

IPCP: Immersive Parallel Coordinates Plots for Engineering Design Processes

Slawomir K. Tadeja*, Timoleon Kipouros†, Per Ola Kristensson‡

University of Cambridge, Cambridge, CB2 1PZ, United Kingdom

Computational engineering design methods and tools are common practice in modern industry. Such approaches are integral in enabling designers to efficiently explore larger and more complex design spaces. However, at the same time, computational engineering design methods tend to dramatically increase the number of candidate solutions that decision-makers must interpret in order to make appropriate choices within a set of solutions. Since all candidate solutions can be represented in digital form together with their assessment criteria, evaluated according to some sort of simulation model, a natural way to explore and understand the complexities of the design problem is to visualize their multidimensional nature. The task now involves the discovery of patterns and trends within the multidimensional design space. In this work, we aim to enhance the design decision-making process by embedding visual analytics into an immersive virtual reality environment. To this end, we present a system called IPCP: immersive parallel coordinates plots. IPCP combines the well-established parallel coordinates visualization technique for high-dimensional data with immersive virtual reality. We propose this approach in order to exploit and discover efficient means to use new technology within a conventional decision-making process. The aim is to provide benefits by enhancing visualizations of 3D geometry and other physical quantities with scientific information. We present the design of this system, which allows the representation and exploration of multidimensional scientific datasets. A qualitative evaluation with two surrogate expert users, knowledgeable in multidimensional data analysis, demonstrate that the system can be used successfully to detect both known and previously unknown patterns in a real-world test dataset, producing an early indicative validation of its suitability for decision support in engineering design processes.

I. Introduction

Computational design methodologies and special optimization tools are common practice in modern engineering. The primary benefit of such approaches is the automated exploration of complex design spaces, which was previously not possible. However, this advantage creates user interface challenges on how to support manipulation and analysis of high volume and high complexity multidimensional datasets. While even formulating a real-world engineering optimization problem is a difficult task, the inherent additional difficulties in extracting understanding, causality, and rationale of discovered information is exacerbating the demands on the user.

The ultimate goal in this work is to assist users in understanding and then communicating the complexities of the engineering design task in question and thereby supporting the decision-making process. As a step towards this goal this paper presents the system design of immersive parallel coordinates plots (IPCP). Prior work¹ has demonstrated that visualizing scientific information in an interactive and engaging environment is beneficial to users' decision-making abilities. We hypothesize that further benefits of visualization can be realized using emerging 3D immersive environments, such as Virtual Reality (VR) and Augmented Reality (AR).

In this work we explore how VR can be used to support visualization and analysis of computational engineering design datasets. In addition, we discuss the positive and negative aspects of the application

*PhD Student, Department of Engineering, University of Cambridge, skt40@eng.cam.ac.uk,

†Senior Research Associate, Department of Engineering, University of Cambridge, t.kipouros@eng.cam.ac.uk

‡Professor, Department of Engineering, University of Cambridge, pok21@eng.cam.ac.uk

of VR in the decision-making process. The overall task we study is the identification of clusters and sets of solutions in an optimization study for the aerodynamic design of compressor blades.^{2,3} However, we do not consider the physical interpretation and understanding of the designed compressor blades. We rather consider the physical nature of multi-objective optimization data. In such datasets, the input parameters express the design parameters of the design optimization study, and the output parameters express the objective functions and constraints.

In this work we do not address the classic challenges in using parallel coordinates plots (PCP) for visualization and analysis of multidimensional scientific data. Instead we explore how to build a system that can support the design process and investigate the usability of emerging immersive technologies for visualization of, and interaction with, scientific information. This allows us to assess whether the same challenges remain, or new ones emerge. Furthermore, the exploration of engineering design computational data is frequently centered on 3D visualization of the designed product. As such, an immersive interface offer a highly promising method for analysis and exploration of such data alongside with simultaneous visualizations of individual product designs. Part of the motivation of this work is to expose and exploit such strengths of VR technology. To this end we present the design of our VR-based IPCP system that is supported by a series of formative studies. A description of the first thereof which briefly discusses the observational user study, results of which we used to trim the vast design space, can be found in Tadeja et al.⁴ The underpinning VR system that included the basis of interaction techniques i.e. the gaze-tracking and racy tracing facilitated by a cross-hair was discussed in Tadeja et al.^{5,6} and is built based on the *Unity VR Samples Pack*⁷ that utilizes the well-know concept of a cross-hair to simulate the user’s gaze (for instance, see orange circle in Fig. 4) and *Oculus Utilities for Unity*⁸ together with some elements from the Unity asset store packages.⁹

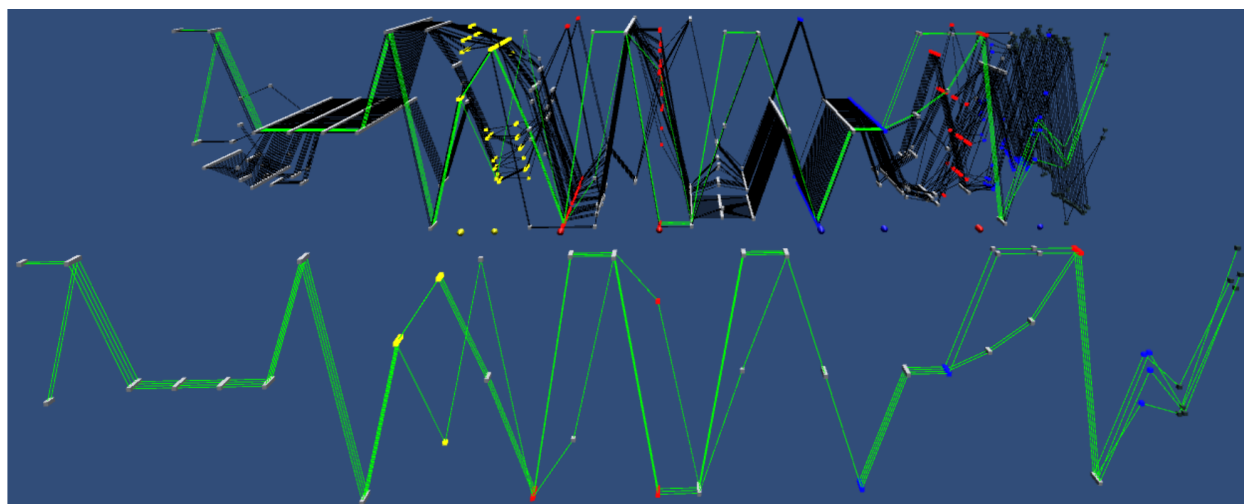


Figure 1. The user’s field of view in the VR environment. Selected data items are copied to a separate view (in green and in front). The copied subsection of data items, the main IPCP and the 3D scatter plots supports the *linking & brushing* principle i.e. the manipulation of one of the elements is constantly reflected on the other elements of the visualization as well.

II. Background and Problem Context

The sample dataset originates from Kipouros et al.¹ and is the result of a design optimization study for the 3D geometrical design of compressor blades subject to many equality and inequality constraints. In this study, we use the data produced for a three-objective function formulation of the optimization problem. Further details can be found in Kipouros et al.² and Kipouros et al.³ This dataset describes the Pareto front, which contains a set of 54 equally optimum design configurations, and the 29-dimensional data items categorized as: *design parameters* (26 dimensions) and *objective functions* (3 dimensions). The 2D view of the data is presented in Fig. 2 (a) whereas Fig. 2 (b) shows the two patterns of interest *A* (in blue) and *B* (in green) composed of 11 and 19 data items respectively that were identified in this dataset by the domain experts. The user’s view through a VR headset in 3D is presented in Fig. 1.

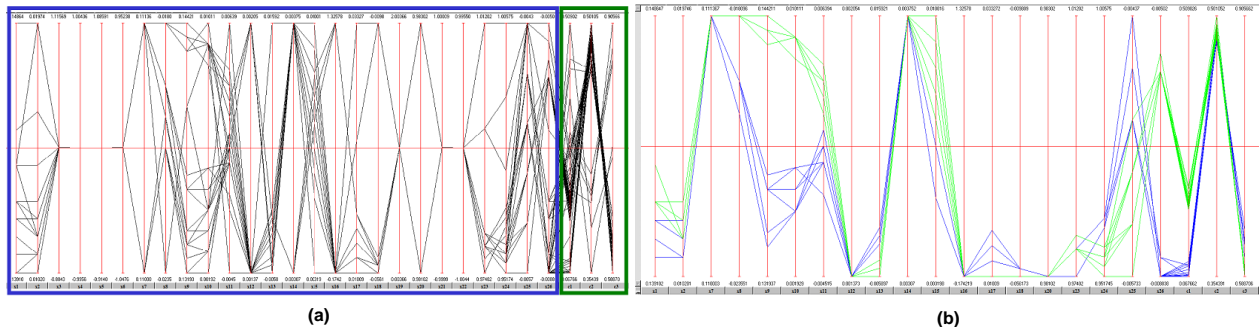


Figure 2. (a) the standard 2D PCP visualization of the dataset. The blue frame covers the *design parameters* (first 26 dimensions) and green frame covers the *objective functions* (last 3 dimensions). (b) the two patterns (in blue and green) originally identified by domain experts. Reproduced from Kipouros et al.¹

A. 3-Dimensional and Immersive Parallel Coordinates

Since the PCP concept was initially proposed by Inselberg¹⁰ in 1959, the number of entries containing the phrase “parallel coordinates”, as returned by Google Scholar search on November 1st 2019, reaches about 1.87 million items. The theory and applications of PCP can be found in Inselberg.¹¹

The wide adoption of 2D PCP has led to multiple attempts of designing a useful and compelling 3D PCP. For example, Wegenkittl et al.¹² and Gröller et al.¹³ describe Extruded PCP, which uses a PCP to visualize a complex surface. A different example of using the surface-type design is a density isosurface,¹⁴ which was used to visualize cytomics data. Falkman¹⁵ describes the *Cube* package applied for swift analysis of medical data.

Parallel Glyphs combine Star Glyphs¹⁶ with 3D PCP.¹⁷ The usability of 3D Multi-Relational PCP in conjunction with High-Precision Textures has been investigated in Johansson et al.¹⁸ Dang et al.¹⁹ describe a few examples of 3D parallel dot displays and stacked PCP, which uses various stacking techniques to augment the standard visualization of dot plots and PCP respectively. In their related work, Chang et al.²⁰ compared joined side-by-side 2D visualizations to PCP with scatter plot matrices, concluding that the PCP-based approach offered clear advantages over individual visualizations for most of their participants for a number of their tasks. Furthermore, Johansson et al.²¹ presented a comparison between 3D PCP visualizations with conventional 2D plots and concluded in favor of the latter. Holten et al.²² discusses different variants and improvements of 2D PCP used to find clusters in the multivariate datasets and remarks that only the use of additional scatter plots has substantial positive impact on the visualization.

ImAxes²³ allows the user to rearrange the axes interactively and supports both 2D and 3D scatter plots. Another example of an immersive PCP is the ART collaborative data analysis tool²⁴ that distributes the user interface across an augmented reality headset and a touch surface. Rosenbaum et al.²⁵ took another approach in which the user was treated as an element of the immersive PCP visualization itself.

Our approach, first described in Tadeja et al.,⁴ differs from other attempts in several key ways. First, we denote values in each dimension with interactive volumetric objects (i.e., unit-size cubes) that automatically highlight the item under the user’s gaze in order to signal that they can be interactive. The same is true for all the other interactive elements, such as selectors (i.e., various size spheres in the interface) used to move or rotate the scatter plots. This in turn opens up new design possibilities²⁶ that can use characteristics of a 3D shape, such as its orientation in the 3D space.²⁶ In addition, users can freely move within and maneuver in the 3D space to explore all parts of the visualization.⁴ Furthermore, our system augments the pattern identification task with 3D scatter plots,⁴⁻⁶ that can substantially aid users in isolating groups of data items that can potentially form patterns.

III. System Design

A. Overview

The overall objective is to design a visual analytics system for IPCP for engineering design processes, supporting efficient identification and comparisons of patterns of interest in a multidimensional dataset. This is



Figure 3. Various VR controllers : (a–b) an Xbox gamepad controller; (c–d) Oculus Touch. The green markers with accompanying letters are showing mapping between the action-buttons on both types of controllers. The description of user’s movement in the 3D space together with interaction details can be found in Tadeja et al.^{5,6}

a high-dimensional design space with many controllable and uncontrollable parameters. Controllable parameters include rendering parameters, such as colors and shapes, visual/auditory feedback design, interaction techniques, and choice of headset and hardware controllers. Uncontrollable parameters include properties about the datasets, such as size, density, and correlation structure, and the degree of domain knowledge and expertise of the user.

As a result of the many controllable and uncontrollable parameters, we have adopted an iterative process of first building an early system that provided direct transplantation of PCP to VR. This early design was then used to carry out qualitative user studies (partially described in a previous poster publication Tadeja et al.⁴), briefly outlined in subsection III B. We use the information distilled from the user study along with knowledge of the tasks, expert advice, and technical constraints to arrive at a set of design principles (see III E), which guide our task analysis in III F. In the task analysis we break down each task into functions and translate each function into a solution based on technical constraints, expert advice and information gathered from the qualitative user study.⁴

B. Qualitative Study

To help inform the design we carried out a qualitative study with seven participants (hereafter referred to as P1-P7). More details about this study can be found in Tadeja et al.⁴ The VR-based IPCP visualization used in the study can be seen in Fig. 1. Users could interact with the system using a mixture of gaze-tracking and Xbox gamepad controller, as shown in Fig. 3 (a–b). The color-coded cones^a were used to denote both the selected dimensions over the IPCP and axes directions at the scatter plots,^{5,6} see Fig. 3 (a, c).

Although the visualized dataset comes directly from the specific domain—the aerodynamic design of turbomachinery components—we wanted to pinpoint the general underlying design principles that can be used in future more advanced versions of the IPCP. Therefore, we broadened our sample of participants to include both fields of expertise and prior familiarity with PCP. The volunteers self-assessed their level of expertise with regards to VR and PCP using a Likert-like scale. As the color-coding has a crucial role in the visualization, participants were pre-screened with Ishihara’s tests,²⁷ which indicated no color deficiency.

On average, participants needed 25 minutes during the training phase and 15 minutes for the study itself. To offer insight into any adverse physiological effects, volunteers were instructed to take the simulation sickness questionnaire (SSQ)²⁸ before and after each phase (i.e., training or study) to see to what extent and if at all the visualization impacted negatively on their perception of comfort. Participants were reminded that they should stop the experiment if they experienced any of the symptoms described in the survey. However, none of the participants decided to stop and discontinue the study.

Before the experiment commenced, each participant sat through a demonstration of the IPCP technique, the tool under study and all the possible ways in which the user could interact with the system. Moreover, while listening to a verbal description, participants could track in real-time the instructor’s actions and their corresponding results on a computer screen.

Our participants had varying levels of prior familiarity with the PCP and VR. To level the playing field, we provided each volunteer with training data visualization. This process took between 18 and 36 min, depending on the individual participant’s needs, to achieve familiarity, as judged by the researcher. The tasks

^aW. Kresse, used under CC BY-SA 3.0; <http://wiki.unity3d.com/>

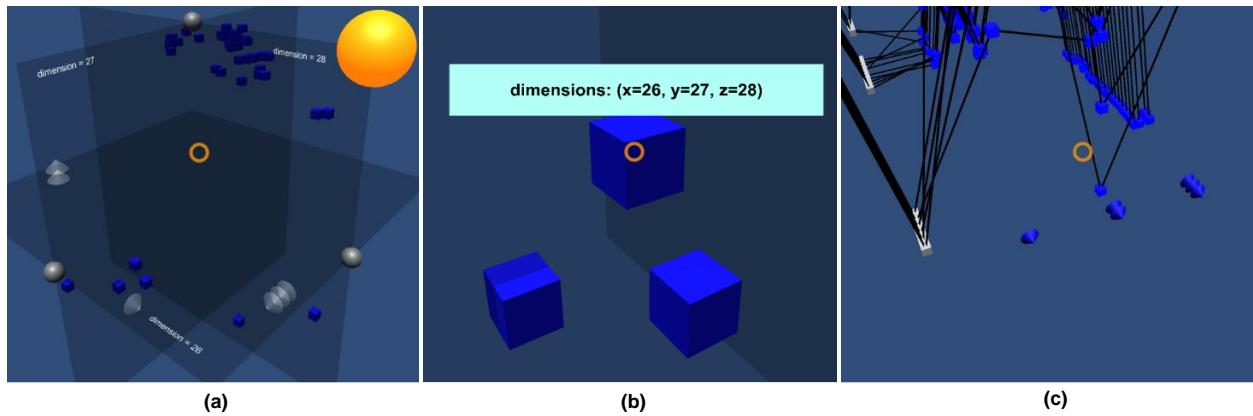


Figure 4. (a) The 3D scatter plot generated from up to three dimensions selected on the IPCP. (b) A few volumetric data points (blue, unit-sized cubes) on the scatter plot. When the user gazes over the cube a text box with additional information pops up. The text box automatically follows the user's gaze. (c) The selected dimensions are colour-coded and marked on the IPCP with a number of cones. The design of these 3D floating scatter plots was adopted from Tadeja et al.^{4–6}

included: a) selection of both data items and dimensions on the PCP; b) selection of data points and clusters on the scatter plots; c) visual analysis of the relation between the visualization elements; and d) rotation and movement of a scatter plot or generation of more than a single such 3D projection at once. Familiarization with the system was especially important, since the training may be very beneficial for the participant's grasp of the PCP technique, as remarked by Shneiderman et al.²⁹ For this purpose, we used a portion of a real-world dataset of 80 multidimensional data points, each consisting of 22 dimensions (20 inputs and 2 outputs). The training was followed up with a semi-structured interview. If deemed necessary by the researcher, the participant was asked to repeat previously completed tasks. Each instruction was marked as "pass" by the study supervisor who constantly monitored the participant's actions on the accompanying screen. During this process, the researcher provided oral instructions upon completion of a task, and whenever the participants requested assistance.

To minimize fatigue, participants were given a break of approximately 15 min after training and before the main experiment commenced. After training, the participants were asked to perform a task within the VR environment without a time constraint. The main task had an identical form compared to the training: *identify patterns in a multidimensional dataset of 26 input parameters and 3 output criteria*, together with the suggested approach to start developing the scatter plot using the criteria dimensions. We adopted a think-aloud protocol and requested participants to discuss what, why and how they are trying to achieve the task as it progressed. The supervisor, who continuously monitored the experiment, could at any moment inquire for further clarifications to gain more insight into the participant's actions and thought processes. In particular, for first-time gamepad users, an inquiry on whether the participant is stuck allowed the supervisor to provide guidance, for example, information about which button to click to invoke a specific action. This is important as the controller is not visible through the VR headset. Moreover, after running two participants (P1-P2), we decided to instruct the remaining participants (P3-P11) that a pattern is constituted by at minimum a group of four points. Audio and video were captured during the study for later analysis.

1. Participants Background and Expertise

All the participants reported having extensive experience and background in the STEM (science, technology, engineering, and mathematics) fields in either industrial or academic settings or both.

Participant 1 (P1) reported approximately six years of experience in engineering at undergraduate and graduate levels and two years spent in the automotive industry. He was not previously exposed to the PCP and had only very limited experience with VR.

Participant 2 (P2) reported less than three months of industrial experience. However, he has a background in aerospace engineering and engineering design spanning over nine years in total. He also considered himself novice with regard to VR with even lesser familiarity with PCP.

Participant 3 (P3) spent seven years in an academic setting and almost a year in the industry. She reported to have not previously used any VR headset. However, she is an expert on the PCP technique and its applications in design.

Participant 4 (P4) reported sixteen years of academic experience and three years of prior work in the industry. He is also an expert on PCP methods applied in modelling of the design processes.

Participant 5 (P5) has twenty one years of academic (15 years) and industrial (6 years) experience in Computer Science and software engineering. She also reported basic awareness with respect to VR and extensive exposure to PCP.

Participant 6 (P6) has a background in Artificial Intelligence and Computer Science (9 years). During that time, he has also worked for three years with the PCP and multidimensional data visualization methods.

Participant 7 (P7) reported that he has an academic background in mathematics and computational fluid dynamics (8 years). However, he was a novice in terms of prior familiarity with VR and PCP.

Participant 8 (P8) was a doctoral student working on the application of analytical methods such as subspace-based dimension reduction to manufacturing deviations in turbomachinery components. He also self-reported having some limited experience with the VR and 2D PCP.

Participant 9 (P9) self-reported an intermediate level of familiarity concerning both 2D PCP and VR. This participant was also an expert in the field of mechanical and manufacturing engineering with vast experience of multi-objective optimization in academia and industry.

Participant 10 (P10) reported having some limited prior exposure to VR and advanced knowledge about the 2D PCP. He used the 2D PCP for conducting research into his master’s thesis in aerospace engineering. He also holds an undergraduate degree in the same area. In addition, he also recognized the dataset in which he looked for the patterns, as he used the same data in his prior research involving 2D PCP. However, he reported not spending a substantial amount of time on the analysis of this data, and he did not previously try to specifically identify the contained patterns.

Participant 11 (P11) self-reported to never before using the VR headset and to have some limited prior exposure to 2D PCP. He also reported to have undergraduate and postgraduate degrees in mechanical and aeronautical engineering and is currently in his first year of doctoral studies in aerospace engineering.

Table 1. The groups of data items identified as potential patterns by the participants (P1-P7). Color-coding of this data indicates the overlap with the two patterns (two bottom rows) identified by Kipouros et al.¹ as can be seen in Fig. 2 in blue and green respectively.

Pattern candidate	P1	P2	P3	P4	P5	P6	P7
1	34, 35, 36, 38, 39	2, 5	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 30	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26	34, 37, 38	33, 34, 35, 36, 38, 40, 41, 42, 43	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26
2	8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28	0, 1	33, 34, 35, 36, 38, 40, 41, 42, 43	33, 34, 35, 36, 38, 40, 41, 42	6, 8, 9	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 30	33, 34, 35, 36, 38, 40
3	9, 10, 12, 13, 14, 15, 16, 17, 18, 24, 25, 26, 27, 28	7, 9, 10, 17, 19, 27, 29, 42	47, 48, 50, 51, 52, 53	8, 9, 10, 11, 12, 13, 14, 19, 20, 21, 22	29, 30, 39	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28, 29, 30	—
4	—	47, 51	—	—	—	46, 47, 48, 49, 51, 52, 53	—
Pattern A	32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43 no. of data items = 11						
Pattern B	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28, 29, 30 no. of data items = 19						

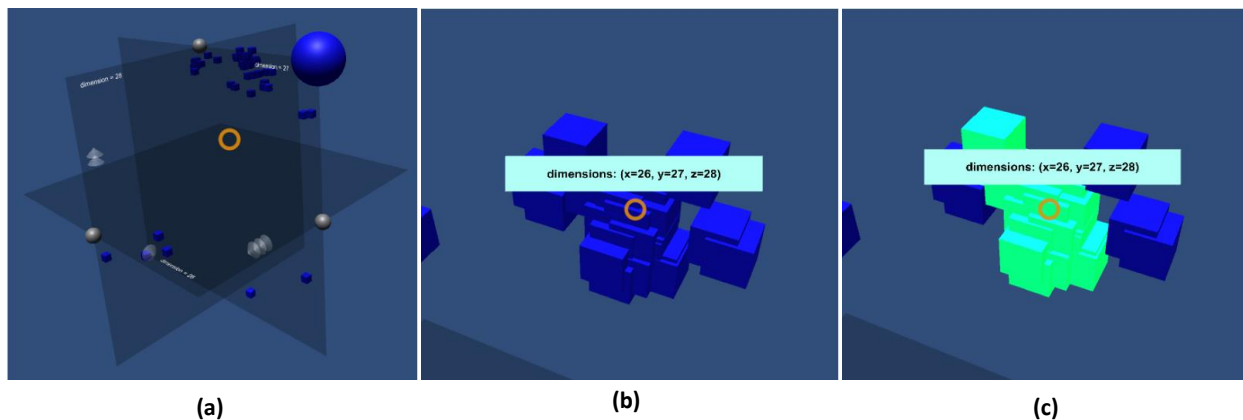


Figure 5. (a) The 3D scatter plots^{5,6} provides built-in simplistic algorithm for automatic clustering of overlapping data points (b-c). If the volumetric data markers (i.e. unit-size cubes) overlap with each other (b) to a certain, predefined degree, the selection of one of them will automatically select the entire cluster as shown in (c). Moreover, these selections will be simultaneously reflected on the other visualization elements, i.e. the IPCP and other scatter plots.

C. Identifying Patterns

Kipouros et al.¹ provided an analysis that discovered two patterns of importance, shown in Fig. 2 (b). As the study involved a real-world dataset, participants were instructed to note that some variability was to be expected in the data. However, they had to decide independently on what level of variability would be acceptable. Only two participants (P2 and P5) failed to recognize most of the data points belonging to the two groups. In three cases, the participants changed their initial pattern candidates by selecting or deselecting elements originally grouped, thus creating subsets or partial repetitions of the previous selection (P1, P4, and P6). The majority of the participants identified almost the complete pattern *A* (with completion ratios for P3:9/11, P4:8/11, P6:9/11 and P7:6/11) or *B* (with completion ratios for P1:16/19, P3:16/19, P4:15/19, P6:19/19 and P7:15/19). These differences most likely come from limited knowledge of the area from which the dataset was sourced and, as such, the allowed levels of noise between the respective values.

D. Observed Participant Behavior

The results of the study involving first seven participants (P1-P7) suggest that the time needed for the analysis was correlated with prior experience with PCP, that is, the more expertise the participant had, the longer it took to finish the training session. Simultaneously, experts took slightly less time to complete the requirements of the main study. However, this data should be interpreted with caution. The participants did not have a clear cut-off point. For example, the study only finished once all the potential patterns were found. There was also an age difference between the two groups; novice participants were in their twenties (P1, P2, P7) whereas experts were in their thirties (P3, P4, P6) or forties (P5). Furthermore, the experts had little or no familiarity with VR or the controller, which put them at a disadvantage compared with younger participants, for instance, P7, who reported extensive experience with this type of gamepad. This may have impacted their approach used during the study.

All the participants quickly became fluent in their chosen movement and manipulation strategies and did not report any comments or suggestions directly linked to these interaction techniques.

None of the participants generated more than a single scatter plot comprised of different dimensions even though they were informed about this possibility existed. Such behavior, however, may have been observed because only three criteria were present in the dataset. Hence, all of them could be mapped onto a single scatter plot. Moreover, the training phase also exposed participants to the methodology of exploring the data mainly through the outputs' scatter plot which was also used by Kipouros et al.¹ in their exploration of this dataset. However, some of the participants—P1, P2, and P6—reflected on the fact that having more than a single scatter plot may be confusing.

We observed that participants P4 and P6 chose to re-position the scatter plot closer to themselves, or the particular part of the main plot, instead of moving towards the scatter plot floating above the center of

the IPCP plot. This behaviour would most likely change if more data items would be visualized at once. Therefore, the movement would take considerably more time, and the perspective would render the distant selection of the elements at the scatter plot almost impossible.

Only participants P2 and P6 chose to rotate the scatter plot. However, we observed that for all participants the exact mapping of the dimensions to the 3D projection was not critical and this was also confirmed later on during the final evaluation study. The participants were selecting clusters on the scatter plot irrespective of the current rotation of the projection. One participant (P3) was constantly selecting the same, largest cluster, instead of exploring other groups or individual points.

E. Design Principles

From the technical constraints of what is achievable in VR, expert feedback,¹ and observations from our study we distilled six design guidelines:

D1–Immersive PCP visualization. The IPCP design should leverage the benefits bestowed by the immersive, spatial environment and remain both useful and compelling to the user.

D2–Pattern candidate identification. The user must be able to identify candidates, i.e. a group of data items, as possible patterns within the dataset.

D3–Pattern candidate inspection. The user has to be able to inspect candidates for potential patterns within the dataset to decide if the particular group of data items forms a pattern.

D4–Pattern candidate comparison. To avoid repeated selections of the same pattern, the user has to be able to effectively compare a group of data items that could potentially form a pattern but also be able to compare such group with previously discovered patterns.

D5–Pattern storage. The possibility of saving and retrieving on-demand information about discovered patterns is necessary for the users to be able to differentiate between potential new candidates and previously discovered patterns.

D6–Visualization exploration. The system should support exploratory visualization by providing mechanisms for zooming and exploring individual visualization elements.

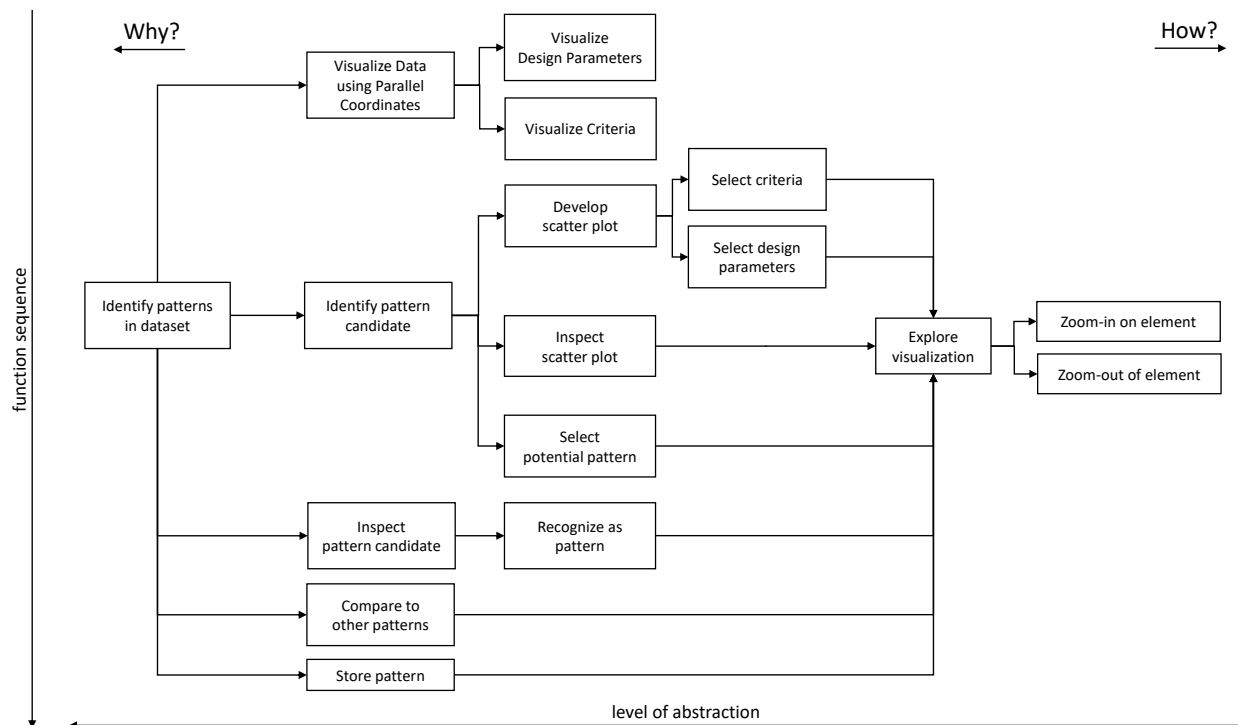


Figure 6. The diagram visualizes the outcome of the Function Analysis Systems Technique (FAST)³⁰ based on which we reasoned about the requirements and functionality of our IPCP system. The diagram should be read from left to right and from top to bottom.

		solutions →		
Sub-functions		1	2	3
functions ↓	<i>Immersive PCP visualization</i>	parallel planes	values spread in the Z-axis	standard 2D view
	<i>Pattern candidate identification</i>	visual analysis of the dataset	scatter plot exploration	iteration over all possible subsets
	<i>Pattern candidate inspection</i>	comparison of values in each dimension	visual inspection	analytical and visual comparison
	<i>Pattern candidate comparison</i>	comparison of the items' metadata (indexes)	visual comparison with the other patterns	analytical and visual comparison
	<i>Pattern storage</i>	storage in the database	pattern duplication in the 3D space	storage in the clipboard
	<i>Visualization exploration</i>	zooming-in and -out	movement in the 3D space	visual inspection of individual elements

Figure 7. Functions captured with the help of FAST technique (see Fig. 6) were translated to the function carriers. The 2nd column (in light blue) contains the actual solution that has been implemented in our system.

F. Task Analysis

We performed a task analysis by first identifying a suitable user interface workflow guided by the previously described user study. Each task is broken down into functions carried out either in series or in parallel. Thereafter, each function was translated into a function carrier (solution) by considering the requirements, technical constraints, and data from the user study. This analysis was carried out by first modeling the function structures with the help of the *Function Analysis Systems Technique* (FAST),³⁰ illustrated in Fig. 6. The translation of functions to function carriers is shown in the conceptual design tables on Fig. 7. For each key requirement (first column on the left) three alternative solutions (columns marked 1-3) were presented. For the benefit of the reader, the selected design solutions are listed in the 2nd column. The decisions were made using information gained through the formative user study in the form of observation and user comments and discussion. The list of these tasks include: (D1) *visualize data using IPCP*; (D2) *pattern identification*; (D2*) *scatter plot generation (augments the D2)*; (D3) *inspection of the selected data items*; (D4) *comparison of the new candidate to other patterns*; (D5) *storage of the identified patterns*; (D6) *exploration of the visualization*. The diagrams flow from left to right; the branching should be understood as two overlapping possibilities that can be carried out together or augment each other.

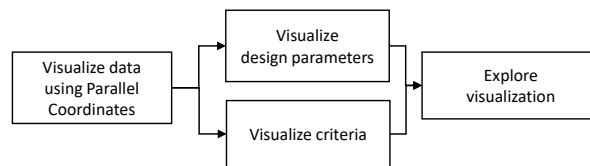


Figure 8. D1: Visualize data using immersive PCP (IPCP).

D1: Visualize data using IPCP. This task can be implemented in many variants in both 2D and 3D settings. For instance, one can use parallel planes^{21,24} to visualize the parallel axes. Instead, in our case, each data item is essentially a series of connected, interactive, cubes spaced in equal intervals over the X-axis. Identically to 2D PCP, the Y-component denotes the value for each dimension. However, as we are spreading the points also in the third dimension, the cube's position in Z-axis signals the data item indexing with the first element starting with the Z-component equal to zero and progressing towards the positive direction of the Z-axis (see Fig. 1). As opposite to other 3D PCP implementations, dropping the parallel planes limits the potential effects of occlusion and allows users to gain a birds-eye overview of the dataset by positioning themselves above the IPCP.

None of the participants reported confusion concerning the way the system visualized PCP. Moreover, all of them quickly understood that the interactive cubes indicate the numerical values in each dimension and that a group of line-connected cubes constitutes a single data item.

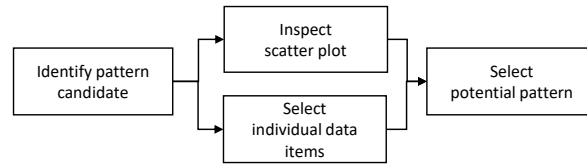


Figure 9. D2: Pattern identification.

D2: Pattern identification. This task can be carried out in two ways: via exploration of the IPCP and visual inspection of a subset of selected data items, or with the scatter plot, as it was strongly suggested to the participants during the user study. However, we observed that the participants preferred a mixture of both approaches, that is, in many cases, a participant augmented a subset of data items selected with the help of the scatter plot by adding or removing selected data items individually from the PCP visualization.

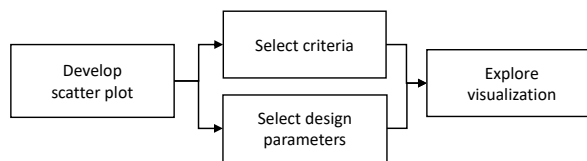


Figure 10. D2*: Scatter plot generation.

D2*: Scatter plot generation. This task provides a feasible approach to pattern identification, especially when handling a large volume of data visualized with IPCP. The IPCP is augmented with the help of scatter plots^{1,22} generated from values in one, two or three dimensions across all the data items. Hence, we gave the participants the ability to create up to five different 3D scatter plots. We also advised our participants to start by selecting the criteria to develop the scatter plot and see if they can use this aid to identify clusters of data points on the scatter plot to help them efficiently select a group of data items that can potentially form a pattern.

The scatter plots were augmented (see Tadeja et al.^{5,6}) with additional functionality, such as the ability to apply rotation in 90° about a chosen axis. In addition, they could be moved into any position in 3D space.

Based on the participants' comments we decided to contextualize the data points and scatter plot axes by providing textual labels, as seen in Fig. 4 (a-b). The labels contained information of which dimensions, that is, which Y-values, were mapped to which coordinates on the scatter plots. However, the two expert participants suggested a re-evaluation of this design, as they found the labels not to be very helpful.

If two or more values in a particular dimension are identical within the floating-point precision of their coordinates, they will be presented as two entirely overlapping cubes on the scatter plot. This effect would interfere with the ray-tracing used for gaze tracking as the game engine has no means to decide with which overlapping object the user wants to interact with. The analogous situation may also happen in the 2D PCP. In such a case, only a single point will be visible on the scatter plot, the selection of which will highlight multiple corresponding items on the IPCP as those points would be automatically clustered with each other.

The simplistic, naïve clustering algorithm (see Fig. 5) relying on the spatial distance of the cubes' centers, was nonetheless quite effective. Certain restrictions, such as a difficulty of differentiating with which object users want to interact with if they are entirely overlapping, cannot be easily solved. We conjecture that different clustering algorithms for multidimensional data, such as one proposed by Wegman et al.³¹ might be applied here. One way to minimize occlusion and cluttering effects in PCP is to use the density information of the data. This is shown in Heinrich et al.³² and also in Artero et al.,³³ where they proposed algorithms that use frequency and density information to decrease the overall number of data points that are shown to the user in PCP. Moreover, giving the user the possibility of zooming in and out of the cluster internal structure may prove to be very convenient in such situations.

D3: Inspection of the selected data items. This task is made possible through visual inspection of the selected items by either rendering not-selected points temporally invisible or by making a copy of the

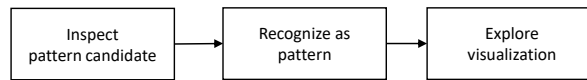


Figure 11. D3: Inspection of the pattern.

subset and placing it in a chosen part of 3D space. In both cases, participants usually first tried to gain an overview of the entire selection by positioning themselves (i.e., zooming out) in front of the subset and then compare and decided if particular data items constitute a pattern.

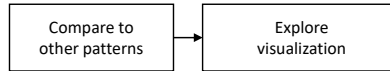


Figure 12. D4: Patterns comparison.

D4–Comparison of the new candidate to other patterns. This task is achievable by placing candidates near each other in 3D space and then visually inspecting possible overlaps. However, one of the participants mentioned that having some analytical aid could be helpful, especially if the data items are confounded by many dimensions. This participant also suggested using color-coding to provide additional information on how particular subsets potentially overlap with each other.

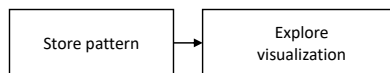


Figure 13. D5: Pattern storage.

D5–Storage of the identified patterns. The system automatically logs a list of data items selected by the user into a file. Second, the sub-selection of data items can be placed and stored in any part of the 3D space throughout the entire duration of the visualization. However, as the IPCP extends along the Z -axis it may be difficult to compare points that are placed on the opposite sides from each other.

In addition, participants also remarked that a combination of both of these methods could potentially be even more beneficial. The logged data can be made re-callable, visualized on-demand to assist the user in understanding, for example, if, and to what extent, patterns, overlap between each other by additional color-coding or temporal shape change.

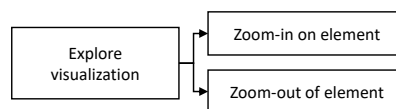


Figure 14. D6: Visualization exploration.

D6–Exploration of the visualization. The movement and manipulation methods allow the user to zoom-in or zoom-out on any visualization element, which, as reported by the participants, was easy to use with either Xbox Controller or Oculus Touch (see Fig. 3).

Based on the gathered qualitative data, all the participants swiftly understood the proposed movement and manipulation techniques. However, participants unfamiliar with the controller had problems remembering which button activated which action. This problem can be rectified by several strategies. First, the training can be extended. Second, the feature set can be limited. Third, the feature set can be initially limited and gradually expanded as the user gains proficiency in the interface. This third strategy is an example of a training wheel interface.³⁴ The ideal solution is probably a combination. In the short-term, extending training is preferable as PCP, even in its 2D version, demands intensive practice to gain fluency in their usage.²⁹ Another possible option would be to use a different VR headset (e.g., HTC Vive) that supports spatial tracking of the wearer. In such a case, instead of using a controller for movement and maneuvering

the user can move around in the real world which would be instantly reflected in the VR environment. This would decrease the number of required interaction methods, although it would raise other issues, such as the increased complexity in tracking set-up and space demands for the system.

IV. Verification, Evaluation and Final Design

A. Interface Verification

To verify if our system offers all the functionality necessary to fulfill the primary user requirements we probed the usability of the system using the widely adopted Nielsen's^{35,36} heuristics.

1. *Visibility of system status*: The system constantly informs the user which visualization elements are in a selected state on both the IPCP (see Fig. 1) and on the scatter plots (see Fig. 4 and Fig. 5). The selected objects remain highlighted (in light green or orange) up to their deselection or when the user invoked visualization reset. This provides the user with continuous feedback about the state of all individual elements. Moreover, the system supports *linking & brushing* interaction techniques by automatically mapping the current state of each of the visualization elements, i.e. scatter plots and the IPCP onto each other. Furthermore, if the user reaches the limit of possible duplicates, the system will disallow this operation and signal this to the user by playing a short sound.
2. *Match between the system and the real-world*: Both the 3D scatter plots and the cross-hair based gaze-tracking (see Tadeja et al.^{5,6}) should be familiar to the majority of users. The use of the cross-hair to represent the user's gaze point and the focus the user's attention on an object directly in front of where the user is looking at the moment was implemented based on a well-known real-world concept.⁷ This, in turn, allows the system to optimize its resources to provided immediate feedback and response to the user's actions.
3. *User control and freedom*: Similarly to Tadeja et al.,⁴⁻⁶ the user can reset (see button [R] in Fig. 3) the visualization to its original state which will preserve the copied data items (see Fig. 1). However, due to efficiency reasons, these duplicates will no longer support any kind of interaction after the system reboots and remain in "a freeze" state based on where they were originally copied. To further foster user control and freedom—before the final evaluation study with the larger dataset—the system was extended to support the *undo & redo* operations i.e. the last selection of data item on the IPCP or data point on the scatter plot can be undone or redone if necessary by the simultaneous press of two *thumb-rest* buttons (one on each of the controllers, see button [H] in Fig. 3).
4. *Consistency and standards*: Despite the recent rapid development and advances in the research, the VR technology is still in its early stage. As a consequence, the standards regarding the interaction techniques, visualization guidelines or the overall design principles are not yet fully understood. We were trying to be consistent with well-known concepts, thus, for instance, the selection of the object requires a double-tap of the gamepad button instead of a single click similarly to what is necessary when selecting an object using a computer mouse.
5. *Flexibility and efficiency of use*: The "accelerator" provided for the benefit of advanced users was the possibility to generate multiple scatter plots. Each of them could map up to three dimensions from the IPCP. Moreover, some users possessed previous experience with the Xbox Controllers (see Fig. 3), and therefore required substantially less time to acquaint themselves with the interaction techniques.
6. *Error prevention*: The most common "slip-type" of an error related to user failing to focus his or her gaze on a chosen point is easily fixable with maneuvering techniques which allow the user to zoom-in the object by just moving towards it. The scatter plots automatically deselect the rotation selector when another one is being selected (see Tadeja et al.^{5,6}) by the user. This prevents the system from confusion about which axis the scatter plot should be rotated. Moreover, the system will omit visualization of the data items on the PCP and the scatter plots if there is a gap in values at the input dataset, hence preventing the user from working with broken or incomplete data.
7. *Help users recognize, diagnose, and recover from errors*: The system uses the audio feedback to inform the user if no more duplicates of the selected data items can be generated immediately after such action was performed.

8. *Recognition rather than recall*: The *linking & brushing* interaction techniques help to minimize the user’s memory load by providing constant feedback about which operations have been previously executed by the user. Moreover, the user can easily duplicate a group of data items for further inspection and comparison with other such groups. This feature can be used, for instance, to keep track on which group of points were previously recognized by the user as a pattern.
9. *Aesthetic and minimalist design*: The systems keeps the information presented to the user to a minimum. Only selected elements are highlighted (see selected: data items on Fig. 1, movement selector on Fig. 4(a) and cluster on Fig. 5(c)). Moreover, the gazed-locked textbox is only shown once the user’s gaze is hovering over a data point on the scatter plot and automatically disappear otherwise. Thus, minimizing the possible additional cluttering and occlusion effects simultaneously providing additional information of which dimensions were used to map into this points’ rearrangement on the particular scatter plot.
10. *Help and documentation*: As it is hard to find users experienced in both the PCP and VR the participants were trained in how to read the PCP and how to interact with our system through the gaze-tracking and the Xbox controller (P1–P7) or the Oculus Touch (P8–P9).

B. Evaluation

It is very challenging to meaningfully evaluate visual analytics tools in short-term lab-based studies. Traditional human-computer interaction evaluation methods, such as A/B testing, rely on the experimenter’s ability to explicitly control all controllable and uncontrollable parameters of the system and the task as much as possible. However, in visual analytics, this parameter space is very large. For example, the data density, the nature of the task, characteristics of the dataset used in the test, and the experience and motivation of the participants, are all variables that are likely to affect the outcome of an A/B test. In addition, easy-to-measure dependent variables, such as task time, are only partially relevant as one of the primary objectives of a visual analytics tool is to provide additional *insight* rather than enabling users to make decisions faster.

As a consequence, we evaluated the system design qualitatively with two independent user experiments. We recruited two pairs of different surrogate expert users to participate in each study. These participants are referred to as P8 and P9 in case of the study involving the same dataset as in the previous study and P10 and P11 with the new, larger, more complex dataset. All four volunteers were also pre-screened with the complete version of the Ishihara’s color deficiency test.²⁷

1. Part I: 54 Data Items with 29 Dimensions per Element

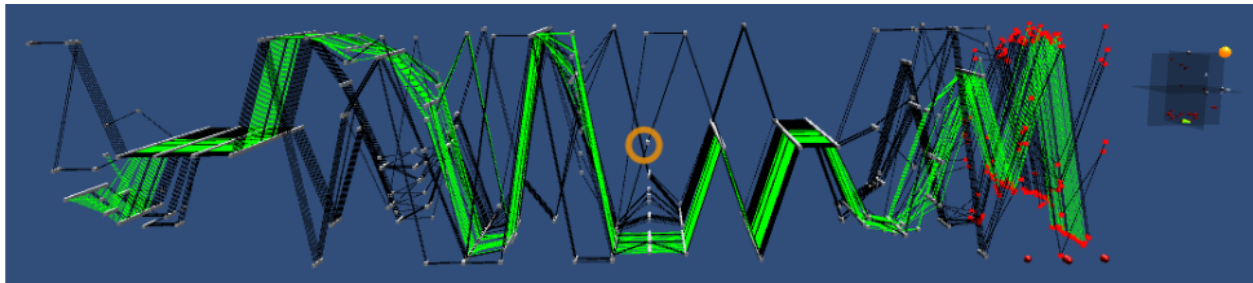


Figure 15. The user’s field of view of the 29-dimensional dataset of 54 equally optimum design configurations describing the Pareto front (26 design parameters and 3 objective functions).

To validate the system design, we used the same three-part procedure and dataset (see Fig. 1) as in the first study reported earlier in this paper. However, based on the previous participants’ feedback, we skipped parts of the presentations related to having multiple 3D scatter plots, although this possibility was not removed from the IPCP. As such, the training phase was truncated. Overall, the results revealed that participant P8 managed to find the majority of the elements in both patterns, *A* with a 9/11 completion ratio and *B* with an 18/19 completion ratio, whereas the second participant P9 correctly identified only the entire pattern *B* with a 19/19 completion ratio.

Table 2. Group of indexes of data points identified as patterns by the participants P8 and P9. The two bottom rows contain the two patterns that were identified by Kipouros et al.¹ as can be seen in Fig. 2 in blue and green respectively. Here we report only the participants' best attempts to identify the two patterns.

Pattern candidate	P8	P9
1	33, 34, 35, 36, 38, 40, 41, 42, 43, 46, 49, 52	8, 9, 10, 12, 14, 15, 16, 17, 18, 27, 28
2	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 27, 28, 30	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28, 29, 30
Pattern A	32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43	
Pattern B	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 26, 27, 28, 29, 30	

2. Part II: 166 Data Items with 39 Dimensions per Element

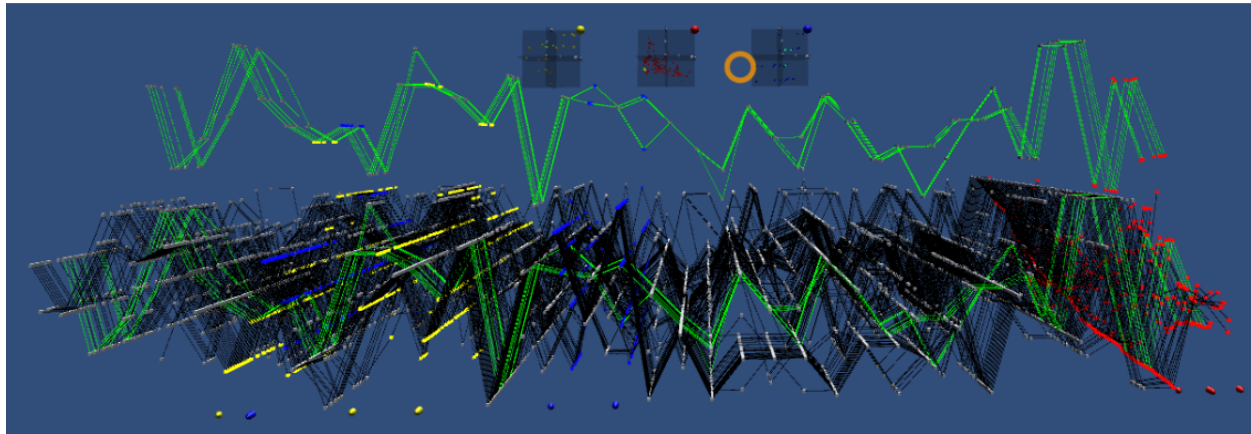


Figure 16. The user's field of view of the 39-dimensional dataset (36 design parameters from FFD parameterization and 3 design criteria: CP, DC60, and Swirl). The mapping i.e. the 3D scatter plot of the dimensions describing the design criteria i.e. *objective functions* is plotted (in red) in the middle scatter plot.

Table 3. Group of indexes of data points identified as patterns by the participants. The four bottom rows contain the four patterns that were identified by domain experts.³⁷ In the case of P11 we excluded repeated pattern selections.

Pattern candidate	P10	P11
1	93,96,102,103,105,119,120	44, 46, 49, 56
2	5, 9, 12, 25	92,95,98,106
3	13, 15, 17, 19, 20, 21, 23, 29, 30, 31, 42, 43	5, 9, 12, 14, 25, 27
4	2, 3, 4, 5, 8, 9, 12, 14, 18, 25, 27	34, 35, 37, 38
5	44, 46, 49, 56, 59	1, 2, 3, 5, 7, 10, 11, 12, 14, 18, 22, 25, 27, 59
6	90, 92, 95, 98, 104, 106	110, 114, 116, 117, 121
7	-	2, 3, 7, 18
Pattern C	44, 45, 46, 48, 49, 51, 56, 59 no. of data items = 8	
Pattern D	57, 61, 62, 63, 64 no. of data items = 5	
Pattern E	13, 15, 17, 21, 29, 30, 31, 34, 35, 37, 38, 42, 43 no. of data items = 13	
Pattern F	2, 3, 4, 5, 8, 9, 12, 14, 18, 22, 25, 27 no. of data items = 12	

After extending the system with additional capabilities, such as the support of basic *undo & redo* operations we reran our validation study with another two domain-experts (hereafter referred to as P10-P11). We also used the Oculus Touch instead of Xbox gamepad controller to facilitate the interaction (see Fig. 3).

We further investigated the feasibility of using IPCP for pattern identification in a substantially larger

and more complex dataset. The user study was carried out over a new dataset containing 166 data items with 39 dimensions per element containing 4 patterns of interest identified by domain experts (see Fig. 16). This dataset describes the results of the optimization study of S-Duct design (see Fig. 17), as presented in D’Ambros et al.³⁷

In addition, the near-misses in terms of selection were the main observable issues with our selection technique. To support the user’s selection task, one can use a form of target acquisition aid, that in our case, has been loosely inspired by the *Bubble Cursor* proposed by Grossman et al.³⁸ We leveraged the features natively supported by the Unity game engine by on-the-fly manipulation of the size of the *sphere collider*³⁹ component attached to the interactive objects (cubes) in the IPCP. The size of the collider, which acts as a target for the raytracer,⁷ is dynamically adjusted based on the size of cross-hair (see the orange circle in Fig. 16) which in turn is dynamically scaled with respect to the user’s distance from the target objects.⁷ Initially, we set the collider’s radius as $\sqrt{3}/2 \approx 0.866$ to encompass the entire cube in the sphere. The further away the user is from the object, the larger the collider becomes.

P10 found 5/8 of pattern *C*, 9/13 of pattern *E* and almost all pattern *F*: 11/12. P11 found 4/8 of pattern *C*, 4/13 of pattern *E* and 9/12 of pattern *F* in his best attempt. However, none of the two participants found pattern *D*. They did, however, identify other potential patterns that were not initially found by the original researchers, which is a concrete demonstration of the utility of the system in enabling users to explore and identify previously unknown patterns in the dataset. Interestingly, some overlapping results were reported by both participants, as can be seen in Table 3, namely pattern candidate 6 for P10 and pattern candidate 2 for P11. Moreover, both participants identified additional groups of points that in their opinion constituted patterns not originally identified by the experts in D’Ambros et al.³⁷ As mentioned before, in the case of visual analytics, we aim to provide additional insight about the dataset, hence we were interested in the new candidates for potential patterns found by the two participants (P10 and P11). To do this, we requested a domain-expert to re-analyze and comment on the new pattern candidates identified. After inspection and preliminary analysis, the expert concluded that indeed these are different patterns sharing some common characteristics with those identified previously and as such they are complementary. However, the domain expert commented that they are not as rich as the dominant four patterns previously identified, i.e. *C*, *D*, *E* and *F*. Nonetheless, this is concrete evidence that VR analytics enable users to detect new patterns.

C. Design Decisions

Here we summarize the final design decisions refined after we ran additional experimental studies with domain experts over two separate datasets (see Fig. 15 and Fig. 16) varying in both number of data items and dimensions per data item as well. Moreover, based on the participants observed behavior and their comments we also identified a number of potential improvements.

D1–Immersive PCP visualization. We have developed immersive PCP (IPCP) as a software system running within the Unity³⁹ game engine. Our approach represents a 3D version of regular 2D PCP but with values of each dimension denoted as volumetric, interactive objects, that is, cubes spread in the *Z*-axis instead of using parallel planes as regular 3D PCP usually does. However, this way of presenting data has its own challenges. For instance, to clearly see the distant objects with respect to the user’s current position, it was necessary to provide adequate means of close inspection of such regions. Here, we give the user the possibility to freely move and maneuver in the 3D space using the controllers (Xbox gamepad in first and Oculus Touches in the following two observational studies). Such approach has additional advantages, especially when we coupled the IPCP visualization with the CAD models of the respective designs (see Fig. 17) which now can be inspected not only from outside but also from within as well (see Fig. 17(c)).

D2–Pattern candidate identification. To identify and highlight a group of data items that can potentially form a pattern, the user either selects data items one-by-one or by generating a 3D scatter plot from up to three dimensions. This data is automatically clustered based on the distance between the data points (see Fig. 5) and the selection of such data points belonging to a cluster simultaneously highlights points belonging to this cluster on the scatter plot and the IPCP alike. The larger the data set and subsequent IPCP visualization, the more the scatter plots seem to aid the user in the task of finding patterns. Here, we noted that the actual orientation of the scatter plot’s local axes is not as important as providing the user with a clear way of distinguishing which points on the scatter plot potentially form a cluster. We observed that even a simple clustering algorithm, however relying on the user’s intuitive feel of the grouping of volumetric points placed in close vicinity can be surprisingly effective. Hence, potential improvements could be threefold. First, other clustering algorithms can be studied here such as *Ordering Points to Identify Cluster Structure*

(OPTICS)⁴⁰ or *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN)⁴¹ as they can, by design, handle noisy data, e.g. containing almost constant dimensional values across all the data items, such as the one used in our experiments. Second, different methods of the 3D scatter plot manipulation can be used. For instance, instead of rotating the scatter plot along one of the three main axes, it could be rotated about any vector selected by the user. These would potentially help the user to swiftly obtain an overview of the entire data present at the scatter plot. Third, to avoid repeated selection of the same cluster, and subsequent repetition of the same group of data items on the IPCP, we could additionally color-code or otherwise previously selected and inspected clusters. Furthermore, additional information could be given in a form of textual labels attached to the data points instead of current axes mapping, such as the scatter plot’s coordinates or number of points within selected cluster which could be more informative to users.

D3–Pattern candidate inspection. To inspect the pattern candidates, the user can temporally disable the unselected data items of the IPCP and use the manipulation and movement capabilities in the 3D space to verify if this particular group of data items forms a pattern. Moreover, if a particular data item is upon inspection deemed by the user to not be part of this pattern, it can be easily deselected as well. This is especially useful the more data items are contained within the dataset. As the data items will be spread in the Z -axis (see Fig. 15 and Fig. 16) the potential occlusion effects are more persistent with the growing number of data items. Here, the ability to toggle between the view with all or only currently selected data items has proven to be very useful.

D4–Pattern candidate comparison. To compare the currently selected group of data items with previously recognized patterns, the user can make a copy of these data items, i.e., the pattern candidate, and place them in a part of 3D space where the other patterns have been stored. The user may then visually compare them to decide if a new pattern has been found. Here, the system should provide the user with the “snapping” mechanism that would automatically align the subsets of the data items that the user wants to compare with each other. To some extent, the participants were trying intuitively to obtain a similar effect by placing the duplicates of previously found patterns next to each other in the 3D space.

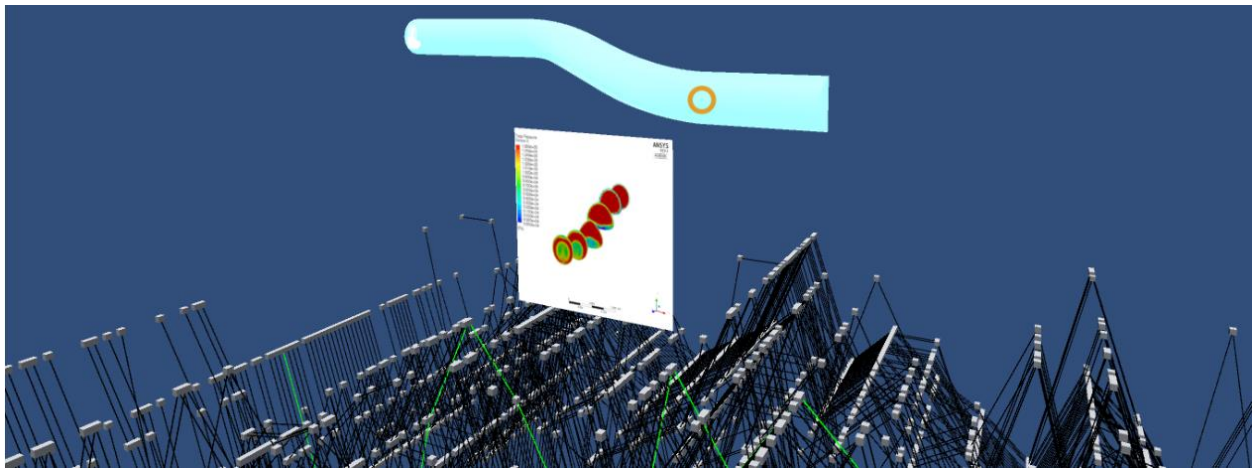
D5–Pattern storage. Previously discovered patterns of data items can be stored freely in any location within 3D space. Here, even though we could keep the “frozen” data items as non-interactive for optimization reasons, we could color-code the data items overlapping across the previously found patterns and the new group of points to inform the user to what extent they share the same subset of data.

D6–Visualization exploration. Exploration of the entire visualization with all of its elements is made possible through the use of gaze-tracking, head-tracking and movement assisted by the controllers. Here, we observed that all our participants swiftly gain fluency in how to move around the virtual space. We can also leverage the tracking of the user in a real space. However, in our case, the IPCP, – pending on the input data size – can be potentially vastly extended in at least the XZ -plane parallel to the user actual ground-plane. As such, it could be unfeasible for the user to walk and see the entire visualization, hence new interaction and transportation methods would have to be included.

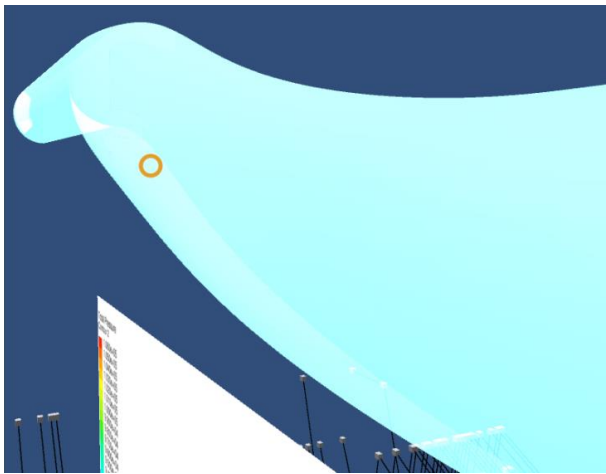
V. Visualization Extensions

After analyzing both the qualitative and quantitative feedback together with the participants’ comments, we have decided to extend our system with additional features. The system contains the visualizations of associated 3D models of the full 3D S-Duct geometries and 2D flow solutions, an example of which can be seen in Fig. 17. In the future we are also planning to give the users the possibility of overlaying the various S-Duct geometries on top of each other to closely inspect and identify potential nuances in the difference between the shapes as demonstrated in the AeroVR system presented in Tadeja et al.^{5,6} and upon which this visualization is built. This allows the users to promptly verify their pattern selections and compare between the S-Duct shapes associated with selected data items. This extension also fosters the connectivity with commercial software and simulation packages that are widely used in the industry.

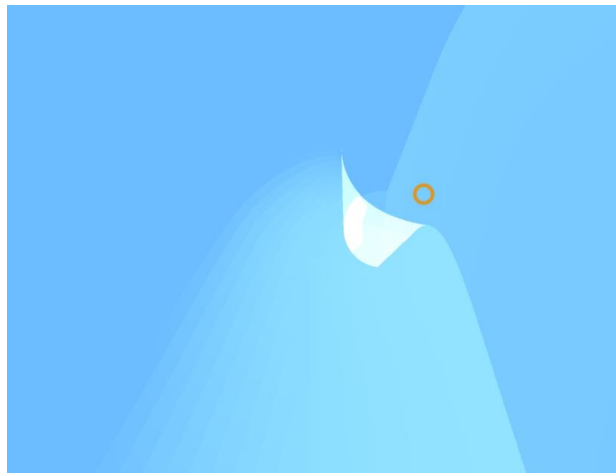
In addition, the tool automatically provides some form of a collaborative design environment. The first-person view of the headset wearer can also be simultaneously streamed on the PC display as well. As a consequence, multiple designers can look at the steps undertaken by the user and easily collaborate having a verbal dialogue between themselves and the person using the headset. This capability was observed by the authors while working on the design and verification of the IPCP tool presented in this paper. The advantages of such setup will be further studied in a series of a more controlled user studies.



(a)



(b)



(c)

Figure 17. After the selection of a data item on the IPCP (a) the user is shown and can inspect additional information associated with that particular design (a) the semitransparent s-duct CAD geometry and accompanying 2D plot representing the flow going through the s-duct. (b) Shows the same CAD geometry closely inspected by the user from one of its sides, and (c) from within the model itself.

VI. Discussion and Conclusions

A previous study conducted by Johansson et al.²¹ indicated that 3D PCP was, in fact, worse than a conventional 2D version in both the metrics (i.e., response times and error rates). However, the study was conducted using a desktop computer connected to an external display and it is unclear whether the findings generalize to truly immersive environment.

The qualitative studies in this paper show tentative evidence toward the possibility that immersing the user within PCP in a VR environment is, at the very least, *complementary* and may provide ways to *enhance* the user's ability to identify patterns. We therefore conclude that immersive 3D PCP can potentially be considered as a feasible way of visualizing high-dimensional data. Most of the participants throughout our qualitative studies managed to identify the same patterns as recognized as interesting in Kipouros et al.,¹ which can be regarded as the reference standard. Moreover, the two participants working with a substantially larger dataset in terms of its complexity and size, i.e. the number of dimensions per data item and the total number of data items, found similar patterns as those given in D'Ambros et al.³⁷ and identified additional candidates for potential patterns. This demonstrates the potential for immersive VR in engineering design processes.

The main contribution of this paper is the task analysis carried out at the functional level and the translation of these functions into a solutions that allow the system to be realized. This scaffolding structure is useful in framing the findings from the preliminary qualitative user study in such a way that the outcomes of the user study could directly shape the refinement of the initial design.

The evaluation with a group of surrogate expert users, P8–P9 for the smaller and P10–P11 for the larger dataset, revealed that users were indeed capable of identifying patterns in both datasets. However, while this evaluation demonstrates the efficacy of the system and serves to provide some evidence to consider the potential merits of 3D PCP, we advise caution against over-interpreting these results given the qualitative aspect of the findings. The sample size is very small and, as a consequence, the findings are highly sensitive to the individual participants’ general ability to use novel immersive technology to analyze a demanding real-world multidimensional dataset. Future work includes conducting a statistically robust quantitative study that is able to tease out the merits and flows of immersive PCP visualization compared with the status-quo 2D PCP visualization. However, such studies are difficult to carry out due to the many controllable and uncontrollable parameters of the system and the experimental task, which makes the design of a robust and meaningful experimental design highly challenging. Prior work has reflected about this problem, both in general for evaluating visualization tools^{42,43} and specifically for the problem of comparing 2D and 3D visualizations.⁴⁴

We believe a promising area of future work is to further refine the system design by evolving the task model and the underpinning functions and revisit the process of translating functions to solutions. This can be achieved by identifying controllable and uncontrollable parameters for each overall function necessary to carry out a task and study the effects of modifying these parameters in isolation. For example, the 3D immersive environment opens up a plethora of methods for selecting data points using input techniques such as head-tracking, gaze-tracking, touch controllers, game controllers or direct manipulation mechanics. At the same time, selection feedback can be provided using a variety of output techniques, such as color-coding, shape modification, sound, haptic feedback, or a combination. Our system implementation can aid such future explorations of how to make informed decisions on, for instance, which combination of inputs and outputs to choose to arrive at an effective operating point for immersive data analysis and thereby assist in gaining a more complete understanding of the trade-offs and their effects that are typically inherent in such design decisions. Such technological capability will be inherently harmonized with the state-of-the-art interactive computational engineering design methodologies⁴⁵ opening new avenues of research for the next generation engineering design software environment and systems.

Acknowledgments

The authors would like to acknowledge the support of the Cambridge European and Trinity Hall and the Engineering and Physical Sciences Research Council (EPSRC-1788814) scholarships. The authors would also like to thank Davide Dal Magro for providing us with the S-Duct CAD geometry and associated flow plots and to John Dudley and Michał Drewniak for proof reading the manuscript. We also appreciate suggestions from Szymon Bobek in regard to the selection of potential clustering algorithms.

References

- ¹Kipouros, T., Inselberg, A., Parks, G. T., and Savill, M., “Parallel Coordinates in Computational Engineering Design,” *AIAA Multidisciplinary Design Optimization Specialists*, Vol. 1750, April 2013.
- ²Kipouros, T., Jaeggi, D. M., Dawes, W N, P. G. T., Savill, M., and Clarkson, P. J., “Bi-Objective Design Optimization for Axial Compressors Using Tabu Search,” *AIAA Journal*, Vol. 46, No. 3, March 2008, pp. 701–711.
- ³Kipouros, T., Jaeggi, D. M., Dawes, W N, P. G. T., Savill, M., and Clarkson, P. J., “Insight Into High-Quality Aerodynamic Design Spaces through Multi-Objective Optimization,” *CMES: Computer Modeling in Engineering and Sciences*, Vol. 37, No. 1, March 2008, pp. 1–44.
- ⁴Tadeja, S. K., Kipouros, T., and Kristensson, P. O., “Exploring Parallel Coordinates in Virtual Reality,” *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA ’19)*, Glasgow, Scotland UK, May 2019.
- ⁵Tadeja, S. K., Seshadri, P., and Kristensson, P. O., “Exploring Aerospace Design in Virtual Reality with Dimension Reduction,” *Proceedings of AIAA SciTech Forum and Exposition*, 7 - 11 Jan 2019 San Diego, Jan. 2019.
- ⁶Tadeja, S. K., Seshadri, P., and Kristensson, P. O., “AeroVR: Immersive Visualization System for Aerospace Design,” <https://arxiv.org/abs/1910.09800>, Oct. 2019 (*under review*).
- ⁷Unity3D Game Engine, “Unity VR Samples pack,” <https://assetstore.unity.com/packages/essentials/tutorial-projects/vr-samples-51519>, Last accessed: November 2018.

- ⁸Oculus VR, “Oculus Utilities for Unity 5,” <https://developer.oculus.com>, Last accessed: November 2018.
- ⁹Greyman studios S.L., “Off Screen Indicator,” <https://assetstore.unity.com/packages/tools/gui/off-screen-indicator-57062>, Last accessed: July 2017.
- ¹⁰Inselberg, A., “The plane with parallel coordinates,” *The Visual Computer*, Vol. 1, No. 2, Aug. 1985, pp. 69–91.
- ¹¹Inselberg, A., *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*, Springer, New York, USA, 1st ed., 2009.
- ¹²Wegenkittl, R., Löffelmann, H., and Groller, E., “Visualizing the behaviour of higher dimensional dynamical systems,” *Proceedings. Visualization '97 (Cat. No. 97CB36155)*, Oct. 1997, pp. 119–125.
- ¹³Gröller, E., Löffelmann, H., and Wegenkittl, R., “Visualization of Analytically Defined Dynamical Systems,” *Scientific Visualization Conference (dagstuhl '97)*, June 1997, pp. 71–71.
- ¹⁴Streit, M., Ecker, R. C., sterreicher, K., Steiner, G. E., Bischof, H., Bangert, C., Kopp, T., and Rogojanu, R., “3D parallel coordinate systems A new data visualization method in the context of microscopy-based multicolor tissue cytometry,” *Cytometry Part A*, Vol. 69A, No. 7, July 2006, pp. 601–611.
- ¹⁵Falkman, G., “Information visualisation in clinical Odontology: multidimensional analysis and interactive data exploration,” *Artificial Intelligence in Medicine*, Vol. 22, No. 2, May 2001, pp. 133–158.
- ¹⁶Ribarsky, W., Ayers, E., Eble, J., and Mukherjea, S., “Glyphmaker: creating customized visualizations of complex data,” *Computer*, Vol. 27, No. 7, 1994, pp. 57–64.
- ¹⁷Fanea, E., Carpendale, S., and Isenberg, T., “An interactive 3D integration of parallel coordinates and star glyphs,” *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, Oct. 2005, pp. 149–156.
- ¹⁸Johansson, J., Ljung, P., Jern, M., and Cooper, M., “Revealing Structure in Visualizations of Dense 2D and 3D Parallel Coordinates,” *Information Visualization*, Vol. 5, No. 2, June 2006, pp. 125–136.
- ¹⁹Dang, T. N., Wilkinson, L., and Anand, A., “Stacking Graphic Elements to Avoid Over-Plotting,” *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 6, Nov. 2010, pp. 1044–1052.
- ²⁰Chang, C., Dwyer, T., and Marriott, K., “An Evaluation of Perceptually Complementary Views for Multivariate Data,” *2018 IEEE Pacific Visualization Symposium (PacificVis)*, April 2018, pp. 195–204.
- ²¹Johansson, J., Forsell, C., and Cooper, M., “On the usability of three-dimensional display in parallel coordinates: Evaluating the efficiency of identifying two-dimensional relationships,” *Information Visualization*, Vol. 13, No. 1, Jan. 2014, pp. 29–41.
- ²²Holten, D. and Wijk, J. J. V., “Evaluation of Cluster Identification Performance for Different PCP Variants,” *Computer Graphics Forum*, Vol. 29, No. 3, 2010, pp. 793–802.
- ²³Cordeil, M., Cunningham, A., Dwyer, T., Thomas, B. H., and Marriott, K., “ImAxes: Immersive Axes As Embodied Affordances for Interactive Multivariate Data Visualisation,” *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, ACM, New York, NY, USA, 2017, pp. 71–83.
- ²⁴Butscher, S., Hubenschmid, S., Mller, J., Fuchs, J., and Reiterer, H., “Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data,” *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, ACM, New York, NY, USA, 2018, pp. 90:1–90:12.
- ²⁵Rosenbaum, R., Bottleson, J., Liu, Z., and Hamann, B., “Involve Me and I Will Understand!: Abstract Data Visualization in Immersive Environments,” *Proceedings of the 7th International Conference on Advances in Visual Computing - Volume Part I*, ISVC'11, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 530–540.
- ²⁶Ribarsky, W., Bolter, J., Op Den Bosch, A., and Van Teylingen, R., “Visualization and Analysis Using Virtual Reality,” *Computer Graphics and Applications, IEEE*, Vol. 14, Feb. 1994, pp. 10–12.
- ²⁷Ishihara, S., *Ishihara's Tests for Colour Deficiency*, Kanehara Trading Inc, Tokyo, Japan, 38th ed., 2017.
- ²⁸Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G., “Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness,” *The International Journal of Aviation Psychology*, Vol. 3, No. 3, July 1993, pp. 203–220.
- ²⁹Shneiderman, B. and Plaisant, C., *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Addison-Wesley Publishing Company, USA, 5th ed., 2009.
- ³⁰Shelfelbine, S., Clarkson, J., Farmer, R., and Eason, S., *Good Design Practice for Medical Devices and Equipment - Requirements Capture*, University of Cambridge Engineering Design Centre and University of Cambridge Institute for Manufacturing, 2002.
- ³¹Wegman, E. J. and Luo, Q., “High Dimensional Clustering Using Parallel Coordinates and the Grand Tour,” *Classification and Knowledge Organization*, edited by R. Klar and O. Opitz, Studies in Classification, Data Analysis, and Knowledge Organization, Springer Berlin Heidelberg, 1997, pp. 93–101.
- ³²Heinrich, J. and Weiskopf, D., “Parallel Coordinates for Multidimensional Data Visualization: Basic Concepts,” *Computing in Science Engineering*, Vol. 17, No. 3, May 2015, pp. 70–76.
- ³³Artero, A. O., Oliveira, M. C. F. d., and Levkowitz, H., “Uncovering Clusters in Crowded Parallel Coordinates Visualizations,” *IEEE Symposium on Information Visualization*, Oct. 2004, pp. 81–88.
- ³⁴Carroll, J. M. and Carrithers, C., “Training wheels in a user interface,” *Communications of the ACM*, Vol. 27, No. 8, 1984, pp. 800–806.
- ³⁵Nielsen, J., “Usability Inspection Methods,” *Conference Companion on Human Factors in Computing Systems*, CHI '94, ACM, New York, NY, USA, 1994, pp. 413–414.
- ³⁶Nielsen, J., “Enhancing the Explanatory Power of Usability Heuristics,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, ACM, New York, NY, USA, 1994, pp. 152–158.
- ³⁷D'Ambros, A., Kipouros, T., Zachos, P., Savill, M., and Benini, E., “Computational Design Optimization for S-Ducts,” *Designs*, Vol. 2, No. 36, October 2018, pp. 1–21.
- ³⁸Grossman, T. and Balakrishnan, R., “The Bubble Cursor: Enhancing Target Acquisition by Dynamic Resizing of the Cursor's Activation Area,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, ACM, 2005, pp. 281–290, event-place: Portland, Oregon, USA.

³⁹Unity Game Engine, “Unity,” <https://unity.com/>, Last accessed: July 2018.

⁴⁰Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J., “OPTICS: Ordering Points to Identify the Clustering Structure,” *SIGMOD Rec.*, Vol. 28, No. 2, June 1999, pp. 49–60.

⁴¹Ester, M., Kriegel, H.-P., Sander, J., and Xu, X., “A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD’96, AAAI Press, 1996, pp. 226–231.

⁴²Tory, M. and Moller, T., “Human factors in visualization research,” *IEEE transactions on visualization and computer graphics*, Vol. 10, No. 1, 2004, pp. 72–84.

⁴³Van Wijk, J. J., “Views on visualization,” *IEEE transactions on visualization and computer graphics*, Vol. 12, No. 4, 2006, pp. 421–432.

⁴⁴Kristensson, P. O., Dahlback, N., Anundi, D., Bjornstad, M., Gillberg, H., Haraldsson, J., Martensson, I., Nordvall, M., and Stahl, J., “An evaluation of space time cube representation of spatiotemporal patterns,” *IEEE Transactions on visualization and computer graphics*, Vol. 15, No. 4, 2009, pp. 696–702.

⁴⁵Hettenhausen, J., Lewis, A., Randall, M., and Kipouros, T., “Interactive multi-objective particle swarm optimisation using decision space interaction,” *2013 IEEE Congress on Evolutionary Computation*, June 2013, pp. 3411–3418.