UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

JORGE ALBERTO WAGNER FILHO

Evaluating Immersive Approaches to Multidimensional Information Visualization

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Advisor: Prof. Dr. Luciana Porcher Nedel
Co-advisor: Prof. Dr. Carla Dal Sasso Freitas

Porto Alegre
March 2018

CIP — CATALOGING-IN-PUBLICATION

Wagner Filho, Jorge Alberto

Evaluating Immersive Approaches to Multidimensional Information Visualization / Jorge Alberto Wagner Filho. – Porto Alegre: PPGC da UFRGS, 2018.

91 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2018. Advisor: Luciana Porcher Nedel; Co-advisor: Carla Dal Sasso Freitas.

1. Immersive visualization. 2. Abstract information visualization. 3. Dimensionality reduction. 4. 3D scatterplots. 5. Virtual reality. I. Nedel, Luciana Porcher. II. Freitas, Carla Dal Sasso. III. Título.

"We can only see a short distance ahead,

but we can see plenty there that needs to be done."

— ALAN TURING

ACKNOWLEDGEMENTS

Foremost, I thank my advisor, Prof. Luciana Nedel, and co-advisor, Prof. Carla Freitas, for their dedicated support and guidance throughout this project, and also for reminding me that it is possible to keep your spirits up despite the sometimes overwhelming nature of life in academia.

Importantly, I thank the Federal University of Rio Grande do Sul, its Institute of Informatics and Graduate Program in Computer Science for their continued efforts in promoting international level research despite all adversities.

I am also grateful to the members of my dissertation committee, Prof. Anderson Maciel, Prof. Benjamin Bach and Prof. João Comba, for their participation and feedback.

I acknowledge the funding from the Brazilian National Council for Scientific and Technological Development (CNPq), without which this project would not have been possible.

Finally, I would like to thank all the 74 participants who kindly volunteered their times to contribute in the different phases of this work.

ABSTRACT

The use of novel displays and interaction resources to support immersive data visualization and improve the analytical reasoning is a research trend in Information Visualization. In this work, we evaluate the use of HMD-based environments for the exploration of multidimensional data, represented in 3D scatterplots as a result of dimensionality reduction. We present a new modelling for the evaluation problem in such a context, accounting for the two factors whose interplay determine the impact on the overall task performance: the difference in errors introduced by performing dimensionality reduction to 2D or 3D, and the difference in human perception errors under different visualization conditions. This two-step framework offers a simple approach to estimate the benefits of using an immersive 3D setup for a particular dataset. As use case, the dimensionality reduction errors for a series of roll calls datasets when using two or three dimensions are evaluated through an empirical task-based approach. The perception error and overall task performance are assessed through controlled comparative user studies. When comparing desktop-based (2D and 3D) with an HMD-based (3D) visualization, initial results indicated that perception errors were low and similar in all approaches, resulting in overall performance benefits in both 3D techniques. The immersive condition, however, was found to require less effort to find information and less navigation, besides providing much larger subjective perception of accuracy and engagement. Nonetheless, the use of flying navigation resulted in inefficient times and frequent user discomfort.

In a second moment, we implemented and evaluated an alternative data exploration approach where the user remains seated and viewpoint change is only realisable through physical movements. All manipulation is done directly by natural mid-air gestures, with the data being rendered at arm's reach. The virtual reproduction of an exact copy of the analyst's desk aims to increase immersion and enable tangible interaction with controls and two dimensional associated information. A second user study was carried out comparing this scenario to a desktop-based equivalent, exploring a set of 9 representative perception and interaction tasks based on previous literature. We demonstrate that our prototype setup, named VirtualDesk, presents excellent results regarding user comfort and immersion, and performs equally or better in all analytical tasks, while adding minimal or no time overhead and amplifying data exploration.

Keywords: Immersive visualization. abstract information visualization. dimensionality

reduction. 3D scatterplots. virtual reality.

Avaliando Abordagens Imersivas para
Visualização de Informações Multidimensionais

RESUMO

O uso de novos recursos de *display* e interação para suportar a visualização imersiva de dados e incrementar o raciocínio analítico é uma tendência de pesquisa em Visualização de Informações. Neste trabalho, avaliamos o uso de ambientes baseados em HMD para a exploração de dados multidimensionais, representados em *scatterplots* 3D como resultado de redução de dimensionalidade.

Nós apresentamos uma nova modelagem para o problema de avaliação neste contexto, levando em conta os dois fatores cuja interação determina o impacto no desempenho total nas tarefas: a diferença nos erros introduzidos ao se realizar redução de dimensionalidade para 2D ou 3D, e a diferença nos erros de percepção humana sob diferentes condições de visualização. Este *framework* em duas etapas oferece uma abordagem simples para estimar os benefícios de se utilizar um *setup* 3D imersivo para um dado conjunto de dados. Como caso de uso, os erros de redução de dimensionalidade para uma série de conjuntos de dados de votações na Câmara dos Deputados, ao se utilizar duas ou três dimensões, são avaliados por meio de uma abordagem empírica baseada em tarefas. O erro de percepção e o desempenho geral de tarefa, por sua vez, são avaliados através de estudos controlados comparativos com usuários. Comparando-se visualizações baseadas em *desktop* (2D e 3D) e em HMD (3D), resultados iniciais indicaram que os erros de percepção foram baixos e similares em todas abordagens, resultando em benefícios para o desempenho geral em ambas técnicas 3D. A condição imersiva, no entanto, demonstrou requerer menor esforço para encontrar as informações e menos navegação, além de prover percepções subjetivas de precisão e engajamento muito maiores. Todavia, o uso de navegação por voo livre resultou em tempos ineficientes e frequente desconforto nos usuários.

Em um segundo momento, implementamos e avaliamos uma abordagem alternativa de exploração de dados, onde o usuário permanece sentado e mudanças no ponto de vista só são possíveis por meio de movimentos físicos. Toda a manipulação é realizada diretamente por gestos aéreos naturais, com os dados sendo renderizados ao alcance dos braços. A reprodução virtual de uma cópia exata da mesa de trabalho do analista visa aumentar a imersão e possibilitar a interação tangível com controles e informações bidimensionais associadas. Um segundo estudo com usuários foi conduzido em comparação a uma versão equivalente baseada em *desktop*, explorando um conjunto de 9 tarefas representativas de percepção e interação, baseadas em literatura prévia. Nós demonstramos que o nosso protótipo, chamado VirtualDesk, apresentou resultados excelentes em relação a conforto e imersão, e desempenho equivalente ou superior em todas tarefas analíticas, enquanto adicionando pouco ou nenhum tempo extra e ampliando a exploração dos dados.

Palavras-chave: visualização imersiva, visualização de informações abstratas, redução de dimensionalidade, scatterplots 3D, realidade virtual.

LIST OF ABBREVIATIONS AND ACRONYMS

HMD	Head-mounted Display
VR	Virtual Reality
AR	Augmented Reality
VE	Virtual Environment
2D	Two dimensions
3D	Three dimensions
3DUI	3D User Interfaces
TUI	Tangible User Interface
CAVE	Cave Automatic Virtual Environment
HCI	Human-Computer Interaction
FOV	Field of view
DR	Dimensionality Reduction
PCA	Principal Component Analysis
MDS	Multidimensional Scaling
t-SNE	t-Distributed Stochastic Neighbor Embedding
SUS	System Usability Scale
SSQ	Simulator Sickness Questionnaire
TLX	Task Load Index
IPQ	Igroup Presence Questionnaire
SD	Standard deviation

LIST OF FIGURES

CONTENTS

1 INTRODUCTION

1.1 Motivation

Three dimensional (3D) representations are known to offer advantages under many circumstances in the context of data visualization. In fact, for inherently spatial data, these representations contribute to a quicker construction of the mental model (MUNZNER, 2014). This is the reason why since the 80's scientific applications have been employing 3D visualizations for conveying information, through mappings from data values, which are associated to spatial positions, to visual attributes like shapes, colours, textures, among others. (WARD; GRINSTEIN; KEIM, 2015).

For abstract information, 3D representations have also been demonstrated to be useful, allowing clearer spatial separation in large graphs (WARE; MITCHELL, 2008), detection of trivariate patterns in scatterplots (SHOVMAN et al., 2015) and more accurate projection of multidimensional data (GRACIA et al., 2016; POCO et al., 2011). Nonetheless, their use has also been long controversial, mostly because they are hindered by the occurrence of known perceptual issues: perspective distortion, foreshortening and occlusion make data exploration cumbersome and error-prone. Moreover, there is a relevant mismatch between 3D visualizations and conventional two-dimensional (2D) interaction devices, such as the mouse.

Along the years, following consecutive breakthroughs in Virtual Reality (VR) research, the visualization community has progressively explored the use of immersive approaches, which combine stereoscopic displays with natural interaction as alternatives to potentially change this scenario. Since the early works in immersive visualization for scientific applications (BRYSON; LEVIT, 1992), there have been multiple favourable results. Volume visualization (COFFEY et al., 2011), analysis of materials (DROUHARD et al., 2015), and neuron tracing (USHER et al., 2017) are only a few examples, demonstrating that for scientific visualization, immersive applications have already achieved a somewhat consolidated usage (GARCÍA-HERNÁNDEZ et al., 2016).

However, in the case of abstract information visualization, research on immersive techniques is still limited, but advantages have already been demonstrated (GARCÍA-HERNÁNDEZ et al., 2016; DONALEK et al., 2014), for example, in graph visualization (HALPIN et al., 2008; KWON et al., 2016; CORDEIL et al., 2017c).

When providing 3D visualizations of a data set within a virtual reality environment, one usually has to provide features for supporting data analysis, and this has been recently called as Immersive Analytics (CHANDLER et al., 2015). Immersive Analytics is a new research field, which combines techniques from Information Visualization, Visual Analytics, Human Computer Interaction and Virtual Reality. Due to this plethora of possible features, evaluation methods and design guidelines are even more needed.

1.2 Research Question and Approach

In this work, we expand this discussion, focusing on a specific abstract information representation: 3D dimensionally-reduced (DR) data scatterplots. Since this particular category of scatterplots, which is commonly applied for multidimensional data visualization, is always analysed in terms of the distances between points, we hypothesize it could benefit from stereoscopic displays, egocentric points of view and more natural user interfaces, characteristics that are inherent to immersive setups. Our research question is studying the use of HMD-based environments for the interactive visual analysis of 3D scatterplots, in comparison to desktop-based alternatives, which correspond to the currently used solutions.

The use of 3D scatterplots has been controversial since long before the first uses of immersive setups, with related studies dating back to the 1970s (FISHERKELLER; FRIEDMAN; TUKEY, 1974). Their application for the representation of dimensionally-reduced data is also often discussed in the literature. Adding an extra component could potentially reduce information loss in the process, but results from studies on quantifying visual analysis gains have been contradictory (POCO et al., 2011; SEDLMAIR; MUNZNER; TORY, 2013; GRACIA et al., 2016). Few authors, however, have investigated how immersion and stereopsis may impact on these issues. Moreover, most of them have only provided preliminary results, based on technologies which have advanced enormously over the past few years (ARMS; COOK; CRUZ-NEIRA, 1999; RAJA et al., 2004). Therefore, we contribute to the investigation of this problem by providing updated results in terms of a new interactive visualization metaphor and evaluation methodology.

To this end, we introduce a new model of the evaluation problem regarding 3D scatterplots visualization, accounting for the two different factors that influence in the final task performance outcome. We argue that the performance gains attained in a task are not just a function of the difference in perceptual accuracy presented by users under different visualization conditions, but rather of its interplay with the difference in errors introduced by reducing a particular dataset to two or three dimensions. This so called *DR error* component is dataset-dependent, depending on the particular complexity of the data structure. This means that, for a given dataset to benefit from a 3D visualization condition, its content must be indeed better mapped to 3D. Moreover, the user must be able to perceive this added information appropriately, what can be challenging given the previously discussed issues associated with 3D representations.

Based on this model, we propose an evaluation framework that aims to separately assess each of these variables. The *maximum potential performance* in 2D or 3D for our datasets is estimated through a task-based empirical approach. The *perception* and *overall task errors*, on the other hand, are assessed through a user study, comparing three alternative visualization conditions: desktop-based 2D, desktop-based 3D and HMD-based immersive 3D. Participants are subjected to a set of analytical tasks for two selected datasets, one with previously detected promised improvements in 3D, and another one that, in theory, allows for similar performance in all representations.

In this first study, we observed that perception errors were similarly low both in desktop-based and HMD-based conditions. Task performance was therefore improved with the addition of the third dimension regardless of immersion, when the data enabled so. Nonetheless, the HMD-based condition required smaller effort to find information and less navigation, besides offering a much larger subjective perception of accuracy and engagement. The use of flying navigation, however, resulted in inefficient times and frequent user discomfort – around 40% of the participants reported significant levels of simulator sickness.

As a consequence of the findings from the first study, we set out to propose, implement and evaluate an alternative data exploration approach that circumvented the main observed limitations. In this novel approach, named VirtualDesk, viewpoint change is only realisable through head movements. All data manipulation is done directly by mid-air gestures, with the data being rendered at arm's reach. To increase immersion and enable the display of two-dimensional associated views and interaction with tangible controls, we also build upon previous work and reproduce in the virtual environment an exact copy of the analyst's desk. Important data exploration resources are provided, including coordinated views, combinable filters and annotation tools.

Within this second study, our main hypothesis was that our immersive setup would enhance user perception and decrease workload, while remaining time-efficient and not inducing cybersickness. A controlled comparative user study was carried out against a desktop-based equivalent 3D environment, implemented with the same functionalities and following typical mouse and keyboard interaction approaches.

The results confirmed our main intuitions: the VirtualDesk metaphor presented excellent results regarding user comfort and immersion, and performed equally or better than the desktop-based solution for all proposed tasks, while adding minimal time overhead and amplifying data exploration and the subjective user perception of accuracy and engagement.

1.3 Summary of Objectives and Contributions

Considering the context of immersive analytics applications based on 3D scatterplots obtained from dimensionality reduction techniques, we can state our main goal as the investigation of immersive visualization techniques that minimize errors in user perception, decrease user's workload and remain time-efficient. As a secondary goal, we focused on defining a task-based evaluation framework for these applications.

In summary, our main contributions are:

1. an improved modelling of the problem of evaluating immersive visualizations of 3D dimensionally-reduced data scatterplots
2. a task-based evaluation framework
3. the identification of a complete set of relevant tasks for evaluation purposes
4. a novel immersive data exploration metaphor (the VirtualDesk)
5. baseline results to be used in future work

1.4 Structure of the Dissertation

This work is organized as follows. Firstly, related work in the relevant areas is reviewed (Chapter 2), and our proposed task-based evaluation framework is introduced (Chapter 3). Our first user study is then presented and discussed in Chapter 4.

These initial results lead to the proposal of an alternative immersive approach, presented in Chapter 5, which is evaluated through a second user study reported in Chapter 6. Finally, Chapter 7 summarizes our conclusions and points directions for future works.

2 RELATED WORK

In this chapter, we briefly revise relevant work in the recent field of Immersive Analytics (Section 2.1). Moreover, since our study is particularly focused on 3D representations of abstract information obtained through dimensionality reduction (Sections 2.2 and 2.3), and their exploration in immersive environments (Section 2.4), we introduce and review the state of the art on these two subjects also.

2.1 Immersive Analytics

Immersive Analytics is a growing research area in the visualization community, concerned with applying novel display and interaction resources in combination to support the analytical reasoning, and to improve the performance of typical tasks. It was recently defined by Chandler et al. (2015), who focused on bringing attention to usability and high-level design questions in the creation of efficient interfaces for data analysis in immersive environments, considering that the advances recently seen in human-computer interaction (HCI) technologies – such as touchable surfaces, immersive VR/AR environments and tracking devices – have not been fully reflected in this context. They argued that, despite extant limitations, these technologies should already be adequate enough to allow the exploration of the design space for immersive data exploration. Immersive setups beyond the classical desktop could be used either by experts, analysts, decision makers or the general public, with great support for distributed collaboration (see Figure 2.1). Other possible research directions include the use of touch screens, desk interfaces and large displays (tiled displays or projectors), coupled with spatial tracking, to support multiple collocated users and enable seamless collaboration. Hybrid 2D/3D visualizations can also be used to allow the presentation of physical components in their natural 3D form and abstract data in 2D. Finally, seven questions for future research have been defined:

1. What collaboration paradigms are potentially enabled with the new interaction modalities and how to evaluate them?
2. Would it be possible, with new technologies, to support more holistic data visualizations, incorporating both 3D spatial information and abstract data?
3. What questions do technologies such as AR raise with relation to data analysis? It

Figure 2.1: A concept sketch for collaboration in Immersive Analytics. Peter and Sarah are joined by a remote user, John, to discuss complex data through visual representations overlaid on the environment.

Source: Chandler et al. (2015)

is discussed, for example, that immersive approaches could shift the natural model from only presenting details on demand (what is known as the Schneiderman's mantra (SHNEIDERMAN, 1996)) to always presenting detailed object information and only presenting the context on demand.

4. What are the interface 'tricks' and affordances that change the user perception from an allocentric data view to a more immersive and egocentric one?

5. What lessons can be taken from previous 3D visualization research?

6. Which areas are more fertile for immersive analysis and which requirements are domain-specific or general?

7. How do we develop generic platforms to support immersive analysis?

Roberts et al. (2014) also discussed the need to adapt visualization research to novel devices and technologies, aiming to create integrated multi-sensorial environments with natural interaction. In their view, important HCI paradigms to be explored include fluid interactions, the development of a holistic theory for multi-sensorial visualization (how to achieve sensorial integration and how cross-modal interference occurs), incorporation of information in the environment (including the definition of "appropriate surfaces" and privacy implications), immersive interfaces and mixed reality. Three potential visions for future visualization approaches were provided for reflection:

1. a room where any tool or object becomes a visualization interface and can communicate with the others;

2. a mixed reality environment where the user receives information (e.g. geomarkers for the stores with the better prices for a given item); and

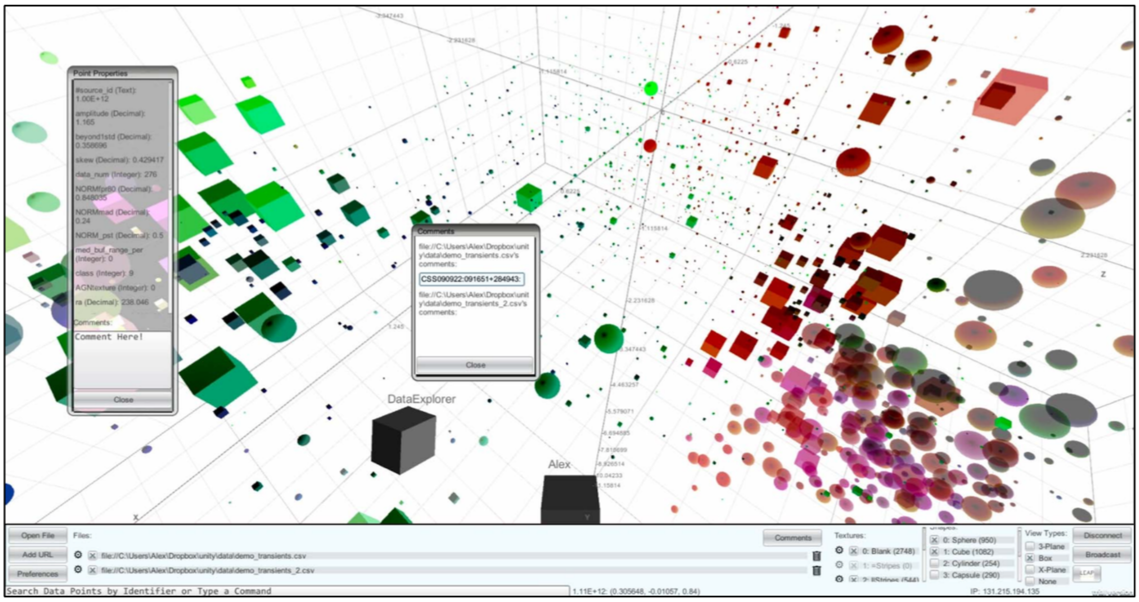
3. a fully-immersive multi-sensorial VR environment where telepresence and collaborative analysis can be performed.

To Nguyen, Marendy and Engelke (2016), multi-modal immersive interfaces have the potential to be broadly used in collaborative visual analytics, allowing the simultaneous investigation of large volumes of data and the visualization of complex data structures in proximity in space. They also believe that future data analysis and decision making infrastructure will be distributed and collaborative, enabling more efficient communication, interaction and coordination between users. Guidelines are discussed for the design of a visual analysis framework to support collaboration (either synchronous or asynchronous and remote or collocated), composed by different visualization approaches for different user interfaces, including AR, VR and desktop, and interaction tools to ease the collaboration between different user levels, from novices to experts. Key features would include the separation between individual and common work areas and between synchronous and asynchronous session environments, techniques for coordination and communication (e.g. user embodiment, motivation and reasoning logging, insight sharing), integration of machine learning and automated analysis components and visualization of different levels of details depending on each user's access rights and output devices capabilities.

2.1.1 Head Mounted Display Based Environments

Early works in the area of immersive visualization explored the use of small spaces surrounded by retro-projected walls, the so-called CAVE environments, which, combined with head-tracking, provided an immersive experience. This kind of structure is still intensively used nowadays (MANJREKAR et al., 2014; KUHLEN; HENTSCHEL, 2014). Here, we are concerned with HMD-based environments, considering that the current technology provides adequate immersive capabilities with much more accessible requirements than CAVEs, both in terms of cost and space, and its exploration in the literature is still incipient.

Figure 2.2: The *iViz* immersive data visualizer maps multidimensional datasets to eight different attributes of points in a 3D scatterplot, such as position, colour and transparency.



Source: Donalek et al. (2014)

Donalek et al. (2014) presented a very interesting early work in this direction. They implemented *iViz*, a platform for visualization of multidimensional data using an Oculus Rift HMD and a Leap Motion sensor for interaction (Figure 2.2). In their application, up to 8 data dimensions of astronomical observation datasets are mapped to different attributes (the X, Y and Z coordinates, size, colour, texture, shape and transparency) of points in a 3D scatterplot. They argued that the more dimensions one can effectually visualize, the greater the chances of identifying potentially interesting patterns, correlations and outliers. Initially, experiments were implemented using off-the-shelf virtual worlds, such as Second Life, OpemSim and vCaltech. However, limitations related to scripting capabilities and the lack of optimization for massive data rendering lead to the development of their own toolkit, capable of visualizing up to a million points with support to collaborative exploration using independent or shared views. Moran et al. (2015) have also implemented an HMD-based tool for the immersive exploration of geospatial information in VR, expecting larger situational awareness. A dataset of geolocated tweets was projected upon a virtual model of the MIT campus, constructed from LADAR data. Characteristics of each tweet determined its visual representation, and interface components allowed filtering by time intervals, searching and changing the opacity of buildings.

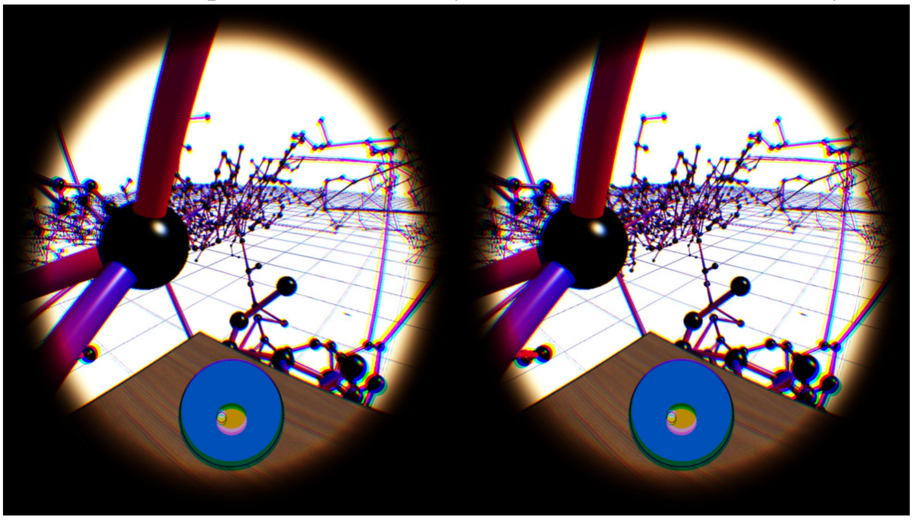
More recently, Cordeil et al. (2017c) presented a comparative study between CAVE-style and HMD-based environments for collaborative analysis of abstract information. A series of differences between both alternatives, such as resolution, presence and freedom of movement, lead the authors to wonder if the latter, much more inexpensive but not inherently designed for collaboration, would still be adequate to this kind of task. Pairs of users were asked to perform tasks related to graph connectivity, including triangles counting and shortest path finding – chosen due to their collaborative potential through divide-and-conquer strategies. Results indicated both conditional presented similar and high accuracies, around 80%, but that users in the HMD condition were 40% faster in the shortest path task, and 30% faster in the triangles one. Possible reasons include the occasional obstruction of the display by the other person and lack of head-tracking for one of the users in the CAVE. No differences were found in the measured shared focus and proportion of oral communication. The capability to see visual representations of the partner's viewing frustum and finger position in the HMD-condition was pointed as crucial for collaboration by the participants.

2.1.2 Immersive Analytics of Abstract Information

A commonly explored 3D representation for abstract data is node-link diagrams representing networks. Halpin et al. (2008) implemented a generic semantic social network visualization software for CAVE-like environments, named *Redgraph*, and, in a user study, observed significant performance improvements for fine-grained questions using the immersive condition. Ware and Mitchell (2008) observed an order of magnitude increase over 2D displays in a path tracing task, using high resolution displays and a mirror stereoscope. Kwon et al. (2016) recently explored different techniques in an HMD-based environment, proposing the use of a new spheric layout that offered performance increase especially for more difficult tasks.

Zielasko et al. (2017), who also explored a use case on graph analysis, presented an interesting discussion about the challenges and opportunities of an immersive analytical scenario named *deskVR*, where the user remains seated in an office chair during the immersive exploration of data. They believe that an immersive solution should be easily integrated to the analyst workplace and workflow in order to be really adopted, and that the transition between real and virtual worlds must be seamless, so that the analyst may combine 2D and 3D environments according to the requirements of each specific task. The issue of cybersickness was addressed with the proposed use of user profiles, which would help to indicate when to limit certain system features, such as the field-of-view, based on his/her personal characteristics and previous experience. Bellgardt et al. (2017)

Figure 2.3: Zielasko et al. proposed an HMD-based setup to immerse the analyst, seated at his/her desk, into a node-link diagram. The issue of cybersickness was addressed with the proposed use of user profiles to restrict system features when necessary.



Source: Zielasko et al. (2017)

also discussed the possibility of a seating immersive scenario, but considered that it would sacrifice the level of immersion and realism, only being useful for short sessions.

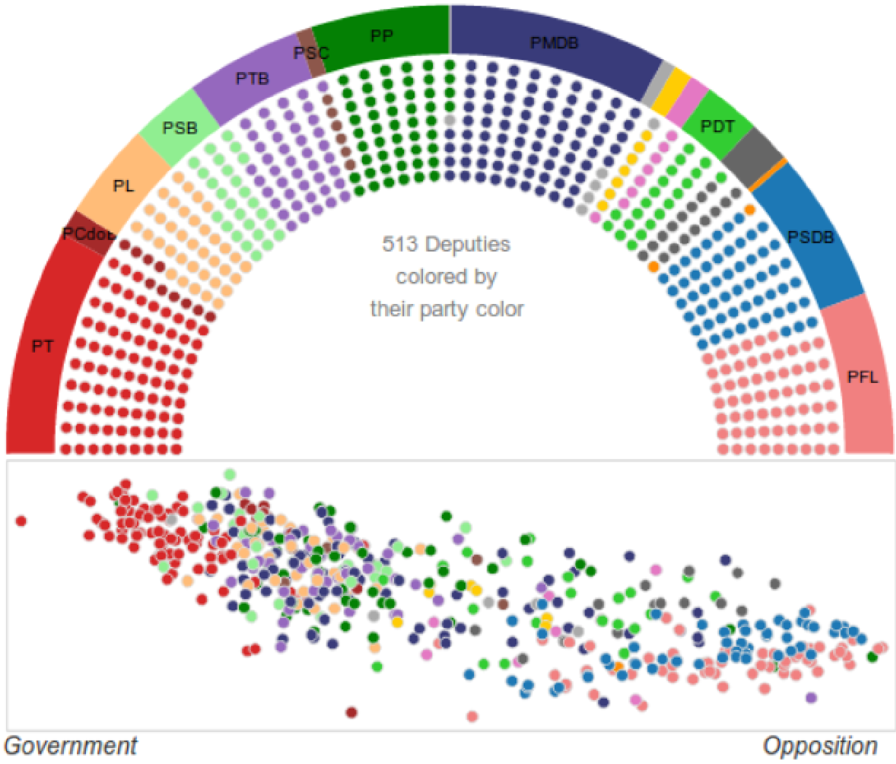
2.2 Dimensionality Reduction and Roll Call Analysis

In order to visualize very high dimensional data, we explore the use of dimensionality reduction (DR) or projection methods. DR methods aim to generate a more compact version of the information, yet maintaining the same characteristics of the original dataset (CUNNINGHAM; GHAHRAMANI, 2015). Some popular examples are Principal Component Analysis (PCA) (HOTELLING, 1933) and Classic Multidimensional Scaling (MDS) (TORGERSON, 1952), linear techniques, which aim to position distant points in the original dataset far apart in the lower dimensional result, and the recent non-linear t-SNE (MAATEN; HINTON, 2008), which aims to maximize the capture of the dataset local structure while also revealing its global structure, such as clusters, and targets specifically data visualization.

Despite presenting several important applications, such as feature selection for algorithmic input, DR techniques are predominantly used for data visualization. In an in-the-wild survey, Sedlmair et al. (2012) observed that 20 out of 27 cases studied used them for generating visualizations with scatterplots, mostly in 2D.

DR has also been extensively applied in the literature to the particular data domain

Figure 2.4: In the *CivisAnalysis* web-based system, Principal Component Analysis is used to automatically compute a political spectrum for the Brazilian Chamber of Deputies through dimensionality reduction.



Source: Borja and Freitas (2015)

we have chosen to target in our analyses – roll call voting data (CARROLL; POOLE, 2014; JAKULIN; BUNTINE, 2004; BRIGADIR et al., 2016). A survey on this specific topic was published by Spirling and McLean (2006). A frequently chosen method is PCA, due to its simplicity and capability of preserving distances to construct a political spectrum. Borja and Freitas (2015) have recently adopted this approach to represent roll call voting data released by the Brazilian Chamber of Deputies, implementing a web-based system called *CivisAnalysis* (Figure 2.4). Non-linear methods, such as the recent t-SNE (MAATEN; HINTON, 2008), may be applied, but prioritize the capture of the dataset local similarities over its global structure – the distances between clusters, for example, may have no meaning (BORJA, 2017; WATTENBERG; VIEGAS; JOHNSON, 2016).

2.3 Evaluation of 3D Scatterplots

The use of 3D scatterplots has been discussed for a long time in the literature. Some precursor works in this field include those of Fisherkeller, Friedman and Tukey

(1974), who explored rotations between 2D plots to generate a parallax effect, and Huber (1987), who analysed the use of interactive 3D scatterplots for arbitrary high dimensional data. Ware (2012), in his thorough discussion on spatial representations and depth cues, argued that the only two cues likely to be useful in a 3D scatterplot are stereoscopic depth and *structure-from-motion* (motion parallax and kinetic depth effect). The first should be more helpful to differentiate depths between near points, while the latter to differentiate more distant ones.

Some authors have specifically investigated the use of monoscopic 3D scatterplots for visualization of dimensionally-reduced data, focusing on multiple categories of analytical tasks, but results have been mixed.

Poco et al. (2011) defended the usefulness of 3D projections to decrease information loss and allow for better cluster discrimination. Acknowledging the associated interaction difficulties, they proposed an exploration framework incorporating features such as predefined optimal 2D views, coordinated 2D and 3D views (using brushing-and-linking), hierarchical 3D cluster exploration and cluster modifications. Furthermore, they also employed selectable enclosing surfaces to assist in data exploration and avoid problems related to 3D cloud points. In their framework, low dimensional representations are always obtained through a generalization of the Least Square Projection DR technique (LSP). This method, which initially applies Multidimensional Scaling (MDS) to a sample of the whole dataset (cluster centroids), and then solves a linear system based on the neighbourhood relationships to position the remaining points, is aimed at handling large datasets. In order to evaluate their 3D exploration framework, three different steps were employed. Initially, two theoretical metrics of neighbourhood preservation were used to quantify the effectiveness of using two or three dimensions. Considering a dataset composed of 2800 scientific papers from 8 areas (clusters) and annotated in terms of the presence of 1200 terms (dimensions), both *neighbourhood hit* (the average between the percentages of each point's neighbours that have been human-assigned to its own class) and *neighbourhood preservation* (the average between the percentages of each point's neighbours, in the low-dimensional space, that belong to the same neighbourhood in the original space) indicated superior performance in 3D. Then, a user study with 12 participants was conducted over a document dataset with 681 objects and 3000 dimensions. Five tasks were assessed: (1) counting clusters, (2) ordering clusters according to density, (3) listing pairwise overlaps between clusters, (4) detecting an object within a cluster using labels, and (5) identifying the closest cluster to a specific point. The results indicated that overall correctness rate went from 64.3% in 2D to 74.4% in 3D, but only task 5 presented statistical significance. Despite a 50% observed time increase in the 3D version (with significance for tasks 1 and 4), it was preferred by all users. Finally, a second user study with the same participants was conducted to compare the so-far used point cloud representation to four different surface-based cluster visualizations. In this study, besides the document dataset, a medical images one with 540 objects and 28 dimensions was also added, and three tasks were assessed: (1) counting clusters, (2) listing cluster overlaps, and (3) identifying the most separable clusters. Only task 2 presented statistical differences in correctness, and only for the document data set. In this case, some of the approaches performed better than the point cloud version, and others worse. Users preferred representations with smaller volumes, and the *non-convex hull* one was considered the best overall.

Sedlmair, Munzner and Tory (2013), on the other hand, argued that user studies are too restrictive in the number of variables examined, and so performed a *data study*, where two annotators evaluated scatterplots with relation to cluster separability, a task similar to those from Poco et al. (2011). To this end, 75 datasets, 4 DR algorithms (PCA, Robust PCA, Glimmer MDS and t-SNE) and 3 visual encodings (2D scatterplots, 3D scatterplots and matrices of 2D scatterplots) were combined, resulting in 816 scatterplots, encompassing 5460 classes. Through this approach, they concluded that the interactive 3D versions never outperformed the 2D scatterplots (individually or in matrices), especially considering the added interaction cost. Although an analysis of subjective preferences showed a recurring preference from one annotator for 2D and from the other for 3D, the inter-coder agreement was calculated as 0.858 using Krippendorff's alpha, claimed to be considered acceptable. They proposed a workflow model, according to which an analyst should always prioritise 2D and experiment with multiple different DR algorithms until finding the one that best fits the dataset. If he/she is restricted to only one specific DR algorithm, then the scatterplot matrix would be helpful in some cases, while the interactive 3D would contribute only under very rare circumstances (observed in artificial entangled datasets synthesised for the evaluation).

More recently, Gracia et al. (2016) performed another user study, targeting different analytical tasks (point classification, distance perception and outlier detection), and also applied metrics from previous literature to reaffirm the advantage of using a third dimension for dimensionality reduction. Forty participants answered tasks through a web visualization of high dimensional DNA micro-array datasets reduced to 2D and 3D using PCA. For the classification and outlier tasks, times were, as expected, considerably lower in 2D, but small precision gains were observed in the 3D version. For distance perception, the error was 80% larger in 2D, while times were similar. Moreover, 55% of the participants considered the 3D scatterplot more useful, and 73% believed making smaller errors in 3D. Nonetheless, 62% felt more comfortable navigating the traditional 2D scenario. In a second moment in their study, 11 different quality criteria, 12 real-world datasets and 12 different DR algorithms were used to quantify the loss of quality in the transition from 3D to 2D in terms of geometric preservation. In an average between all datasets for each algorithm and then between all algorithms for each metric, observed values reached 48.6 of loss in the worst case.

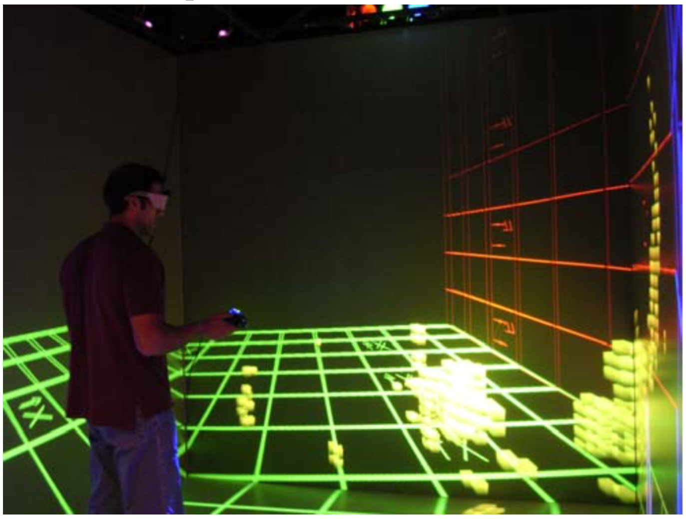
2.4 Immersive Exploration of Scatterplots

Concerning immersive environments, Arms, Cook and Cruz-Neira (1999) performed a comparative evaluation of the visualization of multidimensional data projected to two and three dimensions, achieving better cluster identification results in the virtual environment. However, they explored a CAVE environment, and suffered from heavy technological limitations at the time, especially regarding interaction.

Raja et al. (2004) also explored the application of immersive VR to 3D scatterplots in a CAVE environment (see Figure 2.5), observing favourable results when including large field-of-regard, head-tracking and stereopsis. Six tasks were explored: (1) distance perception in one axis (finding the point with the highest Y value), (2) distance perception in two axes (finding the point with lowest X and lowest Y), (3) determination of trends, (4) identification of clusters larger than 20 points, (5) finding a differently coloured point not visible from the initial position, and (6) identification of the two main outliers. Their user study, however, was very initial, with only four subjects. A later study with 32 users was performed with similar indications, but failed to present statistical significance (RAJA, 2006).

More recently, Etemadpour, Monson and Linsen (2013) continued the previously discussed study by Poco et al. (2011), and compared the performance of tasks in projections of high-dimensional data in a six-sided CAVE environment. The same medical images and scientific papers datasets were used, and the task set was extended to also consider perception of distances between different clusters and identification of outliers. 20 participants performed tasks in two new conditions, which were contrasted to the previous results: one immersive condition using the six sides of the CAVE, and another one using

Figure 2.5: Previous works have explored the application of CAVE-style environments for the exploration of 3D scatterplots and obtained favourable results.



Source: Raja et al. (2004)

only one of its sides. The reduced datasets were again represented both as typical 3D point clouds and as enclosing surfaces. Results indicated better perception of distances between individual objects in the immersive environments, although global analysis tasks, which required the comprehension of the distribution of all points, did not present significant differences. Furthermore, although the one-sided condition never outperformed the six-sided one, in only a few cases it was outperformed by it. The surface-based techniques were noted to profit more from VR than the point clouds. It is also emphasized that the nature of the data played an important role in the results, with significant differences between both datasets observed in several tasks. In later related works, the same authors also defined a user-centric taxonomy for multidimensional data projection tasks, divided in pattern identification, relation-seeking, behaviour comparison and membership disambiguation (ETEMADPOUR et al., 2015), and proposed the comparison between different DR methods through user studies over these tasks (ETEMADPOUR et al., 2015).

Bach et al. (2017), on the other hand, evaluated the effectiveness of Augmented Reality (AR) approaches, using tablets or see-through HMDs combined with tangible markers. They studied four different tasks in generic 3D point clouds: distance estimation, cluster counting, point selection and cutting plane orientation. It was observed that the proposed direct hologram interaction was helpful in highly interactive tasks, but the desktop alternative was still the quickest and most accurate in most cases.

Other recent works have also involved the immersive exploration of scatterplots. Babaee, Datcu and Rigoll (2013), for example, proposed a new metric to compare DR techniques in terms of structure preservation, based on a communication channel model.

They visualized datasets of images reduced to three dimensions in immersive CAVE-like environments. Stenholt (2014) proposed the use of an unnatural mapping for 3D glyph scatterplots, where points always cover the same screen area regardless of their distances, to enhance the perception of structures in cluttered environments, an important issue in visual data mining. He successfully evaluated this approach in an HMD-based immersive environment. Gray (2016) explored HMD-based 3D scatterplot navigation, presenting three different scenes in a public exhibition, and contributing with a recommendation to include a reference plane and some illumination from above, for the sake of orientation. Finally, Cordeil et al. (2017b) proposed an interactive tool, *ImAxes*, where variable axes can be combined, through embodied interaction, to construct different representations, such as 2D or 3D scatterplots and parallel coordinates plots.

2.5 Summary

Immersive Analytics is an ever-growing area in the convergence of Visualization and Virtual Reality research, concerned with applying novel display and interaction technologies in combination to support the analytical reasoning (CHANDLER et al., 2015). Such approaches have already achieved considerable success in the analysis of scientific spatial visualization. For abstract information visualization, however, more research and guidelines are still needed (GARCÍA-HERNÁNDEZ et al., 2016). Some promising results have been demonstrated, for example, in studies regarding node-link diagrams (HALPIN et al., 2008; KWON et al., 2016).

In this work, we expand this discussion to a different representation, commonly applied for the exploration of multidimensional information: 3D scatterplots obtained through dimensionality reduction techniques. In theory, an extra dimension would reduce information loss and allow a more faithful representation of the high-dimensional information. Previous works on this subject targeting monoscopic displays, however, have obtained contradictory results (POCO et al., 2011; SEDLMAIR; MUNZNER; TORY, 2013; GRACIA et al., 2016). Considering immersive displays, some authors have already presented some efforts in this direction, but most have been based in technologies which have progressed substantially over the past few years, and also explored only the application of CAVE-style environments (ARMS; COOK; CRUZ-NEIRA, 1999; RAJA et al., 2004; ETEMADPOUR; MONSON; LINSEN, 2013). Herein, we propose to investigate HMD-based approaches to this problem, considering that these devices present a much more accessible cost and have been recently demonstrated to be even more efficient in the case of graph connectivity tasks (CORDEIL et al., 2017c).

3 TASK-BASED EVALUATION FRAMEWORK

This chapter introduces the task-based framework. Section 3.1 presents a new formal modelling to the evaluation problem, identifying the different factors that determine the overall task performance in an immersive 3D exploration approach: the difference in errors introduced by performing dimensionality reduction to two or three dimensions, and the difference in human perception errors under different visualization conditions. Then, it describes the multidimensional datasets which are targeted as use cases throughout all analyses (Section 3.2) and the different analytical tasks identified for them (Section 3.3). Finally, Sections 3.4 and 3.5 discuss the different methodologies proposed to assess each of two kinds of errors in our model, leading to the next chapters.

3.1 Modelling of the Evaluation Problem

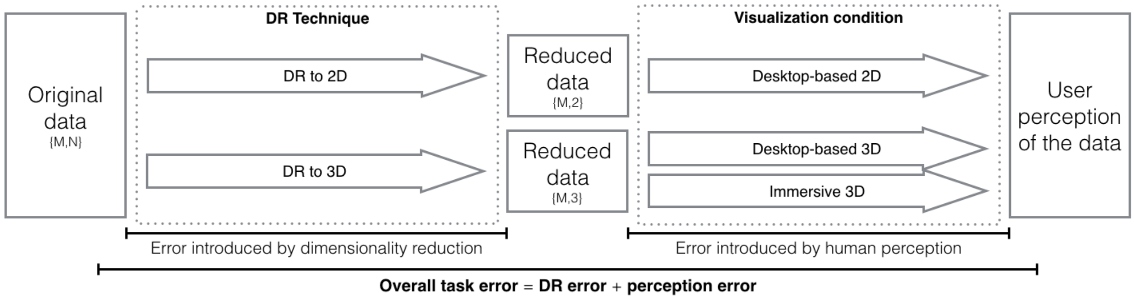
When exploring an immersive 3D dimensionally-reduced data scatterplot, the performance gains attained in a task are not just a function of the difference in perceptual accuracy presented by users under different visualization conditions, but rather of its interplay with the difference in errors introduced by reducing the dimensions of a particular dataset to two or three dimensions.

This new model of our problem in hand, accounting for the two different factors that influence in the final task performance outcome, is presented in Figure 3.1.

The so called *DR error* component is dataset-dependent, depending on the particular complexity of the data structure. This means that, for a given dataset to benefit from a three-dimensional visualization condition, its content must be indeed better mapped to 3D. Moreover, the user must be able to perceive this added information appropriately, what can be challenging given the previously discussed issues associated with three-dimensional representations.

Based on this model, we propose an evaluation framework that aims to separately assess each of these variables. The *maximum potential performance* in 2D or 3D for our datasets is estimated through a task-based empirical approach (Section 3.4). The *perception* and *overall task errors*, on the other hand, are assessed through user studies (Section 3.5), comparing alternative visualization conditions for two selected datasets, one with promised improvements in 3D (*D1*), and another one that, in thesis, allows for similar performance in all representations (*D2*).

Figure 3.1: A visual model of the problem we target. The overall task performance for each of the three scenarios will be a result of the different errors introduced: by reducing the dataset to two or three dimensions and by using a desktop-based 2D, desktop-based 3D or immersive visualization approach.



In order to explain our task-based framework we chose as use case the visualization of roll call voting data, already mentioned in Section 2.2 (and described in the next section), which original scatterplot representation is shown in Figure 2.4.

3.2 Targeted Use Case

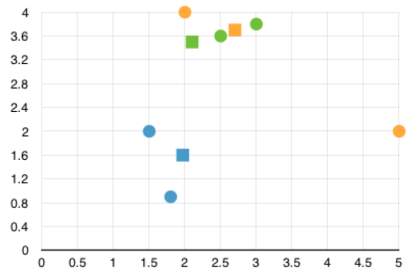
In this work, we visualize roll call voting data from the Brazilian Chamber of Deputies. This dataset is particularly interesting for visualization because the resulting spectrum is composed by more than 20 political parties, with very fuzzy ideological borders. We also consider this domain very appropriate for our goals due to the very high dimensionality of its datasets (each roll call is a dimension), its consistent application of DR techniques in the literature, and the easy definition of semantically meaningful analytical tasks. Moreover, it also appeals to different kinds of public, with potential to engage participants in data exploration during our user tests.

We extracted information about the votes of each deputy and the official vote instruction given by each party represented in the Chamber for every roll call in the last four four-year legislatures from the Brazilian Congress: 52nd (451 roll calls), 53rd (619 roll calls), 54th (428 roll calls) and 55th (493 roll calls). Since it is common for deputies to leave their seats during the term, we avoid calculating positions for deputies with very few votes by following the approach of Borja and Freitas (2015) and selecting only the most present 513 (the official number of seats) in a legislature. For each legislature, we constructed a voting matrix where all deputies and parties are represented by M lines, and roll calls are represented by N columns. Each (i,j) cell is then attributed a value depending on the ith deputy or party vote on the jth roll call: -1 for “no”, 1 for “yes” or 0 for absten-

Figure 3.2: A multidimensional roll call voting dataset can be reduced to two dimensions using Principal Component Analysis (a). The synthesised components can then be plotted as an automatically constructed political spectrum (b). In this representation, Euclidean distances indicate how similarly or differently deputies or parties have voted. Colours encode party affiliation, while shape represents the point category. To obtain a 3D version, the third principal component can be added.

	Roll call 1	Roll call 2	Roll call 3	Roll call 4	Roll call 5	...
Deputy A	Yes (1)	No (-1)	No (-1)	Yes (1)	Yes (1)	...
Deputy B	Absent (0)	Yes (1)	No (-1)	No (-1)	Absent (0)	...
Deputy C	Yes (1)	No (-1)	Yes (1)	No (-1)	Absent (0)	...
Deputy D	No (-1)	No (-1)	No (-1)	No (-1)	No (-1)	...
Deputy E	No (-1)	Yes (1)	Yes (1)	Yes (1)	No (-1)	...
Deputy F	Yes (1)	No (-1)	Yes (1)	Yes (1)	Yes (1)	...
Party A	Yes (1)	No (-1)	Yes (1)	No (-1)	No (-1)	...
Party B	No (-1)	No (-1)	No (-1)	Yes (1)	Yes (1)	...
Party C	Yes(1)	No (-1)	Yes (1)	No (-1)	No (-1)	...
...	...	...	...	...	...	...

	PC1	PC2
Deputy A	1.5	2
Deputy B	1.8	0.9
Deputy C	3	3.8
Deputy D	2	4
Deputy E	2.5	3.6
Deputy F	5	2
Party A	2	1.6
Party B	2.7	3.7
Party C	2.1	3.5
...	...	...



tion or absence. Principal Component Analysis (PCA) by Singular Value Decomposition (GOLUB; REINSCH, 1970) is then applied to this matrix, resulting in $min(N, M)$ principal components (HOTELLING, 1933). For visualization purposes, only the first two or three components are considered, and seen as a political spectrum. Figure 3.2 illustrates the complete process. Euclidean distances in these representations indicate how similarly or differently deputies have voted in the given period.

In the resulting point cloud, party information was encoded by colour, and the category of the point by shape: circles or spheres for deputies and squares or cubes for official party positions. The datasets also encompass associated party and state information for all points, which can be used for filtering.

3.3 Analytical Tasks

In order to assess our visualization conditions under a variety of usage patterns, we have selected a set of nine different tasks, divided into four categories. These tasks were based on both relevant task taxonomies (SARIKAYA; GLEICHER, 2017; ETEMADPOUR et al., 2015) and previous related evaluation studies (BACH et al., 2017; ETEMADPOUR; MONSON; LINSEN, 2013).

The subset of distance perception tasks (Section 3.3.1) will be used independently for our theoretical performance gain analysis (Section 3.4) and in the first user study (Chapter 4). It will then be extended with the remaining five tasks (Sections 3.3.2, 3.3.3 and 3.3.4) for the implementation of the second user study (Chapter 6).

3.3.1 Point-based distance perception tasks

The axes in a DR data scatterplot correspond to artificial, uncorrelated dimensions synthesized by an algorithm and, in general, have no semantic meaning. Instead, much of the information presented is encoded through the distance between points, which quantify the similarities or differences between them in the original data space. This subset is thus composed by tasks relative to different competencies in distance judgements: perception of near, medium and far distances, and of different shape encodings. They were designed to be simple and atomic (i.e., combinable for more complex analyses), but we believe they constitute a representative subset of the typical tasks of a data analysis in this specific domain.

T1 *Selection of a deputy's closest deputy*. In this near-distance perception task, the user is requested to select the closest deputy (sphere) to a given one.

T2 *Selection of a deputy's closest party*. In a more difficult variation of the previous task (since deputies are usually positioned between multiple parties), the user is requested to select the closest party (cube) to a given deputy. It can also be seen as a point classification task, where the user is reclassifying deputies in parties according to vote coherence.

T3 *Selection of a party's furthest member*. In this simplified outlier identification task, the user must select the member of a given party who is furthest located from the official party position.

T4 *Selection of a party's closest party*. Also a variation of T1, but exploring different competencies since parties are more distributed on the spectrum.

In terms of abstract scatterplot tasks, these refer to object comparison, neighbourhood exploration and distances understanding (SARIKAYA; GLEICHER, 2017).

3.3.2 Class-based density perception tasks

Class or cluster density is another important factor in point cloud representations in general, indicating group cohesion. This is one of the behaviour comparison tasks recommended by Etemadpour et al. (2015), and also a key scatterplot analysis task (numerosity/density comparison) in the taxonomy from Sarikaya and Gleicher (2017).

T5 *Density comparison between two parties.* In this task, the user must choose which of two simultaneously visualized given parties is the densest one.

T6 *Density comparison over time.* In this variation of the previous task, the user must choose one between two time periods, which cannot be simultaneously viewed, when the given party was denser.

3.3.3 Clustering task

Clustering is a typical pattern identification (ETEMADPOUR et al., 2015) or known motif search (SARIKAYA; GLEICHER, 2017) task in scatterplots.

T7 *Estimation of the number of clusters in a given point cloud.* This task requires the inspection of different orthogonal points of view (BACH et al., 2017).

3.3.4 Interaction tasks

Interaction efficiency is a main concern in 3D representations, also having been evaluated, for example, in the recent AR proposal by Bach et al. (2017). Tasks were designed to assess the interaction with both associated views and the 3D data points.

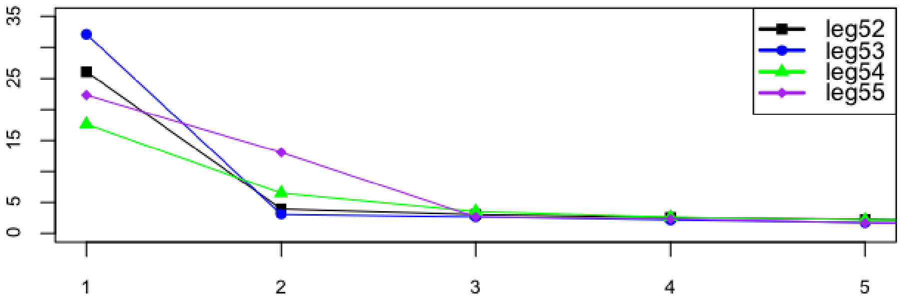
T8a *Filtering of a party-state combination.* The user is requested to select, as quickly as possible, the correspondent filters.

T8b *Selection of all remaining deputies.* Continuing the previous task, the user is requested to select, as fast as possible, all the remaining points in the 3D scatterplot.

3.4 Task-Based DR Error Assessment

The efficiency of a dimensionality reduction method when representing a dataset in a lower dimension is highly dependent on the data geometry. This implies that, while some datasets will benefit from an extra dimension, others will already be well represented in 2D. In fact, several metrics try to quantify the information gain of adding a third dimension to a DR data scatterplot. Gracia et al. applied 11 of these in their study with 12 DR algorithms and 12 real-world datasets to affirm that the loss of quality reducing from 3 to 2 dimensions accounts, in average, for 30.4% of the total DR loss (GRACIA et al., 2016). A simple and commonly used metric in the case of Principal Component Analysis is the proportion of variance contributed by each principal component, given by its eigenvalue. This information is usually plotted in a *scree plot* (see Figure 3.3) and used to estimate the dataset *intrinsic dimensionality* (INGRAM et al., 2010).

Figure 3.3: Scree plot of the proportion of variance contributed by each of the 5 first principal components in our 4 datasets.

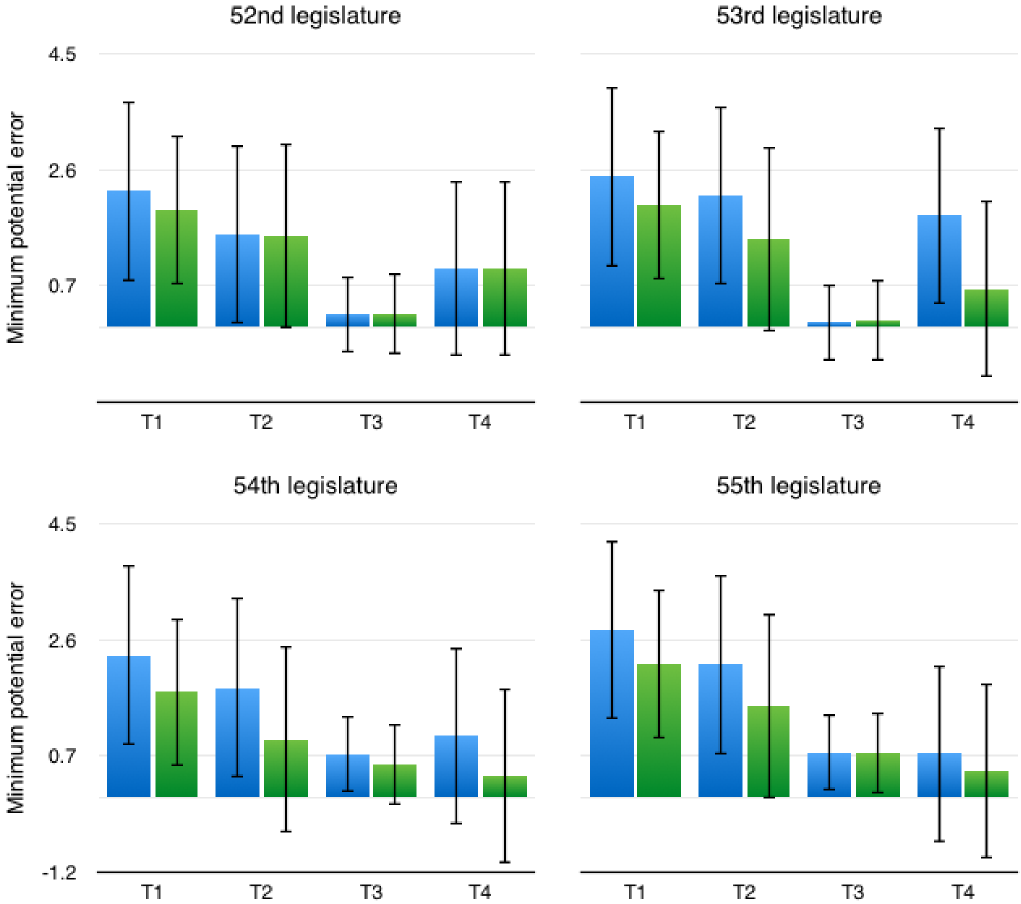


However, from a practical point of view, it is generally difficult to estimate how the information loss from 3D to 2D, even if it exists, will have impact on the user's analytical performance. Moreover, it is hard to conjecture whether the trade-off between information loss and the clearer and simpler visualization provided by 2D is worth it.

One possible approach is to implement a perception-based evaluation in the form of user studies (POCO et al., 2011; ETEMADPOUR et al., 2015). Nonetheless, it is an expensive alternative considering that results obtained for one dataset will not necessarily generalize to others.

We propose to approach this issue in an empirical, task-based way, by computing a user's *maximum potential performance* in 2D and 3D. This is done by simulating the minimum average error a user would achieve in each scenario if he/she were able to perceive the presented information with absolute accuracy, for all possible instances of a task. For example, if the task is selecting the closest deputy to a given one (T1), the

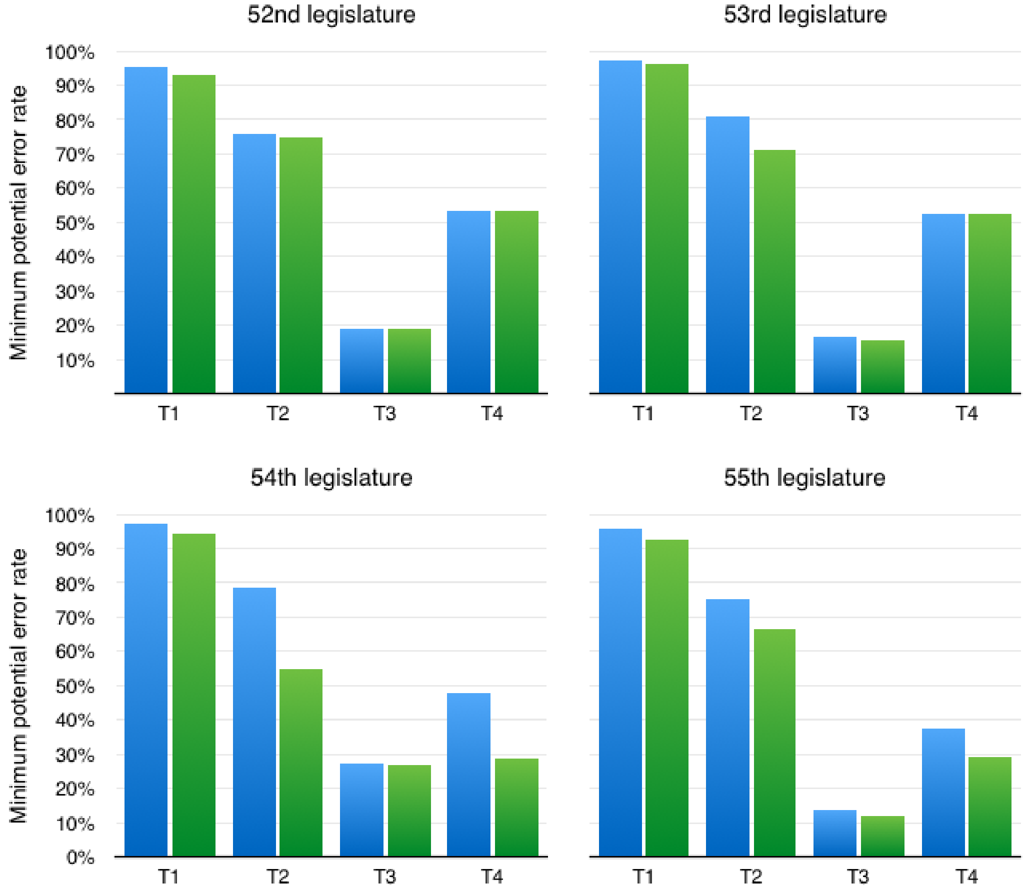
Figure 3.4: Results of our task-based analysis of the minimum potential average error a user could achieve both in 2D (blue) and 3D (green), were he/she always capable of perceiving accurately the distances represented. Error bars present the corresponding standard deviations.



correct answers to all 513 deputies according to the information presented in 2, 3 or all dimensions are calculated and compared. Euclidean distances between points in the corresponding set of dimensions are used, and average errors are always calculated in the original vote matrix.

Figures 3.4 and 3.5 present results for our four datasets over the different distance perception tasks introduced in Section 3.3.1, in terms of minimum average error magnitudes and minimum absolute error rates, respectively. As expected, different legislatures result in different potential contributions for the third dimension. We identify two particular scenarios: for the 54th legislature, all tasks appear to benefit from its inclusion – for T1, T2 and T3, it is the dataset with the largest performance improvement in terms of error magnitude, and notable reductions in error rates are seen for T2 and T4. Observing Figure 3.3, this was indeed the dataset with the smallest variance explained by the two first components combined. For the 52nd legislature, on the other hand, all tasks appear to be equally well performed in both scenarios – for T1, T2 and T4, it is the dataset with the

Figure 3.5: Results of the task-based analysis with respect to minimum potential error rates in two (blue) and three (green) dimensions.



least gain in error magnitudes. From now on, we will refer to these datasets as *D1* (54th legislature) and *D2* (52nd legislature). In the following chapter, we will assess how the task performance is affected by the user perception of the third dimension (under desktop and HMD-based conditions) in both of these cases.

3.5 User Studies for Perception and Overall Error Assessment

While the *DR error* part of Figure 3.1 can be simulated and calculated with exactitude for different datasets (Section 3.4), the *perception error* and, consequently, the *overall task error*, will be different for each individual person. A possible evaluation approach for these components is, therefore, to conduct controlled comparative user studies, aiming to observe patterns across a population for each task and visualization condition. According to Kosara et al. (2003), user studies can help to objectively establish which method, among those being assessed, is the most appropriate for a given situation, and to obtain insight into why that method is effective.

Throughout this work, two main user studies will be reported. They were designed to evaluate different visualization approaches. The first one, described in Chapter 4, used 3D scatterplots and a conventional gaze-directed flying navigation approach, while the second user study is reported in Chapter 6, based on a different visualization metaphor, which is described in Chapter 5.

They follow *within-subjects* protocols, where the same participant is asked to perform a series of tasks in all compared conditions, but in varying orders to compensate possible learning biases. This ensures that the observed differences in results correspond to differences between conditions and not between the participants.

To enable the assessment of subjective factors such as comfort and preference, users are also invited to answer general questions and fill standardized questionnaires. More specifically, the System Usability Scale (SUS) questionnaire was employed to assess system usability (BROOKE et al., 1996), and the NASA Raw Task Load Index (TLX) to assess user workload (HART, 2006). For immersive virtual environments, the Simulator Sickness Questionnaire (SSQ) was applied pre and post VR exposure to evaluate user discomfort (KENNEDY et al., 2003). Igroup Presence Questionnaire (IPQ) was also applied post VR exposure to assess the level of presence experienced by users in the virtual environment (SCHUBERT; FRIEDMANN; REGENBRECHT, 2001).

Finally, results are computed and statistically compared across groups with respect to different criteria, such as completion times and error rates, observing the occurrence or not of significant differences for each task.

3.6 Summary

In this chapter, we introduced a formal modelling proposed to assess the contribution of an immersive 3D approach to the exploration of a given dataset. This model identifies the two separate factors that determine the overall task performance: the difference in errors introduced by performing dimensionality reduction to two or three dimensions, and the difference in human perception errors under different visualization conditions.

We chose as target use case four multidimensional four-year roll-call voting datasets from the Brazilian Chamber of Deputies, and a collected set of nine different analytical tasks, related to distance perception, density perception, cluster identification and interaction.

Different methods are proposed to evaluate each of the errors in our model. A task-based simulation quantifies the *maximum possible performance* a user would achieve in each task using 2D or 3D, were he/she capable of perfectly perceiving the represented distances. Comparative user studies, on the other hand, are employed to assess perception errors. The following chapters report these user studies.

4 USER STUDY 1: CONVENTIONAL FLYING NAVIGATION

This chapter presents our evaluation approach for the error introduced by the human perception of the visual representation (the second factor discussed in Section 3.1), through a comparative user study. This study was informed by a preliminary pilot study (Section 4.1).

In this first user study, a conventional gaze-directed flying navigation approach was implemented. This metaphor is meant to be simple to learn (BOWMAN et al., 2004) and enable an egocentric view, placing the user inside the data representation.

Section 4.2 describes our implementation for each condition – desktop-based 2D (*2D*), desktop-based 3D (*3D*) and HMD-based immersive 3D (*IM*). Section 4.3 then defines the study hypotheses, and Section 4.4 presents the experiment design. Finally, Sections 4.5 and 4.6 report and discuss the study results, respectively.

4.1 Pilot Study

The first experiment in our research consisted in an assessment of our initial ideas through a pilot study. A preliminary user study with 20 participants was conducted to allow a comparative analysis between prototype implementations of desktop-based 2D, desktop-based 3D and HMD-based 3D approaches. Results indicated advantages in accuracy in a point classification task with respect to the original dataset, as well as in distance perception and outlier identification tasks with respect to the principal components being visualized. The proposed immersive framework was also well rated in terms of user perception, with the best scores for accuracy and engagement.

Nonetheless, several issues with the immersive implementation were identified, particularly regarding the hardware setup used, which was based on the outdated Razer Hydra hand controller. During the data analysis, we also observed that a refined problem modelling and study protocol were needed to better evaluate our set of hypotheses.

4.2 Visualization Conditions

The implementations for our three studied visualization conditions were based on those used in our pilot study (Section 4.1), and were updated to include feedback

Figure 4.1: Proposed HMD-based immersive environment for the exploration of dimensionally-reduced data scatterplots. The user is equipped with two position-tracked hand controllers, being allowed to interact with the data through selection pointers.

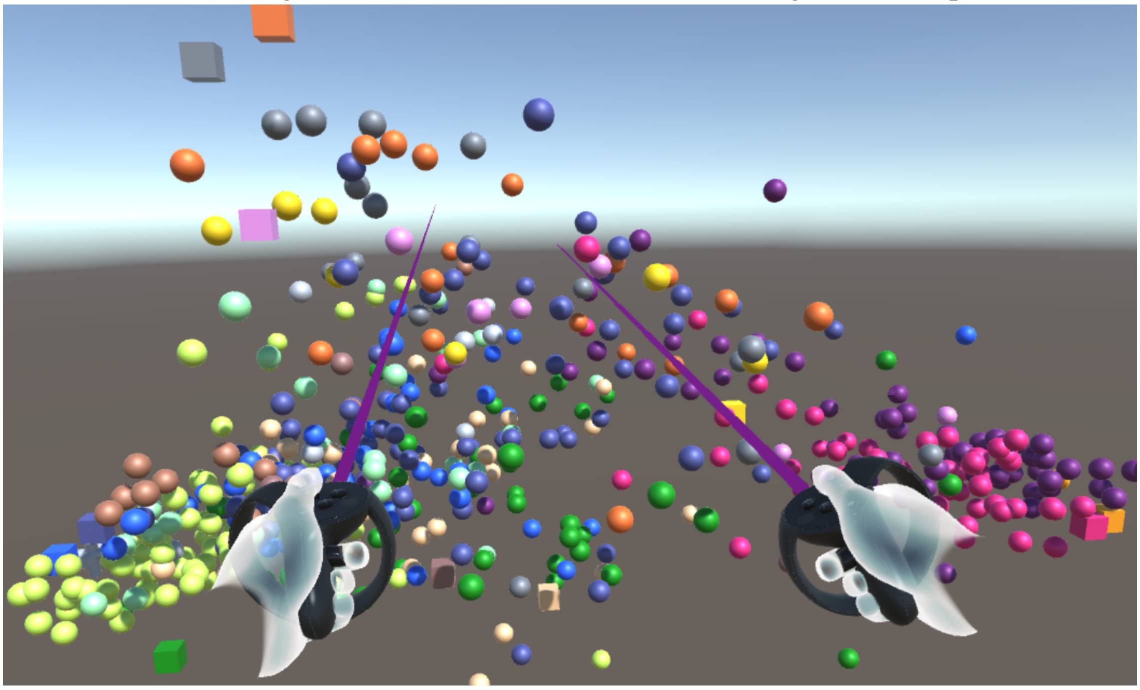


Figure 4.2: Pressing the inner trigger in the controller, emulating a grabbing action, the user can highlight the whole party of any given point to inspect its relative position. The other hand can then be used to highlight another party for comparison or to select individual points. This feature is particularly useful in the party outlier identification task.

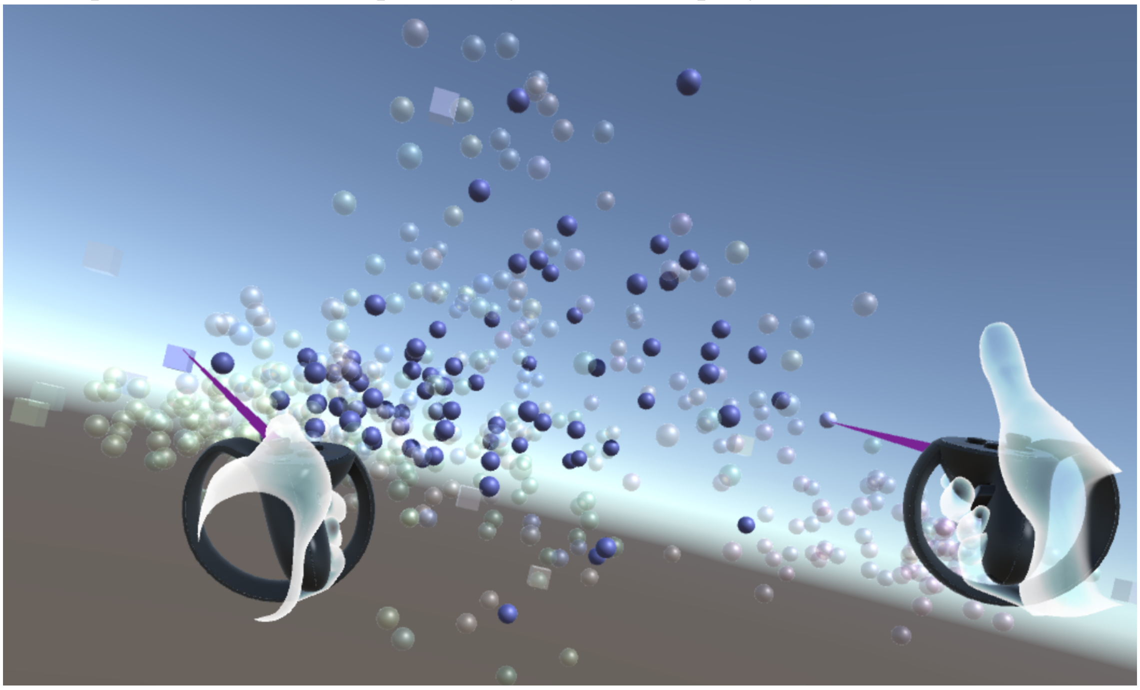
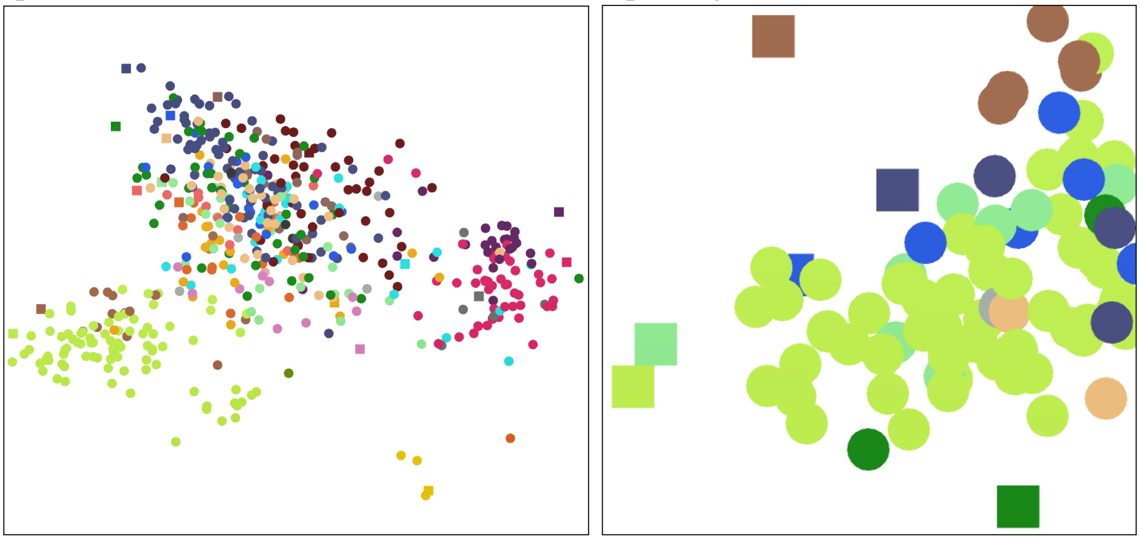


Figure 4.3: In the *2D* visualization condition, data points are distributed along screen space (left), and the user is allowed to zoom and pan (right).



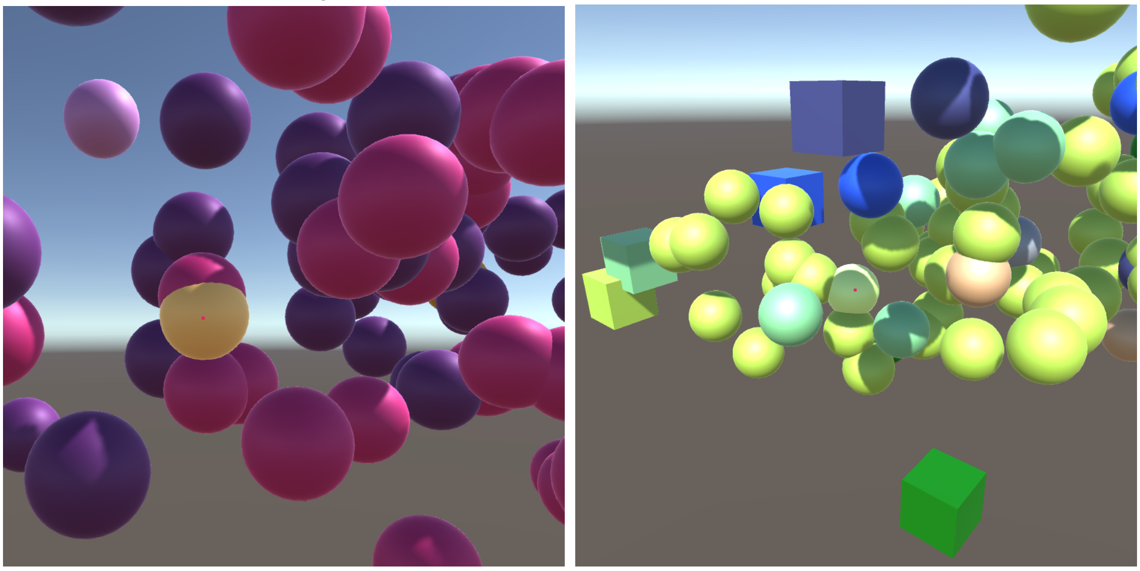
provided by those participants. Both two and three dimensional virtual environments (VEs) were implemented using the Unity game engine. The 3D version can be explored either through desktop-based (monoscopic, non head-tracked) or HMD-based (stereoscopic, head-tracked) setups.

In both desktop-based VEs, explored through a 22” Full HD display, controls were implemented using only mouse and keyboard, as in a traditional data analysis setup. In *IM*, our implementation choice, looking for providing a more natural and immersive interaction, was to use two selection rays, which are controlled by position-tracked Oculus Touch hand controllers (see Figures 4.1, 4.2 and 4.6). Accurate virtual representations of the users’ hands and of the controllers are also shown[1], increasing the feeling of embodiment and serving as anchors to the real world (ZIELASKO et al., 2017; SIMEONE; VELLOSO; GELLERSEN, 2015). This environment is explored through an Oculus Rift CV1 HMD (formed by two 1080x1200 displays), with the user seated in a swivel chair. Several guidelines were employed to minimize possible discomfort: the speed of movement is slow and constant; user control of the camera is maximized; no near ground was included to avoid uncomfortable rapid ground plane changes; and adequate hardware was employed to minimize latency and lag. In the event of teleportation to a new position, such as in the beginning of a task, a camera fade is also applied (YAO et al., 2014).

All VEs explore the same visual encodings: colours for political parties and shape for different categories of points – circles or spheres for deputies and squares or cubes for official parties positions. They also all offer the same set of possible interactions: a

[1]The official Unity Oculus Integration package was used.

Figure 4.4: In the *3D* conditions, the user is allowed to freely navigate through the data, which is distributed along a 3D virtual environment.



user may click on a point to show/hide its name (using double click or a specific button in the controller) and may highlight the whole party of any point to inspect its relative position (using right click or the inner trigger in the controller, to emulate a grabbing action). All versions also support the simultaneous selection of up to two parties for comparison. Labels are shown upon selection during the familiarization phase, to aid in the comprehension of the representation semantics. During the tasks, they remain hidden to avoid potential use of previous knowledge.

The setups differ, however, in the forms of navigation. In *2D*, the user can zoom in/out and pan the screen (see Figure 4.3). In both 3D versions (see Figures 4.1 and 4.4), the user can navigate freely in all directions, through gaze-directed flying (MINE, 1995). He/she is allowed to move forward, backwards, vertically or laterally, using keyboard keys or the left controller joystick, and also rotate the camera, moving the mouse or using the right controller joystick.

Moreover, while, in *2D*, selection is done by the mouse cursor and, in *IM*, by the pointer rays, in *3D*, it is also gaze-directed, implemented by a reticle cursor in the center of the screen, so that the mouse movement can be used to rotate the camera. The 3D environments also include a ground and sky background and illumination from above for orientation purposes (GRAY, 2016).

4.3 Hypotheses

We defined five hypotheses for our evaluation purposes. As defined in Section 3.4, *D1* is a dataset that presents potential information gain in 3D, and *D2* one that does not.

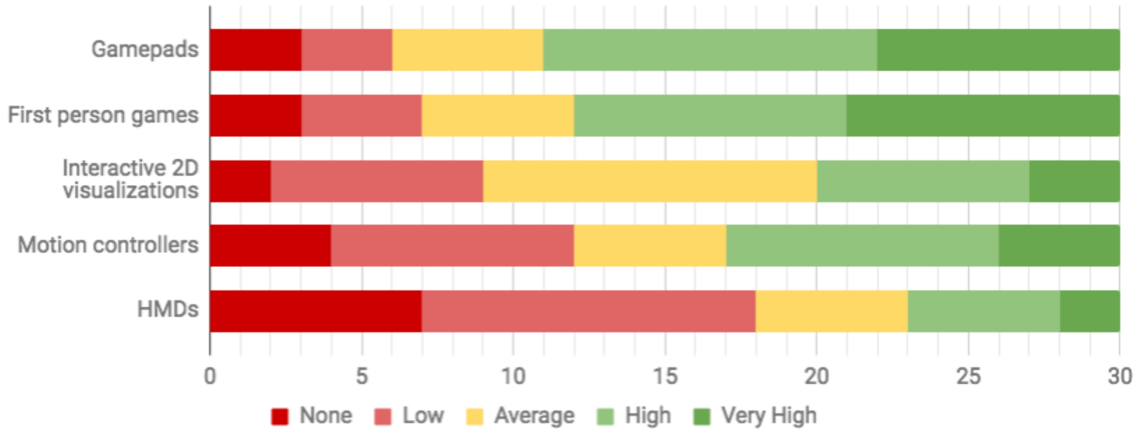
- H1 The perception error will be smaller in *IM* than in *3D*, specially due to the stereopsis (GREFFARD; PICAROUGNE; KUNTZ, 2014).
- H2 The overall task error in *IM* will be smaller than in *3D* or *2D* for *D1*.
- H3 The overall task error in *IM* will be at least as good as in *3D* or *2D* for *D2*.
- H4 *2D* is expected to be the quickest, given its inherent smaller cost for navigation and interaction.
- H5 The benefits provided by immersion, such as a more natural interaction and an egocentric view of the data (CHANDLER et al., 2015), will be reflected on the subjective user evaluations.

4.4 Experiment Design

Our user study was implemented through a within-subjects protocol, combining *3 visualization conditions x 4 tasks x 2 datasets*. The target population, recruited on campus, was composed by 30 subjects (20 male/10 female; average age of 25.2, ranging from 17 to 50), who had not taken part in the pilot study. Regarding previous contact with involved technologies, 76% reported at least average familiarity with first person games and gamepads, and 60% with motion controllers. However, 60% had low or no familiarity at all with HMDs (see Figure 4.5). The testing environment is illustrated in Figure 4.6.

Each participant experienced all visualization conditions in alternating order, to minimize learning biases. The subject was always initially allowed to get familiar with the corresponding controls while exploring the 55th legislature dataset. Then, he/she was asked to perform, as accurately as possible and without specific training, each of the tasks described in Section 3.3 six times in a row, being three in dataset D1 and three in D2. The order of presentation of the datasets in each task is alternated between users, but task order is preserved. Between different conditions, the scatterplots are mirrored with relation to the vertical and/or horizontal axes, so as to minimize the possibility of using previously

Figure 4.5: Participants' familiarities with related technologies. Most reported at least average previous contact with first person games and gamepads, but 60% had low or none with HMDs.



viewed information. The specific task questions presented were selected as follows: for each task and dataset, 10 different sets of 9 points were randomly selected (3 for each condition). Each of these sets was used by three users, alternating the conditions, so that, in the end, every point used once in one condition was also used once in the others. The purpose of selecting multiple sets of random points instead of just one is to maximize the representation of different possible situations in the data, and to cross validate the results (GRACIA et al., 2016). Also to maximize representation, in tasks involving deputies repetition was not allowed even between sets (this way, these tasks explore 90 out of the 513 possible deputy points) – for party tasks, this is not possible due to their smaller number, and so repetition is not allowed just inside the sets.

In all tasks, one point in the scatterplot is shown blinking, and the user must point and click to choose the corresponding answer. Following previous experiences from our pilot study, we opted to block semantically impossible answers (e.g. a party outlier that is not from the given party), so as to reduce noise resulting from accidental clicks or misunderstandings. When this is the case, the user hears a negative audio feedback. Upon an acceptable answer, a positive sound is played, and the camera is teleported back to the initial overview position.

After each technique, subjective opinion questionnaires were applied, including SUS questions (BROOKE et al., 1996). SSQ (KENNEDY et al., 1993) was applied pre and post VR exposure to evaluate eventual well-being effects. In the end, users were also allowed to compare all the techniques according to different criteria. The complete experiment took approximately 45 minutes.

Figure 4.6: Participant performing tasks in the immersive condition.

4.5 Results

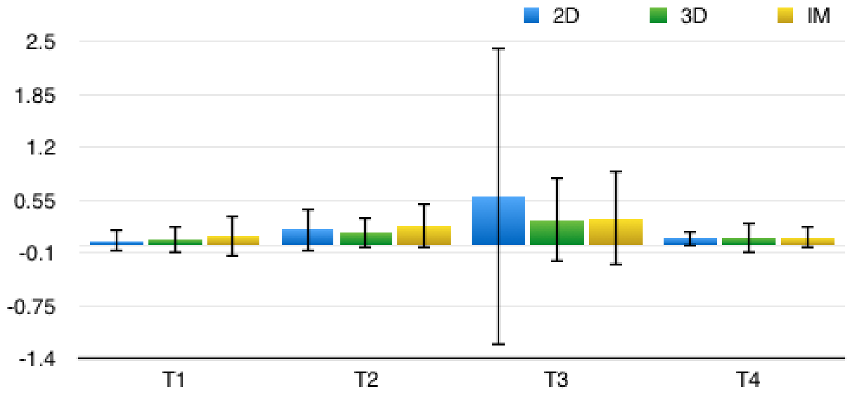
Observations from the user study are reported here in terms of task performance (Subsection 4.5.1), navigation patterns (Subsection 4.5.2), hand usage (Subsection 4.5.3), user feedback (Subsection 4.5.4) and simulator sickness (Subsection 4.5.5). Significance under the adequate statistical tests is indicated in the figures as follows: (*) for $p < 0.05$, (**) for $p < 0.01$ and (***) for $p < 0.001$. Curved lines indicate significant pairwise differences found in post-hoc testing.

4.5.1 Task Performance

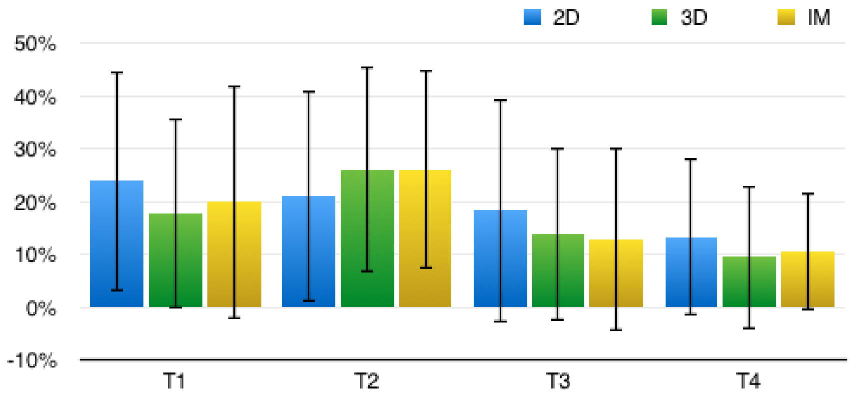
Perception errors were calculated as the differences in Euclidean distances, in two or three dimensions, between the one from the given point to the user's answer and the one to the correct answer in the representation. They refer, therefore, to the errors with relation to the information shown, and not to the original data. The better the user was able to perceive the distances in the representation, the closer to zero this error will be. Figure 4.7 presents results for all tasks (in this analysis, we do not differentiate between datasets). Since we were not able to verify normality under Shapiro-Wilk tests, non-parametric Friedman tests were executed. Post-hoc tests are implemented using the Wilcoxon-Nemenyi-McDonald-Thompson (HOLLANDER; WOLFE; CHICKEN, 2013) test. Surprisingly, no significant differences were observed in any task (p-values .72, .43, .83 and .57, respectively, for error magnitudes), neither between *3D* and *IM* nor between both and *2D*. The same holds true when analysing the error rates (Figure 4.7 (b)), which

Figure 4.7: Results for perception errors under the different conditions and tasks. They are given by the average differences between the Euclidean distances from the task point to the user answer and to the correct one, in two or three dimensions. Surprisingly, no significant differences were observed in any task.

(a) Error magnitude



(b) Error rate



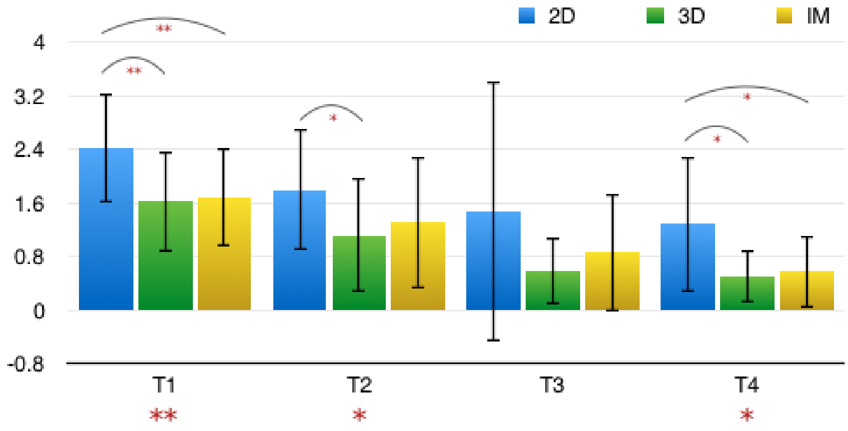
consider only whether answers were correct or incorrect and not the magnitude of the errors – p-values for this metric were .6, .41, .17, .28. H1 was, therefore, not confirmed.

Overall task errors were calculated as the differences in Euclidean distances, in the original vote matrix, between the one from the given point to the user's answer and the one to the correct answer in the real multidimensional data set. They are expected, therefore, to be the combination between the expected DR errors seen in Figure 3.4 and the perception errors seen in Figure 4.7. The results for all tasks in D1 and D2 are shown in terms of magnitude in Figure 4.8. Friedman and the Wilcoxon-Nemenyi-McDonald-Thompson post-hoc tests were again used, and the significant pairwise differences, when found, are indicated with red lines.

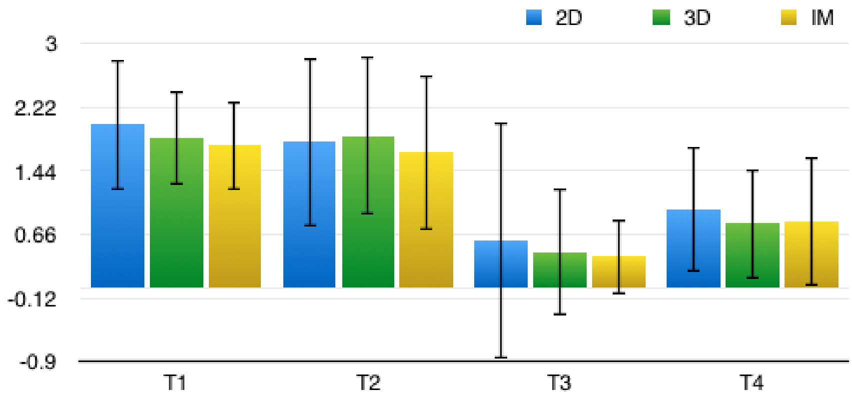
For D2, no task presented significant differences between conditions (p-values .6, .6, .93 and .85, respectively). This confirms our hypothesis H3, i.e., *IM* would be at least as good as *2D* for the dataset that has the least expected information gain with the use of

Figure 4.8: Overall task errors (w.r.t original data) observed for each dataset. Asterisks and curved lines indicate occurrence of statistical significance. Confirming the previous theoretical task-based DR Error simulation, D1 consistently benefited from the inclusion of the third dimension, while D2 did not.

(a) D1



(b) D2

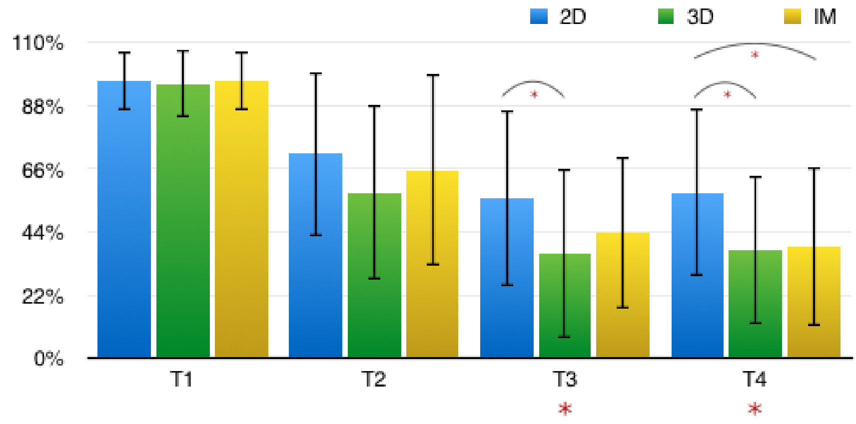


the third dimension. D1, on the other hand, presented significant differences for all tasks except T3, which can be considered almost significant (p-values were .002, .007, .06 and .01). All indicated pairwise differences presented $p < .01$. In T2, *2D* and *IM* presented a trend of significance with $p = .08$. Notably, however, H2 could not be confirmed, since *3D* and *IM* were not found to be different in any case.

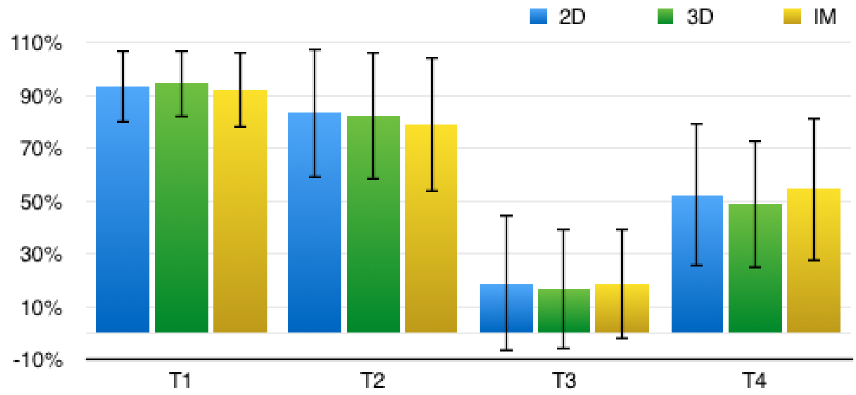
Figure 4.9 also presents the absolute error rates for each dataset. In this analysis, results in D1 for tasks T2, T3 and T4 are similar to the previous one – nonetheless, in this case, T2 did not achieve significance ($p = .14$), while T3 did ($p = .03$ in both the Friedman test and the pairwise *2D-3D* comparison). T1, on the other hand, presented very similar rates under all conditions, despite the clear differences in magnitudes. This suggests that, even in the three dimensional conditions, users were still unable to determine the most appropriate answer in the multidimensional space for this task, but were closer to it than in *2D*. In D2, no task presented significant differences, matching the magnitude analysis.

Figure 4.9: Overall task error rates (w.r.t original data) observed for each dataset. Asterisks and curved lines indicate occurrence of statistical significance.

(a) D1



(b) D2



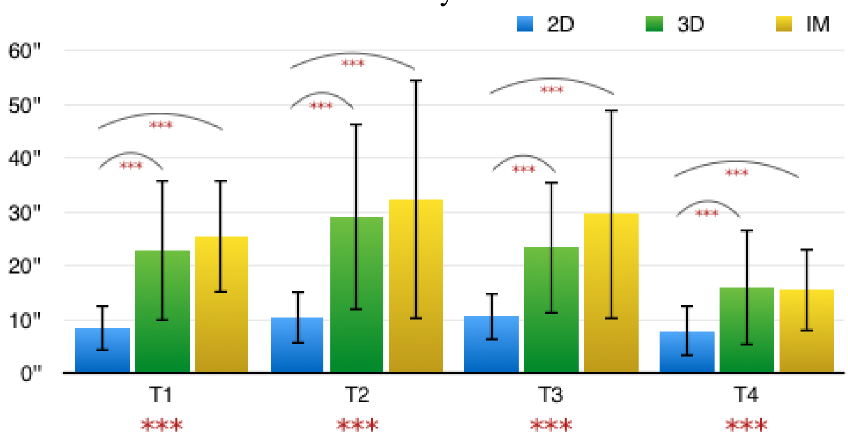
As expected in H4, *2D* was significantly faster in time than the two other conditions in all tasks (always with $p < .001$). *3D* and *IM* did not present significant differences with each other in any case (see Figure 6.3).

Distance perception tasks T1 and T4 were the quickest to be solved, with average times of 8.4s, 23.1s and 26.2s for *2D*, *3D* and *IM* in the former, and 8s, 16.3s and 15.8s for the latter. The outlier identification and classification tasks took longer, especially in the 3D conditions. This was already expected due to their higher difficulty, since frequently there are multiple possible answers (observe Figures 4.2, 4.3 and 4.4). Average times were 10.5s, 29.8s and 33.1s for T2, and 10.5s, 23.6s and 30.2s for T3.

4.5.2 Navigation Patterns

The monitoring of user navigation patterns in both three dimensional conditions showed that navigated distances were consistently longer in *3D* in comparison to *IM*.

Figure 4.10: Task completion time under different visualization conditions. Curved lines indicate significant pairwise differences. *2D* allowed consistently faster task completion, but *3D* and *IM* did not show differences in any case.



More specifically, they were 18% longer in T1 (p = .009, under a paired Wilcoxon signed-rank test), 20% longer in T2 and T3 (p = .03 and p = .07, respectively), and 30% longer in T4 (p = .004). This was not reflected in faster completion times, as seen previously, probably due to the slower navigation speed adopted in the immersive scenario, or because we did not ask users to care about the time. Many users complained about the slow speed and not being able to increase it, but we believe this contributed to minimize the occurrence of simulator sickness. Figure 4.11 illustrates the trajectories of the 30 participants for T4 (the one with the largest observed differences) in both *3D* and *IM* when using dataset D1. Trajectories for other tasks and dataset D2 are available in the supplementary materials.

Similar behaviours were also observed in terms of accumulated camera rotation, which was 48% larger for *3D* in T1 (p = .0004), 23% larger in T2 (p = .1) and 38% larger in T4 (p = .003). The only exception was T3 (rotation 8% smaller, p = .62), what is explained by the different nature of this task (perception of long distances). Considering that our protocol ensures that a task performed in one condition will always be performed by another user in the other conditions, these differences are not related to task difficulty, but to the interaction and visualization techniques themselves. The enabled navigation forms were also similar in both conditions. The rotation difference may be partly due to the different fields-of-view (FOV) in both scenarios (60 degrees in *3D* and 96 in *IM*). Another plausible explanation for navigation and rotation variations may reside, however, in the different depth cues provided. As discussed by Ware (2012), a very important cue for the inspection of clouds of points, besides stereopsis, is structure-from-motion.

Finally, navigated distance was also found to be, as expected, consistently inversely correlated with perception error, particularly for *3D*. Pearson correlations between the two metrics were -0.66 and -0.49 for *3D* and *IM*, respectively. While at least three users were observed to develop the strategy of assuming points positions to obtain egocentric perceptions of distance, most adopted allocentric points of view.

4.5.3 Hand Usage

Considering specifically the immersive condition, we were particularly curious about how users would adapt to the two-handed embodied interaction metaphor. All hand movements and interactions (point selections and party highlights) were thus recorded. Observed right hand use was much more pronounced, as was already expected given that only one participant had reported being exclusively left-handed. Nonetheless, an interesting result was that hand usage varied according to the task requirements. Average numbers of interactions with the left and right hands were, respectively, 1.1 and 7.7 for T1, 1.0 and 8.7 for T2, and 1.0 and 6.3 for T4 (ratios of 6.6, 8.4 and 6.1). For T3, where a common approach was to highlight the party with one hand and select the party outlier with the other (see Figure 4.2), this changed to 4.2 and 13.8 (a ratio of 3.2, less than half of the other tasks). Moreover, while in T1, T2 and T4, less than 30% used both hands to interact, in T3 this was done by 63% of the users.

Another interesting observation was that the differences in average hand movement were much smaller than the ratios in effective interactions, suggesting the users consistently moved both hands together despite using one of them much more frequently. Average hand translation per task was about 1.0 meter for the left hand, and 1.2 meters for the right one.

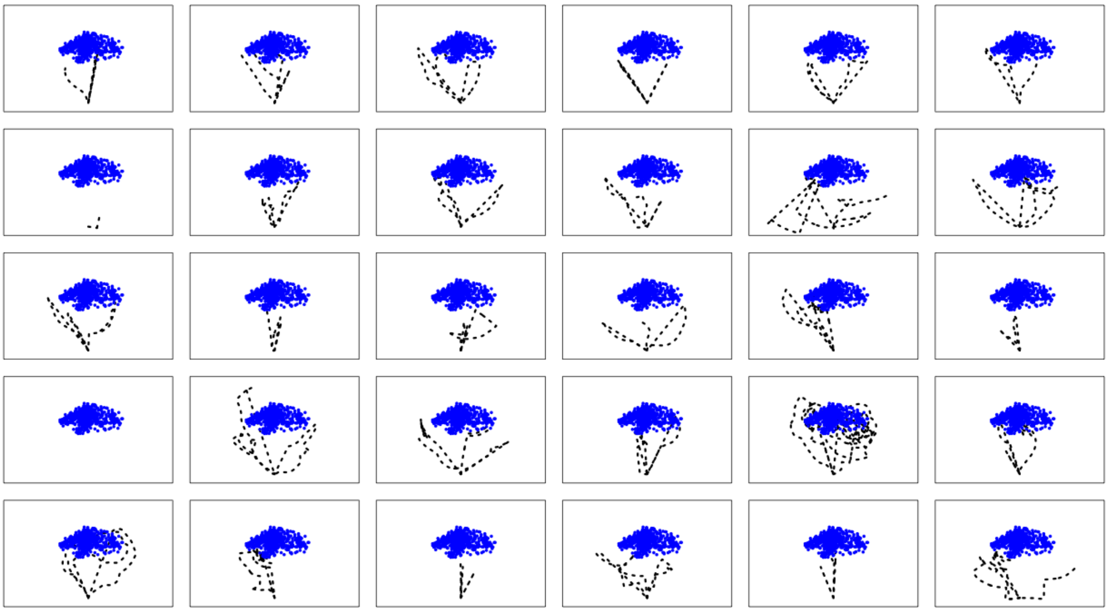
4.5.4 User Feedback

All visualization conditions were well rated with relation to usability, without significant differences (p = .11). SUS questionnaire scores were 81.5 for *2D*, 77 for *3D*, and 76.6 for *IM* (standard deviations 12.6, 15.3 and 20.2). We believe this successfully reflected our efforts to optimize our implementations (in the pilot study, the ratings for the previous versions had been scored 83.1, 61.3 and 68.3, respectively).

In post-technique interviews for all conditions, at least 75% of the participants also agreed it was easy to navigate and interact, achieving 90% in some cases (*2D* navigation,

Figure 4.11: Top view of trajectories for each participant when solving task T4 in dataset D1. Navigated distances and camera rotations were found to be consistently longer in the non-immersive condition, reaching a 30% difference in this task.

(a) 3D



(b) IM

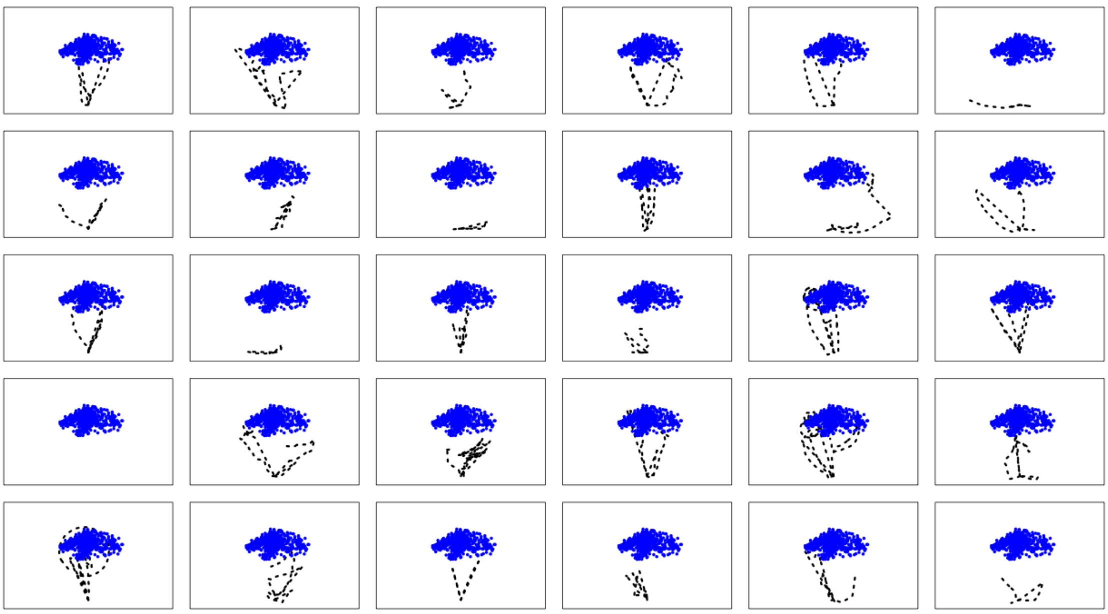


Figure 4.12: User subjective evaluations according to different criteria. *2D* was almost unanimously pointed out as the quickest, and *IM* as the most engaging. Despite similar quantitative results, subjective perception of accuracy in *IM* was much larger than in *3D*.

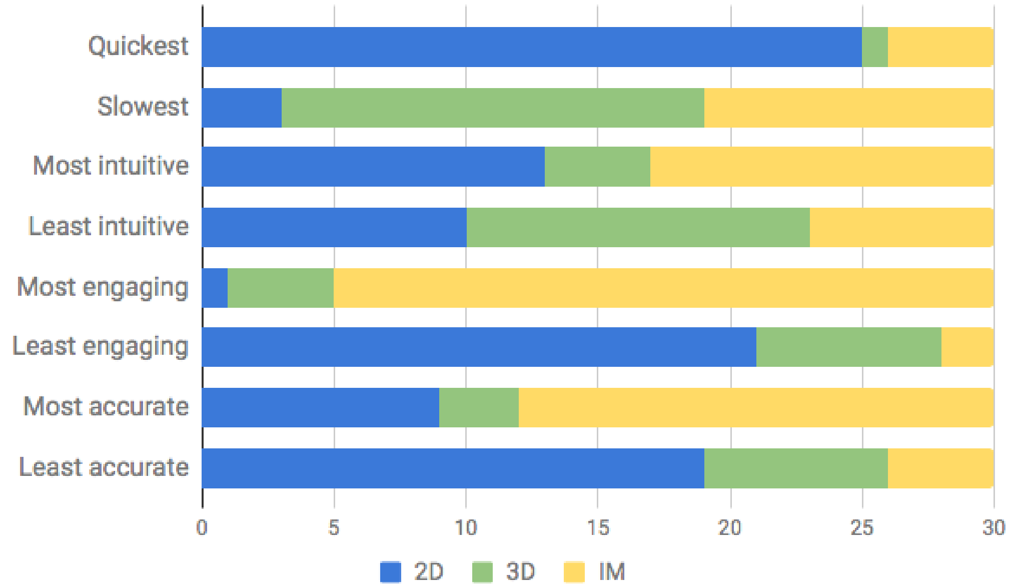
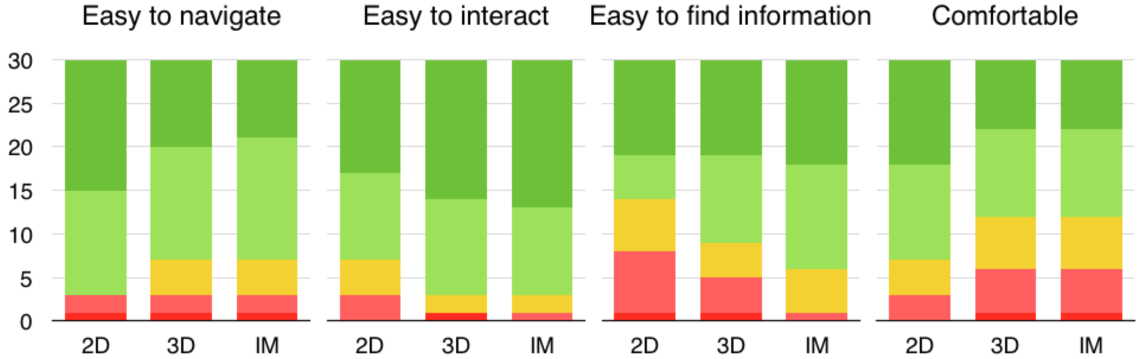


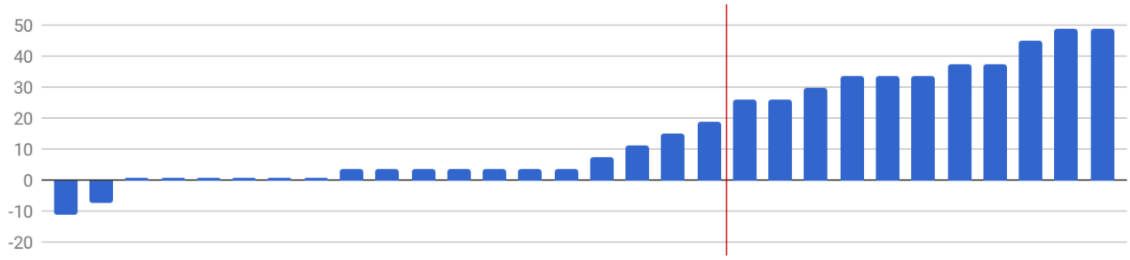
Figure 4.13: User Likert-scale agreements to the different assertions, ranging from completely disagree (dark red) to completely agree (dark green). All techniques were well rated in terms of ease of navigation and interaction. Noticeably, however, *IM* was better rated in terms of ease to find information, and performed as well as *3D* for comfort.



3D and *IM* interactions) (see Figure 4.13). Users appeared to be able to complete the tasks with less effort in *IM*, with 24 agreeing that it was easy to find information in this representation, compared to 21 in *3D*, and 16, in *2D*. However, no significant differences were found in the Likert-scale questions – Friedman tests indicated p = .14 for navigation, .26 for interaction, .15 for information finding and .14 for comfort.

Figure 4.12 presents the user reported perception of each condition according to different criteria, as assessed by the final questionnaires after all three had been experienced. Interestingly, despite the similar quantitative results achieved for both *3D* and *IM*, 18 users perceived *IM* as the most accurate of all, against only 3 for *3D*. 19 users indicated *2D* as the least accurate condition, probably because, with one less dimension,

Figure 4.14: SSQ score impacts post-VR exposure for all 30 participants, ordered from least to most severe. Around 60% of them presented only minor symptoms (to the left of the red line), but others presented quite significant discomfort levels.



points were clearly less well distributed in space. A Friedman test on the mean rankings for accuracy ($p = .005$) indicated significant differences between *IM* and *2D* ($p = .005$) and near significance between *IM* and *3D* ($p = .052$), but no significance between *2D* and *3D* ($p = .71$). *2D* and *IM* tied in the dispute for the title of most intuitive ($p = .16$), with 13 votes each, what is rather surprising considering the ubiquitousness of 2D interfaces (actually, *2D* was also voted 10 times the least intuitive, versus 7 of *IM*). In terms of time, subjective perceptions confirmed the quantitative observations, with *2D* being placed behind *3D* ($p < .001$) and *IM* ($p < .001$), and no significant differences between *3D* and *IM* ($p = .55$). Finally, 25 participants classified *IM* as the most engaging condition, compared to *2D* ($p < .001$) and *3D* ($p = .002$). This is probably related in large part to the novelty of the display and interaction technologies being used, but may also refer in part to its sense of immersion and egocentric point of view. Differences between *2D* and *3D* were not significant ($p = .07$).

4.5.5 Simulator Sickness

Simulator sickness is still a major issue in immersive environments, especially when non-physical navigation is employed. Despite following multiple guidelines (Section 4.2), we still observed significant well being effects on part of the subjects. Figure 4.14 displays, ordered from least to most severe, the observed VR exposure impacts on the SSQ scores (KENNEDY et al., 1993) for all participants.

Noticeably, while around 60% reported only minor symptoms (to the left of the red line), the others presented quite significant discomfort levels. Many users reported that this was minimized (though not avoided) when employing physical movements, for example, to rotate the camera, instead of using the alternative joystick control.

User results did not appear, however, to be impacted by the occurrence of discomfort, with a Pearson correlation of only -0.1 between SSQ scores and average perception errors in *IM*.

4.6 Study Discussion

The results obtained from this user study offered many insights into the hypotheses defined for our evaluation. The most surprising was certainly the absence of significant distance perception differences between *2D*, *3D* and *IM*, contradicting previous beliefs about the suitability of monoscopic 3D scatterplots and also our hypothesis H1. We believe that this is related to the fact that our desktop-based 3D environment, implemented in a powerful game engine, does not resemble typical 3D scatterplots. Designed in an effort to enable a fair comparison with its HMD-based counterpart, it provided game-like first person navigation and a multitude of depth cues (including perspective, occlusion, shading and structure-from-motion). As a consequence, both *3D* and *IM* were able to present the promised information gain for dataset D1, with significant or almost significant differences to *2D* with relation to the original voting data in all tasks. Both techniques were also able to present similar performance to *2D* in dataset D2. These facts confirmed part of hypotheses H2 and H3.

Analysing behavioural and subjective results, however, a series of differences between *3D* and *IM* appears. An equivalent performance appears to have taken considerable less effort in the immersive scenario, given that, under this condition, users were required to navigate up to 24% less, and agreed more often that information was easy to find. This could benefit higher-levels tasks, such as cluster detection, which requires estimating multiple pairwise distances at the same time – nonetheless, this should be verified by future studies.

Subjectively perceived accuracy was also much larger for *IM* than for *3D*, despite their similar results. This was also observed during our post-test interviews, when many participants described being convinced of a better performance within the immersive scenario. However, it could be argued that this might potentially generate over-confidence in incorrect observations. *IM* was also labelled the most engaging, what we believe may be, at least in part, linked to its natural interaction and egocentric point of view, as stated in our hypothesis H5. Despite around 40% of the users presented significant levels of discomfort due to simulator sickness, *IM* was also well rated in terms of usability through the SUS questionnaire, with a similar score to *3D*. Task completion times were, as expected

(H4), around 3 times slower in *IM* than in *2D*, due to the navigation and interaction costs incurred by the third dimension. However, no significant differences were observed between *IM* and *3D*, despite the slower navigation provided.

4.7 Summary

In a comparative user study between three different conditions (desktop-based 2D, desktop-based 3D and HMD-based 3D), we observed that perception errors were similarly low in all conditions. Task performance was therefore improved with the addition of the third dimension regardless of immersion, when the data enabled so. Nonetheless, the HMD-based condition required smaller effort to find information and less navigation, besides offering a much larger subjective perception of accuracy and engagement. A limitation to the proposed immersive approach, however, was the high incidence of simulator sickness, with around 40% of the participants reporting significant discomfort levels.

5 VIRTUALDESK: A NOVEL PROPOSAL FOR MORE COMFORTABLE AND EFFICIENT IMMERSIVE INFORMATION VISUALIZATION

In the previous chapter, we observed that immersive approaches may effectively aid in the exploration of multidimensional information, but that new evaluations and guidelines are still needed. Navigation, especially, is an open topic. Many proposed approaches are impractical for actual usage. Flying metaphors, in particular, are time-consuming and often result in simulator sickness. Other approaches, such as real walking, are also unnecessarily inefficient, both in terms of time and space requirements. Moreover, how to display inherently 2D content and texts in the virtual environment is another known issue.

In this chapter, we propose and implement an alternative data exploration approach where viewpoint change is only realisable through head movements. All data manipulation is done directly by natural mid-air gestures, with the data being rendered at arm's reach. To increase immersion and enable the display of two-dimensional associated views and interaction with tangible controls, we also build upon previous work and reproduce in the virtual environment an exact copy of the analyst's desk. Important data exploration resources are provided, including coordinated views, combinable filters and annotation tools.

5.1 Introduction

Navigation is indeed a key issue in immersive approaches. Most works adopt artificial metaphors such as flying (DONALEK et al., 2014; ZIELASKO et al., 2016; BOWMAN et al., 2004), which frequently induce simulator sickness due to conflicts with the user's real perceived position. Others have also tried to employ physical movements, such as walking, as an alternative (SIMPSON; ZHAO; KLIPPEL, 2017), but this is generally very time and space consuming. Intermediate solutions, such as using physical movements like body leaning to control the artificial navigation, have also been proposed, but with limited success (ZIELASKO et al., 2016). We argue, however, that the best approach would actually be to render the data in smaller scale, at arm's reach, and just manipulate it with natural mid-air gestures to obtain different points of view.

The reproduction of the user's physical desk in the virtual environment, such as done in our work, was firstly seen in Zielasko et al. (2017)'s research. They later also experimented with the inclusion of the user's keyboard into the virtual scene (ZIELASKO et al., 2017). However, both these works still apply artificial flying navigation, with the desk flying coupled to the camera throughout the environment (see Figure 2.3), making our concept and implementation fundamentally different.

Cordeil et al. (2017a) recently defined the concept of *spatio-data coordination* (SD), aiming to lower the user's cognitive workload when exploring information visualizations. They argue for a one-to-one mapping of positions, directions and actions between the physical and virtual environments, and present a design space to categorize novel solutions. Our small-scale dataset rendering is consistent with their sketch of a virtual mid-air design for SD coordinated interaction.

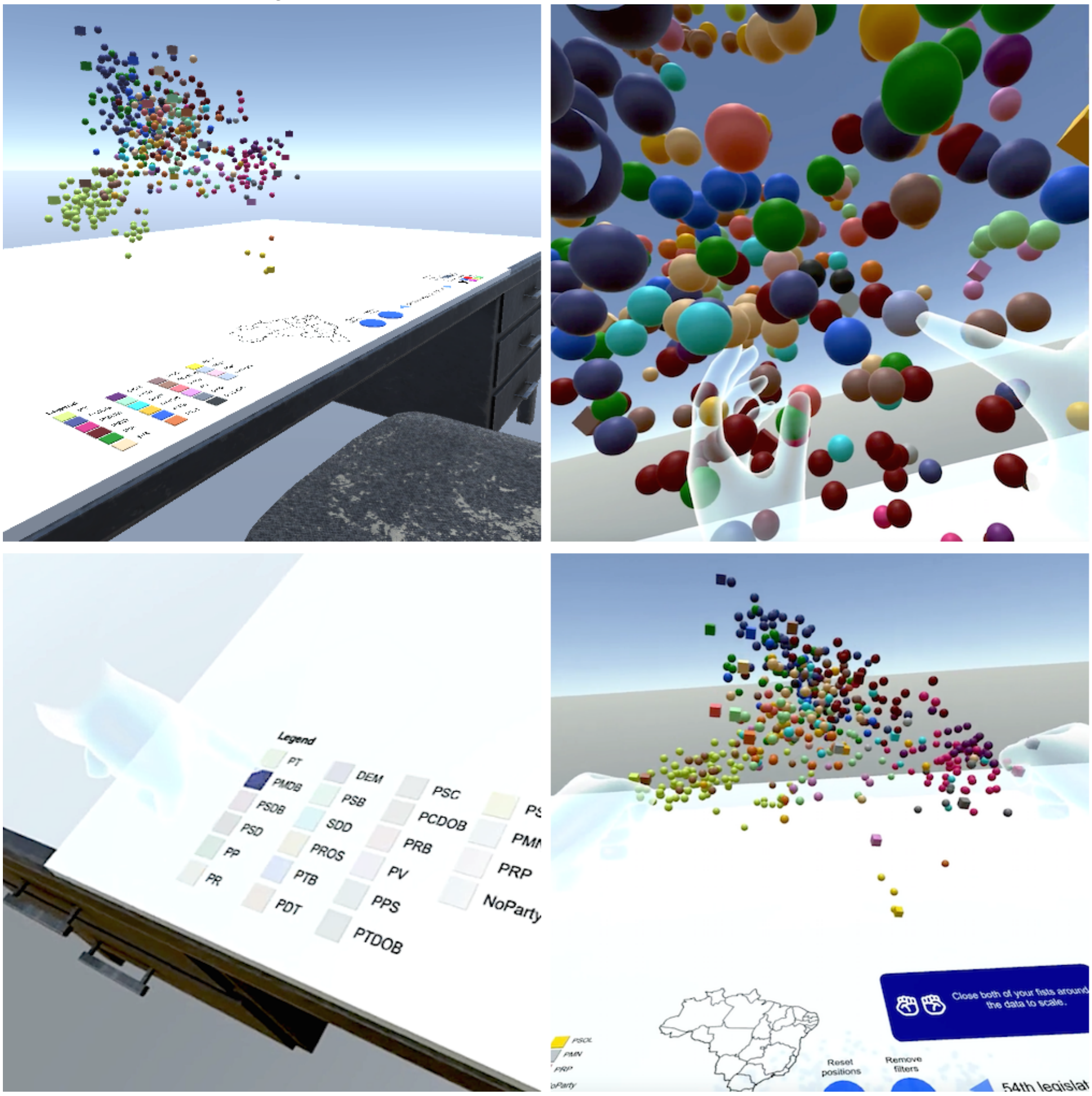
Finally, we also borrow concepts from 3D user interfaces (3DUI) research. Bowman et al. (2004) presented a thorough discussion on 3D interaction techniques. In our scenario, the most relevant is the direct manipulation through simple virtual hands. Mine, Jr and Sequin (1997) also discussed how interacting within arm's reach can take advantage of proprioception to provide a greater sense of position and orientation of manipulated objects. Body-relative interaction also provides higher precision and stronger stereopsis and head-motion parallax cues.

5.2 Data Manipulation and Interaction

In the VirtualDesk prototype , all data manipulation is implemented by natural mid-air gestures, using direct interaction with virtual hands (BOWMAN et al., 2004) (see Figures 5.1 and 5.2). This is expected to minimize the user workload, given the intuitiveness of the actions and also the application of the sense of proprioception during the interaction.

The main actions consist in *grabbing* the dataset and *tapping* data points. After grabbing the dataset with one hand, the user can move it and also rotate it around the hand position. Grabbing with two closed fists allows for the rotation and translation with relation to the central point between hands, and also the scaling of the dataset proportionally to the variation in distance between hands. Data points are selectable by quickly double tapping on their surfaces (see Figure 5.1 – top right). This was chosen instead of single tapping to avoid the selection of undesired points in cluttered regions. Haptic feedback in the form of vibration when touching points contributes to the perception of a tangible

Figure 5.1: In the VirtualDesk prototype, data is rendered at arm's reach and manipulated only by mid-air natural hand gestures. An exact reproduction of the analyst's real desk is included to enable tangible interaction with coordinated views and controls.



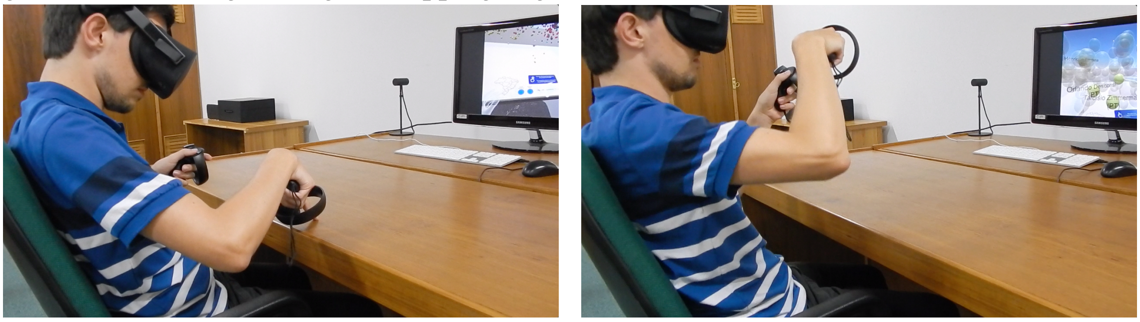
interaction.

In this prototype, we opted to implement different actions for selection with each hand: while the right index finger activates a point (displaying its associated information or choosing it as answer in a task), the left one triggers the supportive action of highlighting a whole set of points for providing context to the user.

5.3 Tabletop Tangible Interaction

Following recent literature (ZIELASKO et al., 2017), we decided to incorporate an exact replication of the user's desk into the virtual environment. The use of tangible

Figure 5.2: In the VirtualDesk prototype, all system control and data manipulation are performed by tabletop tangible interaction (left) or controller-agnostic mid-air natural gestures, such as grabbing and tapping (right).



user interfaces (TUIs) is known to greatly benefit immersion (CORDEIL et al., 2017a). The virtual desk is represented in an exact position (see Subsection 5.5) so as that, when the user touches the surface of the real table, his virtual hand touches the virtual one. We refer to this form of interaction as *tabletop*, to avoid confusion with the term *desktop*. Although the virtual desk is rendered larger than the real one, to provide a greater notion of space, a different marking keeps the user aware of the position of the actual desk (see Figure 5.1 – bottom left).

Several controls are available on the virtual desk's surface in our prototype: buttons to reset the data points to the original position and scale, remove filters and change datasets. These buttons also provide haptic vibration to increase tangibility. Moreover, coordinated filters and visualization tools are also provided (Subsection 5.4). All these components are shown in the frontal part of the desk, for easy access. An important note is that this segment must be free of obstacles (e.g. the user's keyboard) to avoid unintended collisions (ZIELASKO et al., 2017).

By incorporating an element of the real world, VirtualDesk can also be described as a *mixed reality*, or *augmented virtuality* application (MILGRAM; KISHINO, 1994).

5.4 Coordinated 2D Views and Visualization Functionalities

Besides enabling tabletop tangible interaction, we also see the inclusion of the virtual desk as an opportunity to tackle a key challenge in virtual environments: how to display and interact with texts and two dimensional information.

Two views associated to the main dataset were incorporated in the prototype as examples: a legend for categorical information and a map for spatial filtering (see Figure 5.1). Both of them act as combinable coordinated filters, showing or hiding information in the main 3D view.

Additionally, an annotation panel was included as an example of possible extra analytical feature. This panel allows the user to change the colour mapping of points to an uniform colour, and then to mark individual points. These annotations could easily be persisted for future inspection either in VR or in a conventional display.

5.5 Technical Details and Choices

The VirtualDesk prototype was implemented using the Unity3D game engine and the Oculus Rift CV1 HMD (composed by two 1200x1080 stereoscopic displays). Adequate hardware was used to meet the recommendation of a frame rate around 90 FPS (YAO et al., 2014).

An important decision in the implementation was the selection of the hardware for the tracking of the user hands. Several related works that explored mid-air gestures in the past have employed the Leap Motion hand tracker (BURGESS et al., 2015; THEART; LOOS; NIESLER, 2017; ZIELASKO et al., 2017). Based on previous literature and our own experience, however, we felt that this would not match the level of precision and comfort required for a satisfactory user experience. Cordeil et al. (2017c), for example, recently discussed that frustration caused by the Leap Motion tracker losing track of fingers positions during the user study was the main reported downside in their HMD-based condition. Guna et al. (2014) also presented a study on the device's precision and reliability, and concluded that its limited sensory space and inconsistent sampling frequency compromised its suitability for dynamic tracking.

We opted instead to use the recently released Oculus Touch hand controllers. Although these controllers do not track the position of each finger, they are very precise in tracking the overall hand position based on the Constellation tracking technology. Moreover, they apply different touch and near-touch sensors coupled with heuristics to determine the finger positions. The official Unity Oculus Integration Package provided the hand models and the gesture mapping.

The Oculus Touch tracking was also used to implement the virtual desk positioning. Upon the application start, the controllers are placed in a fixed location, and the virtual desk is then rendered in relation to their detected positions, resulting in a very accurate solution.

Another design choice we made was to not use any controller-specific tool, such as buttons, in any action – i.e., the actual controllers are completely abstracted by the users after they learn the gestures. The reason was twofold: we wanted to base interaction only on natural actions, and also to obtain a controller-agnostic framework, which could easily be adapted to any other tracking device.

5.6 Summary

A novel 3D immersive data exploration approach was proposed to circumvent the previously observed limitations related to large completion times and high incidence of simulator sickness. The VirtualDesk prototype is based mainly on embodied natural manipulation and interaction with data rendered at arm's reach, and tabletop tangible interaction with controls and 2D coordinated views positioned on the surface of a virtual desk, which position is synchronized with the analyst's real desk. Next chapter reports the user study we conducted to evaluate this approach.

6 USER STUDY 2: VIRTUALDESK EVALUATION

In order to assess how the VirtualDesk prototype, described in the previous chapter, would perform in comparison both with conventionally used desktop-based approaches and the previously employed flying navigated immersive approach (Chapter 4), a new evaluation study was carried out. A new comparable desktop-based environment was entirely implemented, with the same functionalities and following typical mouse and keyboard interaction approaches (Section 6.1).

Both conditions employ our visualization use case with multidimensional data projected to three dimensions. In order to identify the strengths and weaknesses of each, participants were asked to complete an extended set of 9 representative perception and interaction tasks (Section 3.3), inspired in previous literature.

Our main hypothesis was that our setup would enhance user perception and decrease workload, while remaining time-efficient and not inducing cybersickness.

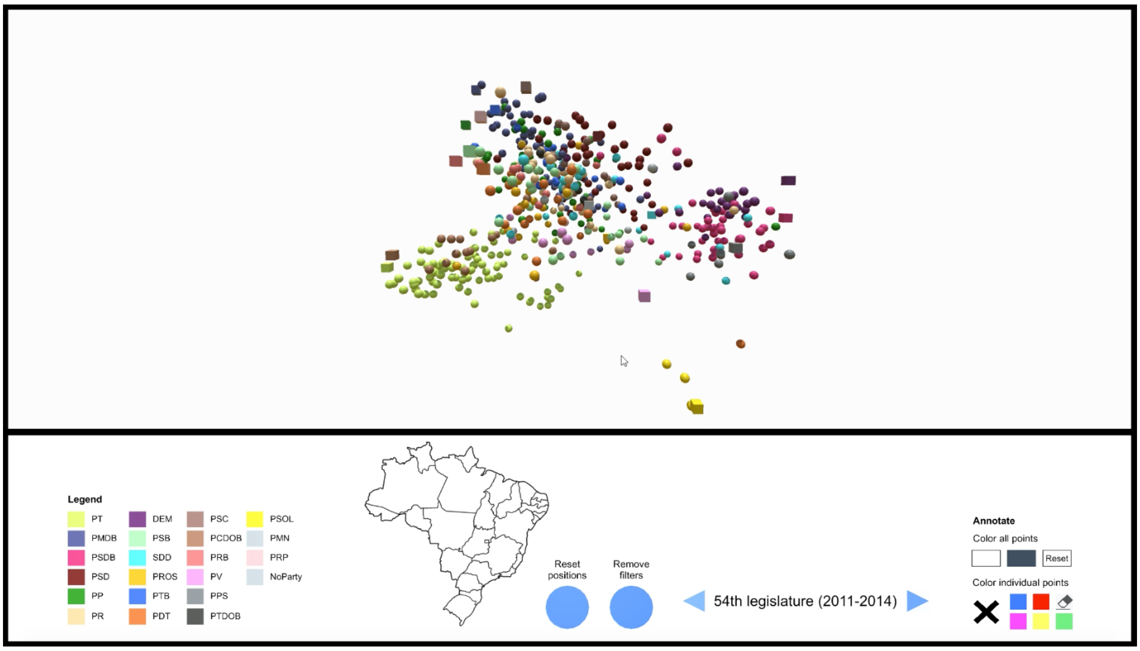
6.1 Comparable Condition

A key limitation in our previous study was the fact that the then-used desktop-based 3D condition was not really representative of a typical 3D scatterplot visualization tool. Its game-like design and interaction resulted in very high perception performance even in a monoscopic display due to the presence of multiple depth cues, but very inefficient task completion times.

This time, we decided to implement a new desktop-based comparable counterpart (*Desktop*) to the VirtualDesk condition based on the Rotate-Pan-Dolly paradigm, a very standard approach employed by almost all 3D modelling environments (JANKOWSKI; HACHET, 2015). Depending on the mouse button being pressed (left or middle), mouse movement is mapped to either rotation around the dataset center or camera translation (*panning*). The scroll wheel can be used to *dolly*, or *zoom*, into the data. Additionally, we also allowed the rotation around any selected pivot point (by holding a keyboard key) in order to enable better local inspection, required in some tasks. The selection of data points is implemented by double-clicks with the left mouse button, while class highlight is associated to the right button. Perspective projection was used as an additional depth cue, increasing similarity to the immersive environment.

This condition is explored in a Full HD 22" monoscopic display. The screen was divided into two areas: the upper 65% are dedicated to the dataset view, while the bottom 35% show a menu panel, with all the components, with the same proportion used on the VirtualDesk's surface (see Figure 6.1).

Figure 6.1: The desktop-based implementation provides all the same functionalities as the immersive environment, but employing a two-panels interface and Rotate-Pan-Dolly interaction for the 3D points cloud exploration.



6.2 Hypotheses

The following specific hypotheses were defined for the evaluation experiment.

H1 Easier data manipulation, proprioception and stereopsis combined will result in enhanced perception of distances and densities in the VR condition.

H2 Consolidated mouse-based interaction will still be quicker and more accurate for the selection tasks.

H3 Natural embodied interaction will decrease user mental workload and increase subjective perceptions of accuracy and engagement.

H4 The VirtualDesk metaphor will be more comfortable and efficient, both in time and task correctness, than previous immersive approaches.

6.3 Experiment Design

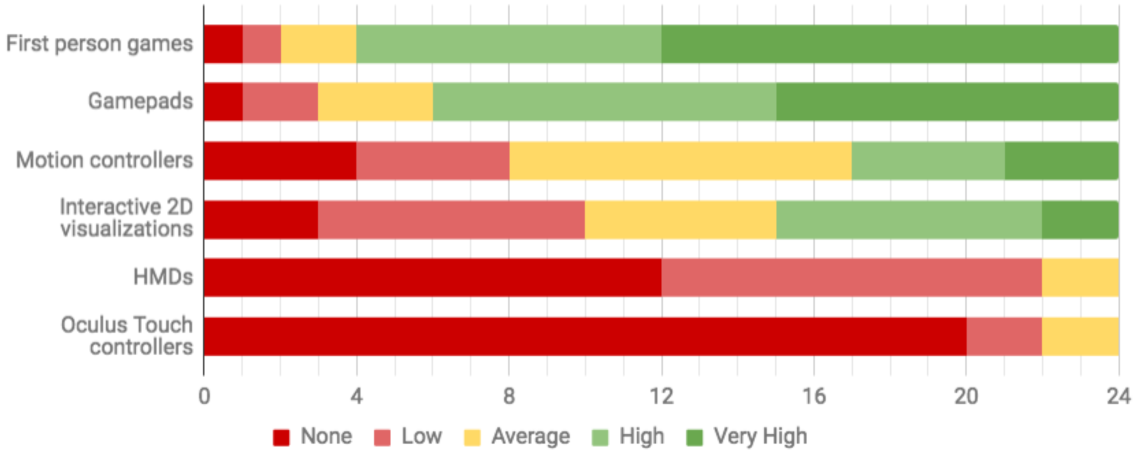
A new population of 24 undergraduate Computer Science students (20 male/4 female, mean age 23.7, SD 2.7) was invited to perform all tasks in the two compared conditions, in a *within-subjects* protocol. Half the users presented some visual condition, and wore glasses in combination with the HMD. Twenty-two of the users reported no or low previous experience with HMDs, and 20 of them had no previous contact at all with the Oculus Touch controllers. Nonetheless, 22 reported at least average experience with 3D computer games, 21 with gamepads and 16 with motion controllers in general (see Figure 6.2).

In the beginning of each condition, users were always presented a tutorial, which guided them through all system functions and exercised the different forms of interaction. Then, they proceeded to execute the tasks, which were always introduced by text accompanied by an illustrative icon (on-screen or close to the surface of the VirtualDesk). Participants were allowed to raise questions at any moment. The condition order was always alternated to compensate for the fact that, in the second condition, tasks would already be familiar to the users, but the task order was always kept the same to avoid confusion. Tasks were also distributed according to their increasing needs for interaction, so that previous tasks contribute to the familiarization with the system. Tasks always started in a data overview position. In Desktop, the monitor was positioned approximately 50cm in front of the users. In VirtualDesk, the center of the point cloud was initially positioned approximately 60cm in front of the users, and points rendered with a 1.5cm diameter.

For tasks T1-T4, one point in the cloud is shown blinking, and the user must select another point as answer. To enable the comparison with our previous study, where these four tasks were also used, the same selection of question points was repeated. These had been selected randomly, forming different sets of points from the *54th legislature* dataset that are repeated only once by a unique participant in each condition (all questions answered by a user in one condition will also be answered by a different user in the other). This maximizes the exploration of different possible situations in the data, and cross validates the results (GRACIA et al., 2016).

In T5 and T6, relevant parties are already shown highlighted (i.e. with the remaining points semitransparent), and the user must select the party cube correspondent to his answer. For each task, two different sets of 3 questions were selected, and were alternated between conditions. In T5, pairs of parties in the *54th legislature* are compared, and in

Figure 6.2: Participants' familiarities with related technologies. Most reported no previous experience with HMDs and the Oculus Touch hand controllers.



T6, the same party is compared between the *53rd* and *54th legislatures*. Only large parties with at least 20 deputies were considered.

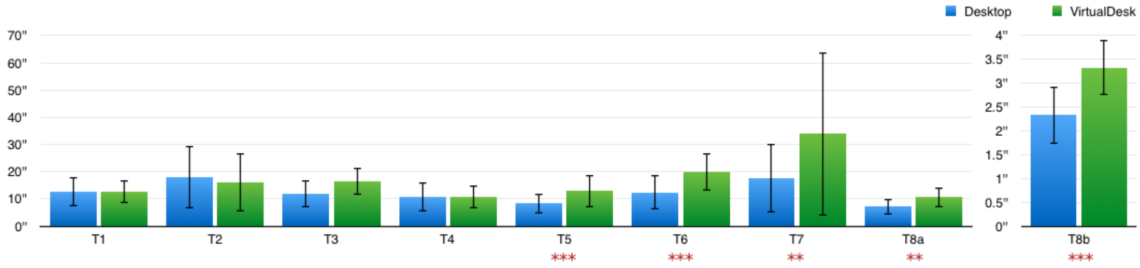
In T7, all points are shown in black to facilitate the perception of clusters and avoid confusion with classes, and the answer is given by an incremental counter positioned in the lower panel of the screen or near the surface of the desk. One of the four collected datasets was presented for each user in each condition, in varying orders.

In T8, the name of the party to be filtered is shown written in the task display, and the filtered state is marked in red on the map (so as to avoid interference of varying previous geographical knowledge). Six different party-state combinations were selected in the *54th legislature* dataset, with 3 being presented in each condition in varying orders. To maximize the representation of real use scenarios, states of different sizes on the map and parties in different positions in the legend were selected. Pairs were also carefully selected to result in the selection of different numbers of points (3, 6 or 9).

For T1-T7 users were asked to be accurate and, for T8, to be fast. Once again, we blocked semantically impossible answers (e.g. a party outlier that is not from the given party), so as to reduce noise resulting from accidental clicks or misunderstandings. When this is the case, the user hears a negative audio feedback. Upon an acceptable answer, a positive sound is played, the image briefly fades and the data returns to its original overview position.

In the end of each condition, users were asked to fill standardized questionnaires and answer general questions. In both parts, the SUS questionnaire was applied to assess system usability (BROOKE et al., 1996), while the NASA Raw TLX was applied to assess user workload (HART, 2006). SSQ was applied pre and post VR exposure to evaluate simulator sickness (KENNEDY et al., 2003). IPQ was also applied post VR exposure to

Figure 6.3: Average task completion times for all tasks and conditions, with standard deviations indicated by error bars. For T8b, reported times are normalized per selected point. The immersive environment was only significantly slower in tasks which required higher amounts of interaction with the tabletop controls.



assess the level of presence experienced by users in the virtual environment (SCHUBERT; FRIEDMANN; REGENBRECHT, 2001).

The complete experiment took approximately 40 minutes.

6.4 Results

Results from the user study evaluation are reported here in terms of task performance (Subsection 6.4.1), user feedback (Subsection 6.4.2) and a comparison with our previous experiment (Subsection 6.4.3). Significance under the adequate statistical tests is indicated in the text and figures as follows: (*) for $p < 0.05$, (**) for $p < 0.01$ and (***) for $p < 0.001$.

6.4.1 Task Performance

Task performance was assessed in terms of task completion times, error rates and error magnitudes. Since, in this study, we are concerned only with the correct perception of the actual representation, and not with the dimensionality reduction accuracy, all tasks are evaluated considering the lower-dimensional space only.

For distance perception tasks T1-T4, pairwise Euclidean distances were computed to determine the correct answers. For density perception tasks T5-T6, we followed Etemadpour et al.'s approach based on the inverse of the average edge length in the Euclidean Minimum Spanning Tree of a class (ETEMADPOUR; MONSON; LINSEN, 2013). For the clustering task T7, ground truth was computed by the X-Means algorithm, and varied between 2 and 3 clusters (PELLEG; MOORE, 2000). Interaction tasks T8a and T8b, on the other hand, are assessed in terms of unintended selections. Times for

Figure 6.4: Average error rates for all tasks and conditions, with standard deviations. For the interaction tasks, errors are given by the number of unintended selections. All perception tasks were performed equally well or better in the immersive environment. Point selection, however, was more accurate in Desktop.

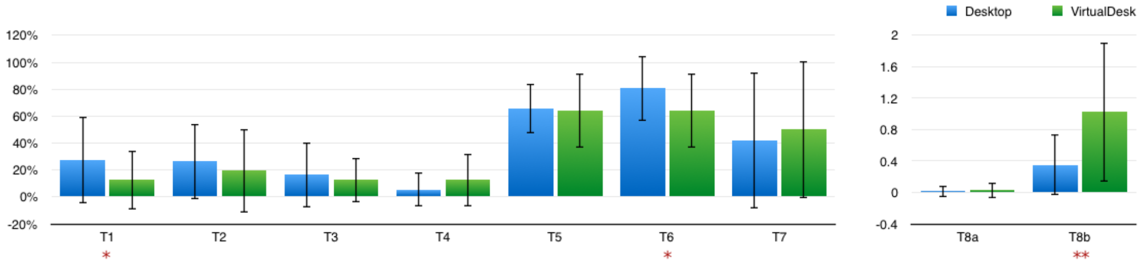
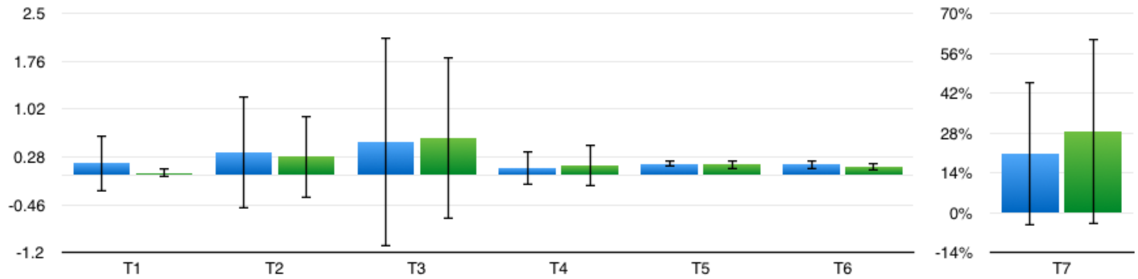


Figure 6.5: Average error magnitudes for tasks T1-T7, with standard deviations. Tasks T1 and T6 in particular also presented improvements in the immersive environment under this metric, but no significant differences were found.

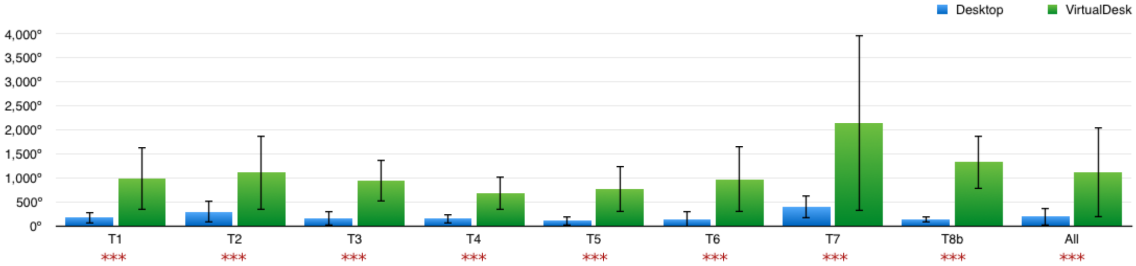


T8b exclusively are averaged per selected point. The error magnitudes were calculated as the absolute differences between the given and expected answers for T1-T6, and as the proportional deviation from the expected answer for T7. Figures 6.3, 6.4 and 6.5 present these results. Since parametric requirements were not met by multiple samples, paired Wilcoxon signed-rank tests were used to determine statistical significances.

Significant differences in time were only found in tasks T5-T8, in which cases the immersive condition was slower. With the exception of T5, all of these were tasks with higher requirement of interaction with the tabletop controls (in T6, the dataset needed to be changed; in T7, the user answer was input through an incremental counter on the desk; in T8, filters should be applied). We believe this is partially related to the fact that some users experienced difficulties with tabletop interaction due to hand sizes (see Subsection 6.5.2). Moreover, the mouse interaction was already expected (H2) to be faster due to its consolidated usage. As opposed to the desktop-based condition, controls and data did not share the user's field of view in the immersive condition, what also required additional time. It is important to note that users were asked to be precise and not fast in tasks T1-T7. Considering task T8b, the point selection time was found to be 43% higher in the VR setup (3.3 vs 2.3s per point), confirming H2.

Hypothesis H1 could be partially accepted, given that tasks T1 (distances between

Figure 6.6: Average accumulated dataset rotations per task question in degrees. This form of exploration was performed 5.8 times more in VirtualDesk, probably due to the intuitiveness of the grabbing action. This increased task accuracy with minimal time overhead.



spheres) and T6 (density comparison over time) obtained significantly smaller error rates in the VirtualDesk condition. When considering error magnitudes, these tasks also presented improvements (82% and 21% reductions, respectively), but, in this case, did not achieve significance (p = .09 in both cases). The VR setup was also never significantly worse than the desktop-based condition in terms of perception (considering either error rates or magnitudes). It was, however, more inaccurate in terms of point selection in task T8b: despite having an extra degree of freedom (DoF), users selected almost three times more unintended points with the virtual finger than with the mouse. We believe this was particularly problematic in cluttered areas of the representation, where it was difficult not to hit adjacent points during selection, especially considering that users had still not mastered the double tap action.

Finally, an interesting difference was observed in terms of dataset rotations. These were performed, on average, 5.8 times more in the immersive condition, probably due to the intuitiveness of the grabbing action. Average accumulated data rotations per task question were 190.4 degrees (SD 179.6) in Desktop and 1,114.8 degrees (SD 935.4) in VirtualDesk. Considering that the observation from different points of view is fundamental in the comprehension of a 3D point cloud, this also partially explains VirtualDesk's advantage in perception tasks such as T1. Figure 6.6 presents results per task. Also note that, despite large differences in T1-T4, these tasks did not present any differences in time. In T7, however, the large difference in exploration between conditions (7.3x) may contribute to explain the difference in completion times. This probably was not reflected on answer accuracy, though, because the task questions turned out to be very easy, with only 2 or 3 clusters per dataset.

6.4.2 User Feedback

The general subjective feedback received from users in post-test interviews was very positive, especially regarding the use of 3D interaction for data manipulation. In terms of usability, both conditions were well rated in the SUS questionnaire. VirtualDesk obtained a 77.2 mean score (standard deviation 16.4) and, Desktop, 72.8 (SD 20.2), but differences were not significant.

For task workload, nonetheless, VirtualDesk's NASA Raw TLX score was significantly higher (*): 30.9 (SD 14.7) compared to 23.2 (SD 15.4). This was especially influenced by two workload components: Physical Workload (37.4 vs 9.7) (***) and Effort (36.1 vs 23.6) (*). This is understandable considering that users were observed to move their left and right hands on average 2.4m (SD 1.2m) and 4.3m (SD 1m) per task question, respectively. Mental Workload was scored at 26.3 (SD 16.2) against 22.2 (SD 20) of Desktop, without statistical difference, partially contradicting H3.

Concerning the immersive environment, SSQ scores were very satisfactory, averaging only 2.18 (SD 9.0), symptoms which can be considered negligible (KENNEDY et al., 2003). No user reported discomfort during or after the tasks. In terms of presence, VirtualDesk was rated in the IPQ (6 points scale) 4.7 (SD 0.88) for Spatial Presence, 4.07 (SD 1.06) for Involvement, 3.11 (SD 0.79) for Experienced Realism and 5.41 (SD 1.17) for the General Item (feeling of "being there"). We provide these results in the expectation of serving as a baseline for future setups. It is important to note that participants were allowed to communicate with the experimenter at any time, keeping them aware of the external environment.

Analysing the users' agreements to different assertions (see Figure 6.7), it becomes clear what were the strengths and weaknesses of our prototype. 46% more participants agreed that it was easy to find information in VirtualDesk. This is probably closely related to the embodied data manipulation, which was not considered difficult by any user. By executing instinctive grabbing and scaling actions, users could easily and rapidly inspect any region of the dataset, as opposed to combining several Rotation-Pan-Dolly actions in the desktop-based version. This was probably what most impressed participants in the experience. Pointing data interaction was rated similarly in both versions, what is very positive considering that the quick double-tap metaphor had just been learned for the experiment, while double-mouse clicking is an universal action. On the other hand, difficulties with the tabletop interaction were the main system weakness: six participants

Figure 6.7: User agreements with different assertions, ranging from completely disagree (dark red) to completely agree (dark green), for Desktop (D) and VirtualDesk (VD). Intuitive embodied data manipulation gestures were well received and allowed easy and rapid inspection and information finding in any region of the dataset.

experienced difficulties due to their real hands being larger than the fixed-size virtual model employed (see Subsection 6.5.2).

In a ranking question after the completion of both parts of the test, VirtualDesk was selected by all participants as the most engaging condition, and by 21 (87.5%) as the most intuitive. This was already expected, and is partially related to the novelty of VR, but also to the experienced immersion and the use of natural gestures for interaction. More importantly, VirtualDesk was perceived by 15 participants (62.5%) as the fastest technique. Both conditions tied in terms of accuracy, with 12 users choosing each. When asked, many reported that Desktop was most accurate for selection, but VirtualDesk for manipulation.

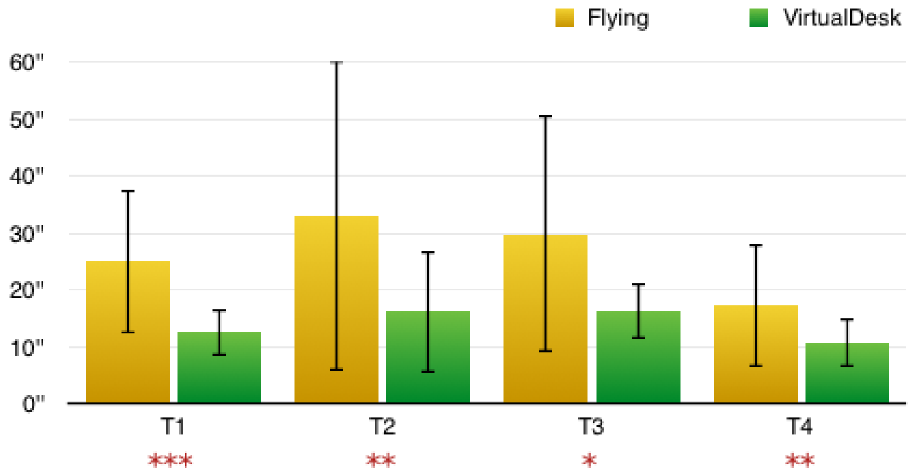
6.4.3 Comparison with Flying Navigation

Figure 6.8 contrasts results between the new and old paradigms for T1-T4 (tasks present in both studies), in terms of completion times (a), error rates (b) and error magnitudes (c). Given that questions for these tasks are not repeated more than once in each condition, only the first 24 users from the previous study and the same dataset are considered, since it would be unfair to consider ones who performed potentially easier or more difficult questions. Mann–Whitney U tests for independent samples were used to compare results.

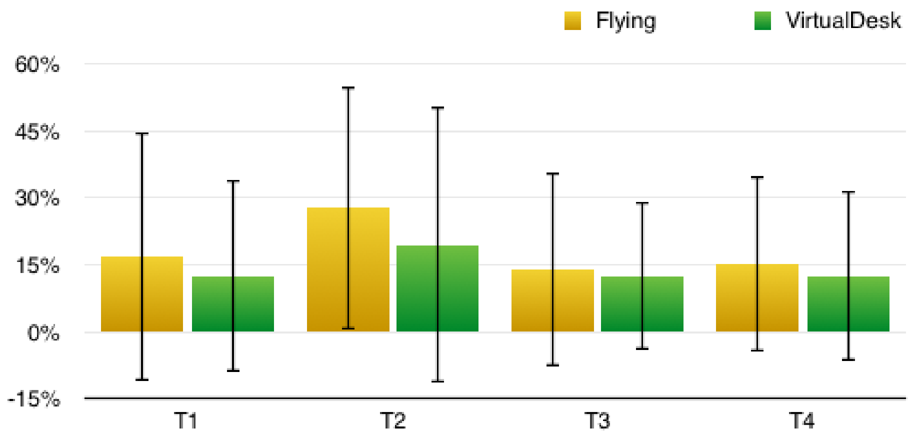
As expected, VirtualDesk was more time-efficient, and all tasks were performed significantly faster, reaching a 51% improvement in T2. In terms of task performance,

Figure 6.8: Comparison between VirtualDesk and the previous implementation employing Flying navigation. Embodied data manipulation resulted in up to 51% shorter average completion times, 30% smaller error rates and 68% smaller error magnitudes.

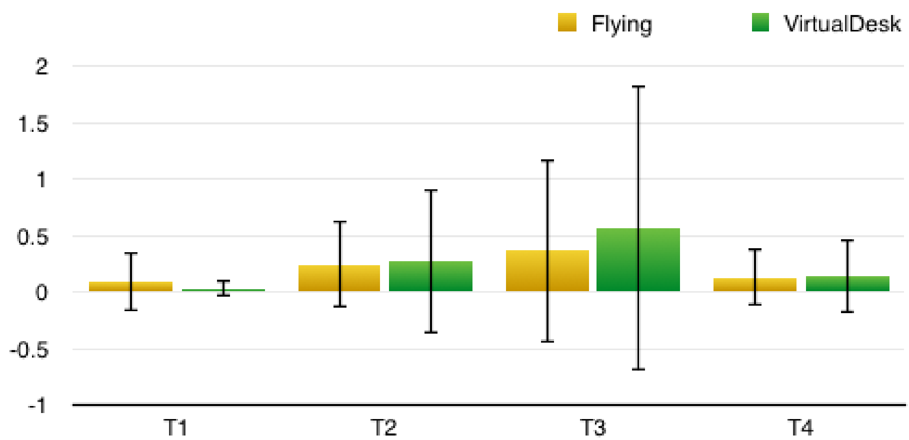
(a) Time



(b) Error rates



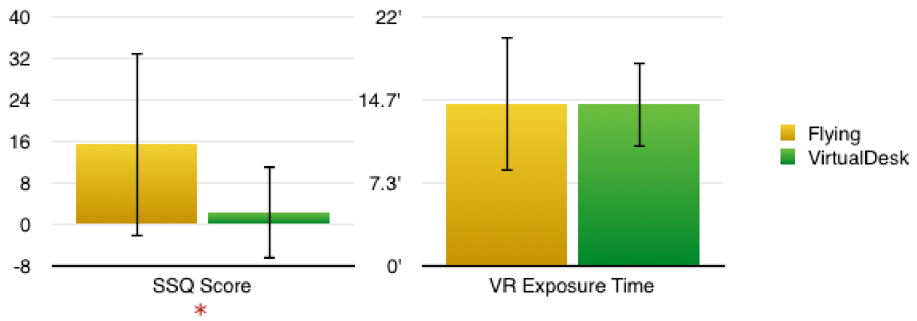
(c) Error magnitudes



collected data has no statistical significance. Although all tasks consistently achieved lower error rates under the new approach (25%, 30%, 10% and 18% reductions, respectively), we are aware that this can be due to a random factor. Just as a speculation, we attribute this to the added notion of proprioception when using embodied manipulation, and the stronger stereopsis and head-motion parallax cues at short distances, as discussed by Mine (MINE; JR; SEQUIN, 1997). However, new tests should be conducted to verify this. In terms of the average error magnitudes, T1 presented a 68% reduction under VirtualDesk, and T2 and T4 were very similar in both approaches. T3, on the other hand, presented a 54% increase. Again, none of the differences were statistically significant.

Another key result, in our opinion, is shown in Figure 6.9. Despite very similar VR exposure times in both studies, the average SSQ score in VirtualDesk was 7 times lower than in the artificially navigated version. Moreover, while in that study 40% of the users had experienced very significant discomfort levels (scores >= 20), now 83% perceived only negligible or minimal symptoms, and the maximum individual score was 18.3. This completed the confirmation of hypothesis H4.

Figure 6.9: Due to its more natural and comfortable navigation paradigm, VirtualDesk achieved a 7x smaller SSQ score than the previous Flying approach, despite very similar VR exposure times.



6.5 Discussion

6.5.1 Findings

Results from this user study confirmed our main intuitions in the proposal of the VirtualDesk approach. VirtualDesk performed equally well or better across all analytical tasks, both in comparison to a standard Desktop interface and to the previous immersive implementation with flying navigation (confirming hypotheses H1 and H4). The added time overhead with relation to Desktop was only significant in tasks with higher requirements for tabletop interaction (which required viewpoint change and also imposed certain difficulties for some users), and was generally only a few seconds. This was despite the fact that data exploration in terms of dataset rotation was found to be 5.8 times higher in the immersive condition.

Tasks T1 (identification of the closest point) and T6 (density comparison over time) in particular presented significant error rate decrease under immersive data exploration. We believe this was related, respectively, to the easier inspection of local areas using 3D interaction, and to a possibly longer persistent obtained mental model of the data in the immersive condition.

The desktop-based 3D condition also performed well across tasks, as had been already observed in our previous study. As discussed by Ware, structure-from-motion cues enable the perception of point positions even without stereopsis. We are convinced that the comparison between the two implementations was fair, and most participants reported that each condition had its pros and cons. In particular, interaction tasks in this condition were still found to be quicker and less error-prone (as expected in H2).

Subjective feedback indicated that VirtualDesk was perceived as quicker, more intuitive and engaging, partially confirming H3. Nonetheless, the mental workload, as measured by the NASA Raw TLX questionnaire, did not present significant variation, and overall workload increased due to higher inherent physical workload and perceived required effort to achieve the task goals.

In terms of interaction gestures, one of our main mistakes, in retrospect, was to assign different selection behaviours to the left and right hands. Despite being familiarized with them in the tutorial phase, even right-handed users intuitively constantly tried to select points with their left hands when they were closer. Meanwhile, the double tap gesture for point selection (as opposed to some controller-dependent action such as button clicking (BACH et al., 2017)), though difficult to master at the beginning for many users, did not affect the correspondent ratings (Figure 6.7), and we believe that, in the long term, would be more intuitive and efficient and reduce workload.

6.5.2 Limitations

The main limitation of the present user evaluation study was the fact that it has only been performed in one specific use case (multidimensional roll call data projected to 3D) and one information representation (point clouds). Nonetheless, we believe this was adequate for our current purposes, which were to investigate and demonstrate the potential of a so-far atypical immersive data exploration paradigm, rather than to propose its mediate adoption in data analysis. It is also important to emphasize that our evaluation is admittedly only concerned with the benefits of different factors, such as stereopsis, tangible interaction, proprioception and embodied data manipulation when combined, and not individually, what could also be assessed in future specific studies.

Considering the prototype implementation, the main identified limitation was that the virtual hand models, obtained from the Unity Oculus Integration Package, were not adjusted accordingly to the participants' real hand sizes. This resulted in difficulties for at least six users who had larger hands and faced difficulties to reach the virtual desk surface despite being touching the real desk. This was always solved by slightly changing the controller position in the user's hand, but negatively affected their overall perception of interaction ease (see Figure 6.7) and partially compromised the evaluation of this aspect of the prototype. We intent to circumvent this limitation in a future version. Nevertheless, we believe the choice for the Oculus Touch controllers instead of other hand tracking solutions was appropriate, and resulted in highly realistic and precise modelling of the hands and hand gestures, what contributed to immersion and user engagement.

Another key limitation in our comparative study is the limited training provided to users in the new technique, as opposed to the ubiquitous familiarity of 2D user interaction, a common issue in the evaluation of novel approaches. Note that only 8% of our users had average previous experience with VR HMDs. After working on this prototype for several months, we are convinced that, upon longer training, interaction times and error rates in tasks T8a and T8b become much lower. In order to demonstrate this, however, long term evaluations will be needed. One such attempt was recently shown by Bach et al. (2017), who reassessed 6 participants across five daily sessions. They observed speed improvements in 22 out of the 24 task-participant combinations, being 9 with significance. However, no significant improvements in precision were observed, and they noted that 5 sessions might have been too few. Finally, our user comfort results refer to an average 14.3 min VR session, and it is still unknown how this would change for longer exposures, also requiring further studies.

6.6 Summary

In order to evaluate the proposed VirtualDesk approach, a new comparative user study was conducted. Results showed that error rates in a series of perception tasks were always equal or lower than in a conventional desktop interface and the previous immersive implementation with flying navigation. The immersive environment also contributed to higher subjective perceptions of efficiency and engagement and much higher data exploration, while incurring minimal time overhead and generating almost no simulator sickness symptoms.

7 CONCLUSIONS AND FUTURE WORK

In this work, in an effort to extend discussions about Immersive Analytics, we presented evaluations on a particular representation commonly for multidimensional data: 3D dimensionally-reduced data scatterplots. Our main motivation was to demonstrate that current off-the-shelf VR technologies may effectively aid in analytical tasks performed on abstract information visualization, even challenging previous beliefs about three-dimensional representations.

7.1 Contributions

We modelled the overall visualization task error in this scenario as the result of the combination between the error introduced by dimensionality reduction and the one introduced by human perception. Through a task-based empirical approach, we simulated the user maximum possible performance assuming the represented distances were perfectly understood, and selected two different datasets in the domain of roll call analysis: one with promising information gain in 3D and one without such gain.

A first user evaluation was then conducted to compare the task performance in desktop-based and HMD-based visualization conditions, and thus evaluate the second factor in our model. Surprisingly, however, no perception differences were observed, and similarly low errors in all conditions resulted in improvements with the addition of the third dimension with or without immersion when the dataset enabled so. In retrospect, we attribute this to implementation choices for the desktop-based 3D scenario, which was not representative of a typical 3D visualization toolkit. Even so, on further inspection of subjective results, it was found that the HMD-based condition had required smaller effort to find information and less navigation, besides offering a much larger subjective perception of accuracy and engagement. Its main limitations, on the other hand, were high reported levels of user discomfort and high task completion times.

In a second moment, we used these initial results to implement a different exploration paradigm, more fit for real usage. Our proposed metaphor, VirtualDesk, combined features from different backgrounds:

1. The dataset representation is rendered in smaller scale at arm's reach, to better benefit from proprioception, more precise body-relative interaction and stronger

stereopsis and head-motion parallax (MINE; JR; SEQUIN, 1997).

2. Embodied natural data manipulation conforms with the recent concept of spatio-data coordination, i.e. a one-to-one mapping between physical and virtual actions, aiming to lower the user cognitive load (CORDEIL et al., 2017a).

3. Following Zielasko et al. (2017), a virtual desk is represented, synchronised with the analyst's real one. This allows tangible interaction with controls and 2D coordinated views, placed on the desk's surface.

In a new comparative user study, error rates in a series of perception tasks were always equal or lower than in the equivalent conventional desktop interface and also the previous immersive implementation, which employed flying navigation. This new immersive environment also contributed to higher subjective perceptions of efficiency and engagement and much higher data exploration in terms of measured dataset rotations.

Furthermore, this metaphor is particularly promising due to its observed results for user comfort and relatively short required completion times. Considering existent implementations for immersive exploration of abstract information (see Chapter 2), most are based in the metaphors of flying through a dataset or walking around it. The former, while providing an egocentric view of the data, commonly results in simulator sickness due to the conflict with the user's real position. The latter successfully avoids this problem, but is very costly in terms of required time and space. VirtualDesk, on the other hand, achieved a very low score for simulator sickness while remaining time-efficient and easily integrable into an analyst's work environment (ZIELASKO et al., 2017). This makes it more convenient for real world usage, requiring only minor improvements.

7.2 Publications

Parts of this dissertation work have been published in the following papers.

- Jorge Alberto Wagner Filho, Marina F. Rey, Carla M.D.S. Freitas, Luciana Nedel (2017). **Immersive Analytics of Dimensionally-Reduced Data Scatterplots.** IEEE VIS Workshop on Immersive Analytics: Exploring Future Interaction and Visualization Technologies for Data Analytics. (*Chapter 4*)

- Jorge Alberto Wagner Filho, Marina F. Rey, Carla M.D.S. Freitas, Luciana Nedel (2018). **Immersive Visualization of Abstract Information: An Evaluation on**

Dimensionally-Reduced Data Scatterplots. 25th IEEE Conference on Virtual Reality and 3D User Interfaces. (*Chapters 3 and 4*)

- Jorge Alberto Wagner Filho, Carla M.D.S. Freitas, Luciana Nedel (2018). **VirtualDesk: A Comfortable and Efficient Immersive Information Visualization Approach.** Under submission. (*Chapters 5 and 6*)

7.3 Future works

Future works include the improvement of the VirtualDesk prototype based on the user study participants' feedback, and further testing under different conditions, including different datasets and representations. We believe that strong candidates to benefit from this exploration metaphor would be node-link diagrams and space-time cube representations. Long term tests and longer VR exposure times would also be important to assess for the real applicability of these techniques.

Moreover, despite being used as two alternative conditions in our study, we believe that one of the main perspectives for the VirtualDesk prototype is its combination with the non-immersive counterpart implementation. This paradigm allows for a direct mapping between all immersive environment contents and a two-panels 2D interface, which reproduce a 2D projection of the 3D data in one part, and the surface of the desk in the other. Combining immersive and conventional data exploration environments becomes thus straightforward. This way, for example, annotations introduced in VirtualDesk could easily be persisted for further inspection in the monocular display if necessary.

Lastly, we encourage further studies of alternative proposals, and provided here multiple results which can be used as baselines, for example, when assessing standard questionnaires for presence and user comfort.

REFERENCES

ARMS, L.; COOK, D.; CRUZ-NEIRA, C. The benefits of statistical visualization in an immersive environment. In: IEEE. **Proc. IEEE Virtual Reality**. [S.l.], 1999. p. 88–95. ISSN 1087-8270.

BABAEE, M.; DATCU, M.; RIGOLL, G. Assessment of dimensionality reduction based on communication channel model; application to immersive information visualization. In: IEEE. **2013 IEEE International Conference on Big Data**. [S.l.], 2013. p. 1–6.

BACH, B. et al. The hologram in my hand: How effective is interactive exploration of 3d visualizations in immersive tangible augmented reality? **IEEE transactions on visualization and computer graphics**, IEEE, 2017.

BELLGARDT, M. et al. Utilizing immersive virtual reality in everyday work. In: **3rd Workshop on Everyday Virtual Reality**. [S.l.: s.n.], 2017.

BORJA, F. G. d. Civisanalysis: exploring representatives' voting behaviour. 2017.

BORJA, F. G. de; FREITAS, C. M. Civisanalysis: Interactive visualization for exploring roll call data and representatives' voting behaviour. In: IEEE. **Proc. SIBGRAPI**. [S.l.], 2015. p. 257–264.

BOWMAN, D. et al. **3D User Interfaces: Theory and Practice, CourseSmart eTextbook**. [S.l.]: Addison-Wesley, 2004.

BRIGADIR, I. et al. Dimensionality reduction and visualisation tools for voting record. In: CEUR WORKSHOP PROCEEDINGS. **24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16), University College Dublin, Ireland, 20-21 September 2016**. [S.l.], 2016.

BROOKE, J. et al. Sus-a quick and dirty usability scale. **Usability Evaluation in Industry**, London, United kingdom, v. 189, n. 194, p. 4–7, nov 1996.

BRYSON, S.; LEVIT, C. The virtual wind tunnel. **IEEE Computer Graphics and Applications**, v. 12, n. 4, p. 25–34, July 1992. ISSN 0272-1716.

BURGESS, R. et al. Selection of large-scale 3d point cloud data using gesture recognition. In: SPRINGER. **Doctoral Conference on Computing, Electrical and Industrial Systems**. [S.l.], 2015. p. 188–195.

CARROLL, R.; POOLE, K. Roll call analysis and the study of legislatures. **The Oxford handbook of legislative studies**, OUP Oxford, p. 103, 2014.

CHANDLER, T. et al. Immersive analytics. In: IEEE. **2015 Big Data Visual Analytics (BDVA)**. [S.l.], 2015. p. 1–8.

COFFEY, D. et al. Slice wim: A multi-surface, multi-touch interface for overview+detail exploration of volume datasets in virtual reality. In: ACM. **Symposium on Interactive 3D Graphics and Games**. New York, NY, USA, 2011. p. 191–198.

CORDEIL, M. et al. A design space for spatio-data coordination: Tangible interaction devices for immersive information visualisation. In: **Proceedings of IEEE Pacific Visualization Symposium (Pacific Vis)**. [S.l.: s.n.], 2017.

CORDEIL, M. et al. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In: ACM. **Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology**. [S.l.], 2017. p. 71–83.

CORDEIL, M. et al. Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display? **IEEE transactions on visualization and computer graphics**, IEEE, v. 23, n. 1, p. 441–450, Jan 2017. ISSN 1077-2626.

CUNNINGHAM, J. P.; GHAHRAMANI, Z. Linear Dimensionality Reduction: Survey, Insights, and Generalizations. v. 16, p. 2859–2900, 2015. ISSN 15337928. Available from Internet: <http://arxiv.org/abs/1406.0873>. Accessed mar. 07, 2018.

DONALEK, C. et al. Immersive and collaborative data visualization using virtual reality platforms. In: IEEE. **2014 IEEE International Conference on Big Data (Big Data)**. [S.l.], 2014. p. 609–614.

DROUHARD, M. et al. Immersive visualization for materials science data analysis using the oculus rift. In: IEEE. **2015 IEEE International Conference on Big Data (Big Data)**. [S.l.], 2015. p. 2453–2461.

ETEMADPOUR, R. et al. A user-centric taxonomy for multidimensional data projection tasks. In: **IVAPP**. [S.l.: s.n.], 2015. p. 51–62.

ETEMADPOUR, R.; MONSON, E.; LINSEN, L. The effect of stereoscopic immersive environments on projection-based multi-dimensional data visualization. In: **IV**. [S.l.: s.n.], 2013. p. 389–397.

ETEMADPOUR, R. et al. Perception-based evaluation of projection methods for multidimensional data visualization. **IEEE transactions on visualization and computer graphics**, IEEE, v. 21, n. 1, p. 81–94, 2015.

FISHERKELLER, M. A.; FRIEDMAN, J. H.; TUKEY, J. W. An interactive multidimensional data display and analysis system. 1974.

GARCÍA-HERNÁNDEZ, R. J. et al. Perspectives for using virtual reality to extend visual data mining in information visualization. In: IEEE. **2016 IEEE Aerospace Conference**. [S.l.], 2016. p. 1–11.

GOLUB, G. H.; REINSCH, C. Singular value decomposition and least squares solutions. **Numerische Mathematik**, Springer-Verlag New York, Inc., v. 14, n. 5, p. 403–420, apr 1970. ISSN 0029-599X.

GRACIA, A. et al. New insights into the suitability of the third dimension for visualizing multivariate/multidimensional data: A study based on loss of quality quantification. **Information Visualization**, SAGE Publications Sage UK: London, England, v. 15, n. 1, p. 3–30, 2016.

GRAY, G. **Navigating 3d Scatter Plots in Immersive Virtual Reality**. Thesis (PhD) — University of Washington, 2016.

GREFFARD, N.; PICAROUGNE, F.; KUNTZ, P. Beyond the classical monoscopic 3d in graph analytics: an experimental study of the impact of stereoscopy. In: IEEE. **3DVis (3DVis), 2014 IEEE VIS International Workshop on**. [S.l.], 2014. p. 19–24.

GUNA, J. et al. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 14, n. 2, p. 3702–3720, 2014.

HALPIN, H. et al. Exploring semantic social networks using virtual reality. In: SPRINGER. **International Semantic Web Conference**. Berlin, Heidelberg, 2008. p. 599–614.

HART, S. G. Nasa-task load index (nasa-tlx); 20 years later. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. **Proceedings of the human factors and ergonomics society annual meeting**. [S.l.], 2006. v. 50, n. 9, p. 904–908.

HOLLANDER, M.; WOLFE, D. A.; CHICKEN, E. **Nonparametric statistical methods**. [S.l.]: John Wiley & Sons, 2013.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of educational psychology**, Warwick & York, v. 24, n. 6, p. 417, 1933.

HUBER, P. J. Experiences with three-dimensional scatterplots. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 82, n. 398, p. 448–453, 1987.

INGRAM, S. et al. Dimstiller: Workflows for dimensional analysis and reduction. In: IEEE. **Proc. VAST**. [S.l.], 2010. p. 3–10.

JAKULIN, A.; BUNTINE, W. Analyzing the us senate in 2003: Similarities, networks, clusters and blocs. **Preprint. Available at http://kt. ijs. si/aleks/Politics/us_senate. pdf**, dec 2004.

JANKOWSKI, J.; HACHET, M. Advances in interaction with 3d environments. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2015. v. 34, n. 1, p. 152–190.

KENNEDY, R. S. et al. Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: Similarities and differences. **Virtual and adaptive environments: Applications, implications, and human performance issues**, CRC Press, p. 247, 2003.

KENNEDY, R. S. et al. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. **The International Journal of Aviation Psychology**, Taylor & Francis, v. 3, n. 3, p. 203–220, jul 1993.

KOSARA, R. et al. Thoughts on user studies: Why, how, and when. **IEEE Computer Graphics and Applications**, v. 23, n. 4, p. 20–25, 2003.

KUHLEN, T. W.; HENTSCHEL, B. Quo vadis cave: does immersive visualization still matter? **IEEE computer graphics and applications**, IEEE, v. 34, n. 5, p. 14–21, Sept 2014. ISSN 0272-1716.

KWON, O.-H. et al. A study of layout, rendering, and interaction methods for immersive graph visualization. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 22, n. 7, p. 1802–1815, Jul 2016. ISSN 1077-2626.

MAATEN, L. v. d.; HINTON, G. Visualizing data using t-sne. **Journal of Machine Learning Research**, v. 9, n. Nov, p. 2579–2605, 2008.

MANJREKAR, S. et al. Cave: An emerging immersive technology–a review. In: IEEE. **Computer Modelling and Simulation (UKSim), 2014 UKSim-AMSS 16th International Conference on**. [S.l.], 2014. p. 131–136.

MILGRAM, P.; KISHINO, F. A taxonomy of mixed reality visual displays. **IEICE TRANSACTIONS on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 77, n. 12, p. 1321–1329, 1994.

MINE, M. R. Virtual environment interaction techniques. **UNC Chapel Hill CS Dept**, 1995.

MINE, M. R.; JR, F. P. B.; SEQUIN, C. H. Moving objects in space: exploiting proprioception in virtual-environment interaction. In: ACM PRESS/ADDISON-WESLEY PUBLISHING CO. **Proceedings of the 24th annual conference on Computer graphics and interactive techniques**. [S.l.], 1997. p. 19–26.

MORAN, A. et al. Improving big data visual analytics with interactive virtual reality. In: IEEE. **High Performance Extreme Computing Conference (HPEC), 2015 IEEE**. [S.l.], 2015. p. 1–6.

MUNZNER, T. **Visualization analysis and design**. [S.l.]: CRC press, 2014.

NGUYEN, H.; MARENDY, P.; ENGELKE, U. Collaborative framework design for immersive analytics. In: IEEE. **Big Data Visual Analytics (BDVA), 2016**. [S.l.], 2016. p. 1–8.

PELLEG, D.; MOORE, A. W. X-means: Extending k-means with efficient estimation of the number of clusters. In: **Seventeenth International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 2000. p. 727–734.

POCO, J. et al. A framework for exploring multidimensional data with 3d projections. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2011. v. 30, n. 3, p. 1111–1120.

RAJA, D. **The Effects of Immersion on 3D Information Visualization**. Thesis (PhD) — Virginia Polytechnic Institute and State University, 2006.

RAJA, D. et al. Exploring the benefits of immersion in abstract information visualization. In: **Proc. Immersive Projection Technology Workshop**. [S.l.: s.n.], 2004. p. 61–69.

ROBERTS, J. C. et al. Visualization beyond the desktop–the next big thing. **IEEE Computer Graphics and Applications**, IEEE, v. 34, n. 6, p. 26–34, Nov 2014. ISSN 0272-1716.

SARIKAYA, A.; GLEICHER, M. Scatterplots: Tasks, data, and designs. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, 2017.

SCHUBERT, T.; FRIEDMANN, F.; REGENBRECHT, H. The experience of presence: Factor analytic insights. **Presence: Teleoperators and virtual environments**, MIT Press, v. 10, n. 3, p. 266–281, 2001.

SEDLMAIR, M. et al. Dimensionality reduction in the wild: Gaps and guidance. **Dept. Comput. Sci., Univ. British Columbia, Vancouver, BC, Canada, Tech. Rep. TR-2012-03**, 2012.

SEDLMAIR, M.; MUNZNER, T.; TORY, M. Empirical guidance on scatterplot and dimension reduction technique choices. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 19, n. 12, p. 2634–2643, Dec 2013. ISSN 1077-2626.

SHNEIDERMAN, B. The eyes have it: A task by data type taxonomy for information visualizations. In: IEEE. **Visual Languages, 1996. Proceedings., IEEE Symposium on**. [S.l.], 1996. p. 336–343.

SHOVMAN, M. et al. Twist and learn: Interface learning in 3dof exploration of 3d scatterplots. In: ACM. **Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems**. [S.l.], 2015. p. 313–316.

SIMEONE, A. L.; VELLOSO, E.; GELLERSEN, H. Substitutional reality: Using the physical environment to design virtual reality experiences. In: ACM. **Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems**. New York, NY, USA, 2015. p. 3307–3316. ISBN 978-1-4503-3145-6.

SIMPSON, M.; ZHAO, J.; KLIPPEL, A. Take a walk: Evaluating movement types for data visualization in immersive virtual reality. In: **2nd Workshop on Immersive Analytics**. [S.l.]: IEEE, 2017.

SPIRLING, A.; MCLEAN, I. The rights and wrongs of roll calls. **Government and Opposition**, Wiley Online Library, v. 41, n. 4, p. 581–588, 2006. ISSN 1477-7053.

STENHOLT, R. On the benefits of using constant visual angle glyphs in interactive exploration of 3d scatterplots. **Proc. ACM TAP**, ACM, New York, NY, USA, v. 11, n. 4, p. 19, dec 2014.

THEART, R. P.; LOOS, B.; NIESLER, T. R. Virtual reality assisted microscopy data visualization and colocalization analysis. **BMC bioinformatics**, Springer, v. 18, n. 2, p. 64, 2017.

TORGERSON, W. S. Multidimensional scaling: I. theory and method. **Psychometrika**, Springer, v. 17, n. 4, p. 401–419, Dec 1952. ISSN 18600980.

USHER, W. et al. A virtual reality visualization tool for neuron tracing. **IEEE transactions on visualization and computer graphics**, IEEE, 2017.

WARD, M. O.; GRINSTEIN, G.; KEIM, D. **Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition - 360 Degree Business**. Natick, MA, USA: A. K. Peters, Ltd., 2015.

WARE, C. **Information Visualization: Perception for Design**. [S.l.]: Elsevier, 2012.

WARE, C.; MITCHELL, P. Visualizing graphs in three dimensions. **ACM Transactions on Applied Perception (TAP)**, ACM, v. 5, n. 1, p. 2, 2008.

WATTENBERG, M.; VIEGAS, F.; JOHNSON, I. How to use t-sne effectively. **Distill**, 2016. Available from Internet: <http://distill.pub/2016/misread-tsne>. Accessed mar. 07, 2018.

YAO, R. et al. Oculus vr best practices guide. **Oculus VR**, 2014.

ZIELASKO, D. et al. buenosdias: Supporting desktop immersive analytics while actively preventing cybersickness. In: **2nd Workshop on Immersive Analytics**. [S.l.]: IEEE, 2017.

ZIELASKO, D. et al. Evaluation of hands-free hmd-based navigation techniques for immersive data analysis. In: IEEE. **2016 IEEE Symposium on 3D User Interfaces (3DUI)**. [S.l.], 2016. p. 113–119.

ZIELASKO, D. et al. Remain seated: towards fully-immersive desktop vr. In: IEEE. **2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)**. [S.l.], 2017. p. 1–6.

APPENDIX A — AVALIANDO ABORDAGENS IMERSIVAS PARA VISUALIZAÇÃO DE INFORMAÇÕES MULTIDIMENSIONAIS

A.1 Introdução

Representações tridimensionais de dados sabidamente oferecem vantagens sob muitas circunstâncias. No caso de dados inerentemente espaciais, elas contribuem para a construção mais rápida de um modelo mental. Para informações abstratas, sua utilidade também já foi demonstrada, permitindo melhor separação espacial em grandes grafos, melhor detecção de padrões tri-variados em *scatterplots* e a projeção mais fiel de dados multidimensionais. No entanto, o seu uso também é controverso, principalmente pois tais representações sofrem com a ocorrência de questões perceptuais conhecidas, como distorção perspectiva, escorço e oclusão, que tornam sua exploração trabalhosa e sujeita a erros. Além disso, há um desencontro entre as visualizações 3D e os dispositivos convencionais de interação 2D, tais quais o mouse. Após consecutivas inovações nas pesquisas em Realidade Virtual, no entanto, a comunidade de Visualização passou a progressivamente explorar a adoção de abordagens imersivas, que combinam *displays* estereoscópicos e interação natural, para potencialmente modificar este cenário.

Neste trabalho, nosso principal objetivo é expandir esta discussão, levando em conta um tipo específico de representação de informações abstratas, *scatterplots* 3D obtidos a partir da redução de dimensionalidade de dados. Em tese, uma dimensão extra permite diminuir a perda de informação neste processo e obter uma representação mais fidedigna dos dados. Além disso, como esta categoria de *scatterplots*, comumente aplicada para a visualização de dados multidimensionais, é sempre analisada em função das distâncias entre os pontos, hipotetizamos que ela seria beneficiada por *displays* estereoscópicos, pontos de vista egocêntricos e interfaces de usuário mais naturais, características inerentes de configurações imersivas de análise. Nós focamos nossa questão de pesquisa especificamente no estudo da aplicação de ambientes baseados em *head-mounted displays* (HMDs) para este fim.

A.2 Trabalhos Relacionados

Análise Imersiva é uma área crescente na convergência entre as pesquisas em Visualização e Realidade Virtual, responsável por aplicar novas tecnologias de *display* e interação combinadamente para suportar o raciocínio analítico (CHANDLER et al., 2015). Tais abordagens já obtiveram sucesso considerável no caso da análise de dados espaciais científicos. No entanto, para a visualização de informações abstratas, novas pesquisas e diretrizes ainda são necessárias (GARCÍA-HERNÁNDEZ et al., 2016). Alguns resultados promissores foram apresentados, por exemplo, em estudos relativos à diagramas de grafos (HALPIN et al., 2008; KWON et al., 2016).

Em relação a *scatterplots* 3D obtidos por redução de dimensionalidade, trabalhos prévios utilizando *displays* monoscópicos apresentaram resultados contraditórios (POCO et al., 2011; SEDLMAIR; MUNZNER; TORY, 2013; GRACIA et al., 2016). Considerando *displays* imersivos, alguns autores já apresentaram esforços semelhantes ao nosso, mas a maioria destes se baseou em tecnologias que progrediram substancialmente nos últimos anos, e trabalharam sempre com ambientes do tipo CAVE (ARMS; COOK; CRUZ-NEIRA, 1999; RAJA et al., 2004; ETEMADPOUR; MONSON; LINSEN, 2013). Aqui, propomos investigar abordagens baseadas em HMD para este problema, tendo em vista que estes dispositivos apresentam um custo muito mais acessível e, recentemente, foram demonstrados inclusive como mais eficientes no caso de tarefas de conectividade de grafos (CORDEIL et al., 2017c).

A.3 *Framework* de Avaliação Baseado em Tarefas

Propomos, neste trabalho, uma modelagem formal para avaliar a contribuição de abordagens 3D imersivas para a exploração de um dado conjunto de dados. Este modelo identifica os dois fatores separados que determinam o desempenho total nas tarefas: a diferença dos erros introduzidos ao se realizar a redução de dimensionalidade para duas ou três dimensões, e a diferença dos erros de percepção humana sob diferentes condições de visualização.

Diferentes métodos são propostos para avaliar cada um destes erros. Uma abordagem de simulação baseada em tarefas quantifica o máximo desempenho possível que um usuário poderia obter em cada tarefa utilizando 2D ou 3D, fosse ele capaz de perceber perfeitamente as distâncias representadas. Estudos comparativos com usuários, por sua vez, são empregados na avaliação dos erros de percepção.

Para tanto, escolhemos, como caso de uso, quatro conjuntos de dados multidimensionais referentes a períodos de quatro anos de votações na Câmara dos Deputados do Brasil, e um conjunto coletado de nove diferentes tarefas analíticas, relacionadas a percepção de distâncias, percepção de densidades, identificação de agrupamentos e interação.

A.4 Estudo com Usuários 1: Navegação Convencional por Voo

Em um estudo comparativo com usuários com três diferentes condições (2D baseado em *desktop*, 3D baseado em *desktop* e 3D baseado em HMD), observamos que os erros de percepção foram similarmente baixos sob todas elas. Assim sendo, o desempenho nas tarefas foi melhor com a adição da terceira dimensão independentemente da imersão, quando os dados assim permitiram. No entanto, a condição baseada em HMD requereu menor esforço para encontrar as informações e menos navegação, além de oferecer percepções subjetivas muito maiores de precisão e engajamento. Uma limitação para a abordagem imersiva proposta, no entanto, foi a alta incidência de enjoo de simulador, com cerca de 40% dos participantes relatando níveis significativos de desconforto.

A.5 VirtualDesk: Uma Nova Proposta para Visualização Imersiva de Informações Mais Confortável e Eficiente

Uma nova abordagem para exploração imersiva de dados 3D foi proposta visando contornar as limitações previamente observadas em relação a altos tempos de completude de tarefas e altas incidências de enjoo de simulador. O protótipo VirtualDesk é baseado principalmente na interação e manipulação incorporada natural com dados renderizados ao alcance dos braços, e interação tangível com controles e visões coordenadas 2D posicionadas na superfície de uma mesa virtual cuja posição é sincronizada com a da mesa real do analista.

A.6 Estudo com Usuários 2: Avaliação da VirtualDesk

Para avaliar a abordagem VirtualDesk proposta, um novo estudo comparativo com usuários foi conduzido. Os resultados obtidos mostraram que as taxas de erro em uma série de tarefas de percepção foram sempre iguais ou menores às de uma interface *desktop* convencional e às da implementação imersiva anterior baseada em navegação por voo. O ambiente imersivo também contribuiu para maiores percepções subjetivas de eficiência e engajamento, e muito mais exploração dos dados, ao passo que incorreu em tempo adicional mínimo e praticamente não gerou nenhum sintoma de enjoo de simulador nos participantes.

A.7 Conclusão

Levando em conta os bons resultados obtidos pela VirtualDesk para conforto e tempos de completude, acreditamos que esta seja uma abordagem promissora para exploração imersiva de informações abstratas, mostrando-se de fato conveniente para uso real e requerendo apenas pequenos ajustes.

Possíveis trabalhos futuros incluem o aperfeiçoamento do protótipo a partir dos comentários dos participantes do estudo, e novos testes sob diferentes condições, incluindo diferentes conjuntos de dados e representações. Nós acreditamos que fortes candidatos a se beneficiarem desta metáfora de exploração seriam diagramas de grafos e cubos espaço-temporais. Testes de longo prazo e maiores tempos de exposição à Realidade Virtual também seriam importantes para avaliar a real aplicabilidade do sistema.

Além disso, embora tenham sido usadas como condições alternativas no nosso estudo, acreditamos que uma das principais perspectivas para o protótipo VirtualDesk é a sua combinação com a sua implementação não-imersiva correspondente. Este paradigma permite um mapeamento direto entre todos conteúdos do ambiente imersivo e uma interface 2D de dois painéis, que reproduzem uma projeção 2D dos dados 3D em uma parte, e a superfície da mesa virtual na outra. Combinar ambientes imersivos e convencionais de exploração de dados se torna, assim, muito simples. Desta forma, por exemplo, anotações introduzidas na VirtualDesk poderiam ser persistidas para posterior inspeção no *display* monocular se necessário.

Por fim, encorajamos também a realização de estudos de propostas alternativas, e oferecemos aqui resultados que podem ser utilizados como *baselines*.