, particularly in Beijing.

In the context of this experiment, the subset used was composed by the trajectories (approximately 40) of one randomly selected anonymized user, during the period of one month (November of 2008). Each entry in the original data is represented as a set of coordinates, namely latitude, longitude, and altitude (spatial information), with a timestamp associated to the moment in time in which those coordinates were recorded (temporal information). Based on these components, the dataset was further enriched with derived information regarding the mover's approximated speed and the estimated stop locations of the mover (i.e., the geographical locations where the user spent more time without moving). From a total of 49680 location points, the dataset was then simplified using a Douglas-Peucker algorithm (Douglas and Peucker, 1973) to approximately 8368 points.

3.1.2 ST-TrajVis Prototype

Figure 3.1 depicts a usage example of ST-TrajVis. As it can be observed, the prototype is divided into four main interactive components, namely a 2D static map, a 3D space-time cube, a data querying component, and a data enhancement component.

2D Static Map and 3D Space-Time Cube

ST-TrajVis combines an interactive two-dimensional map with a space-time cube visualization to represent the spatio-temporal and thematic properties of the data.

Each point's coordinates are represented through the visual variable location and are mapped to their correspondent position within the map and the STC's $x-$ and $y-$axis. Similarly, in the STC visualization, time is mapped as a location over the $z-$axis, with data recency being directly proportional to the STC's height. Therefore, the most recent events are located higher within the STC.

Due to the expected unfamiliarity of the users' with the application, a simple, yet representative, visual style was used to depict the trajectories. Each consecutive trajectory point is connected by a line, emphasizing the evolution of the mover's position over time. In addition, due to the known limitations of 2D maps in representing information from a temporal perspective, and even though this data component is already depicted in the STC, ST-TrajVis represents both the time periods of the day and data recency in the maps. To better emphasize the different periods of

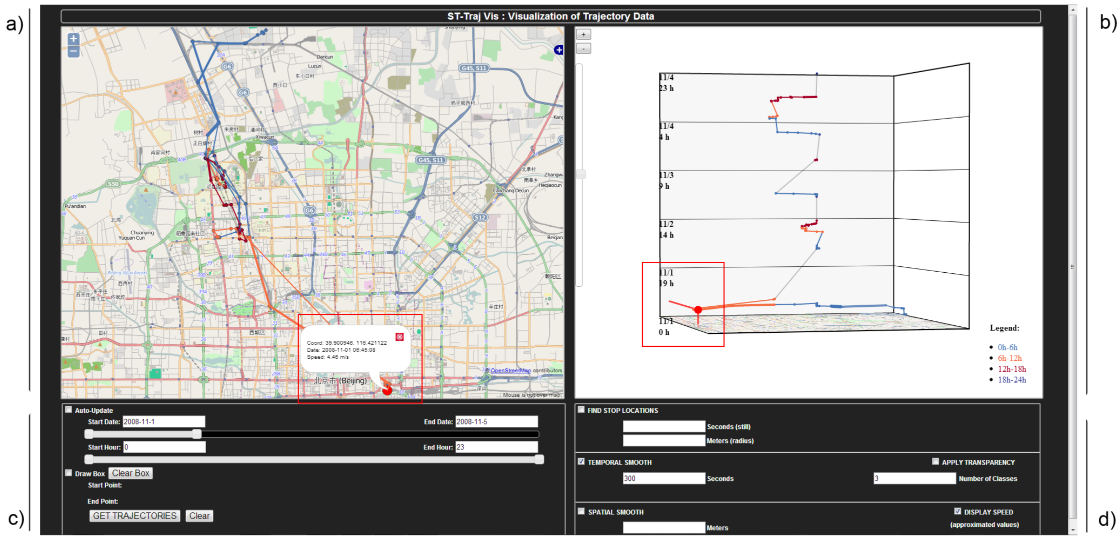

Figure 3.1: ST-TrajVis prototype. a) 2D map b) Space-time cube c) Data query component d) Data enhancement component.

the day, four categories were used, based on the hours of the day (0h-6h, 6h-12h, 12h-18h, and 18h-0h). Considering that colour is considered an useful visual variable for the representation of categories and the fact that people often associate meanings to specific colours (Bertin, 1967; Green, 1998), each point/line is coloured according to each of the four categories, where the periods during the day (6h-12h to 12h-18h) are represented with warm colours generally associated with the sun (orange and red, respectively) and the periods during the night (18h-0h and 0h-6h) are represented with cold colours generally associated with it (dark and light blue, respectively). In contrast, data recency is represented through the points' and lines' transparency, where less transparent (i.e., more visible) points and lines represent more recent data. This way, it should be easier for the user to identify, in the 2D map component, which locations were visited earlier than others, being also able to focus on the most recent information while having an overview of the previous movements displayed. The user can also de/activate this representation by using the data enhancement component. In addition, considering the usefulness of the visual variable size for the representation of numeric values (Bertin, 1967; Green, 1998), the approximated speed detected on each trajectory segment was represented with the point size/line thickness, where wider lines represent higher values and, consequently, faster movements.

For the representation of the mover's stop locations, ST-TrajVis uses point symbols, differently coloured from the trajectory lines, to capture the users' attention (Figure 3.2). In addition to their approximated location, each point also represents the estimated number of times the mover stayed at that location, where the larger the point, the more often that location was visited. In the STC, that information is displayed with vertical lines, whose length is proportional to the amount of time spent at that location. This information could be de/activated with the aforementioned data enhancement component.

Both visualizations provide interaction techniques based on panning and zooming. To min-

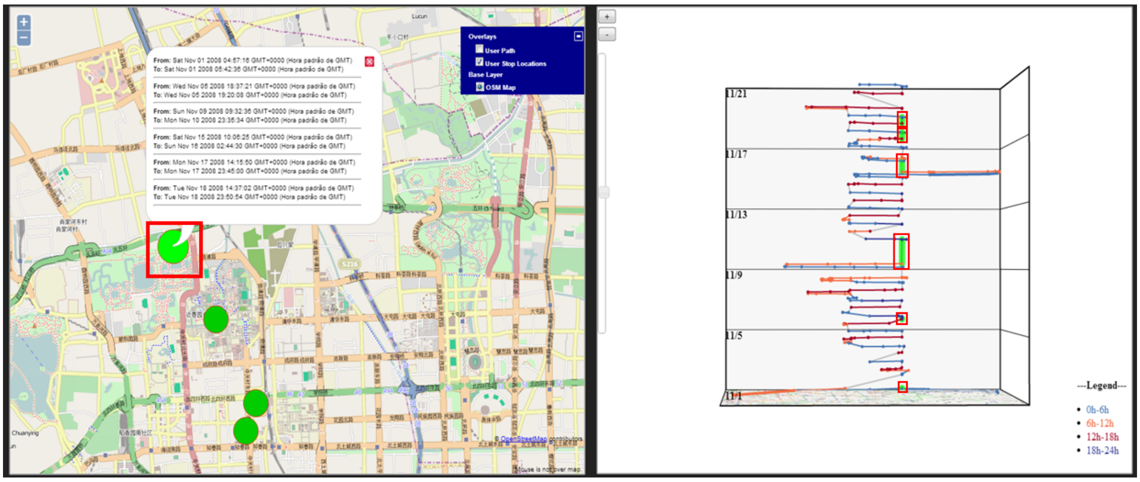

Figure 3.2: Visualization of stop locations with ST-TrajVis.

imize possible perspective issues with the STC, this technique supports also the rotation of the cube, changing its point of view (Kjellin et al., 2010b), and the change of the STC's temporal granularity, through the rescaling of the STC's height, using a slider on the left-side of the STC component.

Finally, both visualizations are linked with highlighting techniques. Specifically, when one point is selected in the STC, line indications will be drawn pointing towards the spatial and temporal locations associated to that point (highlighted with red rectangles in Figure 3.1 b). In addition, regardless of the map in which a point is being selected, the other visualization will highlight the corresponding point, providing additional information about it in a pop-up window pointing towards the selected location. If the selected feature is a trajectory point, ST-TrajVis describes its absolute position, date, and the approximated speed of the mover (highlighted with red rectangles in Figure 3.1 a). Consequently, no data was lost when converting some of the numerical properties of the data into categories. If the selected feature is a stop location, the prototype displays a list of the time intervals in which the mover visited that location (highlighted with red rectangles in Figure 3.2).

Data Querying and Enhancement

The data query component (Figure 3.1 c and Figure 3.3 a) allows filtering the dataset according to spatial and temporal properties before visualizing it. More specifically, by selecting the option *Data Box* and then drawing a rectangular area within the 2D static map component (Figure 3.1 a) the user is able to delimit the geographical area to visualize the trajectories (i.e., any location visited outside the defined area will not be represented). The user can also delimit the period of time in which he/she is interested, by defining the start and end dates of movement, alongside the period of hours to be visualized (e.g., visualize the trajectories between 12h and 17h).

The data enhancement component (Figure 3.1 d and Figure 3.1 b) supports the extraction and representation of additional information, in the form of movement speed and stop locations, as

Figure 3.3: ST-TrajVis a) Data query component and b) Data enhancement component.

well as the customization of some visual properties of the trajectories presented in both map-based visualization techniques. Regarding the mover's speed, the user is able to select how many categories of transparency should be used (e.g., 3 categories, where the first is fully visible, the second is one-third transparent, and the third is two-thirds transparent). Regarding the search for stop locations, the user is able to define the search parameters of maximum radius (in meters) and minimum stop-time (in seconds) – i.e., the minimum time spent within an area of a given radius. Regarding the thematic information of speed, the user can de/activate its representation (explained in the previous section). This component also supports the visual smoothing of the trajectories, based on the temporal and spatial distances between points, for instance, display the visited locations at a minimum distance of 20 meters from each other and at a minimum temporal distance of 5 minutes.

3.1.3 User Study

A usability study with ST-TrajVis was conducted with the participation of 5 volunteers, aged between 23 and 28 (Av: 25.6, SD: 2.5). All participants were experts with computer applications and were familiar with map-based applications, such as Google Maps to search and visualize trajectories for specific points of interest, or Google Earth, for the visualization of points of interest on a 3D perspective. Despite that, none of the participants had any significant experience with the location displayed (China-Beijing) nor with spatio-temporal data analysis and visualization. Consequently, the results of the study should not be affected by the participants' possible bias or previous knowledge.

Tasks

Considering the dataset used for the experiment, ST-TrajVis should support the description and the comparison of the dynamics of the represented mover's characteristics through time. For that, the participants were asked to conduct 6 tasks. The first two tasks aimed to assess the participants' ability to interact with the main search mechanisms to obtain the necessary trajectories for analysis. These tasks 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 four remaining tasks aimed to assess the participants' ability to explore and interpret the different types of information provided by the visual methods being used. Considering the fact that this is the first study in the context of this work and that it is expected that the participants either *identify* or *compare* the various visual features presented in the maps to extract information, these tasks belong to the first level of visual analysis tasks, described in Section 2.3 (Figure 2.2). In the third task (*T3*), the participants had to discover in which days the represented mover had the most activity. The fourth task (*T4*) required the participants to find the regions in which the represented mover had spent the most time over. In the fifth task (*T5*), the participants were asked to list the days in which a faster movement was detected. Finally, in the last task (*T6*), participants had to discover in which hours of the day the mover had been the most active.

The participants were also asked to answer to a small survey, with 8 questions, to collect their opinions. More specifically, they were asked to classify certain sentences on a 0 to 10 scale (0 being not useful, 10 being crucial), namely:

***Q1*:** how useful were the application's features, overall, for the completion of the tasks;

***Q2*:** how useful was the representation of periods of the day (colour);

***Q3*:** how useful was the representation of data recency (transparency);

***Q4*:** how useful was the representation of speed (line thickness);

***Q5, Q6*:** and how useful were the 2D map and the STC for the completion of the tasks.

In addition, the participants were also given open questions about: (*Q7*) the most appreciated features; and (*Q8*) any additional features they thought would help them completing more easily the tasks.

Procedure

All participants performed the tasks and answered to the questionnaire individually in a controlled environment.

At the beginning of the experiment, the participants were informed about the objectives of the study. After that, all of ST-TrajVis' features were demonstrated and explained. Then, each participant had some time to practice with the features of the application, get used to them, and clarify any doubts. After the practice session, each participant performed the six tasks and answered to the questionnaire. During the execution and at the end of the tasks and experiment, the participants were invited to *think-aloud* and share their opinions.

To measure the results, in addition to the participants' comments, the following dependent variables were recorded:

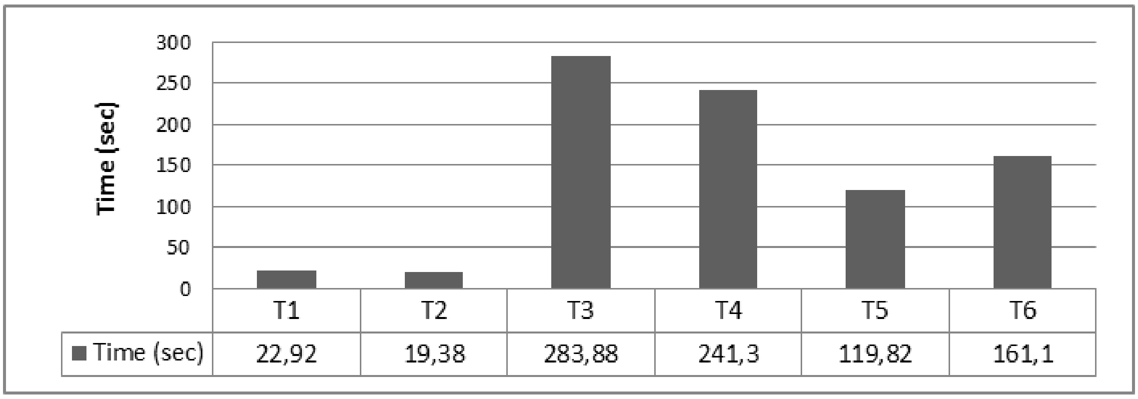

Figure 3.4: Mean task completion times for each task.

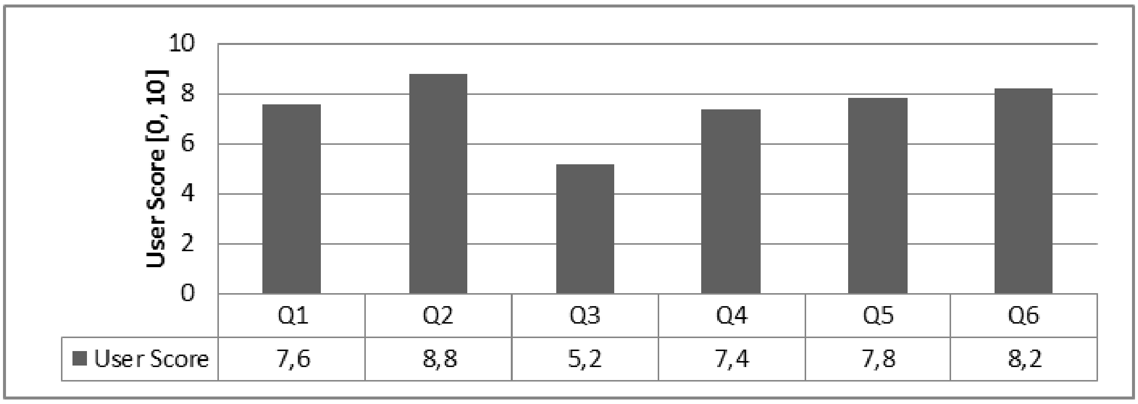

Figure 3.5: Mean participant scores, for the first six questions of the questionnaire.

Task completion times: to measure how efficiently the participants complete the tasks;

Task accuracy: to measure how effectively the participants complete the tasks;

Questionnaire answers: to analyse the participants' acceptance towards the prototype's features and discuss possible ways of improving them.

Results

Given the small number of participants, only simple statistical analyses were possible to be conducted over the results. Figures 3.4 and 3.5 display the participants' mean completion task times, and the mean scores given in the questionnaire at the end of the tasks.

All participants, except one, completed the tasks successfully and accurately. Given the feedback of the participant that did not finished the tasks, it was concluded that he was not able to finish tasks *T3* and *T5* due to technical performance issues and, consequent, perspective issues with the STC component. More specifically, the computer had a slower performance overall, which affected the application and made it interacting with the STC less responsive and noticeable.

From the analysis of Figure 3.4 and the results of tasks *T1* and *T2*, it can be argued that the prototype was intuitive for the configuration and request of trajectory data. Furthermore, the results also suggest an overall increasing performance from task *T3* to *T6*, as the participants were able to complete the tasks generally faster from task to task. This result suggests that, despite the differences between the tasks, there may have been a learning effect, that alongside the knowledge obtained in the training stage, helped the participants performing more easily the tasks. This is also supported by the comments of three participants, which mentioned that, in addition to consider the interface's features easy to use, there was an overall small learning curve associated to the prototype's features.

The analysis of the participants' scores for the first six questions (Figure 3.5) suggests an overall positive feedback towards the application's features, even if less expressive in question *Q3* (*How useful was transparency for the representation of the trajectories' recency?*). Overall, these results suggest that, out of the various visual variables, the use of a colour classification code for the representation of temporal periods was the most useful (*Q2-Q4*).

In addition, despite the small difference in scores, three participants have shown a higher preference for the STC over the 2D map (*Q5-Q6*), commenting that the STC was their *main visualization* to perform the tasks, and that the 2D map, although useful, was mostly used as a complementary tool.

Finally, the participants mentioned that the features they appreciated the most were the display of stop locations, followed by the optional representation of speed, and the filtering and smoothing methods (*Q7*). Nevertheless, they also commented being interested in having more tools to filter the data and interact more with the map visualizations (*Q8*).

3.1.4 Discussion

Even with the small number of participants, the results of this study highlight some relevant aspects and issues associated with the application and its features.

Despite the difficulties felt by one of the participants with the STC, this technique received a very positive feedback by all others, even being preferred by some over a 2D map. This result is both interesting and unexpected, given the aforementioned disadvantages often associated with this technique, alongside the participants' unfamiliarity with STCs and the 2D map's popularity for representing georeferenced data.

In addition to the small learning curve associated with the application, the results suggest the usefulness of the STC for the visualization of trajectory data. Nevertheless, it is plausible to assume that the results may have been influenced by the fact that ST-TrajVis provided also a 2D map, which may have improved the value of the STC to the participants. This is supported by the comments of some participants, which mentioned using each map differently, as with the STC as their main visualization and the 2D map as their secondary visualization technique. Ultimately, the results further encourage the conduction of additional comparative studies, to explore the differences between these two techniques and, by learning how adequate they are individually in the different types of tasks, to assess how they can benefit one another.

Another interesting result is related to the features that were used or suggested. Overall, participants have shown interest in dynamically changing the visualization, namely through the de/activation of the representation of some attributes. However, in comparison, the controls related to the spatio-temporal smoothing features, the use of transparency to represent the data's recency, and the STC's temporal granularity were rarely used. This could suggest that the participants did not feel the need to use them, or that these features must be presented in a more intuitive way.

It should be noted that the dataset used in this experiment describes the trajectories of one mover. This factor is particularly important, since one of the reasons presented by the participants regarding their preference for the STC was the uncluttered visualization of the trajectories, in comparison to the 2D map. However, if the STC displayed the trajectories of more movers, detected at the same time, there would, necessarily, be some visual cluttering. As such, this feedback emphasizes the need to compare these techniques with a larger volume of trajectories and with different complexities, to assess their scalability.

Following the results and feedback obtained in this experiment, the following section describes a second user study focused on the comparison of separated 2D map and 3D STC prototypes for the visualization of trajectory data.

3.2 Visualizing Human Trajectories: Comparing Space-Time Cubes and Static Maps

This section describes a comparative study between two prototypes based on a 2D static map and a 3D STC for the visualization of trajectories (Gonçalves et al., 2014). The main goals of this study consisted in assessing the usability of these techniques with an inexperienced population, detect possible interaction barriers with these techniques and, ultimately, compare how adequate they are in different types of visual analysis tasks. The following sections address the dataset used in the study and the prototypes developed together with their most relevant features. This is followed by the description of the user study conducted with the prototypes. This section ends with an analysis and discussion of the results obtained in this experiment.

3.2.1 Dataset

Similarly to the previous study, the data used in this experiment consists of a subset of the trajectory dataset released as part of the GeoLife project (Zheng et al., 2008, 2009, 2010). However, in this study, the subset was composed by the trajectories of 32 movers randomly selected, recorded during the period of one year (2008-2009). In addition to the spatial and temporal location attributes associated with each trajectory point, some of the recorded trajectories also have a transportation mode associated, namely: walk, bike, bus, car/taxi, train, and airplane. Based on these attributes, the dataset was enriched with additional information, by estimating the movers' speed, at each point. Consequently, based on the average speed obtained along the trajectories, the transportation mode associated to each trajectory point was estimated, when that information was

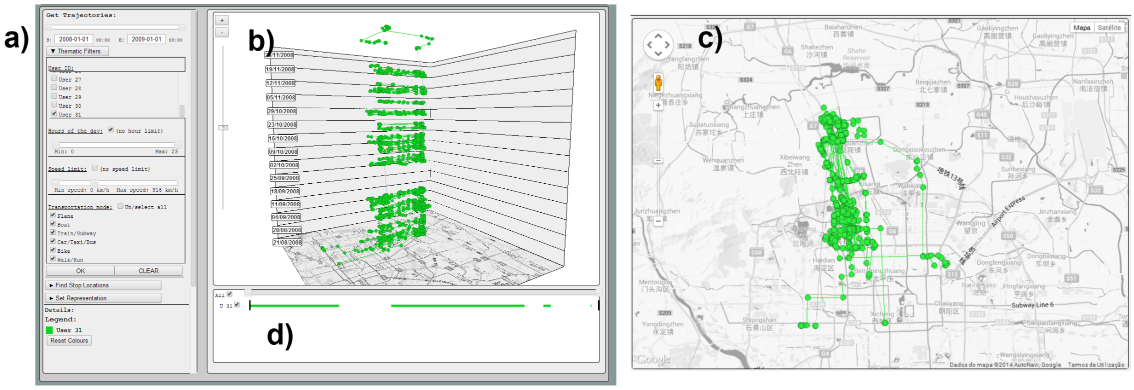

Figure 3.6: HTV prototypes used for the comparative study. a) Control panel b) Space-time cube visualization, c) 2D static map visualization, and d) Interactive timeline.

unknown. Comparatively to the subset used in the previous experiment, the one used in this study was composed by, approximately, 88862 points.

3.2.2 Prototypes

To compare the differences between 2D maps and 3D STCs, in terms of how these techniques help users solving general analysis tasks, two prototypes, called HTV-2D Map and HTV-STC (HTV - Human Trajectory Visualizer), were developed, supporting the visualization and interaction with the data that was previously presented. Both prototypes were developed in HTML5, by taking advantage of the following main libraries:

TRAJMAP2DJS - a JavaScript API developed in the context of this work to help the visualization of spatio-temporal and trajectory data on 2D static maps. This API takes advantage of the Google Maps (Google, 2005) and three.js (Cabello, 2010) APIs to display information, by supporting the creation of additional layers over Google maps based on WebGL, allowing the representation of larger volumes of information comparatively to what Google Maps normally supports.

STCJS - a JavaScript API developed in the context of this work for the visualization of spatio-temporal data in STC, which also takes advantage of the three.js API. Similarly to other APIs for map visualization, this supports the creation of various layers of information, and the attribution of visual styles to the features representing information in those layers (e.g., colour and size). Furthermore, similarly to TRAJMAP2DJS, this API also supports the notification of user interaction events, such as when a user selects a map feature or changes the STC's properties (e.g., height rescaling or map pan/zoom).

Figure 3.6 depicts the interface of the two prototypes, and the two types of maps used. Specifically, both prototypes are divided into three main sections: a control panel, a map visualization, and an interactive timeline. The control panel section (Figure 3.6 a) provides a set of features to select which data will be visualized according to temporal and thematic attributes. Specifically, the user can filter the data based on the start and end dates (and hours) of movement, the identification of the movers', a speed interval, and the movers' estimated means of transportation. The control panel also allows the estimation of stop location events from the observed trajectories (i.e., locations with a given radius, where movers spent a given amount of time). In addition, the user is able to customize how the data is represented, by associating specific visual variables to data attributes, in turn, determining what should be the focus of the visualization. These attributes include: the movers' identification, the time periods of the day in which was detected movement, the movers' estimated speed, the estimated transportation mode, and the number of visits and time spent at a stop location.

Depending on the prototype, the visualization section is composed by either a 3D space-time cube or a 2D static-map (Figures 3.6 b and c, respectively). Similarly to the previous study, due to the expected lack of expertise from the test participants in analysing trajectories, it was necessary to provide simple representations of the data, while still making the prototypes representative of the corresponding groups of techniques and avoiding the inclusion of too many features. This way, the objective consisted in minimizing the cognitive workload associated to the understanding and feature memorization of the prototypes. In addition, though this approach, the analysis of the users' performance is significantly more dependent on the most relevant properties of each visualization rather than on the use of external features.

In both visualizations, the trajectories of movers are represented with sequences of circular points connected with lines, coloured according to the temporal or thematic attributes selected in the control panel section. Considering the fact that colour is considered to be an adequate visual variable for the representation of categories (Bertin, 1967) and due to the possible repetitions of coloured objects across the spatial and temporal representations, it is plausible to assume that representing these attributes categorically may help users to detect movement patterns and cyclic behaviours more easily. On the other hand, stop location events are represented as squares, coloured differently from the trajectory points. Since object size is considered as an adequate variable for the representation of quantitative information (Bertin, 1967), stop location points are also scaled according to the attributes selected in the control panel, i.e., the higher the attribute value, the larger the point. Figure 3.7 depicts an example of the visualization of attribute data and stop locations on both map components. In both visualizations colours represent the time period of the day the points where detected (orange: 6h-12h, red: 12h-18h, blue: 18h-24h). On the left, in the 2D static map visualization, the light blue squares overlaying the trajectory points represent the estimated stop locations from the observed mover (scaled according to the number of visits to each location), which correspond to the vertical line segments of the movement depicted on the space-time cube visualization (on the right).

Considering the limitations generally associated with static maps in the representation of temporal information (Andrienko et al., 2011), the 2D map visualization lacks support for the representation of the temporal dimension of the data, even if with the optional representation of the time periods of the day. This way, the 2D map and the 3D STC could not be considered as *in-*

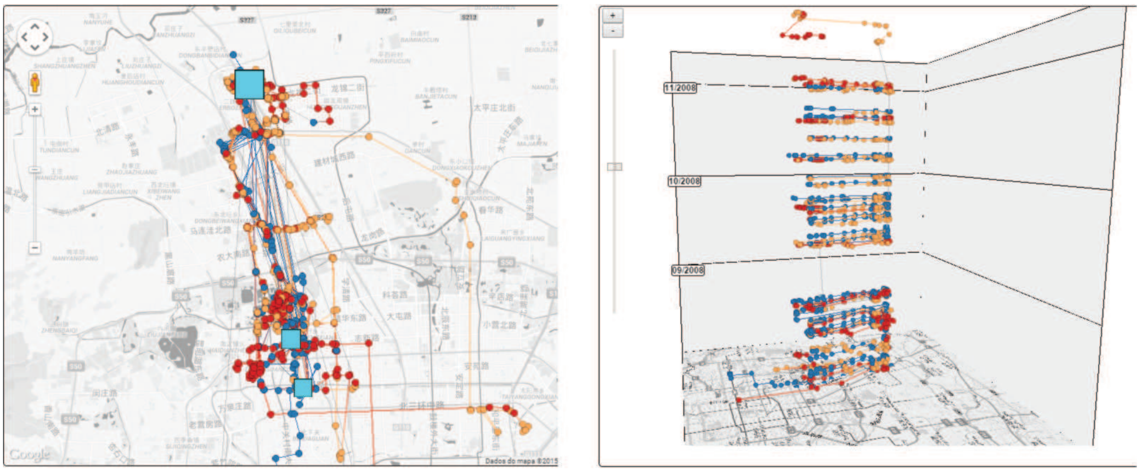

Figure 3.7: Visualization of attribute data and estimated stop location points.

formationally equivalent, i.e., it would not be possible to extract the same information from both visualizations (Tversky et al., 2002). This could compromise the validity of the results obtained from the comparison of both techniques. To prevent that, both prototypes provide a timeline representation to emphasize the temporal dimension of the data and to represent the activity of movers over time (Figure 3.6 d). Each mover and stop location event is represented as a dashed line, where each gap represents a period of inactivity, i.e., when no movement was detected.

Similarly to ST-TrajVis, both visualizations provide common interactive features, such as panning and zooming. In particular, the STC also supports the rotation of the cube, along any of its axes, allowing the user to change its point of view. It is also possible to rescale the cube's height, which allows manipulating the temporal representation of the data (similar to a temporal zoom). Contrarily to what had been suggested in previous works (Kraak, 2003), in this study the STC does not allow moving the spatial plane along the temporal axis. Although, in theory, moving the plane may help users obtaining more easily spatial information about a given point, it will necessarily overlap all data below that point. Consequently, the same effect can be obtained by filtering the data temporally (i.e., showing only the trajectory points since a given date). In addition, further empirical studies, as addressed in Chapter 4, should be conducted regarding this feature to evaluate its effects on the interaction with a STC.

The map visualization techniques and the timeline are connected through highlighting mechanisms (Robinson, 2011). By selecting any representation in the maps, the corresponding time moment (or period, if a stop event is selected) is highlighted over the timeline. Similarly, when the timeline is selected, the maps will highlight the data corresponding to that time. Figure 3.8 depicts an example of this action, where a line between two trajectory points (approximately at the same location) is being highlighted in the STC and, since no movement seems to have been detected during that period, the corresponding period of time is also highlighted in the timeline.

The timeline representation also supports a quick interactive filtering of the data, with a bi-directional drag bar, at the top of the representation, which allows the delimitation of the time

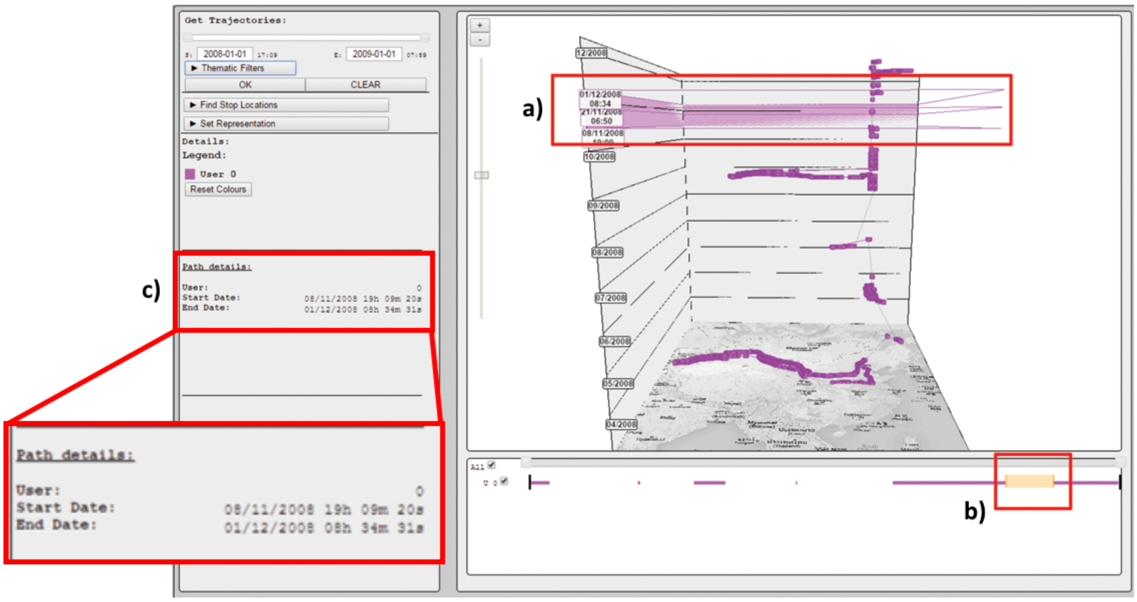

Figure 3.8: Example of a line between two trajectory points being higlighted a) in the STC; b) in the timeline; and c) the available information about the highlighted feature.

period to be visualized in the timeline and in the map. Finally, whenever a visual feature in the map or the timeline is selected, a description of all information associated to that feature is displayed (e.g, where and when was the feature detected, alongside its thematic attributes), bellow the control panel section and the map legend (Figure 3.8 c).

3.2.3 User Study

To assess how adequate 2D static maps and 3D space-time cubes are in helping inexperienced users completing general visual analysis tasks, a usability study was conducted using the aforementioned HTV prototypes to compare the user performance and acceptance towards the techniques.

The study had the participation of 16 volunteers, aged between 19 and 47 (Av: 23.33, SD: 7.7), in which four participants were Ph.D. students, while the others were undergraduate students in computer science. Despite their expertise with computer applications, and their familiarity with applications to search for directions towards specific points of interest (e.g, Google Maps/Earth), none of them were familiar with trajectory data analysis. Moreover, none of the participants were familiarized with the geographical locations displayed, which in turn prevented any previous knowledge regarding the locations from adulterating the study's results.

Based on the features of each prototype and the characteristics of the selected population, this study's hypotheses were the following:

H1 Participants will prefer the 2D static map visualization, due to the high probability of being most familiarized with this technique;

H2 For the same reason, and considering the limitations in human perception of 3D graphics, the 2D map visualization will be more effective; and

H3 Due to the smaller number of points/lines displayed, *elementary tasks*, i.e., tasks that focus in just one mover or just one time moment, will have a better performance and will be easier to solve.

Tasks

As previously stated, the different types of visualization objectives can significantly change the context of a given usage scenario. Based on the previous analysis of this factor (see Figure 2.2), the objectives present in the second level were selected as the basis of this experiment (i.e., *locate*, *identify*, *compare*, and *associate*). For a full comparative study, the third, and most detailed, level of objectives could be used. However, this would imply that a large number of objectives would have to be combined with the different visualizations, and additional independent variables. That would significantly increase the number of tasks and the cognitive workload associated with the experiment, which could cause a higher fatigue to the participants and adulterate the results. In addition, the first level of aforementioned simplified model (i.e., composed by the objectives *identify* and *compare*) would be too general in the context of this study, given the importance of distinguishing the ability of users finding and interpreting information. This can be associated to the ability of distinguishing space from other attributes in cartographic visualization (i.e., within *locate* and *identify* tasks).

To prevent an overwhelming number of tasks in this experiment, the four categories of queries associated with spatio-temporal data, presented by Andrienko et al. (2003), were adapted for this study. Specifically, two search levels (*elementary* and *general*) and two search targets (*what+where* and *when*) are considered, supporting four views over the trajectory data, namely the analysis of one or several trajectories, in a time moment or a time period. It is important to notice that, unlike Andrienko et al. (2003) or Andrienko et al. (2008), these search levels are being considered as directly related to the number of elements involved, thus being more adequate and easier to control in the context of this study, comparatively to, for instance, a data-mining analysis context.

Ultimately, combining the different types of objectives with these four spatio-temporal categories provides a list of tasks to be completed by the participants. Table 3.1 lists some examples of those tasks. As a consequence of this combination and simplification of factors, and as depicted in Table 3.1, the visualization objectives *compare* and *associate* were not considered in the *elementary when elementary what+where* category. By considering the *elementary* search level as the visualization of an individual data feature, *elementary when elementary what+where* tasks refer, necessarily, to a single trajectory in a time moment, thus making it impossible to establish any association or comparison with other features. In order to do that, it would be necessary to either analyse more trajectories (*general what+where*) or analyse more instances of the same trajectory during a time period (*general when*).

Procedure

In this experiment, two independent variables were considered:

Table 3.1: Examples of tasks addressing spatio-temporal data categories and visualization objectives.

Spatio-temporal Category	**Visualization Objectives**	**Example**
Elementary When Elementary What+Where	Locate	Where is mover 1 at 9 AM?
	Identify	Which mover was at region X at 9 AM?
General When Elementary What+Where	Locate	Where was mover 1 between 9-10 AM?
	Identify	How fast was mover 1 travelling between 9-10 AM ?
	Compare	Which transportation mode did mover 1 take the most?
	Associate	Is there a relation between the mover's speed and the day hours? If so, how?
Elementary When General What+Where	Locate	Where are most movers located?
	Identify	Who stayed outside region X?
	Compare	Who travelled the fastest?
	Associate	Is there any relation between the movers' speed and the type of road? If so, how?
General When General What+Where	Locate	Where are the most visited locations?
	Identify	Who travelled outside region X?
	Compare	Did mover 1 travelled more often than mover 2?
	Associate	Is there a relation between the number of visits at region X and the day of the week? If so, how?

Visualization Technique (*Vt*), with two levels, corresponding to the previously described prototypes, 2D map and STC;

Query Category (*Qc*), with four levels: *Elementary what+where elementary when* (*EE*);
Elementary what+where general when (*EG*);
General what+where elementary when (*GE*), and
General what+where general when (*GG*).

The experiment followed a *within subjects* design and all participants carried out each task individually, in a controlled environment. At the beginning of the study, subjects were briefed about the objectives of the experiment, and they viewed a demonstration of the interfaces. Before carrying out the tasks, they were asked to interact with the applications, and encouraged to clarify any doubts.

After the the training phase, the participants performed the tasks (*locate, identify, compare*, and *associate*) under the different categories, as exemplified in Table 3.1. At the beginning of each task, and whenever the data was updated with the control panel, the visualizations would display all available data, i.e., all trajectories would be visible on-screen, to prevent unnecessary interaction to search for off-screen information. In particular, for the STC technique, the view would always be presented with the same initial point of view (similar to Figure 3.8). To mitigate sequence effects, the order of the tasks and the type of visualization technique was counterbalanced based on a *latin-square* design. At the end of the experiment, each participant performed a total of 28 trials: (2 visualization techniques) x (4 *locate* + 4 *identify* + 3 *compare* + 3 *associate*).

To measure the results, a common set of dependent variables employed in usability evaluation were considered, which are derived from the ISO standard of usability $9241 - 11$ (ISO, 1998), namely:

Task completion times: to measure the participants' efficiency in completing the various tasks. This variable was measured in seconds, since the user started the task until a final answer was given.

Task accuracy: to measure the participants' effectiveness in completing the tasks, based on the their answers to the questions. Each answer was classified with a 0-10 scale, depending on the accuracy and detail of the participants' answer. For example, in *location* tasks, an answer with the correct location point of a given trajectory would be ranked higher than a point in the vicinity, which would, in turn, be ranked higher than an approximated area of the correct location point; on the other hand, in *association* tasks, an answer that showed all the areas of the city visited and the corresponding days of the week would be ranked higher than one that failed to point out all relevant areas or simply pointed out the most usually visited locations.

Subjective preferences: to analyse the participants' personal preferences and opinions regarding the techniques. This variable was recorded based on the users' preferred technique and based on their comments throughout and after completing the tasks, as participants were encouraged to *think aloud*, and share their suggestions and opinions about the techniques.

Results

This section describes the results obtained in terms of the participants' efficiency, accuracy, and subjective opinions.

Task Completion Times For the comparative analysis of the participants' task completion times, the results were subjected to the Shapiro-Wilk test of normality. Since no data transformation could provide a normalized dataset, a non-parametric approach was used by applying the Aligned Rank Transform (ART) method, followed by a full factorial ANOVA (Wobbrock et al., 2011). When significant effects were detected, the Bonferroni test for pairwise comparisons was applied.

Figure 3.9 shows the participants' mean completion times for all tasks, with all combinations of the two independent variables. In the *locate* task, the tests revealed a significant effect from

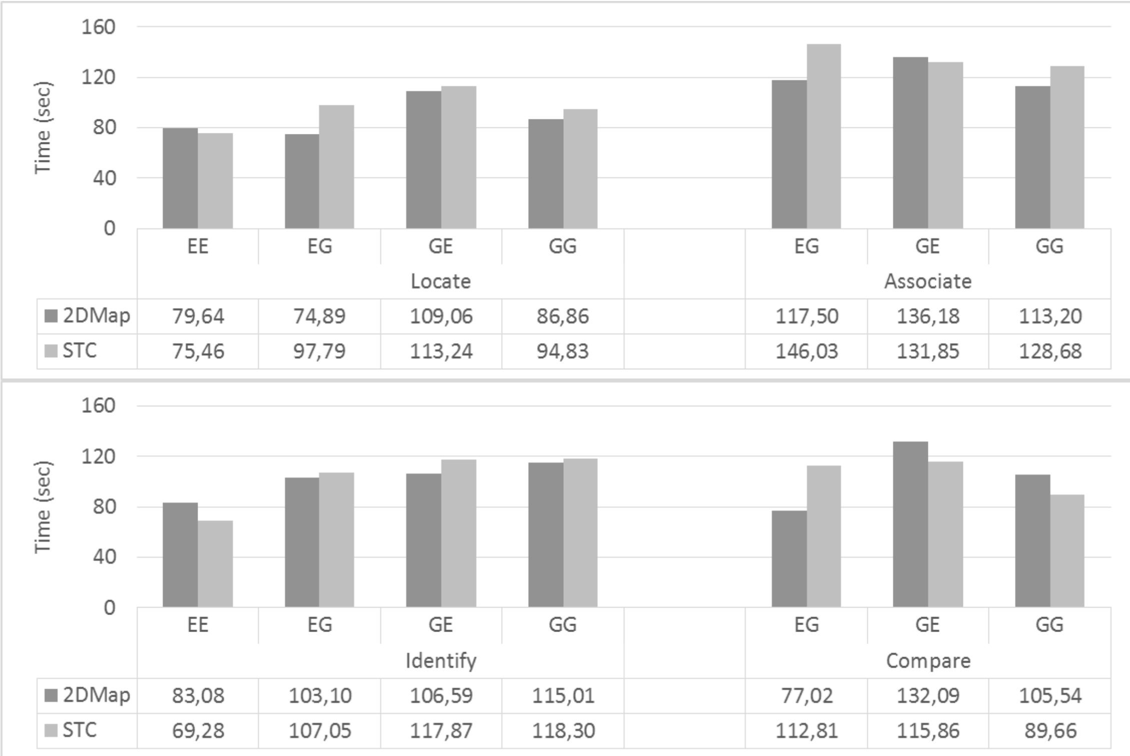

Figure 3.9: Mean participants' task completion times, with the two visualizations, and in all configurations.

Vt ($F(1) = 4.991, p = 0.041$) and from *Qc* ($F(3) = 5.321, p = 0.003$) on task completion times. The pairwise comparisons test revealed significant differences between the level *Qc(GE)* and the *Qc(EE)* and *Qc(EG)* levels ($p = 0.005$ and $p = 0.006$, respectively). A significant interaction between *Vt* and *Qc* was also detected ($F(3) = 2.811$, $p = 0.05$). In the *identify* task, the tests revealed a significant effect of the *Qc* level on task completion time ($F(3) = 10.135$, $p<0.001$), in particular, between the level *Qc(EE)* with all others($p = 0.001$, comparing with *Qc(EG)* and *Qc(GE)*, and $p = 0.01$, comparing with *Qc(GG)*). On the contrary, no significant effects were detected on the *compare* and *associate* tasks.

Task Accuracy To compare the participants' accuracy throughout the experiment, a Friedman's test was used. When significant differences were detected, it was conducted a Wilcoxon Signed Rank test with a Bonferroni correction, for pairwise comparisons. Table 3.2 displays the average scores obtained by the participants in the various tasks. The tests revealed significant results in the *locate* task ($X^2(7) = 56.593$, $p<0.001$), where the participants were generally more accurate with the 2D map visualization, and in the *associate* task ($X^2(5) = 29.446$, $p<0.001$), in which, conversely, the participants were more accurate with the STC visualization. Table 3.3 shows the most statistically significant differences that were detected, and those close to be considered as such (i.e., not significant, but still with a low p-value), between the various factors, in pairwise comparisons tests.

Table 3.2: Mean participants' accuracy in the various tasks, with the two visualizations, and in all configurations.

	Locate				Associate		
	EE	**GE**	**EG**	**GG**	**GE**	**EG**	**GG**
2D Map	9.56	8.94	9.38	9.31	6.25	6.19	6.47
STC	9.13	7.50	7.44	7.75	7.59	7.06	8.06
	Identify				**Compare**		
	EE	**GE**	**EG**	**GG**	**GE**	**EG**	**GG**
2D Map	9.56	9.44	8.63	8.94	8.81	8.50	9.19
STC	9.81	9.31	8.56	9.31	8.88	8.94	8.94

Table 3.3: Most significant pairwise comparisons regarding the participants' accuracy in the *locate* and *associate* tasks.

	Locate			Associate		
	STC EG	**STC GE**	**STC GG**	**STC EG**	**STC GE**	**STC GG**
2D Map EE	$Z = -3.13$ $p = 0.002$	$Z = -3.20$ $p = 0.001$	$Z = -3.06$ $p = 0.002$	-	-	-
2D Map EG	$Z = -2.69$ $p = 0.007$	$Z = -2.36$ $p = 0.019$	$Z = -2.47$ $p = 0.014$	$Z = -2.95$ $p = 0.003$	-	$Z = -3.21$ $p = 0.001$
2D Map GE	$Z = -3.10$ $p = 0.002$	$Z = -3.33$ $p = 0.001$	$Z = -3.26$ $p = 0.001$	-	$Z = -2.04$ $p = 0.042$	$Z = -2.96$ $p = 0.003$
2D Map GG	$Z = -3.09$ $p = 0.002$	$Z = -3.24$ $p = 0.001$	$Z = -3.24$ $p = 0.001$	-	-	$Z = -3.38$ $p = 0.001$
STC EE	$Z = -3.26$ $p = 0.001$	$Z = -3.24$ $p = 0.001$	$Z = -2.65$ $p = 0.008$	-	-	-

Subjective Preferences A similar procedure was applied over the results to compare the differences between the participants' opinions after each task and at the end of the experiment.

Although no significant differences were found between the participants' opinions throughout the tasks, at the end of the experiment the majority shown a significantly higher preference ($Z = -3.384$, $p = 0.001$) towards the 2D map (Av: $8.38/10$) over the STC visualization (Av: $7/10$).

3.2.4 Discussion

The results obtained in this experiment support this study's hypotheses and highlight relevant aspects regarding 2D static maps and 3D space-time cubes.

Despite the participants' lack of experience in terms of trajectory data analysis, all tasks were completed in a generally small amount of time and with a high accuracy. In addition to the positive feedback given by the participants regarding both prototypes and their features, these results highlight the adequacy of these techniques for the visualization of trajectory data by an inexperienced population. Nevertheless, it is important to notice that, despite the STC's positive results, at the end of the experiment the majority of the participants have shown a significantly higher preference towards the 2D map visualization. These results support the first hypothesis, *H1*: participants would show a higher preference for the 2D map visualization.

The participants were able to complete the *locate* tasks with a higher performance using the 2D map rather than the STC visualization, with the exception of the *Qc(EE)* level. Consequently, these results sustain the second hypothesis, *H2*: participants would be more effective with the 2D map visualization. However, contrary to the initial expectations, the participants were significantly more accurate with the STC visualization in *associate* tasks, when searching for possible relations between the data components. Regarding this particular situation, some participants pointed out that, despite not being familiarized with the STC, they considered that its third dimension helped them understanding better what the movers have done, when, and how often they did it.

It can be argued that these results are in agreement with previous research suggesting that, while 3D visualizations enable a better understanding of objects' shapes, 2D visualizations are more appropriate in location-related tasks (John et al., 2001). As previously discussed, not only *location* tasks can be considered as being strongly related to the spatial components of trajectory data, but static maps, by themselves, also emphasize this component over time, which may justify the users' superior performance in these tasks with this technique. By comparison, due to their 3D properties, STCs can be heavily dependent on interaction techniques, like view rotation or, as further discussed bellow, the cube's height adjustment, to allow the user to properly visualize the data. Consequently, regardless of how the view is first presented, this may force an STC user to spend more time interacting with this technique than with a 2D map, which is in agreement with the obtained results. In addition, although STCs may suffer from perspective issues, due to their 3D properties, they share the same visual variable (position) to represent two different data components, space and time. Therefore, considering that the shapes presented in the STC (i.e., trajectory lines and event points) represent information, including the relation between the spatial and temporal components of the data, it is plausible to assume that users will be more effective in understanding that information and find more easily associations between data components with STCs.

Regarding the *identify* and *compare* tasks, the results did not reveal any significant differences between the techniques. These results can be justified given the fact that, in both visualization techniques, the data was encoded with the same visual variables (i.e., the same colours and sizes to represent the same categories and values).

Concerning the *query categories*, the results revealed that the participants were significantly slower in the location tasks with the *Qc(GE)* level, when compared to the *Qc(EE)* and *Qc(EG)* levels. It is plausible to assume that the existence of several points representing different movers spread throughout a map (*GE*) required more time to be analysed than just one point (EE) or a

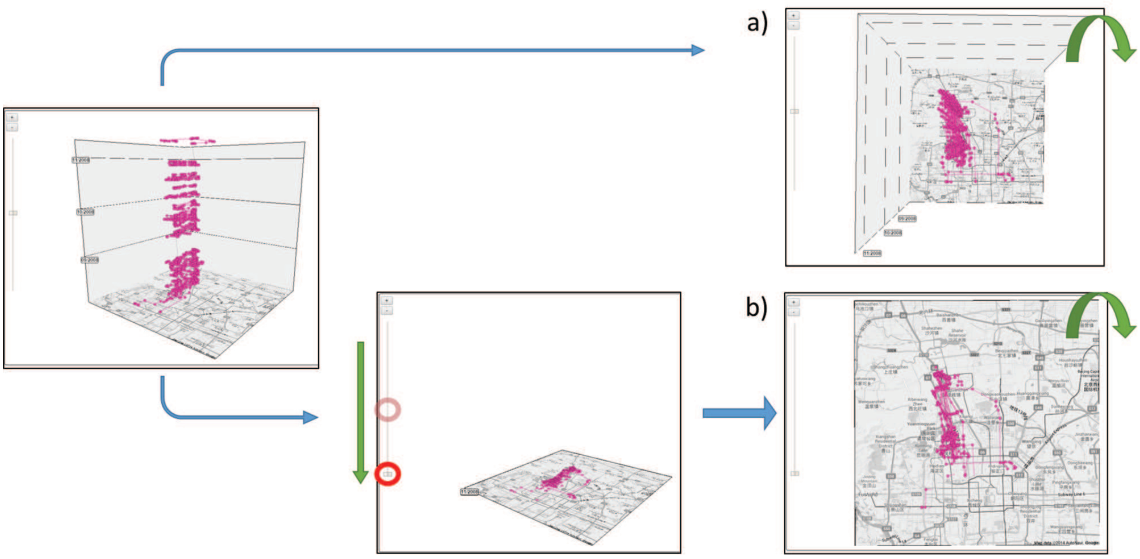

Figure 3.10: Interactive strategies with the STC visualization. a) Look at the data from a *bird's eye* view; b) Rescale the cube's height into a plane

sequence of points from the same mover (*EG*). Similarly, in the *identify* task, the participants were significantly faster in the *Qc(EE)* level, when compared with the other tasks. Although further studies are required, these results partially support the third hypothesis, *H3*: participants would have a better performance in elementary tasks.

Finally, the analysis of the participants' interaction strategies and comments suggests that 2D static maps and 3D STCs should actually be combined into the same visualization. This hypothesis is supported by 2 particular strategies applied by participants in *locate* tasks, in order to acquire a deeper understanding of the spatial distribution of the data (see Figure 3.10). The first strategy, performed by 7/16 participants, consisted in looking at the data from a *bird's eye view* perspective, as an apparent attempt to visualize the data as in a 2D static map (Figure 3.10 a). The other strategy was performed by 8/16 participants and consisted in the rescaling of the space-time cube's height, to compress all temporal information into a plane, thus effectively transforming the visualization into a *common* 2D static map (Figure 3.10 b), even if displayed on a 3D view. Nevertheless, further studies must be conducted to assess the adequacy of this combination (this will be further addressed in Chapters 5 and 6).

3.3 Summary

This chapter described two user studies comparing and assessing how an inexperienced population solves general tasks regarding the visualization of trajectory data with 2D static maps and 3D space-time cubes. The firsts study presented and evaluated ST-TrajVis, an application displaying both a 2D static map and a 3D STC enabling the visualization of the trajectories of one mover. This study consisted of a preliminary assessment of the ability of inexperienced users to explore and analyse trajectory data. The second study presented and compared the HTV prototypes, which also supported the visualization of the trajectories of several users, simultaneously. Each HTV prototype implemented a different map-based visualization technique, one a 2D static map, the other an STC. This study was conducted to compare how users interact with both techniques and explore which ones are most adequate for which types of visual analysis tasks.

Although more noticeable in the second study, both took into account the characteristics associated with trajectory data and, more importantly, the general types of tasks that might be needed to achieve a given goal, based on the simplified task model proposed in the previous chapter.

The results of these studies can be supported by the findings of previous research and, more importantly, they also emphasize relevant aspects regarding the studied visualizations and the types of tasks they may be the most useful, which are directly related to the goals of this research. Moreover, these results also highlight important challenges for the next steps of this research (presented in more detail in the next chapters). In particular, the results suggest that inexperienced *analysts* are able to make use of both techniques, to explore and extract information from trajectory data, while having a general positive experience with them. Nevertheless, the results have shown a higher preference from the users towards 2D static maps, which was also reflected in a higher effectiveness in *locate* tasks. Contrarily, STCs have been shown to be more useful in *associate* tasks, to help users detecting relations between the data. Based on this information and the feedback obtained by the participants, the results suggest that combining these two techniques can be useful for the visualization of trajectories over different types of tasks.

Although some of the obtained results may be considered *expected*, the positive reception by the participants towards these techniques, particularly the STC, is considerably surprising, specially when considering the limitations usually associated with 3D views. According to the results of the second study, STCs are more limited in *locate* tasks and, based on the participants' interactive strategies with this technique, they consider necessary to use a different visualization instead. This suggests the need to further improve this technique and analyse users' strategies when interacting with a given visualization. Indeed, as implied by these results, a possible improvement that should be explored consists in the use of an additional 2D map. However, previous studies have already suggested other alternatives, in particular, allowing the users to change the location of the map plane located at the base of the cube (Kraak, 2003). Nevertheless, no empirical evidence has been given support to the actual adequacy of this method and how it may change the interaction with this technique. The next chapter addresses these issues, by presenting a comparative user study focused on the STC technique and its adequacy in *locate* tasks.

Finally, Figure 3.11 depicts a schematic summary of the contributions of this chapter in the scope of the research reported on this dissertation.

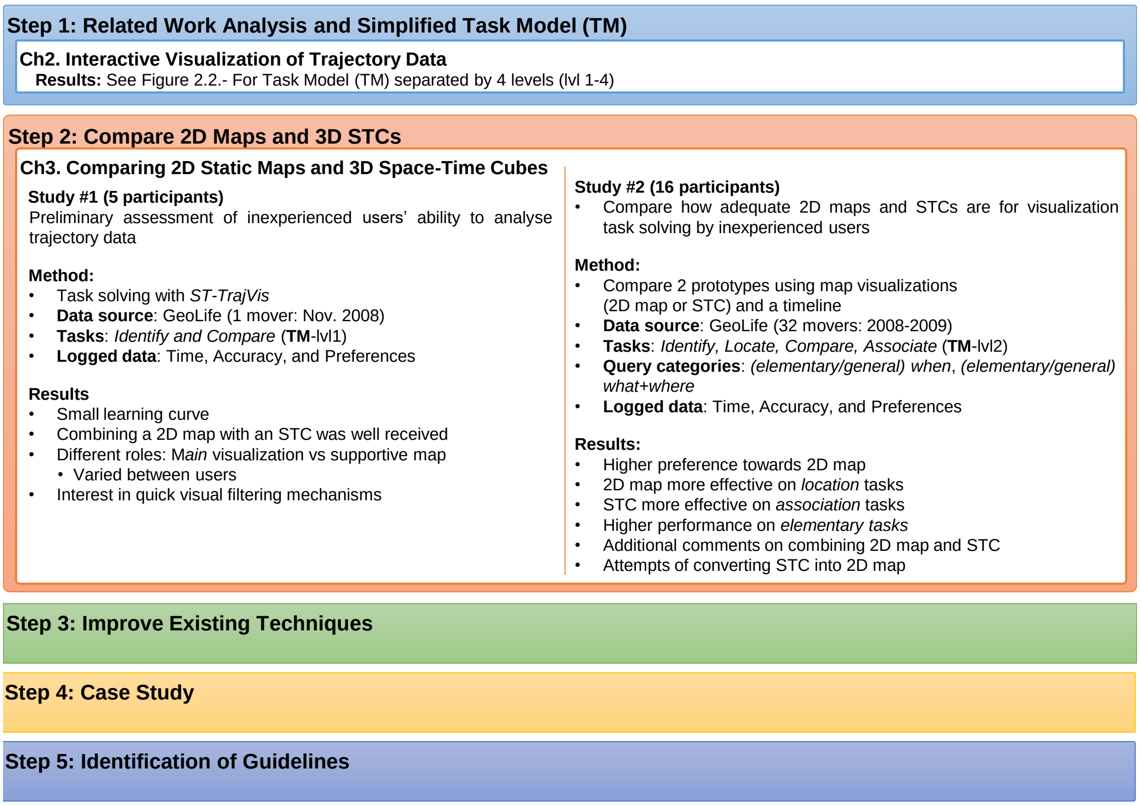

Figure 3.11: Summary of this research and the contributions of Chapter 3.

Chapter 4

Improving Spatial Awareness in Space-Time Cubes

The experiments presented in the previous chapter were successful in two specific aspects. First, they showed a high acceptance from inexperienced users towards 2D static maps and 3D space-time cubes (STCs) for trajectory data visualization. Second, the results also showed the types of visual analysis tasks in which each technique is more useful. In particular, the STC technique was shown to be, potentially, more adequate than 2D static maps, in terms of effectiveness, in *associate* tasks. Nevertheless, it still showed significant limitations in *locate* tasks and it was shown to be the least preferred approach by the participants.

The previous chapter also discussed the importance of studying the different types of interactive strategies used by the participants to acquire information, as a means of better comparing the different types of techniques. In particular, two types of strategies were used by the participants to somewhat convert an STC into a 2D map, by either manipulating the visualization's temporal granularity, or looking at the data from a *birds-eye view*.

Taking advantage of these results, the next two chapters change the focus of this research into another objective. Specifically, the goal consists in analysing possible improvements over existing visualization techniques and their effects over the interaction. As seem in Chapter 2, the analysis of previous studies reveals a significant volume of research conducted with 2D static maps. However, the same cannot be found about the STC, particularly when considering an inexperienced population as its primary target. In addition, considering the results from this research's previous studies, the STC can be considered as the most adequate for improvements.

To prevent the interference of additional factors into the results, in the previous chapter, a baseline STC was used in both studies. However, previous authors, such as Kraak (2003), have already proposed possible improvements to this technique, by allowing the manipulation of the position of the STC's spatial plane. In addition, the results shown in the previous chapter suggested that using 2D maps alongside STCs may be helpful for the users to analyse the data. Consequently, this chapter describes a new user study, comparing three different approaches corresponding to variations of the STC technique (Gonçalves et al., 2015c), and aims to: i) identify strategies that users may adopt to obtain spatial information when using an STC; ii) identify possible interactive techniques to help obtaining spatial information; and iii) assess the users' experience with those methods, while at the same time empirically comparing them.

The following sections present the various steps towards this study. First it describes the dataset used in the experiment, followed by the different visualization prototypes, their most relevant features, and their expected dis/advantages. After that, follows a description regarding the user study conducted with the prototypes, including the procedure, the study's results, and a detailed discussion regarding the effects of the results in relation to the objectives of this work.

4.1 Dataset

The data used for the development and evaluation of the three STC prototypes consists of a subset from the dataset made available by the Taxi and Limousine Commission of New York City (NYC TLC, 2017). This subset was composed by the information regarding the pick-up and drop-off locations of 10 taxis, randomly selected, during the period of one month (January of 2013). Each entry in the dataset represents a trip from a taxi and is composed by: two sets of coordinates (latitude, longitude), representing the pick-up and drop-off locations of the trip; two sets of timestamps, in both numerical and textual formats, representing the instants in time in which the trip started and ended; additional thematic information containing the identification of the taxi and its driver, the total number of passengers, the duration of the trip, its distance, the profits acquired, the type of payment, and a set of values regarding fare taxes, tips, and tolls. In total, the dataset was composed by, approximately, 13460 taxi trips.

As understandable by the aforementioned data description, this dataset is significantly different from the one used in the previous experiments. Although both datasets represent trajectories, the locations in which the data was recorded is different (the first being Beijing, China, the current one being New York, United States). More importantly, the type of movers is also considerably different. Although in both cases trajectories were generated by humans, in this dataset, the movers can be considered the taxis, instead of their drivers. Consequently, the type of thematic information associated with this dataset is dataset is also different and, arguably, more diverse. The usage of this new dataset brings considerable advantages to this work. First, even if the target users of this experiment are not familiar with the location from which the data was collected, they should be able to understand better the map in which the data is displayed. In the previous experiment, some locations (e.g., streets) were labelled with chinese characters, which raised some confusion. Second, even though it is not the focus of this research, the existence of more types of thematic data allows the creation of more visual appealing and informative trajectory maps. Third, and perhaps more important, this change allows the analysis of whether these visualization techniques and the previously obtained results are compatible or comparable with different datasets, in different trajectory visualization contexts.

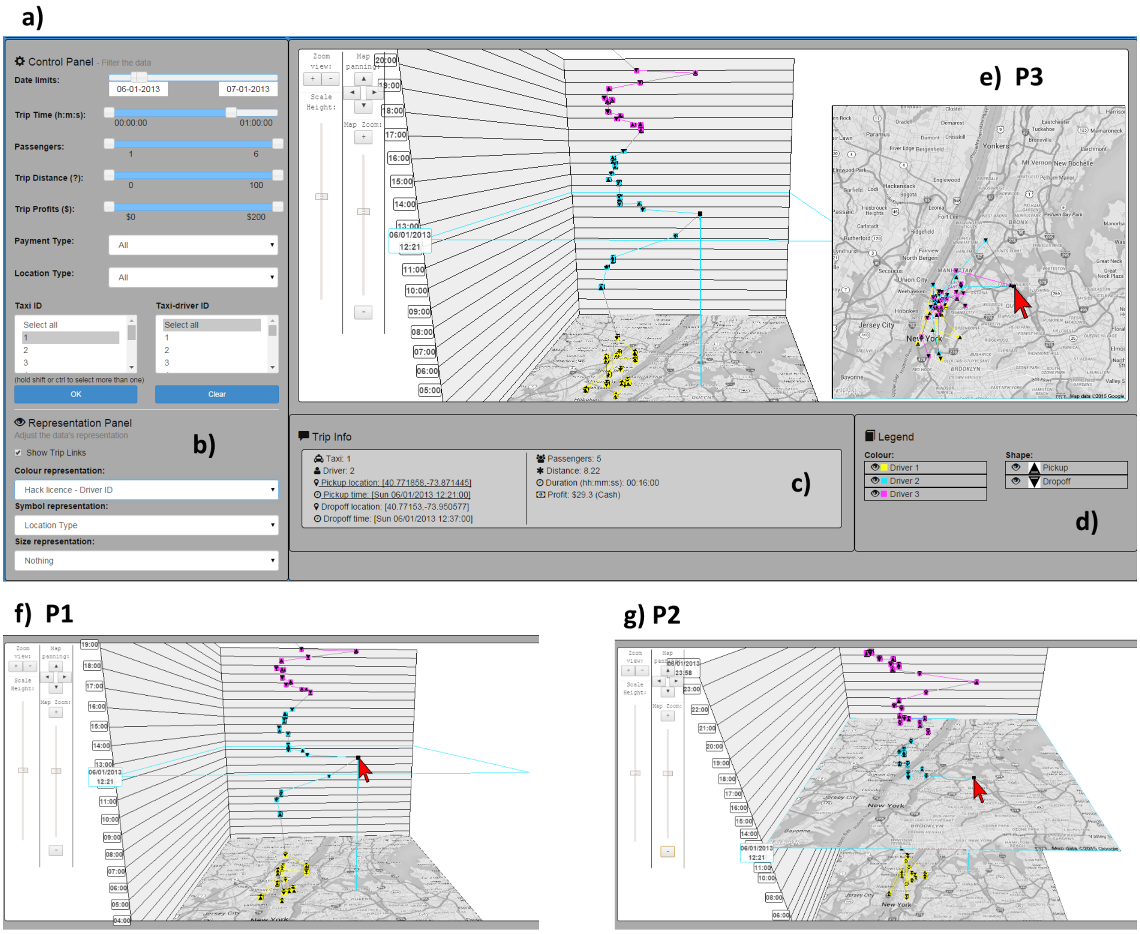

Figure 4.1: ImpSTC prototypes and their components: a) Control Panel; b) Representation Panel; c) Information Panel; d) Legend; e) *P3* STC with overview; f) *P1* basic STC; g) *P2* STC with moveable spatial plane.

4.2 Prototypes

The prototypes developed, named ImpSTC (Improved STC), integrate variants of the STC technique and allow the visualization of the taxi trajectories, described above. The prototypes were developed in HTML5, as local web applications, by taking advantage of the following main libraries and frameworks:

Node.js - a JavaScript runtime environment for developing a diverse variety of tools and applications (Dahi, 2009). Alongside **NW.js** (KDE, 2015), this library supports the creation of (local) web-applications as well as the reading and writing of files, which in the context of this study allowed the manipulation of the dataset previously mentioned and the recording of user interaction logs.

STCJS - an updated version of the same JavaScript API described in the previous chapter, in Section 3.2.2. In addition to various changes in the source code, unrelated to the goals of this work, the API was updated to support the features provided by the prototypes described below.

Figure 4.1 depicts an example of one of the ImpSTC prototypes, showing also the other STC alternatives. As in the last study, the prototypes were designed as similarly as possible, to ensure that the differences detected between the STC alternatives were the result of the differences between the techniques, and not the result of external variables related to the prototypes' aesthetics.

As depicted in the Figure 4.1, the prototypes were composed by five main components. The first component, the control panel (Figure 4.1 a), allows users to select which data to visualize. This selection can be done through various filters, which include:

the dates of movement, i.e., when movement started and ended;

the number of passengers, based on their minimum and maximum values;

trip distance, time and profits, similarly to the filter above;

the payment type, i.e., with money, credit card, or notes of credit;

the location types, i.e., pick-up or drop-off; and

the taxi and driver identifications, since one taxi can have several drivers in different periods of time.

With the second component (Figure 4.1 b) users are able to customize how the information is presented, by associating specific attributes to the visual variables of colour, shape, and size; consequently, supporting several combinations between the data attributes. Similarly to what was discussed in the previous chapter, and considering the usefulness of visual variables of colour and shape to represent qualitative/categorical information (Andrienko and Andrienko, 2005), each colour and shape can be used to represent location types (pick-up/drop-off), payment types, and the periods of the day (i.e., four colours/icons representing periods of six hours each, similarly to what was applied in the previous study). Colour can also be used to represent the identification of the taxi or the drivers. Since size is considered an useful variable for the representation of ordered and quantitative information, in the context of these prototypes, this variable can be used to display information regarding the number of passengers or trip fares (larger icons representing higher values). Figure 4.2 depicts some examples of using this functionality with the same subset of data. In the first example (Figure 4.2 a), it is possible to compare the service of two different drivers and conclude, for example, that the first driver made the trips with the highest profit. The other two examples emphasize the visualization of the data from a temporal perspective and allow the user to conclude, for instance, that the trips with the highest profit happened before 6 am (red), but the taxi carried a higher number of passengers after that (Figure 4.2 b and c, respectively).

The third component (Figure 4.1 c) displays all of the aforementioned attributes that are associated to any given feature highlighted in the STC. On the other hand, the fourth component (Figure 4.1 d) provides a legend describing the meaning of each visual variable illustrated in the

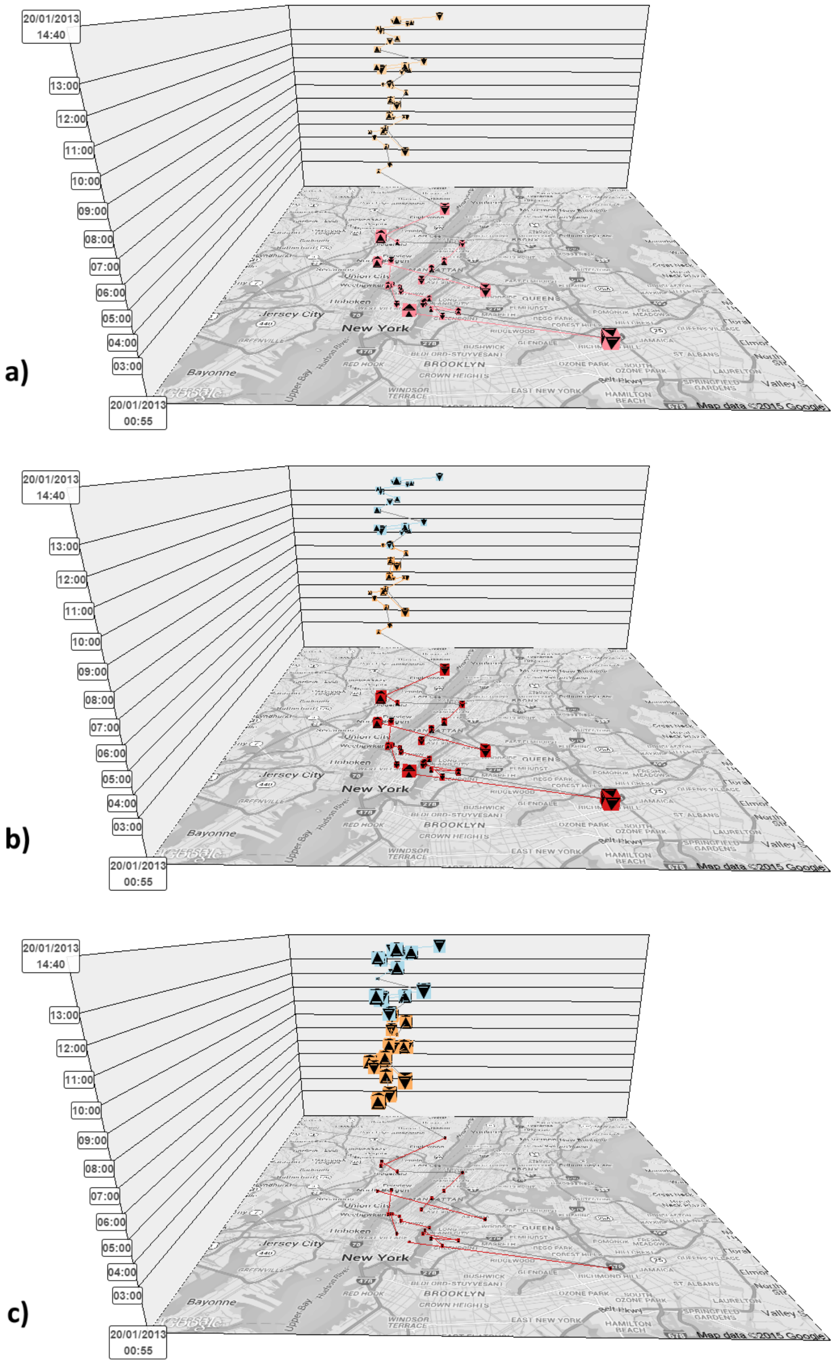

Figure 4.2: STC prototype with different visual variable configurations. a) Colour: driver ID, size: trip profit; b) Colour: time periods, size: trip profit; c) Colour: time periods, size: number of passengers.

STC, and a quick filtering mechanism that allows un/viewing sub-sets of the data, based on those variables.

The last component consists of an STC visualization, as depicted in the Figures 4.1 e, f, and g, depending on the prototype. In all three prototypes, the STC is composed by a 2D map plane, at the bottom of the cube, providing spatial information, and by several labels along the cube's height, providing temporal information. Similarly to the previous studies, trajectories are depicted as a sequence of points connected with lines, coloured and scaled according to the attributes selected in the representation panel.

All prototypes support common interactive features, such as panning and zooming, for the entire 3D view or just the 2D map displayed at the bottom. To pan the view, the user needs to drag the STC with the right mouse button, horizontally or vertically. Scrolling with the mouse or clicking in the respective buttons at the left-upper corner of the STC will allow the user to zoom-in or zoom-out the view. In comparison, to pan or zoom the map displayed at the bottom of the STC, the user would need to click the buttons on the left side menu of the STC, similarly to what is already done with other map-based applications. Furthermore, each visualization also allows the user to rotate the STC along its horizontal and vertical axes, with the left mouse button, and to rescale its height, with the first slider located at the STC's left corner. These, in turn, enable users to, respectively, change the point of view and manipulate the temporal granularity of the data.

As emphasized in Figure 4.1, the differences between the three prototypes are focused in this specific component, namely in how the spatial-component of the data is represented or emphasized. In the first prototype (Figure 4.1 f – *P1*), representing the most simple STC variant, selecting an object in the view will display the time moment in which the object was detected, with a thick line pointing at the object's location in the map plane.

Despite some differences, this approach is mostly similar to the STC used in the last experiment. Specifically, in addition to the prototypes' aesthetics, this new version of the STC makes the temporal granularity controls more visible and supports additional interactive features over the STC's map plane (see Figure 4.3).

In the second prototype (Figure 4.1 g – *P2*), based on the most well known (yet not validated) proposal to improve spatial awareness on STCs (Kraak, 2003), in addition to the previously mentioned methods, a copy of the map plane is displayed at the same height as the selected object. Highlighting the temporal scale of the STC, by hovering the mouse on it, will also move the additional map. Although this approach should simplify how users locate information within the STC, adding a plane to a certain height will, necessarily, occlude all data located bellow the selected object.

Finally, following the results of the previous study, the third prototype (Figure 4.1 e – *P3*) aims to combine the advantages of 2D and 3D views by displaying a 2D map overview, at the right-bottom corner of the visualization. This additional map view displays all information contained within the STC's boundaries, focusing on the spatial location of the data, thus continuously providing some spatial context to the user. Moreover, the selection of an object within any of

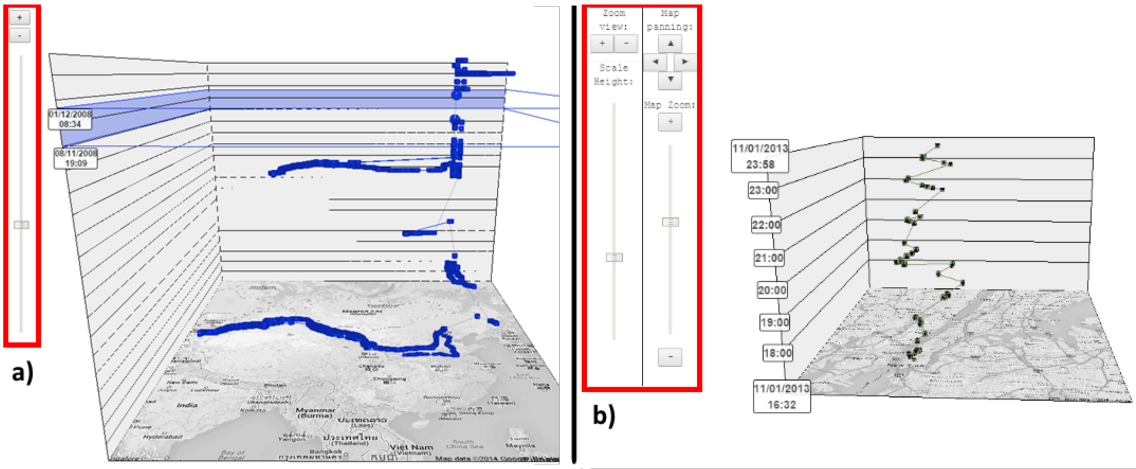

Figure 4.3: Additional interactive controls for the STCs a) from the previous study and b) from the current one.

the views will highlight the corresponding one in the other and, similarly, changing the visual representation of the information within the STC, through the representation panel, or with the left-corner controls, will change the corresponding representations within the 2D overview. Nevertheless, this approach will, necessarily, result in some redundancy, since the same data is being represented in two similar views, and may also lead users to divide their attention between the views (Baudisch et al., 2002).

4.3 User Study

Using the described ImpSTC prototypes, a user study was conducted to compare the usability of the different variants of the STC technique. A total of 20 participants volunteered to the study, aged between 19 and 34 (Av:26.6, SD:4.1). Similarly to the previous experiments, all participants were knowledgeable with computer applications and geographic information systems to search for directions towards specific points of interest. Nevertheless, none of the participants were experts with trajectory data analysis, nor were familiarized with New York's geography, thus preventing previous knowledge over the location to adulterate the study's results.

Considering the properties of each prototype, alongside the characteristics of the participants, the hypothesis of this study were the following:

H1 Participants will prefer the prototypes with complementary spatial information (*P2* and *P3*);

H2 Despite their disadvantages, the additional spatial-context aids provided by *P2* and *P3* will enable the participants to complete tasks with a higher performance; and

H3 Similarly to our previous study, participants will perform better in *elementary* tasks, i.e., tasks that focus in just on object or time moment, due to the lesser amount of information displayed.

4.3.1 Tasks

The experiment conducted in Chapter 3 followed the second level of tasks depicted in Figure 2.2; however, since no significant differences between the techniques were detected the *identify* and *compare* tasks of this level, it can be argued that the most relevant results of that study came only from the *locate* and *associate* tasks. In addition, as previously stated, these tasks can be generalized into a higher level of tasks, composed by more general *identify* and *compare* tasks. For those reasons and for a matter of coherence with the aforementioned analysis and terminology, this user study was conducted based on the two tasks, depicted in the first level of Figure 2.2. The first task, *identify*, required participants to find taxis or locations within the STC according to some spatio-temporal and thematic constraints. The second task, *compare*, required the analysis of similarity and difference relations between data elements in the STCs.

In the context of this study, two independent variables were considered:

Visualization technique (Vt), with three levels, corresponding to the three prototypes, $P1$, $P2$, and $P3$; and

Query category (Qc), with four levels adapted from the spatio-temporal data queries identified by Andrienko et al. (2003), namely:
elementary what+where, *elementary when* (EE);
elementary what+where, *general when* (EG);
general what+where, *elementary when* (GE); and
general what+where, *general when* (GG).
These categories determine whether the focus of the task is in just one (*elementary*) or several (*general*) moving objects (*what+where?*) in a specific time moment (*when?*) or across a time period.

Table 4.1 provides some examples of the task conducted by the participants.

4.3.2 Procedure

The experiment followed a *within-subjects* design and all participants carried out each task individually, in a controlled environment. At the beginning of the study, subjects were briefed about the objectives of the experiment, and they viewed a demonstration of the prototypes. Before carrying out the tasks, they were asked to interact with the applications, as a means of getting used to the STCs' interactive controls. They were also encouraged to clarify any doubts. After the training phase, the participants performed the two tasks, taking into account the different visualizations and query categories. To mitigate sequence effects, the order in which the independent variables were presented was counterbalanced based on a *latin-square* design. However, since *comparison* tasks require necessarily more than one data item (either from the same or from different trajectories), the level *elementary what+where elementary when* ($Qc(EE)$) was not considered for these tasks. Consequently, each participant performed a total of 21 trials: $(3Vt)$ x $(4$ identify $+ 3$ compare).

Table 4.1: Examples of tasks conducted by the participants.

Query Category	Tasks	Example
Elementary What+Where Elementary When	Identify	Where was taxi 1 at 8am?
Elementary What + Where	Identify	Which drivers dropped passengers in airports, with taxi 1?
General When	Compare	Which drivers served the most people outside Manhattam?
General What+Where	Identify	Which taxis were in service around Broklyn, at 4pm?
Elementary When	Compare	To where were heading the taxis with most passengers, by 12am?
General What+Where	Identify	Which taxis picked up passengers in Queens?
General When	Compare	Which taxis picked the most passengers in the centre of Manhattan?

To assess the hypotheses of this study, the following dependent variables were recorded:

Subjective preferences: at the end of the tasks, participants were asked to rate the ease of use of each prototype, on a scale between 0 and 10. This variable allows to evaluate the participants' acceptability towards each prototype individually and comparatively to each other;

Task accuracy: to measure the participants' effectiveness in completing the tasks, the responses for each task were rated between 0 and 10, depending on the detail given by the participants (e.g., saying that a taxi was heading to New York is less detailed than Manhattan, which is less detailed than Central Park);

Task completion time: to measure the participants' efficiency in completing the tasks, the amount of time (in seconds) from the moment the participant started the task until the prototype confirming that his/her answer was recorded;

Number of interactive actions: recorded and calculated during the tasks, this variable takes into account various actions, such as map and view panning/zooming operations, rescaling the STC's height, or highlighting information. This variable can be considered as an additional metric of complexity, allowing to analyse the tasks and prototypes in which the participants required to interact the most to extract information;

3D mitigation strategies: as observed in the previous study, some participants tried to *convert* the STC into a 2D static map, which, in addition to the participants' needs for the completion of the tasks, may be also related with their preferences between the techniques. In particular, two strategies were used, which consisted of reducing the STC's height to its minimum (*FlatV*) and looking at the data from a *birds-eye view* perspective (*BEV*). By recording the

Table 4.2: Statistic test results for the participants' subjective preferences. **n.s.s.** stands for *not statistically significant*.

	Identify	Compare
Friedman	$X^2(2) = 14.8, p = 0.001$	$X^2(2) = 11.7, p = 0.003$
Wilcoxon S.R:		
P1 – P2	$Z = -3.42, p = 0.001$	**n.s.s.**
P1 – P3	$Z = -2.98, p = 0.001$	$Z = -2.96, p = 0.003$
P2 – P3	**n.s.s.**	$Z = -2.56, p = 0.010$

evolution of the positions and rotations of the STCs' point of view, it is possible to estimate when and how often each of these strategies were used.

The information regarding the last two variables were recorded automatically by the prototypes, during the execution of the tasks. In addition, during the study, specially at the end of the experiment, participants were encouraged to *think aloud*, and share their opinions regarding the prototypes.

4.3.3 Results

The following sections present the results obtained in the study according to the considered dependent variables. For a matter of readability, the following sections will focus on the most statistically significant results or those close to be considered as such.

Subjective Preferences

To compare the differences between the participants' opinions after each set of tasks, a Friedman's test was used. When significant differences were detected, the data was subjected to a Wilcoxon Signed Rank test with a Bonferroni correction for pairwise comparisons. Table 4.2 depicts the results of these tests.

Overall, the results revealed significant differences in both tasks. Specifically, in the *identify* task, participants have shown a higher preference towards $Vt(P3)$ (7.3/10) and $Vt(P2)$ (6.5/10) over $Vt(P1)$ (5.5/10). In the *compare* task, participants have also shown a significant higher preference for $Vt(P3)$ (7.2/10) over $Vt(P1)$ (6/10) and $Vt(P2)$ (6.5/10).

Task Accuracy

To compare the participants' accuracy with the visualizations and in the different categories, it was used a similar procedure to the one previously described.

Table 4.3 displays the average scores obtained in the various tasks by the participants, while Table 4.4 shows the results of the statistical analysis conducted over that data. In the *identify* task, the tests revealed significant differences both in terms of visualization technique (Vt) and Query category (Qc). Further pairwise comparison tests confirmed significantly less accurate results with

$Vt(P1)$ comparatively to the other two STCs. Participants were also significantly less accurate in $Qc(GE)$ tasks, comparatively to all others.

In the *compare* task, the tests revealed somewhat similar results, where participants were significant less accurate in $Qc(GE)$ tasks comparatively to $Qc(EG)$ tasks.

Table 4.3: Mean participants' results in terms of task accuracy.

	Identify				**Compare**		
Accuracy [0-10]	**EE**	**EG**	**GE**	**GG**	**EG**	**GE**	**GG**
P1	8.05	8.30	6.40	7.50	7.45	8.05	8.13
P2	8.40	9.20	7.55	8.35	7.75	8.83	7.83
P3	8.65	8.70	7.45	8.63	8.25	9.50	8.85

Table 4.4: Statistic test results for the participants' accuracy. **n.s.s.** stands for not statistically significant.

	Identify	**Compare**
Friedman (Vt)	$X^2(2) = 9.796, p = 0.007$	**n.s.s.**
Wilcoxon S.R.		
P1 – P2	$Z = -2,72, p = 0.006$	**n.s.s.**
P1 – P3	$Z = -2.54, p = 0.011$	**n.s.s.**
P2 – P3	**n.s.s.**	**n.s.s.**
Friedman (Qc)	$X^2(3) = 26.154, p < 0.001$	$X^2(2) = 8.54, p = 0.014$
Wilcoxon S.R.		
GE – EE	$Z = -3.57, p < 0.001$	**n.s.s.**
GE – EG	$Z = -4.81, p < 0.001$	$Z = -2.83, p = 0.005$
GE – GG	$Z = -2.787, p = 0.005$	**n.s.s.**

Task Completion Times

To compare the participants' results in terms of task completion times, their data was subjected to the Shapiro-Wilk test of normality. After that, a repeated measures ANOVA test was used, followed by Bonferroni tests for pairwise comparisons, when significant effects were detected.

Table 4.5 shows the mean results, in terms of task completion times, for all tasks with all combinations of the two independent variables. In the *identify* task, the tests revealed a significant effect from Qc over the results ($F(2.5) = 6.671$, $p = 0.001$). The pairwise comparison tests allowed to confirm significantly lower completion times in the least demanding type of task ($Qc(EE)$), comparatively to the remaining three (with $p \leq 0.05$ in all cases). No significant differences were detected in terms of visualization technique, nor in the *compare* task.

Table 4.5: Mean participants' results in terms of task completion time.

	Identify				**Compare**		
Time (sec)	**EE**	**EG**	**GE**	**GG**	**EG**	**GE**	**GG**
P1	122.90	166.55	203.71	176.33	141.40	116.26	122.38
P2	126.24	168.55	193.24	174.36	152.48	123.10	146.52
P3	127.93	180.05	178.81	142.64	156.17	121.57	107.07

Table 4.6: Mean participants' results in terms of number of interactive actions.

	Identify				**Compare**		
N° Actions	**EE**	**EG**	**GE**	**GG**	**EG**	**GE**	**GG**
P1	26.65	31.95	41.80	51.25	35.20	38.80	46.65
P2	22.85	28.10	35.90	40.15	34.55	32.95	46.55
P3	22.55	45.65	41.45	44.55	34.80	35.35	38.00

Number of Interactive Actions

A similar procedure was used to compare the participants' results in terms of the number of actions performed to complete the tasks.

Table 4.6 shows the mean number of actions performed by the participants, for all tasks with all combinations of the two independent variables. In the *identify* task, significant effects were detected from both Qc and Vt ($F(2.57) = 10.464$, $p < 0.001$ and $F(1.68) = 3.260$, $p = 0.05$, respectively). The pairwise comparison tests also revealed a significantly lower number of actions in $Qc(EE)$ tasks ($p \leq 0.013$, comparing to all cases), and a generally significantly lower number of actions of $Vt(P2)$ comparatively to $Vt(P1)$ ($p = 0.046$). No significant results were detected in the *compare* task.

3D Mitigation Strategies

To analyse the strategies used by the participants to trying to convert the STCs into 2D static maps, the data referring to the ratio of time spent in a specific view (i.e., birds-eye view - *BEV*, and fully compressed STC - *FlatV*) was compared, following a procedure similar to the analysis of the participants' accuracy. Nevertheless, this analysis was focused on the *BEV* approach, considering that only 7 participants used a *FlatV* approach and without any clear pattern.

Table 4.7 shows the mean ratio of time spent by the participants with a *birds-eye view* approach and Table 4.8 provides an overview of the results of the statistical tests. In the *identify* task, the tests revealed significant differences when comparing the visualization techniques (Vt) and the query categories (Qc). Regarding the first variable, the pair-wise comparison tests confirmed a significantly higher usage of the *BEV* approach with the STCs without the 2D map overview (i.e., $Vt(P1)$ and $Vt(P2)$). In regards to the second variable, the results also suggest a significantly lesser usage of the *BEV* approach in $Qc(EE)$ tasks, comparatively to $QC(GE)$ and $Qc(GG)$, but a significantly higher usage of *BEV* in $Qc(GG)$ tasks, comparatively to $Qc(EG)$ tasks.

Table 4.7: Mean participants' results in terms of the ration of time spent by the participants in a *birds-eye view*.

	Identify				**Compare**		
Ratio BEV (%)	**EE**	**EG**	**GE**	**GG**	**EG**	**GE**	**GG**
P1	8.64	32.18	14.03	62.8	19.07	67.55	42.88
P2	8.81	31.55	8.82	64.0	7.88	58.50	39.68
P3	8.53	16.63	7.90	42.8	8.93	43.01	26.03

In the *compare* task, the results suggest significant differences in terms of Qc, by confirming a significantly lesser usage of the *BEV* approach in $Qc(GE)$ tasks, comparatively to $Qc(EG)$ and $Qc(GG)$ tasks.

Table 4.8: Statistic test results for the participants' accuracy. **n.s.s.** stands for not statistically significant.

	Identify	**Compare**
Friedman (Vt)	$X^2(2) = 16.27, p \leq 0.001$	**n.s.s.**
Wilcoxon S.R.		
P1 – P2	**n.s.s.**	**n.s.s.**
P1 – P3	$Z = -3.80, p \leq 0.001$	**n.s.s.**
P2 – P3	$Z = -2.79, p = 0.005$	**n.s.s.**
Friedman (Qc)	$X^2(3) = 28.04, p \leq 0.001$	$X^2(2) = 25.09, p \leq 0.001$
Wilcoxon S.R.		
EE – GE	$Z = -3.27, p = 0.001$	**n.s.s.**
EE – GG	$Z = -4.27, p \leq 0.001$	**n.s.s.**
EG – GG	$Z = -2.98, p = 0.003$	**n.s.s.**
GE – EG	**n.s.s.**	$Z = -4.13, p \leq 0.001$
GE – GG	**n.s.s.**	$Z = -3.40, p = 0.001$

4.4 Discussion

Overall, the obtained results support the main hypotheses of this study.

According to the first hypothesis (H1), it was expected a higher preference towards the prototypes with additional spatial cues (i.e., $P2$ and $P3$). As seen in the results of both tasks, $P3$ was always among the preferred techniques, immediately followed by $P2$.

In the second hypothesis (H2), it was also expected a higher performance with the improved visualizations. Some results point towards a lower accuracy with $P1$ (the simplest approach), comparatively to the others. In addition, although no significant results were found in terms of task completion times, the analysis of the number of actions performed by the participants suggests a significantly lower number, which implies a lower effort when interacting with $P2$, comparatively to $P1$.

In the case the third hypothesis (H3), a higher performance in *elementary* tasks was expected, i.e., tasks focused on the analysis of one time moment or just one mover. The results support this hypothesis, by suggesting a lower accuracy in tasks with a higher visual noise. This was particularly true in GE tasks, in which the participants need to focus on the data from several movers in one specific time moment, implying the need to focus on a single temporal plane, while ignoring or filtering others. In terms of task completion times and number of actions, the participants were significantly faster and performed a lower number of actions in tasks with less information to analyse, namely EE tasks, which focused on the analysis of one mover in a specific time moment.

In addition, the analysis of the participants behaviour and comments also provided interesting insights about the different visualizations. Throughout the experiment, it was possible to detect some attempts to minimize the 3D aspects from the STC and obtain a more detailed spatial context. Similarly to the previous study, the participants focused on two particular strategies, either, more commonly, by looking at the data from a *birds-eye view* (*BEV*), or by using the temporal granularity controls to flatten the STC. The results shown a more frequent usage of these techniques (*BEV* in particular) with $P1$, followed by $P2$, and in tasks that required the analysis of data in a large period of time (i.e., GE and GG tasks). These results suggest that the use of an additional map overview, in addition to improving location accuracy, also minimized the negative effects from 3D visualizations. However, it is unclear if these had any effect over the participants accuracy, as no correlations between the data were detected. Moreover, it is also relevant to notice that although the results show a less frequent usage of these approaches with $P3$, they did not completely prevented them, which is somewhat surprising, considering the existence of the 2D map overview.

Overall, a quantitative analysis over the results points towards the fact that approaches $P2$ and $P3$ are indeed improvements over the STC technique. However, it is also important to take into account the participants' comments. At the end of the study, they mentioned that using a moveable plane, or a 2D map overview on the STC, was helpful to acquire spatially-related information more easily, although, under different circumstances. Regarding the auxiliary moveable plane, the participants mentioned that they considered it more helpful to find the locations of specific points or to analyse the evolution of locations visited over time from a given taxi. Despite these positive comments, as expected, some participants stated that, sometimes, the moveable plane would occlude information, conditioning the interaction. In fact, this situation was particularly problematic when participants moved the plane to a certain height and then looked at the STC from a *birds-eye view*, to better analyse a certain location. In these situations, several users asked if *the information had disappeared*, not having realized, at first, that the information was still there, just bellow the auxiliary map. Figure 4.4 depicts two examples of this scenario when the STC is in a *normal* perspective (Figure 4.4 a), in which time and space are visible, and when the information is visualized from a *birds-eye view* perspective (Figure 4.4 b). Regarding the 2D map overview, the participants stated that it was helpful to have a better overall perception of the geographical space,

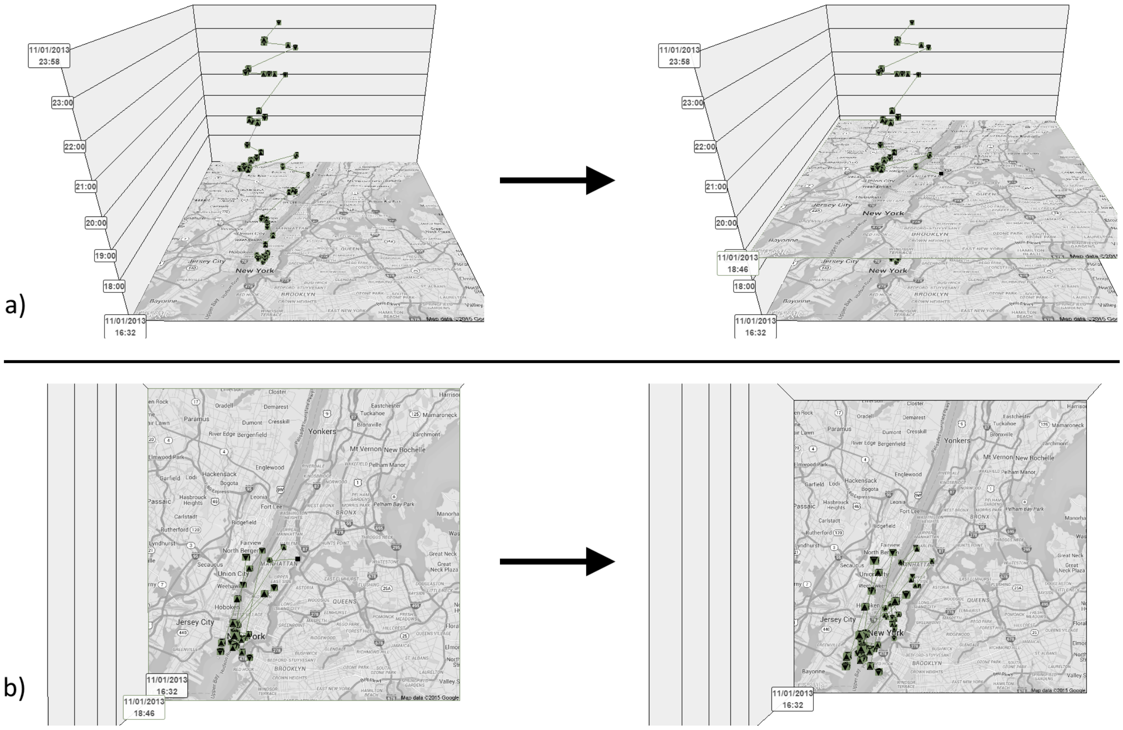

Figure 4.4: Moveable map plane overlapping existing information. a) in a *normal* perspective; and b) in a *birds-eye view* perspective.

and that, alongside the STC, allowed them to focus better on spatial and temporal information simultaneously. Despite that, some participants expressed their interest in having more control over the 2D map overview, given their familiarity with 2D maps. Although the participants were able to select or highlight information in this overview, the remaining of the interactive controls were associated to the rest of the STC view.

These later comments are particularly important, since they emphasize the need to further study the combination of 2D static maps with 3D STCs, instead of turning one of the techniques into a *complement* of the other. Besides, they also emphasize the critical role of interactivity in the visualization of trajectories. This fact can help justifying why some participants, despite having a 2D overview of the map, still tried to look at the STC's data from a *birds-eye view*. Despite helpful, its interactivity was rather limited, which may have encouraged the participants to sometimes pay more attention to the STC, when using the 2D overview would have been more helpful.

4.5 Summary

This chapter is a direct follow up from the previous one, in the sense that, while the focus of the Chapter 3 consisted in understanding the main limitations with 2D static maps and 3D space-time cubes, from an inexperienced user perspective, the current one aimed at minimizing those limitations. In particular, two variations of the STC visualization were studied, to improve the adequacy of this technique in *location*-based tasks, which have been shown as the ones users may complete with a lower performance. The first improvement, adapted from previous (not validated) work (Kraak, 2003), consisted in using an additional map plane that the user could move anywhere along the STC's height, to help in the location of information. The second followed the premise of somewhat combing 2D maps and 3D STCs, to complement each other, by using a 2D static map overview over a normal 3D space-time cube. These approaches were compared with a baseline STC approach, taking into consideration some of the main types of visual analysis tasks and trajectory data categorizations, already described in Chapter 2 and used in Chapter 3.

In addition to the comparative analysis of these approaches, this chapter also addressed the importance of taking into account the users' general strategies and actions, when evaluating the visualizations. In particular, the number of actions performed by the participants was taken into consideration, as an additional measure of effort to complete the tasks, alongside the occurrence or persistence of actions focused in trying to convert 3D STCs into 2D static maps.

The results of this experiment suggest that both approaches are well received by the users, as they were preferred over simpler approaches, and they also significantly improve the users' performance, particularly in *identify* tasks. In addition, combining the space-time cube with a 2D static map overview is likely to reduce the attempts at looking at the data in a space-time cube from a *birds-eye view* perspective. Nevertheless, as discussed in the previous experiment (Section 3.2.4), the volume of data under analysis shown a negative effect over the users' performance, and despite their positive comments towards these improvements, they may have caused some confusion, by hiding information through the moveable plane, or by lacking interactivity (2D overview).

Figure 4.5 depicts a schematic summary of the contributions of this chapter in the scope of the research reported on this dissertation.

Ultimately, the results of this experiment highlight two important aspects. The first is that, as suggested in the first chapter, the improvement of 2D maps and 3D STCs can be achieved by combining both techniques. However, it is necessary to understand the role of each technique in the interaction. Second, the usage of the participants' interactive actions as an evaluation metric, from a qualitative and quantitative point of view, has a critical role in the understanding of the issues and capabilities of the various techniques.

The next chapter continues to address the overarching issue of improving these visualization techniques, by focusing, in particular, on a more robust combination between 2D maps and 3D STCs alongside the evaluation of the effects over the interaction and performance of this combination, comparatively to using the techniques individually. For that, three new prototypes were developed, implementing a 2D map, a 3D STC, and a combination of both techniques.

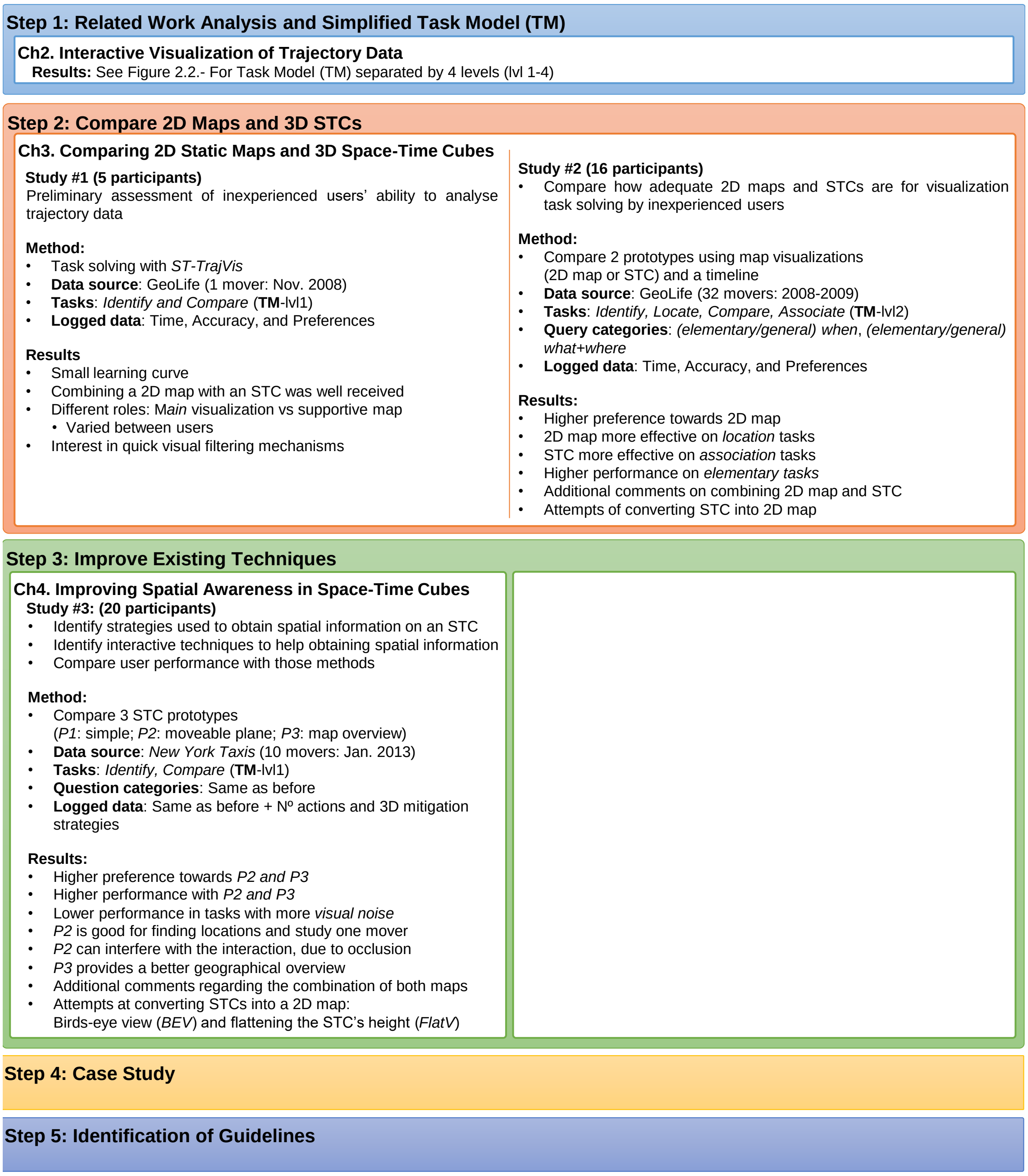

Figure 4.5: Summary of this research and the contributions of Chapter 4.

Chapter 5

Combining 2D Static Maps and 3D Space-Time Cubes

As discussed in the previous chapters, inexperienced users in terms of data visualization and analysis are able to work out conclusions related to the visualization of trajectory data, when using map-based techniques, such as 2D (static) maps and 3D space-time cubes (STCs). However, as suggested by the different characteristics of these techniques and later emphasized by the results presented in Chapter 3, the adequacy of a given map technique to visualize trajectories is significantly dependent on the types of visual analysis tasks required to achieve a given goal.

Nevertheless, it is unlikely that the analysis of trajectory data will depend on just one type of task. This motivates the need to understand how to improve existing techniques by minimizing their known limitations. As analysed in Chapter 4, even the inclusion of simple additional spatial-related cues in STCs (e.g., a moveable map plane) supports a significant performance increase by the participants in location-based tasks. More importantly, the results presented in Chapter 4 also highlight the possibility of improving both techniques through their direct combination.

This particular topic has already been somewhat addressed in some studies, by suggesting the ability to switch between 2D and 3D maps to enable the exploration of the data from different points of view (Amini et al., 2015), or by reporting a positive reception from domain and non-domain expert users towards using both 2D maps and STCs to visualize spatio-temporal data (Kveladze et al., 2015). Despite the valuable contributions of these works, they are still mostly focused in distinguishing the different map components, rather than exploring the actual consequences of their combination in terms of performance and interaction strategies.

This chapter continues to address the challenges, first mentioned in the previous two chapters, that were related to the comparison and improvement of 2D maps and STCs for the visualization of trajectory data, in particular by focusing on the issue of combining both visualization techniques. For that, a new user study is presented, comparing three interactive map prototypes, this time however, with one using a 2D map, one using a 3D STC, and one implementing both techniques in the same visualization (Gonçalves et al., 2015b, 2016b). Specifically, this study was conducted to: i) expand the existent knowledge that distinguishes the benefits of using 2D maps compared to STCs (including the results obtained in Chapter 3); ii) explore whether it makes sense to combine both techniques and analyse the dis/advantages of combining both techniques, comparatively to using them separately; iii) empirically validate the combination of these techniques to help the interaction, taking into account the types of visual analysis tasks a user might need to perform to achieve a given goal; and iv) understand and analyse the types of strategies used by inexperienced users to acquire information with the different techniques, alongside their performance and acceptance.

As described in Chapter 2, the results of some studies suggest also a possible effect of gender over user performance (Vecchi and Girelli, 1998). Although this factor is beyond the scope of this research, the analysis of the results of this study take also this variable into account. Similarly, considering the results of the experiments presented in the previous chapters, the volume of information on display has also a significant effect over user performance. For that reason, this variable is also taken into consideration for this study.

The remainder of this chapter is organized as follows: the next section describes the dataset used, followed by the prototypes developed in the context of this study, the user study and the results obtained with the prototypes. The chapter concludes with a detailed discussion of the results, and their effect in relation to the objectives of this work.

5.1 Dataset

Similarly to the experiments presented in Chapter 3, in which the second study was a direct follow up of the first, the study presented in this chapter can also be considered as a follow up of the study presented in Chapter 4. Even though with distinctive approaches, one of their main goals consists on the optimization of map-based visualization techniques. In addition, the dataset used in these experiments was considerably useful for an efficient creation of the prototypes as well as the design of the various tasks. As such, in this chapter, for a matter of coherence and simplicity, the data used for the development and evaluation of the prototypes consists, again, in a subset of the data made available by the Taxi and Limousine Commission of New York City (NYC TLC, 2017).

Similarly to the previous study, this subset was composed by the information regarding the pick-up and drop-off locations of 10 randomly selected taxis, different from those used in the previous study, during the period of one month (January of 2013). In total, the dataset was composed by, approximately, 13430 taxi trips. In addition to the spatio-temporal and thematic properties associated with the data, which were described in the previous chapter (section 4.1), to further optimize the users' recognition of the observed locations, this dataset was enriched with more detailed results from reverse geocoding the stop location coordinates (i.e., acquiring the name of the locations based on their geographical coordinates).

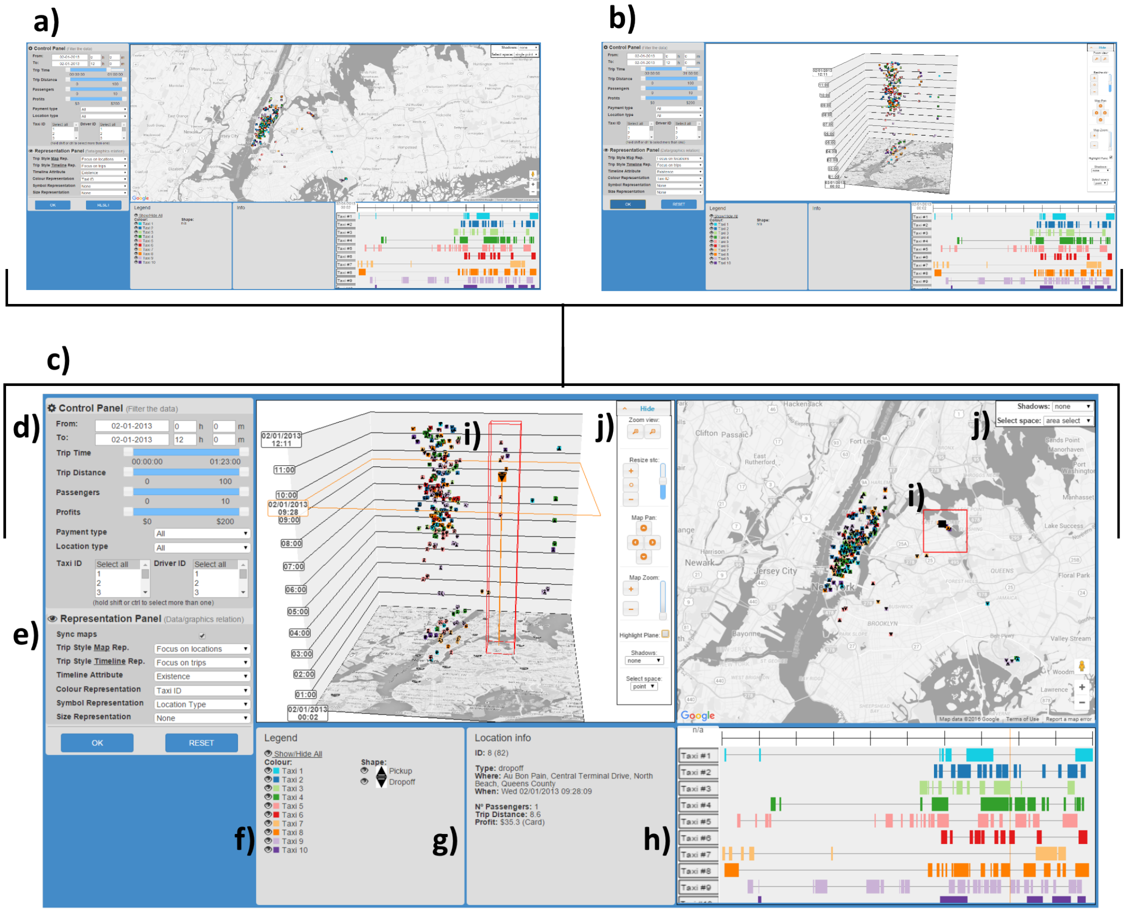

Figure 5.1: TTV prototypes a) 2D map prototype; b) STC prototype; c) 2D map + STC prototype; d) Control panel; e) Representation panel; f) Legend; g) Information panel; h) Timeline; and i) Spatial highlight reflected in both maps.

5.2 Prototypes

To achieve the goals presented in this chapter, it was necessary to compare 2D static maps (Figure 5.1 a), 3D STCs (Figure 5.1 b), and their combination (Figure 5.1 c). As such, three prototypes, called Taxi Trajectory Visualizer (TTV), were developed and then compared, integrating different interactive map-based visualization techniques. The three prototypes were developed in HTML5, as local web applications, and took advantage of the following main libraries and frameworks:

Node.js and **NW.js** - as described in the previous chapter, in Section 4.2;

STCJS - an updated version of the JavaScript API described in the previous chapters, in Sections 3.2.2 and 4.2. In addition to updated interactive controls, such as the ones depicted in Figure 5.1 j, and further explained below, this version supports various performance upgrades.

TRAJMAP2DJS - an updated version of the JavaScript API described in Section 3.2.2. Similarly to the one above, this version supports additional visualization options (Figure 5.1 j) and various performance upgrades.

TIMELINEJS - a JavaScript API developed over the three.js library, that supports the creation and interaction with timelines for the visualization of data from a temporal perspective. This API was developed based on STCJS and TRAJMAP2DJS to simplify the development of the prototypes and, thus, it supports the creation of various layers with information, represented with customizable styles (e.g., colour and size settings). Consequently, this API also supports the notification of various user interaction events, such as when a user selects a time point/period feature or changes the timeline's properties (e.g., pan and zoom).

Figure 5.1 displays the three prototypes and their main components. As previously stated, an important detail regarding the comparison of these visualization approaches is to ensure that the prototypes are informationally equivalent, i.e., it is possible to extract similar information from the different techniques. Therefore, as observed in Figure 5.1, all three prototypes are composed by six main components, sharing a similar visual style and interactive controls, ultimately having the map-based visualizations as their distinguishing factor.

The first component (Figure 5.1 d and Figure 5.2 a), the control panel, provides temporal and thematic filters that support the selection of the data subset that is to be visualized. In particular, the user can select the data based on: the time period of movement, expressed in terms of days and hours/minutes of the day; the maximum/minimum values of time, distance, number of passengers, and profit obtained in each trip; and the identification of the taxi or the driver, since it is possible for one taxi to be used by several drivers in different periods.

The second component (Figure 5.1 e and Figure 5.2 b), the representation panel, allows users to associate the thematic attributes described above to a set of visual variables (i.e., colour, shape, and size), in the map-based visualization(s) and in the timeline (explained below). Colour can be

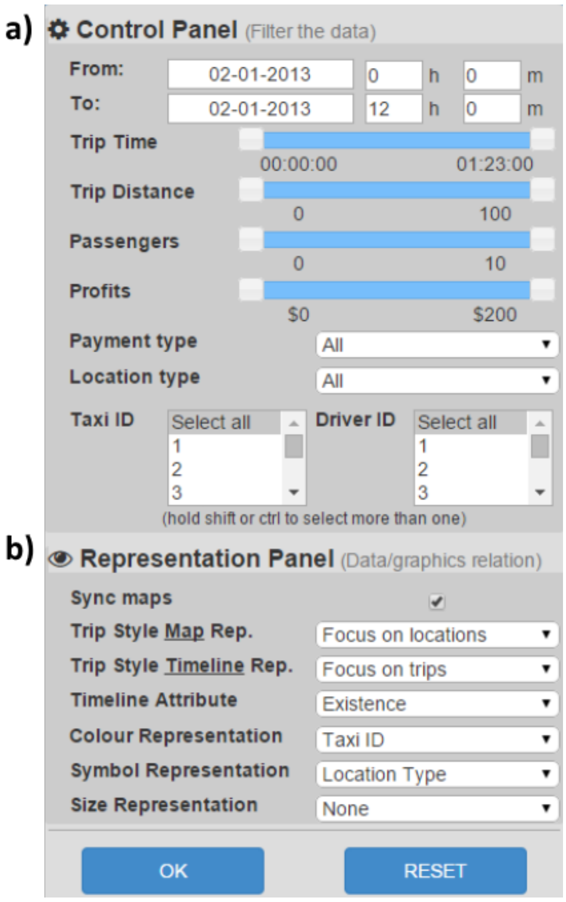

Figure 5.2: Control (a) and Representation (b) panels common on all three prototypes.

used to represent sets of categorical data, including: the time periods of the day (i.e., four colours, each representing a period of six hours), the location type (i.e., pickup/droop-off), the type of payment (i.e., cash, credit card, or note of credit), and the taxi or driver identifiers (i.e., one colour for each identifier). The point shape/symbol can be used to emphasize the same types of attributes described above, with the exception of the taxi and driver identifiers. Size can be used to represent, essentially, quantitative information, namely the number of passengers, the distance, duration, or the profit associated to each trip. The higher the value associated to the attribute, the larger/thicker the points and lines used to represent the data.

The third component (Figure 5.1 f and Figure 5.3 a) consists simply of an interactive legend, that informs the user about the meaning of the visual variables being used (e.g., which colour corresponds to which taxi). This component also allows users to quickly toggle the visibility of subsets of the data associated to those variables (e.g., hide/show the trajectories of a given taxi).

The fourth component (Figure 5.1 g and Figure 5.3 b) consists of a textual information panel, where all data associated to a single location or to the pickup/dropoff locations associated to a trip are displayed, depending on what was selected within the visualizations. That information includes: the name of the location (*where?*), the temporal instant for when the taxi passed through the location (*when?*), and the thematic information regarding the number of passengers, duration,

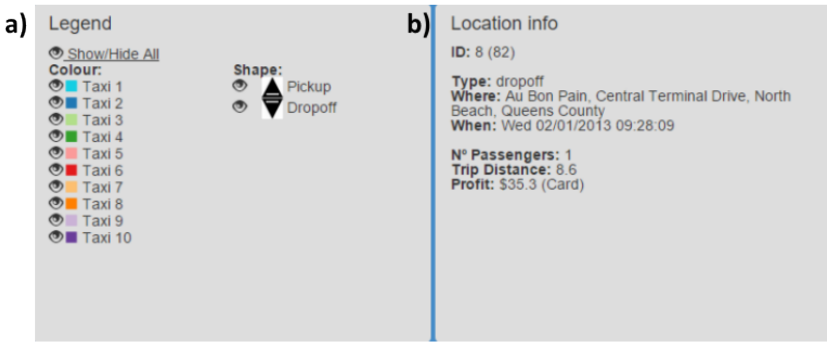

Figure 5.3: Legend panel (a) and Information panel (b) common on the three prototypes.

distance, and profit associated to a trip (*what?*).

The fifth component (Figure 5.1 h and Figure 5.4), the timeline, consists of a chart focused on displaying data from a temporal perspective. As seen in Figure 5.1 h and Figure 5.4 a, the timeline is composed by several rows, each one representing the movement data of one taxi. Each row contains several area icons, distributed horizontally, to indicate the periods of time in which movement was detected. In these icons, the larger the area, the longer the duration of the trip. However, this can be changed with the representation panel, where the user can select points (i.e., one per detected pickup and dropoff location) instead of area icons. This feature can be particularly helpful when the user is just interested in visualizing the temporal distribution of a certain type of location. The user can also associate quantitative thematic attributes to the timeline, like number of passengers, distance, duration, or profit associated to a trip, to visualize their evolution through time, in an alternative timeline view (Figures 5.4 a and b). Regardless of the type of view used, it is important to mention that the representation styles used for the time points and periods are similar to the ones used in the map-visualizations described below. Moreover, similarly to the study presented in Chapter 3, the inclusion of this component was considered necessary for two main reasons: first, the timeline provides an additional, and often neglected, point of view over the data (time); second, due to the objective of this research in making the prototypes informationally equivalent, this component was also included to compensate the limitations with the 2D map in representing time. Indeed, temporal information can be acquired through the legend component or when the time periods of the day are represented on the map; however, these methods can still be considered limited compared to a timeline that also provides an overview of the temporal distribution of the recorded events.

The last component consists of the map visualization technique, which can either be a 2D map, a 3D STC, or a visualization using both techniques side-by-side, depending on the prototype being used (Figures 5.1 a, b, and c, respectively). The approach followed by the prototype combining both techniques is comparable to ST-TrajVis, presented in Chapter 2. However, despite their similarities in terms of how the techniques are displayed (i.e., one STC juxtaposed with a 2D map), this prototype provides an additional set of interactive controls, further explained below, that do not exist in ST-TrajVis and that support a more adequate interaction with both views separately

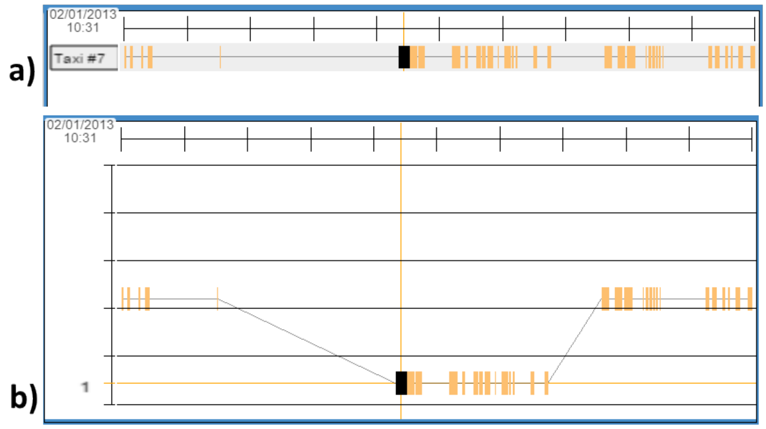

Figure 5.4: Alternative timeline views with the data of one taxi a) Display of event trips and their duration; b) Display of the event trips, their duration, and variations on quantitative thematic information.

or together. Furthermore, since this study focuses on the visualization and analysis of trajectories by inexperienced users, it was necessary to use data representations that should be both simple to use and understand. In all prototypes, the pickup and drop-off locations are represented with point icons, where their colour, size, and shape are directly related to the attributes selected in the representation panel, described above. In addition, to better illustrate the evolution of the taxis' positions, the points can also be connected with lines using the representation panel. Similarly to the point icons, line colour and thickness is determined according to the attributes that are selected. Specifically, lines connecting a pick-up and a drop-off locations, representing a trip, are coloured and sized according to those attributes, and lines connecting a drop-off to pick-up locations (i.e., when the driver is looking for passengers) are represented in a faded grey colour with minimum thickness. Figure 5.5 depicts an example of this representation, where a *non-trip line* is being selected in the STC and further highlighted in the timeline, with a similar representation.

To support interaction with the prototypes, both the map visualization and the timeline components allow the users to pan and zoom their views. In particular, the 2D map uses similar controls to common map-based applications (i.e., Google Maps) and allows the user to pan and zoom the map view with the left-mouse button and the mouse wheel, or the interactive buttons located at the bottom-right corner of the map, respectively. On the other hand, the STC and the timeline, require the right-mouse botton to pan and also the mouse-wheel to zoom the views (similarly to already common 3D design applications). Using the control panel located at the edge of the view, it is also possible to zoom the STC view and to pan/zoom the map displayed at the STC's base (see Figure 5.1 j). In addition, given its three dimensional properties, similarly to the prototypes described in the previous chapter, the STC also allows the user to rotate the visualization along its horizontal and vertical axes, with the left mouse button.

As shown in the previous figures, particularly Figures 5.5 and 5.6 a and b, the STC component used in this experiment is considerably different from those used in previous studies of this work,

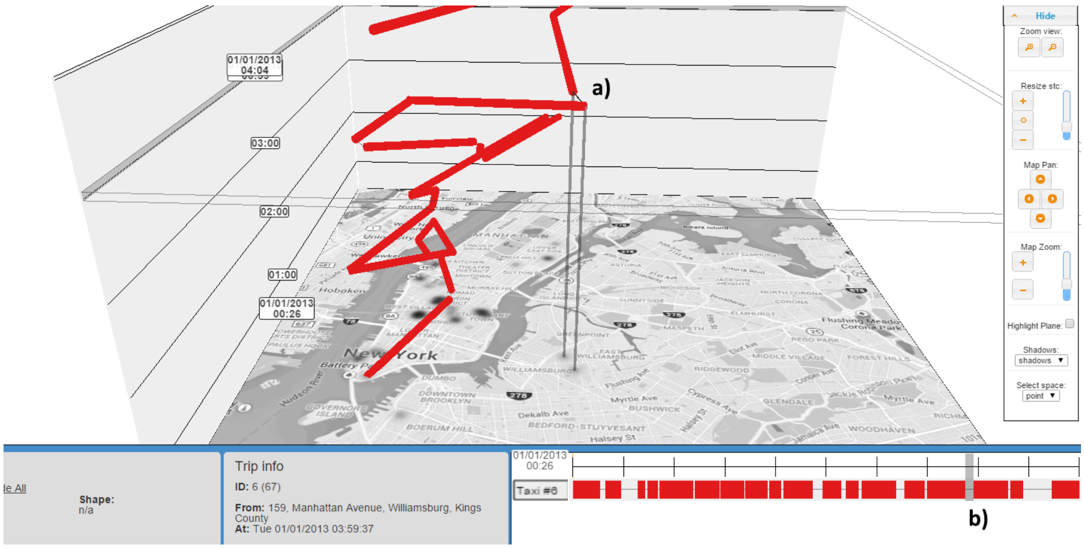

Figure 5.5: Selection of segment of a trajectory representing the search for a passenger in the STC (a), further highlighted in the timeline (b).

most noticeably in regards to its interactive control menu. Overall, the STC controls used in this experiment present some stylistic changes, now grouping the various interactive options around the same space and allowing the user to hide/reveal them if needed. Nevertheless, despite the new presentation style and some new interactive controls, this menu still incorporates the same features described in previous STCs. In particular, the user is able to zoom in/out the point of view in the STC or, pan/zoom the map plane displayed on the STC's base, or to rescale the STC's height. As discussed in the previous chapters, this later feature has been shown to be useful to the users, in the sense that it is comparable to an additional zoom operation. However, this type of zoom is focused on the representation of the STC's temporal dimension by allowing the manipulation of the temporal granularity of the data. In addition to help visualizing temporal information with a customizable level of detail, this feature can also help minimizing problems with information overlap, e.g., when stretching the STC to better distinguish different locations close in space and time.

Regarding the new types of controls, to improve the users' understanding of the spatial component of the data, the STC also provides an additional map plane, that can be dragged and locked anywhere along the STC's height. This feature was made optional, since, as discussed in the previous study, some users have shown some confusion when using it or did not consider it always necessary for the analysis of the data.

Furthermore, in both map components, the users are able to request the additional drawing of *shadows* or heatmaps over the map plane, such as depicted in Figures 5.5 and 5.7 a and b. This feature has two main purposes: first, it can show the areas with a higher density of information (such as what is, usually, intended when using a heatmap); second, in the STC view, it can be used as an additional spatial cue, since the highlighted information within the map plane will be located directly bellow the object within the STC. For a matter of simplicity and to minimize the

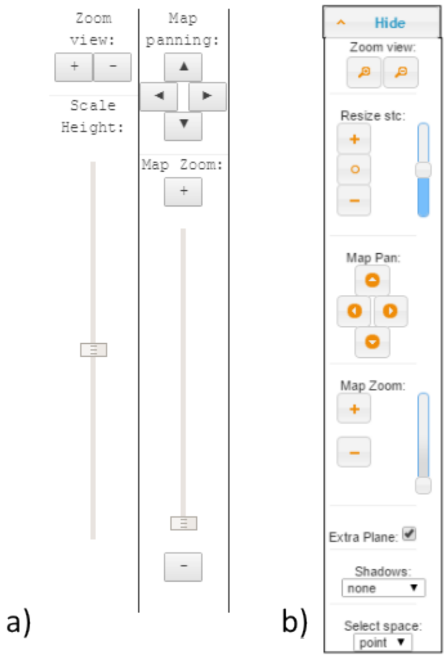

Figure 5.6: Evolution of the additional controls used in the STCs a) in Chapter 4 and b) current experiment.

possibility for a higher visual noise, this feature is deactivated by default.

Double clicking on the map plane of the prototypes allows the user to draw *points/lines* or *areas* inside the map views, based on the geographic location(s) selected (Figure 5.1 i). Overall, these can be used as a spatial reference when analysing the data or, more importantly, to minimize the risk of inexperienced users *loosing their location*, when changing the map view through panning or zooming operations. It must be acknowledged the fact that this feature could be used to support a more complex system for selecting and filtering information. However, in addition to being outside of the scope of this study, it could also significantly increase the complexity of use of the prototypes.

As usual in applications with multiple views, the selection of a trip or location point, in any of the map views or in the timeline, will highlight the correspondent data element in the other views. For instance, selecting a pick-up location in the map will highilight the corresponding time moment and trip in the timeline, and vice versa. Similarly, in the prototype combining both the 2D map and the STC techniques, any change or selection made to one map is immediately reflected in the other, further emphasizing the combination of both techniques and, expectedly, enabling an easier transition between views. In particular, panning/zooming the map or drawing a reference point or area, will cause the other map to be panned/zoomed or have a reference point/area drawn into the correspondent geographical area (see Figures 5.1 i, and 5.5). Nevertheless, the synchronization between the two maps can be deactivated through the representation panel, allowing the user to further customize how to visualize the data in the two different views.

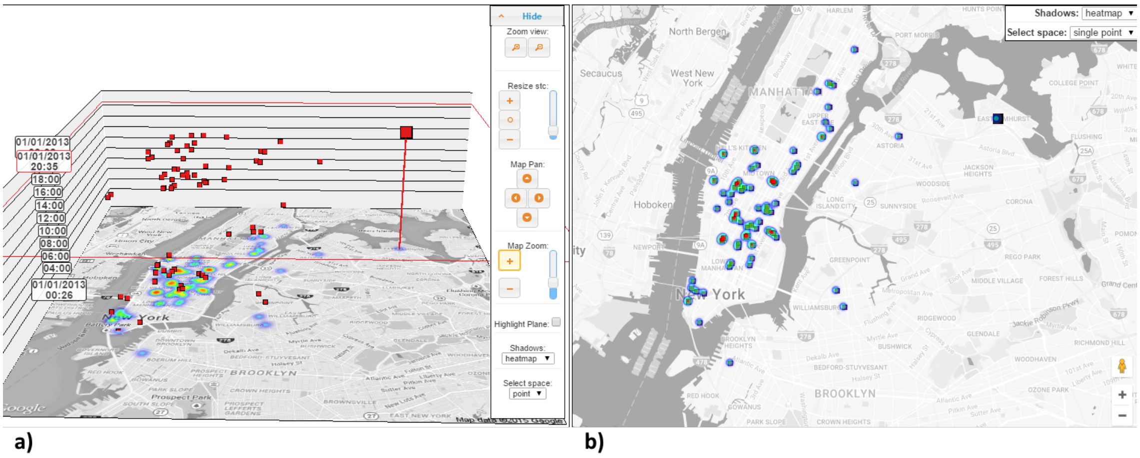

Figure 5.7: Heatmap functionality over the a) STC and b) 2D map.

5.3 User Study

To evaluate the viability of combining 2D static maps and 3D space-times cubes in the same visualization and to compare the differences in user performance and interaction strategies of when using a combined approach, against both visualization techniques individually, a user study was conducted with the participation of 30 volunteers (20 male, 10 female), aged between 19 and 40 (Av: 25.6, SD: 4.9). All participants were knowledgeable with computer applications to search for directions towards specific points of interest (e.g., Google Maps) and with applications that used 3D as a means of conveying or manipulating information (e.g., mapping applications, such as Google Earth, computer assisted design and modelling applications, such as Blender, and most commonly video games). Among all participants, four mentioned that despite being knowledgeable, they were not *fully comfortable* with 3D graphic applications. Finally, no participant was familiarized with New York's geography, nor had any significant experience with spatio-temporal nor trajectory data analysis. Consequently, all users were exposed to a similar context, since the possible effects of subjective bias and prior knowledge over the location were minimized, which could interfere in how users would look for information within the visualizations and, consequently, unbalance the results.

Based on the characteristics of the studied techniques and on the overall objectives of the experiment, this study's hypothesis were the following:

H1 It was expected a significant preference towards the prototype combining the 2D map and the STC visualizations. Despite the visual similarities with the other prototypes, and the redundancy of the data being displayed (i.e., same data in two similar maps), this prototype provides more choices to the participants, in terms of visualization and interaction, which could affect their preferences;

H2 A higher performance and a more significant usage of the 2D map in spatial tasks was also expected. This hypothesis is supported by previous studies that point towards the more advantageous use of 2D maps for these types of tasks, in addition to the fact that 2D maps are widely known and used visualization techniques, even by users with no experience with Geographic Information Systems (GIS);

H3 A higher performance with the STC is expected in tasks that require the analysis of the data's temporal component. Although the 2D map prototype also provides temporal information, through the use of the timeline component, the participants need to change their focus between the two different views, which, in turn, may affect their analysis. The STC, on the other hand, combines spatial and temporal components within the same view, expectedly allowing an easier analysis;

H4 It was expected an overall lower performance when more trajectories were displayed, due to the increasing data to be visualized and, consequently, the complexity of the visualization.

5.3.1 Tasks

Similarly to the previous experiment, to test these hypotheses the participants were asked to perform two tasks, *identify* and *compare*, which are based on the most common types of cartographic visualization objectives described in the literature (Roth, 2013), and correspond to the first level of tasks described in Figure 2.2. Comparatively to the previous experiments, both tasks were divided into two categories: *spatial* and *spatio-temporal*, depending on the main types of constraints associated with the task. This division was considered necessary for two main reasons. First, as initially described in Chapter 2, one of the most distinguishing factors between 2D maps and 3D STCs consists in how temporal information is presented. This may affect how the data is analysed, including spatial information, such as discussed in the results of Chapters 3 and 4, in which participants were less accurate at locating information with STCs and often transformed the STC based on their temporal representation to better visualize spatial information. Second, as observable from the results of Chapters 3 and 4, the most relevant differences between the techniques were detected in conditions where the the user had to analyse data recorded in a time period or focused on the search and analysis of a location. Consequently, this subdivision results into four types of tasks, described as follows:

Spatial identification: required participants to detect information (e.g., *point of interest or taxi/driver id*) in which one or several taxis performed a certain action (e.g., *drop a passenger*), based on a spatial restriction (e.g., *which drivers dropped passengers outside of New York?*);

Spatio-temporal identification: required participants to identify information based on a spatio-temporal constraint (e.g., *which driver was the first to drop a passenger outside of New York, in a certain day?*);

Spatial comparison: required participants to analyse relations, similarities, or differences between data elements based on a spatial constraint (e.g., *which driver covered a wider geographical area?* or *which points of interest were the most visited?*);

Spatio-temporal comparison: required participants to analyse relations, similarities, or differences between data elements based on a spatio-temporal constraint (e.g., *which taxi had the most service in New York during a given time period?* or *in which time periods was a given location visited the most?*).

Furthermore, for each task, two independent variables were considered, namely: Visualization technique (*Vt*), with three levels, corresponding to the three previously described prototypes, namely 2D map, STC, and 2D map + STC; and number of moving objects (*Nt*), with two levels, 1 and 10 objects (taxis). Although in this experiment the variable *query category* is not explicitly considered, which would determine the focus of the task on different numbers of objects in a time moment or period, given the results of the previous studies, it was still considered important to analyse the effects of the volume of information over the user's analysis.

5.3.2 Procedure

The experiment followed a *within-subjects* design and all participants carried out each task individually, in a controlled laboratory environment. At the beginning of the experiment, the participants were briefed about the objectives of the study and the types of visualization techniques being used. Then, they viewed a demonstration of the prototypes that covered all of their functionalities. After that, the participants were asked to interact with the prototypes, in order to get used to the interactive controls, and were encouraged to clarify any doubts.

After the training phase, the participants were asked to perform the two tasks, first the *identify* and then the *compare* tasks. This order was decided based on two main factors: first, as a result of their characteristics, a higher difficulty in *compare* tasks was expected, since more than just detecting information, the participants would also have to analyse it; second, in order to compare different data elements it is usually necessary to identify those elements first. For instance, to detect which point of interest in New York was the most visited, it may be necessary to identify/locate where New York is and what is each point of interest. Consequently, this could produce a negative effect over the results, due to an over-simplification of the *identify* tasks.

Nevertheless, to mitigate sequence effects, the order in which the independent variables and the type of task (spatial or spatio-temporal) were presented was counterbalanced based on a *latin-square* design. As a result, each participant performed a total of 24 trials: 2 tasks x 2 task types x 3 *Vt* x 2 *Nt*.

In addition, to measure the results and assess the hypothesis, the following dependent variables were recorded:

Subjective preferences: at the end of each task, the participants were asked to rate each prototype on a 0 to 10 scale, based on their preferences regarding the use of each prototype for that specific task. Similarly to the previous study, this variable allows the evaluation of the participants' acceptability towards each visualization, individually and comparatively to the others;

Task accuracy: each task was rated on a 0 to 5 scale, based on the correction and level of detail of the answers given by the participants (e.g., if the task was to name the taxis that visited a certain location, the score assigned would be lower if the participant did not mention all taxis or named incorrect ones in the answer). The analysis of this variable allows to evaluate the participants' effectiveness and, thus, verify if they are able to reach satisfactory responses for the various tasks;

Task completion time: measured in seconds, this was recorded from the moment the participant started the task (by pressing a *START TASK* button) until selecting the option to save a final answer;

Number of actions: these were recorded automatically from the moment the participant would start the task until the task was finished and included map and view panning/zooming operations, rescaling the STC's height, highlighting information, or changing any other property in the visualization (e.g., filter a given subset of information). Alongside task completion time, these variables allow to evaluate the complexity of each task and how the prototypes help in their completion;

3D mitigation strategies (specifically for the prototypes using the STC component): based on the current view rotation of the STC and its height, two different strategies were taken into account for the mitigation of the STC's third dimension. These consist of the same strategies identified in the previous studies, specifically looking at the data from a bird's eye view (*BEV*) and reducing the STC's height to its minimum, effectively, converting it into a plane/flat view (*FlatV*). Similarly to the previous study, this variable allows to analyse when or how often the participants may feel the necessity to convert the STC into a 2D map. More importantly, however, this also allows to verify if participants considered necessary to apply these strategies even if with a fully functional 2D map already available;

Ratio of use of the visualization components (specifically for the third prototype): it consists in the ratio between the total number of actions and the number of actions on each visualization component (2D map, STC, or timeline). This variable can also be used when analysing the first two prototypes, since these contain a timeline component which can be the focus of attention from the user when analysing time. However, the main focus of this study is related to the map component and thus, this variable, like the one above, allows to analyse the participants' visualization needs throughout the tasks and learn if or when one map component is used more than the other.

With the exception of the first two, all of the dependent variables described above were automatically recorded by the prototypes, during the execution of the tasks. In the context of this study, all prototypes generated log files representing the various user's interactive actions during the tasks. These actions included any specific operation over the views, such as resizing the STC,

Table 5.1: Mean participants' results in terms of subjective preferences.

	Identify		Compare	
S. Pref [0-10]	**Sp**	**SpT**	**Sp**	**SpT**
2DM	7.97	7.20	7.40	6.77
STC	5.40	6.10	6.27	6.83
3VC	7.97	8.23	8.03	8.20

panning/dragging, zooming (in/out), or rotating the views, and selecting (either by clicking or hovering) any feature with information within any of the views. Each recorded action was associated with a set of additional information, including, the mouse position within the application, timestamps (both in real-time and in relation to the beginning of the task), and information specific to the event, like the identification of the item selected within the view, or the position and rotation of the camera used when dragging or rotating the STC. Moreover, during the study, specially at the end of the tasks, the participants were encouraged to *think aloud*, share their opinions regarding the prototypes and answer to some questions depending on their actions, such as: which features they preferred the most, or in which situations they considered a certain feature, or a specific action they did, as being more useful or needed. These answers were registered manually during the study.

5.3.3 Results

This section describes the results obtained in the user study, particularly focused on the most statistically significant results and those close to be considered as such (i.e., not statistically significant but still with a low p-value). For a matter of simplification and readability, in the remainder of this chapter, the prototypes that implement the 2D map, the STC, and both techniques simultaneously will be, respectively, addressed as *2DM*, *STC*, and *3VC*. In addition, throughout the various tables depicting the results of this study, the type of task, spatial and spatio-temporal, will be, separately, addressed as *Sp* and *SpT*. The factors of the variable *number of moving objects* will be addressed as *Nt(1)* and *Nt(10)*, referring to 1 and 10 objects (taxis), respectively.

Subjective preferences

To compare the differences between the participants' subjective preferences, a Friedman's test was used, followed by a Wilcoxon Signed Rank test, using a Bonferroni correction, for pairwise comparisons (see Table 5.1 and Table 5.2).

In both the *spatial identification* and *comparison* tasks, significant differences were detected, in particular a lower preference towards *Vt(STC)* over *Vt(2DM)* and *Vt(3VC)*. In both *spatio-temporal* tasks, significant differences were also detected, as the results reveal a higher preference towards *Vt(3VC)* over *Vt(2DM)* and *Vt(STC)*.

Table 5.2: Statistic test results for the participants' subjective preferences. **n.s.s.** stands for *not statistically significant*.

	Identify		**Compare**	
	Sp	**SpT**	**Sp**	**SpT**
Friedman	$X^2(3) = 38.697$, $p < 0.001$	$X^2(3) = 28.625$, $p < 0.001$	$X^2(3) = 25.59$, $p < 0.001$	$X^2(3) = 19.049$, $p < 0.001$
Wilcoxon S.R:				
2DM - STC	$Z = -4.24$, $p < 0.001$	**n.s.s.**	$Z = -3.30$, $p = 0.001$	**n.s.s.**
2DM - 3VC	**n.s.s.**	$Z = -3.02$, $p = 0.003$	**n.s.s.**	$Z = -3.38$, $p = 0.001$
STC - 3VC	$Z = -4.32$, $p < 0.001$	$Z = -4.58$, $p < 0.001$	$Z = -3.64$, $p < 0.001$	$Z = -3.64$, $p < 0.001$

Table 5.3: Mean participants' results in terms of task accuracy.

	Identify				**Compare**			
Accuracy [0-5]	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
2DM	4.67	4.20	4.17	3.07	4.60	4.00	4.13	3.73
STC	4.73	3.80	4.53	4.17	4.73	3.83	4.53	4.37
3VC	4.77	4.27	4.40	4.00	4.60	4.00	4.43	4.40

Task Accuracy

A similar procedure was used to compare the differences in terms of the participants' accuracy (see Table 5.3 and Table 5.4).

In both the *spatial identification* and *comparison* tasks, the tests reveal a significantly lower accuracy in conditions with more trajectories (*Nt(10)*) when using *Vt(STC)*, comparatively to conditions with just one (*Nt(1)*). On the other hand, in the *spatio-temporal* tasks, the results suggest a lower accuracy with *Vt(2DM)* comparatively to the *Vt(STC)* and *Vt(3VC)*. In particular, in the *spatio-temporal identification* task, the test results reveal also a lower accuracy from participants in *Nt(10)* conditions, when using *Vt(2DM)*.

Task completion time

To assess the differences in terms of the amount of time taken by the participants to complete the various tasks, the data was subjected to a repeated measures ANOVA, followed by Bonferroni tests for pairwise comparisons (see Tables 5.5 and 5.6). In the *spatial identification* and *comparison* tasks, significant effects were found in the number of trajectories (*Nt*), in the visualization technique (*Vt*), and in both factors. In the *spatio-temporal identification* task, significant effects were also detected from *Nt* and *Vt*.

In all tasks, pairwise comparison tests revealed longer completion times with *Nt(10)* condi-

Table 5.4: Statistic test results for the participants' accuracy. **n.s.s.** stands for *not statistically significant*.

	Identify		Compare	
	Sp	**SpT**	**Sp**	**SpT**
Friedman	$X^2(5) = 31.16$ $p < 0.001$	$X^2(5) = 51.38$ $p < 0.001$	$X^2(5) = 34.37$ $p < 0.001$	$X^2(5) = 19.20$ $p = 0.002$
Wilcoxon S.R.				
***Vt* in *Nt(10)*:**				
2DM - STC	**n.s.s.**	$Z = -2.63$, $p = 0.009$	**n.s.s.**	$Z = -2.37$, $p = 0.018$
2DM - 3VC	**n.s.s.**	$Z = -2.66$, $p = 0.008$	**n.s.s.**	$Z = -2.62$, $p = 0.009$
Nt(1) - Nt(10):				
with 2DM	**n.s.s.**	$Z = -2.90$, $p = 0.004$	**n.s.s.**	**n.s.s.**
with STC	$Z = -3.33$, $p = 0.001$	**n.s.s.**	$Z = -3.03$, $p = 0.002$	**n.s.s.**

Table 5.5: Mean participants' results in terms of task completion times.

	Identify				Compare			
Time (sec)	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
2DM	96.23	120.85	103.53	183.55	71.52	117.32	112.33	144.23
STC	142.13	262.80	106.62	230.23	70.85	171.75	139.33	131.33
3VC	143.62	155.43	104.80	174.42	76.70	130.57	130.52	107.72

tions, comparatively to *Nt(1)*. On the other hand, in the *spatial* tasks, particularly with a higher number of trajectories, the results show significantly faster completion times with *Vt(2DM)*, comparatively to *Vt(STC)*.

Number of actions

To analyse the differences in terms of the number of interactive actions used in each task by the participants, a similar procedure to the one used in the previous section was applied (see Tables 5.7 and 5.8).

In both *spatial* and *spatio-temporal identification* tasks, similar differences were detected. In particular, the results show a significant effect from *Nt*, where participants performed a significantly higher number of actions in *Nt(10)* conditions. The results have also shown a significant effect from *Vt*. Pairwise comparison tests revealed a significantly smaller number of actions conducted by participants using *Vt(2DM)*, comparatively to the other visualizations. An identical result also was detected in the *spatio-temporal comparison* task, with the pairwise comparison tests revealing a significantly smaller number of actions with *Vt(2DM)*, comparatively to *Vt(3VC)*.

Table 5.6: Statistic test results for the participants' task completion times. **n.s.s.** stands for *not statistically significant*.

	Identify		Compare	
	Sp	**SpT**	**Sp**	**SpT**
Nt	$F(1,29) = 15.61$, $p < 0.001$	$F(1,29) = 56.9$, $p < 0.001$	$F(1,29) = 44.51$, $p > 0.001$	**n.s.s.**
Vt	$F(2,53) = 10.71$, $p < 0.001$	$F(2,53) = 3.43$, $p = 0.043$	$F(2,58) = 3.040$, $p = 0.05$	**n.s.s.**
2DM-STC	$p < 0.001$	**n.s.s.**	$p = 0.048$	**n.s.s.**
Nt x Vt	$F(2,51) = 15.61$, $p < 0.001$	**n.s.s.**	$F(2,43) = 44.51$, $p = 0.014$	**n.s.s.**

Table 5.7: Mean results in terms of the number of interactive actions performed by the participants to conduct the tasks.

	Identify				Compare			
N° Actions	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
2DM	18.43	40.43	21.87	45.37	21.60	41.07	22.80	29.50
STC	55.03	93.03	46.83	105.43	32.67	71.00	90.80	71.90
3VC	51.80	71.10	43.83	92.47	33.27	60.53	29.50	59.03

3D mitigation strategies

To study the use of 3D mitigation strategies by the participants and the differences between them in *spatial* and *spatio-temporal* tasks, when using *Vt(STC)* and *Vt(3VC)*, the time spent in *BEV*(*bird's eye view*) and in *FlatV*(*flat view*) was recorded and used for the conduction of a repeated measures ANOVA, followed by Bonferroni tests for pairwise comparisons (see Tables 5.9 and 5.10).

Regarding the use of *BEV* with *Vt(STC)*, the results show a significant effect on both tasks from *Nt* and the type of task, where participants used this strategy significantly more in conditions with more trajectories (*Nt(10)*) and in *spatial* tasks, respectively.

With respect to the use of *FlatV* with *Vt(STC)*, the tests reveal similar results, with a significant effect on both tasks from *Nt* and the type of task.

On the contrary, no significant differences were detected regarding the use of any of these strategies with *Vt(3VC)*.

Based on that information, followed an analysis of the differences in accuracy, based on the strategies used by the participants, when using *Vt(STC)*, on each task (see Table 5.11). For that, a Mann-Whitney U Test was conducted, which revealed significant differences in the *spatial identification* task with the condition *Nt(10)* ($U = 29$, $p = 0.01$), where participants were significantly more accurate when taking advantage of a *FlatV* approach over a *BEV* one.

Table 5.8: Statistic test results for the participants' number of actions.**n.s.s.** stands for *not statistically significant*.

	Identify		Compare	
	Sp	**SpT**	**Sp**	**SpT**
Nt	$F(1,29) = 15.49$ $p < 0.001$	$F(1,29) = 24.88$ $p < 0.001$	$F(1,2) = 14.48$ $p = 0.001$	**n.s.s.**
Vt	$F(1.7,49.2) = 8.74$ $p < 0.001$	$F(1.9,55.1) = 8.20$ $p < 0.001$	$F(1.6,48.9) = 4.05$ $p = 0.030$	$F(1.4,42.87) = 5.32$ $p = 0.015$
2DM-STC	$p = 0.006$	$p = 0.004$	$p = 0.027$	**n.s.s.**
2DM-3VC	$p = 0.006$	$p = 0.004$	$p = 0.049$	$p = 0.019$

Table 5.9: Mean participants' results in terms of a) time spent in *BEV* b) time spent in *FlatV* during the completion of the tasks.

	Identify				Compare			
Time spent (%)	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
a) BEV								
STC	65.88	189.83	22.55	84.60	39.12	91.72	24.28	35.35
3VC	29.07	6.63	5.12	6.77	6.92	1.45	3.22	4.00
b) FlatV								
STC	37.60	84.55	13.30	38.35	24.43	57.92	10.27	12.75
3VC	11.85	4.43	0.0	0.0	0.0	0.0	0.0	1.03

Ratio of use of the visualization components

To understand which visualization components were the most used by the participants during each task, an ANOVA test was conducted over the results, followed by a Bonferroni test for pairwise comparison (see Tables 5.12 and 5.13).

The tests revealed significant differences in both *spatial* tasks, where participants used the 2D map component generally more often than the STC and timeline components. Contrarily, in both *spatio-temporal* tasks, the results reveal that participants used the STC component more often than the 2D map and timeline components.

Gender comparison

Comparatively to the previous experiments, this study had the participation of a higher percentage of female subjects (30%). In addition, as addressed in Chapter 2, previous research suggests that the gender of the user can have an effect over the interaction, particularly when visualizing information in a 3D view (Vecchi and Girelli, 1998). As such, although not being the focus of this study, it was decided to complement its analysis by comparing the differences of results between male and female participants.

Table 5.14 shows the average accuracy results obtained by the participants. The data was subjected to Mann-Whitney U Tests over each task condition for the two genders. The results

Table 5.10: Statistic test results for the participants' time spent when flattening the STC (**FlatV**) or looking at the data from a *birds eye view perspective* (**BEV**), using *Vt(STC)*. **TT** stands for the factor task type. **n.s.s.** stands for *not statistically significant*.

	FlatV		**BEV**	
	Identify	**Compare**	**Identify**	**Compare**
Nt	$F(1,29) = 9.70$ $p = 0.004$	$F(1,29) = 9.73$ $p = 0.004$	$F(1,29) = 13.51$ $p = 0.001$	$F(1,29) = 8.51$ $p = 0.007$
TT	$F(1,29) = 12.25$ $p = 0.002$	$F(1,29) = 4.53$ $p = 0.042$	$F(1,29) = 52.23$ $p < 0.001$	$F(1,29) = 9.11$ $p = 0.005$
Nt x TT	**n.s.s.**	**n.s.s.**	$F(1,29) = 4.61$ $p = 0.04$	**n.s.s.**

Table 5.11: Mean participants' results in terms of accuracy (0 to 5 scale) obtained with both strategies with the STC.

	Identify				**Compare**			
Accuracy [0-5]	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
STC - BEV	4.54	**3.20**	4.50	4.38	3.85	4.86	4.86	4.38
STC - FlatV	4.92	**4.05**	4.33	3.91	3.75	4.25	4.25	4.43

Table 5.12: Mean results in terms of the ratio of use of the visualization components by the participants, when interacting with *Vt(3VC)*.

	Identify				**Compare**			
Comp. use (%)	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
2DM	57.23	54.77	21.19	27.56	41.45	52.51	17.81	19.08
STC	25.42	12.42	39.73	49.07	15.03	15.60	47.51	51.62
Timeline	14.64	6.25	28.78	11.23	3.43	5.47	25.65	21.26

of the tests revealed a significantly higher accuracy from male participants in *spatial comparison* tasks, when using *Vt(STC)* (*Nt(1)*: $U = 80, p = 0.042$; *Nt(10)*: $U = 54.5, p = 0.030$). No statistically significant results were detected when comparing task completion time, number of actions, nor subjective preferences.

5.4 Discussion

This section addresses the results presented above, commenting the relation with the hypothesis they support, the results of previous studies (Section 2.5), and, more importantly, how the results highlighted relevant aspects regarding the visualization techniques studied and the interaction strategies used with them.

5.4.1 Hypotheses and Independent Factors Analysis

The results obtained in this study support the first hypothesis (H1), in which a higher preference towards *Vt(3VC)* was expected. As seen in the results, *Vt(3VC)* was constantly among the preferred techniques by the participants, both from the analysis of average values and the results of the statistical difference analysis. Although not significantly different from *Vt(2DM)* in *spatial* tasks, *Vt(3VC)* was the preferred visualization in *spatio-temporal* analysis tasks.

With the second hypothesis (H2), a higher performance and usage of *Vt(2DM)* in *spatial* tasks was expected. For both cases, the results partially support this hypothesis. Analysing task completion times and number of actions in *spatial* tasks shows that the participants were able to complete the tasks faster and with a smaller number of actions using *Vt(2DM)*, even though this fact had no significant effect over the participants' accuracy. The analysis of the component usage on *Vt(3VC)* shows that for *spatial* tasks, the 2D map component was the most used by the participants. These results are particularly interesting when compared to those of the second experiment, presented in Chapter 3, in which participants using the STC prototype were significantly less accurate in *location* tasks. Although the argument can still be made that 2D maps are preferable to STCs in *location*-based tasks, the lack of evidence towards one being better than the other in terms of accuracy, further emphasizes the results presented in the previous chapter, after the study of possible improvements over the STC technique, which resulted in an overall better performance than when using a simple STC.

In the case of the third hypothesis (H3), it was expected a higher performance and usage of *Vt(STC)* in *spatio-temporal* tasks. Similarly to the previous case, the results also go in agreement with this hypothesis. Despite the fact that the participants had to perform a higher number of actions to complete the tasks with *Vt(STC)*, analysing the results in terms of accuracy shows that they were, generally, more accurate in both *spatio-temporal* tasks, yet, not significantly slower as in *spatial* tasks. In addition, the analysis of the component usage on *Vt(3VC)* also shows that the STC component was used more frequently in *spatio-temporal tasks*.

Table 5.13: Statistic test results for the participants' technique/component usage in *Vt(3VC)*. **n.s.s.** stands for *not statistically significant*.

	Sp Nt(1)	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
Identify	$F(2,84) = 20.16$, $p < 0.001$	$F(2,84) = 53.75$, $p < 0.001$	$F(2,84) = 3.16$, $p = 0.048$	$F(2,84) = 14.01$, $p < 0.001$
2DM - STC	$p < 0.001$	$p < 0.001$	$p = 0.049$	$p = 0.006$
2DM - Timeline	$p < 0.001$	$p < 0.001$	**n.s.s.**	**n.s.s.**
STC - Timeline	**n.s.s.**	**n.s.s.**	**n.s.s.**	$p < 0.001$
Compare	$F(2,84) = 6.14$, $p = 0.003$	$F(2,84) = 36.67$, $p < 0.001$	$F(2,84) = 6.99$, $p = 0.002$	$F(2,84) = 10.55$, $p < 0.001$
2DM - STC	**n.s.s.**	$p < 0.001$	$p = 0.002$	$p < 0.001$
2DM - Timeline	$p = 0.003$	$p < 0.001$	**n.s.s.**	**n.s.s.**
STC - Timeline	**n.s.s.**	**n.s.s.**	$p < 0.013$	$p < 0.001$

Table 5.14: Mean participants' results in terms of accuracy, separated by gender.

	Identify				**Compare**			
Accuracy [0-5]	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**	**Sp Nt(1)**	**Sp Nt(10)**	**SpT Nt(1)**	**SpT Nt(10)**
2DM - M	4.65	4.35	4.40	3.15	4.60	4.20	3.95	3.85
2DM - F	4.70	3.90	3.70	2.90	4.60	3.60	4.50	3.50
STC - M	4.75	4.10	4.80	4.30	**4.99**	**4.10**	4.70	4.25
STC - F	4.70	3.20	4.00	3.90	**4.20**	**3.30**	4.20	4.60
3VC - M	4.70	4.50	4.45	4.10	4.80	4.05	4.55	4.45
3VC - F	4.90	3.80	4.30	3.80	4.20	3.90	4.20	4.30

In addition to the type of task (*spatial* or *spatio-temporal*), the results allow to discuss the scalability of the different techniques, as they point towards the fact that increasing the number of trajectories (*Nt*) has an effect on the interaction with the techniques. Expectedly, in terms of task completion times and number of actions, the results show that, with more trajectories, participants took more time and had to perform more actions to complete the tasks. More interestingly, however, in terms of accuracy, the effects of this variable were not only noticeable in *Vt(STC)* in *spatial* tasks, but also in *Vt(2DM)* in *spatio-temporal identification* tasks. These results thus support the fourth hypothesis (H4), where it was expected a higher performance in tasks with less trajectories displayed.

As previously stated, although it was not the focus of this study, it was also considered relevant to compare the results obtained in terms of gender. Interestingly, the results only show significant differences in the *spatial comparison* task, where male participants were more accurate than female participants with *Vt(STC)*. By association, looking at the average results suggests the possibility of male participants being, generally, more effective in these tasks, particularly with 3D visualizations like *Vt(STC)*. However, it is important to acknowledge that the compared groups were unbalanced (20 males, 10 females), thus preventing from taking any definitive conclusion, but still allowing to highlight the possible differences.

5.4.2 Interaction Strategies and Feedback

To further illustrate the results presented above and comment on the participants' strategies and feedback, several heatmaps were generated representing the distribution of all participants' interactive actions, during the *spatial* and *spatio-temporal* tasks, respectively, for all three prototypes. For the remainder of this document these will be referred to as *interaction maps*, in opposition to the the heatmaps available in the prototypes.

The comparison of both interaction maps obtained with *Vt(2DM)* (see Figure 5.8) shows a noticeable higher usage of the timeline component in *spatio-temporal* tasks. This goes in agreement with the the participants' feedback, who commented that it was more difficult to work with the 2D map when temporal information was necessary and, for that reason, they considered that the timeline component was very useful to compensate the map's limitations. Consequently, in

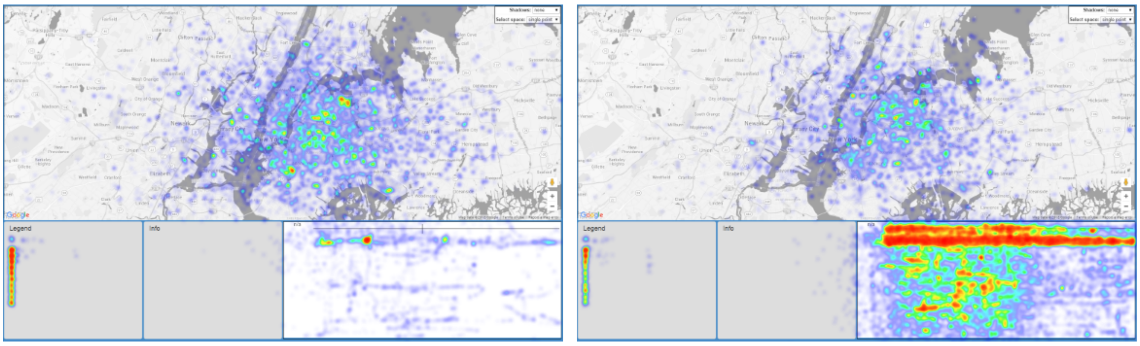

Figure 5.8: Interaction maps regarding *Vt(2DM)* in all spatial (left) and spatio-temporal tasks (right).

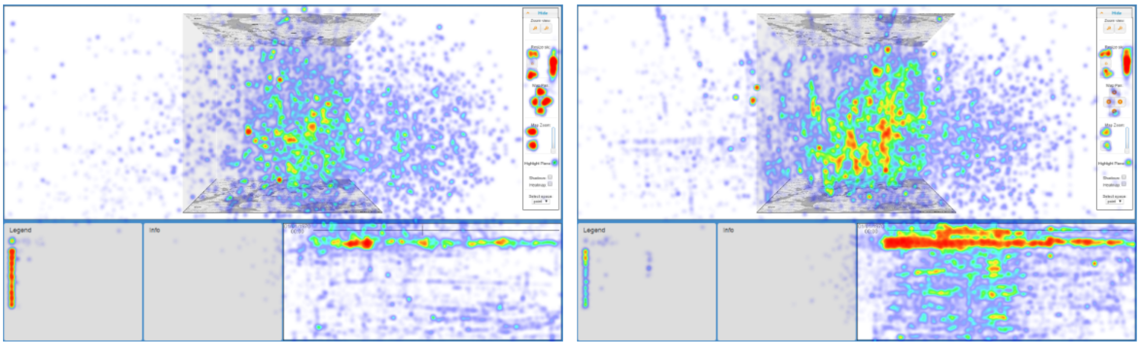

Figure 5.9: Interaction maps regarding *Vt(STC)* in all spatial (left) and spatio-temporal tasks (right).

spatio-temporal tasks, it was possible to observe that participants switched often between the two components. While some would focus first on the analysis of a geographical area, with the 2D map, and then try to understand it temporally with the timeline, others focused mostly on analysing the timeline and then looked for confirmation in the 2D map. These last participants mentioned that, due to the fact that highlighting one view would highlight the others, and that the data was temporally ordered within the timeline, they were able to understand better the movement of the taxi/driver(s). These last comments can also help explaining the fact that, as visible in Figure 5.8, the timeline component was also used in *spatial* tasks, even if less noticeably. Although the tasks did not require a precise understanding of time (e.g., knowing the date of a certain event), the linear representation provided by the timeline associated to the highlight of the map component and the sequential nature of trajectories can make the timeline a simpler and more apelative tool to interact with, allowing the user to follow the movements of one or several taxi/driver(s).

The analysis of the interaction maps obtained with *Vt(STC)* (see Figure 5.9) emphasizes the dependency of STCs towards proper interaction mechanisms, as these are noticeably more disperse and have areas with a higher concentration of events.

Comparing both types of tasks shows, once again, that the timeline was used more often in *spatio-temporal* tasks, while the STC's map controls (i.e., pan and zooming mechanisms at the right-side of the STC) were used more often in *spatial* tasks. Such as before, the timeline

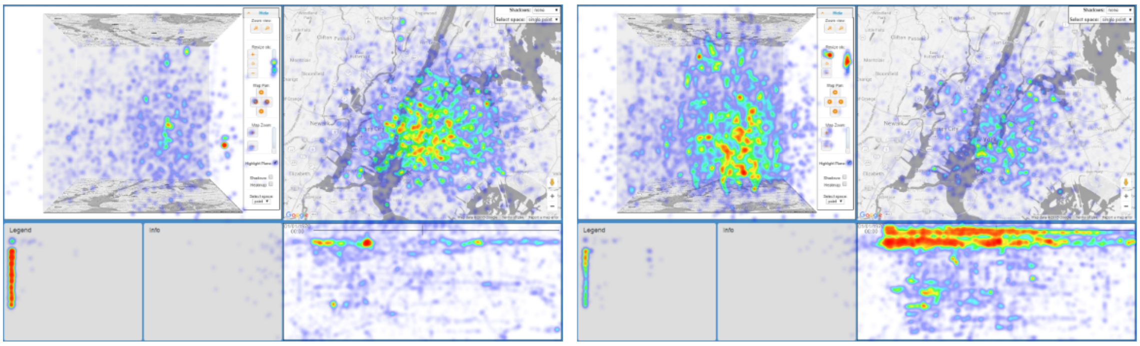

Figure 5.10: Interaction maps regarding *Vt(3VC)* in all spatial (left) and spatio-temporal tasks (right).

component was used in both types of tasks, which can be surprising, considering the fact that the STC already displays information from a temporal perspective. However, in addition to the reasons presented above regarding the use of the timeline from *Vt(2DM)* in *spatial* tasks, some users have mentioned that dealing with third dimension of the STC, although useful, could be challenging, particularly when a larger number of trajectories was displayed.

Another important set of features highlighted by these interaction maps refers to STC's controls regarding its additional map plane and temporal granularity. Regarding the additional map plane, the participants praised this feature, since it allowed them to focus it on a specific time location or slide it up/downwards to understand better the evolution of the taxi/driver(s) position over time. Regarding the temporal granularity controls, the participants justified its usefulness based on the fact that it allowed them to visualize, with more detail, variations of information in close periods of time and, as previously discussed, to convert the 3D STC into a 2D plane. This later strategy was mostly used in *spatial* tasks; however, some participants have also used it in *spatio-temporal* tasks, alternating between a flattened STC view and a *normal* one, depending on their interest on spatial or temporal information, respectively. This usage can be also seen, to a certain extent, in the interaction maps, in which the controls (buttons) referring to the STC's height were noticeably more used in *spatial* tasks (see Figure 5.6 for reference).

The analysis of the interaction maps with *Vt(3VC)* (see Figure 5.10) further emphasizes how the 2D map component was used the most in *spatial* tasks, while the STC and the timeline components were used more often in *spatio-temporal* tasks. Despite the predominance of specific components for different types of tasks, it is worth noticing that both maps were used.

In both types of tasks, it is also possible to see that the STC controls were used less than with *Vt(STC)*. More relevant, however, is the representation regarding the usage of the STC's temporal granularity controls. It is possible to see that in *spatial* tasks, the slider controlling the STC's height is mostly highlighted on its lower half, showing that the majority of the interaction with that control consisted in reducing the STC's height (see Figure 5.11). Comparatively, in *spatio-temporal* tasks, a wider area around these controls is highlighted, suggesting that users did not just reduce the STC's height or visualized it with a very low scale.

Figure 5.11: Interaction maps with *Vt(3VC)*, particularly with the STC controls in spatial (left) and spatio-temporal tasks (right).

Another important component highlighted by the various interactions maps refers to the legend component and, by extension, the customizable representation features. Through these mechanisms, the participants often used the following strategies:

1. quickly remove all information, to analyse or find a specific geographical area;

2. filter the data from one mover and compare it with another or a group of movers;

3. remove information considered to be irrelevant after an analysis (e.g., trajectories/locations that after being analysed were considered to go against some criteria used to solve a task);

4. insert information, previously removed after an analysis (e.g., compare the geographical distribution of two movers by alternating their visibility); or

5. emphasize certain aspects of the data the participant considered useful for the completion of the task, even if in a redundant way (e.g., set visual variables to represent temporal information to help comparing different time periods).

The participants also mentioned that being able to customize how the data should be displayed, helped them highlighting information they considered relevant and, more importantly, made them feel more control over the interaction.

Finally, the participants also appreciated the possibility of highlighting a specific geographical area (Figure 5.1 i), having used this feature to mark an area of interest as relevant during the analyses. The participants mentioned that this feature was more helpful in *Vt(STC)* than in *Vt(2DM)*, since it helped minimizing the negative perspective effects of the third dimension, while not losing

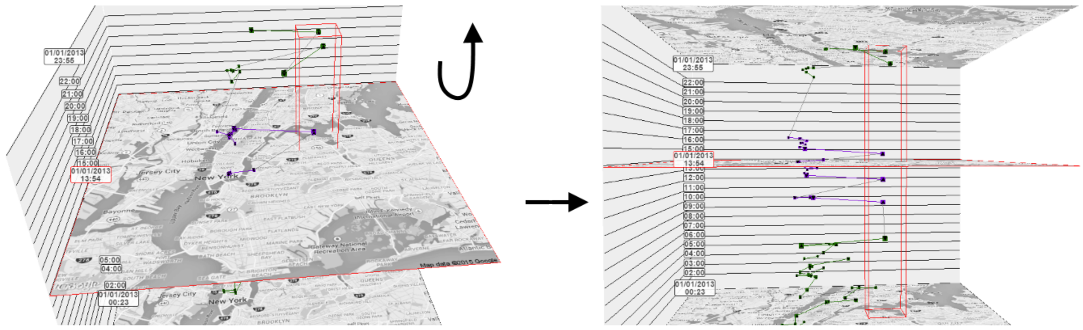

Figure 5.12: Example of the combination of the auxiliary map plane and the spatial area highlighting features.

temporal information, ultimately, helping them understanding better how the information was distributed over the geographical area. Interestingly, some participants combined this functionality with the auxiliary moveable plane, when trying to limit the information they wanted to focus on, while still having an overview of the remaining data (see Figure 5.12). They also mentioned appreciating this functionality in *Vt(3VC)*, since they could highlight an area in the 2D map component and then analyse the correspondent area in the STC, further helping them transitioning between one component to the other.

5.4.3 Combining 2D maps and STCs

The results of this study point towards the advantages of *Vt(3VC)*. First, as discussed above, the participants have shown a higher preference towards this technique. Throughout the study, the participants have justified their preference based on the fact that they considered that the 2D map helped overcoming the STC's limitations when analysing spatial information. As expected, some participants have also praised the fact that this prototype provided more choices to analyse the data. In addition, they mentioned that the synchronization between the two maps was helpful, since it allowed them to easily change the focus between views. Despite being given the possibility to deactivate this synchronization, none of the participants considered necessary to do it, highlighting the usefulness of combining both techniques.

Analysing the results in terms of task accuracy also shows that this prototype was never significantly outperformed by the others, suggesting that, at worst, this visualization is as useful as the better technique from *Vt(2DM)* and *Vt(STC)*.

As discussed above, the analysis of the component usage on *Vt(3VC)* shows that the participants have used both maps in the various tasks, suggesting the interest and necessity of having both techniques available to the user in scenarios where the focus of the task may vary between spatial and temporal information. The analysis of the 3D mitigation strategies used also help supporting this claim, since, in *spatial identification* tasks, participants that flattened the STC were more accurate than those that tried to identify the data from a *bird's eye view* perspective. In turn, these results further emphasize the advantages of being able to switch between both visualizations, the advantages of extracting 2D information from a 3D STC, and how that can be achieved.

Based on the differences that were detected in terms of gender, it may be appropriate to provide both techniques to the users as a means of minimizing the possible effects of this factor over the interaction. Nevertheless, further studies should be conducted to properly access this new hypothesis.

It is important to notice, however, that combining 2D maps and STCs will increase the complexity of the visualization, possibly requiring users to perform a higher number of actions to complete a given task. Interestingly, as evidenced by the results, particularly in terms of subjective preferences, users may still feel more comfortable with this combined approach, comparatively to just using one technique. In addition, combining the two techniques, as seen in this study, will also support a somewhat significant redundancy in the visualization. Despite their differences, 2D maps and STCs still share several visual similarities (e.g., the spatial map plane). Although the majority of the participants did not find that redundancy detrimental for their interaction, 3 participants raised their concern towards this factor in less complex tasks, such as *identification* tasks focused on one trajectory.

Finally, the results of this experiment suggest the interest in further studying different approaches to combine 2D maps and STC visualizations, like those suggested by Kveladze et al. (2015), which shows an application that supports the user switching between a 2D map and a 3D STC by pressing a button external to the map view. Arguably, an identical approach was already used with *Vt(STC)*, as the temporal granularity controls allow the user to flatten the STC into a 2D plane. However, as seen in this study, the consequence of this functionality was not clear to all participants, as several did not use it, preferring or only considering necessary to look at the data from a *birds-eye* perspective. This suggests that this functionality should either be better emphasized to the user or a different approach should be explored.

5.5 Summary

This chapter continued to address the challenges of improving existing map-based techniques for the visualization of trajectory data. In particular, this chapter was focused on the study the adequacy of combining 2D static maps with 3D space-time cubes within the same visualization, under the hypothesis that using both techniques would make one technique minimize the limitations found over the other. To achieve this goal, three prototypes were developed, one implementing a 2D static map, another implementing a 3D space-time cube, and a third combining both techniques in the same view. All prototypes included an interactive timeline, also connected to the map visualizations. Moreover, all prototypes supported various interactive features, in part, based on the results of the studies presented in Chapters 3 and 4.

The three prototypes were compared in a user study, focused on an inexperienced population in terms of trajectory data analysis. Considering the most distinguishing difference between 2D maps and 3D STCs consists in their temporal representation, the study took into account the main types of visual analysis tasks and the effects of focusing on the spatial or spatio-temporal components of the data. In addition to the analysis of the users' performance and opinions regarding each visualization, the various strategies and sequences of actions performed were also considered to evaluate the prototypes when completing the tasks.

Overall, the results of this experiment address the objectives presented at the beginning of this chapter, namely to: i) expand the existent knowledge that distinguishes the benefits of using 2D maps compared to STCs; ii) explore whether it makes sense to combine both techniques; iii) empirically validate the combination of these techniques; and iv) understand and analyse the types of strategies used by inexperienced users to acquire information with the different techniques.

Despite being focused on the combination of both techniques, this experiment allowed to further support and increase the knowledge of the differences between 2D static maps and 3D space-time cubes. The results point, once again, towards 2D maps being preferred over STCs and being, generally, more adequate in tasks heavily focused on the spatial components of the data. In comparison, STCs were shown to be more adequate in tasks that also with temporal information. Interestingly, some of the results, namely in terms of accuracy, were less significant or expressive than those of previous experiences, further suggesting the importance of the STC improvements addressed in Chapter 4.

The obtained results also suggest that combining 2D maps and STCs can be a successful improvement over just using one of the techniques. Not only this approach was always among the preferred by the participants, it was also never significantly outperformed by the other two. The analysis of the participants' actions also support this achievement. Notable examples include the users' attempts at converting the STC prototype into a 2D map, similarly to what was discussed in the previous chapters, or the usage of the different maps depending on the focus of the task, namely spatial or spatio-temporal information. Nevertheless, these results also suggest the study of additional ways to combine these techniques, particularly by being able to switch between them, alternatively to having them both visible at the same time (as used in this study).

Based on the participants' comments and their interactive actions, the results of this experiment also support the usefulness of various interactive features and components present in the prototypes, including: the timeline component, which helped the participants in tasks dependent on the analysis of temporal information, particularly with the 2D map prototype; the ability to highlight specific geographical areas, particularly with the prototypes implementing the STC technique; and the interactive controls in the STC prototypes, particularly related to the additional spatial cues and temporal granularity control.

Figure 5.13 depicts a schematic summary of the contributions of this chapter in the scope of the research reported on this dissertation.

Ultimately, the experiments described throughout this and the previous two chapters provide a set of relevant results that clarify and expand on the existing knowledge associated to 2D static maps and 3D space-time cubes. In particular, these results address the main similarities and differences between the two techniques, based on various metrics, their adequacy in different types of tasks, and, perhaps more importantly, various types of interactive features and controls to improve their adequacy.

Nevertheless, it is important to acknowledge the fact that, throughout these studies, no specific context of use was considered. This decision was taken based on two main reasons. First, to better focus on the study of these techniques in terms of their spatial and temporal representations. Second, to avoid a possible bias effect from the study participants, which could adulterate the results. The next chapter addresses this lack of a specific context of use and advances the study of these techniques. By taking advantage of the knowledge acquired so far, a new user study is described, assessing the adequacy of using 2D static maps and 3D space-time cubes for the visualization of game-related trajectory data, by players likely inexperienced in terms of trajectory data analysis, yet interested in understanding their own data.

Step 1: Related Work Analysis and Simplified Task Model (TM)

Ch2. Interactive Visualization of Trajectory Data

Results: See Figure 2.2.- For Task Model (TM) separated by 4 levels (lvl 1-4)

Step 2: Compare 2D Maps and 3D STCs

Ch3. Comparing 2D Static Maps and 3D Space-Time Cubes

Study #1 (5 participants)

Preliminary assessment of inexperienced users' ability to analyse trajectory data

Method:
- Task solving with *ST-TrajVis*
- **Data source**: GeoLife (1 mover: Nov. 2008)
- **Tasks**: *Identify and Compare* (**TM**-lvl1)
- **Logged data**: Time, Accuracy, and Preferences

Results
- Small learning curve
- Combining a 2D map with an STC was well received
- Different roles: *Main* visualization vs supportive map
  - Varied between users
- Interest in quick visual filtering mechanisms

Study #2 (16 participants)
- Compare how adequate 2D maps and STCs are for visualization task solving by inexperienced users

Method:
- Compare 2 prototypes using map visualizations (2D map or STC) and a timeline
- **Data source**: GeoLife (32 movers: 2008-2009)
- **Tasks**: *Identify, Locate, Compare, Associate* (**TM**-lvl2)
- **Query categories**: *(elementary/general) when*, *(elementary/general) what+where*
- **Logged data**: Time, Accuracy, and Preferences

Results:
- Higher preference towards 2D map
- 2D map more effective on *location* tasks
- STC more effective on *association* tasks
- Higher performance on *elementary tasks*
- Additional comments on combining 2D map and STC
- Attempts of converting STC into 2D map

Step 3: Improve Existing Techniques

Ch4. Improving Spatial Awareness in Space-Time Cubes

Study #3: (20 participants)
- Identify strategies used to obtain spatial information on an STC
- Identify interactive techniques to help obtaining spatial information
- Compare user performance with those methods

Method:
- Compare 3 STC prototypes (*P1*: simple; *P2*: moveable plane; *P3*: map overview)
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: Same as before
- **Logged data**: Same as before + Nº actions and 3D mitigation strategies

Results:
- Higher preference towards *P2 and P3*
- Higher performance with *P2 and P3*
- Lower performance in tasks with more *visual noise*
- *P2* is good for finding locations and study one mover
- *P2* can interfere with the interaction, due to occlusion
- *P3* provides a better geographical overview
- Additional comments regarding the combination of both maps
- Attempts at converting STCs into a 2D map: Birds-eye view (*BEV*) and flattening the STC's height (*FlatV*)

Ch5. Combining 2D Static Maps and 3D Space-Time Cubes

Study #4 (30 participants)
- Expand existing knowledge regarding using 2D maps and STCs
- Explore if it makes sense to combine both techniques
- Empirically validate the combination of these techniques
- Analyse the types of strategies used to acquire information, alongside user performance and acceptance.

Method:
- Compare 3 prototypes using map visualizations (2D map, *STC*, and *3VC*: using both techniques) and a timeline
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: *spatial* and *spatio-temporal*
- **Density of information**: 1 or 10 movers
- **Logged data**: Same as before + % of component use

Results:
- Combining both techniques is useful
- *2D Map* is faster and less interactive demanding on *spatial identification* tasks, but not significantly more accurate then *STC*
- *STC* is better on *spatio-temporal* tasks
- *3VC* is preferred in *spatio-temporal comparison* tasks
- *3VC* reduces the need to transform STCs into 2D maps
- *FlatV* is a better approach than *BEV*

Step 4: Case Study

Step 5: Identification of Guidelines

Figure 5.13: Summary of this research and the contributions of Chapter 5.

Chapter 6

Case Study: Visualization of Trajectory Data from Online Games

The previous chapters allowed to retrieve a set of relevant results associated with 2D static maps and 3D space-time cubes (STCs), from the perspective of an inexperienced population in terms of trajectory data analysis.

As pointed out in Chapter 1, despite acknowledging the importance of thematic information in the representation and analysis of trajectory data, this work is essentially focused on the simultaneous representation of the spatial and temporal information. Consequently, the studies presented in the previous chapters cannot be considered as being particularly focused on a specific context of use. This is reflected by the fact that the datasets used throughout the previous experiments have always been different, even if some came from the same source. Nevertheless, this approach has some advantages since, in addition to help validating these techniques with different types of datasets, this approach also helped preventing possible bias effects that the participants could have. Regardless of that bias being over the data or the tasks, ignoring this possibility could adulterate the validity of the results.

Despite these factors, it is still necessary to validate the knowledge and results acquired in the previous studies using a more specific application-based context. As such, it was decided to apply this knowledge into the area of visual-game analytics, specifically associated with the visualization of trajectories generated in games of the MOBA genre (Multiplayer Online Battle Arena), such as *League of Legends* (*LoL*) (Riot Games, 2009) .

The following sections will expand upon these topics. Section 6.1 presents some background and motivation for the selection of this context as the case study. Section 6.2 describes the *LoL* game alongside its importance and validity for this research. Section 6.3, describes VisLoL, a prototype created based on the results obtained in the studies presented in the previous chapters. Following this description, Section 6.4 describes the user study conducted with VisLoL, which aimed at: i) studying if inexperienced users are able to reach meaningful conclusions by using 2D static maps and 3D space-time cubes when visualizing personal trajectories, in a game-related context, similarly to what is expected with already existing tools; ii) understanding how the analysis of the data differs from using a map-based application and already existing tools, e.g., which different metrics and interactive strategies are used; and iii) studying the possible effects of knowing whose data is being visualized over the analysis, due to the possible personal motivations associated with the interaction. Section 6.5 presents a discussion of the results obtained in the study.

Finally, it is important to acknowledge that part of this work was developed in the context of an MSc work (Carreiro, 2016), already introduced in Chapter 1 (see Section 1.3.4).

6.1 Visual-Game Analytics Background

In recent years, competitive online video games have become one of the main sources of entertainment for both active and passive users, i.e., players and spectators. In particular, the MOBA genre is among the most popular, having registered an increasing participation in both player base and event viewership (Riot Games, 2014; Walker, 2016), such as in games like *League of Legends* (*LoL*) (Li et al., 2017).

The increasing complexity associated with video games has also been accompanied by various technical improvements, which has made the collection of information from a given session a much more simple, common, and reliable task (Wallner and Kriglstein, 2013). Nowadays, large volumes of data are recorded daily, including spatio-temporal data related to the game, such as the players' trajectories over the game environment, the spatio-temporal events in which they participated, and the actions they performed throughout the game.

Similarly to more common applications, such as traffic analysis, in order to properly extract relevant knowledge from these, often voluminous, datasets it is fundamental to explore adequate interaction and visualization mechanisms.

Fortunately, the information captured by players' movements in a virtual environment is not entirely dissimilar to the information obtained with the movement of real people (Zheng et al., 2010), cars or even boats (Riveiro et al., 2008), and has already been the subject of several studies, as discussed in Chapter 2.

Although, moving in the physical world is considerably different from doing the same in a virtual world, location-based information and, by consequence, movement play a very important role in both contexts. Most modern games, ranging from simple 2D platform games to more complex 3D open world games, are dependent of movement, which constitutes one of the fundamental aspects from which gameplay derives and, thus, one of the basic elements in the playing experience (Drachen and Schubert, 2013b). Both contexts use necessarily a set of coordinates to classify space, regardless if they are geographic latitude, longitude, and altitude values or the $x/y/z$ coordinates of the screen. It is also important to notice that the gaming sessions do not occur in an isolated point in time and that the order of the players' actions or the occurrences of specific game events can also affect and help analysing the playing experience. As such, the temporal aspects of the data can be also significantly relevant for the analysis of the players' activities (Miller and Crowcroft, 2009; Drachen and Schubert, 2013a), as these can be compared to those made in the physical world (e.g., moving from one point of interest to the other, or revisiting a specific location to acquire something).

The challenges related with the analysis of video-game data can be associated to those faced in sports analytics and, consequently, in the analysis of geographic spatio-temporal data. This fact is particularly true when considering MOBA games, as these can also be considered as team sports. Consequently, the spatio-temporal information associated to these contexts is considered extremely important, since the comprehension and analysis of this data can be significantly advantageous for various groups of users, including players, spectators, developers, and analysts (Pingali et al., 2001; Hoobler et al., 2004).

Although the most significant body of work presented in the literature is dedicated to game developers, there has been an increasing interest in showing spatio-temporal data to players (Wallner and Kriglstein, 2013), the user group focused in this study. These users, despite knowledgeable in regards to a specific game, in some cases even experts, are often inexperienced in terms of trajectory data visualization and analysis. Nevertheless, they may still be interested in analysing their personal data, even if in a less complex and casual scenario (Huang et al., 2015; Pousman et al., 2007), to improve their gameplay performance and, consequently, their experience. This growing trend has already been addressed in several projects, including in professional contexts. Researchers, such as Drachen and Schubert (2013b), have already identified the lack of software tools to support the visualization of spatio-temporal data by inexperienced users as a relevant open issue in the area of visual-game analytics. In addition, some game developers and independent groups have also showed interest in providing visualization tools and APIs that grant, nearly any user, access to player generated data (Wallner and Kriglstein, 2013). Notable examples include the *Sc2gears* (Belicza, 2017), OPGG (2017), LoLKing (2017), and the *Rockstar Games Social Club* (Rockstar Games, 2016), which provide general game statistics referring to the players' performance or a gaming session. Although some of these applications provide complex map or graph-based visualizations depicting the evolution of a gaming session, the majority is focused on the display of aggregated statistics, which makes the detection of valuable spatio-temporal details difficult, if not impossible (Li et al., 2017) .

Ultimately, the objectives of this Ph.D. research go in agreement with some of the most relevant challenges associated to this area, namely the current interest in showing spatio-temporal information to a wider (inexperienced) population, associated with the uncertainty regarding the usefulness of existing map visualization techniques.

To address these objectives, the game *League of Legends* was used as a case study. This choice was motivated by three main reasons. First, out of all explored official data sources, the one provided by *Riot Games* (developers of *LoL*) was one of the few that included detailed information regarding to the players' spatio-temporal positions alongside their most relevant game activities. Second, MOBAs in general, and *LoL* in particular, are games in which the positioning of players throughout the matches is extremely important. Finally, among the various MOBAs, *LoL* is ranked among the oldest and most popular in the genre, thus, providing a representative case study for this work. The following section describes in detail this game and the data extracted.

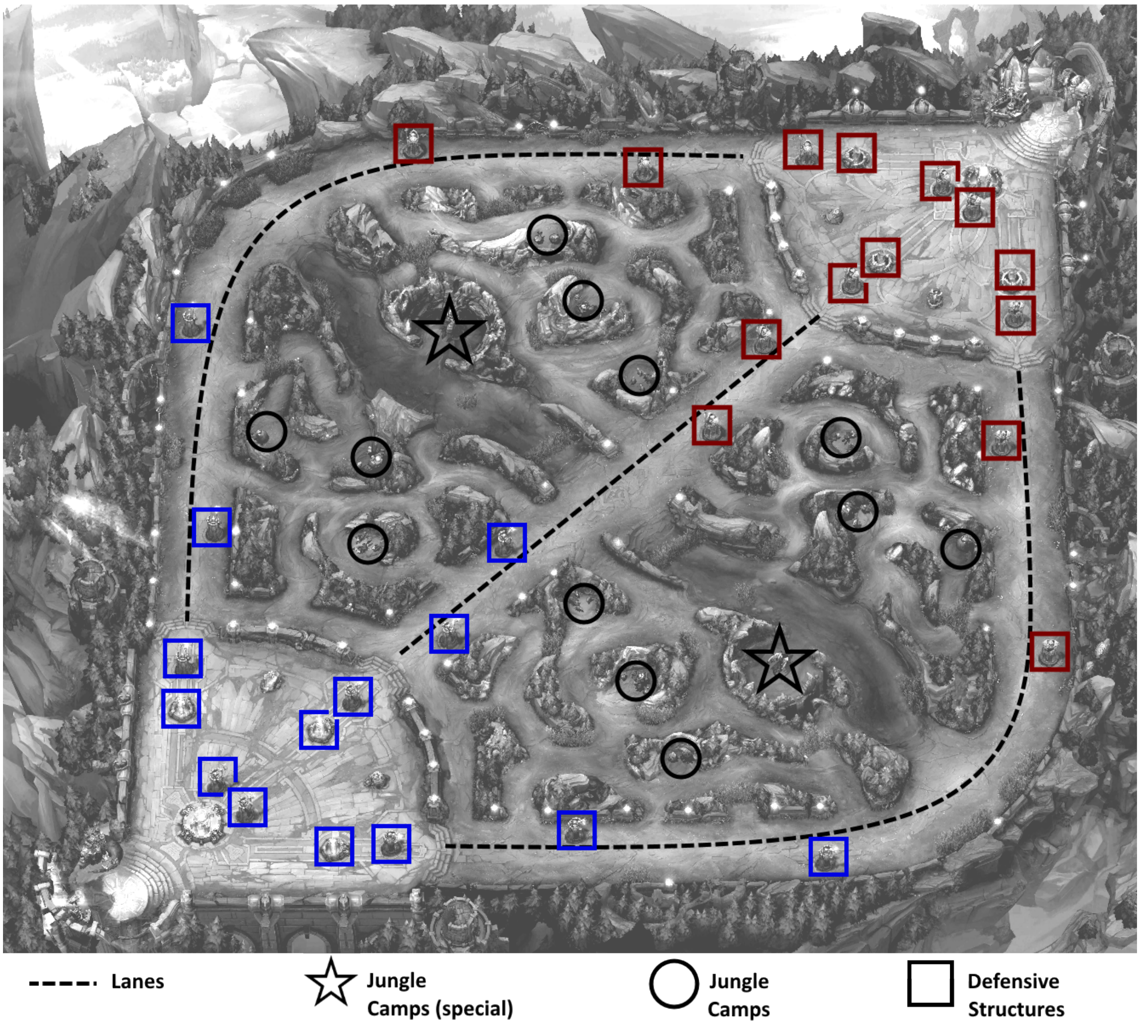

Figure 6.1: League of Legends map (source: https://i.imgur.com/HRgupVM.jpg).

6.2 Overview on League of Legends

League of Legends is a game that features the most common characteristics associated to the MOBA genre, involving two opposing teams, traditionally of 5 players each. At the start of the game, each team begins at one side of the map, and the objective of the game is to destroy the other team's base, located at the opposite side of the map (see Figure 6.1). The game world is composed by three main pathways connecting both team bases, usually known as *lanes*, or *top*, *mid* and *bot lanes*, depending if they are at the top, middle, and bottom of the map, respectively (Figure 6.1). Automatic defensive structures, known as *turrets* or *towers*, are placed along the *lanes* and need to be destroyed by the players, before they can reach the enemy's base (Figure 6.1). Non-player character (NPC) units, also known as *minions*, are created periodically from the beginning of each team's base and move along each *lane* towards the enemy's base, helping the players in the destruction of the enemy's structures. Between those *lanes*, the players can also conquer various points of interest, *jungle camps* (Figure 6.1). Out of the various camps available in the *jungle*, 2 of them, the *baron* and the *dragon*, are labelled as *special* (Figure 6.1) since these provide more valuable rewards than the remaining camps.

To facilitate these goals, each player will have to control one game character and collect various bonuses, in the form of *experience* and *gold* points, to improve their character's abilities and equipment, thus maximizing its capabilities in helping their team winning the game. During the match, the players will be rewarded according to various actions performed, which include, among others: killing or help killing a character controlled by the enemy team or various NPCs; or destroying the mentioned structures. Furthermore, in addition to various in-game items that the players can acquire at their bases, given a set amount of *experience* and *gold* points, they can also acquire temporary improvements (*buffs*) to their characters or their team, by destroying certain creatures/NPCs located at the *jungle camps*.

Each player is given a specific role in the team, such as *Attack Damage Carrier* (ADC) or *Support*, which will often determine where the player needs to be or move to. Therefore it is expected that some players, at least in the early stages of the game, roam in a specific lane or move between the lanes (*rotate*) to help their teammates. In order to access this information, the game allows any player to place certain items in the map, *wards*, which will temporary monitor a limited area and warn the players, and their teammates, about any enemy passing through that location.

Although a match does not have a time limit, players and analysts tend to identify three specific periods within a match: *early*, *mid*, and *end-game*, which go, approximately, until the 10, 20 minutes, and the end of a match, respectively.

Finally, to provide a more balanced experience to its players, *League of Legends* ranks each player with a 7-scale categorization, according to their performance and win-rate. These rankings include: *bronze*, *silver*, *gold*, *platinum*, *diamond*, *master* and *challenger*.

Based on this general description, it is possible to affirm that spatial and temporal information have a critical role in analysing and understanding this game. Similarly to a *real-world* scenario, where monuments or shopping centres would be considered points of interest, in this context, these would be the defensive structures or the various *jungle* locations that grant *buffs* to the players. Similarly, whereas in *real-world* scenarios important events to consider would be, for example, when/where a taxi driver dropped a passenger, or when/where a football player scored, in this context, important events depict character deaths or the destruction of structures. Consequently, these facts can affect the player's mobility throughout the match, forcing them to decide between safer, but most likely longer, routes (i.e., the *lanes*) or shorter routes, possibly much more dangerous and with lesser visibility (i.e., the *jungle*). Based on this information, several types of analysis questions can be made that, when answered, can help the player improve his/her performance for future matches.

6.3 VisLoL Prototype

In the context of this study, the VisLoL prototype was created to support the visualization of spatio-temporal game data for players likely familiar with the domain knowledge (the game), yet inexperienced in terms of data analysis, specifically, video-game analytics. As previously mentioned, data from the game *League of Legends* was used for this study, obtained through the *Riot Games* API (Riot Games, 2016c).

Comparatively to the previous experiments, this study is focused on a more specific, and somewhat unexplored context of use, the visualization of personal game-related data. As such, it was necessary to further interact with the main target users of this prototype. Considering the expected experience from the users regarding the game, it was also expected a possible biased view over the data, particularly when regarding the player. Therefore, prior to the development of the prototype, a questionnaire was released into *League of Legends* forums and *Reddit* (Riot Games, 2016a,b), to gather information regarding the players' perception of the importance of their spatio-temporal information in their games alongside the types of metrics and functionalities they consider the most necessary to understand what happened in a match. The following section provides an overview regarding this questionnaire's structure and results.

6.3.1 Questionnaire

The questionnaire was composed by 3 main sections. The first consisted of a set of profiling questions, e.g., types of games similar to *LoL* that the user plays, in-game ranking, or how often the user plays *LoL*. The second section was composed by a set of affirmations regarding the importance of various spatio-temporal factors during the match (e.g., positioning of players and events). The user would need to rank them, in a 5 point scale, according to their level of agreement. The third section inquired the users about their interest in studying the game (e.g., watching other players, inspecting their status in dedicated web-pages) and the types of information they extract from this task.

A total of 270 players volunteered to answer the questionnaire, distributed by 8 specific game-specific ranking categories (34 Unranked, 33 Bronze, 62 Silver, 56 Gold, 48 Platinum, 28 Diamond, 5 Challenger, and 4 Master players). Out of all participants, 85% played *LoL* daily.

In regards to the importance given by the players to spatio-temporal events, 88%, 90%, and 81% of the players considered, respectively, that their own positioning, the one of their team, and of the enemy team were *Important* or *Very Important* (the two highest ranks in the questions' classification). Similarly, 27% and 62% of the players also considered *Important* and *Very Important* the position of deployable items into the game environment that notify the player's team regarding the enemy's movement in the map. In addition, 65% and 88% of the participants answered that they regularly visit web pages/applications that report their performance in LoL and watch other players' matches, respectively, to improve their performance at the game. Among the most mentioned metrics of interest when watching or inspecting these past games, players have mentioned

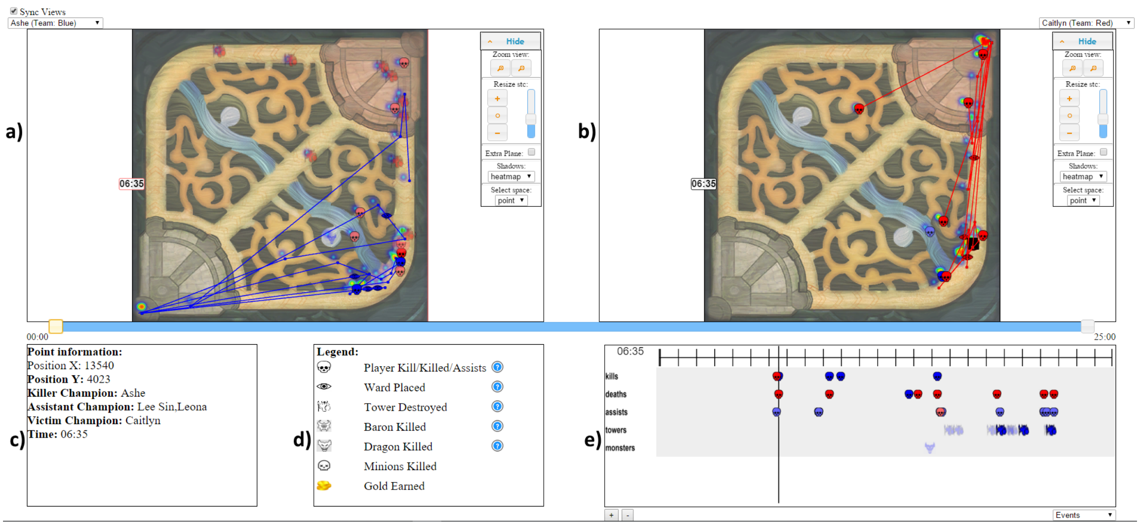

Figure 6.2: VisLoL prototype. a) Player 1 event map; b) Player 2 event map; c) Information panel; d) Legend panel; and e) Timeline.

kill/death/assists (79%), team strategies (66%) and positioning (62%).

Based on this questionnaire, in addition to the confirmation of the metrics seen as relevant by the players, it is understandable that this community of users is interested in analysing their own spatio-temporal information, which goes in agreement with the main motivations of this Ph.D., in general, and this study, in particular.

6.3.2 Data and Prototype Features

Based on the results of the questionnaire, alongside the analysis of previous experiments of this work, it was developed the VisLoL prototype (see Figure 6.2). Similarly to the prototypes described in the previous chapters, VisLoL was developed in HTML5, as a local web application, taking advantage of code already developed for the TTV prototypes described in the Chapter 5, while also using the following libraries and frameworks:

Node.js, NW.js - as described in the previous chapters, these support the creation of (local) web-applications;

STCJS - an updated version of the JavaScript API described in the previous chapters. In addition to various performance upgrades, the STCs created by this API allows the user to convert this visualization into a 2D map (further explained below);

TIMELINEJS - an updated version of the JavaScript API described in the previous chapter, with various corrections and performance upgrades.

All data used by the VisLoL prototype was made available thanks to the aforementioned Riot Games API and a custom Python application to request and store the data locally. This data is composed by five lists of spatio-temporal information associated to each match, which include:

Player position and attributes: recorded at every minute since the beginning of the match, each entry is composed by the spatio-temporal locations of the players, alongside other player attributes like, *gold*, *experience* points, and NPCs killed;

Player kills: each entry is composed by the spatio-temporal locations of the event and the identifications of the players involved, namely, who died, who the kill was attributed to, and who assisted/helped;

Special jungle monsters killed: each entry is composed by the spatio-temporal location of the event, alongside the identification of the player responsible for its destruction;

Tower destruction: similarly to the events above, it is composed by the spatio-temporal locations of the towers destroyed and the identification of the players involved;

Ward placement: composed by the time when a *ward* was placed, and the identification of the player that did it. Based on this information, an approximated location where the *ward* was placed is estimated, through the linear interpolation of the player's movement and event data.

All of this information is depicted in VisLoL through the use of three main sections; the map component (Figures 6.2 a and b), the timeline component (Figure 6.2 e), and a textual information and legend components (Figure 6.2 c and d). These components are described with more detail in the following sections.

Map Component

The map component is composed by one or two map visualizations, depending on the number of players under analysis (see Figure 6.2 a and b). Using two dropdown buttons, located at the top of the views, the user is able to select the player(s) that wants to visualize. If the user selects *None*, in the second map, the corresponding map view will not be displayed and this component will be filled by expanding the other map (see Figure 6.4 c).

Both maps display the trajectory and the events associated to each player, using simple, yet recognizable visual representations, since despite they unfamiliarity with the application, these users should be considerably knowledge regarding the game, and its symbology. In particular, the locations visited by the player throughout the match are represented as simple point symbols, coloured according to the player's team (i.e., red or blue). The events in which the player participated are represented with a similar symbology as the ones used in the game and already existing tools, e.g., skulls to represent deaths, towers to represent the destruction of towers. All location points and events directly related to the player's location are connected with lines, coloured accordingly. An important distinction is made between the events where the player provided assistance and those that he or she directly participated. Although both types of events are likely to be relevant to the player, the visualization emphasizes direct participations over assistances, representing these with full visibility, i.e., with no transparency applied. Since in assistance events the player was not the

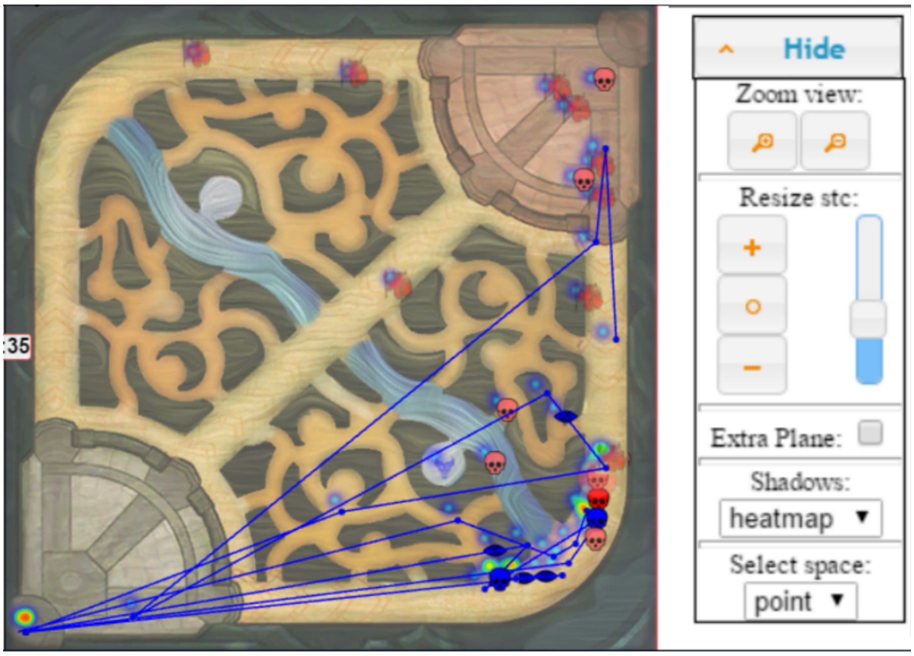

Figure 6.3: Player 1 event map and controls (highlight from Figure 6.2 a).

ultimate responsible, e.g., the player did not deliver the killing blow, so the opponent could still have escaped, their representations are displayed with a reduced visibility, i.e., symbols with a higher transparency.

Although Figure 6.2 depicts two 2D static maps, the user can also work with 3D space-time cubes. Comparatively to the previous experiments, in which both types of maps were displayed side-by-side, VisLoL applies a different approach. Considering that, throughout the previous studies, 2D maps were shown as being the preferred by the participants, by default, the map component displays a 2D map (see Figure 6.3). However, by rotating the view vertically, the user can transform this map into a 3D space-time cube (see Figure 6.5). It is important to mention that this default approach is different from simply looking at the data from a *birds-eye view* (BEV). Contrarily to the STCs implemented in the previous studies, when the STC is rotated into a BEV, the application will switch the map's projection from *perspective* to *orthogonal*, implying that the symbology used to represent information will not be affected by the third-dimension. It can be argued that the same result could be achieved by using the temporal granularity controls, such as it was used before. However, based on the results of the previous studies, it is important to notice three factors: first, this strategy may not be always used; second, analysing the data from a *birds-eye view* can hinder user performance; and third, this perspective takes barely any advantage of the third dimension to visualize time.

Timeline Component

Similarly to the prototypes presented in the previous experiment, VisLoL also provides a timeline component, which consists of an interactive point/line chart focused on displaying data from a

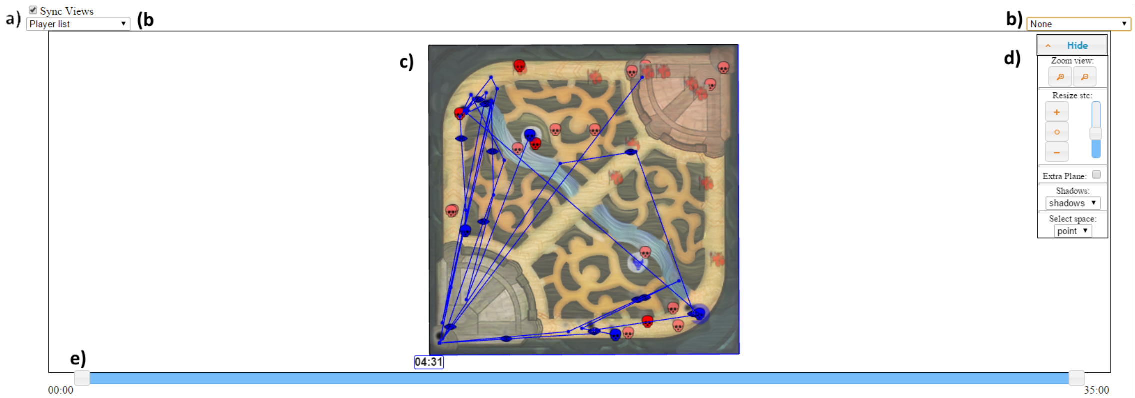

Figure 6.4: VisLoL map component (in detail). a) *Sync Views* option; b) Player selection menus; c) Map view as a 2D static map; d) Map control panel; e) Temporal slider.

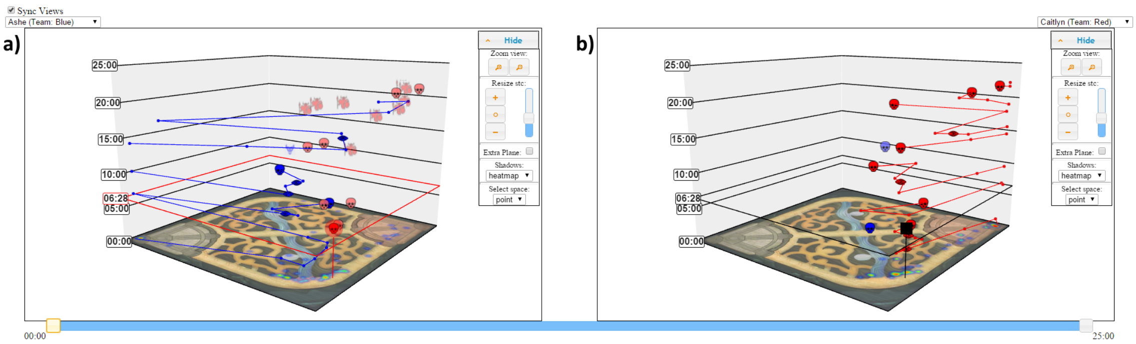

Figure 6.5: VisLoL - STC map visualization a) Player 1; and b) Player 2 event maps.

temporal point of view (see Figure 6.2 e). The implementation of this component was considered necessary due to the fact that it provides an additional point of view (often neglected) over the data. However, this was also motivated by the considerably positive feedback received from the previous study, particularly when combined with the 2D static map.

Considering that the players' position is retrieved at each minute and the player is always expected to be in movement (unless his/her character is killed), this component focuses mostly on the visualization of the players' events, their *gold* and *experience* points. This is different from the timeline representations from the previous experiments, which were focused on displaying when movement was detected, or the duration in which a trip took place. By default, this view is com-

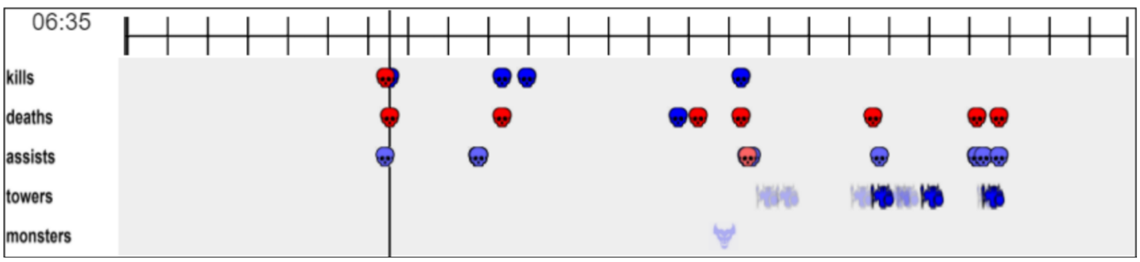

Figure 6.6: VisLoL event timeline (highlight from Figure 6.2 e).

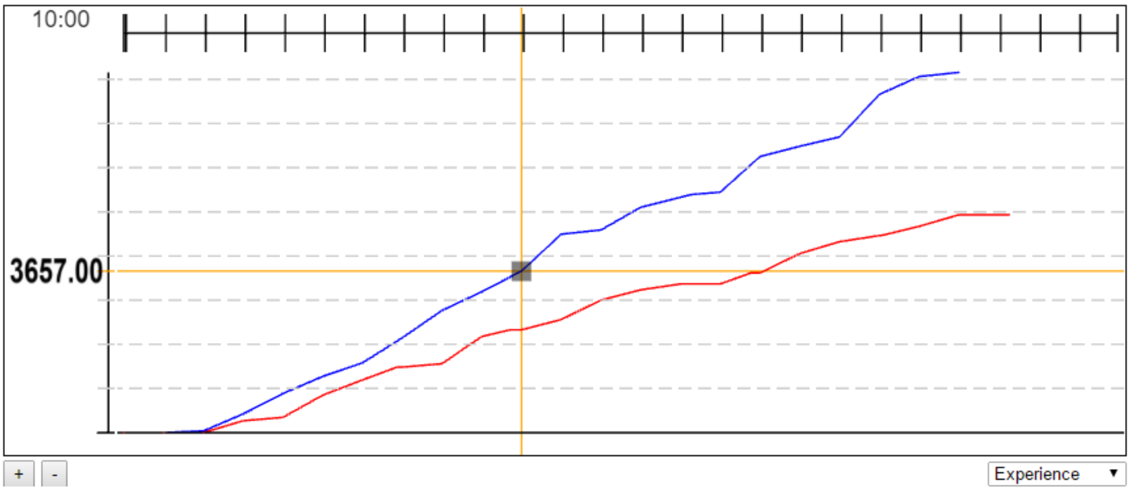

Figure 6.7: Timeline component as a value line graph, showing the evolution of the players' experience over time.

posed by several rows, each one associated to a specific type of event, such as player and enemy deaths, or towers destroyed. On each row, various icons are distributed horizontally to indicate the specific moments in which they were detected during the match (see Figure 6.6). In order to simplify the transition of the user's attention between the map and the timeline components, the symbols used in the timeline are similar to those used in the map component. Furthermore, using a dropdown menu, located below the timeline, it is possible to select a different quantitative attribute, namely: *gold* earned, *minions* killed, *experience*, and total *wards* placed during the match. This will change the representation of the timeline into a common $x - y$ value line chart (see Figure 6.7).

Information Components

The third component consists of panels with textual information and a legend (see Figure 6.2 c and d and Figure 6.8 a and b). Whenever a user selects an event from the timeline or the map component, regardless if it is a 2D map or an STC, all data associated with that event is displayed on the first container. Specifically, this information includes: the coordinates and timestamp of the event, the player responsible for it, the player's team, and the type of event. If the event involves multiple players, their character names are also displayed. On the other hand, the legend component explains the meaning of each icon on the map and timeline components.

Interactive Controls

The interaction with VisLoL is fundamentally done through mouse controls, and considering its use of APIs and frameworks similar to those used in the previous experiments, several interactive controls are the same.

Both map and timeline components provide common pan and zoom operations. As described before, the map component also allows the user to rotate the view horizontally and vertically,

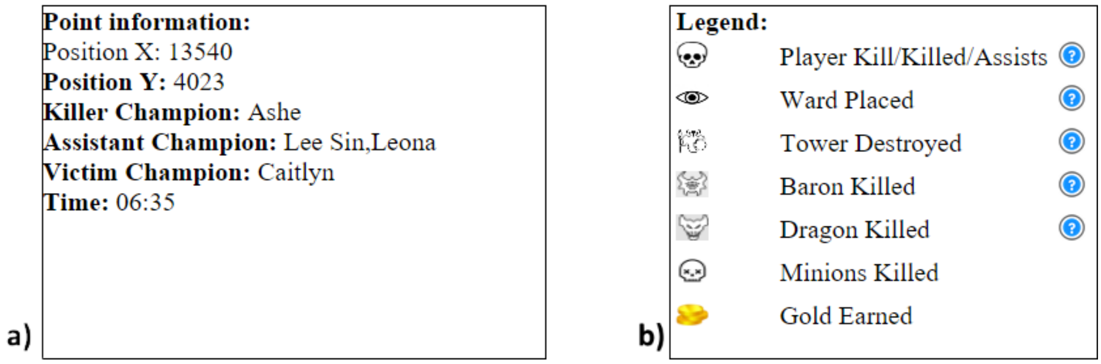

Figure 6.8: VisLol components a) Information and b) Legend panels (highlight from Figure 6.2 c and d).

thus controlling the point of view within the STC, and even switch between a 2D map and a 3D STC. When using the STC, it is also possible to control its temporal granularity, by changing its height, with the control panel displayed at the right-side corner of the map (Figure 6.3). To minimize the STC's limitations in location-based tasks, this technique provides an additional map plane, controllable by the user, that can be dragged and locked anywhere along the STC's height. However, this feature was, once again, made optional to the user, given the results of the previous studies have reported some confusion when using this type of approach.

Comparatively to the previous experiments, VisLoL can use two maps representing different, but complementary, information about the same overarching event (the game). As such, following similar concerns from the previous studies, it was necessary to emphasize the direct relation between the different components. Therefore, when the map component consists of two map views, any change made to one map is immediately reflected into the other (e.g., rotating the map or rescaling the STC's height). This functionality can, however, be de/activated, by toggling the option *Sync Views*, located on top of the map component (see Figure 6.4 a). In a similar way, hovering an event point in the map view or in the timeline, will highlight the correspondent data element in the other views, e.g., selecting a kill event in the map will highlight the matching time moment in the timeline and on the other map. Moreover, VisLoL also supports the heatmap and spatial point/area highlight selection features, described in the previous chapter. Finally, below the map component, a temporal slider (Figure 6.4 e) allows the user to filter information based on a specific period of time, on both maps and the timeline components.

6.3.3 Application Scenario

This section describes the analysis of the events from a match conducted with VisLoL. Figures 6.2, 6.5, and 6.6 depict the same game played by two members of opposing teams in the *ADC* (Attack Damage Carry) role. By looking at the 2D maps (Figure 6.2) it is possible to understand that, although both players spent the majority of the time in the *bot lane*, the player from the blue team (left map) covered a higher geographical area, and participated in more events throughout the game. Then, analysing the data with the timeline component (Figures 6.2 e and 6.6) allows to complement the previous information and verify that, no significant events were recorded, in the first 6 minutes of the game (approximately). However, after an approximated period of 10 minutes, in which various confrontations resulted in several deaths, the blue team proceeded to destroy all enemy structures and win the game.

Converting the map into an STC (Figure 6.5) allows to see, for example, the advantage gained by the player of the blue team, due to being able to participate in the deaths of two opponents in the enemy's side of the map, one of them the player's direct opponent in *lane*. Approximately 4 minutes later, despite being killed, the player contributed again in the deaths of two enemy players, around the same location and, once again, one of them was the player's direct opponent. Later, particularly from minute 16 onwards, the user is able to see the player from the red team covering a smaller area with its movements, whereas the player from the blue team is covering more. This fact also coincides with the destruction of the red teams' defensive structures mentioned earlier.

6.4 User Study

A user study was conducted to address the objectives of this chapter, which include: understanding if inexperienced users, in terms of visualization and analysis, are capable of reaching meaningful conclusions with VisLoL, and if so how, comparatively to already existing tools. The study compared the VisLoL prototype with the official match-history web-page used by *Riot Games* to present information to *LoL* players regarding the various matches they participated. An example of this application can be found in Figure 6.9.

Comparatively to VisLoL, the match-history web-page is not focused on the representation of the spatio-temporal information associated with the players but rather their thematic information, e.g., items bought, or the total amount of *gold* collected or *minions* killed. In addition, while VisLoL focuses on the visualization of just one or two players, the web-page supports the visualization of individual and aggregated information regarding the two teams. Unlike VisLoL, however, the web-page does not depict any information regarding the player's movement, their approximated *ward* placement, nor associates this information with the other game-related events. Despite that, both applications display similar types of information and use similar types of visualizations, namely a map displaying player deaths, and a timeline displaying player generated events or *gold* collected over time.

It is important to notice that the objectives of this study do not encompass the comparison of VisLoL with the match-history page to analyse which one is better. Instead, the objective consists in comparing both visualizations to learn how differently people analyse the data, to better understand the importance given to spatio-temporal data by players, and learn the role these techniques may have in the data analysis. In addition, contrarily to the studies presented in the previous chapters, it was expected that the participants were familiarized with the game and, thus, had a significant domain knowledge. This is likely to cause some bias effects over the results, particularly if the participants analyse their own game data. To minimize the impact of this factor, an additional objective of this experiment consists in studying the possible effects over the analysis from a user's knowledge of the data that is being visualized.

A total of 21 participants volunteered to the study. All participants played *League of Legends* frequently and were, at least, familiarized with existing applications to review and analyse their performance after matches, such as the official web-page, or other *independent* web-pages such as op.gg (OPGG, 2017). Out of all participants, their ranking distribution was as follows: 4 were ranked in *bronze*, 7 were ranked in *silver*, 3 in *gold*, and 7 in *platinum*. No *diamond* nor *master & challenger* players, i.e., ranks above *platinum*, were available to participate in this study.

6.4.1 Tasks

A common motivation towards the use of applications like the official match-history web-page and, ultimately, VisLoL consists in the willingness for the players to improve their performance. This implies the understanding of the relation between their actions and their outcome. In addition, due to the competitive nature of the game, comparisons between players or the same player through different moments are fundamental. For these reasons, in the context of this experiment, the participants were asked to perform two different tasks, namely *associate*, and *compare*. These are correspondent to the second level of the hierarchy of tasks, described in Section 2.3, with the omission of the *identify* and *locate* tasks.

In the *associate* task, the participants were asked to evaluate and classify, between 0 and 10, the performance of a given player in the match, based on any set of criteria they considered relevant. The participants were also asked to justify that classification based on their selected criteria. This task aims to simulate one of the most common tasks performed by players, for example after a match, in which they are able to visualize a set of statistics describing the various results.

In the *comparison* task, the participants were asked to evaluate the performance of two players, from opposing teams, on a similar role. Similarly to the *association* task, the participants were free to select any criteria they considered appropriate and, at the end of the task, they were asked to classify both players, between 0 and 10, and justify those scores. Considering the competitive nature of MOBAs, this task simulates a common practice both by casual players and even in professional environments, in which players on the same role are compared.

It is important to acknowledge that, unlike the previous experiments, presented in Chapter 3, that used this level of tasks, the ones proposed in this study do not include the tasks *locate* and *identify*. The primary reason for this is related to the significantly different design of VisLoL, comparatively to the web page, alongside the different sets of information displayed or focused by the applications. For instance, while VisLoL shows the positions of players throughout the game, the web-page shows the players' equipment. Even though it would be possible, it is important to notice that the comparison of these applications in terms of how specific types of information are located or identified is not relevant towards the objectives of this study. In addition, it can be argued that, in order to *compare* or *associate* information, the user will still need to *locate* or *identify* it first (e.g., comparing two different movers, implies that those movers were first located

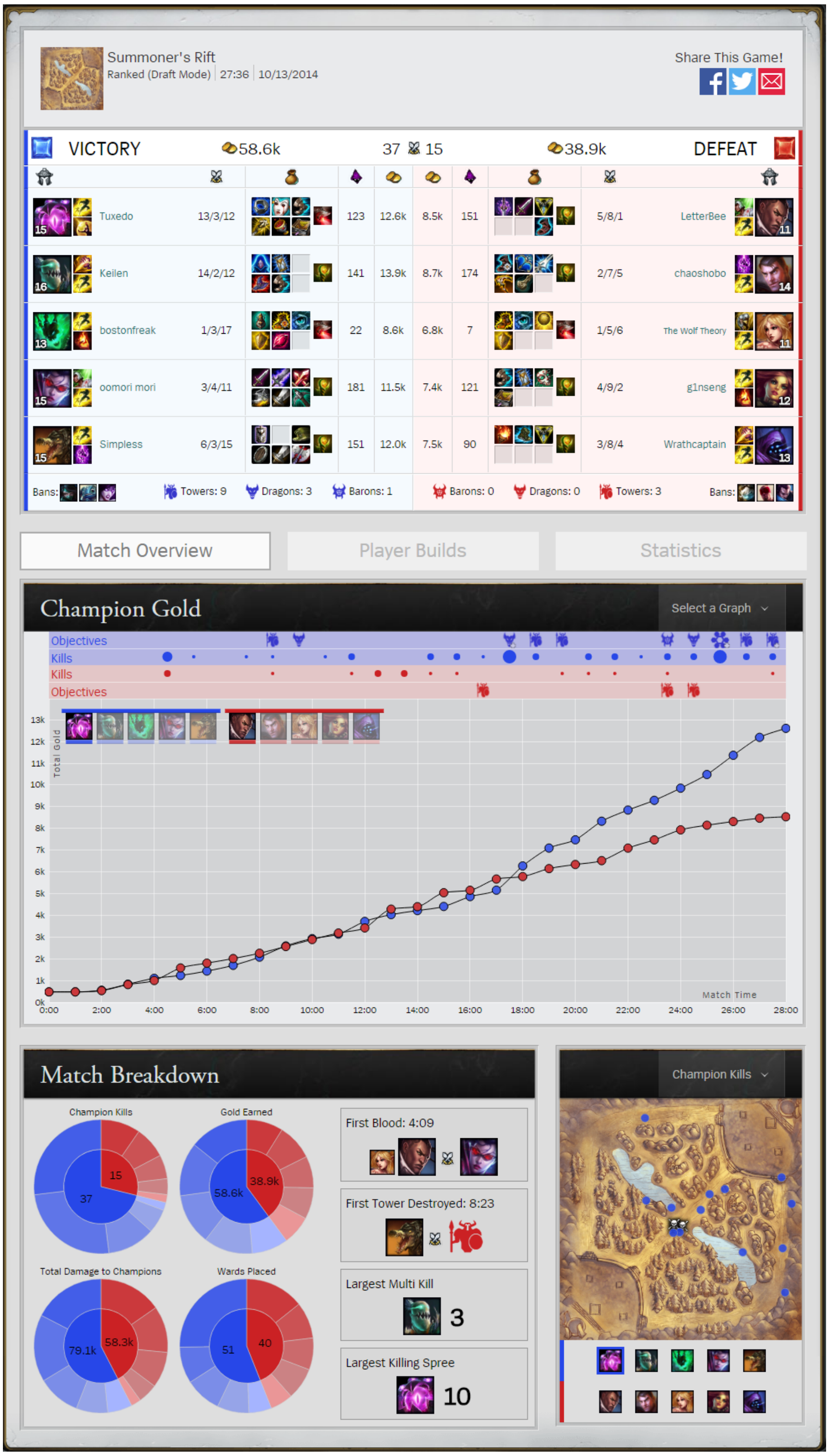

Figure 6.9: Example of the official match-history web-page from Riot Games' League of Legends (source: http://tinyurl.com/h9568gu).

and that their properties were identified). Consequently, although not necessarily relevant as the main tasks of this study, these are still eligible for the training phase of this study, described in the next section.

In addition to the aforementioned tasks, two independent variables were also considered, namely:

Application (*App*) - this variable is composed by two levels, corresponding to the prototype VisLoL and the official match-history web-page provided by *Riot Games* (web-page);

Target Analysis (*TA*) - this variable is also composed by two levels. In the first, the player is told that the displayed data is from one of their matches. In the second, the player is told that the information displayed is not theirs (even though it is).

6.4.2 Procedure

The experiment followed a *within-subjects* design and all participants carried out each task individually. At the beginning of the experiment, the participants were briefed about the nature of the study and the applications being used. Then, they performed a tutorial with VisLoL that covered all of its functionalities and, in particular, included a set of *identification* and *location* training tasks, requiring the participants to detect specific sets of information, based on spatial or spatio-temporal constraints. These questions include:

Was the player part of the destruction of any game objective?

Where did the player die for the last time during the match?

Indicate a location where the player killed an enemy player; and

Where and when did the player used his first ward?

The participants were then encouraged to further interact with the prototype, so they would get used to the interactive controls and be able to clarify any doubts. This stage took, approximately, between 15 to 20 minutes.

After the training phase, the participants were asked to perform the two tasks, first the *association* and, then, the *comparison* task. This was decided due to the increasing complexity of one task to the other and neglecting this factor could produce a negative effect over the results, due to an over-simplification of the *association* task.

During the training tasks, VisLoL displayed anonymized data from one of the most recent matches from one professional *LoL* player. For the *associate* and *compare* tasks, anonymized data was also displayed; however, recorded from games of the study participant. It is important to mention that the data visualized in the VisLoL prototype and in the official web-page was from the same match; however, the player was never informed about that. To avoid possible sequence effects, the order in which the independent variables were presented in both tasks was counterbalanced based on a *latin-square* design. As a result each participant performed a total of 8 trials: 2 *App* $*$ 2 *TA* $*$ 2 tasks.

The following dependent variables were considered to measure the study's results:

Task completion time: recorded from the moment the user was instructed to begin the task until s/he gave a final answer. This variable is used as a common metric to compare the performance between applications;

Ratio of component interaction time: recorded when using VisLoL, it is the ratio of time spent interacting with the map and the timeline components, comparatively to the total amount of time spent completing the task. This variable is an additional measure of task complexity, allowing the comparison of the users' interactive needs with the application and with its core components;

Ratio of map mode interaction: recorded when using VisLoL, it is the ratio of time spent interacting with the map component, as a 2D map or a 3D STC view, comparatively to the total amount of time spent completing the task. This variable is used to understand which type of map-visualization technique the users prefer on each task and, consequently, to better understand how users interact with VisLoL;

Coherence of ratings: consists in the comparison of the ratios of participants that increased, decreased, or kept the players' ratings after analysing the data with both applications. This variable is used to help comparing the conclusions reached by the users' analyses with the different applications when these are or not informed of whose data they are analysing;

Coherence of decisions: used in the *comparison* task, it is the ratio of participants that gave similar ratings when using both applications, e.g., ranking one player better than the other using VisLoL and doing the same with the web-page, even if with different values. Similarly to the previous variable, this metric allows the study of the users' analyses; however, while the above focuses on the individual values provided by the users, this variable focuses on the overall conclusion reached with those values, e.g., player A performed better than player B;

Similarly to the prototypes described in the previous chapters, VisLoL generated log files recording the task completion times, both interaction ratios, and representing the various user's interactive actions during the tasks. These actions were the same as those recorded from the last study, which include any specific operation over the views, such as resizing the STC, panning/dragging, zooming (in/out), or rotating the views, and selecting (either by clicking or hovering) any feature with information within any of the views.

In addition to these variables, the various criteria used by the participants to rate and compare players in the *associate* and *compare* tasks were also listed.

At the end of the tasks, the participants were invited to answer to a small questionnaire covering their opinions regarding the VisLoL prototype, its features, and when the participants considered them more useful or difficult to work with.

Table 6.1: Mean participants' task completion times (in seconds), in the *association* and *comparison* tasks with both applications and target analysis (TA) configurations.

	Association		**Comparison**	
	Unknown	**Known**	**Unknown**	**Known**
VisLoL	274,60	284,95	352,90	303,74
Web-page	212,00	225,67	211,57	197,29

6.4.3 Results

The following sections describe the main results obtained in the study. For a matter of simplification and readability, these focus on the most statistically significant results obtained, those close to be considered as such, or that may reveal important patterns.

Task Completion Times

To compare the differences in terms of task completion times in the different conditions of the *associate* task, the results were subjected to the Shapiro-Wilk test of normality. Since no valid data transformation could provide a normalized dataset, a non-parametric approach was used by applying the Aligned Rank Transform (ART) method, followed by a full factorial ANOVA (Wobbrock et al., 2011). Although the mean participants' completion times suggest that participants took longer to complete the task using VisLoL, no statistically significant results were detected (see Table 6.1).

A similar procedure was used to analyse the results in the *comparison* task. The tests revealed a significant effect of *App* over the results ($F(1, 80) = 11, 4; p = 0, 001$), as participants took longer to complete this task when using VisLoL, comparatively to when using the official web-page (see Table 6.1).

No significant results or detectable patterns were found in regards to player ranking and target analysis configuration.

Ratio of Component Interaction Time

The results in terms of the ratio of component interaction time were also subjected to the Shapiro-Wilk test of normality. Since no data transformation could provide a normalized dataset, non-parametric approaches were used to analyse the results.

For the context of this variable, the results obtained from the participants' training period were also included, and labelled as *identification* tasks. For those, a Wilcoxon Signed Rank Test shows significant differences, where there was a higher usage of the map component, comparatively to the timeline component ($Z = -3, 77; p < 0, 001$). In the *association* and *comparison* tasks, the application of the ART method, followed by a full factorial ANOVA did not reveal any significant results (see Table 6.2).

Table 6.2: Mean participants' ratio of interaction time (%) when using the map and the timeline component, for all task configurations.

	Identification *(training)*	**Association** Unknown	Known	**Comparison** Unknown	Known
Map	62,44	44,91	37,96	36,64	39,92
Timeline	16,12	36,66	37,96	38,15	30,33

Table 6.3: Mean participants' ratio of map mode interaction time (%) when using the map as a 2D or as a 3D STC, for all task configurations.

	Identification *(training)*	**Association** Unknown	Known	**Comparison** Unknown	Known
2D	48,16	64,95	72,29	71,92	67,29
STC	51,84	35,05	27,71	28,08	32,71

Ratio of Map Interaction Time

A similar procedure to the one previously described was applied over the results in terms of ratio of map interaction time. While no significant differences were detected in the *identification* tasks, the tests revealed a significant effect from *App* in the remaining tasks (*associate*: $F(1,19) = 8,591$; $p = 0,009$; *compare*: $F(1,20) = 9,833$; $p = 0,005$). In both tasks, the results allowed to confirm that participants spent more time using the map component as a 2D view, rather than an STC view (see Table 6.3).

Coherence of Ratings

To analyse the coherence of ratings, it was necessary to compare the classifications given to the players by the study participants, after (unknowingly) analysing the same match data with VisLoL and the official web-page. The results were categorized into three groups, depending if the participant gave a higher, lower, or kept the same rating with the web-page comparatively to VisLoL. Tables 6.4 and 6.5 show the distribution of results throughout the *association* and *comparison* tasks, respectively, taking into account the different target analysis (*TA*) configurations.

Although the results can be considered inconclusive in the *association* task, in the *comparison* task they suggest that, after analysing the data with both applications, a higher percentage of participants changed their classifications when they were not informed that one of the players under evaluation was him/herself. By contrast, when participants were informed that they were visualizing their data, the majority kept the same classifications.

Based on this observation, a Chi-Square test was used to assess the possible relation between these patterns and the participants' knowledge on whom they were evaluating (*TA*). Regardless of the target of the observation (i.e., the player or the opponent), both tests suggest an association between knowing whom the data belongs to and the coherence of the ratings (Rating the player:

Table 6.4: Distribution of participants (%) that reduced, kept, or increased the player's ranking, after visualizing the data with both applications, in the *association* task.

	Unknown	**Known**
Reduced	23,81	38,10
Kept	38,10	33,33
Increased	33,33	23,81

Table 6.5: Distribution of participants (%) that reduced, kept, or increased the two players' ranking, after visualizing the data with both applications, in the *comparison* task.

TA	**Unknown**		**Known**	
	Player	**Opponent**	**Player**	**Opponent**
Reduced	42,86	52,38	19,05	23,81
Kept	9,52	19,05	61,90	57,14
Increased	47,62	28,57	19,05	19,05

$X(2) = 12,56; p = 0,002$; Rating the opponent: $X(2) = 6,65; p = 0,036$).

Coherence of Decisions

To analyse the coherence of decisions, it was necessary to first compare the ratings given to both players. The results were, then, compared based on the type of application used (i.e., VisLoL or the official web-page). Table 6.6 illustrates the coherence of the decisions. The results show that, regardless of being told who was being analysed, 81% of the participants that considered one player's performance better than the other with VisLoL, had the same opinion when using the web-page.

6.5 Discussion

The quantitative results obtained, alongside the participants' subjective feedback highlight some relevant aspects regarding the VisLoL prototype, its components, and the way participants interacted with it.

In terms of task completion times, the results suggest that users may take more time to analyse information with VisLoL, comparatively to the official web-page from *Riot Games*. This difference, however, was only shown to be significant when comparing different players. These results were actually expected since, in addition to the novelty of the application, the VisLoL prototype encourages the user to interact more with the data (e.g., through map zooming, or view rotation operations), comparatively to common web-pages that focus on the display of general aggregated statistics. Nevertheless, despite the possibility of taking more time to compare the performance of different players with VisLoL, ultimately, the participants' main conclusions in terms of who

Table 6.6: Distribution of the participants' coherence (%) in judging one player better than the other in the *comparison* task.

	Unknown	**Known**
Different	19	19
Same	81	81

performed better were mostly the same, regardless of the application.

These results suggest that, at least, VisLoL may support the same conclusions as existing tools, thus making it, at worst, an acceptable alternative. However, the study of the participants' comments alongside their strategies and metrics used in the tasks provide further relevant information. With both applications, the participants tried to use similar metrics, including: *KDA* (i.e., number of opponent kills, own deaths, and assists/participation in opponent kills), objective taking (i.e., the destruction of points of interest that benefit the player's team), and player farming (i.e., the ability to acquire in-game resources faster than other players). However, although the metrics used and the conclusions reached by the participants were generally similar, the process in which these were used was considerably different. When using the web-page, the participants focused mostly on commenting the values of the various metrics at the end of the match, rarely taking into consideration the temporal evolution of the values of those metrics. Based on that information, some participants would also propose hypothesis to explain the game results. On the contrary, when using VisLoL, participants often looked at the evolution of the various metrics in time and, more importantly, at their spatial distribution.

Based on this, it is possible to conclude that when using the web-page, participants tried to hypothesize some details, from a data overview; however, when using VisLoL, most participants tried to tell a story, describing the state of the match during the *early*, *mid*, and *late-game* periods. This a very important difference, since it suggests that the first approach is dependent on the players' previous experience with the game and their own ability to hypothesize something, while the second approach is essentially based on observed facts, ultimately, providing more believable conclusions.

Considering the preponderance of the map and timeline components of the VisLoL prototype, it makes sense that the participants gave considerable importance to the spatio-temporal data that was presented, when analysing the information. However, it is important to notice that, although some of the information was also available in the web-page (e.g., a timeline describing important events, or a map with the spatial distribution of player deaths), the participants did not use this information as much, if at all.

At the end of the experiment, all participants have shown their interest towards VisLoL, having answered that they would like to have VisLoL's functionalities present in the web-pages they already use to study their performance. The participants also justified this preference based on them considering that the spatio-temporal information displayed was useful and that the visualization components allowed them to learn *what* happened during the game and, more importantly, understand *why* something happened (e.g., the participant understanding that a player died since s/he had no map visibility, through *wards*).

The analysis of the remaining results highlights how participants have used VisLoL for the completion of the various tasks. In terms of ratio of component interaction, the results have only shown a significantly higher usage of the map component, comparatively to the timeline, in the training tasks. This result was expected, due to the spatial constraints associated with these tasks. On the other hand, the *association* and *comparison* tasks had more open ended answers, which allowed the participants to look for and provide more information not necessarily restricted by, but still related with, spatial constraints. Therefore, the lack of empirical evidence showing a higher usage of a particular component suggests that the participants considered necessary to interact with both components to extract information.

In terms of ratio of time spent in map mode, the results showed that in the *association* and *comparison* tasks the participants used the map component mostly as a 2D map view, rather than a 3D STC. Once again, this result was not completely unexpected, since it goes in agreement with some of the conclusions reached in the previous chapters, namely, that users tend to prefer 2D static maps over 3D STCs. Interestingly, during the tasks, the direct observation of the participants' actions and the analysis of their interaction logs supported the identification of 3 high-level strategies used to extract information from VisLoL, which include:

2D map + Timeline : Some participants focused their interaction on these two components, only using the STC occasionally. A possible justification for this may be the fact that, by default, the map was shown in a 2D view. This, alongside the familiarity towards this type of representation, may have led participants to not consider necessary or remember to switch to a different view. Some participants have justified their reduced or lack of interaction with the STC component due to their familiarity with the 2D map, which resembles the mini-map present in the game;

2D map then STC: Some participants focused their attention on the map component and avoided using the timeline. In these scenarios, after obtaining an overview of the data in the 2D view, the participants would switch the view into a 3D STC in order to obtain more detailed information, taking into account specific events in time and space. Some participants justified this behaviour saying that, by visualizing the spatio-temporal distribution of their related events, they considered the STC view better to understand the evolution of the players' position in space and time, and the impact they had in the game;

Timeline to 2D map and STC: In alternative, a third group of participants focused their attention on the timeline to guide their analysis with the map visualization. These participants would either limit the amount of visible information based on a temporal window, or select specific game events, and then visualize that information with more detail with the map component as a 2D map or 3D STC view.

Finally, the results of this study also suggest a possible effect of the participants' knowledge of whom the data under analysis belongs to. During the tasks, when informed that the data displayed was from one of their games, some participants mentioned that they were trying to remember that particular match. In a similar context, the analysis of the results in terms of the coherence of classifications given by the participants suggests that they had a higher tendency to keep the same answers, from one application to the other. Even with the possibility of this tendency for coherence being derived from the players remembering their analysis from previous tasks, this last observation is still particularly relevant, since it suggests a possible bias effect over the participants' evaluation. Due to the limited sample size, it is difficult to evaluate the amplitude of the effect caused by participants being sure that they were analysing their own data. However, it is plausible to assume that, by knowing that it was their own data being analysed, the participants may have been more motivated to conduct a more detailed or assertive analysis. More importantly, this result goes in agreement with the responses of the aforementioned questionnaire, prior to the development of VisLoL, and, thus, it further emphasizes the interest (and perhaps even need) that players have towards the understanding of their own spatio-temporal data.

6.6 Summary

This chapter continued the study of the possible combination of 2D static maps and 3D space-time cubes for trajectory data visualization by an inexperienced population, in terms of data analysis and visualization. More importantly, the knowledge and results acquired from the previous chapters were applied on a more specific context of use.

Given its increasing popularity, both in the academia and in personal usage applications, this chapter focused in the context of video-game analytics, particularly the visualization of player trajectories and spatio-temporal events generated in virtual worlds, in the game *League of Legends*.

A questionnaire was released into *League of Legends*' forums and a comparative user study was conducted to assess the importance given by the players to their spatio-temporal information during or after the game, their interest towards visualization techniques for that context, and the types of metrics they use to evaluate their performance.

In this user study, the official match-history web-page from *Riot Games* was compared with the VisLoL prototype. The first provides general aggregated statistics to the players regarding their matches, based on various thematic attributes and events occurred during the game; these include, among others, game characters death or gold/experience accumulated. The second, VisLoL, implements a map visualization that allows the user to switch between a 2D static map and a 3D space-time cube view for the visualization of the players' trajectories and the events they (or their team) participated, alongside a timeline component, focused on the representation of the information from a temporal perspective. Through this study, it was also possible to compare the types of strategies used in both applications and explore how the results of an analysis may be affected if the users are aware of the origin and context of the observed data (namely, knowing that they are analysing their own data).

Overall, the results obtained can be considered positive. In both the questionnaire and the user study, the participants have shown a noticeable concern regarding their spatio-temporal information. In addition, they have also expressed their interest in having access to that information and being able to properly visualize and analyse it, as a means of improving their game performance. More importantly, it was possible to verify a considerable acceptance towards VisLoL and its functionalities, focused on the representation of their spatio-temporal data. Although the metrics employed may be similar to those used when interacting with existing tools, the techniques used in VisLoL helped keeping an historic context of the game and, consequently, supported an analysis more focused on observable facts rather than on hypothesis derived from the players' past experiences. Consequently, the participants commented that, by using VisLoL, they were able to know *what* happened in a game and, more importantly, they were able to understand *why* something happened.

These results are very important, since they further emphasize one of the main motivations of this work, i.e., the interest towards the visualization of personal trajectory and spatio-temporal information, by inexperienced users. Moreover, they also emphasize the acceptance and capability of 2D static maps and 3D space-time cubes for the visualization of these data, by this type of users. In particular, these results suggest the adequacy of combining both 2D static maps and 3D STCs, by allowing the user to switch between both views.

An important difference from this study to the previous ones consists in the fact that these participants had a significant domain knowledge associated with the data and context of use of the application, since they were already familiarized with the game itself, the various concepts and attributes associated with it, and had also to analyse their own data. Regarding this particular factor, the results obtained point towards an effect of this knowledge over the analysis; however, not necessarily in terms of accuracy and impartiality (e.g., being more benevolent towards the own player's errors), but instead, when knowing that they were analysing their own data, the players may be encouraged to perform a more *assertive* analysis. Interestingly, despite the detection of this possible effect, the types of interactive strategies used, during the tasks, with the 2D maps and STCs, alongside their comments at the end of the tasks, were not significantly different from those obtained in the previous studies. Thus, further emphasizing the results presented in Chapters 3 to 5.

Figure 6.10 depicts a schematic summary of the contributions of this chapter in the scope of the research reported on this dissertation.

Ultimately, by the end of this study, a significant amount of knowledge was collected regarding: the differences between 2D static maps and 3D space-time cubes; possible ways of improving them, including through their combination into the same visualization; and their adequacy to represent different types of trajectory datasets and in relevant contexts of use, such as visual game analytics. Based on this knowledge, the next chapter will address the remaining main research goal of this work and take the *lessons learned* with these studies to present a set of general design guidelines.

Step 1: Related Work Analysis and Simplified Task Model (TM)

Ch2. Interactive Visualization of Trajectory Data

Results: See Figure 2.2.- For Task Model (TM) separated by 4 levels (lvl 1-4)

Step 2: Compare 2D Maps and 3D STCs

Ch3. Comparing 2D Static Maps and 3D Space-Time Cubes

Study #1 (5 participants)
Preliminary assessment of inexperienced users' ability to analyse trajectory data

Method:
- Task solving with *ST-TrajVis*
- **Data source**: GeoLife (1 mover: Nov. 2008)
- **Tasks**: *Identify and Compare* (**TM**-lvl1)
- **Logged data**: Time, Accuracy, and Preferences

Results
- Small learning curve
- Combining a 2D map with an STC was well received
- Different roles: M*ain* visualization vs supportive map
  - Varied between users
- Interest in quick visual filtering mechanisms

Study #2 (16 participants)
- Compare how adequate 2D maps and STCs are for visualization task solving by inexperienced users

Method:
- Compare 2 prototypes using map visualizations (2D map or STC) and a timeline
- **Data source**: GeoLife (32 movers: 2008-2009)
- **Tasks**: *Identify, Locate, Compare, Associate* (**TM**-lvl2)
- **Query categories**: *(elementary/general) when*, *(elementary/general) what+where*
- **Logged data**: Time, Accuracy, and Preferences

Results:
- Higher preference towards 2D map
- 2D map more effective on *location* tasks
- STC more effective on *association* tasks
- Higher performance on *elementary tasks*
- Additional comments on combining 2D map and STC
- Attempts of converting STC into 2D map

Step 3: Improve Existing Techniques

Ch4. Improving Spatial Awareness in Space-Time Cubes

Study #3: (20 participants)
- Identify strategies used to obtain spatial information on an STC
- Identify interactive techniques to help obtaining spatial information
- Compare user performance with those methods

Method:
- Compare 3 STC prototypes (*P1*: simple; *P2*: moveable plane; *P3*: map overview)
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: Same as before
- **Logged data**: Same as before + Nº actions and 3D mitigation strategies

Results:
- Higher preference towards *P2 and P3*
- Higher performance with *P2 and P3*
- Lower performance in tasks with more *visual noise*
- *P2* is good for finding locations and study one mover
- *P2* can interfere with the interaction, due to occlusion
- *P3* provides a better geographical overview
- Additional comments regarding the combination of both maps
- Attempts at converting STCs into a 2D map: Birds-eye view (*BEV*) and flattening the STC's height (*FlatV*)

Ch5. Combining 2D Static Maps and 3D Space-Time Cubes

Study #4 (30 participants)
- Expand existing knowledge regarding using 2D maps and STCs
- Explore if it makes sense to combine both techniques
- Empirically validate the combination of these techniques
- Analyse the types of strategies used to acquire information, alongside user performance and acceptance.

Method:
- Compare 3 prototypes using map visualizations (2D map, *STC*, and *3VC*: using both techniques) and a timeline
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: *spatial* and *spatio-temporal*
- **Density of information**: 1 or 10 movers
- **Logged data**: Same as before + % of component use

Results:
- Combining both techniques is useful
- *2D Map* is faster and less interactive demanding on *spatial identification* tasks, but not significantly more accurate then *STC*
- *STC* is better on *spatio-temporal* tasks
- *3VC* is preferred in *spatio-temporal comparison* tasks
- *3VC* reduces the need to transform STCs into 2D maps
- *FlatV* is a better approach than *BEV*

Step 4: Case Study

Ch6. Case Study: Visualization of Trajectory Data from Online Games

Study #5 (21 participants)
- Understand how the analysis of the data differs from using a map-based application and already existing tools
- Study the effects of knowing whose data is being visualized

Method:
- Compare *VisLoL* with *LoL*'s official match-history web-page
- **Data source**: *Riot Games* (1 or 2 movers)
- **Tasks**: *Identify, Locate, Compare, Associate* (**TM**-lvl2)
- **Target analysis**: the user is told if the data displayed is theirs
- **Logged data**: Time, Ratio of component interaction, Ratio of map mode, Coherence of grades and decisions, and Preferences

Results:
- Analysis with *VisLoL* is slower
- Metrics used and conclusions reached are similar with both applications
- Web-page: analysis based on hypothesis and previous experience
- *VisLoL*: analysis based on the data displayed (*tell a story)*
- 2D Map is more used than the STC
- Participants were more coherent when told that they were analyzing their own data

Step 5: Identification of Guidelines

Figure 6.10: Summary of the various steps of this research and the contributions of Chapter 6.

Chapter 7

Discussion and Design Guidelines

Throughout the previous chapters, several user studies were conducted with prototypes implementing 2D static maps and 3D space-time cubes (STCs). Through these experiments a relevant set of results was collected regarding: i) the differences between 2D maps and STCs, particularly their adequacy based on different types of visual analysis tasks; ii) the possible improvements that can be applied over 2D maps and STCs, including the addition of supplementary visual components or visual cues (e.g., timeline or moveable map planes), or even their combination into the same visualization; and iii) the adequacy of 2D maps and STCs to represent and interact with different types of trajectory datasets, for an inexperienced population in terms of visualization and analysis.

This chapter identifies a set of general design guidelines regarding these two techniques and aims to minimize the effort and time in the design process. The identified guidelines are based on: the overall discussion of the results of previous studies; the direct observation of the participants' behaviour during the various studies; and the analysis of the results obtained, throughout the experiments, both from a quantitative (e.g., differences in accuracy or interaction times) and a qualitative (e.g., participants comments) point of view. These guidelines are organized in four categories addressing: relevant criteria that should be taken into consideration prior to the selection of a given technique, particularly user domain knowledge and interactivity capabilities, discussed in the next session; the reasons to select a given technique over another (Section 7.2); and the types of interactive features that may be appropriate to include in said technique(s) from a temporal (Section 7.3) and spatial perspective (Section 7.4).

7.1 Domain Knowledge and User Control

As suggested by the results from the study with the VisLoL prototype, alongside some results existing in the literature (Kveladze et al., 2015), it is important to consider the effects of the users' knowledge in regards to the domain and the source of the displayed data. Despite their possible limited knowledge in terms of data analysis, if the user is familiar with the domain regarding the data displayed, the more likely for him/her to conduct a better analysis. Similarly, although it may not significantly affect the users' performance, the awareness of whom the data displayed belongs to may affect how the analysis is conducted, particularly if the data displayed belongs to the actual user. This fact is supported by the results of the previous chapter, in which participants have shown a higher coherence in terms of the results of their analysis when they were informed that the data displayed was generated by them.

Considering that, in the context of this research, the expected users are likely to be inexperienced in terms of data visualization and analysis. It is also expected that the domain and the data used is of interest to the user on a personal level. Nevertheless, regardless of the map visualization used, it is important to *emphasize the personal aspects of the data* and to provide tools and explanations to the users, as a means of *minimizing the negative effects of their limited knowledge* over the analysis. A very simple example of this consists in the inclusion of tutorials/instructions regarding the visualization and its features, e.g., describing a feature's use when the user hovers it. A somewhat similar example, implemented in VisLoL, consists of the inclusion of extended map legends providing additional explanations regarding the represented information.

In addition, it is also important to consider how to access and control the data displayed. Nowadays, regardless of the context of use and the target users, data filtering mechanisms are widely popular. Naturally, throughout the previous studies, the participants often gave a positive feedback towards the existence of these features in the prototypes. However, and more interestingly, they have also shown a considerable appreciation towards data customization mechanisms. These features allowed the user to associate a specific visual variable (e.g., colour, size or shape) to a specific temporal or thematic attribute (e.g., time period or speed). The participants of these studies have stated that, in addition to help them focus more easily on relevant information (or at least, information that they considered relevant), these functionalities also made them feel more in control of the interaction and analysis. These later comments are particularly interesting and relevant, specially considering the personal motivations that the user may have already when interacting with the data. As such, in addition to already common querying and filtering mechanisms, it is important to consider the design of interactive controls that *support the customization of data representations*.

7.2 Techniques for Trajectory Data Analysis in Visual Analysis Tasks

Throughout the previous experiments, there has been a focus on two-dimensional spatial data, i.e., not considering spatial attributes like altitude/depth; however, there is no reason to believe that 2D static maps or 3D space-time cubes would not be able to represent three-dimensional data. As already discussed in Chapter 2, this can be achieved by *converting* altitude into a thematic attribute and using visual variables like colour, or icon shape, to represent this attribute, instead of position (Gonçalves et al., 2016a).

The results of the previous studies have also shown that user performance is significantly affected by the volume of information displayed. With a limited amount of information or when the focus of the task is dependent or limited to a small subset of the data, test participants were able to perform the tasks adequately with both techniques. However, increasing the volume of information will emphasize each techniques' limitations. Consequently, this shows that the selection of a given technique is, at least in part, dependent on the type of expected tasks that may need to be conducted with them.

Overall, considering the results obtained regarding 2D static maps, it can be argued that the question is not necessarily *if*, but rather *when* this type of technique should be used. Based on the types of tasks, suggested in the literature and confirmed here, *the 2D map technique should be used when the expected tasks are of a spatial nature*. As supported by the results of Chapter 5, although this type of technique was not significantly better than the other approaches, in terms of accuracy, participants were generally faster at completing the tasks and were able to perform those tasks in a smaller number of steps/actions. In addition, users have also shown a higher preference towards 2D maps comparatively to STCs. This happened both directly, when asked about their preferences, and indirectly, when they converted the STCs into a 2D map or used the 2D map as their primary visualization component in spatial tasks.

By comparison, *the STC should be used when the expected tasks require the analysis of time, and its relation with the geographical area*. As discussed in Chapters 3 and 5, the STC is more helpful to users, than common 2D maps, on *comparison* and *association* (spatio-temporal) tasks. In fact, despite not being preferable to common 2D maps, the test participants' comments were often positive, given the STCs' ability to provide both spatial and temporal information in the same *shape*. However, despite these advantages, it is important to acknowledge the fact that *a STC should only be used with interactive controls*. As evidenced by the results of Chapters 4 to 6, and further emphasized in the remainder of this chapter, the STC technique is significantly dependent on various interactive features, including, the rotation of the STC, or the ability to highlight its information. If this technique is expected to be used in a context that does not ensure the means for a user to properly interact with it, a different approach should be considered instead (e.g., 2D map with additional temporal views).

Moreover, *when the analysis of both spatial and temporal information is needed, a combination of the two techniques is a viable option*, similarly to the approaches used in Chapters 4, 5, and 6. As evidenced by the results, this approach is generally preferred by users and effectively supports the completion of spatio-temporal tasks. Nevertheless, despite the positive results with this approach, it is important to acknowledge the increase in visual and interactivity complexity. The following sections address these issues, focusing on various types of interactive necessities and strategies that can be included.

7.3 Interactive Features for Temporal Visualization of Trajectory Data

When time is a relevant metric, it should be visible as often as possible, regardless of the type of map being used.

When using a 2D map this can be a challenging issue, since this technique is limited in representing temporal information. Nevertheless, it is possible to mitigate this issue, for example, by representing time similarly to a thematic attribute, or using additional views, within or outside the map. Throughout the experiments presented in the previous chapters, both approaches were used in two distinctive ways. First, by allowing the *customization of visual variables*, namely colour and shape, of the displayed trajectories in function of the time period of the day and, second, by providing *interactive time graph* visualizations, like timelines, alongside the map view. Based on the interaction logs recorded throughout the various tasks, it was possible to see the frequent usage of these features in tasks that required the analysis of the temporal dimension. Moreover, based on the participants' feedback, it is possible to see that these features are useful to, in the case of the visual variable customization, emphasize repetitions/patterns of movement in space and time. In the case of the timeline, this feature was useful to constantly display temporal information, that would otherwise be too confusing, if not impossible to represent on just a 2D map plane. It is important to consider that using multiple separated views to display similar or complementary information can make users having to divide their attention between the different components of the visualization (Baudisch et al., 2002). As such, should a timeline be used, it is important to *use a similar representation to the ones used in the map visualization(s)*, as a means of supporting a better transition of the user's attention from one view to the other.

When using an STC visualization, distributing information in the third dimension allows users to quickly recognize which events happened before others. However, position by itself cannot properly display more precise information, like how many days are being represented, or where/when does one day starts and other ends. The most common solution would be to place *textual labels* along the STC's height, together with timestamp information. However, this information can be occluded if the user zooms in the view too much or significantly changes the point of view over the visualization. As such, any type of temporal representation should always be visible, for example, by constantly updating the textual labels positions, assuring that at least one is always visible on screen. This particular functionality was implemented in the most recent STC visualizations described in the Chapters 5 and 6, being praised by the participants that noticed it. Specifically, they mentioned that *dynamic labels* allowed them to understand when something happened and, more importantly, to regain context operations over the STC that would change the visual state of the map, namely pan, zoom, and STC rotations.

Similarly to what was discussed regarding the 2D static map technique, *customizable temporal representations* can also be useful when interacting with STCs. As discussed in Chapter 5, despite having a visualization technique displaying information from a temporal point of view, some participants considered useful to emphasize temporal information within the STC (i.e., display the time periods of the day), to emphasize spatio-temporal patterns. This is particularly interesting, considering the argument presented by the literature (Kjellin et al., 2010b) regarding the existence of at least one additional visual variable to represent thematic information, since time is already being represented as a location within the STC. Although this is true, this argument can be further expanded by emphasizing instead the enhanced temporal perception provided by STCs, due to the possibility of being able to use, at least, one more visual variable to represent time.

In addition, *when using the STC technique, it is important to provide controls over the temporal granularity of the data* and to *make it noticeable to the user*. These controls can be compared with common *geographic zooming* techniques present in most map applications, although in this case dealing with temporal information. The approach followed in this work consisted of *changing the STC's height*. As previously discussed, this functionality was included in all prototypes using the STC technique. In addition, throughout the evolution of the prototypes implementing this technique, there was an effort into making this functionality noticeable enough or more understandable to the users. Figure 7.1 depicts the evolution of the STC's additional controls used throughout the experiments described in the previous chapters. Considering the irregular usage of these controls in the studies presented in Chapter 3 (Figure 7.1 a), in addition to some small stylistic changes, labels were included describing the function of each feature (Figure 7.1 b). This was also necessary, given the inclusion of more controls, visually similar, such as the map zoom slider. After the results of Chapter 4, in which more features were added (extra plane, shadows/heatmaps, and location or region highlights), it was necessary to further compact and group these controls, so that they would not take too much space, but would still be visible (Figure 7.1 c). Considering the possibility that some users could consider using these controls unnecessary, or even intrusive, it was also implemented the ability to hide these controls.

Ultimately, considering the participants' comments and actions, this was considered a very important feature. Throughout the experiments, several participants commented on the usefulness of these controls to improve their understanding of the data's temporal dimension. In spatio-temporal tasks, this functionality allowed participants to analyse smaller periods of time in more detail, when increasing the STC's height, or progressively compress data temporally, when reducing the STC's height.

7.4 Interactive Features for Spatial Visualization of Trajectory Data

As evidenced by the results of the previous studies, particularly those from Chapter 4, *when using an STC, it is important to provide additional spatial context cues*, to improve the analysis of this dimension, particularly in *location* based tasks.

During this work, different approaches to minimize this problem were implemented in the various STC prototypes. The simplest approach consisted in drawing lines pointing towards the

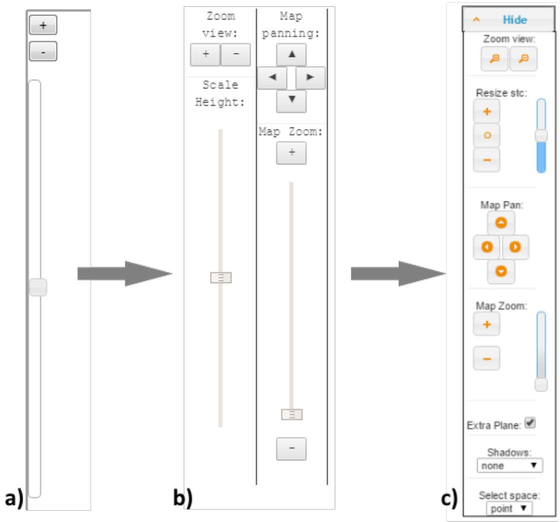

Figure 7.1: Evolution of the STC's additional controls throughout the various experiments described in a) Chapter 3; b) Chapter 4; and c) Chapters 5 and 6.

location in the map, located at the base of the STC, whenever an object was highlighted/selected in the view. Although helpful, this strategy is not enough, as shown in the results from Chapters 3 and 4, in which test participants using the prototypes implementing only this type of representation performed worse in *location*-based tasks.

To solve this problem, Chapter 4 studied the use of an *additional map plane* that can be moved along the STC's height at any point of the interaction (Kraak, 2003; Gonçalves et al., 2015c). This functionality allows users to complete tasks with a higher performance than when using a simpler STC. Nevertheless, this approach still caused some confusion, due to the possibility of this plane occluding information, depending on the point of view within the STC (see Figure 4.4). Moreover, in the study described in Chapter 5, some participants deactivated this functionality in various tasks, and mentioned that they considered this less useful when they were more interested in the temporal dimension of the data. As a consequence, these results also suggest that *the auxiliary map plane in the STC should be made an optional feature*. However, if this functionality is implemented, it is important to take into consideration the effects it may have over the visualization when combined with other features, namely when changing the temporal granularity of the data. If the user decreases the data's temporal granularity, the auxiliary plane will necessarily be moved closer to the STC's base, since it will, most likely, be associated to a specific time moment. As a consequence, this can effectively hide all information located bellow it. In some scenarios, this may be easily solved by simply rotating the STC; however, if the map plane is too close to the STC's base, it will become impossible to see it (Figure 7.2 a). It is plausible to assume that expert analysts already familiar with STCs or that become familiarized with this functionality may actually use it as a quick visual filtering mechanism. However, such behaviour is not expected from inexperienced users, who are more likely to be confused by this, similarly to what already happened when using the moveable map plane. Another possible approach to minimize this prob-

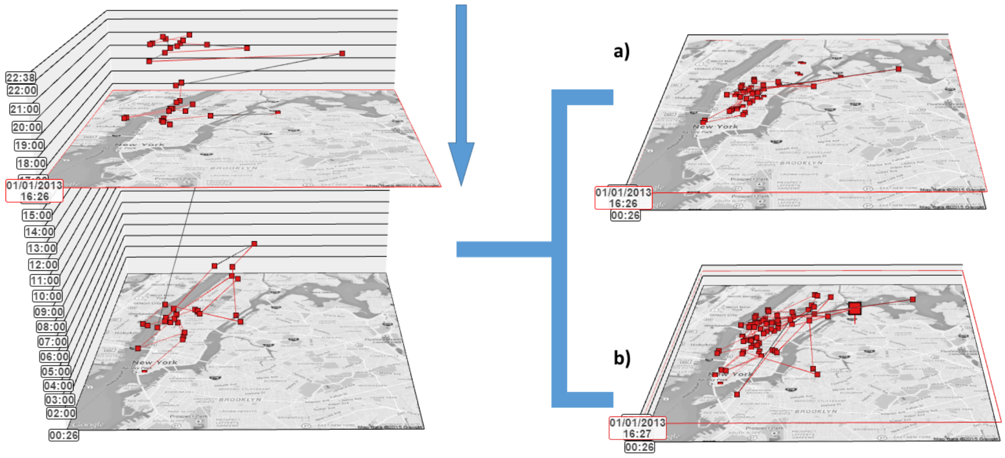

Figure 7.2: Examples of using an auxiliary map plane to highlight a given location, while adjusting the STC's temporal granularity controls, through height change: a) the auxiliary plane remains visible; b) the auxiliary plane is hidden after resizing the STC.

lem consists in (optionally) disabling the auxiliary map plane when it approaches the STC's base map given a certain threshold (Figure 7.2 b). In this work, it was used a threshold of 5% of the STC's original height; nevertheless, future implementations should support the customization of this value, through the use of an additional menu option.

An additional approach consists of combining the STC with a 2D map. Three possible alternatives were studied in this work, namely: i) providing a 2D overview of the data presented in the STC (Figure 7.3 a), such as in the *P3* prototype described in Chapter 4; ii) juxtaposing an interactive 2D map (Figure 7.3 b), present in the *3VC* prototype described in Chapter 5, and iii) allowing the user to switch between the techniques (Figure 7.3 c), present in the *VisLoL* prototype described in Chapter 6.

The first two alternatives are similar, in the sense that both of them require additional screen space to display the auxiliary map. However, they differ in the role they have in the visualization and in how the user is able to interact with them. Overall, the 2D map overview is comparable to common *mini-maps*, present in traditional large screen map applications or even games, thus, implying this map to be, generally, smaller than the STC. As such, with a 2D map overview, the STC is implied to be the *main* map visualization and, thus, the main focus of the user's attention and interaction. Comparatively, by juxtaposing both map techniques, the user has the opportunity of choosing which one should be the focus of their attention/interaction, even if it may take more screen space than the previous alternative. Considering the generally higher preference towards 2D maps over STCs, if the designer can use enough screen space for both types of maps, a design guideline should be to *use juxtaposed maps*.

On the contrary, if it is impossible to adequately accommodate more than one map view within the same visualization, an additional alternative consists of *supporting the conversion of one technique into the other*. In Chapter 3, two types of strategies were identified, applied by the participants trying to achieve this feature. In particular, they rotated the STC to look at the data from a

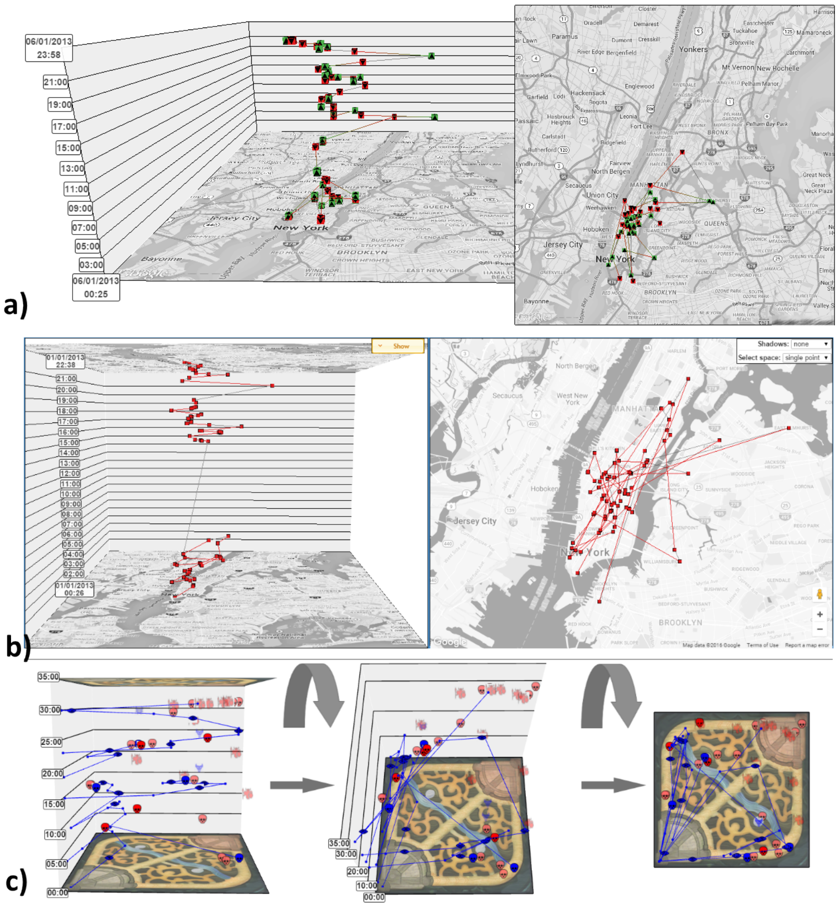

Figure 7.3: Examples of combinations between 2D static maps and 3D space-time cubes a) 2D map overview; b) map juxtaposition; and c) conversion between techniques.

bird's eye view (Figure 7.4 a), or reduced the temporal granularity to its minimum, compressing the data into the STC's base plane (Figure 7.4 b). Consequently, and due to the stylistic similarities between the two maps used, e.g., the same map planes, the same symbols to represent the same information, the approach followed in this work consisted in converting one technique into the other by automatically changing projections when the user rotates the STC to look at the data from a *bird's eye view*. In addition to not taking more screen space to represent both techniques, an important advantage of this alternative consists in the natural way in which the conversion is made. As discussed in Chapters 5 and 6, looking at the data within an STC from a *bird's eye view* perspective does not take advantage of the map's third dimension to represent time. In fact, during the experiments, participants that applied this strategy were usually not interested in this dimension. The results of this research's studies have also suggested that this strategy can have a negative effect over the analysis, as the significant results detected show that the participants that used this approach performed, generally, worse than those that compressed the STC's temporal dimension. Although flattening the STC into a plane provided better results, it can be argued that this is not an intuitive approach, as not all participants considered using it, in addition to also requiring, at least, an additional interaction step. As such, an important design guideline, when

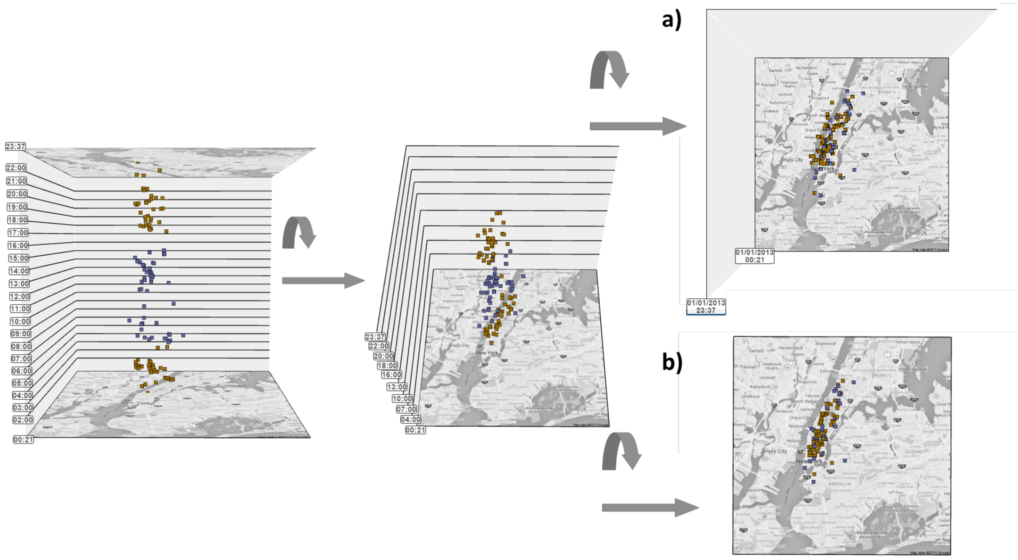

Figure 7.4: Perspective change on an STC when looking at the data from a *birds-eye view* a) no change; b) change to an orthographic perspective.

using an STC, even if already complemented by a 2D static map, is to *include the possibility of transitioning between an STC and a 2D map, when looking at the data from a birds-eye view*. In addition, *this functionality should be noticeable*, otherwise users may forget about its existence, particularly when interacting with the visualization for the first times.

Nevertheless, regardless of the approach to combine the STC and the 2D map, it is important to consider the effects of the interaction with one technique over the other. Namely, an important design guideline is to *use of similar controls for the different types of map techniques*. Traditionally, as seen in common map view applications (e.g., Google Maps), in a keyboard+mouse interaction context, the left mouse button is used to pan the map view. However, common 3D design applications (e.g., *Maya* (Corporation, 1998)) use this button to rotate the view. These analogies were used to decide the controls for the 2D map and STC visualizations and, when using the STC separately, these controls did not brought up any detectable problem. However, when using this technique alongside a 2D map, it was noticeable that participants often tried to drag/pan the STC the same way they did with the 2D map, therefore, implying the necessity to use similar controls.

In addition, if both map visualization techniques are visible, another guideline is to *keep both map views synchronized* or, if the views are alternated, to *preserve/translate the state of interaction when transitioning between the map views*. Based on the interaction logs recorded during the tasks and the feedback provided by the participants, they never deactivated the synchronization between the 2D map and the STC views, having also praised this feature for making it easier to transition from one view to the other. In fact, this guideline is also applied when using additional views or maps representing different sets of data from the same event, such as the timeline and the *second player* map implemented in the VisLoL prototype, described in Chapter 6.

7.5 Summary

In this chapter, a set of general guidelines were proposed regarding the design of 2D static maps and 3D space-time cubes, to be used by an inexperienced population in terms of visualization and analysis of trajectory data.

In particular, this chapter focused on the *lessons learned* throughout the various experiments, by taking into consideration the quantitative differences in the results obtained throughout the different tasks with the various prototypes (e.g., differences in task completion times, or number of actions), the various comments made by the participants of the studies, and the various strategies applied by them when interacting with the techniques.

The guidelines were divided into four overarching topics addressing: important factors that should be taken into account independently of the technique chosen, relevant criteria when choosing a technique, and features that should be included in the visualization techniques from a temporal and spatial perspective, respectively.

Figure 7.5 provides an overview of the proposed guidelines. As understandable by the analysis of the figure, regardless of the criteria used, a 2D static map representation of the data is always advised. There may be some scenarios in which the analysis of the data's spatial components may not be necessary; however, these are out of the scope of this research work. In addition, despite the suggestion of using an STC in a limited screen space interaction, it is still important to support the visualization of the data from a 2D map perspective.

By comparison, the scenarios in which the STC technique may be used and how it should be used is significantly more diverse. If time is an irrelevant metric for the visualization or, even if it is, the developer cannot ensure a proper interactive environment to the user (e.g., software or hardware limitations), then there is no clear advantage in using the STC technique. On the other hand, if time is a relevant metric and it is possible to interact with the visualization with enough screen space, juxtaposing both techniques is the most adequate approach. If that is not possible, however, the developer may consider an alternative in which the 2D map is less *emphasized* but still visible simultaneously to the STC, or simply allowing the conversion between one view into the other. Furthermore, whenever time is a relevant metric, it is also advised the use of additional temporal graphs, such as the timelines used in the aforementioned experiments. Although this is mostly an optional measure when an STC is also used, this feature can be particularly helpful, when a user is more focused in visualizing the data with a 2D map.

In addition, using an STC also implies the consideration of various features to simplify user interaction from both a spatial and temporal perspective, including among others, the use of temporal granularity controls or auxiliary moveable map planes. Nevertheless, regardless of which technique the users focus their attention the most, it is important to take into account how they are used and how the data is displayed, to support an easier transition between them and, consequently, maximize the users' attention, experience, and performance.

Finally, Figure 7.6 finishes the summary of the contributions of this research, first introduced in Chapter 1 (Figure 1.3).

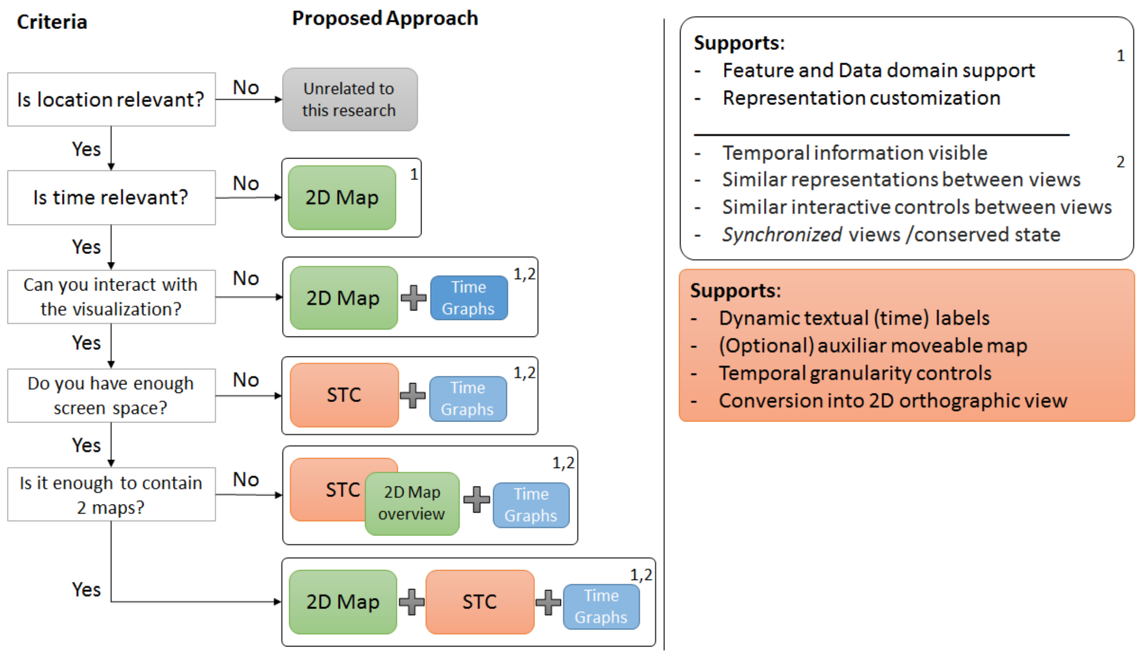

Figure 7.5: Overview of the guidelines presented in Chapter 7.

Step 1: Related Work Analysis and Simplified Task Model (TM)

Ch2. Interactive Visualization of Trajectory Data
Results: See Figure 2.2.- For Task Model (TM) separated by 4 levels (lvl 1-4)

Step 2: Compare 2D Maps and 3D STCs

Ch3. Comparing 2D Static Maps and 3D Space-Time Cubes

Study #1 (5 participants)
Preliminary assessment of inexperienced users' ability to analyse trajectory data

Method:
- Task solving with *ST-TrajVis*
- **Data source**: GeoLife (1 mover: Nov. 2008)
- **Tasks**: *Identify and Compare* (**TM**-lvl1)
- **Logged data**: Time, Accuracy, and Preferences

Results
- Small learning curve
- Combining a 2D map with an STC was well received
- Different roles: M*ain* visualization vs supportive map
  - Varied between users
- Interest in quick visual filtering mechanisms

Study #2 (16 participants)
- Compare how adequate 2D maps and STCs are for visualization task solving by inexperienced users

Method:
- Compare 2 prototypes using map visualizations (2D map or STC) and a timeline
- **Data source**: GeoLife (32 movers: 2008-2009)
- **Tasks**: *Identify, Locate, Compare, Associate* (**TM**-lvl2)
- **Query categories**: *(elementary/general) when*, *(elementary/general) what+where*
- **Logged data**: Time, Accuracy, and Preferences

Results:
- Higher preference towards 2D map
- 2D map more effective on *location* tasks
- STC more effective on *association* tasks
- Higher performance on *elementary tasks*
- Additional comments on combining 2D map and STC
- Attempts of converting STC into 2D map

Step 3: Improve Existing Techniques

Ch4. Improving Spatial Awareness in Space-Time Cubes
Study #3: (20 participants)
- Identify strategies used to obtain spatial information on an STC
- Identify interactive techniques to help obtaining spatial information
- Compare user performance with those methods

Method:
- Compare 3 STC prototypes (*P1*: simple; *P2*: moveable plane; *P3*: map overview)
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: Same as before
- **Logged data**: Same as before + Nº actions and 3D mitigation strategies

Results:
- Higher preference towards *P2 and P3*
- Higher performance with *P2 and P3*
- Lower performance in tasks with more *visual noise*
- *P2* is good for finding locations and study one mover
- *P2* can interfere with the interaction, due to occlusion
- *P3* provides a better geographical overview
- Additional comments regarding the combination of both maps
- Attempts at converting STCs into a 2D map: Birds-eye view (*BEV*) and flattening the STC's height (*FlatV*)

Ch5. Combining 2D Static Maps and 3D Space-Time Cubes
Study #4 (30 participants)
- Expand existing knowledge regarding using 2D maps and STCs
- Explore if it makes sense to combine both techniques
- Empirically validate the combination of these techniques
- Analyse the types of strategies used to acquire information, alongside user performance and acceptance.

Method:
- Compare 3 prototypes using map visualizations (2D map, *STC*, and *3VC*: using both techniques) and a timeline
- **Data source**: *New York Taxis* (10 movers: Jan. 2013)
- **Tasks**: *Identify, Compare* (**TM**-lvl1)
- **Question categories**: *spatial* and *spatio-temporal*
- **Density of information**: 1 or 10 movers
- **Logged data**: Same as before + % of component use

Results:
- Combining both techniques is useful
- *2D Map* is faster and less interactive demanding on *spatial identification* tasks, but not significantly more accurate then *STC*
- *STC* is better on *spatio-temporal* tasks
- *3VC* is preferred in *spatio-temporal comparison* tasks
- *3VC* reduces the need to transform STCs into 2D maps
- *FlatV* is a better approach than *BEV*

Step 4: Case Study

Ch6. Case Study: Visualization of Trajectory Data from Online Games

Study #5 (21 participants)
- Understand how the analysis of the data differs from using a map-based application and already existing tools
- Study the effects of knowing whose data is being visualized

Method:
- Compare *VisLoL* with *LoL*'s official match-history web-page
- **Data source**: *Riot Games* (1 or 2 movers)
- **Tasks**: *Identify, Locate, Compare, Associate* (**TM**-lvl2)
- **Target analysis**: the user is told if the data displayed is theirs
- **Logged data**: Time, Ratio of component interaction, Ratio of map mode, Coherence of grades and decisions, and Preferences

Results:
- Analysis with *VisLoL* is slower
- Metrics used and conclusions reached are similar with both applications
- Web-page: analysis based on hypothesis and previous experience
- *VisLoL*: analysis based on the data displayed (*tell a story)*
- 2D Map is more used than the STC
- Participants were more coherent when told that they were analyzing their own data

Step 5: Identification of Guidelines

Ch7. Discussion and Design Guidelines
Results: See Figure 7.5. – For summary of guidelines

Figure 7.6: Summary of the various steps of this research and the contributions of Chapter 7.

Chapter 8

Conclusions and Future Work

The Ph.D. research reported in this dissertation addressed the study of various properties of 2D static maps and 3D space-time cubes (STCs) for the visualization of trajectory data, particularly focused on their utilization by an inexperienced population, in terms of data visualization and analysis.

The main objectives proposed for this research were three-folded:

- Obtain empirical data supporting the dis/advantages of a given technique;
- Propose improvements to existing techniques focusing on inexperienced users, and the dis/advantages of combining different techniques;
- Identify design guidelines for map-based trajectory data visualization, based on the two points above.

To address these goals, in addition to the analysis of related work described in the literature, several prototypes were created, implementing variants of techniques based on 2D static maps and 3D STCs to support the visualization of different trajectory datasets. These prototypes were subjected to usability tests, to compare the different sets of variants and evaluate their adequacy in representing movement data. Consequently, a series of quantitative and qualitative results were obtained, which support the basis of the various contributions.

The following sections revisit and discuss the most significant contributions presented and address relevant and unexplored challenges in the context of this research.

8.1 Primary Contributions

The primary contributions of this work can be divided into four main topics, further discussed bellow.

8.1.1 Simplified Task Model - Literature Analysis

Based on the proposed objectives, it is possible to understand that this research is directly associated with multiple disciplines, such as human-computer interaction, visualization and cartography.

Consequently, a requirement for this work consisted in the analysis of various relevant sources existing in the literature, associated to these areas of knowledge, as seen in Chapter 2. This analysis was focused on three main factors: i) main components and characterizations associated with trajectory data; ii) types of user objectives/tasks that can be conducted when using any given visualization; and iii) characterization of the most general types of map visualization techniques to represent trajectory data.

The analysis of the second factor (visual analysis tasks) revealed a significant redundancy in the various models/taxonomies present in the literature. By one hand, the same tasks were identified by several different studies, which shows some agreement between the various sources. However, several similar or overlapping tasks were given different denominations. Considering the dependency of this research on the identification and understanding of these tasks to study which map-based techniques are the most adequate, it was necessary to clarify and organize the already existing lists os tasks. Therefore, the main contribution presented in Chapter 2 consisted in the organization of the most significant types of visual analysis tasks present in the literature.

This contribution is particularly important due to three main reasons. First, this systematization provides a much needed compilation and simplification of the proposed tasks, making it usable as a possible reference for future research. Second, this model acknowledges the inherent hierarchy existing in the various proposed tasks, in terms of their specificity. Finally, as a result of the previous two reasons and as shown throughout this document, it was possible to conduct various (successful) usability studies using these tasks as one of the basis for comparing the various developed approaches.

8.1.2 Comparisons Between Map-based Visualization Techniques

After the study of the main properties related with trajectory data and the most relevant types of tasks associated with data visualization, the second main contribution from this work addressed the comparative evaluation of various map-based prototypes supporting the visualization of trajectory datasets. This contribution was primarily addressed in Chapter 3 and, based on the results obtained, they were also followed up between Chapters 4 to 6. The focus of this contribution consisted in the comparison of 2D static maps and 3D STCs. The literature also supports the existence of other techniques, namely animated and small-multiple maps. However, it was decided to exclude them from the focus of this work, based on the observation that exists an inherent dependency between these techniques and the former, since both animated and small-multiple maps can be displayed over 2D maps and 3D STCs.

It was hypothesized that the participants' unlikely familiarity with STCs, associated with the known limitations in human perception of 3D graphics, would cause the 2D static map prototypes to be favoured by the participants and to support a higher performance. Through the conduction of usability studies, the main conclusions drawn from the experiments associated to this topic were the following:

- Inexperienced users in terms of data visualization and analysis are, indeed, able to interact and work out conclusions, regarding the analysis of trajectory datasets, with the help of 2D static maps and 3D STCs;

- Both techniques are adequate in general visual analysis tasks, particularly tasks in which the analysis is focused on one trajectory or on one specific moment in time;
- Users prefer 2D static map visualizations;
- 2D static map visualizations support a higher performance in *locate* tasks and tasks focused on the spatial components of the data;
- STC visualizations support a higher effectiveness in *associate* tasks and *comparison* tasks dependent on the analysis of temporal information;
- STC-based visualizations require a higher degree of interactivity, e.g., a higher number of actions, to conduct various tasks.

2D static map prototypes were always among the preferred approaches by the participants. This was one of the most consistent results across the various experiences, as 2D static maps, or other prototypes supporting 2D static map features were often ranked higher, in terms of preferences, than other approaches (sometimes, regardless of the participants' performance).

Despite not being the preferred technique, the STC visualizations received, overall, positive feedback from the participants and were shown to be more adequate than 2D static maps, typically in temporal-related tasks. Even though some previous works partially support these results, these positive findings obtained with the STC are considerably particularly surprising and worth discussing.

Regardless of the expected advantages of STCs, given their combined representation of spatial and temporal information in 3D, and despite the concerns in providing some practice time to the participants, prior to the beginning of the user studies, it is important to acknowledge the fact that the participants' familiarity with this technique was significantly lower than the one they had with 2D static maps. In turn, it was expected that this factor could have a stronger effect in the participants' abilities in interacting with the technique and analysing the data, which could ultimately be reflected in the results. As further discussed in the next section, although these results also emphasized some of the limitations present in STCs they also open new and interesting opportunities for the use of this technique.

8.1.3 Improvements over Existing Techniques

The comparative analysis of 2D static maps and STCs, addressed by the aforementioned studies, helped exposing some relevant limitations associated with these techniques. Consequently, as detailed throughout Chapters 4 to 6, the objective was to explore different ways to improve these techniques.

Overall, the results obtained suggest that STC visualizations were less adequate in tasks dealing with the location of information and spatial data. Comparatively, the 2D static map visualizations were less adequate in tasks that required a more detailed analysis of the data's temporal dimensions. In addition, throughout some of the experiments, several participants have mentioned their interest in not having to chose between either technique but, rather, have access to both of them. More interestingly, based on the analysis of the participants' interactive strategies, it was possible to notice that, throughout some of the experiments, several participants tried to somewhat convert the 3D space-time cube into a 2D static map. Ultimately, this fact emphasizes the limitations associated with STCs and the willingness of being able to use both techniques.

When studying possible improvements over the existing techniques, this research focused on two main approaches. First, by studying possible representational and interactive features to improve spatial awareness in STC visualizations. Second, by studying the adequacy and the consequences of combining 2D static maps with STCs. To achieve this goal the following alternatives were compared:

- the continuous inclusion of time-graph visualization components to be used alongside the map visualizations, particularly those implementing 2D static maps;
- the use of simple spatial cues (e.g., points and lines), pointing towards the location of any highlighted/selected information within the STC;
- the use of an additional moveable map plane, to help reducing possible perspective issues, when examining information within the STC;
- the use of a 2D overview of the information present in the STC;
- the juxtaposition of interactively/logically connected 2D static maps with STCs;
- the transition between a 2D static map and an STC, through the alteration of the visualization's projection (*orthogonal* into *perspective*).

It was hypothesized that both user performance and satisfaction would increase with these improvements, despite their possible disadvantages. The conduction of usability studies, the analysis of the obtained results and the participants' behaviour allowed to conclude that:

- Users have a higher preference towards improved STCs, i.e., with a moveable map plane feature and a 2D overview, comparatively to more simple approaches;
- User performance is generally higher with the improved STCs, comparatively to more simple approaches;
- The use of an additional moveable map plane, within the STC, is more likely to help users finding specific locations and analysing the evolution of one specific mover. However, the occlusion caused by this feature can interfere with the user interaction;

- The use of a 2D map overview, within the STC, is more likely to provide a better synopsis of the geographical distribution of the data;
- Although outclassed by 2D static maps, in terms of (temporal and interactivity) efficiency, improved STCs may not differ significantly from these techniques, in terms of accuracy;
- The combination of 2D static maps with STCs is a viable approach for inexperienced users to visualize spatio-temporal data, as it was never outperformed by other approaches focused on just one of the visualizations;
- The combination of both techniques (through juxtaposition) is preferred in spatio-temporal comparison tasks, comparatively to (isolated) 2D static maps and STCs. Similarly, the ability to convert one technique into the other is well accepted by inexperienced users (even if with prior knowledge over the data);
- The combination of both techniques (through juxtaposition) minimizes the tendency of users trying to convert STCs into 2D static maps;
- Out of the two strategies identified to ignore/minimize the third dimension in STCs, namely look at the data from a *bird's-eye view* and reduce the STC's height/temporal granularity to its minimum, the second approach supports a better performance than the first.

As explained throughout the previous chapters and considering the motivations of this research in regards to this topic, the obtained results make sense. However, it is particularly interesting to analyse the positive results that were obtained, in terms of subjective preferences, regarding the combination of both visualization techniques through juxtaposition. An important limitation associated to this approach consists in the redundancy resultant of showing both maps, i.e., somewhat similar techniques, displaying some common data. Furthermore, given the inclusion of an STC component in the visualization, this approach will, most likely, require a larger amount of interactivity by the user to properly analyse the data. Consequently, and considering the reduced experience from the participants, in terms of data visualization and analysis, it would be plausible to assume that these would prefer a simpler or minimalistic design, which was not the case. Although further exploration is advised, these results open the possibility for developers considering the usefulness of more complex visual approaches in more casual/personal visualization contexts.

8.1.4 Design Guidelines

The quantitative and qualitative results of the usability studies conducted throughout this research, described in Chapters 3 to 6, alongside their development process, allowed to identify several patterns. These support and suggest the existence of a series of *best practices* when designing applications for the visualization of trajectory data, based on 2D static maps and 3D space-time cube visualizations, for inexperienced users.

Based essentially on this information, Chapter 7 described a set of design guidelines, addressing three main subjects. First, it was discussed the importance of considering the effects of the users' previous knowledge regarding the data alongside their possible motivations over the analysis. Although further studies should be conducted, current results suggest a more assertive research when the users are aware that they are analysing their own data. In addition, given their constantly positive feedback, alongside their interactive strategies, it was also discussed the interest and possible advantages of enabling the customization of the data representations by the users, both as a means of improving their analysis of the data and to make them feel in control over it.

Secondly, the focus changed to various criteria needed for deciding which visualization technique should be implemented. Overall, the results of the studies allowed to conclude that a 2D static map should be always included in the visualization, even if not as the primary map component. On the other hand, the main decisions are related to the usage of an STC and how it should be incorporated alongside the 2D static map. For these decisions, the analysis of this work allowed to identify three main factors: i) Relevance of temporal information over the analysis – if there is none, the STC is, most likely, unnecessary; ii) Ability to interact – if the users are limited in how they can interact with the visualization, the STC is not advisable; and iii) Screen space – depending on the available space on screen, this can determine the use of an STC convertible into a 2D static map (no space), an STC with a 2D overview (limited space), or the simultaneous use of both techniques in the same view.

Finally, this contribution ended with the discussion of various interactive features and practices that should be taken into account when using these techniques, from a spatial and temporal perspectives. The most significant ones include:

- Temporal information visible (e.g., timeline, dynamic time labels);
- Similar representation between views;
- Similar interactive controls between views;
- Synchronized views and conserved/translated state at visualization transition;
- Optional moveable map plane on STCs;
- STC temporal granularity controls (e.g., STC height resize);
- Support of orthographic projection with STC.

A significant number of the proposed features are directly associated to STC visualizations. Although this can, in part, be justified by the relevance that this technique had in the aforementioned studies, it also significantly emphasizes the critical role of interactivity when using this technique. Despite its various advantages, this type of visualization still has several drawbacks that can be minimized through adequate interactive features.

8.2 Complementary Contributions

In addition to the results described in the previous sections, the development and evaluation stages of the conducted usability studies provided an additional set of contributions.

Given the necessity of creating several different prototypes based on the same base techniques, three complementary APIs were developed for the creation of HTML 5 applications for the visualization of trajectory data on interactive 2D static maps, 3D space-time cubes, and timelines. Overall, the creation of these APIs significantly simplified the development process of the various prototypes, allowing for a better focus on the remaining components of the work.

On the other hand, even at the early stages of the research, it became clear that common usability metrics, namely task completion time, accuracy, and subjective preferences, would not be enough to properly evaluate the various visualization techniques. As such, it was necessary to adopt additional metrics to complement the analysis, including:

Number of actions: considering the already mentioned importance of interactivity with some of these approaches, being able to understand the amount of actions necessary to answer to spatial or temporal related questions can work as a complementary metric to time, to analyse how complex a given task may have been to the user;

Ratio of visualization used: considering that the developed approaches often involved using more than just one visualization component, such as the timelines or the additional map views, this metric allowed to analyse (and confirm) which components were the most used (or when they where used) during the various tasks;

3D mitigation actions/visualization type transitions: considering the perceptual limitations associated to 3D visualizations, alongside the, later, possibility of converting one map technique into another, similarly to the previous case, this metric supported the analysis and confirmation of the need and usefulness of specific interactive features for the STC (e.g., temporal granularity controls or projection conversion).

8.3 Future Work

Despite the various results and contributions that were obtained, there are still several topics related to those addressed by this work that should be further investigated. The following sections list relevant and promising topics for future investigation.

8.3.1 The Role, Effects, and Interactivity Requirements of Animated and Small-multiple Maps

The focus of this research consisted in the study of visualization techniques based on 2D static maps and 3D STCs. However, such as discussed in Chapter 2, the literature presents additional types of techniques that can be taken into consideration, namely animated and small-multiple maps.

Existing works in the literature show some contradictions in regards to the actual usefulness of animation for data visualization and analysis (Tversky et al., 2002). Nevertheless, they have shown the interest and the necessity in expanding the knowledge on this type of technique (Harrower and Fabrikant, 2008), and suggest a generally high user acceptance towards the use of animated techniques (Robertson et al., 2008). Considering the focus of this research on inexperienced users, in terms of data visualization and analysis, it is plausible to assume that the contexts in which applications for trajectory data visualization are used are mostly *personal* or *informal*. Consequently, even if not significantly better than static approaches, providing the user with a visualization considered to be more *fun* and appellative can lead to promising results.

In addition, the results of this research emphasize significantly the role of interactivity when dealing with these techniques, particularly with STCs, alongside the advantages and disadvantages of combining different types of visualization techniques. As addressed in Chapter 2, animated maps can be projected over 2D static maps or STCs. However, while STCs are significantly dependent on interaction, considering the inherent properties of animated maps, it is plausible to assume that an animated map would promote a somewhat more passive interaction (Gonçalves et al., 2016a).

Consequently, future work will address the study of animated maps for trajectory data visualization, also focused on an inexperienced population and exploring:

- The role of animated maps in the visualization, depending on the type of technique into it is projected, e.g., if animation is used over a 2D map or a 3D STC;
- The types of tasks in which this technique may support a better performance, similarly to what had been conducted in this research with 2D static maps and STCs;
- The users' appreciation towards this type of visualization, alongside the results of their analysis with them;
- The type of interactive controls required to control the flow of the animation and, thus, explore the data.

At the time of this documents' writing, some of these issues are already being addressed by one of the *Legacy* projects, described in Section 1.3.4. In particular, the last project, *Visualization of Spatio-temporal Information for Personal Performance Analysis in Games*, is studying the adequacy of using animated 2D maps for the representation of trajectory data to video-game players, likely familiar with the domain associated with the data, but highly inexperienced in terms of spatio-temporal data visualization and analysis.

To a certain extent, the use of this technique was also addressed in the first project of the same section, *Visualization of Human Trajectories on Mobile Devices* (Vieira, 2016). Although focused on the visualization of data on a mobile device context and, despite the small sample of users approached, the results of this research suggest that animated 2D maps may be less interactivity demanding and more appellative to the users, while not differing significantly, in terms of effectiveness, from static map alternatives.

Similarly, small-multiple maps are also a visualization approach promising of interesting results. Some of the approaches followed in this work considered the use of more than one map visualization, either by displaying two different map techniques with similar information (see Chapter 5), or multiple instances of the same technique, but displaying different sub-sets of data regarding the same event (see Chapter 6). Considering that small-multiple maps follow the same overall principle, it would be interesting to evaluate how the conclusions and guidelines obtained are applicable over this technique. However, it is important to notice that, in this work, at maximum, only two maps were displayed simultaneously, while small-multiple maps are likely to require a higher number.

As discussed in Chapter 2 and later shown in Chapter 7, when deciding on the type of map approach to use, screen size is also an important factor to consider. This factor is also particularly important when working with small-multiple maps, thus making it also an important topic for discussion regarding this technique.

Consequently, in addition to the study of animated maps, it is also important to study small-multiple map approaches for trajectory data visualization, in order to understand:

- The effects of screen-size and number of maps over the visualization (i.e, analyse how much the complexity of the visualization can scale in function of the users' data and visualization capabilities);

- The role of small-multiple maps in the visualization, depending on the technique into it is projected (i.e., displaying various 2D static maps or STCs);

- The types of tasks in which this technique may support a better performance;

- The type of interactive controls required to properly connect the information between the various views.

Moreover, alongside the various additional controls implemented with the studied visualizations, the previous experiments also allowed the user to customize how the information was displayed over the maps. According to the comments from the study participants, this is a very important feature and supports one of the identified design guidelines. Although no noticeable differences were detected when comparing and analysing 2D static maps and 3D STCs, it is still important to take into consideration how users may work with this functionality to visually codify the displayed information and to study how this may change in function of the task and, more importantly, the visualization technique.

8.3.2 User Bias and Context Studies

Throughout this work, one of the main concerns with the design of the various studies and prototypes was to avoid any prior knowledge and personal subjective effects over the data and applications that were to be used. This allowed to focus better on the study of these techniques in terms of their spatial and temporal representations. Moreover, this also prevented the effects of preconceived notions over the various analyses conducted by the users.

As previous discussed, although lacking experience in terms of visualization and analysis, it is plausible to assume that interested users in this type of visualizations are, at least, knowledgeable in terms of the domain or the data they are exploring. As discussed in Chapter 6, although this factor does not seem to affect *how* users interact with the visualization techniques, it may have an effect on the final conclusions obtained through the analysis. As such, it is necessary to study the possible effects of personal bias over the analysis and how to prevent or take advantage of them, depending on their negative or positive effects, respectively. In particular, it is important to consider that the target audience is likely to have a significantly varied level of domain expertise, such as observed in the study discussed on Chapter 6.

In addition, as discussed in Chapter 6, the users' awareness of whom the data being visualized belongs is likely to affect the users' analysis, in terms of assertiveness and interest. This fact supports one of the identified design guidelines, namely *to emphasize the personal aspects of the data*. To a certain degree, this issue is closely related to the challenges faced by the research areas focused on social networks and persuasion, in the sense that both take advantage of the users' personal information, whether it is public or private, to capture the users' interest towards specific applications, to optimize their social interactions, or change/improve on old habits. As a consequence, it will be advantageous to study and understand how the knowledge acquired in these areas of research can be applied to improve the users' capabilities with tools for personal analysis, particularly with spatio-temporal data.

Moreover, this research used the area of (personal) video-game analytics as a case study, a topic with increasing relevance due to its combination of *traditional* visual analytics research practices with the interests of various groups of typically non-data analysts. As such, future work should include the continuous application of the results associated to these techniques into this context of use, while not ignoring others in equally relevant domains, such as the analysis of personal training exercises (e.g., running or biking) and other information associated with personal health (e.g., propagation of diseases, social networking). Although these areas share similar concepts (space/time/objects), it is important to compare how the different personal contexts may affect the analysis and the types of visualization techniques and their interactive features.

Bibliography

Aigner, W., Miksch, S., Müller, W., Schumann, H., and Tominski, C. (2007). *Visualizing time-oriented data*. Springer, 1st edition.

Amar, R., Eagan, J., and Stasko, J. (2005). Low-level components of analytic activity in information visualization. In *Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 1–15.

Amar, R. and Stasko, J. (2004). A knowledge task-based framework for design and evaluation of information visualizations. In *Proceedings of the 2004 IEEE Symposium on Information Visualization*, pages 143–150.

Amini, F., Rufiange, S., Hossain, Z., Ventura, Q., Irani, P., and McGuffin, M. (2015). The impact of interactivity on comprehending 2D and 3D visualizations of movement data. *IEEE Transactions on Visualization and Computer Graphics*, 21(1):122–135.

Andrienko, G. and Andrienko, N. (2008). Spatio-temporal aggregation for visual analysis of movements. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pages 51–58.

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. *International Journal of Geographical Information Science*, 24(10):1577–1600.

Andrienko, N. and Andrienko, G. (2005). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer, 1st edition.

Andrienko, N. and Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. *Cartographica: the International Journal for Geographic Information and Geovisualization*, 42(2):117–138.

Andrienko, N. and Andrienko, G. (2013). Visual analytics of movement: a rich palette of techniques to enable understanding. In *Mobility Data: Modeling, Management and Understanding*, chapter 9, pages 1–27. Cambridge University Press, 1st edition.

Andrienko, N., Andrienko, G., Bak, P., Keim, D., and Stefan, W. (2013). *Visual Analytics of Movement*. Springer, 1st edition.

Andrienko, N., Andrienko, G., and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6):503–541.

Andrienko, N., Andrienko, G., Pelekis, N., and Spaccapietra, S. (2008). Basic concepts of movement data. In Giannotti, F. and Pedreschi, D., editors, *Mobility, Data Mining and Privacy*, chapter 2, pages 15–38. Springer, 1st edition.

Andris, C. and Hardisty, F. (2011). Extended abstract: Visualizing migration dynamics using weighted radial variation. In *Proceedings of GeoViz Hamburg*, pages 1–3.

Bach, B., Dragicevic, P., Archambault, D., Hurter, C., and Carpendale, S. (2014). A review of temporal data visualizations based on space-time cube operations. In *Proceedings of the Eurographics Conference on Visualization*, pages 1–19.

Bach, B., Dragicevic, P., Archambault, D., Hurter, C., and Carpendale, S. (2016). A descriptive framework for temporal data visualizations based on generalized space-time cubes. *Computer Graphics Forum*, pages 1–26.

Baudisch, P., Good, N., Bellotti, V., and Schraedley, P. (2002). Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the 20th SIGCHI Conference on Human Factors in Computing Systems*, pages 259–266.

Belicza, A. (2017). Sc2gears. Retrived from https://sites.google.com/site/sc2gears/. Last access: 09/03/2017.

Bertin, J. (1967). *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press, 1st edition.

Bisantz, A., D'Arcy, J.-F., Kerker, D., Hegde, S., Guan, P., Voshell, M., and Kilgore, R. (2014). Maps, space-time cubes, and meta-information for understanding path information. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1):370–374.

Bischoff, J., Maciejewski, M., and Sohr, A. (2015). Analysis of berlin's taxi services by exploring gps traces. In *Proceedings of the 2015 International Conference on Models and Technologies for Intelligent Transportation Systems*, pages 209–215.

Bleisch, S. (2012). 3D Geovisualization – Definition and Structures for the Assessment of Usfulness. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume I-2, pages 129–134.

Blok, C. (2000). Monitoring change: Characteristics of dynamic geo-spatial phenomena for visual exploration. *Spatial Cognition II, Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, pages 16–30.

Booker, J., Buennemeyer, T., Sabri, A., and North, C. (2007). High-resolution displays enhancing geo-temporal data visualizations. In *Proceedings of the 45th Annual Southeast Regional Conference*, pages 443–448.

Buard, E. and Brasebin, M. (2011). Visual exploration of large animal trajectories. In *Proceedings of 25th International Cartographic Conference*, pages 1–7.

Cabello, R. (2010). ThreeJS. Retrived from: https://threejs.org/. Last accessed: 19/07/2017.

Carreiro, N. (2016). Técnicas de visualização para melhor o desempenho em jogos online. Master thesis, Faculdade de Ciências Universidade de Lisboa, Departamento de Informática.

Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16.

Corporation, A. S. (1998). Autodesk maya. Retrived from: http://www.autodesk.com/products/maya/overview/. Last access: 19/07/2017.

Crampton, J. W. (2002). Interactivity types in geographic visualization. *Cartography and Geographic Information Science*, 29(2):85–98.

Cutler, M. E. (1998). *The effects of prior knowledge on children's abilities to read static and animated maps*. PhD thesis, University of South Carolina.

Dahi, R. (2009). Nodejs. Retrived from: https://nodejs.org/en/. Last access: 19/07/2017.

de Oliveira, M. G. a. and de Souza Baptista, C. (2012). Geostat-a system for visualization, analysis and clustering of distributed spatiotemporal data. In *Proceedings of 13th Brazilian Symposium on GeoInformatics*, pages 108–119.

Demissie, B. (2010). *Geo-Visualization of Movements: Moving Objects in Static Maps, Animation and The Space-Time Cube*. VDM Publishing.

Demsar, U. and Virrantaus, K. (2010). Space-time density of trajectories: exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science*, 24(10):1527–1542.

Deng, K., Xie, K., Zheng, K., and Zhou, X. (2011). Trajectory indexing and retrieval. In *Computing with Spatial Trajectories*, chapter 2, pages 35–60. 1st edition.

Ding, Y., Liu, S., Pu1, J., and Ni, L. M. (2013). Hunts: A trajectory recommendation system for effective and efficient hunting of taxi passengers. In *Proceedings of the 14th International Conference on Mobile Data Management*, pages 107–116.

Dodge, S., Weibel, R., and Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information Visualization*, 7(3):240–252.

Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2):112–122.

Drachen, A. and Schubert, M. (2013a). Spatial game analytics. In *Game Analytics, Maximizing the Value of Player Data*, chapter 17, pages 365–402. Springer, 1st edition.

Drachen, A. and Schubert, M. (2013b). Spatial game analytics and visualization. In *IEEE Conference on Computational Intelligence in Games 2013*, pages 1–8.

Drecki, I. and Forer, P. (2000). Tourism in New Zealand - International visitors on the move (a1 cartographic plate); tourism, recreation research and education centre: Lincoln University, Lincoln.

Eccles, R., Kapler, T., Harper, R., and Wright, W. (2008). Stories in geotime. *Information Visualization*, 7(1):3–17.

Elmqvist, N. and Yi, J. S. (2012). Patterns for visualization evaluation. In *Proceedings of the 2012 BELIV Workshop: Beyond Time and Errors - Novel Evaluation Methods for Visualization*, pages 12:1–12:8.

Endomondo LLC (2007). Endomondo. Retrived from https://www.endomondo.com/. Last access: 19/07/2017.

FitnessKeeper (2009). RunKeeper. Retrived from https://runkeeper.com/. Last accessed: 19/07/2017.

Flanagin, A. J. and Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, 72(3):137–148.

Frank, A. U. (1998). Different types of times in GIS. In *Spatial and Temporal Reasoning in GIS*, chapter 3, pages 40–62. Oxford University Press, Oxford, 1st edition.

Gao, M., Zhu, T., Wan, X., and Wang, Q. (2013). Analysis of travel time patterns in urban using taxi gps data. In *Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pages 512–517.

Garlandini, S. and Fabrikant, S. I. (2009). Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In *Proceedings of the 9th international conference on Spatial information theory*, pages 195–211.

Girardin, F., Calabrese, F., Fiore, F., Ratti, C., and Blat, J. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4):36–43.

Goldsberry, K. (2012). Courtvision: New visual and spatial analytics for the NBA. In *Proceedings of the 2012 MIT Sloan Sports Analytics Conference*, pages 1–7.

Gonçalves, T. (2017a). STCJS. Retrived from https://github.com/Encait/STCJS. Last access: 19/07/2017.

Gonçalves, T. (2017b). TIMELINEJS. Retrived from https://github.com/Encait/TimelineJS. Last access: 19/07/2017.

Gonçalves, T. (2017c). TRAJMAP2DJS. Retrived from https://github.com/Encait/TrajmapJs. Last access: 19/07/2017.

Gonçalves, T., Afonso, A. P., Ferreira, A., and Vieira, A. R. (2016a). Trajectory data visualization on mobile devices with animated maps. In *Proceedings of the 37th Annual Conference of the European Association for Computer Graphics: Short Papers*, pages 9–12.

Gonçalves, T., Afonso, A. P., and Martins, B. (2014). Visualization techniques of trajectory data: Challenges and limitations. In *Proceedings of the 2nd AGILE PHD School 2013 CEUR Workshop*, pages 1–10.

Gonçalves, T., Afonso, A. P., and Martins, B. (2015a). Cartographic visualization of human trajectory data: Overview and analysis. *Journal of Location Based Services*, 9(2):138–166.

Gonçalves, T., Afonso, A. P., and Martins, B. (2015b). Combining and assessing 2D maps and space-time cubes for trajectory data. In *Proceedings of the 29th 2015 British HCI Conference*, pages 273–274.

Gonçalves, T., Afonso, A. P., and Martins, B. (2015c). Improving spatial awareness for human trajectory visualization in space-time cubes. In *Proceedings of the 15th IFIP TC 13 International Conference – INTERACT 2015, Part IV*, pages 327–334.

Gonçalves, T., Afonso, A. P., and Martins, B. (2016b). Why not both? – Combining 2D maps and 3D space-time cubes for human trajectory data visualization. In *Proceedings of 30th British BCS Human Computer Interaction Conference*, pages 1–10.

Gonçalves, T., Vieira, A. R., Afonso, A. P., and Ferreira, A. (2015d). PATH - visualização de percursos pessoais com mapas animados em dispositivos móveis. In *Proceedings of INFORUM - Simpósio de Informática*.

Gonçalves, T., Afonso, A. P., and Martins, B. (2013a). Visual analysis of mobility data. In *IEEE 14th International Conference on Mobile Data Management (MDM)*, volume 2, pages 7–10.

Gonçalves, T., Afonso, A. P., and Martins, B. (2014). Visualizing human trajectories: Comparing space-time cubes and static maps. In *Proceedings of 28th British BCS Human Computer Interaction Conference*, pages 207–212.

Gonçalves, T., Afonso, A. P., Martins, B., and Gonçalves, D. (2013b). ST-TrajVis: Interacting with trajectory data. In *Proceedings of the 27th International BCS Human Computer Interaction Conference*, pages 48:1–48:6.

Google (2005). Google Maps. Retrived from: https://developers.google.com/maps/. Last accessed: 19/07/2017.

Green, M. (1998). Toward a perceptual science of multidimensional data visualization: Bertin and beyond. *ERGO/GERO Human Factors Science*, 8:1–30.

Griffin, A. L., MacEachren, A. M., Hardisty, F., Steiner, E., and Li, B. (2006). A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers*, 96(4):740–753.

Hägerstrand, T. (1970). What about people in regional science? *Papers of the Regional Science Association*, 24(1):6–21.

Harrower, M. and Fabrikant, S. (2008). The role of map animation for geographic visualization. In *Geographic Visualization Concepts, Tools and Applications*, chapter 4, pages 49–65. Wiley, 1st edition.

Heckbert, P. S. and Garland, M. (1997). Survey of polygonal surface simplification algorithms. Carnegie-Mellon University Pittsburgh PA School of Computer Science.

Hoobler, N., Humphreys, G., and Agrawala, M. (2004). Visualizing competitive behaviors in multi-user virtual environments. In *Proceedings of the Conference on Visualization 2004*, pages 163–170.

Huang, D., Tory, M., Aseniero, B. A., Bartram, L., Bateman, S., Carpendale, S., Tang, A., and Woodbury, R. (2015). Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics,*, 21(3):420–433.

Hurter, C., Tissoires, B., and Conversy, S. (2009). FromDaDy: Spreading aircraft trajectories across views to support iterative queries. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1017–1024.

ISO (1998). *ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability*.

Ivanov, Y., Wren, C., Sorokin, A., and Kaur, I. (2007). Visualizing the history of living spaces. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1153–1160.

John, M. S., Cowen, M. B., Smallman, H. S., and Oonk, H. M. (2001). The use of 2D and 3D displays for shape-understanding versus relative-position tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1):79–98.

Kapler, T. and Wright, W. (2004). Geotime information visualization. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 25–32.

Kaufman, M. M. (2004). Using spatial-temporal primitives to improve geographic skills for pre-service teachers. *Journal of Geography*, 103(4):171–181.

KDE (2015). Node webkit. Retrived from: https://nwjs.io/. Last access: 19/07/2017.

Keim, D., Andrienko, G., Fekete, J.-D., Gorg, C., Kohlhammer, J., and Melançon, G. (2008). Visual analytics: Definition, process, and challenges. *Lecture notes in computer science*, 4950:154–176.

Kessler, F. C. and Slocum, T. A. (2011). Analysis of thematic maps published in two geographical journals in the twentieth century. *Annals of the Association of America Geographers*, 101(2):292–317.

Kjellin, A., Pettersson, L. W., Seipel, S., and Lind, M. (2010a). Different levels of 3D: an evaluation of visualized discrete spatiotemporal data in space-time cubes. *Information Visualization*, 9(2):152–164.

Kjellin, A., Pettersson, L. W., Seipel, S., and Lind, M. (2010b). Evaluating 2D and 3D visualizations of spatiotemporal information. *ACM Transactions on Applied Perception*, 7(3):19:1–19:23.

Knapp, L. (1995). A task analysis approach to the visualization of geographic data. *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, 83:355–371.

Koua, E. L. and Kraak, M.-J. (2004). A usability framework for the design and evaluation of an exploratory geovisualization environment. In *Proceedings of the 8th International Conference on Information Visualisation*, pages 153–158.

Koua, E. L., MacEachren, A. M., and Kraak, M.-J. (2006). Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science*, 20(24):425–448.

Koussoulakou, A. and Kraak, M. J. (1992). Spatio-temporal maps and cartographic communication. *The Cartographic Journal*, 29(2):101–108.

Kraak, M. J. (2003). The space-time cube revisited from a geovisualization perspective. In *Proceedings of the 21st International Cartographic Conference*, pages 1988–1995.

Kraak, M.-J. (2008). Geovisualization and time - new opportunities for the space-time cube. In *Geographic Visualization Concepts, Tools and Applications*, chapter 15, pages 294–306. Wiley, 1st edition.

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. (2007). The trade-offs with space time cube representation of spatiotemporal patterns. *Computing Research Repository*, vol. abs/0707.1618.

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.

Kveladze, I., Kraak, M.-J., and Van Elzakker, C. P. J. M. (2015). The space-time cube as part of a geovisual analytics environment to support the understanding of movement data. *International Journal of Geographical Information Science*, 29(11):2001–2016.

Lee, W.-C. and Krumm, J. (2011). Trajectory preprocessing. In *Computing with Spatial Trajectories*, chapter 1, pages 3–33. Springer.

Li, Q., Xu, P., Chan, Y. Y., Wang, Y., Wang, Z., Qu, H., and Ma, X. (2017). A Visual Analytics Approach for Understanding Reasons behind Snowballing and Comeback in MOBA Games. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):211–220.

Liu, S. and Lai, S. (2009). Schematic visualization of object trajectories across multiple cameras for indoor surveillances. In *Proceedings of the 5th International Conference on Image and Graphics.*, pages 406–411.

Lobben, A. (2008). Influence of data properties on animated maps. *Annals of the Association of American Geographers*, 98(3):583–603.

LoLKing (2017). LoLKing. Retrived from http://www.lolking.net/. last access: 19/07/2017.

Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. (2005). *Geographical Information Systems and Science*. John Wiley & Sons, Ltd, 2nd edition.

Lorenzo, G. D., Sbodio, M. L., Lopez, V., and Lloyd, R. (2013). Exsed: an intelligent tool for exploration of social events dynamics from augmented trajectories. In *Proceedings of the 14th International Conference on Mobile Data Management*, pages 323–330.

Lucey, P., Bialkowski, A., Monfort, M., Carr, P., and Matthews, I. (2014). Quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. In *Proceedings of the 8th Annual MIT Sloan Sports Analytics Conference*, pages 1–9.

MacEachren, A. M. (1995). *How Maps Work: Representation, Visualization, and Design*. Guilford Press, 1st edition.

MachEachren, A. M., Coscoe, F. P., Haug, D., and Pickle, L. (1998). Geographic visualization: Designing manipulable maps for exploring temporally varying georeferenced statistics. In *Proceedings of the 1998 IEEE Symposium on Information Visualization*, pages 87–94.

MachEachren, A. M., Wachowicz, M., Edsall, R., Haug, D., and Masters, R. (1999). Constructing knowledge from multivariate spatiotemporal data: integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographical Information Science*, 13(4):311–334.

Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141.

Miller, J. L. and Crowcroft, J. (2009). Avatar movement in world of warcraft battlegrounds. In *Proceedings of the 8th Annual Workshop on Network and Systems Support for Games*, pages 1:1–1:6.

Naaman, M. (2011). Geographic information from georeferenced social media data. *SIGSPATIAL Special*, 3(2):54–61.

Neumann, A. (2005). Navigation in space, time and topic, interdependencies of spatial, temporal and thematic navigation in 2D interactive maps. In *Proceedings of the 22nd International Cartographic Association Conference*, pages 1–16.

Norman, D. A. (2002). *The Design of Everyday Things*. Basic Books, Inc., 1st edition.

NYC TLC (2017). NYC Taxi and Limousine Commission - TLC Statistics. Retrived from http://www.nyc.gov/html/tlc/html/about/statistics.shtml. Last access: 11/11/2016.

Ogao, P. J. and Kraak, M.-J. (2002). Defining visualization operations for temporal cartographic animation design. *International Journal of Applied Earth Observation and Geoinformation,*, 4(1):23–31.

OPGG (2017). OPGG. Retrived from https://euw.op.gg/. last access: 19/07/2017.

Ostermann, F. O. (2010). Digital representation of park use and visual analysis of visitor activities. *Computers, Environment and Urban Systems*, 34(6):452–464.

Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M. L., Gkoulalas-divanis, A., Macedo, J., Pelekis, N., Theodoridis, Y., and Yan, Z. (2012). Semantic trajectories modeling and analysis. *ACM Computing Surveys*, pages 42:1–42:32.

Patton, D. K. and Cammack, R. G. (1996). An examination of the effects of task type and map complexity on sequenced and static choropleth maps. In *Cartographic Design: Theoretical and Practical Perspectives*, pages 237–252. Wiley-Blackwell, 1st edition.

Perin, C., Vuillemot, R., and Fekete, J. D. (2013). Soccerstories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515.

Perše, M., Kristan, M., Kovačič, S., Vučkovič, G., and Perš, J. (2009). A trajectory-based analysis of coordinated team activity in a basketball game. *Computer Vision and Image Understanding*, 113(5):612–621.

Peuquet, D. J. (1994). It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Association of American Geographers*, 84(3):441–461.

Phan, D., Xiao, L., Yeh, R., Hanrahan, P., and Winograd, T. (2005). Flow map layout. In *Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 29–34.

Pileggi, H., Stolper, C. D., Boyle, J. M., and Stasko, J. T. (2012). Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828.

Pingali, G., Opalach, A., Jean, Y., and Carlbom, I. (2001). Visualization of sports using motion trajectories: providing insights into performance, style, and strategy. In *Proceedings of the Conference on Visualization*, pages 75–82.

Pousman, Z., Stasko, J., and Mateas, M. (2007). Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152.

Pu, J., Liu, S., Ding, Y., Qu, H., and Ni, L. (2013). T-watcher: A new visual analytic system for effective traffic surveillance. In *Proceedings of the 14th International Conference on Mobile Data Management*, pages 127–136.

Qian, L., Wachowicz, M., Peuquet, D., and MacEachren, A. M. (1997). Delineating operations for visualization and analysis of space-time data in gis. In *Proceedings of GIS/LIS '97 Annual Conference and Exposition*, pages 872–877.

Reichenbacher, T. (2007). The concept of relevance in mobile maps. In *Location Based Services and Telecartography*, pages 231–246. Springer, 1st edition.

Renso, C., Baglioni, M., Macedo, J. A. F., Trasarti, R., and Wachowicz, M. (2012). How you move reveals who you are: understanding human behavior by analyzing trajectory data. *Knowledge and Information Systems*, pages 1–32.

Riot Games (2009). League of legends. Retrived from http://euw.leagueoflegends.com/. Last access: 9/3/2017.

Riot Games (2014). League players reach new heights in 2014. Retrived from http://www.riotgames.com/articles/20140711/1322/league-players-reach-newheights-2014. Last access: 11/11/2016.

Riot Games (2016a). League of legends boards. Retrived from https://boards.euw.leagueoflegends.com/en/. Last access: 9/3/2017.

Riot Games (2016b). League of legends reddit. Retrived from https://www.reddit.com/r/leagueoflegends/. Last access: 9/3/2017.

Riot Games (2016c). Riot developer - league of legends api. Retrived from https://developer.riotgames.com/. Last access: 9/3/2017.

Riveiro, M., Falkman, G., and Ziemke, T. (2008). Visual analytics for the detection of anomalous maritime behavior. In *Proceedings of the 12th International Conference Information Visualisation*, pages 273–279.

Robertson, G., Cameron, K., Czerwinski, M., and Robbins, D. (2002). Polyarchy visualization: Visualizing multiple intersecting hierarchies. In *Proceedings of the 20th SIGCHI Conference on Human Factors in Computing Systems*, pages 423–430.

Robertson, G., Fernandez, R., Fisher, D., Lee, B., and Stasko, J. (2008). Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics,*, 14(6):1325–1332.

Robinson, A. C. (2011). Highlighting in geovisualization. *Cartography and Geographic Information Science*, 38(4):373–383.

Rockstar Games (2016). Rockstar games social club. Retrived from https://socialclub.rockstargames.com/. Last access: 9/3/2017.

Roth, R. E. (2012). Cartographic interaction primitives: Framework and synthesis. *The Cartographic Journal*, 49(4):376–395.

Roth, R. E. (2013). Cartographic interaction: What we know and what we need to know. *Journal of Spatial Information Science*, (6):59–115.

Scheepens, R., Willems, N., van de Wetering, H., and van Wijk, J. (2012). Interactive density maps for moving objects. *IEEE Computer Graphics and Applications*, 32(1):56 –66.

Scheepens, R., Willems, N., van de Wetering, H., and van Wijk, J. J. (2011). Interactive visualization of multivariate trajectory data with density maps. In *Proceedings of the IEEE Pacific Visualization Symposium*, pages 147–154.

Schuessler, N. and Axhausen, K. W. (2009). Processing raw data from global positioning systems without additional information. *Transportation Research Record: Journal of the Transportation Research Board*, 2105(1):28–36.

Seipel, S. (2013). Evaluating 2D and 3D geovisualisations for basic spatial assessment. *Behaviour & Information Technology*, 32(8):845–858.

Seipel, S. and Carvalho, L. (2012). Solving combined geospatial tasks using 2D and 3D bar charts. In *Proceedings of 16th International Conference on Information Visualisation*, pages 157–163.

Shao, L., Sacha, D., Neldner, B., Stein, M., and Schreck, T. (2016). Visual-interactive search for soccer trajectories to identify interesting game situations. *Electronic Imaging*, 2016(1):1–10.

Slocum, T., Sluter, R., Kessler, F., and Yoder, S. (2004). A Qualitative Evaluation of MapTime, A Program For Exploring Spatiotemporal Point Data. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 39(3):43–68.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., and Vangenot, C. (2008). A conceptual view on trajectories. *Data and Knowledge Engineering*, 65(1):126–146.

Sports Tracking Technologies Ltd. (2004). Sports tracker. Retrived from http://www.sports-tracker.com/. Last access: 19/07/2017.

Tani, T., Huang, H.-H., and Kawagoe, K. (2014). Sports play visualization system using trajectory mining method. *Procedia Technology*, 18:100 – 103.

Thakur, S. and Hanson, A. (2010). A 3D visualization of multiple time series on maps. In *Proceedings of the 14th International Conference Information Visualisation*, pages 336–343.

Thudt, A., Baur, D., and Carpendale, S. (2013). Visits: A spatiotemporal visualization of location histories. In *Proceedings of the Eurographics Conference on Visualization*, pages 1–5.

Tobler, W. (1987). Experiments in migration mapping by computer. *The American Cartographer*, 14(2):155–163.

Tominski, C. and Schulz, H.-J. (2012). The great wall of space-time. *Vision, Modeling, and Visualization*, pages 199–206.

Tominski, C., Schulze-Wollgast, P., and Schumann, H. (2005). 3D information visualization for time dependent data on maps. In *Proceedings of the 9th International Conference on Information Visualisation*, pages 175–181.

Tominski, C., Schumann, H., Andrienko, G., and Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2565–2574.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, 1st edition.

Tversky, B., Morrison, J. B., and Betancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57(4):247 – 262.

Valiati, E. R. A., Pimenta, M. S., and Freitas, C. M. D. S. (2006). A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. In *Proceedings of the 2006 AVI Workshop on Beyond time and errors: novel evaluation methods for information visualization*, pages 1–6.

Vasiliev, I. (1997). Mapping Time. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 34(2):1–51.

Vecchi, T. and Girelli, L. (1998). Gender differences in visuo-spatial processing: The importance of distinguishing between passive storage and active manipulation. *Acta psychologica*, 99(1):1–16.

Vieira, A. R. (2016). Visualização de trajectórias humanas em dispositivos móveis. Master thesis, Faculdade de Ciências Universidade de Lisboa, Departamento de Informática.

Walker, A. (2016). More people watched league of legends than the nba finals. Retrived from https://www.kotaku.com.au/2016/06/more-people-watched-league-of-legends-than-the-nba-finals/. Last access: 11/11/2016.

Wallner, G. and Kriglstein, S. (2013). Visualization-based analysis of gameplay data - A review of literature. *Entertainment Computing*, 4(3):143–155.

Wang, H., Zou, H., Yue, Y., and Li, Q. (2009). Visualizing hot spot analysis result based on mashup. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, pages 45–48.

Ware, C., Arsenault, R., Plumlee, M., and Wiley, D. (2006). Visualizing the underwater behavior of humpback whales. *IEEE Computer Graphics and Applications*, 26(4):14–18.

Wehrend, S. and Lewis, C. (1990). A problem-oriented classification of visualization techniques. In *Proceedings of the 1st conference on Visualization*, pages 139–143.

Wei, X., Sha, L., Lucey, P., Morgan, S., and Sridharan, S. (2013). Large-scale analysis of formations in soccer. In *2013 International Conference on Digital Image Computing: Techniques and Applications*, pages 1–8.

Willems, N., van de Wetering, H., and van Wijk, J. J. (2011). Evaluation of the visibility of vessel movement features in trajectory visualizations. *Computer Graphics Forum*, 30(3):801–810.

Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. (2011). The align rank transform for nonparametric factorial analysis using only ANOVA procedures. In *Proceedings of the 29th SIGCHI Conference on Human Factors in Computing Systems*, pages 143–146.

Wolfe, J. M. and Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6):495–501.

Xie, K., Deng, K., and Zhou, X. (2009). From trajectories to activities: a spatio-temporal join approach. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, pages 25–32.

Yi, J. S., Kang, Y. a., Stasko, J., and Jacko, J. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231.

Yuan, J., Zheng, Y., and Xie, X. (2012). Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, pages 186–194.

Zheng, Y., Li, Q., Chen, Y., Xie, X., and Ma, W.-Y. (2008). Understanding mobility based on GPS data. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, pages 312–321.

Zheng, Y., Xie, X., and Ma, W.-Y. (2010). Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data Engineering Bulletin*, 33(2):32–39.

Zheng, Y., Zhang, L., Xie, X., and Ma, W.-Y. (2009). Mining interesting locations and travel sequences from gps trajectories. In *Proceedings of the 18th international conference on World Wide Web*, pages 791–800.

Zheng, Y. and Zhou, X., editors (2011). *Computing with Spatial Trajectories*. Springer, 1st edition.

Zhou, M. X. and Feiner, S. K. (1998). Visual task characterization for automated visual discourse synthesis. In *Proceedings of the 16th SIGCHI Conference on Human Factors in Computing Systems*, pages 392–399.