Interactive Data Exploration and Knowledge Discovery

Elena Zudilova-Seinstra  
University of Amsterdam  
Computational Science Department, Informatics Institute  
Science Park 107, 1098 XG Amsterdam, The Netherlands  
e.v.zudilova-seinstra@uva.nl

Jean-Bernard Martens  
Eindhoven University of Technology, User-Centered Engineering Group, Faculty of Industrial Design  
P.O. Box 513, 5600 MB Eindhoven, The Netherlands  
j.b.o.s.martens@tue.nl

Tony Adriaanssen  
PROIMIM  
PO Box 527, Cannon Hill, QLD 4170, Australia  
tony.adriaanssen@gmail.com

1. INTRODUCTION
People have always relied on visual tools such as maps, charts and diagrams to better understand problems and solve them in less time. Continuous improvements in computer processing power and graphics capabilities have made it possible to incorporate a wide range of advanced visualization techniques in most computing application domains, including business, medicine, engineering and science.

Visualization technologies empower users to perceive important patterns in data, identify areas that need further scrutiny and make sophisticated decisions. However, without interactivity, visualization is often considered as an end point of the workflow, or as a way of communicating observations. Massive increases in data sizes have created an urgent need for new systems that enable users to explore complex datasets and computations in an intuitive and flexible manner. The way people perceive and interact with visualizations can strongly influence their understanding of the data as well as the usefulness of a visualization system in general.

2. INTERACTIVE DATA EXPLORATION
The intuitive exploration of complex data is only possible when advanced interaction and visualization techniques are integrated. When combined properly the ability to visualise and interact with the data can aid analysis and understanding in many areas, such as: scientific experiments, manufacturing process control, financial data analysis, etc.

With the rapid development of advanced visual interfaces our work and learning have become more efficient and fascinating. The maturing virtual reality (VR) techniques, together with the emerging haptic interfaces and multimedia networking technologies provide ground for new enabling tools addressing computer-supported activities. They open the way to new forms of collaborative work and new domains of multi-participant systems. Today’s VR-based collaborative visualization and exploration environments are becoming the established ways of facilitating human interaction with large amounts of information [1]. They allow opinions to be shared, removing the data bottleneck of individual analysis and reducing the time to discovery.

Multimodal interaction systems have become very popular as well. Based on the combination of different interaction techniques (i.e., direct manipulation, speech recognition, haptics, real time video and audio, etc.), multimodal systems aim to provide efficient, convenient and natural interaction and communication between computer systems and users in a seamless way and will ultimately enable people to interact more fully within an exploration environment using everyday skills [4].

Another important requisite in the interaction design is aimed at emphasizing ‘human-to-human’ properties, so called social user interfaces. This kind of design considers human emotions and personality, e.g. face-to-face communication between users, embodied agents, etc. Embodied agents usually interact with users or each other via multi-modal communicative acts, which can be non-verbal (virtual embodied agents) or verbal (conversational embodied agents). They permit building a kind of relationship with an interactive environment as well as with other users to assist in the exploration process.

3. CHALLENGES
Important progress has been made recently toward the improvement of a computer-supported information analysis cycle. However, many challenges still remain.

The greatest challenge is how to efficiently deal with a community of scientists who generate more data than they can possibly look at and understand [2]. This requires novel feature extraction techniques and high-performance visualization algorithms. The extreme growth in data sizes has resulted in research on combining real-time computing and visualization.

With the expanding volume of visual data, there is an urgent need to develop novel interaction methods that maximize the user's understanding of the visual information [6]. In this respect, the challenge is to leverage what is known about human perception so as to design user interfaces in a way that a typical viewer can intuitively extract the greatest amount of accurate and relevant information.

Finding effective visual idioms for direct user interaction is yet another challenge [3]. A major problem here is that user interfaces...
that seem natural from the developer’s point of view often feel unnatural to users; and those that feel natural to users seem peculiar and full of special cases to visualization experts.

Even though advanced interaction technologies are applied more and more often to the field of scientific and information visualization, the integration of visualization and interaction methods is still an open question. The relatively new concept of interactive visualization [7] introduced recently aims to address this research concern. The purpose of interactive visualization is not only to provide users with a possibility to view the data and modify representation parameters but also to permit them using interaction abilities for the interrogation and navigation through datasets and communication these insights with others.

To permit users' intuitive, as well as collaborative exploration, interaction and visualization capabilities need to be optimized so that access to the data and associated features will become apparent. Hence, it is becoming crucial to develop new guidelines and metrics to be able to efficiently evaluate interactive visualization environments and systems. Another concern is how to choose from a wide plethora of available input devices and display configurations to ensure a good user experience and 'human-to-human' interaction between scientists while exploring complex data spaces [5].

4. WORKSHOP DESCRIPTION

The goal of the AVI 2010 Workshop on Interactive Data Exploration and Knowledge Discovery is to bring together academic researchers and practitioners from computer science, human computer interaction (HCI), software engineering, social sciences and psychology to discuss the state-of-the-art research in the field of interactive data exploration and analysis and to exchange ideas to help overcome existing challenges.

The workshop will serve as an international forum for the information exchange on theoretical, generic and applied aspects of advanced visual interaction within the broad scope of scientific and information visualization. It will also offer a possibility to discuss existing interactive visualization tools and systems. The developers will have an opportunity to show their products and research prototypes, while potential users can pose their questions and requests.

The main questions to be discussed during the workshop are as follows. How can interactive visualization methods and tools be augmented to address both concerns of scientific computing and HCI? What are the criteria for choosing between advanced display and input devices? Which research questions need to be considered when aiming to achieve efficient HCI and 'human-to-human' interaction between scientists while exploring complex data spaces? In what way users and the environment of use should be modeled when designing interactive visualization systems? Which factors contribute to user preference and acceptance? How can we define effective abstractions for the visualization and user interaction processes? What is the impact of task and application field-orientation on interactive visualization? To what extent are usability problems independent of the context of use and need to be taken into account when designing interactive exploration environments?

The workshop will be organized in individual presentations, problem-oriented group activities and a final round-table discussion of results. All workshop submissions were peer-reviewed and judged based on their scientific quality and application value. Eight extended abstracts accepted for the research presentations during the workshop are included in the proceedings and can be found below.

This workshop is the fifth in a sequence of events around the theme that have run over the last few years. The previous related workshops include:

- The HCI 2006 Workshop on Combining Visualisation and Interaction to Facilitate Scientific Exploration and Discovery (London, UK);
- The ICMI 2005 International Workshop on Multimodal Interaction for the Visualization and Exploration of Scientific Data (Trento, Italy);
- The ICCS 2004 International Workshop on Interactive Visualization and Interaction Technologies (Krakow, Poland);
- The ICCS 2003 International Workshop on Scientific Visualisation and Human-Machine Interaction in Problem Solving Environments (Melbourne, Australia).

The detailed information about the AVI 2010 Workshop on Interactive Data Exploration and Knowledge Discovery (including the workshop program) can be found at: [8].

5. ACKNOWLEDGEMENTS

We would like to thank all authors for submitting their work to the workshop, the members of the Workshop Program Committee for their efforts in the reviewing process and the AVI 2010 Workshop Chair Professor Luciano Gamberini for his help in dissemination.

REFERENCES

Exploring Concepts Collaboratively: Considering how Wii interact

Christopher Foster
Durham University
School of Engineering and Computer Sciences
South Road, Durham, UK
Tel: +44191 33 41704
c.r.foster@durham.ac.uk

Liz Burd
Durham University
School of Engineering and Computer Sciences
South Road, Durham, UK
Tel: +44191 33 41719
liz.burd@durham.ac.uk

Andrew Hatch
Durham University
School of Engineering and Computer Sciences
South Road, Durham, UK
Tel: +44191 33 41706
andrew.hatch@durham.ac.uk

EXTENDED ABSTRACT
This abstract describes ongoing research on human to human interaction between early career Computer Scientists as they explore a complex collaborative concept mapping task performed by a collocated group using a large wall-screen projected display. We have investigated the effects of input configuration and mode of input on human to human computer interaction through the use of gesture based controllers. Application of Bales’ Interaction Process Analysis (IPA) [1] supports identity trade offs in choosing from a plethora of available input devices and displays when investigating interaction and knowledge discovery.

Increasingly, Higher Educational departments are encouraged to plan and deploy technologies that facilitate interaction as a fundamental principle of moving into the Interaction Age, specifically ‘new tools are needed to support informal learning activities, in particular processes associated with conceptual development’ [2]. This need for new tools resulted in the creation of WiiDraw; software through which single or multiple users could interact with and manipulate conceptual mapping diagrams using gestural interaction concurrently. Gaming interfaces like the Nintendo Wii provide options for creating gesture-based input beyond the move-click capability of a mouse, offering new modes to groups who interact with and create conceptual knowledge with large screens.

Other example systems have used mouse input [3] and laser pointers [4] as a means through which to explore large-screen based systems, yet none of these have emerged as a clear choice for a range of applications, and there may not be a single best fit option. However, we now present a study of eleven groups who completed a conceptual mapping task on a shared wall-display to determine how the configuration and mode of input influenced the amount of interaction. The experiment consisted of a single between-groups factor of input configuration of two levels (one controller and two controllers) and a single within-groups factor of interaction style, consisting of two levels (controller with no gestures enabled and controller with gestures enabled). IPA was applied to the data that was obtained from video recordings made of each experiment. ANOVA indicated a main effect of number of controllers, F(1,18)=6.38, p<0.02, with a higher number of interactions when dyads had one controller (M=432, SD=93) than two controllers (M=310, SD=140). A main effect of gestures was evident, F(1,18)=5.08, p=0.04 with more interactions occurring with gestures (M=420, SD=119), than without (M=310, SD=129) (see Figure 1). The interaction effect was not statistically significant.

Results indicate that one controller afforded higher levels of human to human interaction, with gestures also increasing the number of interactions seen. Further analysis describes the differences in type of interaction and its impact upon knowledge discovery. However it appears Wii interact more when gesturing with concepts.

REFERENCES
Generation of Roadside Panoramic Images without Obstacles

Shunichiro Tsuji
School of Science & Technology
Kwansei Gakuin University
2-1 Gakuen, Sanda, Japan
coi29394@kwansai.ac.jp

Yasuyuki Kono
School of Science & Technology
Kwansei Gakuin University
2-1 Gakuen, Sanda, Japan
kono@kwansai.ac.jp

Categories and Subject Descriptors
I.4 [Image Processing and Computer Vision]: Applications; H.5 [Information Interfaces and Presentation]: Multimedia
Information Systems - Hypertext navigation and maps

General Terms
Design feature, Human Factor

Keywords
SIFT feature, real-world map, panoramic image generation, obstacle exclusion

1. INTRODUCTION
This paper presents the method for generating panoramic image(s) from plural in-vehicle camera images. Our real-world map system employing the proposed method has the following features: 1) it puts the generated panoramic images of the real-world to an existing web map, e.g., GoogleMaps, 2) partial regions in the images are classified according to the depth from the camera by tracking each partial region employing the SIFT feature, so that 3) it can obtain panoramic background images that exclude obstacles, e.g., parked vehicles. Our updatable real-world map system can offer up-to-date information to map-based systems, e.g., car navigation system or web interface any time.

2. DEPTH-BASED CLASSIFICATION
A monocular video camera is mounted in the center of the left side window of a test vehicle and faced towards the left side of the car. The in-vehicle camera captures the scenery of the road side while driving. A GPS receiver is also mounted in the vehicle and each captured image is annotated by its location information. Figure 1 shows sample images of every 5 frames. Two pickup trucks are parked on the roadside and occlude the background buildings as shown in Figure 1.

Our idea is to cluster feature points into similar depth regions according to their movement speed between frames.

The distance of corresponding feature points between two frame images reflects the depth to the region they belong to, i.e., a feature point moves faster if it is nearer to the camera. We can classify the projected objects in the captured images into similar depth regions by grouping feature points according to their movement speed. We employ the SIFT feature extraction method to correctly associate corresponding feature points between frames.

Figure 1. In-vehicle camera images

Figure 2. Generated panoramic image

Figure 2 shows the panoramic image of the background buildings generated by unifying the slowest regions according to the amount of movement of feature points. The unification is overall successful, although the system could not recover the region where the obstacles fully occluded. Also some part of the background was misclassified as vehicle region and some of the vehicle as the background mainly around the border of the regions.

3. FUTURE WORK
This paper presented the method for generating panoramic image by classifying regions in frame images according to the amount of movement. The method is simple and was overall successful. However, it is necessary to accurately divide regions on the border of the subjects. Also to update the generated images by following the change of the real-world, the system should detect the changes of the background.
1. INTRODUCTION AND BACKGROUND
We use, work, and think with information to be able to make decisions, solve problems, plan, analyze, forecast, and learn. These information-based activities can be classified as epistemic activities (EAs), as they all involve or are related to knowing. Often, we use computational tools in the form of Visual Cognitive Tools (VCTs) to help us carry out these EAs. VCTs maintain and display information in a digital, visual form to mediate our EAs. They sit at the interface between the mind and an information space and participate in the ‘interplay’ between the two. This paper presents a framework for a systematic exploration of the anatomical structure of the space between the mind and a given information space—that is, the space in which the interplay between the mind and information takes place. The motivation behind the creation of this framework is to facilitate the analysis and design of VCTs that can effectively support our EAs.

Before presenting the framework, some conceptual and terminological background regarding EAs, VCTs, and human-information interaction (HII) is provided here.

Epistemic Activity: An example of an EA is exploring a genome information space to make sense of chromosomal relationships. An EA is broader than a task and may include other activities, tasks, and subtasks—e.g., the EA of triaging a set of documents to discover if they are conceptually related involves tasks such as scanning the documents, extracting information, building associations among similar information items, and comparing these items. These in turn may consist of other information-based actions and further low-level events.

Visual Cognitive Tool: There are many different kinds of VCTs, e.g., information visualization systems, virtual science museums, geographic information systems, and mathematical visualization tools. All VCTs provide users with representations of the elements, properties, and internal relationships of a given information space in structural, logical, causal, and visuospatial forms and allow users to interact with these to carry out their EAs.

Human-Information Interaction: The epistemic locus of a VCT is at its interface, where users interact with representations of an information space. Hence, at its core, any VCT has two closely-related components: representation and interaction, both of which can be regarded as systems. HII design is concerned with what is ‘done’ with the represented information and the provision of mechanisms for ‘acting’ upon the representations. This component involves discourse between the user and the represented information. Although some fundamental concepts regarding representation and interaction design have been in place for a while, we do not really have a principled, scientific understanding of how to design these two components for diverse information spaces, EAs, and VCTs. Several frameworks have been proposed in the context of HCI, information visualization, visual analytics, and interactive visualizations [1,2,3,4]. Some deal with action cycles, some with levels of interaction and its cost, and some with taxonomic categorizations. However, we need frameworks that integrate different components to better support the design of HII for EAs.

2. PROPOSED FRAMEWORK
The proposed framework integrates structural, procedural, and taxonomic components of HII design to support different types of visual representations and EAs in the context of different VCTs. Briefly, the procedural component of HII design is captured through a sensing-interpreting-acting cycle. In this cycle, a user senses a representational system, interprets its current state and acts upon that system through an epistemic interaction. The structural component is captured through the incorporation of anatomical elements that affect the quality of epistemic interactions—including such elements as affordance/constraint, event granularity, focus, action flow, and propagation. Finally, the taxonomic component is captured through classification of the elements of the different subsystems. This component categorizes and characterizes different representational systems (e.g., tables, diagrams, etc.), different epistemic interactions (e.g., annotating, collecting, filtering, probing, etc.), different techniques that instantiate these interactions (e.g., spatial proximity and zooming as instances of probing), different EAs (e.g., problem solving, decision making, etc.), and different VCTs (e.g., visual analytics tools, games, virtual museums, etc.). The integration of all these different components not only allows designers of VCTs to think about the design space more systematically, but also provides them with a palette of design possibilities.

3. REFERENCES
Role of Externalization and Composite Interactions in the Exploration of Complex Visualization Spaces: A Usability Study

Hai-Ning Lia
National ICT Australia (NICTA)
Queensland Research Lab
St. Lucia, Queensland, Australia
+61 73300 8624
hai-ning.liang@nicta.com.au

Kamran Sedig
The University of Western Ontario
London, Ontario, Canada
1-519-661-2111
sedig@uwo.ca

ABSTRACT
Complex visualization spaces are not easy to explore. This is especially true of exploring 4D mathematical structures due to occlusion and visual clutter. In this paper, we present preliminary results involving a usability study of two strategies for designing interaction strategies, namely, externalization and composite interactions, to aid in the exploration of 4D mathematical structures. Study results show the potential benefits of these two strategies for exploring complex mathematical structures, in particular, and other complex visualization spaces, in general.

1. INTRODUCTION
Complex visualization spaces, by necessity, incorporate much implicit information, hidden within a vast amount of other visual elements. One way of facilitating the exploration of implicit, hidden relationships and structures is to add interaction to these visualization spaces.

Interaction adds a temporal dimension to visualization spaces and can potentially extend their epistemic and communicative power [1, 2]. Interaction involves both the actions users perform on the spaces and reactions offered by these spaces in response to those actions. It is interaction which allows users to conduct dynamic probes into the deeper aspects of the spaces. Allowing these probing activities may be conducive to discovering implicit, hidden information in a dialectic, progressive manner. The temporality of the spaces can lead to enhanced exploration, as interaction can greatly increase users’ abilities to navigate through, query, transform, and manipulate different elements of the spaces—actions which are essential for knowledge-oriented activities. There is need to develop interaction strategies which can optimize users’ exploration and, in the process, their understanding of the spaces. The presented research investigates two strategies for designing interactions capable of supporting users’ exploratory and knowledge activities, namely, externalization and composite interactions. Succinctly, by externalization we refer to two things: 1) visualizing available interactions according to the structural possibilities of the space, and 2) propagating the effects of an action in future possible interaction. The composite interactions strategy refers to providing users with multiple interactions that complement each other in order to allow more coordinated and integrated exploratory activities.

2. METHOD
A software application which allowed users to explore 4D mathematical visualization spaces was implemented, and a usability study was conducted with it. The application (see figure) provided users with two main interaction strategies, designed based on strategies of externalization and composite interactions. Twenty-two (n=22) graduate students participated in the study. Using the software application, they worked in pairs and were given exploratory activities involving a set of tasks to complete.

3. PRELIMINARY RESULTS
Our general findings show that the design strategies were supportive of users’ exploration activities. Participants stated that without these design features exploring these complex spaces would not have been possible. On the one hand, externalization of interaction possibilities and their effects assisted many users in the manner which they explored the spaces, helping to reduce their cognitive load and making exploration more strategic. On the other hand, composite interactions enabled participants to perform diverse, yet complementary, operations, offering users the possibility of switching between exploration styles.

In short, the results of this study reveal potential benefits of these two strategies and can inform the design of the interactive features of other complex visualization spaces, besides 4D mathematical structures.

4. REFERENCES
Evaluating the Utilization of Clustering Methods Connected with Multivariate Visualizations

Barbara Scheuner
ETH Zurich
Universitätsstrasse 6
8092 Zurich, Switzerland
+41 44 632 22 11
barbara.scheuner@inf.ethz.ch

EXTENDED ABSTRACT

Visualization software in all fields becomes increasingly interactive with more and more elements to support users. Elements that help an observer to interpret given data sets by computing certain properties of the data to enrich their visualization we call “visual scouts”.

Many visualization programs implement visual scouts, but few is known about how users apply them in a real setting. In order to learn more about how users work with interactive visualization software we provide them with clustering methods (k-mean and hierarchical clustering), as one possibility to manipulate the data. Subsequently we evaluate how natural science students use these clustering methods as provided by an interactive comparative visualization software for high dimensional data (VisuLab®). The students are introduced to four different visualization methods (figure 1) and given the possibility to apply four clustering methods to help them when interpreting two sample datasets (table 1 first two rows). This is part of a guided instruction after which they independently interpret two new but related real data sets (table 1 last two rows). In total they work for about six hours.

During our evaluation we were able to record the clustering activities of ~250 students, by directly logging the user’s activities within the software. Based on this data we computed how many times the data was clustered and with which visualization methods the results were visualized. In total we registered over 57’000 clustering activities of which 30% where recorded during the instruction period and the remaining 70% during the time they worked independently.

Based on the data gained we investigate relationships between clustering methods, the type of data analyzed and the visualization method chosen. This is shown in figure 2, where the number of clustering in regard to the visualization method is presented in a boxplot.

Table 1. Table captions should be placed above the table

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Complexity (dimensions/ data points)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>socioeconomic</td>
<td>7 / 35</td>
<td><a href="http://www.ggobi.org">www.ggobi.org</a></td>
</tr>
<tr>
<td>flea</td>
<td>7 / 74</td>
<td></td>
</tr>
<tr>
<td>iris flowers</td>
<td>4 / 90</td>
<td>Fisher’s iris data</td>
</tr>
<tr>
<td>air pollution</td>
<td>16/ ~90</td>
<td><a href="http://www.bafu.ch">www.bafu.ch</a></td>
</tr>
</tbody>
</table>

![Figure 1. VisuLab® graphics: 1) Andrews’ Curves, 2) Scatterplot Matrix, 3) Permutation Matrix, 4) Parallel Coordinates](image1.png)

![Figure 2. Mean number of clustering in each visualization](image2.png)
Virtual Environment for the Navigation of Ideas and Concepts in Education (V.E.N.I.C.E)

Simon Foley
Royal Holloway, University of London
Egham, Surrey TW20 0EX, UK
01784 (41)4296
simon.foley@rhul.ac.uk

This paper reports on current research into the development of an interactive visualisation tool for postgraduate research students. The aim of the research is to construct a virtual environment that allows students to navigate ideas, which they can then relate to conceptual structures. V.E.N.I.C.E offers users a virtual learning environment, which is at once a concept map, a file manager, and a memory palace. Modeled loosely on the infrastructure of the medieval city of Venice the application encourages users to build metaphorical relationships between conceptual islands of ideas, that are separated by canals, but linked by bridges, which in turn lead to further islands. On each island the user constructs a memory palace, which acts as a repository for data associated with a single idea. The palace is located within a confluence of discursive pathways that allow the student to position, affirm, challenge and expand upon a line of related argument.

Whilst there have been extensive advances in applications that aim to facilitate deeper learning techniques through the utilisation of visualisation, pedagogic agents, concept mapping, mnemonic recall, and data organisation [1, 2, 3, 4, 5, 6, 7, 8], there is as yet no single application, which brings these techniques together. The experience of many supervisors and tutors of research methodology courses is that students at the early stage of their research can struggle to convert dissertation ideas into projects that are theoretically robust. There is a need therefore for a tool that can, not only provide intuitive guidance, but also allow the student to dynamically structure their project.

V.E.N.I.C.E differs from existing concept mapping tools in that instead of encouraging users to map hierarchical relationships between different conceptual domains, it prompts the user to establish and expand upon a position generated from an initial idea. For example, a user may have a vague project idea to explore the meaning of happiness. The user is prompted by a pedagogical agent (PA) to explain why they have chosen this idea. The PA requests they establish an initial position e.g. “happiness is related to the accumulation of wealth”. The user is then asked to affirm their position by completing the following prompt “the literature supports the view that...”and the user may argue, “people in richer nations are happier because they live longer”. A further prompt encourages the student to challenge this statement with “however, there is evidence to suggest...” and the counter argument may be “poorer nations have lower suicide rates”. The user is then encouraged by the PA to re-assess their original position with the prompt “it can be argued therefore...” and the student concludes, “the pursuit of happiness is related to lifestyle”.

At this point, the student can juxtapose their original position “happiness is related to the accumulation of wealth” with “the pursuit of happiness is related to lifestyle”. The PA then offers a bridge to a different conceptual island where they can expand on their original idea “to explore the meaning of happiness” with the idea to explore how “lifestyle factors influence the perception of happiness in developed countries”. The student is thus encouraged to acknowledge that by choosing to research a deceptively straightforward idea, the succeeding journey is not without intellectual challenges. Indeed the route they must follow often demand they negotiate a range of previously unforeseen variables. In this way the student is encouraged to read, critically evaluate and establish a theoretical position on which to base their research.

The paper will be of particular interest to those interested in constructivist approaches to education. It will not only demonstrate how visualisation techniques can promote the construction and development of theoretical frameworks, but also how visualisation can aid recall, and encourage dynamic interaction with applications, which promote deeper learning.

REFERENCES
Many real-world networks are multivariate, i.e., they have attributes associated with nodes and/or edges. Examples include social networks whose nodes represent people and edges represent relationships. There is usually information about each person (such as name, age, and gender) and the relationship (such type, duration, and strength). Besides common graph analysis tasks (such as identifying the most influential or structurally important nodes), there are more complex analyses for multivariate networks. One of these is the multivariate graph clustering, i.e., identifying clusters formed by nodes that have similar attributes and are close to each other in terms of graph distance. For instance, in social network analysis, it is interesting to sociologists whether or not people with similar characteristics (node attributes) are also connected to each other. Currently there are very few visualization methods available for such analysis.

Graph and multivariate visualization have been well studied separately in the literature. Herman et al. summarized the recent work on graph visualization [3], and Wong and Bergeron covered the development in multivariate visualization [4]. However, there is relatively less work available on multivariate network visualization. Two types of approaches are commonly used. The first one is the mapping approach, which maps attributes to visual elements of a node or edge. A simple example is to map one attribute to node size and another to node color [2]. A more advanced mapping approach uses glyphs to represent node or edge attributes. One such example is to use the length and width of a rectangle node glyph to represent two node attributes [1]. The second one is the 2.5D approach: it uses the third dimension to present the multivariate information, while the graph is shown on a 2D plane. Examples include the recently proposed GraphScape [5], which adopts a landscape metaphor: each attribute is represented by a two-and-a-half-dimensional surface, whose height indicates its value.

Each approach has its strength and weakness. The mapping approach is effective of showing numerical value using visual element such as size, but it can be difficult to compare the value of attributes represented by different elements such as size and color. The problem is alleviated by a carefully designed glyph, but visual complexity increases quickly as the number of attributes that a glyph needs to represent grows. The 2.5D approach is good at showing the distribution of attribute values over the network, but the attribute surface could introduce occlusion and affect the visibility of underlying network.

In this paper, we present a study evaluating the effectiveness of these two approaches for different analysis tasks. We compare the performance of mapping and 2.5D approach in a controlled lab environment. We included both simple tasks (such as identifying nodes with the largest attribute value) and complex tasks (such as multivariate graph clustering). The performance is measured both in terms of accuracy and completion time. The results indicate that statistically mapping approach performs better for the simple tasks, while the 2.5D approach is favored in the complex task. The outcomes from this study provide some guidelines for the design of effective multivariate graph visualization for different analysis tasks.

REFERENCES
An Interactive Exploration Environment for Complex Process Design

Eric S. Fraga
Centre for Process Systems Engineering
University College London (UCL), UK
e.fraga@ucl.ac.uk

1. INTRODUCTION
Process design requires the solution of complex models. These models are typically nonlinear and non-convex. The design optimisation problem is also often combinatorial in nature. An example is the design of thermally integrated process plants for energy minimisation where excess heat from one part of a process plant may be used to meet heating demands in other parts of the same plant (cf. e.g. [2]). Due to the complexity of the models and the combinatorial nature of the search, design is often undertaken as a series of steps with increasing detail from step to step [1]. Information gained in any step can be useful in subsequent steps. This information may come from data generated during design. Therefore, data analysis and visualisation are key elements in effective design.

2. DESIGN VISUALISATION
A prototype Java3d based interactive graphical environment has been developed which presents data corresponding to discrete alternative designs generated in the earliest steps. The visualisation of these designs, residing in a high dimensional space, is addressed by the generation of an undirected graph, where individual discrete designs are connected to other designs based on their proximity in the design domain. This differs from the traditional approach of using the design variables directly without any interpretation. The graph can be displayed in a two-dimensional view, highlighting the connectivity, and also in a three-dimensional view, emphasizing the objective function values for the alternative designs.

The graphical environment is based on a model-view-control pattern; the figure shows the control window on the left with the new three-dimensional view of the design space on the right. This particular design space was initially explored with the aim of minimising an annualised cost based without consideration of thermal integration possibilities. The post-optimisation interactive stage allows the engineer to explore discrete designs incorporating the potential impact of thermal integration into the objective function.

Each processing unit for each alternative design for this problem is represented by a sphere. Neighbouring designs are selected based on their similarity in design variable values. The current design is highlighted through relative sizing; neighbouring designs can be identified which are better or worse than the current design. The engineer can move the focus to any neighbouring solution. The engineer can explore the space and select designs which are worth further consideration. The automatic resizing and re-centring of the graph caters for large design spaces effectively.

Figure 1. The left control window [3] presents the heating and cooling requirements for the currently selected design point; the right shows the 3-d view of the discrete design space with connectivity and objective function values.

3. CONCLUSIONS
An interactive environment with a novel graph based mapping of a high-dimensional space to a 3d representation has been developed. This allows an engineer to identify potentially good designs quickly and effectively, even for large and high dimension design problems. The concept of neighbouring solutions is used as the basis of the visualisation. This allows an engineer to see how different designs are related and how the objectives may be better met.

REFERENCES