3D OBJECT CLOUDS: VIEWING VIRTUAL OBJECTS IN INTERACTIVE CLOUDS ON MOBILE DEVICES

by

Xiaoting Hong

Submitted in partial fulfilment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
December, 2016

Dedication

I dedicate this thesis to my mother and father for raising me up and making me who I am with their endless love, support, and trust.

Table of Contents

List of Tables

List of Figures

Abstract

Given the expanding use of 3D Objects in a variety of fields and the “find rather than create” trend, we proposed a novel system for object browsing and searching. Specifically, the system packs 3D objects into an interactive cloud view for browsing and searching. We designed it as a novel way to increase searching efficiency and engagement while providing a visually-engaging layout, and we later evaluated this by conducting user studies. We presented that our system can significantly decrease search time compared to the classic grid-based layout, and it has been suggested by subjects that they feel cloud-based searching is more interesting and visually-engaging.

List of Abbreviations Used

ANOVA	Analysis of Variance
CV	Cloud-based layout system
GV	Grid-based layout system
SA	Staggered animation
NSA	Non-staggered animation

Acknowledgments

First, I wish to express my sincere gratitude to my supervisor Dr. Stephen Brooks. He provided me the valuable opportunity to study at Dalhousie University, and he always inspires and guides me with brilliant ideas and endless patience. I won't accomplish this research without his guidance.

I would like to thank all the lab members in the GEM lab. Special thanks to Dong Gao, Muhammad Salimian, Yun Cai, Huiyuan Zhou and Aziz Abdurexit, for being the greatest friends and always be there when I need help.

Thank you to my partner Qihao Zhang, for all his love and support.

Thanks to my thesis committee Dr. Bonnie Mackay and Dr. Kirstie Hawkey, for their inspiring questions and comments.

Finally, I would like to thank the Boeing Company for providing the financial support for me to conduct my research.

Chapter 1

Introduction

1.1 Background and Motivation

The usage of 3D objects has become wide and prevalent in both commercial and academic fields within the current technical environment. Current trends such as inexpensive and practical graphic devices and the rapid expansion of the web enables the use of 3D models for a variety of applications and users. These include visual reality applications (Yu, et al., 2011), complex 3D scene construction (Merrell & Manocha, 2011), and 3D printing.

Considering that constructing 3D models consumes a lot of time and effort, and requires professional skills, many users prefer to find the desired 3D object rather than create it themselves. This trend leads to the desire for an effective and efficient approach to the browsing and searching of 3D objects. Several academic and commercial systems have explored this area. In terms of academic work, some shape-based search engines for 3D models have been proposed (Funkhouser, et al., 2003) (Chen, Tian, Shen, & Ouhyoung, 2003) where users can search their desired 3D models by providing natural queries (such as 3D sketching, 2D sketching, etc.). These focused on providing a search engine based on the “shape” feature because it’s considered to be the most important factor from a human’s perspective. They also provide methods to extract and match the features to achieve the real-time searching goal. On the other hand, in terms of commercial usage, there many browsing and searching services of 3D models, including online web browsing (Yobi, 3D Warehouse, etc.), 3D model galleries in game engines (Unity, etc.), and 3D modelers (Google Sketchup, etc).

In both the academic and commercial works described above, they typically adopt a classic grid-based view to present the intermediate searching results which is a collection of 3D objects. We found there is little that has been explored about how to better organize a set of 3D objects to facilitate the comparison and recognition so that users can locate their desired targets more easily and efficiently. The idea of "cloud view" inspired us as it has been widely used in tag clouds, and it has been suggested that it can support other usages including searching, browsing, recognition and matching (Rivadeneira, Gruen, Muller, & Millen, 2007). Additionally, it can provide an impression formation about the underlying patterns in the presented objects, and can guide a user's attention by following certain design principles. Thus, we proposed a cloud-view based prototype where users can interactively browse and search the 3D object datasets, with the aim of decreasing search times and increasing user engagement. We also evaluated its effectiveness through user studies.

Moreover, current trends suggest that users will be browsing and searching 3D object datasets under various scenarios anywhere and at any time. Imagine the scene in which an engineer is examining/constructing an airplane and wishes to check if this is the correct component in the database or explore if there are better choices for the current component (maybe a highly similar one with minor differences). The user might search through simulated datasets which stores the 3D models of all available components. Although the conducting a searching in an office and then returning to the work environment may be sufficient, there is the potential to save time and effort by finishing the search on mobile devices which can be carried within many work environments. These considerations led us to develop our prototype on the mobile devices.

1.2 Research Questions and Objectives

As described above, we proposed a cloud-view based system where users can interactively browse and search 3D object datasets on the mobile devices. We hypothesized that our system will benefit the user by providing a more efficient and engaging approach.

To implement such a system, there are some aspects that need to be considered:

i) How to best organize the collection of 3D objects within the limited screen space?
ii) What features should be chosen as the principle to weight the importance of objects to form the organizing patterns?
iii) How to maximize the information for 3D objects?

While considering these aspects, we proposed that “shape” should be chosen as the principle feature as it has been considered the most critical searching feature and has been adopted in several searching engines (Funkhouser, et al., 2003) (Chen, Tian, Shen, & Ouhyoung, 2003). Based on this feature, the 3D objects are arranged in a radial cloud form based on the shape-similarity. In other words, the central object should be the users’ choice or the most relevant searching results, and the less similar objects are further from the center. Finally, we maximize the valid information of 3D objects by using canonical views which has been described in section 2.4 and 3.3.

With these design considerations, we aimed to provide a visually-engaging layout which can assist the browsing and searching of 3D object datasets by increasing both efficiency and engagement.

1.3 Challenges

The interactive 3D object clouds system is a novel approach that supports the iterative browsing and searching of 3D object datasets. Designing and implementing such a novel and relatively complex system has posed many challenges.

First, considering the above objectives we introduced, the critical challenge is that how to pack a collection of 3D objects within the limited screen space without occlusions and preserve the underlying patterns. These two conditions are mutually exclusive in the extreme: avoiding occlusions tends to push the objects away while we want to preserve similar objects close by. Similar research by Reinert *et. al* (Reinert, Ritschel, & Seidel, 2013) projects the 2D objects’ features onto a 2D matrix, and then use an algorithm based on an improved central Voronoi diagram relaxation to position the objects according to the retrieved feature matrix. Another research inspired us are Wordles (Feinberg, Wordle - Beautiful Word Clouds, 2008) which are a prevalent approach to generating visually-pleasing and attractive tag clouds. In Wordles, the words are generally weighted by their frequency, and the underlying patterns are implied by sizes rather than positions. In other words, a large size indicates a high frequency of the word while it could be randomly positioned anywhere on the screen. Considering the implementation complexity and our design purpose, we decided to use an improved randomized greedy algorithm to position the objects which generate an occlusion-free layout while preserving an approximate underlying pattern. Meanwhile, we varied the sizes and color-tones of objects to emphasize the underlying connections which is based on shape-similarity.

Another challenge is that how to show the most informative viewpoints of 3D objects with the purpose of assisting the comparison and recognition of similarities and differences between them. For a single 3D object, the presenting results can be significantly different by viewing it from different angles. Thus, the concept of a "canonical view" becomes relevant. Based on the throey of Blanz *et al.* (Blanz, Tarr, Bulthoff, & Vetter, 1995), familiar and geometry properties have the most effect on determining the canonical view. In other words, a canonical view should present the 3D object in a way that people are used to observing it, and maximize the significant geometry information. Since it will be inefficient and unpractical to obtain a ground truth for each object's canonical view, there are a few explored algorithms that can generate canonical views of 3D objects, such as including the view area (Dutagaci, Cheung, & Godil, 2010), the ratio of visible area (Polonsky, Patane, Biasotti, & Gotsman, 2005), the surface area entropy (Vazquez, Feixas, Sbert, & Heidrich, 2001), the silhouette length (Polonsky, Patane, Biasotti, & Gotsman, 2005), and the silhouette entropy (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003). Due to the massive and complex calculation of the canonical view selection algorithm, we decided to pre-calculate the canonical views of all the 3D objects by using existing algorithms. Another issue is that it will be disorganized to show all the objects with their own canonical view in the same layout, and it's not very helpful for comparison. We solved this problem and described it in section 3.3 in detail.

1.4 Contributions

We proposed and implemented a prototype system which supports the iterative browsing and searching of 3D datasets on mobile devices. We aimed to provide a novel way to increase the searching efficiency and engagement while offering a visually-engaging layout and we later evaluated this hypothesis by conducting user studies. We showed that our system can significantly decrease search times compared to the classic grid-based layout, and it has been suggested by subjects that they feel cloud-based searching is more interesting which may address "more engagement" design goal. Moreover, it has also enjoyed a higher subjective score on the "visually pleasing" measure.

By proposing this system, we presented a way to arrange the 3D objects in the cloud-based view, and summarized some design principles to benefit further explorations in this area. We addressed many challenges and adjusted our initial designs iteratively to achieve our final goal. We shared our research design and suggested line of future research on the presentation of 3D objects in a way that facilitates browsing and searching.

Our system is a prototype having very specific goals, and the off-line components (e.g. calculating canonical views) are highly flexible for further adjustments or replacements. Thus, it can be embedded into any other 3D search engine to benefit the searching time and engagement.

1.5 Thesis Outline

The remainder of this thesis is organized as follows. A brief description of recent theories and approaches most related to our work are discussed in Chapter 2. Due to the multi-faceted nature of our proposed system, we can only present the most relevant work here. Chapter 3 presents the main proposed model and the implemented details for the interactive 3D object cloud system. Our evaluations of the proposed system are illustrated in Chapter 4 and Chapter 5 from different aspects. Chapter 4 describes the user study that compares our system to the grid-based layout system to evaluate the effectiveness of our design, while Chapter 5 depicts the evaluation of the staggered-animation design in our system. Finally, we summarized the thesis and highlight the contributions, limitations, and suggest potential directions for future work in Chapter 6.

Chapter 2

Related Work

The 3D object clouds system includes elements from a wide range of fields. And as the proposed designs and implementations of our system involves aspects from various research fields, there are many different challenges. The main challenges of presenting the 3D objects in an interactive cloud view are deciding which viewpoint to be presented for the 3D object, proposing a meaningful arrangement of the objects to effectively assist browsing and searching, packing objects without overlaps while maintaining this structure, and highlighting the underlying relationships between the objects. For these critical challenges, we reviewed and summarized the most related work in this chapter.

2.1 Tag Clouds

Tag clouds are a visual presentation of a set of words selected based on some rationale, and the visual properties (such as font sizes, positions, and colors) are applied to the words to encode the underlying information such as each word's frequency and popularity. Some examples of tag clouds with different layouts are presented in Figure 2.1.

With the prevalent usage of tagging culture in social media, tag clouds have been widely used to present a collection of tags for purposes including navigation, abstractions, and recommendations. Generally, there are four distinct usages of tag clouds (Rivadeneira, Gruen, Muller, & Millen, 2007), namely: i) Searching. A specific tag or tag group can be located using tag clouds; ii) Browsing. Using tag clouds to casually explore without specific target; iii) Impression Formation and Impression Presentation. A general impression (e.g. the most frequent words) can be formed by looking at tag clouds; iv) Recognition and Matching. Recognizing the matching source used to generate the tag clouds. Among the tag-cloud approaches, Wordle (Feinberg, Wordle - Beautiful Word Clouds, 2008) stands out which provides aesthetic layouts (Figure 2.1(a)). The tag clouds generated by Wordle has been used in various domains suggested by Fernanda *et al.*

(Viegas, Wattenberg, & Feinberg, 2009), including mainstream media, personal usage (e.g. finding people with similar interests), and education.

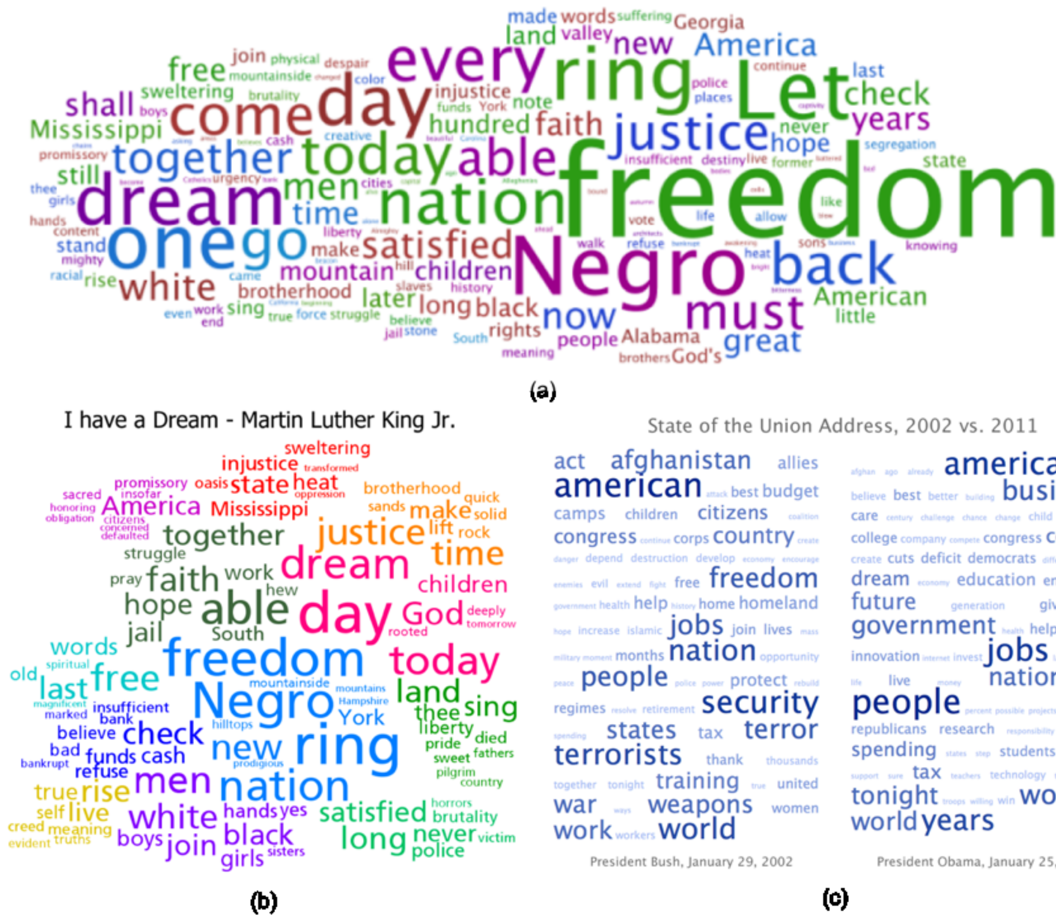


Figure 2.1: Some examples of tag clouds with different layouts:
(a) A spatial layout generated by Wordle (Feinberg, Wordle - Beautiful Word Clouds, 2008); (b) A clustered layout generated by Clustered Word Clouds (Clark, 2008); (c) A sequential layout retrieved from the Tag Cloud definition page on Wikipedia (Wikipedia, 2016).

There is also research exploring the design effects of tag clouds to provide some general design guidelines on both the layouts and visual properties. Martin *et al.* (Halvey & Keane, 2007) has considered the effects of the different properties of tag clouds, and concluded that alphabetization (ordering the words based on ascending/descending alphabetize order) can assist users to search information more efficiently and easily, font size and position of tags are very important for searching efficiency and cognitive load, and users tend to scan the clouds rather than read them. The effects of these properties of tag clouds have also been evaluated by Rivadeneira *et al.* (Rivadeneira, Gruen, Muller, & Millen, 2007) using different approaches, they suggested that the font size has a strong influence on recognition, and a simple ordered list slightly, but significantly, performs better by providing users with a more accurate impression about the tags.

Although these systems have provided some views about tag cloud design, Bateman *et al.* (Bateman, Gutwin, & Nacenta, 2008) provides a deeper insight by exploring how visual properties (e.g. font size, font weight, and intensity etc.) affect people. According to their conclusions, font size, font weight, and intensity are critical visual properties which have consistent and strong influences. Specifically, tags with larger font size, thicker font weight, and higher intensity (darker, higher contrast and saturation) will more easily capture a viewer's attention. The position though seems to show less influence, and the center and areas near it are safe zones for attracting a subject attention. By conducting a task-related study with eye tracking, Lohmann *et al.* (Lohmann, Ziegler, & Tetzlaff, 2009) studied the influences of different layouts and properties on various usages including specific searching (find a specific tag), general searching (find the most popular tags), and cluster searching (find tags belong to a certain topic). These tasks are performed with different tag clouds layouts (sequential layout, circular layout, and clustered layout) along with a reference layout as shown in Figure 2.2. They found that sequential layouts with alphabetical sorting is the best for specific searching, the circular layout is favored for general searching and the clustered layout is the optimal choice for clustered searching. In other words, there is no uniform optimal layout can be found to satisfy all usages well. Some of the previous research findings are also supported by their study: Tags with larger font size will be memorized and found more easily and quickly; Users scan rather than read tag clouds; The central area is the zone to attracts more attention, and its influence is especially obvious in the circular layout.

Figure 2.2: Tag-cloud layouts have been evaluated in the study of Lohmann et al. (Lohmann, Ziegler, & Tetzlaff, 2009):
(a) Sequential (alphabetical sorting); (b) Circular (decreasing popularity); (c) Clustered (thematic clusters); (d) Reference (sequential, alphabetical sorting, no weighting of tags). Quadrant separations and ordering principles are indicated by the colored lines, circles, and arrows which is invisible to participants, and the weights of words are determined by the popularity.

2.2 Image Packing

Beyond packing tags, arranging a collection of images in a visually-engaging and informative form has also been explored. One interesting form to pack a large amount of images is mosaics, which means a source image is converted into tightly filled tiles. Hausner (Hausner, 2001) proposed an algorithm to pack small square tiles to form synthetic mosaics using the centroid Voronoi diagram to position the tiles on the plane evenly, and Kim *et al.* (Kim & Pellacini, 2002) use arbitrary images rather than the square tiles to fill shapes by packing image tiles compactly while matching the color tones as shown in Figure 2.3. A centroid Voronoi diagram which is a relaxation of a Voronoi diagram, also known as Lloyd's algorithm (Lloyd, 1982). Hiller *et al.* (Hiller, Hellwig, & Deussen, 2003) bring this algorithm into the image packing field by positioning arbitrary 2D objects on the plane to form a stippling painting instead of being limited in positioning points. Finally, Reinert *et. al* (Reinert, Ritschel, & Seidel, 2013) expand Hiller's approach by explicitly including the boundary distances into the objective function to produce a more balanced layout, and this algorithm can also provide an inverse layout by parsing the user's inputs to extract the key features to arrange the elements in the layout as shown in Figure 2.4.

Figure 2.3: An example of Jigsaw Image Mosaics algorithm (Kim & Pellacini, 2002)

Figure 2.4: An example of how the layouts are reorganized based on the user's inputs in Reinert's approach (Reinert, Ritschel, & Seidel, 2013).
Starting with a default layout (Left), the layout is changed to be organized by size vertically after user's pins and movements (Middle), and then it rearranged additionally by brightness horizontally after a user's direction (Right).

These approaches explored a way to pack a collection of images in an evenly and visually-engaging way, and it aligns our purpose – to present a collection of 3D objects in a plane while taking most advantage of the available space to provide an attractive and informative layout. On the other hand, these works are more focused on producing a static layout, while our work is trying to implement an interactive 3D browsing and searching system which benefits from the packing layouts. Little has been explored regarding how a well-packed layout can assist users to perform iterative search in 3D datasets, and our work bridges this gap.

2.3 3D Model Applications

The usage of 3D models has become wide and prevalent in both commercial and academic fields. According to Funkhouser *et. al* (Funkhouser, et al., 2003), there are three recent trends accelerating the popularity of 3D models: i) the new practical and effective tools to construct detailed 3D models; ii) inexpensive graphic hardware provides support for working with 3D models; iii) the web provides accessibility of existing 3D objects. Based on the trend of increasing demand for 3D models and the reasons causing it, they proposed a shape-based search engine for 3D models which can search 3D models according to 3D sketching, 2D sketching or by text (as shown in Figure 2.5). Chen *et. al* (Chen, Tian, Shen, & Ouhyoung, 2003) also proposed a similar system which is more focused on search interactively and iteratively based on the drawing of 2D shapes, and they also proposed a different 3D descriptor to retrieve and store the shape features of 3D objects. As with these two systems and their results, we also assume that shape is a critical feature for browsing and searching 3D objects since sketching shapes as query source has been effective.

Further evidence of this is found in the 3D search engine proposed by Funkhouser *et. al* which is applied to the work "Modeling by Example" (Funkhouser, et al., 2004). This work presents an approach to constructing a new 3D object which consists of different parts being cut from other models by the users after they query and gather the desired 3D objects, and an example schema is shown in Figure 2.6.

There are also other prevalent and interesting applications of 3D models, such as 3D printing, virtual reality, and augmented reality. For example, Yu et. al (Yu, et al., 2011) introduced a system to place 3D furniture in a room automatically and optimally, while Merrell *et. al* (Merrell & Manocha, 2011) proposed an algorithm to provide the automatic generation of complex 3D models based on the user-defined inputs to facilitate the building of large 3D scenes (as shown in Figure 2.7).

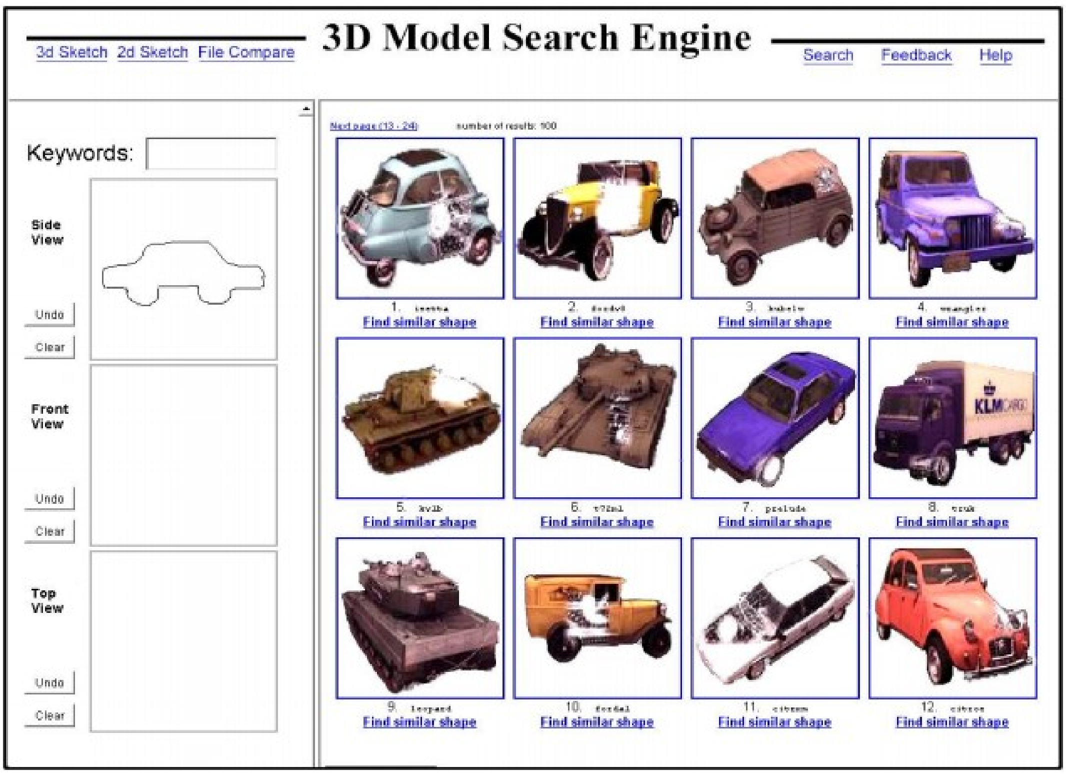


Figure 2.5: The query interface of the 3D search engine proposed by Funkhouser et. al (Funkhouser, et al., 2003).
On the left side, the user can specify a query via any combination of keywords, 2D sketches, or 3D sketches. On the right side, a ranked set of thumbnails of the 16 best matching 3D models are presented based on the provided searching query. Each model can be retrieved by clicking the thumbnail or the link below it.

The 3D object applications are applied in various commercial and academic fields by wide range of users, and the increasingly important question of “how do I find this 3D object” is raised. Our proposed system is designed to provide a visually-engaging and efficient layout for users to crawl through the dataset to browse and search for the desired 3D objects. We also noted that it can be integrated with any other 3D search engine as an assisting function to benefit more 3D applications.

Figure 2.6: An example schema of Modeling by Example (Funkhouser, et al., 2004): The center large brown chair is generated from the circled parts of the other objects.

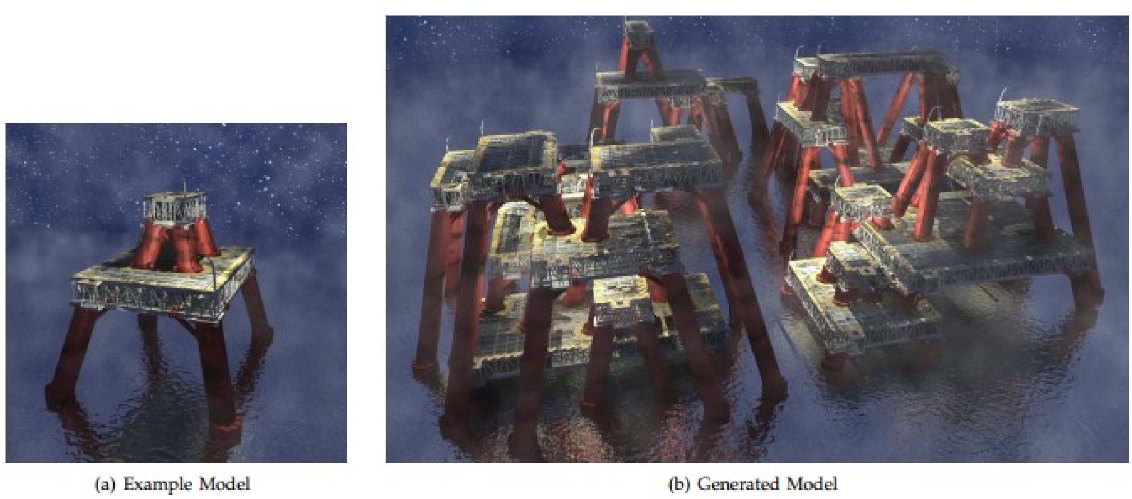


Figure 2.7: Model Synthesis (Merrell & Manocha, 2011):
An example of constructing a large complex model from the user-defined example. (a) An example model given by the user; (b) A model of several oil platforms generated automatically in about half a minute.

2.4 3D Model Canonical View

Presenting 3D objects is a critical issue in our proposed system. A central issue is which viewpoint should be initially presented since there are unlimited options for a 3D model. Thus, the concept “canonical view” which is also known as the best view (in some sense of the term) is involved. Although there is no consensus about how to define the best canonical view, Blanz *et al.* (Blanz, Tarr, Bulthoff, & Vetter, 1995) did provide some good criteria to provide a better understanding. Based on their experimental results, there are three factors contributing to human choice of canonical views: experience, task, and geometry. Specifically, the familiarity of the view and the geometric properties have the most effect on determining the canonical view. In other words, a canonical view should present the 3D object in a way that the people are used to observing it, and also maximize the significant geometry information that is visible.

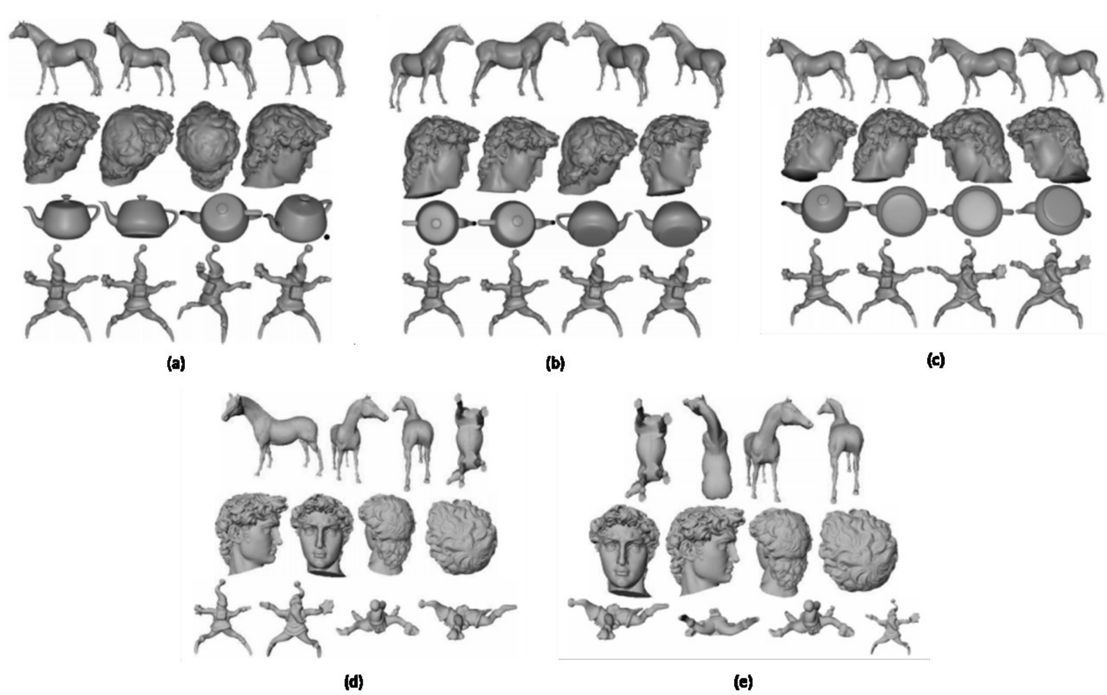


Figure 2.8: Four top-ranking (left to right) views among candidate views generated by algorithms (a) to (f):
surface area entropy, ratio of visible area, curvature entropy, silhouette length, and silhouette entropy. The images are collected from (Polonsky, Patane, Biasotti, & Gotsman, 2005).

Since it is inefficient and unpractical to obtain ground truth for every object's canonical view, there are a few existing algorithms that can generate reasonable canonical views of 3D objects. These algorithms, according to Polonsky et al. (Polonsky, Patane, Biasotti, & Gotsman, 2005), are based on different principles for generating canonical views consistent with human observation habits. Among them, some algorithms are based on maximizing the geometric complexity (e.g. shape, surface area) (Vazquez, Feixas, Sbert, & Heidrich, 2001) (Polonsky, Patane, Biasotti, & Gotsman, 2005) (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003), some of them are based on view-dependent features (e.g. silhouette) (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003), and some based on other proposed principles. The various *a priori* principles lead to different generated results, and some work focus on evaluating these algorithms by trying to find which one is the best and provide quantitative benchmarks. Polonsky *et al.* (Polonsky, Patane, Biasotti, & Gotsman, 2005) reviewed some canonical view algorithms and summarized some general principles to address the "automatic generation of canonical view" challenge, and some generated results via different algorithms in this paper are predicted in Figure 2.8. Instead of a semantic comparison, Dutagaci *et al.* (Dutagaci, Cheung, & Godil, 2010) proposed a benchmark to evaluate canonical view selection algorithms quantitatively, and measured seven algorithms with ground truth. The seven methods that have been measured including view area (Dutagaci, Cheung, & Godil, 2010), ratio of visible area (Polonsky, Patane, Biasotti, & Gotsman, 2005), surface area entropy (Vazquez, Feixas, Sbert, & Heidrich, 2001), silhouette length (Polonsky, Patane, Biasotti, & Gotsman, 2005), silhouette entropy (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003), curvature entropy (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003), and mesh saliency (Lee, Varshney, & Jacobs, 2005). Their experiment result reveals that "None of the methods is consistently the best (or worst) over all the objects" while they all mostly did a reasonable job. Based to their results, we applied some best view selection algorithms in our system to generate the initial views of 3D objects in each layout (the details are introduced in section 3.3).

2.5 Staggered Animation

Our proposed system is an interactive application, and it involves relatively complicated transitions (including transforming, rotating, scaling of multiple objects) when users select a new central object. Animations are a promising approach to retaining coherence of changes between different layout states.

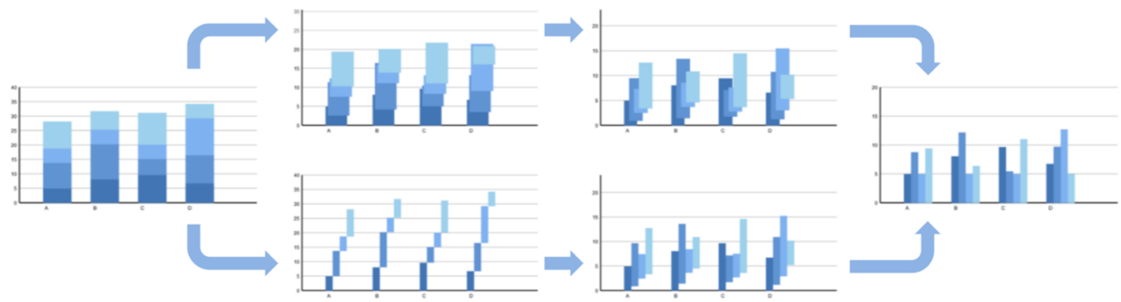


Figure 2.9: Animated transition from stacked bars converted to grouped bars (Heer & Robertson, 2007).
The top path illustrates the non-staggered animation procedure, while the bottom one shows the corresponding staggered procedure: the first stage changes the width and x-positions of bars, and the second stage drops the bars to the baseline.

Staggered animation, also known as staged animation, is the kind of animation that have multiple stages where the whole transitions are separated into each stage rather than happen simultaneously. It has been adopted into the designs of some visual applications, and suggested to help keep the coherence when a transition happens by providing less overwhelming visual transitions. Previous research has explored the effectiveness of staggered animation and its design principles. Heer and Robertson (Heer & Robertson, 2007) have investigated the effectiveness of different staggered and non-staggered animations between statistical data graphics (as shown in Figure 2.9), and they also suggested some principles for designing effective transition animations. Their results showed that the staggered animation is highly preferred by the participants for an improved understanding of transitions and it also increased engagement. Additionally, in their view, a well-designed staggered animation should have simple stages as intricately staged animations will lead to increasing errors. They also suggest minimizing the occlusion during the transition, and the transition time should be around 1 second. Chevalier *et al.* (Chevalier, Dragicevic, & Franconeri, 2014) have attempted to provide a more specific understanding about what factors have influence on the design of staggered animation and how they effect it. Their experiment is mainly conducted with the collection of dots within 2D plane, asking subjects to track multiple dots during transitions, and the staggered animation in this case is presented by moving dots to new positions in separated groups with multiple animation stages. To better define the potential influence factors and examine them, they proposed task complexity metrics for the dots' transitions, involving target crowding, inner crowding, and deformation (as illustrated in Figure 2.10). Their result shows that staggered animation can reduce the transition complexity under some conditions, and target crowding is the main factor that benefits the staggered animation.

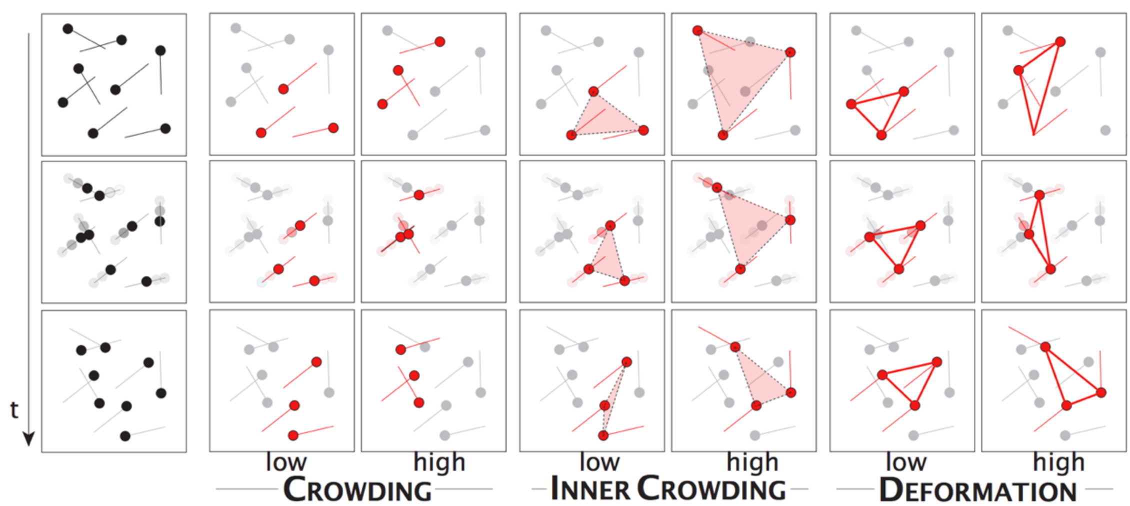
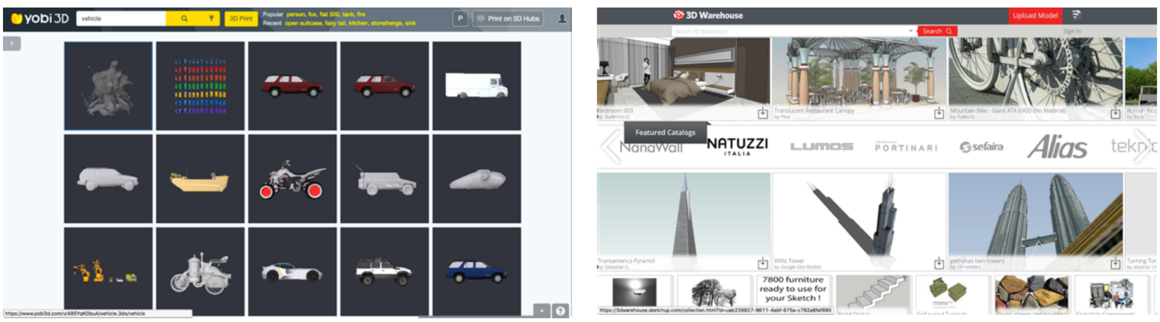


Figure 2.10: Illustration of the complexity metrics proposed by Chevalier et al. (Chevalier, Dragicevic, & Franconeri, 2014)

We applied staggered animation in our system during the transition expecting it provide an instructive introduction to the system and how it works. Although, on the other hand, these prior studies provide good insights about the staggered animation's effectiveness and design, little is known about how it will affect complex 3D transitions as in our system. Thus, after we adopted the staggered animation to our transition procedure, we conducted a user study to evaluate it as described in section 5.

2.6 Commercial Systems

Along with the increasing demand of searching and finding desired 3D objects and the wide usage of them, there currently are many browsing and searching services for 3D models, including online web browsing, 3D model galleries in game engine (Unity, etc.) and within 3D modelers.



(a) (b)

Figure 2.11: The typical grid-based view of online 3D model dataset searching websites: (a) The view of presenting a collection of searching keyword "vehicle" on the Yobi website (Yobi3D, 2016); (b) The home page of 3D Warehouse website (Warehouse, 2014).

We reviewed 16 popular 3D model dataset websites recently (2016) to provide a better understanding of the typical layout design in commercial systems. We found that these 16 websites use the grid-based layout (with minor differences) to present the collection of 3D objects (as shown in Figure 2.11), and they usually only vary in their approach to browsing objects in close detail (zooming in). We categorized their close browsing approaches into four types: i) Click an object in the list to enter a larger view of the object which can be free-rotated and scaled; ii) Click an object to enter a large view with several images to switch between; iii) Click an object to view a big image; iv) Or simply no way to browse objects in detail and only can be viewed in the list. The details of which websites correspond to which approaches are shown in the Table 2.1. The grid-based view, as the most widely used form of layouts, has its advantages of being simple and clear, and we evaluated the effectiveness of our approach by comparing with the grid-based view as specifically described in section 4.

Table 2.1: The summarization of the interactive approaches to observe 3D objects in detail on some popular online 3D object websites.

Browsing Approach	Website
Free Rotation	Yobi 3D
Switch several images	Flying Architecture, Kinnarps, TurboSquid
View a big image	Google 3D Warehouse, Resources.blogscopia, Kolo, Model3D.biz, Dark GDK, Archive 3D, 3D Model Free
Null	ArchiBit Generation, Free the models, Creative-3D.net, Klicker

Chapter 3

Viewing 3D Objects in Interactive Clouds on Mobile Device

In this chapter, the design and implementation of the interactive 3D object cloud on mobile devices are described. First, we provide an overview of the framework and interface design in Section 3.1. Then, we introduce some main features proposed in Section 3.2, 3.3. 3.4, 3.5 and 3.6, separately. Finally, we introduce few additional features designed to further support the browsing and searching of 3D objects in Section 3.7.

3.1 Overview

The interactive 3D object cloud is a system for the browsing, selection and comparison of a collection of 3D objects on mobile devices. The main contributions of the proposed system are mainly focused on the design for browsing and searching within 3D object datasets in an interactive view.

From the perspective of system design, we focused on providing an interactive technique to let users crawl through the space of objects with simple gestures. To maximize the salient information, we adapted the canonical view for every center object. We also varied the colors and sizes of objects when arranging them to assist users to understand the underlying similarity patterns.

From the perspective of development and implementation, our system was developed in the Unity Game Engine. For generating layouts in real-time on the mobile devices, we pre-calculated the canonical views and the similarities between models considering the limited abilities of mobile devices. The algorithms that were involved in the pre-calculation are developed in Matlab, and are flexible for future improvements or related applications.

We began with an overview of the system interface to situate the reader in our discussion. The interface consists of several featured parts: the main layout (A), the center object (B), the menu button (C), the interaction instructions (D). Figure 3.1: shows the layout of the overall system interface, with corresponding labels identifying different parts, and we will now provide a brief introduction about the function of each part and how to use them.

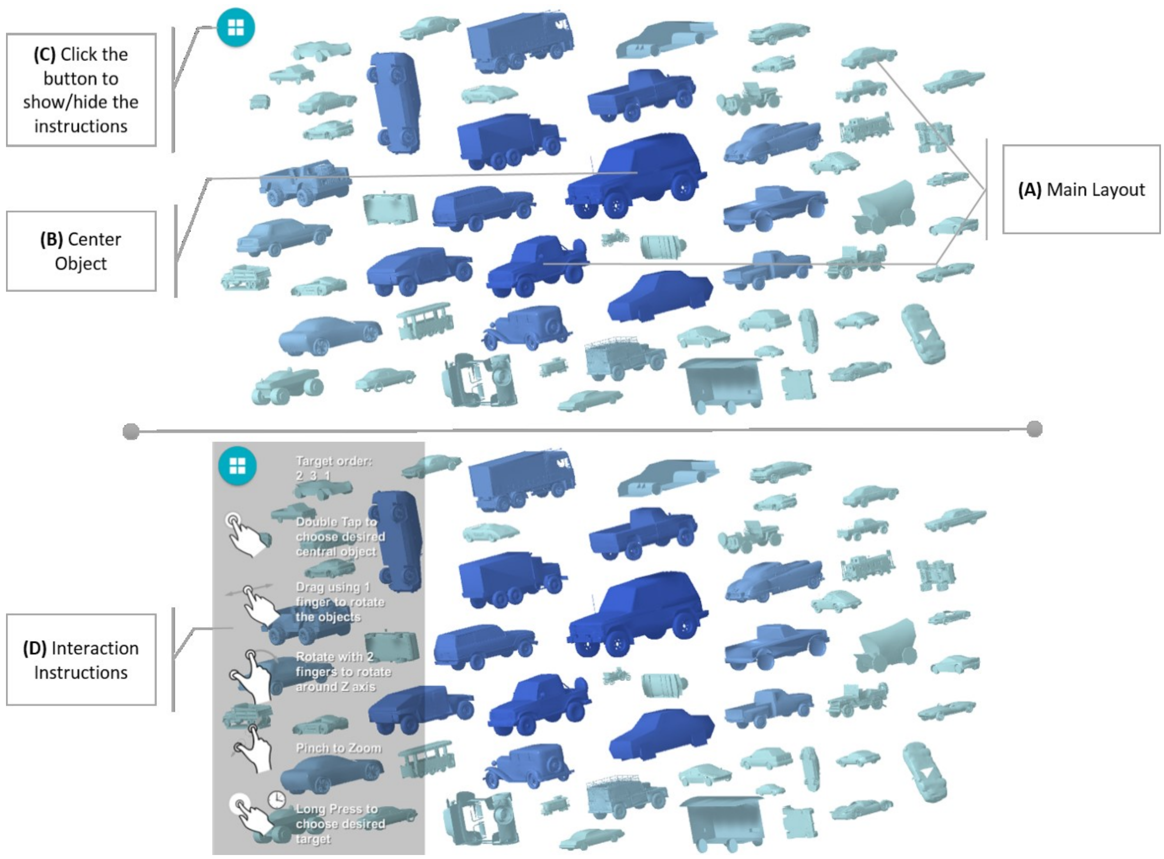


Figure 3.1: System interface design

(A) Main Layout:

In the main layout, similar objects to the center object are packed in circular cloud shape. They are shown in the same viewpoint with the center object to parallel them and facilitate comparison. Based on the similarity level, all objects' sizes decrease and colors fade. That means the more similar the object is with the center object, the larger the size of the object will be and the color will be darker. Otherwise, its size will be smaller and the color will be more faded. The user can observe the main layout and interact with it using the gestures shown in (D) to find their desired objects.

(B) Center Object:

Initially, the center object is chosen randomly, but subsequently is always chosen by the user. Once an object is chosen as the new object of interest, it will move to the center of the screen and change its size and color to be the biggest and darkest one. Then, the most-related objects within the dataset will be packed around it using the packing algorithm in section 3.4 and using the principles in section 3.5 to vary their sizes and colors.

(C) Menu Button:

The menu button is always shown on the upper left corner of the screen. The user can tap the button to show/hide the interaction instructions, so that they can always maintain an awareness of how to interact with the system without blocking the main layout.

(D) Interaction Instructions:

The interaction instructions are optionally shown in a semi-transparent form on the left side of the screen once user taps the menu button an odd number of times, and will be hidden with an even number of taps. This contains the instructions for all the interaction gestures, namely:

- *Double Tap:* The user can double tap on any object within the screen to make it the center object. If the tapped object is already the central one, all the objects will be reset into the initial viewpoints of this layout. Because there will be overlaps between some objects after users rotate them, a reset of the viewpoints will help them get remove overlaps quickly.
- *One Finger Drag:* The user can use one finger to perform the drag motion, all the objects within the screen will be rotated with this dragging movement.
- *Two Finger Drag:* The user can use two finger to perform the drag motion – hold one finger, and move another finger around. All the objects within the screen will be rotated in the depth direction.

- *Pinch:* The user can use two fingers to zoom-in/zoom-out.
- *Long Press:* The user can long press on an object to choose it as the desired one.

3.2 3D Object Similarity

Just as the technologies and applications of 3D models have been widely developed and implemented (as introduced in section 2.3 and 2.6), research on content-based 3D model retrieval have been explored to provide various approaches as well. We decided to use the visual similarity approach proposed by D.Chen *et. al* (Chen, Tian, Shen, & Ouhyoung, 2003) determine the similarity between the 3D objects in the preprocessing stage. This method has been evaluated with better performance (precision-recall evaluation) than other previous work in this field, and it is robust against transformations such as translation, rotation, scaling, noise, decimation and model degeneracy.

The main idea of visual similarity based 3D model retrieval is summarized with the statement – "If two 3D models are similar, they also look similar from all viewing angles" (Chen, Tian, Shen, & Ouhyoung, 2003). According to this principle, this approach proposed a 3D model representation called a Lightfield Descriptor to summarize the similarities of many different viewpoints of 3D objects to obtain the overall similarity. The procedure for retrieving the similarity between two objects is comprised of 3 steps: i) Translate and scale the objects to make them entirely contained in the retrieval field; ii) Generating the Lightfield Descriptors for both objects; iii) Calculate the differences between their Lightfield Descriptors to generate the dissimilarity value.

To generate one Lightfield descriptor for an object, a set of contour images are produced by setting the light field camera on the intersection point of a sphere, and then rotating the sphere around the object to rotate the camera, as shown in Figure 3.2.

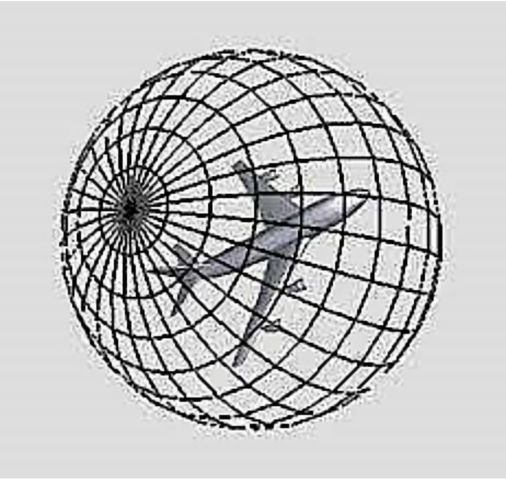


Figure 3.2: The scheme of the measuring sphere, and the camera is on the intersection point (Chen, Tian, Shen, & Ouhyoung, 2003).

To enhance the efficiency and robustness, two simplifications have been adapted to the above steps by D.Chen *et. al*: i) Move the camera along 20 uniformly distributed vertices of a dodecahedron to yield 20 rendered images; ii) Apply orthogonal projection to contour image procedure. The 20 rendered images have been suggested as being sufficient to represent the shape of a 3D model, and the number of rendered images is reduced to 10 because the silhouette projections of two opposite vertices on the dodecahedron are identical because of applying the orthogonal projection. Thus, the Lightfield descriptor is obtained by getting and storing features of 10 images rendered from vertices of dodecahedron over a hemisphere. For extracting the features of every single rendered image, an integrated method proposed by Zhang and Lu (Dengsheng & Guojun Lu, 2002) is used, and it mainly retrieves the diagram which represents the distance of points on the contour's edge to the shape's center. Finally, a set of Lightfield descriptors are generated and stored for an object to make it robust to the rotation. Figure 3.3 demonstrates the final set of Lightfield descriptors of two objects and how to use them to summarize the overall similarity.

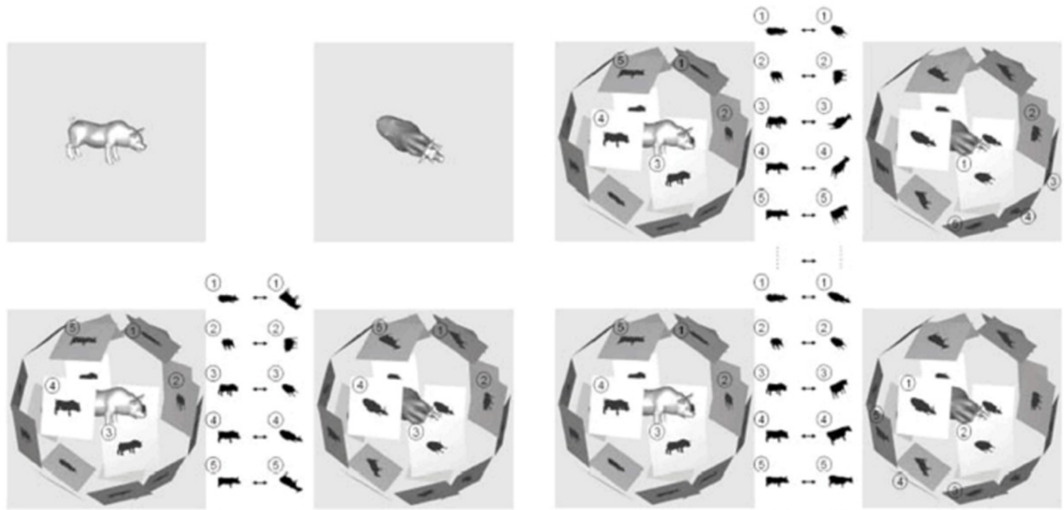


Figure 3.3: A set of Lightfield Descriptors for two 3D objects (Chen, Tian, Shen, & Ouhyoung, 2003)

To compare the Lightfield Descriptors between two 3D objects, we simply find the match of descriptors that minimizes the dissimilarity. The dissimilarity between every two descriptors can be obtained by summarizing the differences of corresponding images. To decide how to match the images between two descriptors, the match with the smallest difference has been chosen among the permutations of all matches. The difference between two rendered images is generated by calculating the distance between two images' shape descriptors. More details of this visual similarity retrieving approach can be found in the original paper (Chen, Tian, Shen, & Ouhyoung, 2003).

Since retrieving the similarity among a large number of objects by comparing the Lighfield descriptors can take about 2 seconds on a server, we assume it will be time consuming and impractical on the mobile devices. Thus, we used the executable code from the above paper to pre-calculate the dissimilarity between every two 3D objects, and then generate a similar-objects list for every object according to the descending order of the dissimilarity between this object and all the other objects. This pre-calculated result is stored in a table and applied as a resource file in the interactive 3D cloud system.

3.3 Canonical Views

As described in section 2.4, we need to calculate the canonical views for the 3D objects to maximize the valid information in the initial state of each layout. Generating canonical views is an off-line computation in our system because of the massive amount and high complexity of the calculations.

There are a few canonical view selection methods, but there is not a method can consistently provide best result as stated in section 2.4. In our system, we applied the view area algorithm (Dutagaci, Cheung, & Godil, 2010) based on the result we tested on our application that tends to provide the most acceptable results, and we also used other best view selection algorithms in some cases (e.g. In the “switching canonical view” function described in section 3.7.2). All the implemented source code of the best view selection algorithms is from Dutagci *et al.* (Dutagaci, Cheung, & Godil, 2010). Figure 3.4 illustrates the originally imported viewpoints and generated canonical views of some 3D objects in our system. The pre-calculated viewpoints are saved as 3D vector in an excel file, and then are loaded when the application runs.

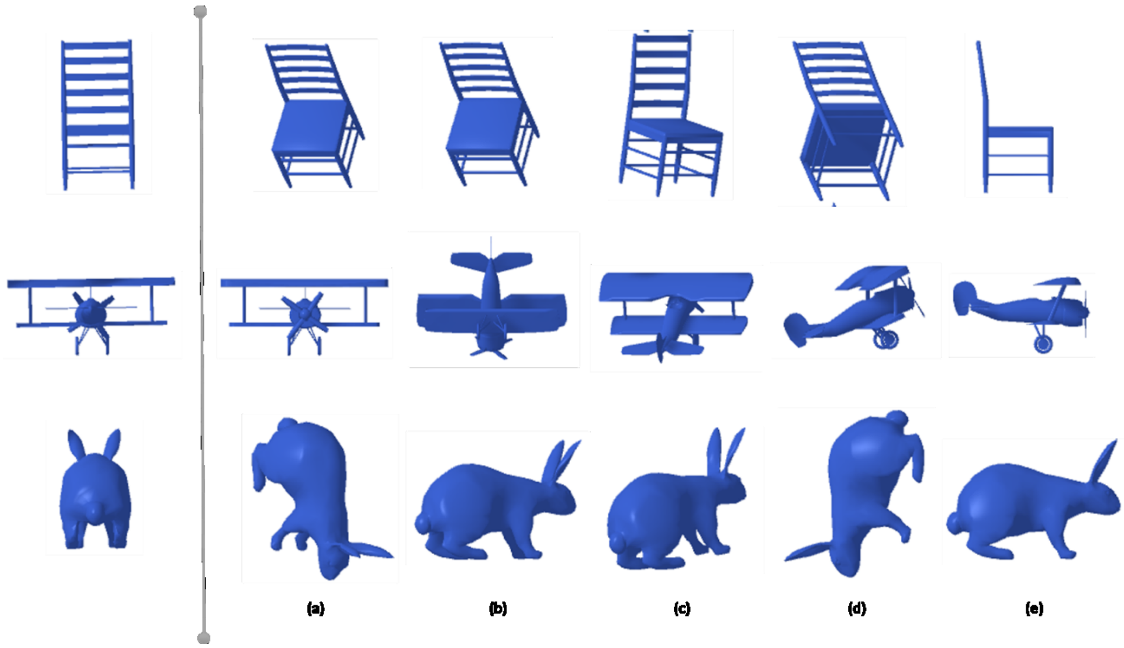
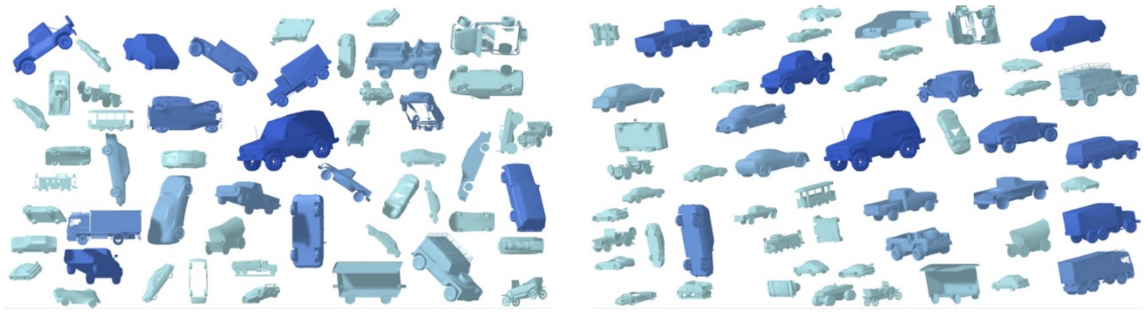


Figure 3.4: A set of examples shows the originally imported viewpoints and the generated canonical views of some objects.
On the left side are the original viewpoints, while on the right side are the canonical views generated by the best view selection algorithms (a)-(e): (a) surface area entropy (Vazquez, Feixas, Sbert, & Heidrich, 2001);(b) view area (Dutagaci, Cheung, & Godil, 2010);(c) silhouette length (Polonsky, Patane, Biasotti, & Gotsman, 2005); (d) silhouette entropy (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003); (e) mesh saliency (Lee, Varshney, & Jacobs, 2005).

We initially wished to present separate canonical views for all objects individually within a layout, but we found it yielded a confusing layout which is neither visually engaging, nor useful. When every object is shown in its own unique viewpoint without considering the relationship with other objects it is difficult to make direct comparisons. Considering we put the current chosen object (user selected object) in the center, we decided to adjust our original design by enforcing the other objects to be parallel to the central object's own viewpoint, and this does produce a much more organized layout as shown in Figure 3.5(b). Because the remaining objects should be similar with the central one, they tend to be presented in good viewpoints as well.



(a) (b)

Figure 3.5: The packing layouts without parallel (a) and with parallel (b)

3.4 Packing Layout Design

In this section, we discuss the design and development of the approach for packing the objects into the main layout. We designed our algorithm to generate a layout that makes full use of the space, preserves the “sizes and colors” principles in section 3.5 and is visually-engaging which is supported by the evaluating result in section 4.4.4.

Generating the non-overlapping packing layouts is the essential foundation of the design. The design goal of this layout is providing a balanced, non-overlapping arrangement of the included 3D objects. A further development goal is to make this layout can be generated in real-time on the mobile devices.

To achieve the design goal, we proposed the following method to pack 3D objects which is inspired by the randomized greedy algorithm used in Wordle (Feinberg, Wordle - Beautiful Word Clouds, 2008). To simplify the calculation, we firstly divided the available packing space into a grid matrix $G[N, N]$, so the space will have N elements in each row and N grids in each column. Second, we used the rectangular bounding boxes of objects to check for collisions with any other objects and the boundaries of the packing space. Then, we built our method based on these simplified definitions.

First, we randomly choose an object to be the central one and load its similar-objects list which is generated by the approach in the section 3.2. Then, we place the central object in the center of the screen. Next, all the other objects will be packed one by one based on the list order which is also the similarity-level order. For every object, it will be rotated to the parallel viewpoint with the central object, and then it will be put into a semi-random initialized position. If the object can fit in current position without any overlap with all the previous objects and stay within the screen, the object will stay in the current position and never move in this layout until a new central object is selected. Otherwise, it will move along a spiral path until it finds a position that it can fit in, then we will move on to the next object. Finally, the packing layout will be finished when all the related objects are packed or there is not enough space and we cannot find a position for the current object. Figure 3.6 demonstrates how objects move along spirals and find their final positions.

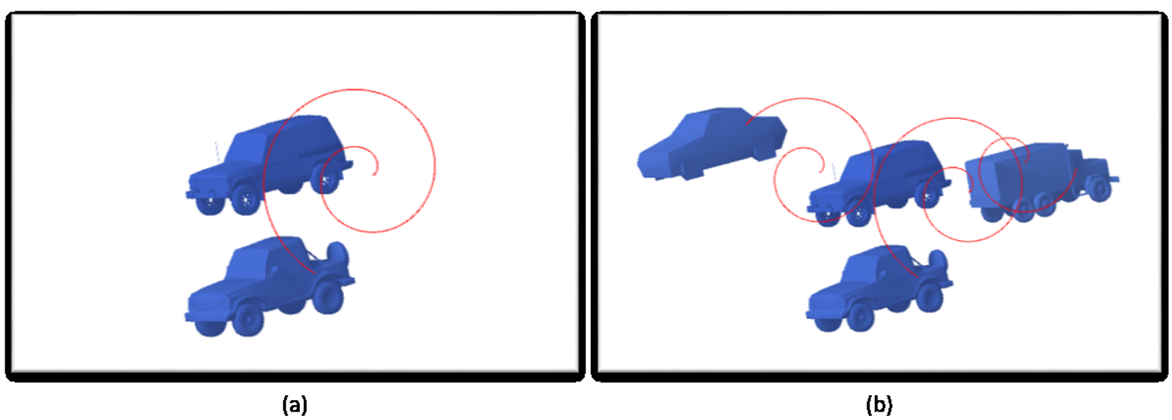


Figure 3.6: An example of moving objects along spirals:
(a) One object moves along a spiral path to find its final position around the center object in this layout; (b) Three objects move along spiral paths respectively.

The spiral path applied in our system is the Archimedean spiral which is described by the following equation in polar coordinates (r, θ):

$$r = a + b\theta$$

In the above formula, a and b are real numbers where parameter a controls the direction of the spiral while b controls the distance between successive turnings (Wikipedia, 2016). In the implementation of our system, we set value of a equals to 0 and b equals to 0.01. Every starting a new spiral, θ will be initially given the value 0, and we limited its maxim value to be 6π. In other words, if the value of θ reaches its upper bound and the object still hasn't found a fit-in position, it will start moving along a new spiral which has a different beginning position.

To find a semi-random position for the beginning point of every spiral, we use the breadth-first search algorithm which starts from the center of the grid matrix G and tries to find the first free. Once this free grid is found we will make it the beginning position of the spiral. As we tried to make the beginning position to be as close as possible to the center to have a loose neighborhood structure, there is the possibility that an object will not fit in this initial position and it will start moving along a spiral. We therefore say this is a semi-random position because it's not totally randomized but it does not guarantee that the object will end up with this beginning position. Also, to optimize the packing speed, we applied another randomized factor. Depending on whether the current initial position found by the BFS failed to find a fit position for the object, there is a high possibility that the grids near it will also fail to find the position. Thus, it will waste a lot of time to check all these potential failed initial positions one by one by moving the object along a spiral until it reaches the threashold. To speed up this process, we randomly will skip 2 to 5 initial positions which is near to the failed position.

Since we defined the packing space as a grid matrix $G[N, N]$, the values of parameter N will have influence on the final layout. In practice, we manually set the value of N. Figure 3.7 illustrates the layouts with different values of N. In the final version of our system, we set N to 50.

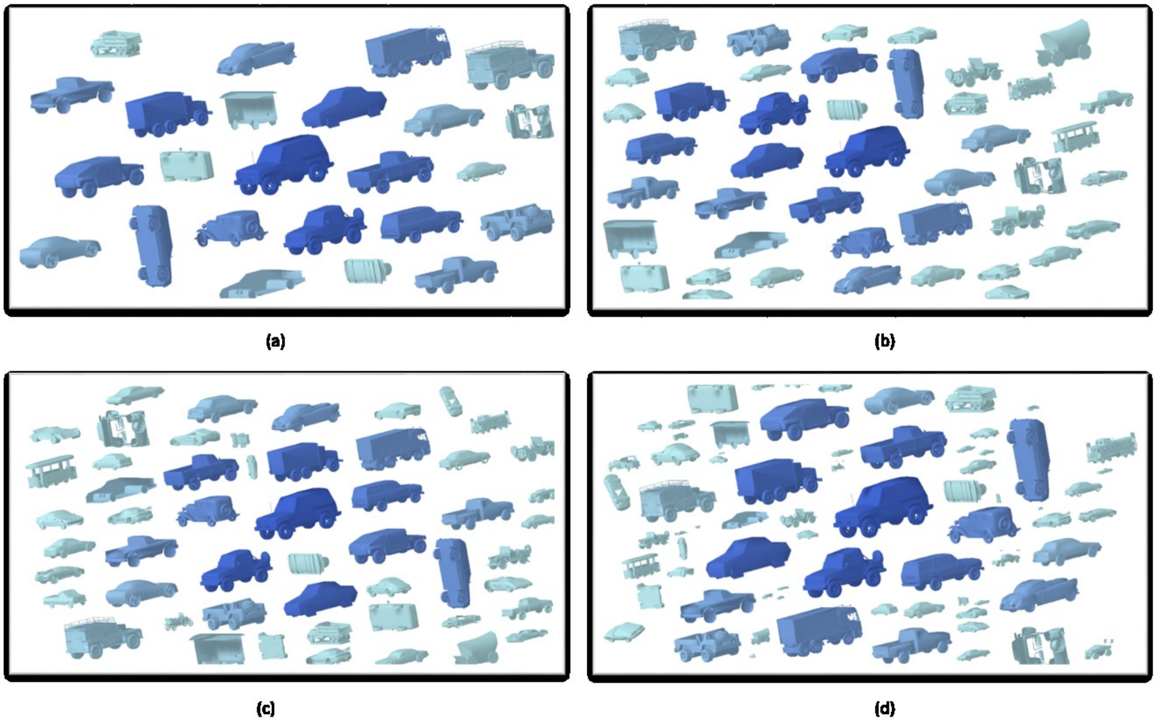


Figure 3.7: The layouts with different values of grid numbers (N).
(a) N = 10; (b) N = 30; (c) N = 50; (d) N = 70.

Below we provide the pseudocode for the above packing algorithm as Algorithm 1.

Algorithm 1 PackObjects (G, O). Pack all objects within the screen preserving the "size and color" principle and without overlaps.

Input: G is the grid matrix where the status of all grids are initially free. O is the set of objects that need to be packed.

Output: P, which is a set of packed objects which contains their updated information such as positions, sizes and colors, etc.

P = *Empty*;

for each object *obj* in *O*: **do**

    newObj = PackSingleObject(*G*, *obj, P*);

---

if *newObj* is not **NULL**: **then**

Render *newObj* with the generated color based on its index in *O*

Add *newObj* to *P*;

UpdateGrids(*G*, *newObj*): *Update the grid matrix since the status of grids that is taken by*

newObj should be updated to be "occupied".

else

Finish the packing process and **return** the current *P*;

end if

isEnoughSpace = checkIfThereIsEnoughSpace(*G*): *Check if there is still enough space for*

continuing the packing process according to the predefined limitation. E.g. If there is less than

10% space left, return false.

if *isEnoughSpace* is **true**: **then**

continue;

else

Finish the packing process and **return** the current *P*;

end if

end for

return *P*;

For further clarity, we also provide the pseudocode for the packing process of a single object as Algorithm 1.1.

Algorithm 1.1 PackSingleObject (*G*, *object, P*). Pack one single object, find its final position within the given grid space.

Input: *G* is the grid matrix with the current grids' status. *object* is the object that needs to be packed. *P* is the set of objects which have been successfully packed.

Output: *newObj*, which is the current packing object with updated position and size information. Return NULL if cannot find a fit-in position in the given space.

---

diffSizeCounter = 0;

while (*diffSizeCounter* < *diffSizeCounterUpperBound*): **do**

    tempG = *G*;

    thisSizeCounter = 0;

    while (*thisSizeCounter* < *thisSizeCounterUpperBound*): **do**

        startGrid = findClosestFreeGrid(*tempG*): *Use the BFS algorithm to search the grid matrix*

        to find a free grid which is closest to the center. If cannot find any free grid, return ***NULL****.*

        if *startGrid* is **NULL**: **then**

            tryCounter ++;

            countinue;

        end if

---

startPosition = convertGridToWorldPosition(*startGrid*): *Get the world position of the*

center of the start grid.

tempG = MoveAlongSpiral(*object*, *tempG*, *startPosition*): *Set all tried grids' status to be*

"occupied" and return the updated temp grid map. Object's position will be out of the

screen if it didn't find a fit-in position.

if object is within the screen: **then**

return object;

end if

thisSizeCounter++;

end while

object.size *= *sizeDecreaseRatio*;

diffSizeCounter++;

end while

return **NULL**;

3.5 Vary Sizes and Colors

To stress the underling similarity pattern of the layout, first of all, we decided to vary the sizes of packed objects according to their similarity levels with the center object. In other words, the more similar the object is with the center object, the bigger the object's size.

To accomplish this design goal in our implementation, we firstly varied the initial sizes of objects before actually starting the packing process. To achieve this, in the per-packing part, we estimate an initial size of all objects, and then we apply different size decrease ratios to different similarity groups of objects. For example, the size decrease ratio of the center object is 1, and the ratio of the top 10 percent similar objects is 0.85, etc. We provide the pseudocode for the initializing sizes process as Algorithm 2. Second, in the packing process, we will preserve the size-reduction principle by decreasing the size of object if it cannot find a fit-in position with the current size within the upper-bound of trying times. For more details of the implementation of this part, it can be found in the section 3.4 Algorithm 1.1.

Algorithm 2 EstimateInitialSizes (*O*, *screenWidth*, *screenHeight, decreaseRatio*): Estimate the initial sizes of all the related objects according to the varying size principle.

Input: *O* is the set of objects which is similar to the chosen center object. *screenWidth* and *screenHeight* are the width and height of the current device's screen, respectively. *decreaseRatio* is the size decrease ratio for different similarity levels of objects.

Output: *S*, the new set consisting of the objects in *O* with the updated sizes.

S = *Empty*;

objectNumber = *O.size*;

devideNumber = Square Root of *objectNumber*;

if (*devideNumber* > *devideNumberUpperBound*): **then**

  devideNumber = *devideNumberUpperBound*; *This part of check can prevent the initial sizes of*

  objects to be too small if there are a large amount of related objects.

end if

---

horizentalUnit = *screenWidth* / *devideNumber*;

verticalUnit = *screenHeight* / *devideNumber*;

unit = **Max** (*horizentalUnit*, *verticalUnit*);

screenUnitVector = **Vector**(*unit*, *unit*, *unit*);

unitVector = convertScreenPositionToWorldPosition(*screenUnitVector*);

counter = 0;

levelDivide = 0.1: *we use 0.1 as an example because this is the value we used in our final version, but it's still flexible.*

for each *object* in *O*: **do**

  if *counter* == 0: **then**

    object.scale = *unitVector*;

  else if *counter* is between (0, *levelDevide* **objectNumber*): **then**

    object.scale = *unitVector* * (1-*decreaseRatio*);

  else if *counter* is between (*levelDevide objectNumber*, 2**levelDivide***objectNumber*): **then**

    object.scale = *unitVector* * (1-2* *decreaseRatio*);

  …

  else

    object.scale = *unitVector* * (1-9* *decreaseRatio*);

  end if

  Add *object* to *S*;

end for

return S;

By reducing the sizes of objects based on the similarity, we not only stressed the most similar objects with the center, but also make better advantage of using the space. Because by making less-similar objects smaller, we can make them to fill in the empty gaps between the most-similar and bigger objects rather than waste those negative spaces. Figure 3.8 demonstrates the layouts with and without varying the sizes (the figure hasn't included the varying-color feature yet).

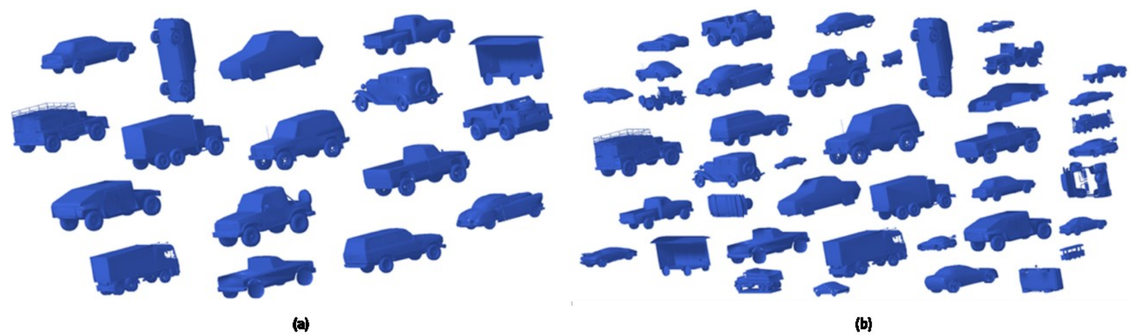


Figure 3.8: The layouts with applied varying size feature (b) and without (a).

The purpose of our system is to assist people to browse and search the 3D object dataset. Given this purpose, we proposed several design principles for the sizes of objects: 1. Provides good visual cues to help users understand the similarity relationship between the objects; 2. Make the sizes of similar objects big enough to observe and compare their features; 3. Make the sizes of less similar objects to be big enough in the case of the users' desired targets are different from current group. In terms of choosing a size decrease ratio to meet those requirements, there is a trade-off between principle 1 and principle 3, while size decrease ratio doesn't have influence on the principle 2. Figure 3.9 shows the layouts with different size decrease ratios for their initial sizes. Based on the observations, as the size decrease ratio increases, the similarity relationship (principle 1 is satisfied) will be more obvious, while the sizes of less similar objects will decrease (principle 3 is violated); and vice versa. In our implementation, we choose the value 0.28 to be the size decrease ratio.

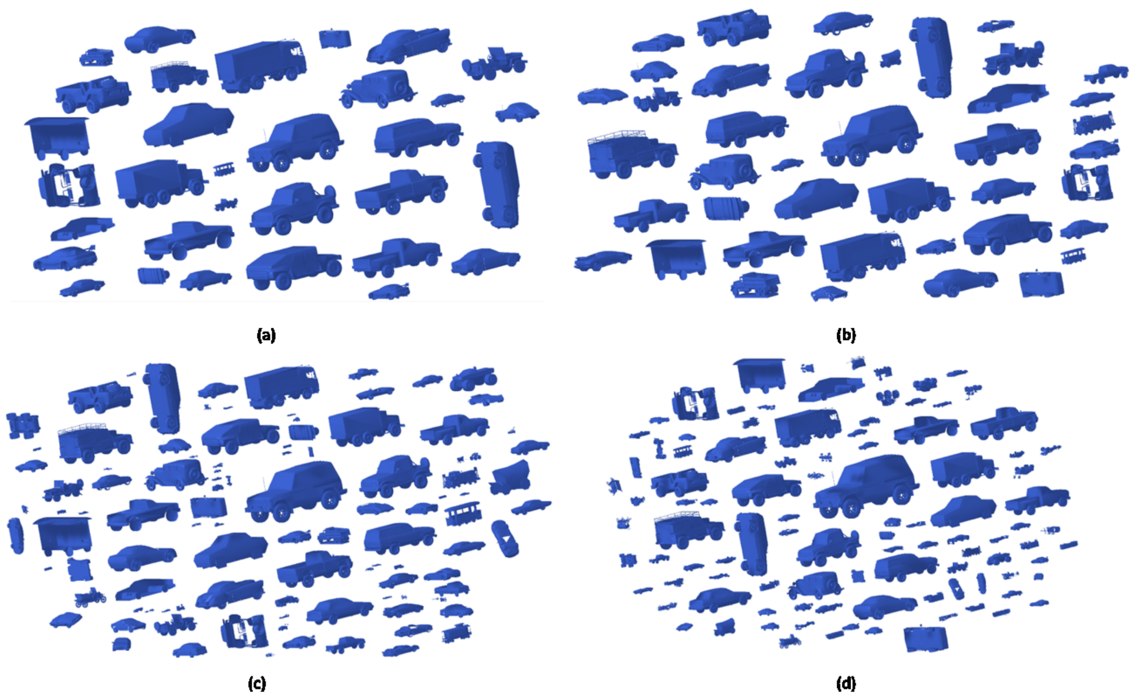


Figure 3.9: The example of layouts with different size decrease ratio for their initial sizes: (a) decrease ratio = 0.1; (b) ratio = 0.2; (c) ratio = 0.3; (d) ratio = 0.4.

Secondly, we combine the varying-color feature with the above size principle. The varying-color feature, similarly, means that the similar with the center, the color of the object will be more opaque, and vice versa. We expect that the combination of the size and color features will enhance the expression of the underlying pattern. Figure 3.10 illustrates the layouts applying the color feature to all the layouts in Figure 3.9.

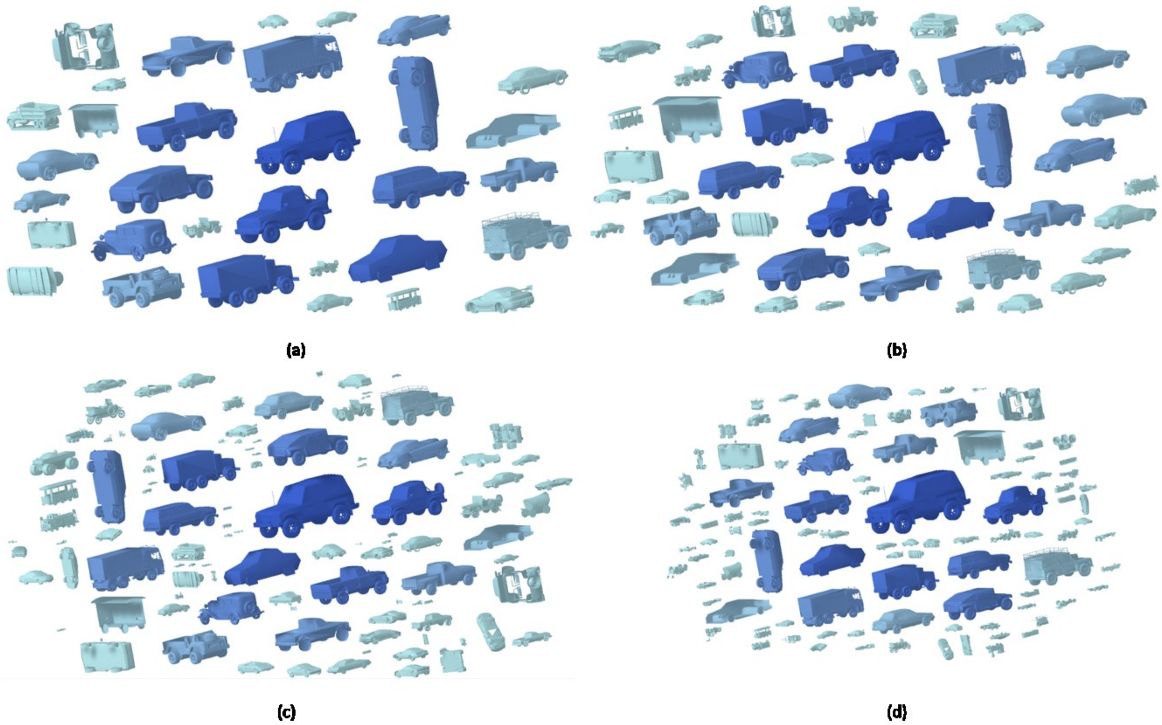


Figure 3.10: The example of layouts after being enforced color-tone feature.

By bounding these two features – size and color – with the similarity levels between objects, we believed it would assist users to have a better insight with the underlying pattern, so that they can search and browse the 3D object dataset more easily.

3.6 Staggered Animation

Based on the previous academic works on staggered animations (details in section 2.5), we decided to apply staggered animation for the transition process in the interactive 3D cloud view system, with the aim of helping users to retain visual coherence during the transition. In section 5, an user study has been conducted to evaluate this hypothesis.

In the interactive 3D cloud view system, users need to choose new center object constantly to browse the 3D object dataset. An animation transition will happen to all visible objects when switching to the new center, and the staggered animation design has been adapted to this transition process. According to the design principles in (Heer & Robertson, 2007) and (Chevalier, Dragicevic, & Franconeri, 2014), we decided to separate the animation transition into five stages, namely: 1. The non-related objects fade out; 2. Then, the remaining objects move to the new position and change their colors at the same time; 3. Third, all the objects rotate to the new angles; 4. Fourth, all the objects scale to the new sizes; 5. Finally, the new related objects fade in. We demonstrate this process in Figure 3.11.

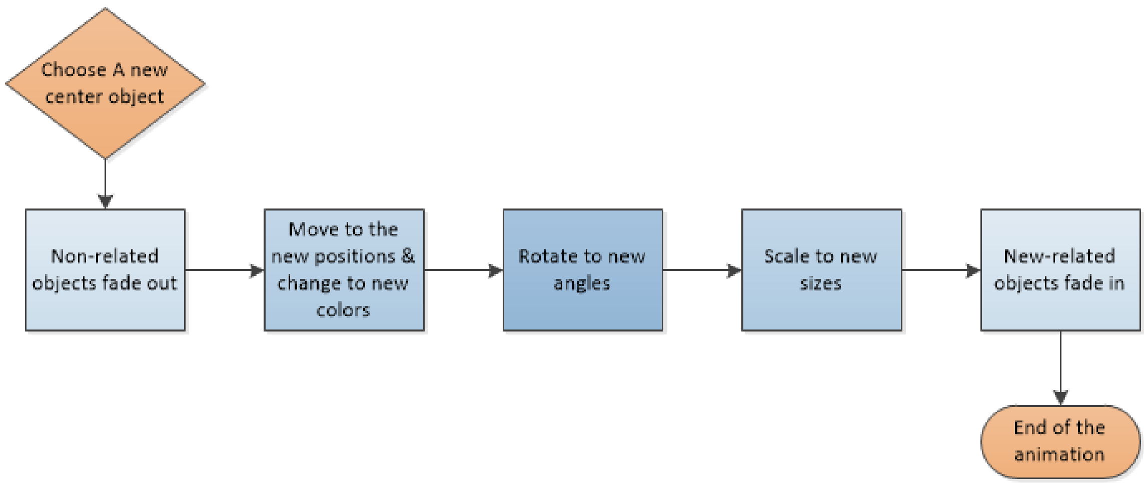


Figure 3.11: Staggered Animation Procedure Diagram

In every staggered animation stage, we used slow-in, slow-out timing instead of straight linear timing for the animation. By applying this timing, the animation will slowly start, then accelerate smoothly, finally decelerate in the end. According to the evaluation result in Dragicevic *et. al* 's work (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011), the slow-in, slow-out animation can effectively assist people to predict the motion compared with other types of temporal distortions of animated transitions (including the straight linear timing), because most of movements occur in the middle third of the animation duration. Figure 3.12 shows a diagram of how the animation progress changes over time by using a segment of the arctangent function.

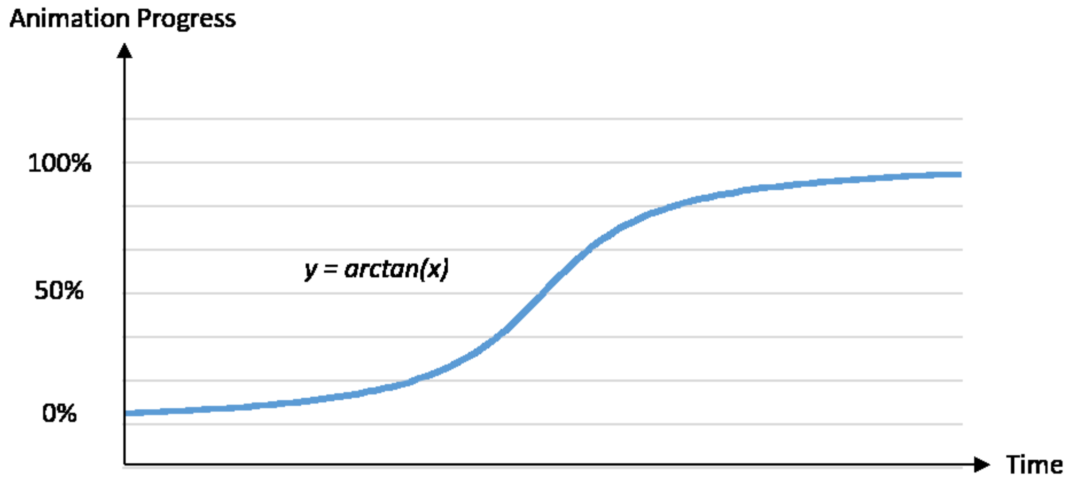


Figure 3.12: A schema of the slow-in, slow-out timing (Yee, Fisher, Dhamjia, & Hearst, 2001)

In terms of the implementation, we developed the animation effect by applying the iTween (iTween, 2014) package in Unity3D.

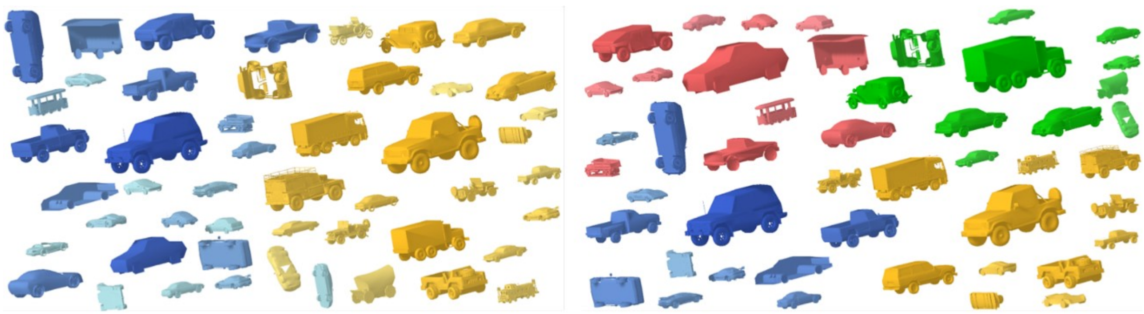
3.7 Additional Features

In addition to the above, our system provided some features to assist browsing and searching a 3D object dataset. However, based on the initial informal pilot studies, these additional features are not suitable to be included in the formal studies described in chapter 7 and 8. In the pilot study, it revealed that adding these features may cause too much of a learning burden for the participants (novice users). Thus, we introduce these additional features in this section as potentially serving more advanced users in a commercial system.

3.7.1 Multi-group Packing Layouts

We proposed multi-group packing layout to increase the sizes of the high-related objects and reveal the underlying similar pattern more clearly. In this layout, the related objects can be packed into different numbers of groups whose center is the most similar objects.

Our design purpose for this feature is to divide the packed objects into N balanced groups whose centers are the N most similar objects, and arrange these groups in a distributed way. The value of N can be up to 10, Figure 3.13 illustrates the packing layouts with different values of N (2 & 4)



(a) (b)

Figure 3.13: The layouts with different number of groups: (a) two groups; (b) four groups.

To implement this design, the packing algorithm for every group is similar to the packing algorithm described in the section 3.4. First of all, we put the N objects in the center of each group, the position of these centers are based on the results of the best circle packing in a rectangle with variable aspect ratio (Specht, 2012). In this case, the variable aspect ratio is decided by the ratio of the screen width and screen height. Second, we calculate the available packing area for each group based on the centers. The definition for each group's available packing area is the group of grids that are closest to this center rather the other centers. Third, we start packing the remaining objects one by one by putting them into the corresponding packing area, and then follow Algorithm 1.1 in that area. For example, after putting the first Nth objects into the optimal centers, the (N+1)th object will be put into the 1st group, the (N+2)th object will be put into the 2nd group, …, the (2N)th object will be put into the Nth group, then repeat this procedure until all objects are packed or the space is full.

We conducted an informal pilot study with fellow lab members to get an initial idea about how well our system works, and it is the same pilot study that is described in section 4.3 (the first one). Based on the results of pilot study, this layout has been suggested to be helpful to compare some highly similar objects because it makes these most similar objects become bigger and more detailed which may be easier to compare. We tested different numbers of groups in a simple pilot study which are 2, 4, 6. The result suggest that the 2 groups are more preferred, but the different colors between groups can be confusing because it seems to imply they're not similar but actually they are. We suggest in this type of groups, applying the same color and only let the sizes and color saturation to distinguish the groups will be more reasonable and less misleading. Another result from the pilot study is that the groups should not be more than 6, which will be too much for the limited screen size of mobile devices.

3.7.2 Switching Canonical Views

In the current version of this system, central objects in every new layout are initially shown the canonical views while the other objects are shown in the parallel angle, and then they can be free-rotated by the users. However, we also developed a "canonical views switching" feature, in addition to the free-rotation. Here users can switch between multiple canonical views of the central objects by horizontally swiping the screen, and then the other objects will change the viewpoints according to the current viewpoint of the center object. Then, the whole layout will be repacked in real time by applying the algorithm described in section 3.4.

The purpose of designing this feature is to assist users to use the shortest time to get the maximal valid information. By only browsing the multiple canonical views, we posited that the valid information can be delivered effectively since each canonical view is attempts to maximize the valid information according to the description in section 3.3. Although some canonical view algorithms will generate highly similar viewpoints, it's common that the other algorithm will generate significantly different viewpoints for the same object. Based on this observation, we decide to provide 5 different canonical views to ensure that users can get enough visual information for the object by browsing these views. Switching the canonical views only needs a simple swipe gesture which saves users time to drag through the screen and go through all the angles to find a suitable one.

To implement this feature, multiple canonical views have been pre-calculated by using different canonical view generating algorithms. For each object, it has 5 canonical views generated by the different algorithms – view area (Dutagaci, Cheung, & Godil, 2010), surface area entropy (Vazquez, Feixas, Sbert, & Heidrich, 2001), silhouette length (Polonsky, Patane, Biasotti, & Gotsman, 2005), silhouette entropy (Page, Koschan, Sukumar, Roui-Abidi, & Abidi, 2003), and mesh saliency (Lee, Varshney, & Jacobs, 2005) respectively . An example for multiple canonical views of an object is shown in Figure 3.4.

A small pilot study has been conducted on this feature which involved 6 lab members. We asked them to find a specific target within the dataset. In this version of the system, the objects can only be switched the canonical views rather than freely rotated. The participants all found the desired target, which suggests that multiple canonical views are providing enough information for browsing and searching. They also suggest that switching instead of rotating can save their searching time because every different view is canonical, but they also think it will be beneficial if the objects can be free-rotated as well.

We suggested that this feature may be useful when combined with the free rotation. But exactly how they might be combined is the subject of future work.

Chapter 4

User Study 1

The main purpose of the interactive 3D cloud system is to assist users to browse and search through a 3D object dataset in a more efficient and interesting way. To test that we have reached our goals, the effectiveness and error rate are key factors we evaluate in our experiment. Thus, combined with the pilot study's result which will be described in section 5.2, we choose three objective dependent factors to be evaluated, namely: i) Time of finishing the tasks; ii) Number of wrong targets chosen; iii) Confidence level of making decisions.

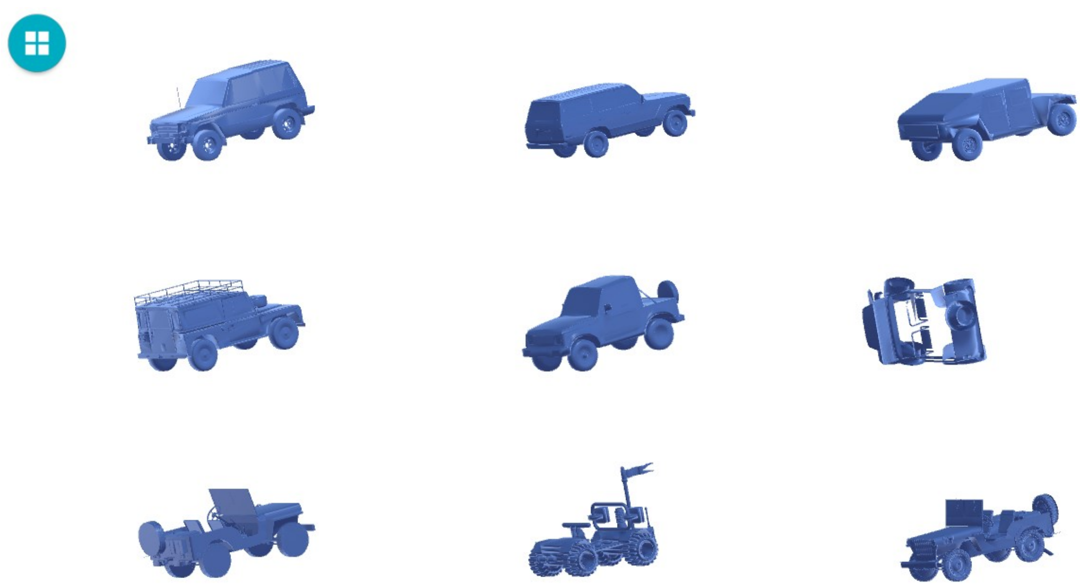


Figure 4.1: A typical layout screenshot of grid-based layout

Since there is no highly similar technique to browse and search a 3D object dataset, we decided to conduct a controlled user study comparing our approach with the typical commercial browsing technique that is commonly available – a grid view based approach. The typical layouts of the grid view based system are demonstrated in Figure 4.1, the available interactions of both cloud-view and grid-view based system are shown in Figure

4.2. Based on the observations described in section 2.6, we designed and developed the grid-view based system to compare with the cloud-view based system. In the grid-view based system, the 3D objects are arranged in grids in comparable sizes with the cloud-view based system's center object's size. If all objects are present in their own canonical views, the whole layout will look chaotic, so we use the adjusted method introduced in section 3.3 to parallel all the other objects with the up-left corner object. Every object can be double tapped to enter a full screen view, and then it can be free-rotated and scaled in this enlarged view. The order organizing the objects in the list is according to the categories which is the typical grouping method in commercial systems. For example, the chairs will gather together in an area, while the tables will be grouped together in another area. Both the cloud-view based approach and the grid-view based approach do not involve text searching and tag information. We were focusing on evaluating the key features of the cloud view based approach by comparing with the grid view based approach, and revealed some design principles for the cloud view based browsing 3D objects techniques on mobile devices.

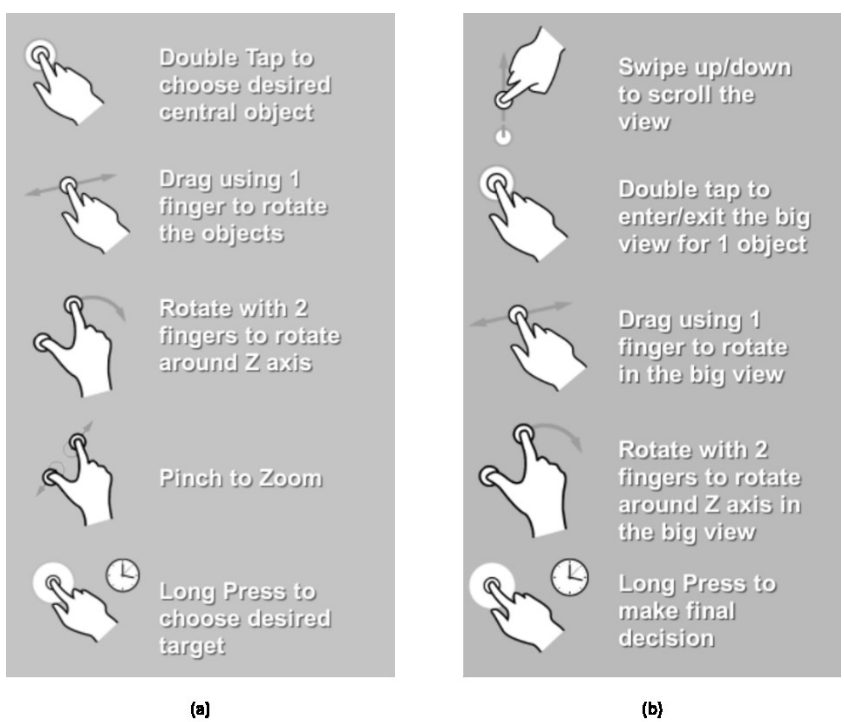


Figure 4.2: The lists of available interactions of (a) cloud-based system; (b)grid-based system.

The first study involves 48 participants, and it compared our system with the classical grid view system to evaluate which is better for browsing and searching the 3D-object datasets. The classical grid view was created based on the reviews of popular 3D model searching websites. This study only compared these two techniques' layouts in the specific searching, so the tags and text-based searching were not involved.

4.1 Specific Research Questions

As discussed in the section 1.2, the goal of our research is to provide a technique to improve the efficiency and reduce the error rates of users engaged in tasks involving browsing and searching of 3D objects. By conducting a user study, we want to test the following research hypotheses:

- The cloud view based (CV) approach is more efficient for the specific searching than the grid view based (GV) approach;
- The cloud view based (CV) approach is easier for comparing the similarity and differences between the objects and the targets than the grid view based (GV) approach;
- The cloud view based (CV) approach is more interesting/attractive to use compared to the grid view based (GV) approach;

4.2 Initial Study Design

The goal of this study is to compare the effectiveness between two types of techniques used to browse and search the 3D object dataset on the mobile devices. Hence, the measurements initially only consist of one independent variable – technique.

We designed a within-subject study with two different browsing technique (CV, GV). Considering the variation and universality of 3D object datasets, we include four different types of datasets which are mammals, furniture, vehicles, and aircrafts. The mammal dataset is used for the practice task before the formal tasks, while the other three datasets are used to perform the formal tasks. Because we regarded the different types of datasets only as a variation, we fully randomize the order of showing the three different datasets.

For each participant, he/she needs to perform three tasks that varied in types of datasets using both techniques as shown in Table 4.1. In every task, the participant needs to find a specific object in the corresponding dataset by using 2 techniques respectively.

Table 4.1: The corresponding task ID of each 3D object dataset

Dataset Type	Furniture	Vehicle	Aircraft
Task ID	Task 1	Task 2	Task 3

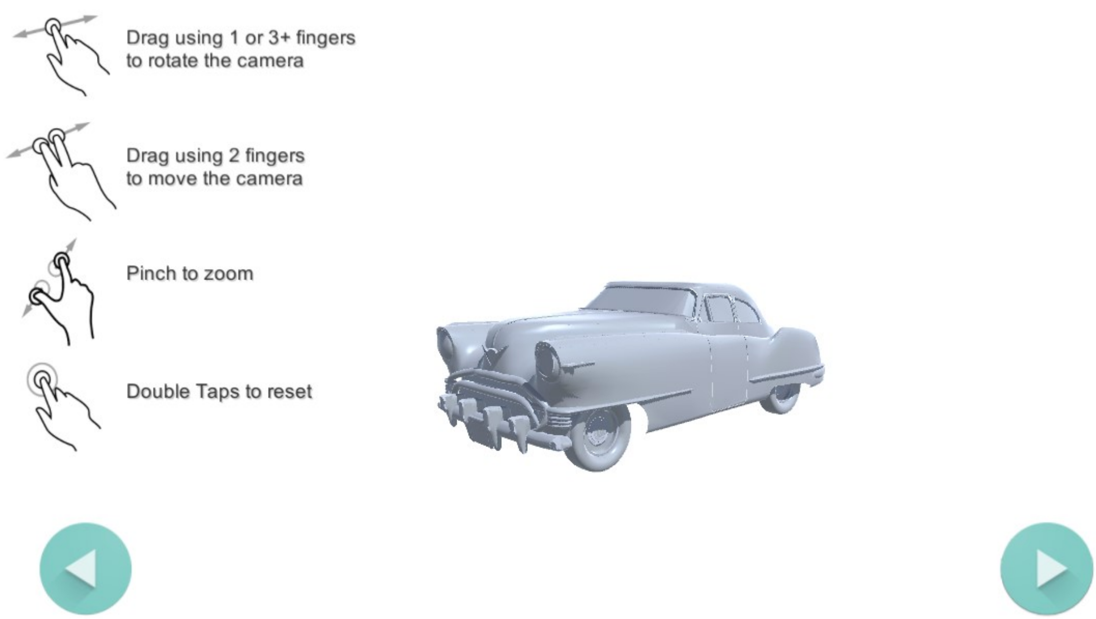


Figure 4.3: An example layout of the reference interface

All the models are downloaded from the Princeton Shape Benchmark (Princeton University, 2005) in the .off file format, and have been converted to the .obj format via the 3D model converter "meshconv" (Min, 1997). The models are all without inner parts, so only the visible parts will influence the model-similarity calculation. The number of objects in the four datasets are listed as follows: 51 objects in the mammal dataset; 198 objects in the furniture dataset; 103 objects in the vehicle dataset; 216 objects in the aircraft dataset. All the tasks were performed on a Windows Tablet. The target objects were shown as a reference by the researcher before every task on a separate Android tablet, and then the participants can browse the target object before and during the tasks. In the reference interface, the target object can be rotated, transformed, and scaled. Figure 4.3 shows an example of the reference interface.

4.3 Pilot Study

For formulating and refining our formal user study, we conducted two informal pilot studies with the lab members from the GEM Lab at Dalhousie University. Because we were not sure if there will be strong learning effect in the study, we first conducted a simple pilot study involves four participants. Each participant performed all three tasks by using both browsing techniques, so 6 tasks need to be performed in total. In each task, the participant tried to find the same target with both techniques. The order of using the techniques was counterbalanced, and the order of showing the different types of dataset was fully randomized.

In this pilot study, some general reactions and performance were observed, and led to some changes on the study design. First, as we can see from Table 4.2, whichever technique was used first took longer than the second, participants who used either technique in the first place spent more time than participants who previously finished the same tasks using another. Another observation is that participants tended to not choose the wrong object in the second system regardless of which technique it is. Based on the interview with the participants, the reason caused this learning effect was that the task become much easier since they already paid close attention to the details of the targets and the similar objects with the targets. This revealed that participants learned about the task itself and we need to control the effect of the learning effect. Second, participants also felt that they didn't explore the CV system sufficiently because only finding one object is fast and relatively easy. Third, another observation we learned from this study is that the participants hardly used the “multi-group packing” feature (section 3.7.1) in the CV system. In this version of the CV system, users can dynamically change the number of groups by vertically swipe gesture which means swiping up will increase the number of groups and swiping down will decrease it. Participants suggested that this function is kind of hard to understand and remember. Besides, it could be confusing that several groups of similar objects are rendered into different colors, because they thought only different objects should be highlighted in a different color. We needed to adjust this feature to decrease the learning burden and the chance of misunderstanding. Fourth, participants can make a random guess when they thought they just cannot find the exact same one, and this kind of action generally led to choosing the wrong object. We therefore should have a measurement to distinguish when the participant makes a random guess, and this will provide an insight as to whether the technique is really assisting users to browse and search.

Table 4.2: Average task finishing time using different orders of explosion techniques

Experiment Order	First Time	Second Time	Average
CV	109	81	95
GV	103	96	99.5

Based on the above observations in the first pilot study, we made a few changes on the study design, and then we conducted another pilot study to verify whether these adjustments are working. First of all, to control the learning effect, we changed the task from finding the same object in both techniques to finding different objects. To let participants fully explore the systems, we also increased the number of targets in each task from 1 to 3. Hence, each participant needed to find 3 objects in a dataset using one technique, and then find 3 different objects in the same dataset using another technique. The order of showing the datasets remained random, and the order of finding the 3 objects is also fully random. Another change was that we decided to not include the "multi-grouping layout" in the experiment since it was too complex for novice users, and we changed the design of this feature from highlighting different groups in different colors into the similar color as described in section 3.7.1. Finally, to distinguish the situation when the participant makes a random guess, a pop-up window was added as demonstrated in Figure 4.4. This window will pop-up each time after participants have made a decision, and they will need to choose how confident is this decision to move on.

Figure 4.4: The screenshot of pop-up "choosing confidence level" window

After these adjustments, we conducted our second pilot study with 6 participants. Each of the participants performed 6 tasks using both browsing techniques with counterbalancing. Based on the result of second pilot study, we found that there is no strong learning effect anymore as the targets are different. Another aspect we learned from this pilot study is that since the participants now can fully explore the system with an increasing number of targets, the strategies varied between different participants. We thought that the development of different strategies may be related to the spatial ability of the participants, so we also needed to measure the spatial ability of the participants to see if there are connections. We also found that there was a noticeable difference in performance when finishing tasks in different datasets. For example, the time for finishing the tasks in the aircraft dataset tends to always be the longest. Thus, we need to counterbalance the influence of different types of datasets. Finally, some participants suggested that the function "switching multiple canonical views" (section 3.7.2) is highly useful for quickly browsing the objects, while the other thought that it's not very helpful and can waste their time when trying to figure out how this function actually works.

4.4 Final Study Design

According to the results and observations of the second pilot study, we made some additional changes in our study design, as follows:

- We added a “spatial mental ability test” before the participants formally start performing the tasks to warm them up and test their spatial ability. The full questionnaire of the spatial ability test is attached as Appendix I.

We decided to take the type of datasets as another independent factor. Thus, the measurements are categorized into two independent variables which are: technique and type of datasets. The design of study becomes a 2×3 within-subject study, with browsing technique (CV, GV) and the type of datasets (furniture, vehicle, and aircraft). Each participant needed to perform 3 tasks that varied in the types of datasets using both techniques as shown in Table 4.1, and each task consists of 3 sub-tasks with 3 objects needing to be found. As shown in the Table 4.3, G1 and G2 represent the different groups consist of 3 objects within each type of dataset, the reason for this design is that based on the pilot study’s observation that there will be no strong learning effect as long as the participants need to find different objects in both techniques

Table 4.3: The sub-tasks index table

Dataset Type	Group ID	Technique Type	Task ID
Vehicle	G1	CV	1-C
		GV	1-G
	G2	CV	2-C
		GV	2-G
Furniture	G1	CV	3-C
		GV	3-G
	G2	CV	4-C
		GV	4-G
Aircraft	G1	CV	5-C
		GV	5-G
	G2	CV	6-C
		GV	6-G

- We removed the feature "switching multiple canonical views" from the CV system, because it may cause too much of a learning burden for some participants. We wanted to leave this feature for the usage of the advanced users and may be evaluated in the future.

After adjusting our study design, we did another pilot study of 4 participants to make sure the changes worked well. We permutated the three object groups and two techniques, such as Group1 has been matched with both CV and GV and been tested with two participants separately. Specifically, we wanted to verify that the complexity between different object groups are comparable. We interviewed them about whether they think the complexity of tasks of different groups are fair, and they did think the difficulty of object groups are fair and the reason causing the task difficulty's difference is different techniques not different objects. The remaining aspects of study design stayed the same as described in the initial study design, and more detailed procedure and design are introduced in the following sections.

4.4.1 Study Task and Protocols

The purpose of this study is to compare our techniques with the grid-based technique by efficiency, error rates and user experience of browsing and searching 3D objects datasets. In this study, we mainly focused on evaluating specific searching rather than general searching. In each task, the participant was asked to find three specific objects within a 3D objects dataset using either technique.

The general procedure of conducting experiments is shown in Figure 4.5. First, we presented slides to give the participant an introduction to the purpose of this study, the whole procedure of the study, the content of tasks, the features and functions of both techniques, and how to use them to finish the tasks. Second, to obtain basic information of the participants and let them warm up, we asked them to finish a couple of pre-task questionnaires, including the background questionnaire [see Appendix E] and the spatial mental test questionnaire [see Appendix I]. The spatial mental test questionnaire is comprised of the questions chosen from S.Neuburger et. al's (Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011) and Vandenberg.S.G. and Kuse.A.R's researches' (G & R, 1978) materials. Third, we trained them with 1 training task with each technique. The training task was conducted in the mammal dataset, and there were two objects that need to be found in the given order in each task. Once finished the practice session, we made sure that participants' questions have been satisfactorily answered, and there was no misunderstanding or confusion. Then, we proceeded with the formal tasks. There are three formal tasks using one of the two browsing techniques, so each participant performed 6 tasks consisting of 18 sub-tasks in total. In each task, there were three specific objects that need to be found in the given order which is fully randomized. Two tablets were involved in every task: one tablet showed the three targets as references to let participants browse them before and during the tasks; another tablet was where the tasks are performed. The participants used both techniques on the latter tablet to browse the dataset and search for the targets. After finishing each task, the participants were asked to fill a post-task questionnaire [see Appendix F1] to compare the two techniques. Finally, they filled an interview questionnaire [see Appendix G1] followed by a short interview with the researcher.

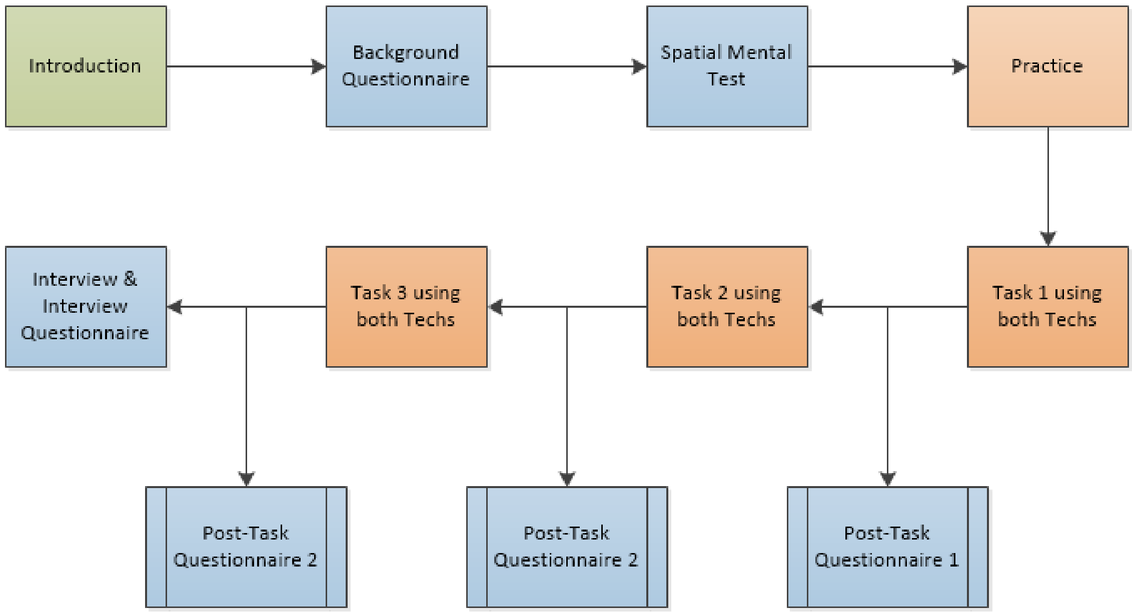


Figure 4.5: The general procedure of the user study

Before the above study procedure, all participants involved in the study read and signed an informed consent form [see Appendix D1] which introduced the details of the study, the participants' right to withdraw from the study without consequence, what data is collected and the assurance of confidentiality and anonymity of personal data, and the compensation [see Appendix H] they will receive.

4.4.2 Participant Recruitment

After we received the research ethics board approval letter [see Appendix J], we recruited 48 participants from the Dalhousie University campus through different recruitment platforms, including Notice Digest, the Computer Science Mailing List, the Dal Student Email lists and physical bulletin boards. The email announcement and the poster recruitment are given in Appendix A1 and Appendix B1 respectively. When there is a potential participant who replies to our notice/poster, we sent a screening email [see

Appendix C] to ensure that the potential participants meet the inclusion criteria (not color blind, English fluency, and experience with touchscreen devices) and are aware of the important aspects (audio and video recording). Not being color blind is important because the CV technique involve the gradient-change of color brightness (section 3.5), and the participants need to be able to distinguish the color brightness to use this feature. Second, all our resources including the prototype application, electrical materials and interview are in English which makes English fluency a critical. Third, having the basic experience with touchscreen devices is necessary since our experiment is carried out on tablets and requires interactions on touch screens. Finally, participants need to be aware of some personal data will be collected (detailed in section 4.4.4) so they can be involved in the experiments without misunderstanding.

4.4.3 Counterbalance Measure

On the purpose of counterbalancing the conditions in the experiment, we made the full permutations of the independent factors (2 techniques) and the potential influence factors (3 types of datasets, and 2 different object groups), and we fully randomized the order of finding the objects within each object group. Hence, we need $2! \times 3! \times 2! = 24$ participants to fully conduct all cases, and we repeated every case twice which leads to $24 \times 2 = 48$ participants and $48 \times 3\, subtasks = 144\, cases$. We decided to fully randomized the order of finding three objects because it would require 168 participants to achieve a full permutation which could be infeasible. The full task arrangement for this user study is demonstrated in Table 4.4. Participant ID is a sequence number we gave to each participant. For example, P1 means the first participant, P2 means the second participant, etc. Technique order means the order of techniques used to perform each pair of tasks (there is a post-task questionnaire after each pair of tasks as described in 6.4.1). Dataset order indicates the order of performing the three types of datasets, and V, F, and A stands for Vehicle, Furniture, and Aircraft respectively. Task IDs represents each individual task as illustrated in Table 4.4.

Table 4.4: The task arrangement of the first user study

Participant ID	Technique Order	Dataset Order	Task IDs					
			Task1		Task2		Task3	
P1, P13	1. CV 2. GV	V, F, A	1-C	2-G	3-C	4-G	5-C	6-G
P2, P14		V, A, F	1-C	2-G	5-C	6-G	3-C	4-G
P3, P15		F, V, A	3-C	4-G	1-C	2-G	5-C	6-G
P4, P16		F, A, V	3-C	4-G	5-C	6-G	1-C	2-G
P5, P17		A, V, F	5-C	6-G	1-C	2-G	3-C	4-G
P6, P18		A, F, V	5-C	6-G	3-C	4-G	1-C	2-G
P7, P19	1. GV 2. CV	V, F, A	2-G	1-C	4-G	3-C	6-G	5-C
P8, P20		V, A, F	2-G	1-C	6-G	5-C	4-G	3-C
P9, P21		F, V, A	4-G	3-C	2-G	1-C	6-G	5-C
P10, P22		F, A, V	4-G	3-C	6-G	5-C	2-G	1-C
P11, P23		A, V, F	6-G	5-C	2-G	1-C	4-G	3-C
P12, P24		A, F, V	6-G	5-C	4-G	3-C	2-G	1-C
P25, P37	1. CV 2. GV	V, F, A	2-C	1-G	4-C	3-G	6-C	5-G
P26, P38		V, A, F	2-C	1-G	6-C	5-G	4-C	3-G
P27, P39		F, V, A	4-C	3-G	2-C	1-G	6-C	5-G
P28, P40		F, A, V	4-C	3-G	6-C	5-G	2-C	1-G
P29, P41		A, V, F	6-C	5-G	2-C	1-G	4-C	3-G
P30, P42		A, F, V	6-C	5-G	4-C	3-G	2-C	1-G

P31, P43	1. GV 2. CV	V, F, A	1-G	2-C	3-G	4-C	5-G	6-C
P32, P44		V, A, F	1-G	2-C	4-C	3-G	6-C	5-G
P33, P45		F, V, A	3-G	4-C	1-G	2-C	5-G	6-C
P34, P46		F, A, V	3-G	4-C	5-G	6-C	1-G	2-C
P35, P47		A, V, F	5-G	6-C	1-G	2-C	3-G	4-C
P36, P48		A, F, V	5-G	6-C	3-G	4-C	1-G	2-C

4.4.4 Data Collection

As we described at the beginning of this section, we have 3 dependent factors that need to be evaluated, namely: i) Time of finishing the tasks; ii) Number of incorrect targets; iii) Confidence level of making decisions.

Based on this design, we recorded the timings of participants' attempts at finding and selecting a target object, the number of errors they made during this process, and the confidence levels when they made decisions. The timing of finishing the task is defined as the duration between the timestamp of the participants from making the first touch and the timestamp when the participants make the last correct decision. The number of errors is defined as when participants mistakenly select objects that are not the target objects, and we eliminated the possibility of accidently making wrong choices by presenting a decision window (as shown in Figure 4.6). The confidence level is defined as a score chosen by participants from 1 to 5 from a pop up window as illustrated in Figure 4.4. Additionally, all the interactions made by the participants were also recorded in the log files, and we also video-recorded the full procedure of every experiment to obtain detailed knowledge about how participants browsed and searched the 3D object dataset on mobile devices and where they might have difficulty (difficulties to find a promising group, hard to observe the objects in detail, etc.). After finishing each pair of tasks (e.g. 1-C and 2-G in Table 4.4), participants were given a 1-minute break and then filled out a post-task questionnaire [see Appendix F1] in order to provide subjective quantitative data, so there were 3 post-task questionnaires in total. We also asked participants to fill out a post-session interview questionnaire [see Appendix G1] and gave them an audio-recorded interview to obtain additional feedback, which might not be apparent in the previous data and may help us to more comprehensively understand the preferences of a participant in terms of the two browsing and searching techniques.

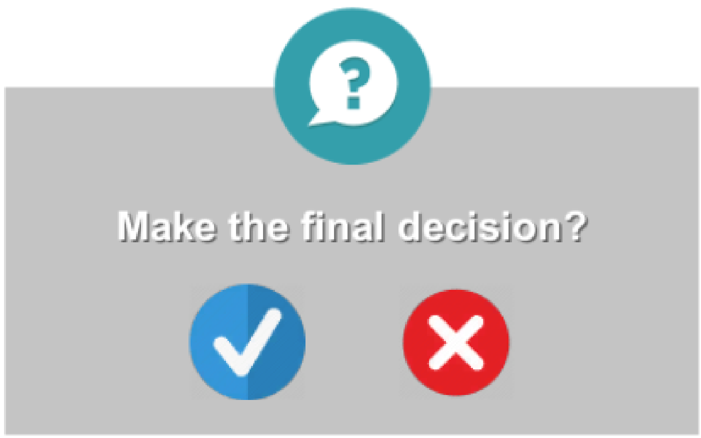


Figure 4.6: The screen shot of pop-up "making the final decision" window

4.5 Study Results

As discussed above, our log system recorded each participant's task finishing time, wrong target selection times, and the confidence level of making every decision. Participants also filled out three post-task questionnaires after finishing each pair of tasks and a post-session interview questionnaire after finishing all 6 tasks, and we conducted a short interview with voice recording. We had two independent variables and three dependent variables as listed below:

- *Independent variables*
    - Techniques: CV (cloud-based technique) and GV (grid-based technique)
    - Datasets with three types: furniture (number of objects is 198), vehicle (number of objects is 103), and aircraft (number of objects is 216)
- *Dependent variables*
    - Time of finishing tasks
    - Number times choosing the wrong targets
    - Confidence level of making decisions

4.5.1 Task Finishing Time

Task finishing time is defined as the duration from when the participant started first-touching the screen to the final selection of the correct target. As we described in section 4.3 and 4.4, we decided to take the type of dataset as a potential independent factor based on the observations in pilot study, so we used a two-way repeated-measures ANOVA to analyze the effects of the two independent factors' (techniques and datasets) interactions. Table 4.5 shows the descriptive statistical results of these two independent variables' interaction. Because this experiment is a within-subject design study, the sample size *N* is equal to 48 in all cases.

Table 4.5: Descriptive statistics of task finishing time in seconds combined with techniques and type of datasets

Type of Datasets	Technique	Mean	Std. Deviation
Furniture	CV	54.0140	18.20628
	GV	49.1215	19.74183
Vehicle	CV	70.9514	24.42928
	GV	86.5972	39.18827
Aircraft	CV	125.0764	49.96501
	GV	163.2292	64.97842

Figure 4.7 illustrates the plot for two-way interaction between the techniques and types of datasets. The number of objects in the datasets furniture, vehicle, and aircraft is 198, 103, and 216 respectively. With the observations in both the pilot and the formal experiment, we suggest that the complexity of datasets in this experiment has the relation: $Complexity(furniture) < Complixity(vehicle) < Complexity(aircraft)$. We also suggest the reasons that cause this difference involves many aspects, including the similarity between the objects, the number of objects, and how familiar the subjects are with the objects in datasets. For example, the objects in the aircraft dataset are highly similar because they're all airplanes with different components and people are not very familiar with different kinds of airplanes, while the objects in the furniture dataset are not that similar with various sub-categories (chairs, tables, and shelves, etc.) and people are familiar with various types of furniture. As seen from the graph, there is an intersection of the two techniques' lines. On the furniture dataset, the GV's finishing time is less than the CV's. Then, on the both vehicle and aircraft datasets, the CV's time performance is over the GV's. The influence effects of two different browsing techniques, three types of datasets, and their interactions have all been found to be significant (P-value < 0.005). Moreover, the mean task-finishing time of CV (M = 83.347s) is less than the one of GV (M = 99.649s), and the mean finishing time of three datasets is increasing along with the complexity increases (M = 51.568s, 78.774s, and 144.153s). These give further reason that both techniques and datasets have primary effect on the task-finishing time.

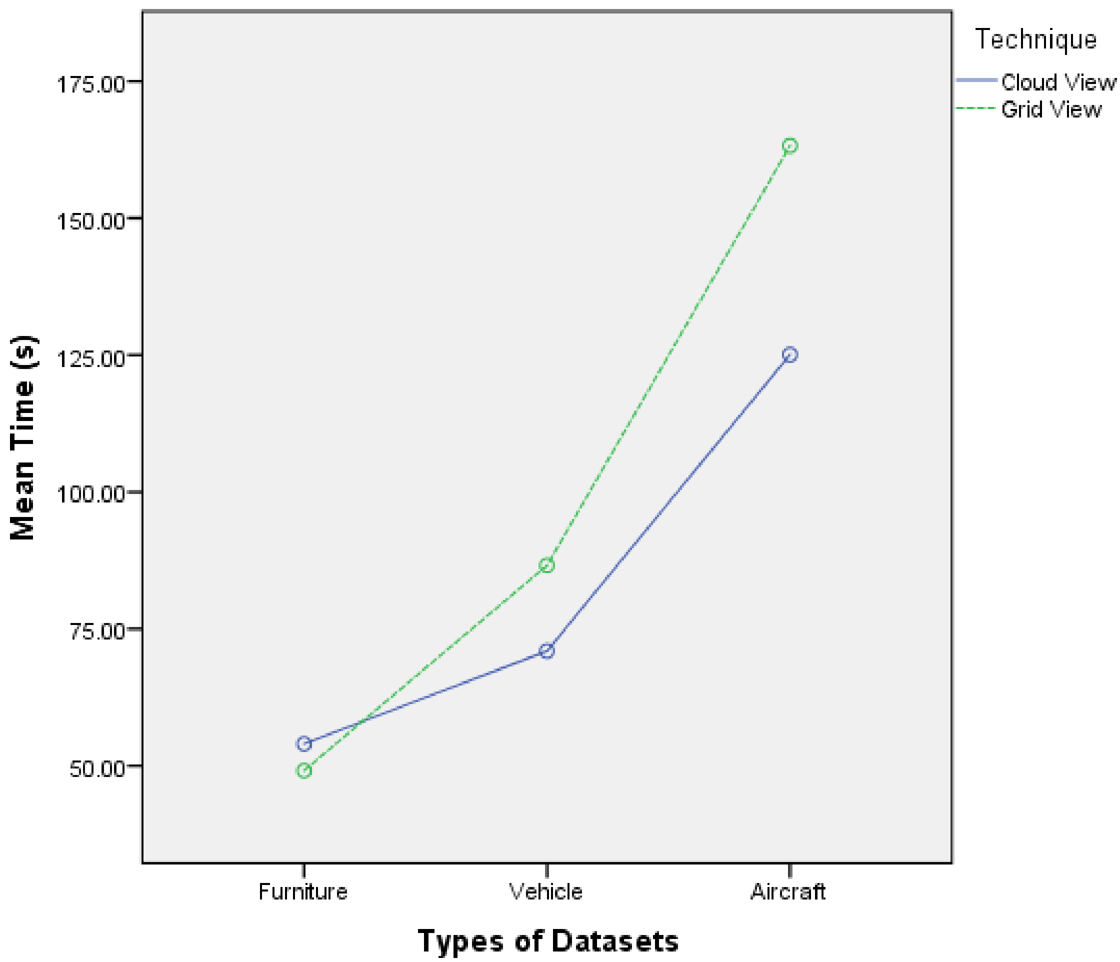


Figure 4.7: Interactions between techniques and type of datasets on mean task finishing time were found to be significant.

As observed from analyzing results when the complexity of the datasets increases, more time is required to finish tasks and this applies to both techniques. Moreover, the average task completion time has a small difference at the low complexity dataset – furniture. As the datasets become more complex (vehicle and aircraft), the difference of average task finishing time between two techniques increases significantly, and the CV's time performance is significantly over the GV's. In other words, our technique – the 3D interactive cloud view – has a stronger performance advantage when the datasets become more complex and there are higher similarity levels within the dataset.

We analyzed the result with two-way repeated-measures ANOVA to see if any factor had significant effects on the average task finishing time. The analyzing results indicates that there were statistically significant effects for the techniques, datasets, and their interaction ($F1 = 14.918$, $P1 < 0.001$; $F2 = 136.670$, $P2 < 0.001$; $F3 = 10.735$, $P3 < 0.005$) because the P-values for both factors were less than 0.05. Hence, there was a statistically significant difference between the two techniques in the average task finishing time. In other words, the tasks using CV costs less time than the tasks using GV. Also, the different types of datasets also have the significant influence on the task finishing time, and on the performance of two techniques. Combined with the observations in the above graph, it takes more time to finish the task as the complexity of the datasets increases. The CV shows a strong performance gain when the similarity level within the datasets is high, and doesn't show this advantage when the similarity level is relatively low. In addition, based on the Partial Eta Squared (J, 1988), the value obtained for techniques, datasets and their interaction were 0.241, 0.744 and 0.186 respectively, which suggested a very large effect size (0.01 = small effect, 0.06 = moderate effect, 0.14 = large effect).

4.5.2 Number of Wrong Choices

The number of wrong choices is defined as the average number of instances when participants mistakenly select objects that are not the target within one task. Table 4.6 shows the descriptive statistical results of these two independent variables' interaction. The sample size *N* is equal to 48.

Table 4.6: Descriptive statistics of Mean number of wrong choices combined with techniques and type of datasets

Type of Datasets	Technique	Mean	Std. Deviation
Furniture	CV	0.271	0.077
	GV	0.625	0.145

Vehicle	CV	0.292	0.084
	GV	0.667	0.109
Aircraft	CV	0.458	0.103
	GV	0.875	0.224

Figure 4.8 shows the plot for two-way interaction between the techniques and types of datasets over the estimated marginal means of number of wrong choices. There was no significance found in the interaction of these two factors (F = 0.037, P-value = 0.934).

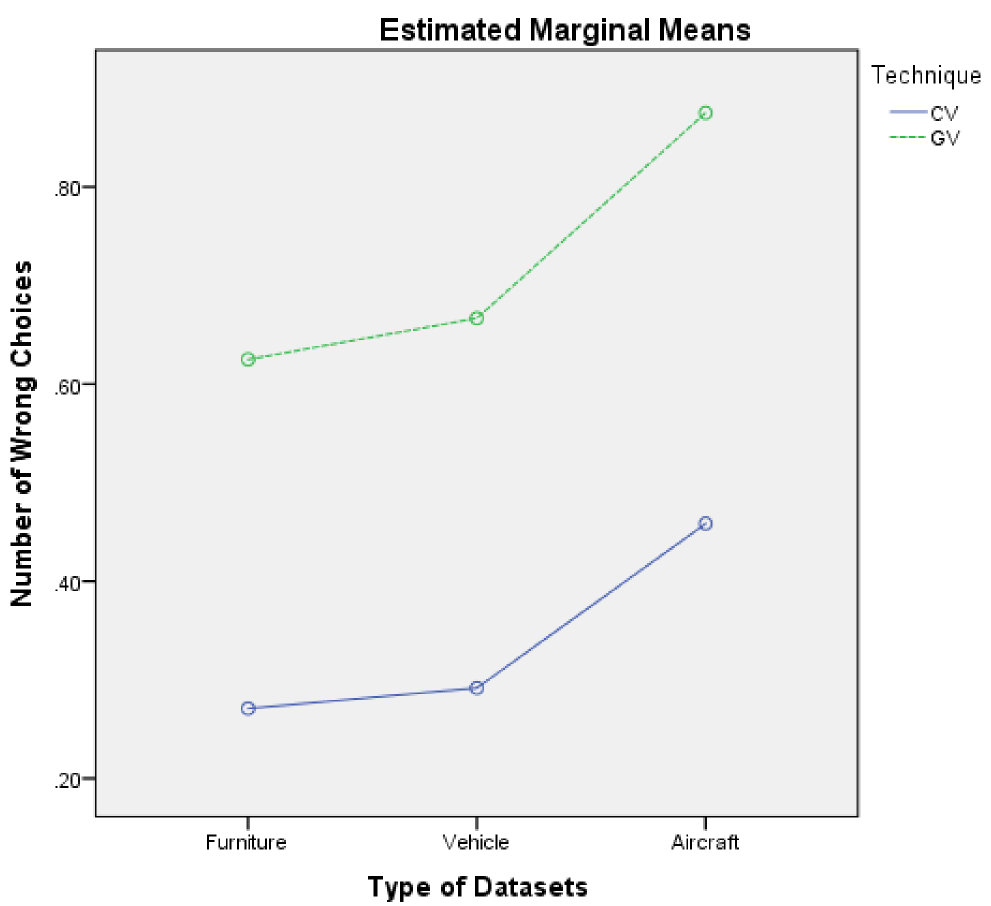


Figure 4.8: Interactions between techniques and type of datasets on mean number of wrong choices were found not to be significant

Figure 4.8 indicates that the tasks using the CV technique yields on average fewer wrong choices than the tasks using the GV technique over all three types of datasets without intersection. In addition, the mean number of wrong choices of CV (M = 0.340) is less than the mean of GV (M = 0.722), which provides further evidence that the CV technique has a better performance over the number of wrong choices than the GV technique. Also, we can see from the above graph that the number of wrong choices follows a consistent pattern over the different datasets: furniture (M = 0.448), vehicle (M = 0.479), and aircraft (M = 0.667).

To observe if there is any factor having significant effects on the mean number of wrong choices, we applied the two-way repeated-measures ANOVA to analyze the result. We found that there is statistically significant difference (F = 24.283, P-value < 0.001) between the two browsing techniques in this case since the P-value was smaller than 0.05. This reveals that the tasks using the CV technique lead to fewer wrong choices than the tasks using GV. Moreover, as we discussed before, the effect size on the different techniques was large according to the Partial Eta Squared (J, 1988) value which was 0.341 (> 0.14).

On the other hand, we discovered that there is no statistically significant influence from the types of datasets, and its interaction with the different techniques (F1 = 2.071, P1 = 0.140; F2 = 0.037, P2 = 0.964). By observing this, we know that the different types of datasets did not have significant influence on the number of wrong choices, and the above conclusion that the CV technique will lead to less wrong choices works on all three types of datasets.

4.5.3 Level of Confidence

The level of confidence is defined as the average confidence score of participant's decisions (including both correct and wrong decisions) within a task. Every single confidence score was chosen by the participants after each decision by clicking on the score of confident-window (as shown in Figure 4.4) from 1 point for "not confident" to 5 points for "very confident". Table 4.7 shows the descriptive statistical results of these two independent variables' interaction. The sample size N is still equal to 48 in all cases as well.

Table 4.7: Descriptive statistics of Mean level of confidence combined with techniques and type of datasets

Type of Datasets	Technique	Mean	Std. Deviation
Furniture	CV	4.413	0.086
	GV	4.477	0.081
Vehicle	CV	4.418	0.072
	GV	4.333	0.085
Aircraft	CV	4.146	0.099
	GV	4.162	0.116

Figure 4.9 shows the plot for two-way the interaction between the techniques and types of datasets over the estimated marginal means of level of confidence. There was no significance found in the interaction of these two factors (F = 1.531, P-value = 0.222).

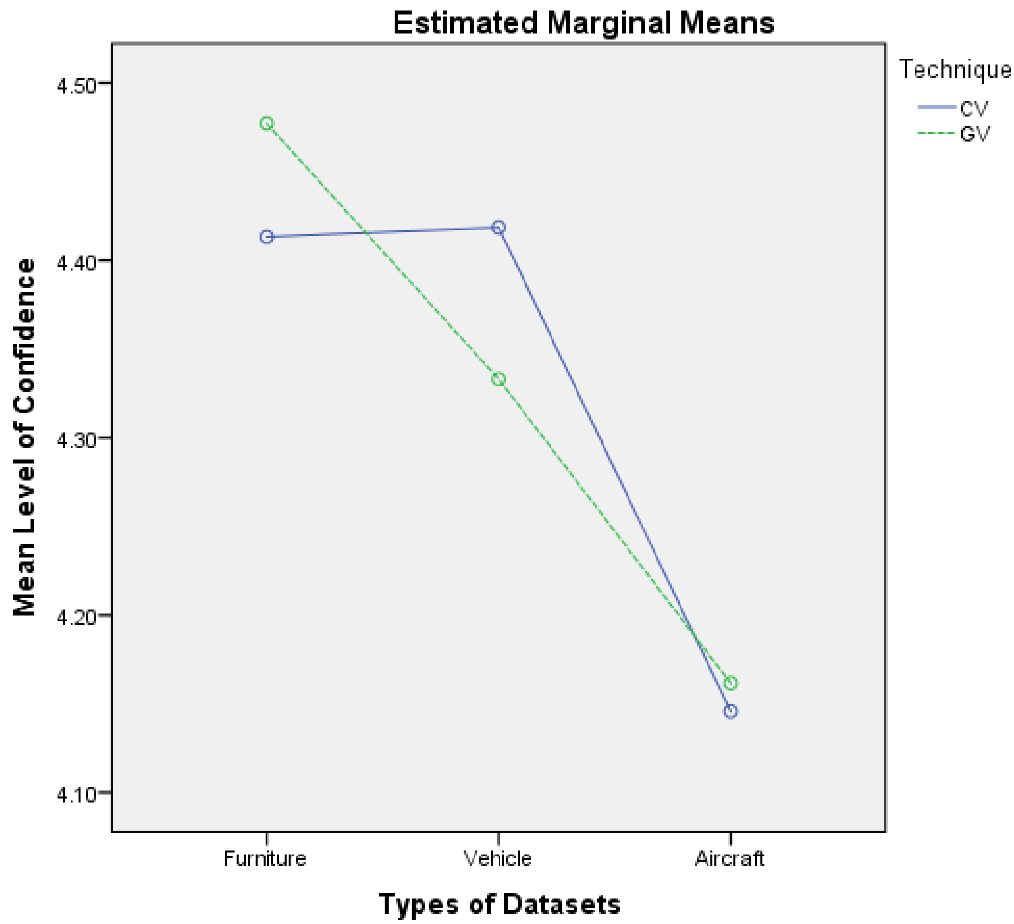


Figure 4.9: Interactions between techniques and type of datasets on mean level of confidence were found not to be significant

As we can see from the Figure 4.9, there is no clear trend of the difference of the level of confidence between two browsing techniques. Also, the average confidence level of CV and GV are 4.326 and 4.324 respectively, which are almost the same. In terms of the types of datasets, the confidence level decreases as the complexity of datasets increases using the GV technique with the values of mean confidence 4.477, 4.333. and 4.162 respectively. With the CV technique, there is a similar phenomenon except the vehicle dataset, and the average confidence values of three datasets are 4.413, 4.418, and 4.146 respectively in this case.

By analyzing the result with two-way repeated-measures ANOVA, we found out that there was no statistically significant difference on two browsing techniques ($F = 0.004$, P-value

= 0.951) since the P-value is greater than 0.05. This means that the average confidence level of participants generally stays the same using different techniques, and the value of mean confidence level is good in general (near 4.3 out of 5). Conversely, the influence of different types of datasets turns out to be statistically significant with the F = 15.020, and P-value < 0.001. It indicates that the mean confidence level significantly decreases with the dataset's complexity increasing, but the Partial Eta Squares (Cohen, 1988) value is only 0.032 which shows only a moderate effect. The interaction of these two factors (techniques and datasets) has no significant influence on the confidence level with the analyzing value F = 1.531, and P-value = 0.222 whose P-value is greater than 0.05.

4.5.4 Questionnaire

We asked every participant to fill out a background questionnaire [see Appendix E] and a spatial ability test [see Appendix I] before starting the tasks, and then the participant needed to fill out a post-task comparison questionnaire [see Appendix F1] after finishing each task. Thus, the participant needed to fill out two pre-task questionnaires and three post-task questionnaires in total. In the background questionnaire, we mainly collected information about gender, age, and how familiar the participant was with 3D models and touch devices. Then, we evaluated the mental spatial rotation ability of participants in the spatial ability test, which also warmed them up. Finally, in the post-task questionnaire, we included the questions about evaluating the control interface, the difficulty level of finishing tasks, and the difficulty of browsing and comparing objects.

First of all, we will present the background information of the 48 participants in this experiment. The experiment involves 34 males and 14 females from the Dalhousie University with the ages range from 18 to 47 with average of 25.4. The other results of the background questions are demonstrated in the Figure 4.10 (a) to (f). As the background results indicated, all of our participants are familiar with the touch-screen devices and mobile devices, and most of them are familiar with 3D models. 54.1% and 56.3% of participants had hardly interacted with 3D models on the computers and mobile devices respectively. Also, only few of them has experience with tag clouds.

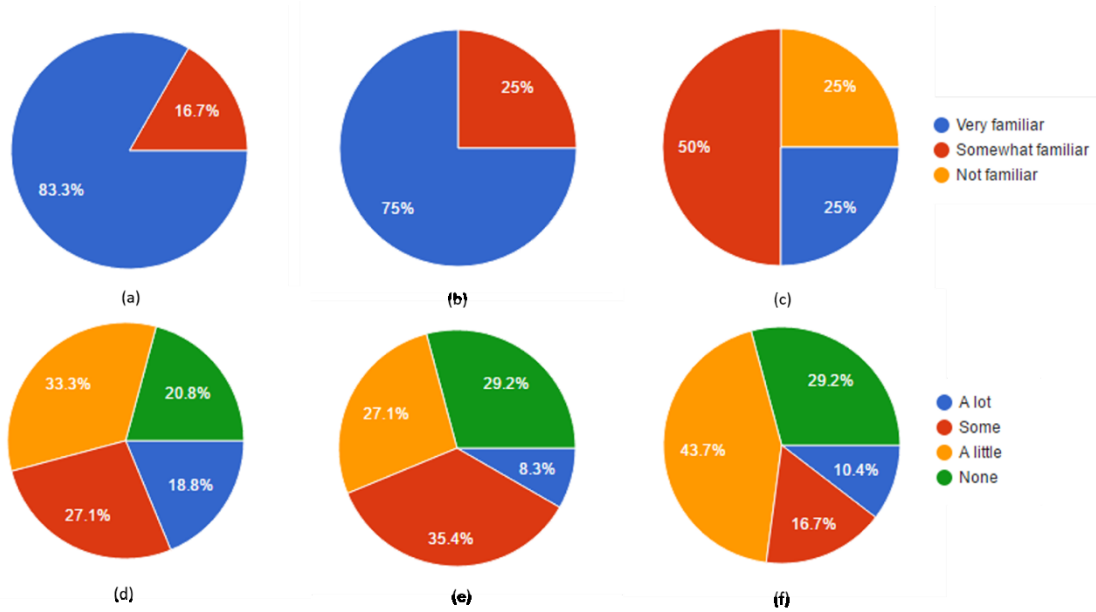


Figure 4.10: The pie diagram of the results of background questions: (a) How familiar are you with mobile devices; (b) How familiar are you with touch-screen devices; (c) How familiar are you with 3D models; (d) Do you have experience on interacting with 3D models on a computer; (e) Do you have experience on interacting with 3D models on a tablet or other mobile devices; (f) Do you have experience with tag clouds.

We also listed the results of different questions of post-task questionnaires in Figures 4.11 – 4.18.

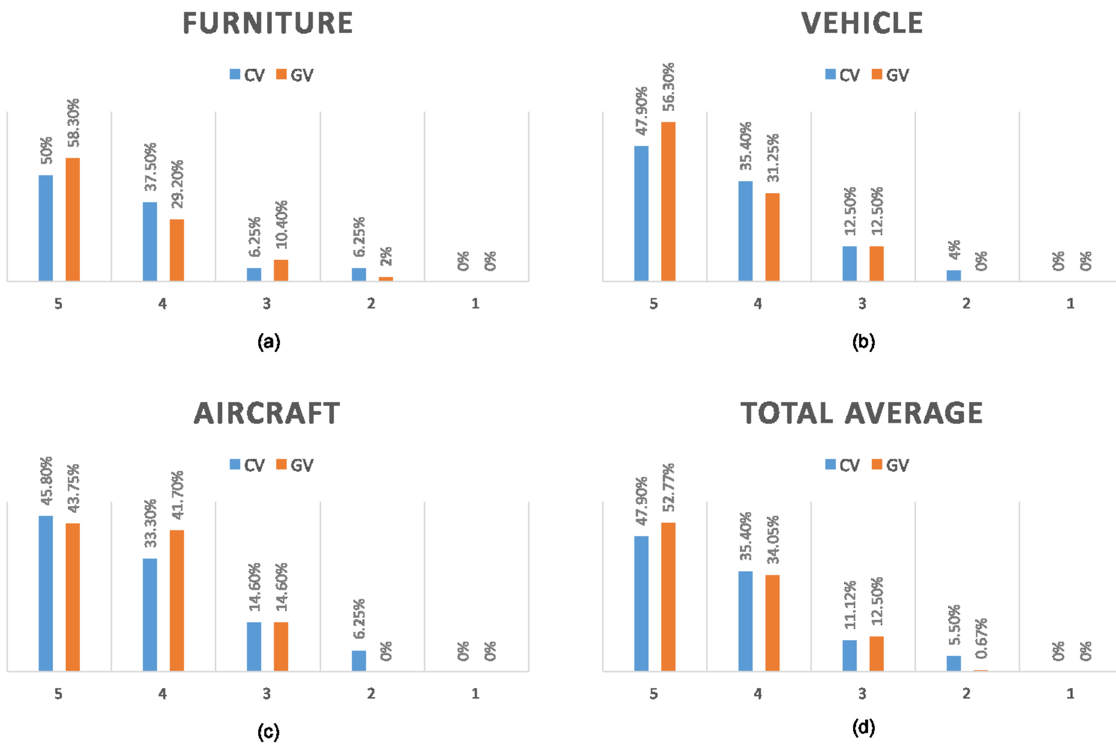


Figure 4.11: CV and GV comparison for question 1 of the first user study (“The controlling interface of application is easy to understand. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

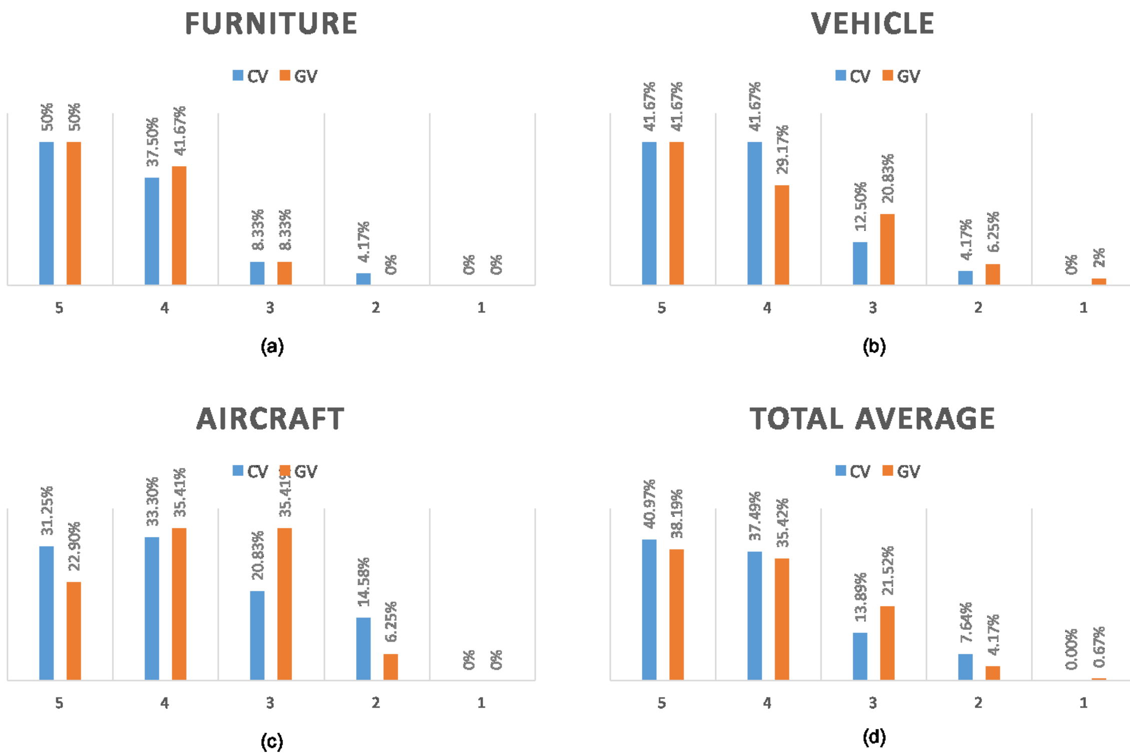


Figure 4.12: CV and GV comparison for question 2 of the first user study (“It was easy to finish the task without any difficulty. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

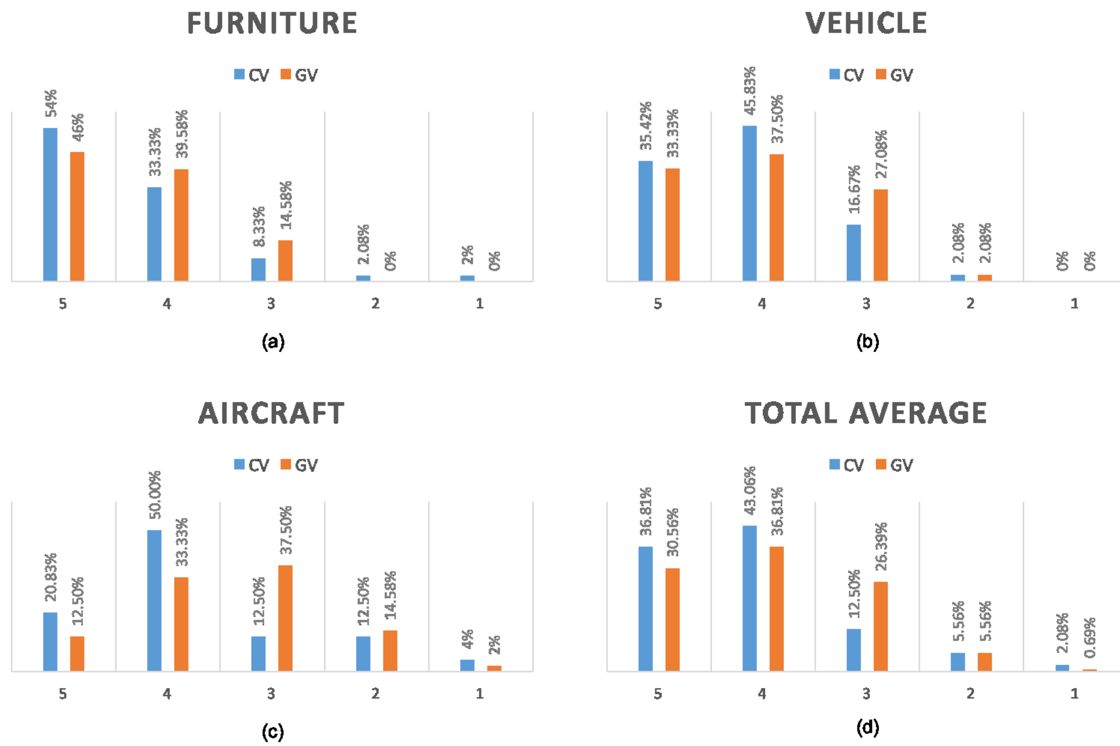


Figure 4.13: CV and GV comparison for question 3 of the first user study ("It was easy for me to find the targeted objects. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree"): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

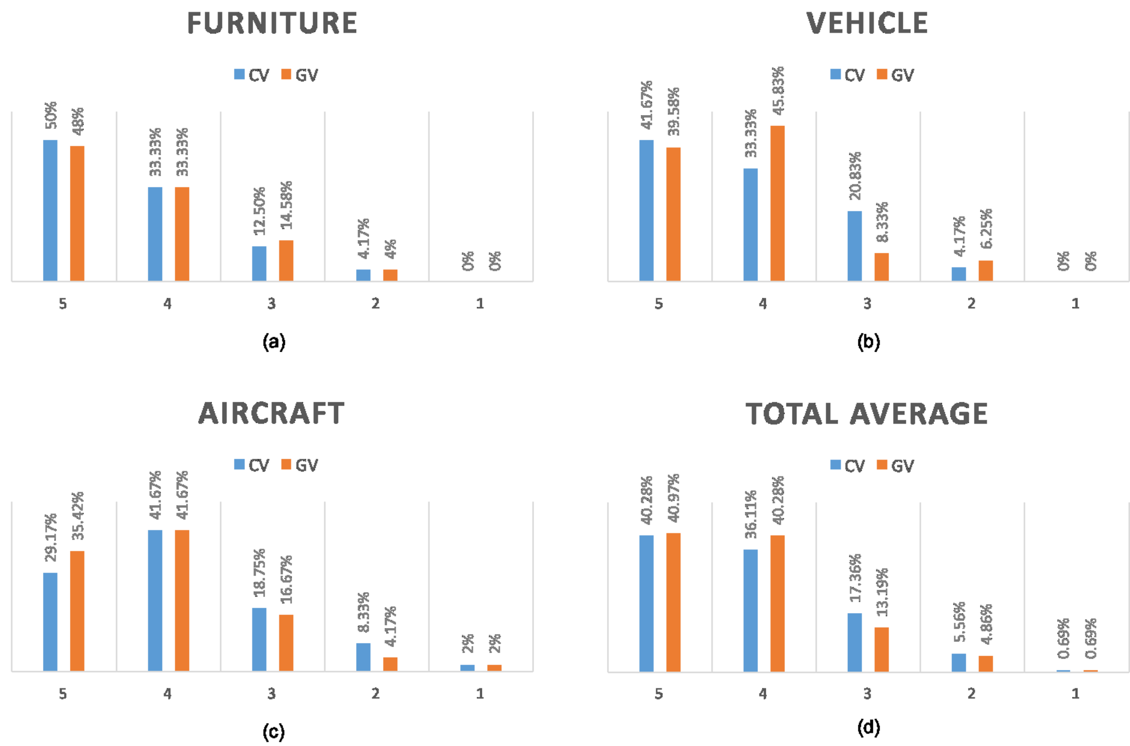


Figure 4.14: CV and GV comparison for question 4 of the first user study ("It was easy to browse the models. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree"): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

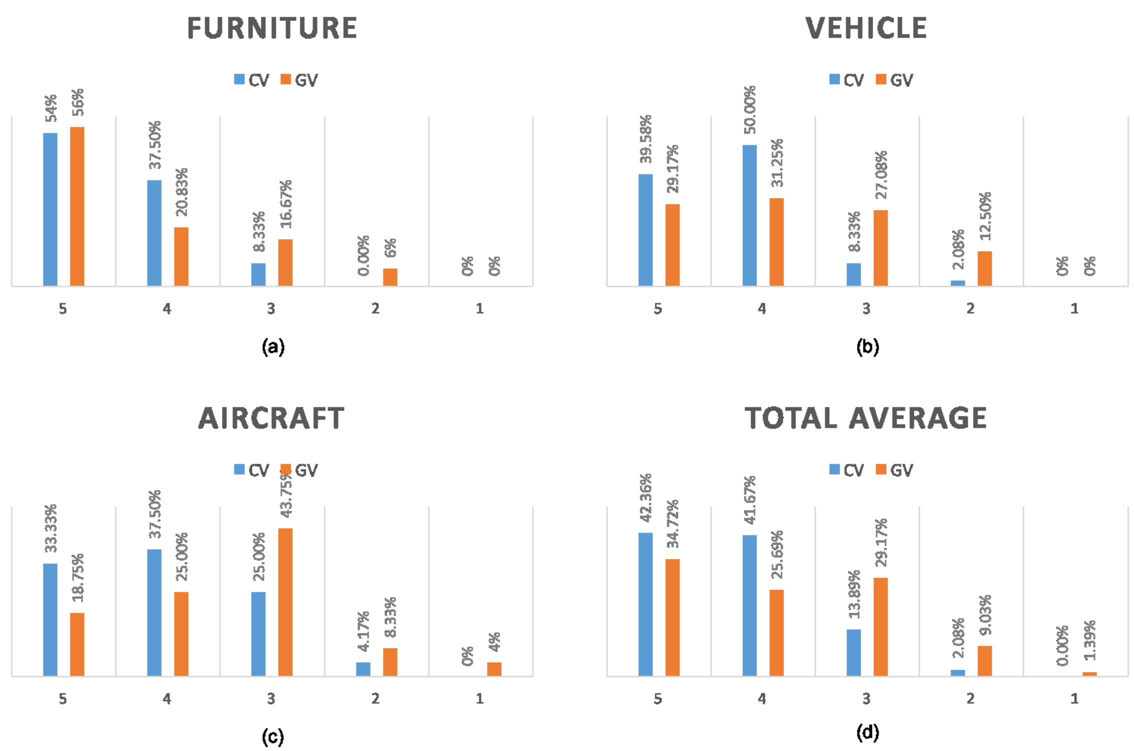


Figure 4.15: CV and GV comparison for question 5 of the first user study (“It was easy for me to identify the differences and similarities between the target and other objects. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

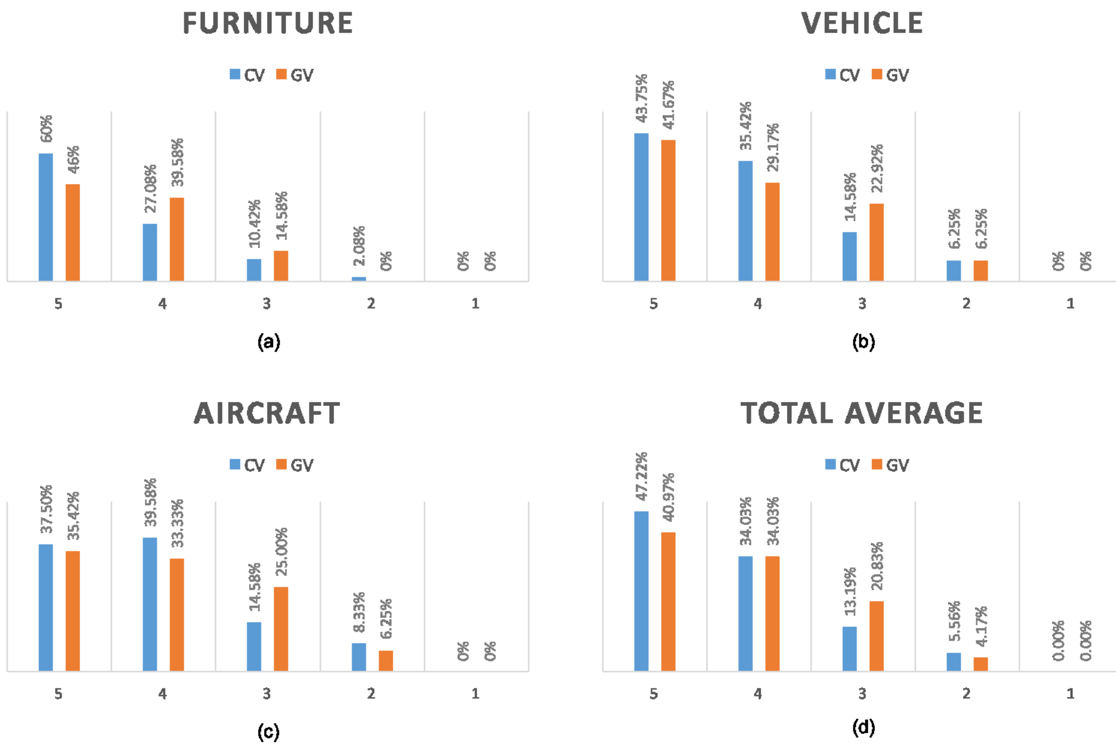


Figure 4.16: CV and GV comparison for question 6 of the first user study (“The layout is ‘visually pleasing’. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

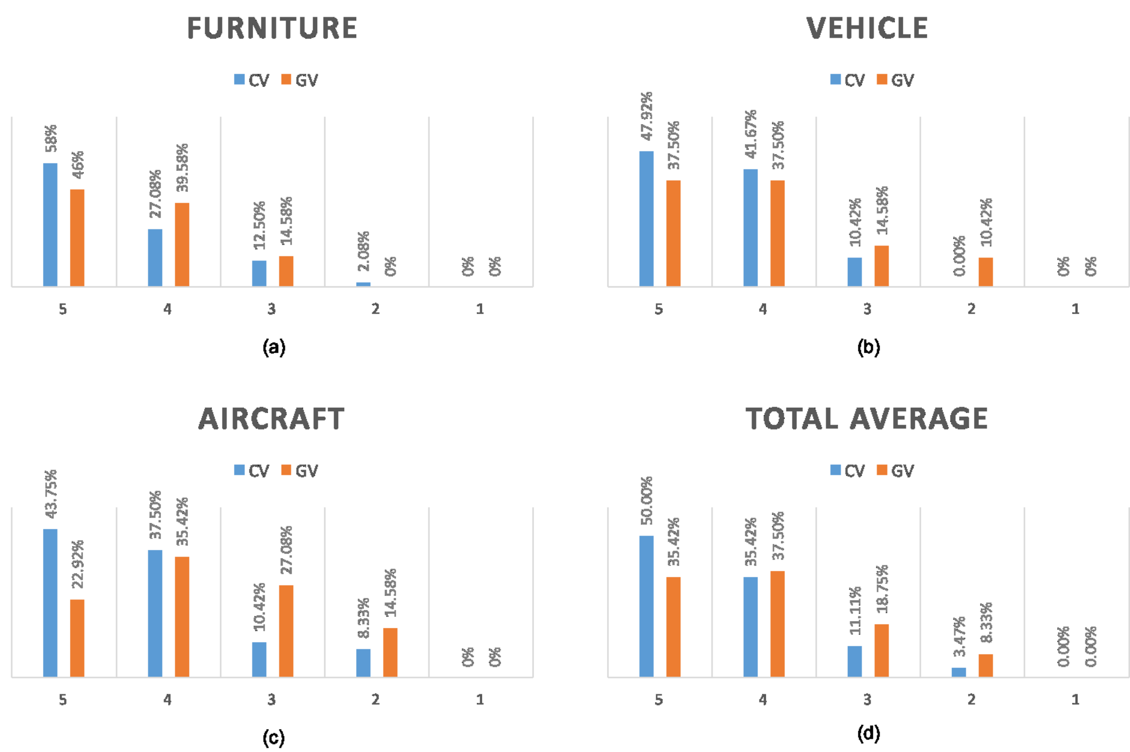


Figure 4.17: CV and GV comparison for question 7 of the first user study (“The orientations (viewpoints) of the objects were informative. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

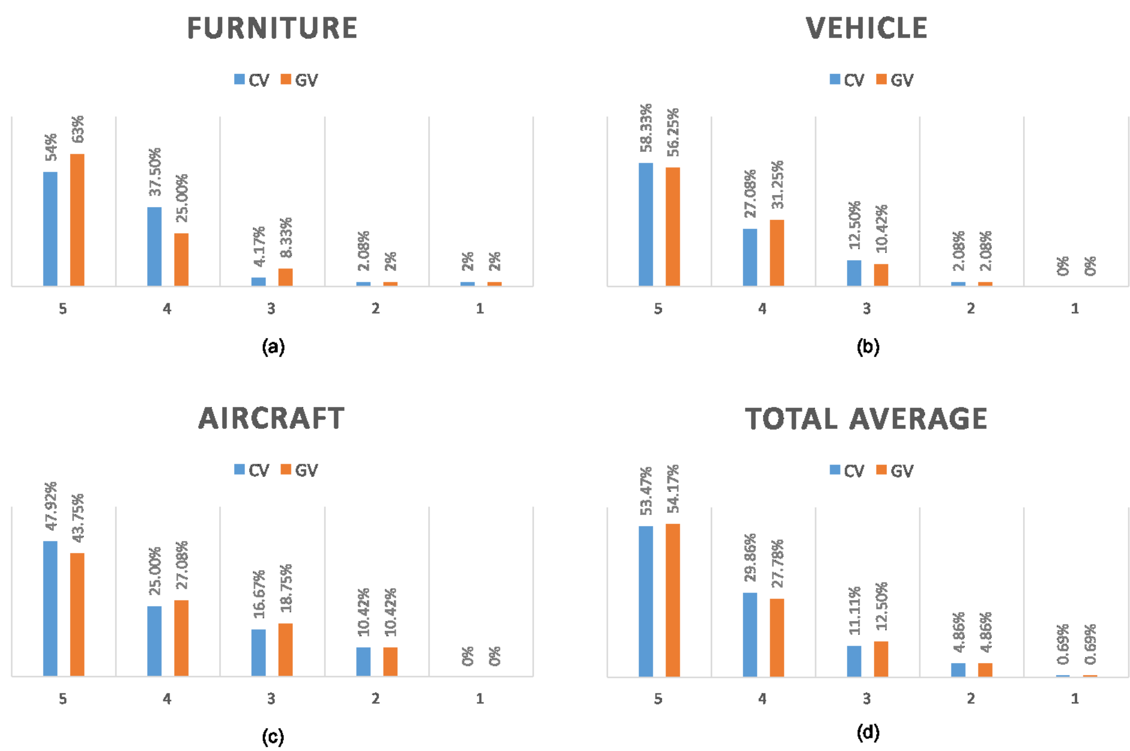


Figure 4.18: CV and GV comparison for question 8 of the first user study (“The arrangement of the 3D objects was understandable. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

After running T-test for all the above cases shown in Figure 4.11 to Figure 4.18, the P-values for questions 1 to 8 on three different datasets and the average value are listed as follows (the P-values are presented with the order of Furniture, Vehicle, Aircraft, and Average per row):

- Q1: P1-1 = 0.453; P1-2 = 0.298; P1-3 = 0.535; P1-4 = 0.329
- Q2: P2-1 = 0.578; P2-2 = 0.331; P2-3 = 0.753; P2-4 = 0.704
- Q3: P3-1 = 0.801; P3-2 = 0.449; P3-3 = 0.136; P3-4 = 0.269
- Q4: P4-1 = 0.812; P4-2 = 0.725; P4-3 = 0.404; P4-4 = 0.675
- Q5: P5-1 = 0.266; **P5-2 = 0.006; P5-3 = 0.007; P5-4 = 0.006**
- Q6: P6-1 = 0.340; P6-2 = 0.585; P6-3 = 0.663; P6-4 = 0.466
- Q7: P7-1 = 0.502; **P7-2 = 0.041; P7-3 = 0.013; P7-4 = 0.023**

- Q8: P8-1 = 0.815; P8-2 = 1.000; P8-3 = 0.768; P8-4 = 0.963

The analyzing result indicates that there are statistically significant differences for question 5 and question 7 on the vehicle, the aircraft datasets, and the average score of all three datasets according to the above values of P5-2, P5-4, P5-4, P7-2, P7-3, P7-4. The mean score of these two questions over two browsing techniques are:

$$M5_{CV_{vehicle}} = 4.27, M5_{GV_{vehicle}} = 3.77, M5_{cv_{average}} = 4.24,$$

$$M5_{GV_{average}} = 3.83, M7_{cv_{vehicle}} = 4.37, M7_{GV_{vehicle}} = 4.02,$$

$$M7_{cv\_aircraft} = 4.17, M7_{GV\_aircraft} = 3.67, M7_{CV\_average} = 4.32, M7_{GV\_average} = 4.00$$

This indicates that the two different browsing techniques have significant influences on the aspects “identify the differences and similarities between the target and other objects” and “The orientations (viewpoints) of the objects were informative” on the more complex datasets, and the CV technique performed better in these aspects than the GV technique in these cases. Combined with the participants’ suggestions from the interview section (section 5.4.6), the reasons causing the difference in question 5 may be that:

- The CV technique gathers similar objects within one screen and the most related objects are shown in similar sizes and colors. These features assist participants to narrow down their searching focus to the similar ones and easily compare them with the potential targets. This could be difficult in the GV technique because similar ones may be not on the same screen as the potential targets which may make it hard to make comparisons and require more interaction to ensure the potential one is correct;
- It’s also been suggested that the sizes of objects in the CV technique are relatively smaller than the ones in GV technique. Although the zoom in/out function has been provided, it’s still harder to observe the objects in a comfortable size than in the GV technique. This disadvantage may have been overcome by the advantage noted in i).

We also suggest that the reason for the difference in question 7 is that although the initial viewpoint is not ideally canonical in both techniques, it’s easier to observe the different viewpoints in CV technique since all objects can be rotated together, rather than only rotating one object once in GV technique. On the other hand, there is no significant difference for the furniture dataset, which supports the above observations in section 4.5.1 to 4.5.3 that there is barely a significant influence of different techniques on the furniture dataset.

While no significant results were found in the remaining questions (1, 2, 3, 4, 6, 8) the graphs still reveal some interesting patterns. According to Figure 4.11 (d), about 87% participants think the controlling interface is easy to understand in technique GV, where about 83% think it's easy in technique CV. Although the percentage of CV is less than GV, it's still at a similar level. As expected, we think the controlling interface of CV will be harder to understand because it's novel, but it should still be acceptable for participants since we removed some additional features (described in section 3.7). The results in the graph support our hypothesis. The Figure 4.12 (d) and Figure 4.13 (d) reveals that in general more participants think it's easier to finish the task and find the target in the CV technique with percentage 78.46% vs. 73.61% and 79.87% vs. 67.37%. Also, the CV technique was rated better when it comes to "visually pleasing" with a higher preferred percentage (as shown in Figure 4.16). This suggests that participants visually favored the layout of CV technique. The Figure 4.18 (d) shows that there are comparable percentage (about 82%) of participants of both techniques think the arrangement of 3D objects was understandable. Since the GV technique is a traditional browsing layout, we speculate that the arrangements of objects in the CV technique did not cause too great of a learning burden or mislead the participant even though it is a novel layout.

To analyze if there is significant connection between the participants' spatial abilities and their preferences, we categorized the participants based on their scores of the spatial ability test questionnaire as demonstrated in Table 4.8.

Table 4.8: The categories of participants' spatial abilities

	Low (<= 10)	Moderate (11 to 14)	High (>= 15)
Number of Participant	7	21	20

We ran a two-way repeated-measures ANOVA over the participants' spatial abilities and their post-task questionnaires' average scores on every question. Table 4.9 shows the descriptive statistical results of these two independent variables' interaction. The P-values for the interactions between the levels of spatial ability and the different techniques over the post-task question 1 to 8 are namely: P1 = 0.186, P2 = 0.450, P3 = 0.176, **P4 = 0.002**, P5 = 0.298, **P6 = 0.038**, P7 = 0.162, P8 = 0.709. This indicates that there is a significant difference for these two variables' interaction on the question 4 and 6 with P-values less than 0.05.

Table 4.9: Descriptive statistics of mean scores on post-task questions 1 to 8 combined with the spatial ability and techniques

Spatial Ability	Technique	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Low	CV	4.095	3.857	3.952	4.048	3.952	4.476	4.190	4.429
	GV	4.333	4.000	3.857	4.095	3.905	4.429	4.143	4.476
Moderate	CV	3.714	3.905	3.810	3.714	4.286	4.095	4.000	3.476
	GV	4.333	3.857	3.857	4.381	3.714	3.905	3.714	3.714
High	CV	4.429	4.524	4.381	4.524	4.429	4.286	4.095	4.381
	GV	4.286	4.000	3.429	3.524	3.429	3.429	3.381	4.381

To observe the interaction trends more clearly, we plot the estimated marginal means of question 4's and question 6's scores over the two variable's interactions in Figure 4.19 and Figure 4.20. Figure 43 reveals that the participants with moderate spatial ability favored the GV technique on the question "Easy to browse the models", while the participants with high spatial ability preferred the CV technique. In Figure 4.20, in the question "the layout is 'visually pleasing'", it can be observed that the preference on the CV technique is stronger as the participants' spatial ability levels increase. Also, the Partial Eta Squared (J, 1988) value of question 6 is 0.464 which means it has a large effect. In other words, we suggest that the CV provides a visually engaging layout for the potential users who have relatively strong spatial ability.

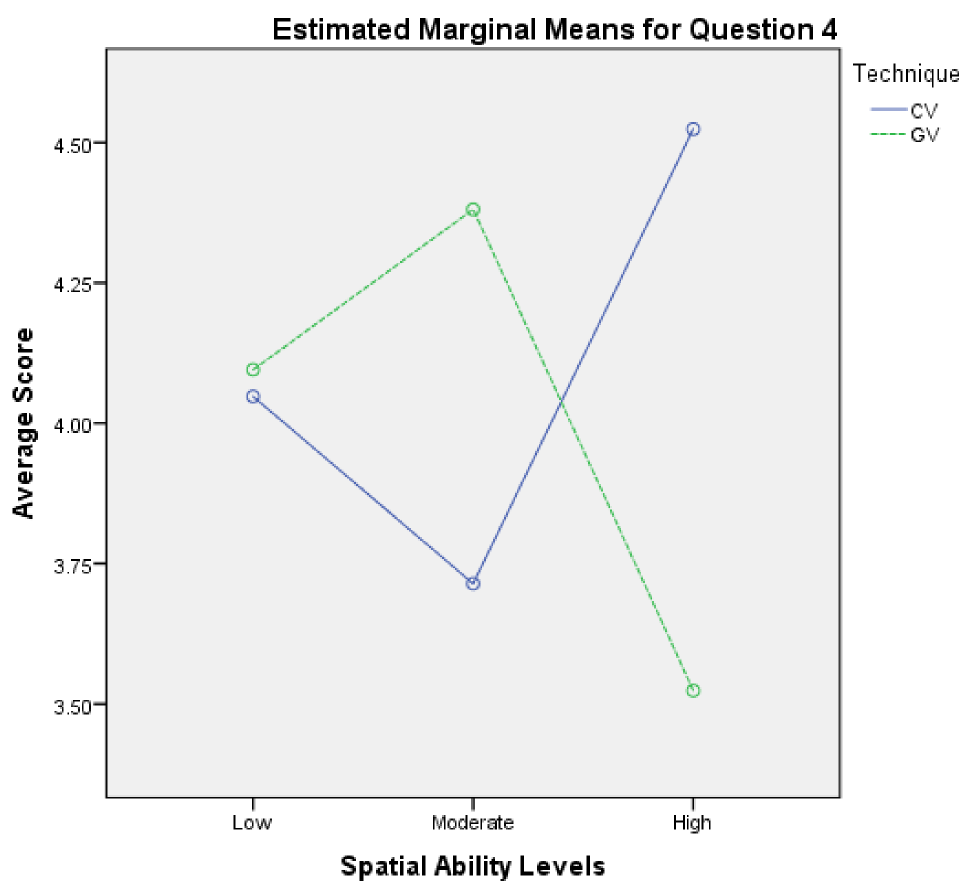


Figure 4.19: Interactions between techniques and type of datasets on mean score of for question 4 ("It was easy to browse the models")

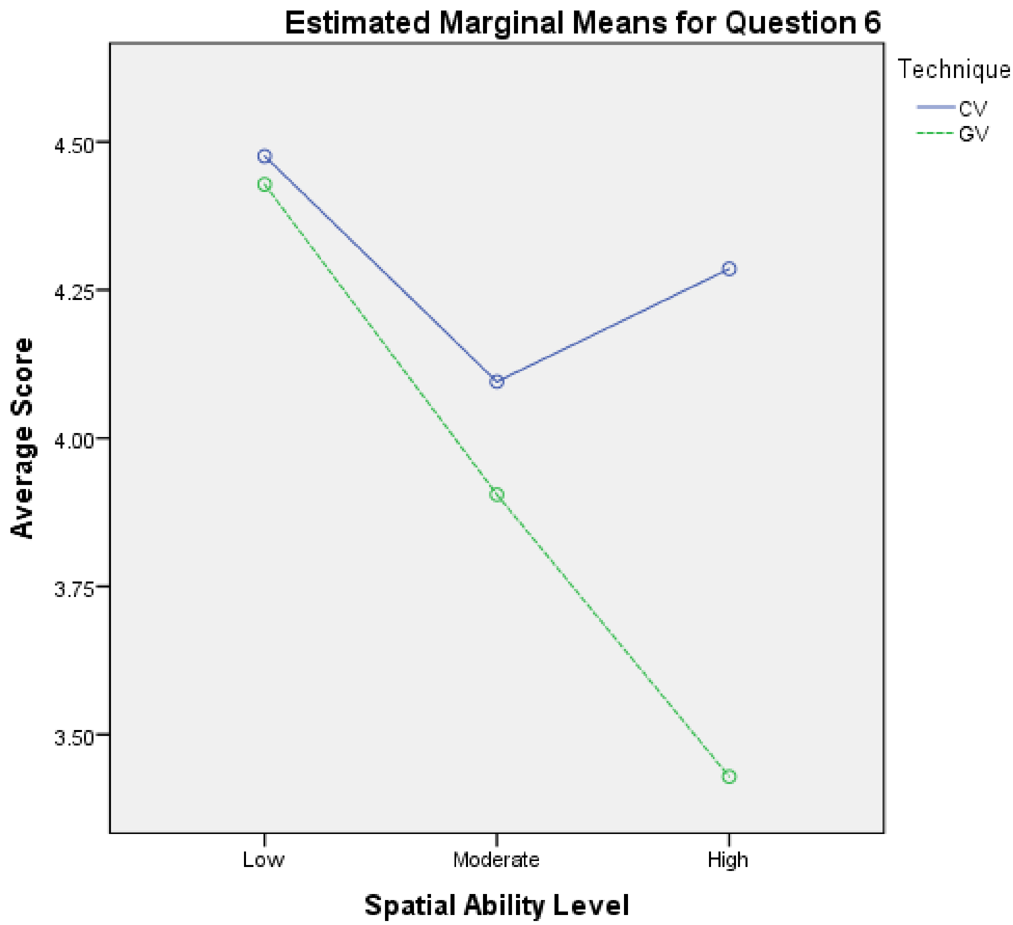


Figure 4.20: Interactions between techniques and type of datasets on mean score of for question 4 ("The layout is 'visually pleasing'")

4.5.5 Interview

We asked each participant to filled an interview questionnaire [see Appendix G1] at the end of session and analyzed the results based on six key questions.

When asked question about "Which system you prefer to browse the 3D-object dataset? Why?", 56.25% of participants said that they would prefer CV compare to 27.08% of participants who prefers GV, and the remaining participants (16.67%) preferred different techniques depending on the situation. Here we present some responses from the participants who prefer CV: Participant 16 said, "Cloud view because of how you see all the similar objects together, also for the feature of making the ones less like is smaller and fainter"; Participant 19 said, "Overall, I preferred the cloud view over the grid view because I was able to select a model I thought was very similar to the target, make it the center of the cluster, and then refine my search by looking at the models that were clustered around the center"; Participant 31 suggested "Cloud View. Once you find a unique characteristic of the target, cloud view helps find the target fast by comparing the similar objects"; Participant 35 said "Cloud view, I have more fun in Cloud view then the Grid view!!!!"; Participant 45 said "The Cloud View of course. Easier to find targets, finding more similar results to decide which is the true one and enjoying the surprise". Then, we also present some answers from the participants who prefer GV: Participant 18 said, "Grid view because it was much simpler to look and compare the different objects because they were all present"; Participant 24 said. "Grid view. Even though it takes time it looks more organized and I don't have to worry about selecting a wrong object by mistake and messing up the whole layout like I have to in the cloud view"; Participant 30 said, "Due to the easier visual mechanism of grid view, I choose this option even at the cost of time taken in scrolling"; Last, some suggestions from the participants who has different preference under certain conditions are listed: Participant 26 said, "It would depend on what kind of 3D models I am browsing. If I am looking at fairly simple and distinct model I would choose cloud but for complex models with lot of details I would prefer grid view"; Participant 29 said, "With animals, furniture and cars I preferred the cloud view as it was easier to navigate and the differences between objects were significant. With aircraft, I preferred the grid view because it was easier to see the minor differences between the objects"; Participant 43 suggested, "If the sample size is small, I prefer to use grid view; If there are many graphs, I would prefer Cloud view. It is more efficient".

The participants also provided some valuable opinions on the question "Did you find the task easy in the Cloud View / Grid View approach? Why?", and we listed some of these opinions: Participant 4 suggested, "It depends upon the object you are looking up in these two views. the object with the limited detailing and corners like furniture was easy to find in both the views whereas object like plane were very difficult to find in both views. However, I prefer cloud view as it shows all object together and you can catch out the difference at single glimpse"; Participant 11 said, "The cloud view would be very easy when compared to the grid. Since the cloud view would sample multiple object into one search space the search operation is optimized"; Participant 27 said, "I found the Cloud View helpful in quickly finding similar objects; at times it was a little more difficult because it was hard to guess what you'd need to select to find something similar. (i.e. choosing a desk just to see something similar to a shelf, but not a shelf itself)"; Participant 46 said, "I think using Cloud View is easier to find the task, because when I found a similar object and I could choose this one, then the target would be around it, so it is a good way to find objects. On the contrary, when I used the Grid View to look for the target, I had to search the all objects in the screen on by one, which need more time to scan it".

About the question "What advantages and disadvantages does each technique have?", we demonstrated some answers from participants: Participant 6 said, "The Cloud View is very user friendly, and people can do their categories and figure out the target easily. The Grid view will always waste time to go through all the objects and then make the decision"; Participant 8 said, "The Cloud View has the advantage of making comparisons easier, considering that a large number of objects can be simultaneously manipulated. This view is disadvantageous in that it is inconsistent in functionality and hard to understand in its methodology of reorganization. The Grid View is advantageous in its simplicity and ease of understanding. It is disadvantageous in its inefficient usage of screen space, where objects need to be isolated to be manipulated (thus making comparisons trickier)"; Participant 29 said, "Cloud view takes less time to identify an object but it looks overcrowded sometimes. The grid view looked neat but required time to browse through the objects"; Participant 32 said: "Cloud View: Advantages: if find the similar, it helps find the right object fast; easy to see the difference between similar objects; rotation a group of similar object at the same time. Disadvantages: if find the wrong one, it takes more time to find the object; sometimes looks messy because overlay. Grid View: Ad: easy to find the right group of similar object| Dis: hard to see difference between similar objects"; Participant 38 said "The cloud view worked very well and was easy to just go step by step clicking things that were similar, they were grouped very well. The grid view did group together well for the most part and once you found the right area you were looking for you could just keep looking through there until you found it. There was also an advantage with the cloud view where you could see and move many objects at once making it easy to track down which one you were looking for".

For the question “What improvements can our Cloud View / Grid View technique benefit from in the future?”, the participants provided some interesting opinions for the CV technique: Participant 8 said, “The way in which the Cloud View reorganizes objects should be more understandable. To be more specific, I'd like to know how this mode compares different objects when reorganizing them. In knowing that, I'd have a better idea of which object to double tap. Better yet would be to have different ways to organize objects (e.g., by sharpness of angles, edge count, et cetera)”; Participant 27 suggested, “Having a toolbar for the cloud view would be useful: include buttons with functions like "Random Object" or "Dissimilar/not similar object" to get out of situations where you're looking at a bunch of similar objects and you want something completely different, instead of having to click through a few models”.

Even though we already have a quantitative measure of the task finishing time, in the audio-recorded interview session, we also asked participants “Which technique they think they finished tasks faster?” to get a better insight from the subjective view. 58.3% participants suggested that they thought the cloud view makes them finish the tasks more efficient, while 35.4% participants favor the grid view (the remaining participants think it depends on the specific tasks). Among those who prefer the CV technique, we listed some opinions here: Participant 8 said “I think it was cloud view, because in grid view sometimes switching objects you have to go up and down when you search the target.”; Participant 12 suggested “It was cloud view. Because I could see pretty much the whole dataset at once, and I could click on the model that I thought was very similar to the target, even if sometimes I was not sure. If it was very similar to the target, chances it will bring the target closer to me, and I was able to click on the target and find target quicker. With list view, you would have to all of the models in order to be sure you didn’t miss something.”; Participant 38 said “The cloud one absolutely. Because if you knew what you’re looking for, it would instantly show you the things very similar, and also it let you change the directions of multiple objects at the same time. You can very easily break it down what exactly you need.”. We also listed some opinions from the participants who thought the GV technique is more efficient: Participant 14 said “Honestly, initially, the cloud was something more interesting for me. But for some specific shapes, the size of the shapes become too small. And changing from one shape to another one was hard for me.

But, in terms of grid view, it was something good for me that I can see all the categories. Since I can choose each of them and rotate them fully, it helps me a lot. So, I will say it's the grid one."; Participant 27 suggested "Probably the grid. I just get it done faster, but I think did find things in the cloud view very quickly too. It might be actually faster in the practice, but as far as like whatever feels better to move through, the grid view obviously feels more organized".

In the middle of this experiment, we thought it would be helpful to get an initial insight about their engagement by asking "Which technique you think is more interesting". Among the later 24 participants, 23 of them suggested the CV technique is definitely more interesting, while only 1 participant thought it depends on the particular tasks. We listed some opinions here to present the reasons the 95.83% participants favor the CV technique: Participant 29 said "I'd say the cloud view is more interesting. … Because you don't have to scroll. When it's a really long list of objects, then you need to scroll a lot. But with cloud view, if you just find one thing that is similar, it brings all the other similar things to you. It's like Google search, when you type one word, they find most relevant things for you. So the cloud view is more interesting in that way."; Participant 35 suggested "Be honest, I think the cloud view is more interesting. Because the interaction, and also the way you find more similar objects by clicking one object."; Participant 38 said "The cloud one as well. I thought it looks a lot nicer than the grid view. And grid view worked, you can find what you were looking for, but it wasn't as effective as the cloud one did, the nice job."; Participant 44 said "I really like the cloud view. … I'm thinking the cloud view has a lot of really neat potential, as it could be useful for, say, a website, or when you browse something. I think the cloud view is definitely, its ability to group things that similar to it, it's definitely more attractive to me." On the other hand, there were some other participants who thought the CV was more interesting, but they were not sure about the reasons or they thought because it's a "new" technique.

4.5.6 Task Complexity Verification

Based on the pilot study's observation, we decided to use different groups of objects as targets within each group since this can mitigate the learning effects for our within-subject design experiment. Since we assumed that the different groups of objects have comparable complexity, we verified this assumption in this section.

First of all, when we designed tasks, we aimed to make the two different task groups have comparable similarities. To achieve this goal, we controlled the complexity between the task objects following these two principles: i) The targets are similar but not the same. For example, if the target in group1 is a square table, the corresponding target in group2 will be a rectangle table; ii) We made sure that the target will not be too easy to find in both techniques, especially in the CV system. For example, the target won't be the initial center object in the CV system.

Then, we applied a T-test on the three dependent factors over the different groups of objects within each dataset combined with different techniques. Table 4.10 to Table 4.12 demonstrates the descriptive statistical results of the task finishing time, the number of wrong choices, and the level of confidence over the different groups of objects within each dataset combined with the two techniques.

Table 4.10: The descriptive statistical results of the task finishing time over the different groups of objects within each dataset combined with the two techniques

Dataset	Technique	Group ID	Mean	Std. Deviation	N
Furniture	CV	1	50.6806	18.12317	24
		2	57.3474	18.04745	24
	GV	1	52.6205	25.44641	24
		2	45.6181	23.75702	24
Vehicle	CV	1	69.0417	54.28754	24
		2	71.8611	42.49518	24

	GV	1	92.5556	21.83343	24
		2	80.6389	17.14964	24
Aircraft	CV	1	128.1250	42.10806	24
		2	122.0319	35.93095	24
	GV	1	170.4176	70.43786	24
		2	155.6250	56.04831	24

In terms of the task finishing time, the P-values for each pair of groups within each dataset with different techniques are listed as the order in the Table 4.10: P1 = 0.208, P2 = 0.594, P3 = 0.666, P4 = 0.223, P5 = 0.297, P6 = 0.425. In addition, the corresponding P-values generated by Levene's test for Equality of variances are: P1 = 0.758, P2 = 0.829, P3 = 0.220, P4 = 0.184, P5 = 0.603, P6 = 0.315.

Table 4.11: The descriptive statistical results of the number of wrong choices over the different groups of objects within each dataset combined with two techniques

Dataset	Technique	Group ID	Mean	Std. Deviation	N
Furniture	CV	1	0.2917	0.55003	24
		2	0.2500	0.53161	24
	GV	1	0.5833	1.10007	24
		2	0.6667	0.91683	24
Vehicle	CV	1	0.3333	0.63702	24
		2	0.2500	0.53161	24
	GV	1	0.7083	0.75060	24
		2	0.6250	0.76967	24
Aircraft	CV	1	0.5000	0.78019	24
		2	0.4167	0.65386	24
	GV	1	1.0833	1.88626	24
		2	0.6667	1.12932	24

Considering the number of wrong choices, the P-values for each pair of groups within each dataset with different techniques are listed as the order in the Table 4.11: P1 = 0.791, P2 = 0.625, P3 = 0.690, P4 = 0.777, P5 = 0.706, P6 = 0.359. In addition, the corresponding P-values generated by Levene's test for Equality of variances are: P1 = 0.667, P2 = 0.326, P3 = 0.542, P4 = 0.799, P5 = 0.781, P6 = 0.314.

Table 4.12: The descriptive statistical results of the level of confidence over the different groups of objects within each dataset combined with the two techniques

Dataset	Technique	Group ID	Mean	Std. Deviation	N
Furniture	CV	1	4.3819	0.57991	24
		2	4.4444	0.61908	24
	GV	1	4.3750	0.53726	24
		2	4.5792	0.57045	24
Vehicle	CV	1	4.3508	0.49187	24
		2	4.4861	0.51056	24
	GV	1	4.3014	0.48344	24
		2	4.3646	0.68936	24
Aircraft	CV	1	4.0868	0.64853	24
		2	4.2049	0.72814	24
	GV	1	4.0913	0.77265	24
		2	4.2319	0.83985	24

As to the level of confidence, the P-values for each pair of groups within each dataset with different techniques are listed as the order in the Table 4.12: P1 = 0.720, P2 = 0.355, P3 = 0.556, P4 = 0.208, P5 = 0.715, P6 = 0.549. In addition, the corresponding P-values generated by Levene's test for Equality of variances are: P1 = 0.376, P2 = 0.179, P3 = 0.465, P4 = 0.577, P5 = 0.118, P6 = 0.379.

The above results indicate that there is no significant difference between the different groups of objects on all three dependent variables (the task finishing time, the number of wrong choices, and the level of confidence). Also, the standard deviations between different groups show no significant difference. Thus, statistically, we suggest that there is no significant influence of the different groups of objects. In other words, the two groups of different objects within each dataset statistically have comparable complexity.

In terms of the subjective view, we asked all the participants about "Do you think the tasks having the different objects are fair?" or "Do you think what cause your preference? The different tasks or the different techniques?". There was no participant that suggested that the tasks are unfair. Although some of them thought their preference depends on the different datasets, they still said that this preference is not related to find different objects in the same dataset. Hence, we suggest the two groups of different objects within each dataset subjectively have similar complexity. This result supports the conclusion that two different task groups have comparable complexity.

4.5.7 Discussions and Summary

In this chapter, we analyzed the study results of quantitative data: task finishing time, number of wrong choices, level of confidence, and method preference on questionnaire, and we summarized the main effects for each independent factor in the Table 4.13. We also presented subjective data in the interview session (section 4.5.5) with some interesting and valuable opinions.

Table 4.13: Summary table of main effects for each of independent factors (N.S. = Not Significant and S. = Significant)

Dependent Factors	Independent factors	
	Technique	Type of Datasets
Task Finishing Time	S.	S.
Number of Wrong Choices	S.	N.S.
Level of Confidence	N.S.	S.

The technique turned out to be a main effect on the task finishing time and number of wrong choices which are the two critical dependent factors in our design, and this result supports our conjecture that: The interactive 3D object cloud view technique will provide a more efficient approach with leading to less mistakes.

In terms of the task finishing time, because the type of datasets and its interaction with techniques both have significant influence, we suggest that the CV technique will generally provide a more efficient performance in more complex datasets which have many objects and the objects have relatively high similarity levels. Considering the number of wrong choices, the CV technique will lead to less wrong choices because of it provides an easier way to compare the similarities and differences between objects. Thus, the CV technique is shown to be a technique providing more accurate choices on mobile devices, and we suggest that less wrong choices will also lead to save time on finding objects. Although there is no significant difference between techniques on the level of confidence, the CV as a novel technique can provide a confidence level (approx. 4.3 out of 5) as good as the GV technique.

According to the subjective opinions from the participants described in section 4.5.5, the CV technique is also preferred by more participants due as it helps participants gather the similar objects together with different similarity levels distinguished by colors and sizes, and this assists participants to compare the objects easier and faster. Although it has been suggested that the sizes of objects in CV could be sometimes too small to observe details, the statistical results of number of wrong choices and level of confidence shows that this potential disadvantages did not mislead the participants' searching and browsing.

Overall, the CV technique meets our expectations throughout the evaluation in this experiment. In other words, our approach provides a more efficient way to search and browse the relatively complex 3D-object datasets, and leads to less wrong choices with providing an acceptable confidence level.

Chapter 5

User Study 2

We applied staggered animation to transitioning in our system as described in section 3.6. The main purpose for this is to assist the users with the coherence of the layouts since there are complicated transitions (including changing positions, scales, and viewpoints, etc.) involved when a new center is chosen. Staggered animation might be useful for an instructive introduction to the system and how it works. To see if we meet the expectations, the correctness rate of tracking objects when transitions happens are key factor we wish to evaluate in this experiment. Based on the previous work on the evaluation of the staggered animation, the accuracy metrics (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011) (Chevalier, Dragicevic, & Franconeri, 2014) (Heer & Robertson, 2007) and the error metrics (Heer & Robertson, 2007) (Chevalier, Dragicevic, & Franconeri, 2014) are suggested to be two criteria which provide good insights of the tracking correctness rate. Also, as described in the user study 1, we want to know if the participants have been misled from the subjective view by recording the confidence scores of decisions. Hence, we choose three objective dependent factors to be evaluated, namely: i) Accuracy metrics of choices; ii) Error metrics of choices; iii) Confidence level of making decisions.

We conducted a controlled user study comparing our staggered animation design with the non-staggered animation design. Here we only evaluate situations when the targets stay on the screen after transformations (including positions, scales, colors, and viewpoints). In other words, we did not evaluate situations such as the targets fading out after the transition.

The study involved 24 participants, and it compared our staggered animation design with the non-staggered animation design with the same parameters to evaluate which assists participants to maintain coherence better.

5.1 Specific Research Questions

As discussed in section 3.6, the purpose of the staggered animation design is to improve the coherence of the participants when the transitions occur to help them track the objects and understanding the movement. By conducting the study, we wish to evaluate the following research hypotheses:

- The staggered animation (SA) will lead to higher accuracy of tracking objects than the non-staggered animation (NSA) when the transition happens;
- The SA will lead to lower error rate than the NSA;
- The SA will provide higher confidence level of choices than the NSA.

5.2 Initial Study Design

The purpose of this study is to compare the effectiveness of two types of animation design for maintaining coherence in a 3D cloud interface. Like the first user study, we also want to see if there is influence from the different types of datasets. Thus, we categorized the measurements into two independent variables which are: i) Animation types; ii) Datasets types.

We designed a 2×3 within-subject study, with animation types (SA, NSA) and datasets types (furniture, vehicle, and aircraft). All the tasks are performed on a Windows Tablet. The participants perform the same tasks as described in section 4 in the CV technique with different animation types and we evaluate the score of "easy to track the transitions movement" in the questionnaire and the interview questions about their user experience.

5.3 Pilot Study

On the purpose of formulating and refining our formal user study, we conducted an informal pilot study with lab mates from the GEM Lab in Dalhousie University. We first conducted a pilot study involving 4 participants. Each participant performed all three tasks as shown in Table 4.1 by using the CV technique with different types of animation, so 6 tasks need to be performed in total. After the participants finished each pair of tasks with both types of animation, we ask them to rate the animation movement in the aspect of easy to track, and interviewed them about their preference and reasons.

In this pilot study, the participants could not notice the differences of the transition while they're mostly focusing on searching objects, even though they were told that they need to pay attention to the different types of transitions. Correspondingly, they provided similar scores for both types of animations, and suggested that they could only recall a vague memory of the differences.

Based on this observation, we decided to adjust our study design to only focus on the transition instead of evaluating the transition while simulating the actual usage process. Combined with the user study design in the previous related work (Chevalier, Dragicevic, & Franconeri, 2014) (Heer & Robertson, 2007) (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011), we designed our second version of study – the study's task now becomes tracking two target objects using two types of animations when the transition happens. According to the previous work, tracking multiple objects while in transition is more reasonable than only tracking one object because of these three reasons: i) Tracking one object is too easy to distinguish differences because participants tend to be always correct; ii) Tracking one object is too easy to develop any interesting strategy to track; iii) Tracking one object will lead participants staring at the target and ignore what happens in the whole layout. Thus, we decided to ask participants to track multiple objects in the experiment. Due to the complexity of 3D transitions, we tested the number of tracking targets of 2 and 3 in our pilot study to determine the number of targets for the formal study. In the adjusted task, two target objects were highlighted in the initial layouts for 8 seconds, and then the participants needed to find these two targets in the new layout after finishing the transitions. In addition, to allow the participants to develop the strategies, we decided to also highlight the next central object in a different color in the initial layout.

With these adjustments, we conducted the second pilot study of 8 participants to make sure the changes worked well. As described in the first user study, we used different targets within each dataset to avoid strong learning effects. Four of the participants tracked two targets in all tasks, while the other four participants tracked three. We recorded the information of participants' choices including the name of the chosen object, the position of the chosen object, and the confidence score of the decisions, and gave them a short interview about which type of animation is preferred, the user experience of using both animations, and their strategies of tracking objects. The remaining aspects of the study design stayed the same as described in the initial study design, and more detailed procedure and design are introduced in the following section.

The results of the second pilot study shows that tracking two objects is reasonable while tracking three objects is almost impossible to track. Also, the current values of parameters turned out to provide a reasonable animation design because the participants suggested the animation is trackable and understandable and they can notice the difference between the two types of animations. Furthermore, the participants who tracked two objects developed various and interesting strategies to track objects, and some of them also suggested that showing the next center will help them to understand the transitions of the whole layout better. By reviewing the participants' interviews, we noticed that the changes of form of the tasks have pros and cons. The advantages are that the participants can focus on the transition and notice the differences between two types of animations. Therefore, we can evaluate the effectiveness in both the quantitative and subjective measurements by recording the information of users' choices and analyzing their questionnaires and interviews rather than just evaluating a subjective view. The disadvantage is that users may just track targets rather than observe the transitions of the whole layouts. After considering these factors, we decided on a trade-off which keep the tasks focusing on the transition and tracking two targets.

5.4 Final Study Design

According to the results and observations in the above two pilot studies, we made some changes in our study design. These changes are mainly focused on allowing participants to focus on the transition stage by tracking two target objects rather than trying to rate the user experience of the transition through browsing and searching tasks. More details about the adjusted tasks are described in the sections below.

5.4.1 Study Task and Protocols

The purpose of this study is to compare the two types of animation design (SA vs. NSA) based on the accuracy, incorrect selection ratio, confidence level, and user experience of tracking the transitions in the CV technique. According to the study design described in the beginning of this section, each participant needs to conduct tasks with two kinds of animations combined with three types of datasets. In each type of dataset using either animation type, the participant needs to conduct a task of tracking two targets during the transitions with three repetitions. In other words, every participant finished 18 tasks whose targets are varied in every task.

In each task, at the beginning, the screen starts masked and the participants can double click the screen to remove the mask and start the task whenever they are ready. Then, two objects will be highlighted in red in the initial layout, while the next central object will be highlighted in green. After 8 seconds, these highlighted objects will return to the original color, and then they will shake, which is the sign of starting the transition. Once the transition has finished, a "transition finished" notice will be temporally shown on the screen. Finally, the participants will need to find the two previously highlighted red objects. The different procedures of SA and NSA in this experiment is demonstrated in Figure 5.1. We also made both starting and ending layouts in every task have enough similar objects with the two targets to eliminate the situations that the participants find the targets by only memorizing their geometric features.

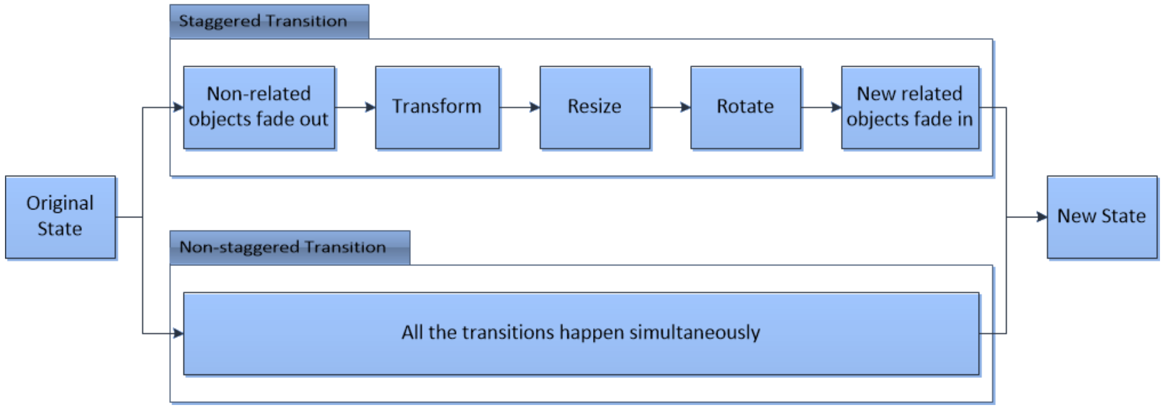


Figure 5.1: The procedures of SA and NSA transitions in the experiment

The general procedure of conducting the experiment is the same as described in the first user study (section 4.4.1). Hence, the participants firstly need to fill out two pre-task questionnaires [see Appendix E and I] after given the introduction and sign the informed consent form [see Appendix D2]. Then, they need to conduct three groups (three tasks in each group) of tasks with every dataset using both kinds of animations, and fill out a post-task questionnaire [see Appendix F2] after finishing each group of tasks to compare SA and NSA. Finally, they will fill an interview questionnaire [see Appendix G2] followed by a short interview.

5.4.2 Participant Recruitment

After obtaining the research ethics board approval letter [see Appendix J], we recruited 24 participants from the Dalhousie University campus through different recruitment platforms, including Notice Digest, the Computer Science Mailing List, the Dal Student Email lists and physical bulletin boards. The email announcement and the poster recruitment are given in Appendix A2 and Appendix B2 respectively. When there is a potential participant who replied to our notice/poster, we sent a screening email [see Appendix C] to ensure that the potential participants meet the inclusion criteria (not color blind, English fluency, and experience with touchscreen devices) and are aware of the important aspects (audio and video recording). Finally, participants need to be aware that some personal data will be collected (detailed in section 5.4.4), so they can be involved in the experiments without misunderstanding. We also made sure that the participants involved in this study were not the same group with the first study.

5.4.3 Counterbalance Measure

To counterbalance the conditions in the experiment, we made the full permutations of the independent factors: 2 kinds of animations and 3 types of datasets. Hence, we need $2! \times 3! = 12$ participants to fully conduct all cases, and we repeat every case twice which leads to $12 \times 2 = 24$ participants. For each combination of animations and datasets, we have 3 repetitions (similar tasks) which make $24 \times 3 = 72$ $cases$ in total. The full task arrangement for this experiment is shown in Table 5.1. The meaning of notations in the table stay the same as explained in section 4.4.3, the only difference is that the animation order means the order of animation types used to perform each group of tasks

Table 5.1: The task arrangement of the second user study

Participant ID	Animation Order	Dataset Order
P1, P13	1. SA 2. NSA	V, F, A
P2, P14		V, A, F
P3, P15		F, V, A
P4, P16		F, A, V
P5, P17		A, V, F
P6, P18		A, F, V
P7, P19	1. NSA 2. SA	V, F, A
P8, P20		V, A, F
P9, P21		F, V, A
P10, P22		F, A, V
P11, P23		A, V, F
P12, P24		A, F, V

5.4.4 Data Collection

As we described at the beginning of this section, we have 3 dependent factors that need to be evaluated, namely: i) Accuracy metrics of choices; ii) Error metrics of choices; iii) Confidence level of making decisions.

According to this purpose, we recorded the names and positions of objects that participants chose to calculate the accuracy metric and error metric respectively. The detailed definitions of these two matrixes can be found in section 5.5.1 and 5.5.2. We applied the same approach described in section 4.4.4 to obtain the values of confidence. In addition, all the interactions made by participants are recorded in the log file, and we also video record the whole session to obtain the detailed knowledge for the potential analysis. After finishing every 6 tasks in each type of dataset using both animations, participants were given a 1-minute break and then filled out a post-task questionnaire [see Appendix F2] for providing subjective quantitative data. There were three post-task questionnaires in total per participant. In the end, we asked participants to fill out a post-session interview questionnaire [see Appendix G2] and answer some face-to-face interview questions which were audio recorded.

5.5 Study Results

As introduced above, we recorded the name and positions of each choice made by every participant, and the confidence level of every decision. Participants were also asked to fill out three post-task questionnaires after finishing each group of tasks and a post-session interview questionnaire after finishing all 18 tasks, and complete a short interview with the researcher. In summary, we have two independent variables and three dependent variables as follows:

- Independent variables
  - Animation types: SA (staggered animation) and NSA (non-staggered animation)
  - Datasets with three types: furniture (number of objects is 198), vehicle (number of objects is 103), and aircraft (number of objects is 216)
- Dependent variables
  - Accuracy
  - Error
  - Confidence level of decisions

5.5.1 Accuracy

The accuracy is defined as the number of participant's correct choice divided by the total number of targets. We defined the set of participants' selections as S, while the set of targets is defined as T. To make the accuracy rate meaningful, we enforce the size of S to be equal to the size of T by asking participants to choose as many objects as targets to finish the task.

As described in section 5.3 and 5.4, we have two independent factors: animation and datasets. We applied a two-way repeated-measures ANOVA to analyze the effects of these two factors and their interactions. Table 5.2 shows the descriptive statistical result of these two independent variables interaction. The sample size N is equal to 24 in all cases.

Table 5.2: Descriptive statistics of accuracy rate combined with animations and type of datasets

Type of Datasets	Animation	Mean	Std. Deviation
Furniture	SA	0.729	0.044
	NSA	0.674	0.042
Vehicle	SA	0.583	0.056
	NSA	0.687	0.045
Aircraft	SA	0.479	0.047
	NSA	0.500	0.040

Figure 5.2 illustrates the plot for two-way interaction between the animations and types of datasets on the estimated marginal means of accuracy rate. There was no significant influence found in the interaction of these two factors (F = 1.695, P-value = 0.196).

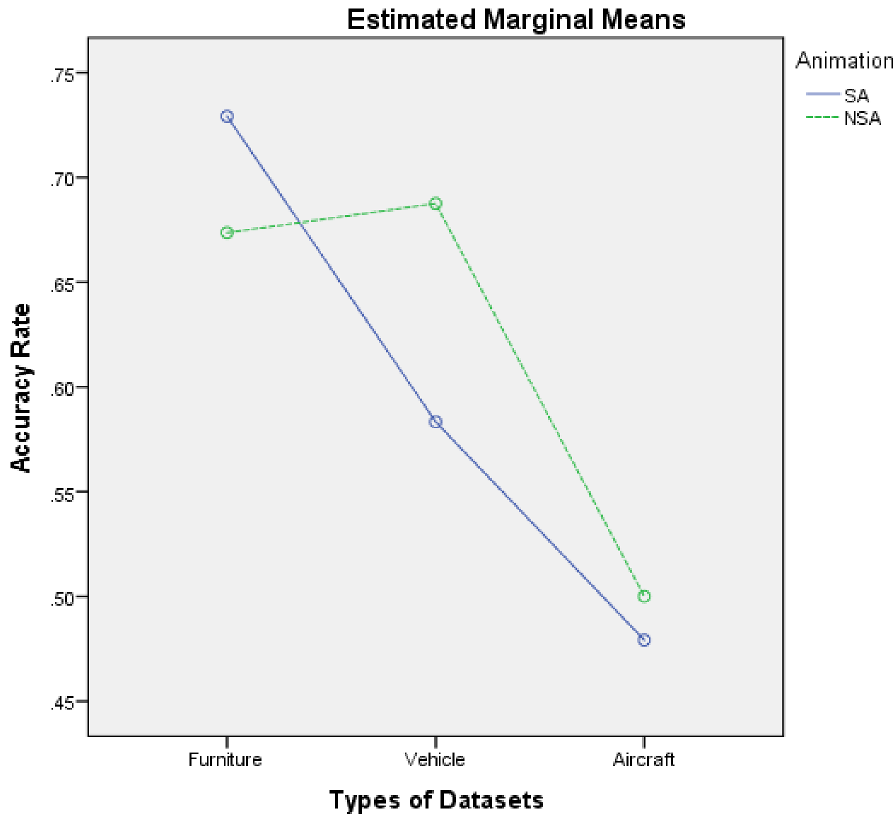


Figure 5.2: Interactions between animations and types of datasets on accuracy rate were found to be not significant

As we can see from the Figure 5.2, the accuracy rate of NSA is higher than the SA in both vehicle and aircraft dataset, while the SA has higher accuracy rate in the furniture dataset. The SA's accuracy rate is decreasing along with the similarity levels of dataset is increasing, while there is no such trend with the NSA. We speculated that the participants relied more on the geometry features of objects when using SA.

By analyzing the result with two-way ANOVA, we found out that there is no statistically significant difference on two kinds of animations (F = 0.666, P-value = 0.423) since the P-value is greater than 0.05. This reveals that the accuracy rate remains similar ($M_{SA} = 0.597,\ M_{NSA} = 0.620$) while using different types of transitions. On the other hand, it turns out there is significant difference between different types of datasets (F = 15.988, P-value < 0.001). Since the mean accuracy of datasets furniture, vehicle, and aircraft are

0.701, 0.635, and 0.490 respectively, we suggested that the accuracy rate will decrease when the similarity levels of datasets increases. Also, the Partial Eta Squared (J, 1988) value of the datasets is 0.410 which represents a large effect.

5.5.2 Error

The accuracy metrics provide a good and simple view of the correctness rate of participants' selections, but it does not reveal the distances between the selections and the targets. In our experiment, the selection which is far from the target should be penalized more than the one that is close to the target. Hence, we used the error metrics that evaluate the distance from the selections to the correct answers.

The error metric was first proposed by Dragcevic *et. al* (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011) for single object tracking, and expanded to evaluate multiple objects by Chevalier *et. al* (Chevalier, Dragicevic, & Franconeri, 2014). We applied the definition introduced by Chevalier *et. al* in our experiment to calculate the error metrics of tracking multiple objects. According to their work, the error between a selection set S and a target set T is defined as:

$$error(S,T) = \frac{\sum_i err(s_i, t_i^1)}{E(err(P,T))}, \qquad err(a,b) = ||a-b||$$

The numerator is the total distance between selected objects and targets, and it is calculated to as the optimal result which means it yields the lowest possible error. t_i^1 is the position of target i after finishing transitions, while s_i is the matching selection which has been matched optimally to get lowest error. As described in (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011), the denominator is a normalizer which represents the expected error that would have been measured with participants selecting random targets. Also, according to (Dragicevic, Benzerianos, Javed, Elmqvist, & Fekete, 2011), the value of $E(err(P,T))$ is equal to the average distance of all points from the target, so we estimated our denominator in every task using this definition to get the sum of average distances of all objects from the two target objects. Therefore, an average error of 1 means participants had a random guess of the answers, while a value greater than 1 means the participants have been misled by the animation.

Table 5.3 demonstrates the descriptive statistical results of the interaction of two independent factors (animation and dataset) on the error metric. The sample size N is still equal to 24 in all cases as well.

Table 5.3: Descriptive statistics of error combined with animations and type of datasets

Type of Datasets	Animation	Mean	Std. Deviation
Furniture	SA	0.256	0.047
	NSA	0.310	0.043
Vehicle	SA	0.316	0.043
	NSA	0.241	0.037
Aircraft	SA	0.421	0.038
	NSA	0.441	0.043

Figure 5.3 illustrates the plot for two-way interaction between the animations and types of datasets on the estimated marginal means of *error*. There was no significant influence found in the interaction of these two factors (F = 1.416, P-value = 0.254).

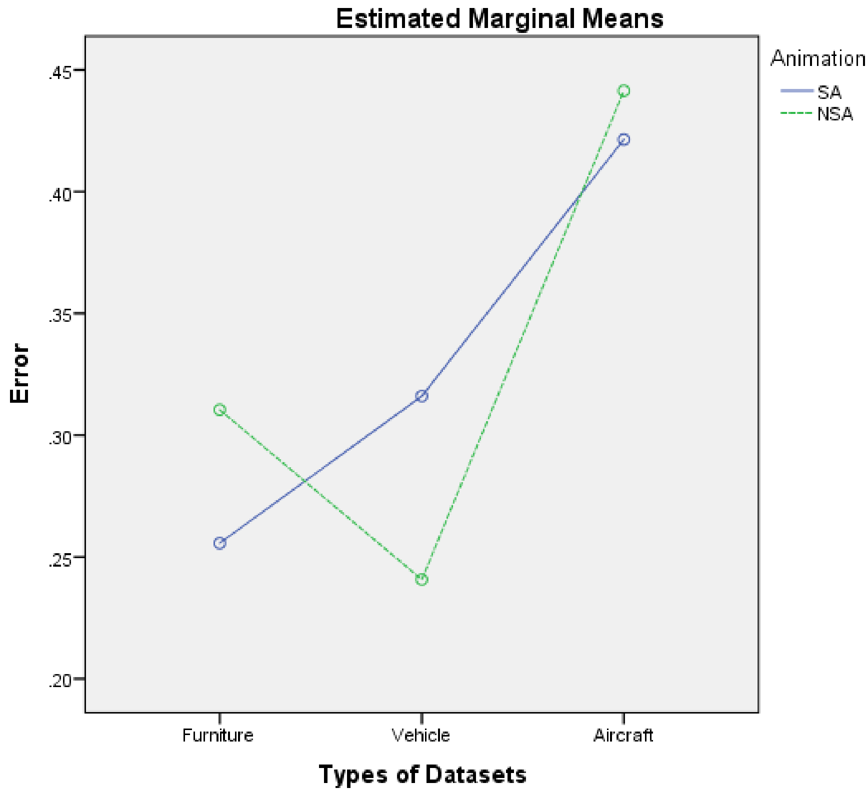


Figure 5.3: Interactions between animations and types of datasets on error *were found to be not significant*

Figure 5.3 shows that the *error* of SA is lower than the NSA in both the furniture and aircraft datasets. Like the trend in the accuracy rate, the SA's *error* rate increases as the similarity levels of dataset increases, while there is no such trend with the NSA.

To test if there is any factor that has significant influence on the *error*, we used the two-way ANOVA to analyze the results. It turned out the animations did not have a significant influence on the final result with F < 0.001, P-value = 0.994, and the two types of animations have the same average *error* which is 0.331. We suggest that two animations generally have no influence on the *error* of participant selections, but they both did provide some guidance to assist participants to track the objects with mean *error* less than 0.5. With regards to datasets, different types of datasets proved to have a statistically significant influence on the *error* metric (F = 10.338, P-value < 0.001) with the average *error* of three datasets $M_{furniture} = 0.283$, $M_{vehicle} = 0.278, M_{aircraft} = 0.431$. It reveals that the *error* is similar when the complexity of dataset is relatively low, while *error* will show a relatively high value with the datasets having high similarity level. Also, the Partial Eta Squared (J, 1988) value of the datasets is 0.310 which represents a large effect. Because of these observations on the datasets, we suggested that the participants are somehow relying on geometric features to track objects.

5.5.3 Level of Confidence

As described in section 4.5.3, the level of confidence is defined as the average confidence score of the participant's decisions (including both correct and wrong decisions) within a task. Every single confidence score is chosen by the participants after each decision by clicking on the score of confident-window (Figure 4.4) from 1 point for "not confident" to 5 points for "very confident". Table 5.4 demonstrates the descriptive statistical results of the interaction of two independent factors (animation and dataset) on the level of confidence. The sample size N is still equal to 24 in all cases as well.

Table 5.4: Descriptive statistics of level of confidence combined with animations and type of datasets

Type of Datasets	Animation	Mean	Std. Deviation
Furniture	SA	3.556	0.159
	NSA	3.576	0.190
Vehicle	SA	3.257	0.183
	NSA	3.417	0.153
Aircraft	SA	3.097	0.137
	NSA	3.278	0.150

Figure 5.4 illustrates the plot for two-way interaction between the animations and types of datasets on the estimated marginal means of confidence level. There was no significant

influence found in the interaction of these two factors (F = 0.537, P-value = 0.588), but the Partial Eta Squared (J, 1988) value is 0.023 which indicates it only has small effect.

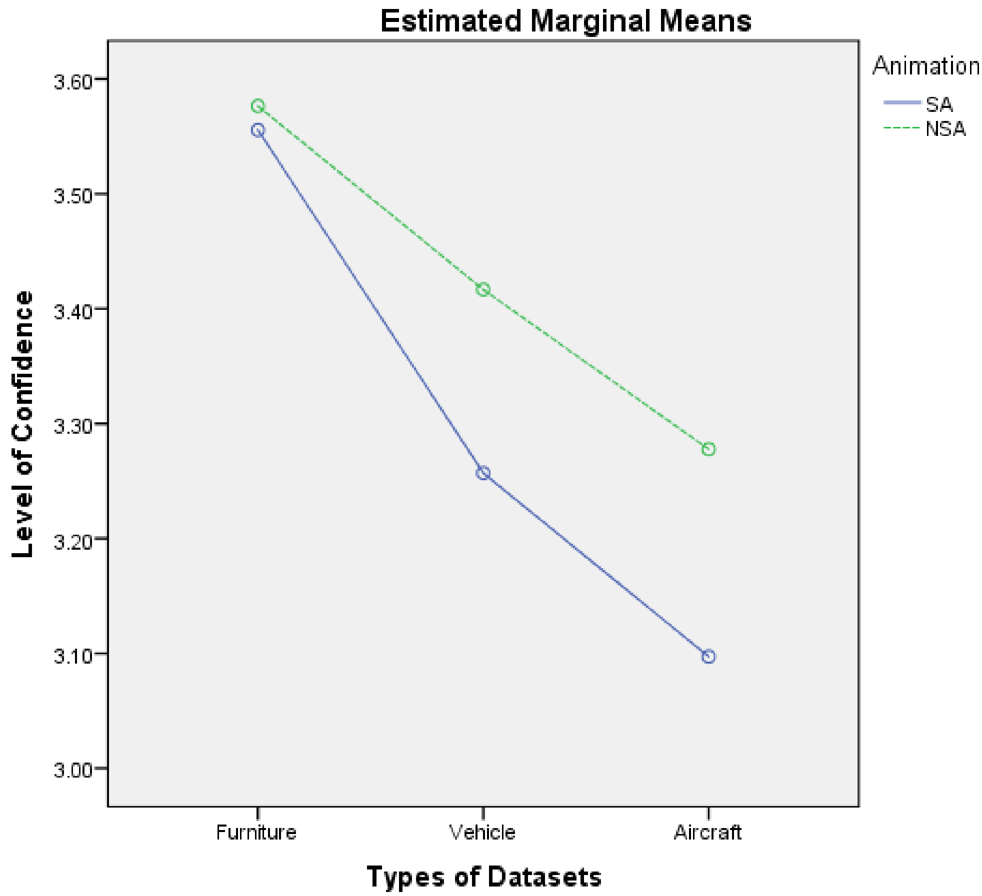


Figure 5.4: Interactions between animations and types of datasets on error were found to be not significant

As we can see from Figure 5.4, the confidence level of NSA is higher than SA over all three datasets. Since the interaction between two factors is not significant, we speculate that NSA provided higher confidence level than SA.

Applying a two-way ANOVA on the result, we found that both animations and datasets have a significant influence on the confidence level (F1 = 4.993, P1 = 0.036; F2 = 7.209, P2 = 0.003). The mean confidence level of SA and NSA are 3.303 and 3.424 respectively. Thus, the NSA can lead participants to gain more confidence on selections than SA, and this is also supported by the observations in Figure 5.4. Moreover, the Partial Eta Squared (J, 1988) value is 0.177 indicating it has large effect. The average confidence levels of three datasets furniture, vehicle, and aircraft are 3.566, 3.337, and 3.187. This shows that the higher similarity level of datasets can lead to lower confidence level while tracking objects. It also has large effects since the Partial Eta Squared (J, 1988) value is 0.239.

5.5.4 Questionnaire

As with the experimental procedure described in section 5.4.1, each participant needed to fill out 5 questionnaires in total, namely: a background questionnaire [see Appendix E], a spatial ability test questionnaire [see Appendix I], and three post-task comparison questionnaires [see Appendix F2]. We present the analysis on these questionnaire data in this section.

First of all, we show the background information of the 48 participants in this experiment. The experiment involved 15 males and 9 females from the Dalhousie University with the ages ranging from 20 to 53, with average of 30.3. The other results of the background questions are demonstrated in the Figure 5.5 (a) to (f).

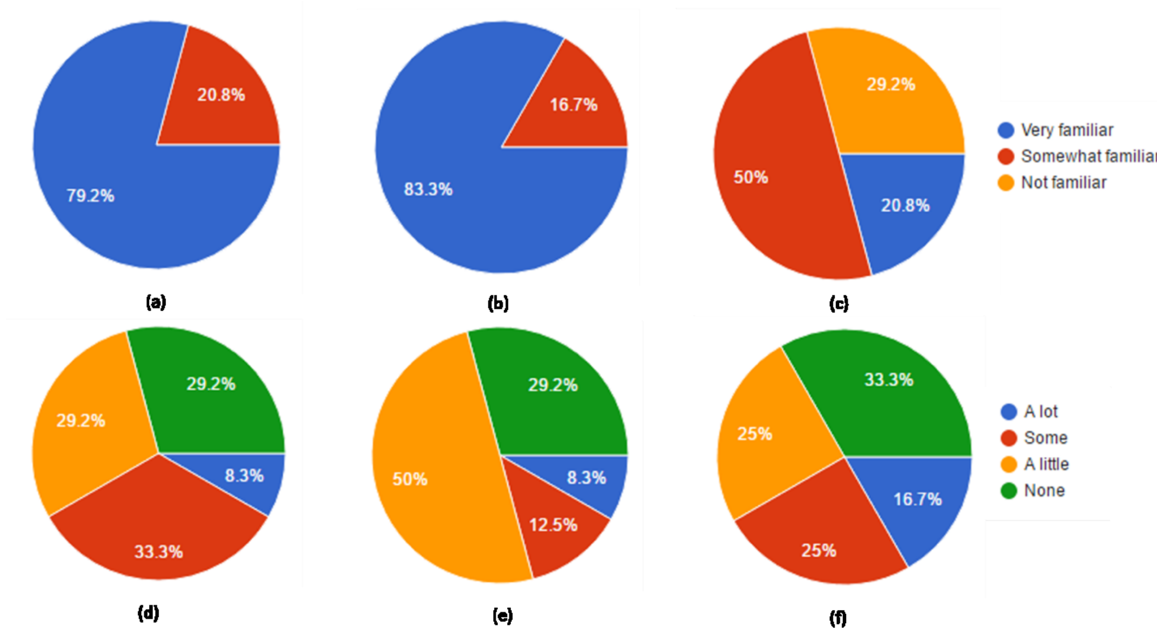


Figure 5.5: The pie diagram of the results of background questions:
(a) How familiar are you with mobile devices; (b) How familiar are you with touch-screen devices; (c) How familiar are you with 3D models; (d) Do you have experience on interacting with 3D models on a computer; (e) Do you have experience on interacting with 3D models on a tablet or other mobile devices; (f) Do you have experience with tag clouds.

We also list the results of different questions of post-task questionnaires in Figures 5.6 – 5.10.

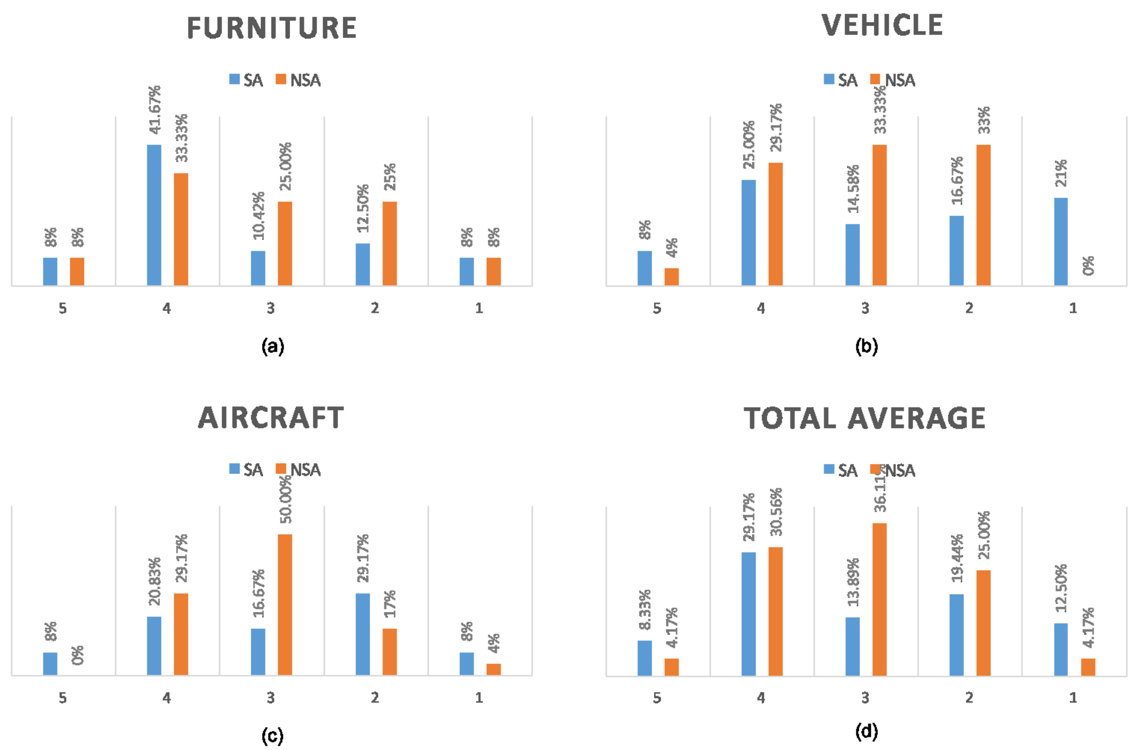


Figure 5.6: CV and GV comparison for question 1 of the second user study ("It was easy to finish the task without any difficulty. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree"): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

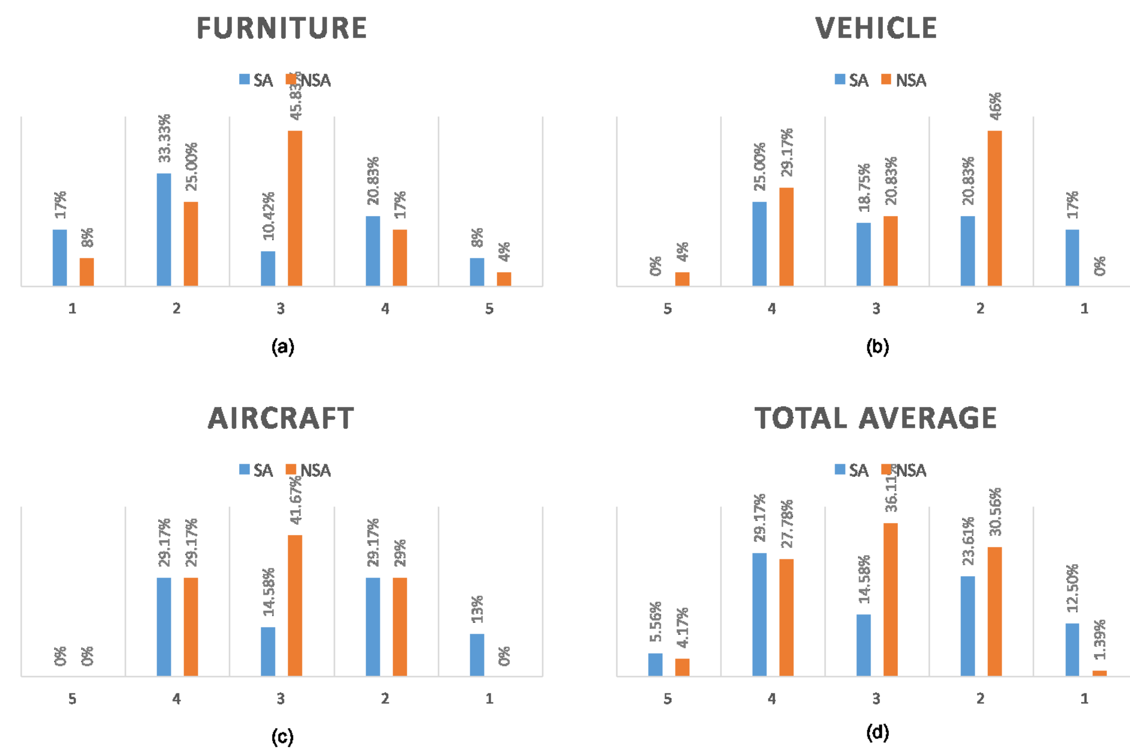


Figure 5.7: CV and GV comparison for question 2 of the second user study (“It was easy for me to track the targeted objects. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

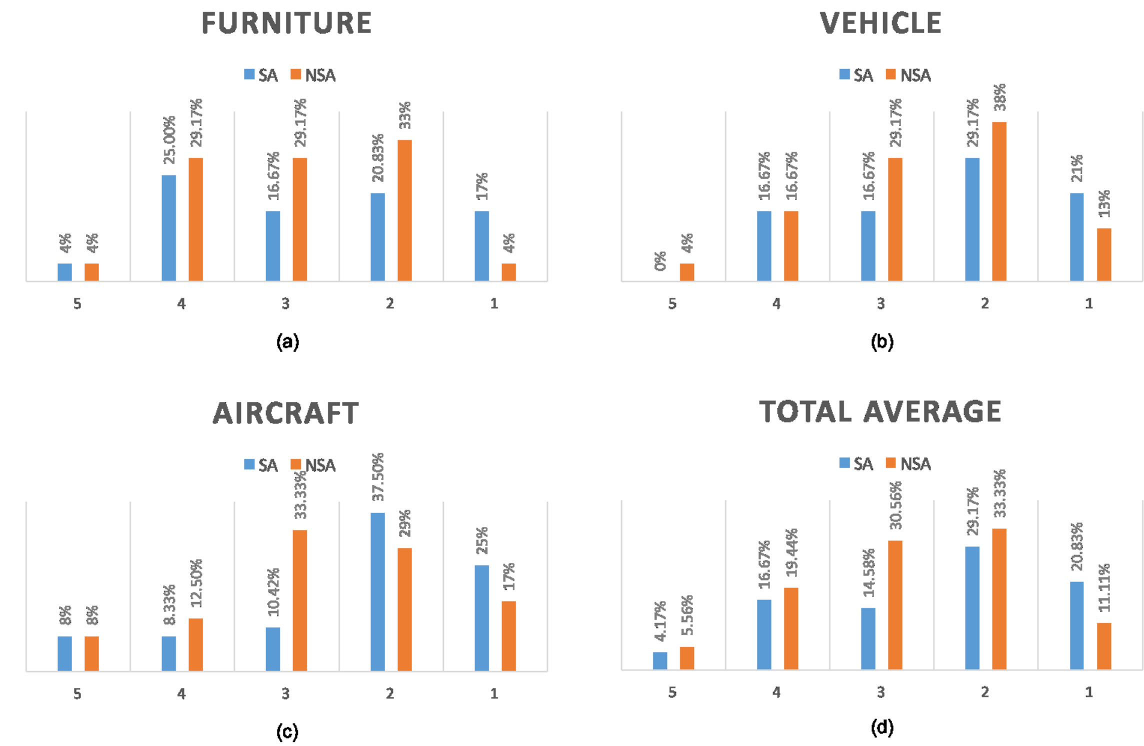


Figure 5.8: CV and GV comparison for question 3 of the second user study (“The movement of objects is easy to follow. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree”): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

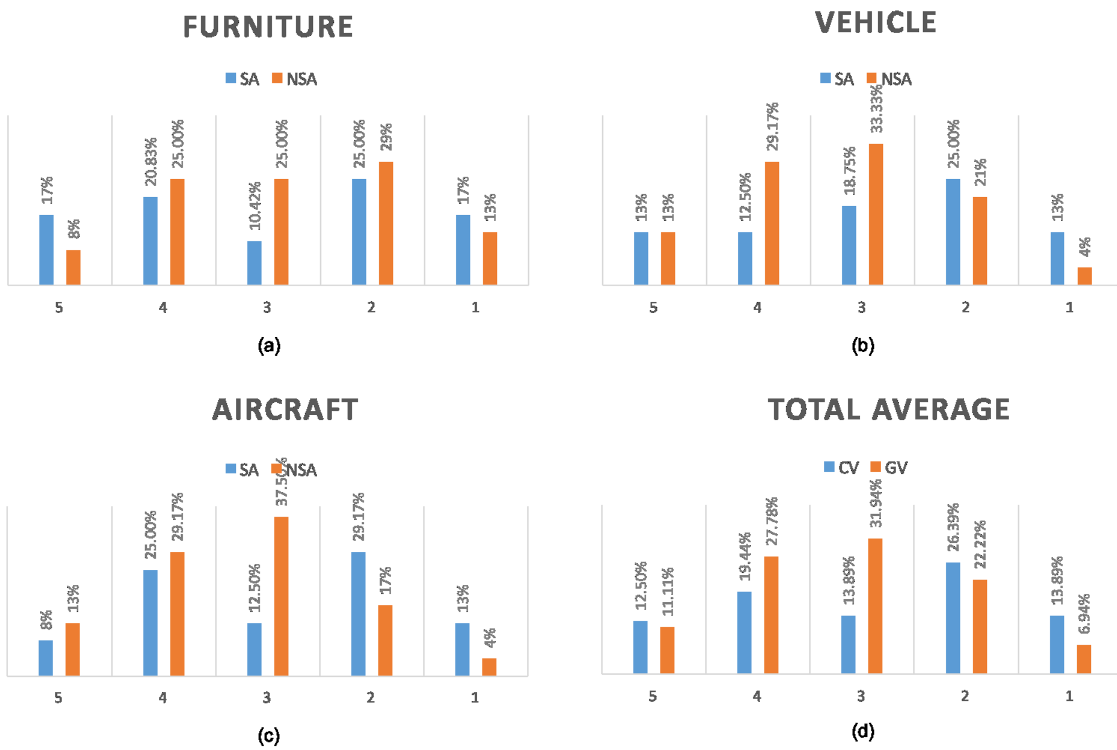


Figure 5.9: CV and GV comparison for question 4 of the second user study ("The movement of objects is understandable. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree"): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

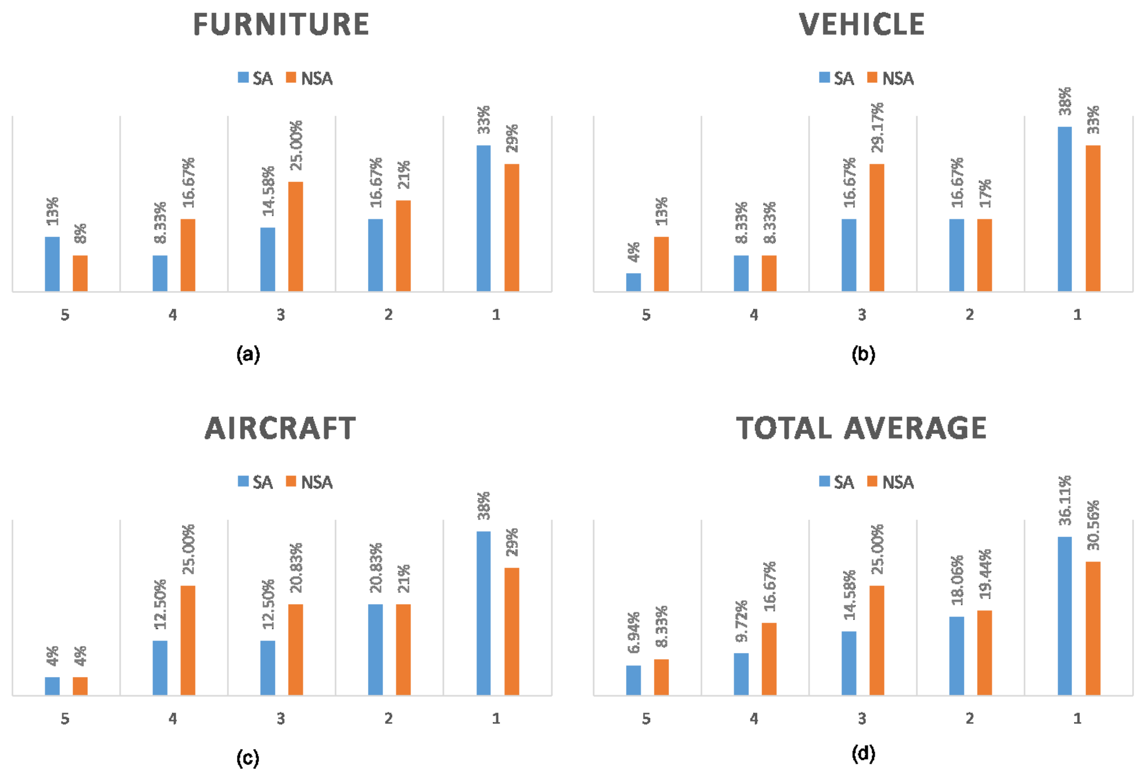


Figure 5.10: CV and GV comparison for question 5 of the second user study ("The movement of objects is predictable. – 1. Strongly Disagree; 2. Somewhat Disagree; 3. Neutral; 4. Somewhat Agree; 5. Strongly Agree"): (a) result of the furniture dataset; (b) result of the vehicle dataset; (c) result of the aircraft dataset; (d) the total mean result of all three datasets

After running T-tests for all the above cases shown in Figure 5.6 to Figure 5.10, the P-values for questions 1 to 5 on three different datasets and the average value are listed as follows (the P-values are presented with the order Furniture, Vehicle, Aircraft, and Average per row):

- Q1: P1-1 = 0.905; P1-2 = 0.518; P1-3 = 0.656; P1-4 = 0.701
- Q2: P2-1 = 0.697; P2-2 = 0.478; P2-3 = 0.349; P2-4 = 0.630
- Q3: P3-1 = 0.593; P3-2 = 0.581; P3-3 = 0.400; P3-4 = 0.410
- Q4: P4-1 = 0.823; P4-2 = 0.258; P4-3 = 0.204; P4-4 = 0.426
- Q5: P5-1 = 0.915; P5-2 = 0.505; P5-3 = 0.425; P5-4 = 0.551

The result shows that there is no significant difference between two different techniques. While no significant results were found, the graphs reveal some interesting patterns. Although the scores are similar on question 1 and question 2, the NSA presents higher score on question 3, 4, and 5. Specifically, Figure 5.8 (d) illustrates that about 25% of participants thought the movement of is easy to follow with NSA compared to about 20% of participants thought using SA makes the movement is easy to follow. Similarly, as shown in Figure 5.9 (d), fewer participants (31.94%) thought SA's movement is understandable than in the NSA's condition (38.89%). Furthermore, NSA has been rated higher (25%) than the SA (16.66%) over the question "The movement of objects is predictable".

Overall, we speculate that NSA is favored by participants for tracking objects. The reasons for the preference could be that although the SA separates the transitions into different stages to let participants focus on one kind of movement per stage, it also decreases the duration for each type of transition while NSA provides a longer duration for all transition types since they happen simultaneously. The durations of some stages maybe too short causing confusion in the animation movements which leads to more cognitive burden of participants when tracking object movements.

5.5.5 Interview

We interviewed each participant at the end of session and analyzed results based on 4 key questions. In addition to the quantitative data above, this section provides a good insights of the user experience from the qualitative subjective aspects.

When asked question "Which type of transitions you prefer? Why?", 20.8% of participants favored SA while 66.6% of participants favored NSA (the remaining participants have no preference). We presented some responses from participants who preferred SA: Participant 9 said, "I like staggered animation more than the other. Though it is hard to track, and may decrease the accuracy, it is more natural and fit my preference". Participant 22 said, "Staggered. Easier to track and I could focus on the effect of individual transformations on the appearance of the object". We also presented some responses of participants favored NSA: Participant 2 said, “The non-staggered animation because it is slower than the staggered one”. Participant 6 said “non-staggered animation. This type gives me more time to track the movement”. This reveals that SA has the main advantage of enabling users to focus on the individual transformation, but the main disadvantage is that it does not provide enough duration as the NSA does.

There are also some interesting responses from participants on the question “What kind of strategy you have developed to track/find the targets with the Staggered / Non-staggered animation?”, we presented them here: Participant 3 described, “1. got the two red objects then compare them and find which one is hard to recognize. 2. Focus on the one that is hard to see, and track it. 3. when you find it then find the another one”. Participant 13 said, “It is hard to focus entirely on both objects. My strategy was to focus more on one and try to follow the other one in my side vision”. Participant 14 said, “I was tracking the less similar one with both non-staggered and staggered tasks because the other red one would be near the center.” Participant 19 suggested, “Depending on what kind of targets and new center, if the targets are similar with the new center, I compare them to find the differences and predict the position of the target. If they are not similar, I try to follow one of the target's movement and watch the other target's by side vision, of course, they should be located in the far distance of the new center”. These comments suggest that it is hard to follow multiple objects in our system since the screen size is too wide for the users’ visual field but participants can develop strategies to track them. They usually tracked one of the two objects by focusing on its transitions, and applied their various strategies to another object to locate it.

The above question is followed by the question “When this strategy was challenged with the Staggered / Non-staggered animation?”, the participants provided interesting opinions to help us understand the difficulties can happen in both animation types. We list some of opinions here: Participant 11 said, “In the staggered I was able to see the properties before rotation, which was good. But there was problem during the rotation process because the size and the location of objects was changed, and it was difficult to follow”. Participant 14 said, “It is found that with the staggered task was more challenging because their movements are unexpected”. Participant 21 suggested, “I think non-staggered animation is more intuitive for me to trace. When huge rotation changes applied to staggered objects, I usually get lost”. The rotation issue has been stressed by many participants, so we argue that the duration for the rotation stage of SA should be longer since the viewpoint is a critical part of presenting 3D objects. We suggest that the transformation and rotation stages should be given enough time to let users to track the key features of the 3D objects to maintain coherence.

5.5.6 Discussions and Summary

In this chapter, we analyzed the study results of quantitative data: accuracy metric, error metric, level of confidence, and method preference in the questionnaire. We summarized the main effects for each independent factor in the Table 5.5. We also present subjective data in the interview session (section 5.5.5) with some interesting and valuable opinions.

Table 5.5: Summary table of main effects for each of independent factors (N.S. = Not Significant and S. = Significant)

Dependent Factors	Independent factors	
	Technique	Type of Datasets
Accuracy	N.S.	S.
Error	N.S.	S.
Level of Confidence	S.	S.

The types of dataset were a main factor as we can see from Table 5.5. For the level of confidence, the technique is also a main factor influencing how confident are the participants in their decisions. There was no significant interaction found on any dependent factor between two independent variables.

Overall, the types of datasets seemed to be the most influential factor. We speculated that this is due to the animation transitions not being good enough to assist the participants with tracking the objects, so that the participants had to rely on the objects’ geometric features. Although the technique seems to be the less influential one, it does have some influence on the confidence level, and also there are some interesting trends showed over other dependent aspects. The NSA provides a little higher score than the SA on both accuracy metrics and error metrics, and it even shows a significantly higher rating than SA on the level of confidence. This also validated some of the opinions given by the participants in the interview session that the NSA provides longer and slower transitions and seems to be a more natural approach.

There are many more parameters that might have an influence on the final result, including duration of the animation, rotation of each stage, the combination of stages, and the animation dwell time. In our case, we left the values of these parameters fixed. Further study will be needed to obtain a better insight into animation design for 3D objects.

Chapter 6

Conclusions

In this chapter, we summarize the thesis and highlight the contributions, limitations, and suggest potential directions for future work.

6.1 Approach Summary

We proposed a cloud-view based system where users can interactively browse and search 3D object datasets on mobile devices. Specifically, we expected that our system will benefit the query process by providing a more efficient and engaging presentation. Based on this goal, we designed and developed the interactive 3D object cloud system which provides a visually-engaging layout while avoiding occlusion and preserving underlying patterns. Additionally, for the purpose of providing a more convenient application, we implemented our prototype on mobile devices while pre-processing computationally intensive aspects.

6.2 Study Results Summary

By conducting user studies, we evaluated the above hypothesis and we highlight our contributions here.

The first user study compared our system to the commonly found grid-based browsing and searching interface. According to the results, our system is shown to have statistically significant performance gains for decreasing the searching time, especially when the 3D dataset tends to have high similarity levels. Our system also leads to fewer wrong choices, and we suggest this may be because it provides an easier way to compare the similarities and differences between objects. In other words, the CV technique provides more accurate choices on mobile devices, and we suggest that fewer wrong choices will also lead to saving time while finding objects. It also has been shown that our novel system does not mislead the subjects by having a comparative confidence score with the grid layout, and most subjects considered it to be a more interesting way to browse and search the 3D objects. Overall, the proposed approach met our expectations in the evaluation of this system, providing a more efficient way to search and browse relatively complex 3D-object datasets, and leads to fewer wrong choices while maintaining a good confidence level.

The second user study focused on examining the effectiveness of the staggered animation design in our system. In the experiment, we compared the staggered animation to the non-staggered animation by asking subjects to track multiple targets during the transition which happens when users choose a new central object. The results showed that the different types of transition seem to be less influential, but it does have significant influence on the confidence level, and also there are some interesting trends in the other dependent aspects. Most subjects suggested that non-staggered animation offered a more natural way to observe the transition. On the other hand, the similarity level of datasets seemed to be the most influenced factor. Generally, this experiment is just a preliminary evaluation since there are many parameters that might have an influence on the final result, including the duration of the animation, the combination of stages, and the animation dwell time. In our case, we left the values of these parameters fixed in our application. Further study will be needed to obtain a better insight into the animation design for 3D objects.

6.3 Limitations and Future Works

As summarized above, our approach meets our expectations by providing a more efficient way to search and browse the relatively complex 3D-object datasets, and leads to fewer wrong choices while providing a good confidence level. On the other hand, there are still some limitations we want to address.

In terms of system's functions, we suggest the following potential directions of the future work. First of all, though we provide an occlusion-free layout in the initial state, overlaps can happen during the free rotation phase. This could block the subjects' observations, slow down their searching, and potentially influence their emotions and strategies. Addressing this would be a necessary improvement for future work by adjusting the packing algorithm, such as involving the force-directed graph. Second, using shape as the only feature to organize the objects seem to be good enough as shown in our results, but we think more features may narrow down the search. On the other hand, it may also have potential issues for novice users and would require a careful design about how to organizing objects based on multiple features. Third, this prototype does not involve text searching. Jumping from one group to another very different group (e.g. change from the table group to the chair group) can be difficult and time-consuming. The improvements can be easily added, such as providing a feature generating 10 very different objects, showing the typical object in each group, or letting users search by the text, tag or sketch drawing (Halvey & Keane, 2007). Also, this advantage can impact the performance of our system when the approach scales to large datasets. Because when the dataset is very large and the desired target is very different from the current shown group, it might be time-consuming to find that target.

In terms of evaluation, these aspects might be interesting to look at. First, our first study only evaluates the comparison for the specific searching case (finding a given target), and so a study that evaluates general searching (e.g. find your favorite cars) performance remains to be conducted. Second, exactly which features of our layout design promotes efficiency are still not comprehensively known, and further studies may be needed to look into the specific influential factors. Third, the definition of a dataset's complexity mixes a set of potential factors including the similarity level within the dataset, how familiar is the subject with the dataset's object, and the number of objects within the datasets. Which factor is the most influential factor being unclear, further study to evaluate this part can be done in the future. Lastly, since our approach delivered a good result for mobile devices, whether it will also offer similar performance gains on desktops can be look at. We speculate that the CV technique can also benefit the 3D dataset browse and search on desktops with the proper design of the interactions.

Bibliography

Bateman, S., Gutwin, C., & Nacenta, M. (2008). Seeing Things in the Clouds: The Effect of Visual Features on Tag Cloud Selections. *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia. HT '08*, pp. 193-202. Pittsburgh, Pennsylvania, USA: ACM.

Blanz, V., Tarr, M. J., Bulthoff, H. H., & Vetter, T. (1995). What object attributes determine canonical views? *Perception, Suppl.*, 24:199c.

Chen, D.-Y., Tian, X.-P., Shen, Y.-T., & Ouhyoung, M. (2003). On Visual Similarity Based 3D Model Retrieval. *Computer Graphics Forum, 22*, Number 3.

Chevalier, F., Dragicevic, P., & Franconeri, S. (2014). The Not-so-Staggering Effect of Staggered Animated Transitions. *IEEE Transactions on Visualization and Computer Graphics*, 1077-2626.

Clark, J. (2008). *Clustered Word Clouds*. Retrieved 2016, from Clustered Word Clouds (2008): http://www.neoformix.com/2008/ClusteredWordClouds.html

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* USA: Psychology Press.

Dengsheng, Z., & Guojun Lu. (2002). An Integrated Approach to Shape Based Image Retrieval. *Proc. of 5th Asian Conference on Computer Vision (ACCV)*, (pp. 652-657). Melbourne, Australia.

Dragicevic, P., Benzerianos, A., Javed, W., Elmqvist, N., & Fekete, J.-D. (2011). Temporal Distortion for Animated Transitions. *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 209-2018). Vancouver,BC,Canada.

Dutagaci, H., Cheung, C. P., & Godil, A. (2010). A Benchmark for Best View Selection of 3D Objects. *3DOR '10 Proceedings of the ACM workshop on 3D object retrieval* (pp. 45-50). Firenze, Italy: ACM.

Feinberg, J. (2008, June 1). *Wordle - Beautiful Word Clouds*. Retrieved 2015, from Wordle - Beautiful Word Clouds: http://www.wordle.net/

Feinberg, J. (2010). Wordle. In J. Steele, & N. Iliinsky, *Beautiful Visualization: Looking at Data through the Eyes of Experts* (Vol. Chapter 3, pp. 37-58). Sebastopol: O'Reilly Media.

Funkhouser, T., Kazhdan, M., Shilane, P., Min, P., Kiefer, W., Tal, A., . . . Dobkin, D. (2004, Augest). Modeling by Example. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2004 , 23*(3), 652-663.

Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., & Dobkin, D. (2003, January). A Search Engine for 3D Models. *ACM Transactions on Graphics (TOG), 22*(1), 83-105.

G, V. S., & R, K. A. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills,*, 599 - 604.

Halvey, M., & Keane, M. T. (2007). An Assessment of Tag Presentation Techniques. *Proceedings of the 16th international conference on World Wide Web. WWW '07*, pp. 1313-1314 . Banff, Alberta, Canada: ACM.

Hausner, A. (2001). Simulating Decorative Mosaics. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (pp. 573-580). Los Angeles, CA, USA: ACM.

Heer, J., & Robertson, G. G. (2007, November). Animation Transition in Statistical Data Graphics. *IEEE Transactions on Visualization and Computer Graphics, 13*(6), 1240-1247.

Hiller, S., Hellwig, H., & Deussen, O. (2003). Beyond Stippling - Methods for Distributing Objects on the Plane. *Computer Graphics Forum 22, 22*(3), 515-522.

iTween. (2014). (Pixel Placement TM and LLC) Retrieved 2015, from http://itween.pixelplacement.com/index.php

Kim, J., & Pellacini, F. (2002, July). Jigsaw Image Mosaics. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2002, 21*(3), 657-664.

Lee, C. H., Varshney, A., & Jacobs, D. W. (2005, July). Mesh Saliency. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2005, 24*(3), 659-666.

Lloyd, S. P. (1982, March). Least squares quantization in PCM. *IEEE Transactions on Information Theory, 28*(2), 129-137.

Lohmann, S., Ziegler, J., & Tetzlaff, L. (2009). Comparison of Tag Cloud Layouts: Task-Related Performance and Visual Exploration. *Proceeding INTERACT '09 Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I* (pp. 392-404). Berlin, Heidelberg: Springer-Verlag.

Merrell, P., & Manocha, D. (2011). Model Synthesis: A General Procedural Modeling Algorithm. *IEEE Transactions on Visualization and Computer Graphics. 17*, pp. 715-728. IEEE.

Min, P. (1997). *meahConv*. Retrieved 2015, from http://www.patrickmin.com/meshconv/

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011, March 12). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, 1238 - 1242.

Page, D. L., Koschan, A. F., Sukumar, S. R., Roui-Abidi, B., & Abidi, M. A. (2003). Shape Analysis Algorithm based on Information Theory. *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. 1*, pp. I - 229-32 vol.1. IEEE.

Polonsky, O., Patane, G., Biasotti, S., & Gotsman, C. (2005). What's in an Image? Towards the Computation of the "Best" View of an Object. *The Visual Computer*, (pp. 840-847).

Princeton University. (2005, March 15). *Princeton Shape Benchmark*. Retrieved 2015, from Princeton Shape Retrieval and Analysis Group: http://shape.cs.princeton.edu/benchmark/

Reinert, B., Ritschel, T., & Seidel, H.-P. (2013, November). Interactive by-example design of artistic packing layouts. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2013, 32*(6), No. 218.

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007). Getting Our Head in the Clouds: Toward Evaluation Studies of Tagclouds. *CHI 2007 Proceedings*, *Tags, Tagging & Notetaking*, pp. 995-998. San Jose, CA, USA.

Specht, E. (2012, Sep 12). *The best known packings of equal circles in a rectangle with variable aspect ratio*. Retrieved 2015, from The best known packings of equal circles in a rectangle with variable aspect ratio: http://hydra.nat.uni-magdeburg.de/packing/crc_var/crc.html

Vazquez, P.-P., Feixas, M., Sbert, M., & Heidrich, W. (2001). Viewpoint Selection using Viewpoint Entropy. *VMV '01 Proceedings of the Vision Modeling and Visualization Conference 2001*, (pp. 273-280). Stuttgart, Germany.

Viegas, F. B., Wattenberg, M., & Feinberg, J. (2009). Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics. 15*, pp. 1137 - 1144. IEEE.

Warehouse, 3. (2014). *Homepage*. (Trimble Navigation Limited) Retrieved 2016, from 3D Warehouse: https://3dwarehouse.sketchup.com/index.html

Wikipedia. (2016, November 29). *Archimedean spiral*. (I. Wikimedia Foundation, Producer) Retrieved November 2016, from Wikipedia: https://en.wikipedia.org/wiki/Archimedean_spiral

Wikipedia. (2016, September 20). *Tag cloud*. Retrieved September 2016, from Wikipedia: https://en.wikipedia.org/wiki/Tag_cloud#cite_note-stateunion-24

Yee, K.-P., Fisher, D., Dhamjia, R., & Hearst, M. (2001). Animated Exploration of Dynamic Graphs with Radial Layout. *Information Visualization, 2001. INFOVIS 2001. IEEE Symposium on* (pp. 43-50). San Diego: IEEE.

Yobi3D. (2016). *Search result of keyword "vehicle"*. Retrieved 2016, from Yobi: https://www.yobi3d.com/#!/search?q=vehicle

Yu, L.-F., Yeung, S.-K., Tang, C.-K., Terzopoulos, D., Chan, T. F., & Osher, S. J. (2011, July). Make it Home: Automatic Optimization of Furniture Arrangement. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2011, 30*(4), Article No. 86.

Appendix A1: Email Recruitment Notice for the First-group study

We are recruiting participants to take part in a research study that evaluates a new object selection method called Interactive 3D-object Clouds for selecting objects on a tablet. We are looking for students, faculty or staff members from Dalhousie University who are not color blind, can speak and read English fluently and have some experience with touch screen devices.

The study will be conducted in a quiet room on the campus of Dalhousie University, and will take approximately 40 minutes. Firstly, a researcher will go through an introduction section of the study and ask you to sign a consent form to do the study. Then you will need to fill out a background questionnaire and a 3D warmup exercise. After that, you will perform several tasks related to the selection of 3D models on a tablet. Finally, there will be a post-task questionnaire and a short interview. The whole process will be video recorded, and the interview will be voice recorded. All personal data from the study will be kept strictly confidential. Compensation is $10 for participation in the study.

If you are interested in participating, please contact Xiaoting Hong (x.hong@dal.ca).

Appendix A2: Email Recruitment Notice for the Second-group Study

We are recruiting participants to take part in a research study that evaluates the animation design of a new object selection method called Interactive 3D-object Clouds for selecting objects on a tablet. We are looking for students, faculty or staff members from Dalhousie University who are not color blind, can speak and read English fluently and have some experience with touch screen devices.

The study will be conducted in a quiet room on the campus of Dalhousie University, and will take approximately 40 minutes. Firstly, a researcher will go through an introduction section of the study and ask you to sign a consent form to do the study. Then you will need to fill out a background questionnaire and a 3D warmup exercise. After that, you will perform several tasks related to the selection of 3D models on a tablet. Finally, there will be a post-task questionnaire and a short interview. The whole process will be video recorded, and the interview will be voice recorded. All personal data from the study will be kept strictly confidential. Compensation is $10 for participation in the study.

If you are interested in participating, please contact Xiaoting Hong (x.hong@dal.ca).

Appendix B1: Poster Recruitment Notice for the First-group Study

Recruitment Notice

3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices

We are recruiting participants to take part in a research study that evaluates a new selection method for selecting 3D models on a tablet. We are looking for students, faculty or staff members from Dalhousie University who are not color blind, can speak and read English fluently and have experience with touch screen devices.

The study will be conducted in a quiet room on campus of Dalhousie University, and will take approximately 40 minutes. Firstly, a researcher will go through introduction section of the study and ask you to sign a consent form to do the study. Then you will need to fill out a background questionnaire and a 3D warmup exercise. After that, you will perform several tasks related to the selection of 3D models on a tablet. Finally, there will be 3 single page multiple choice questionnaires that will be filled out (one for each two of the 6 tasks) and a short interview. The whole process will be video recorded, and the interview will be voice recorded. All personal data from the study will be kept strictly confidential. Compensation is $10 for participation in the study.

If you are interested in participating, please contact Xiaoting Hong (x.hong@dal.ca).

Appendix B2: Poster Recruitment Notice for the Second-group Study

Recruitment Notice

3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices

We are recruiting participants to take part in a research study that evaluates the animation design of a new object selection method called Interactive 3D-object Clouds for selecting objects on a tablet. We are looking for students, faculty or staff members from Dalhousie University who are not color blind, can speak and read English fluently and have experience with touch screen devices.

The study will be conducted in a quiet room on campus of Dalhousie University, and will take approximately 40 minutes. Firstly, a researcher will go through introduction section of the study and ask you to sign a consent form to do the study. Then you will need to fill out a background questionnaire and a 3D warmup exercise. After that, you will perform several tasks related to the selection of 3D models on a tablet. Finally, there will be 3 single page multiple choice questionnaires that will be filled out (one for each two of the 6 tasks) and a short interview. The whole process will be video recorded, and the interview will be voice recorded. All personal data from the study will be kept strictly confidential. Compensation is $10 for participation in the study.

If you are interested in participating, please contact Xiaoting Hong (x.hong@dal.ca).

Appendix C: Screen Email

Thank you very much for your reply to the recruitment notice. I have a few questions to ask before you can participate in our study.

1. Are you color blind?
2. Do you speak and read English fluently?
3. Do have any experience with touchscreen devices, such as tablets?
4. Are you willing to be audio and video recorded in the study? (All personal data from the study will be kept strictly confidential.)

Please answer each question and return them in an email to me. Again, thank you for showing interest in this study. Contact information is Xiaoting Hong (x.hong@dal.ca).

Appendix D1: Informed Consent for the First-group Study

3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices

Principle Investigator:

Xiaoting Hong, MCS Thesis Student, Faculty of Computer Science (x.hong@dal.ca)

Dr. Stephen Brooks, Faculty of Computer Science (sbrooks@cs.dal.ca)

Contact Person:

Xiaoting Hong, MCS Thesis Student, Faculty of Computer Science (x.hong@dal.ca)

We invite you to participate in a research study being conducted by Xiaoting Hong and Dr. Stephen Brooks at Dalhousie University. There are 48 positions available in total. This study is entirely voluntary and you can withdraw from it anytime you wish. Your academic (or employment) performance evaluation will not be affected in any way by whether or not you take part in the study. If you have any further questions about the study, please contact Xiaoting Hong.

The study is about evaluating 3D object selection methods on mobile devices and it could take approximately 40 minutes. The researcher will guide you through the beginning with an introduction to the study. You will be given several tasks related to the usability of the application. During the study, your voice, times and finger interaction with the tablet will be recorded and your whole body (including head and hands movements) will be videotaped after you log in. Although there is a chance that may not achieve some tasks, the researcher will always be close to providing guidance and encouragement. After finishing each task, you will need to fill out a multiple choice post-task questionnaire (6 in total) and participate in a short interview with the researcher. The researcher will be always available to answer any questions you might have.

All personal data from the study will be kept strictly confidential. We will use pseudonyms (e.g., P1, P2) to denote you and all your direct quotes. Your pseudonyms will be random and not related to your name. All the data will be kept in a secure location under confidentiality.

Every participant will receive compensation of $10 and it will be given directly from the researcher after the study. The compensation will be given even if the participant does not finish the study.

In the event that you have any difficulties with, or wish to voice concerns about, any aspect of your participation in this study, you may contact Catherine Connors, Director, Office of Research Ethics Administration at Dalhousie University's Office of Human Research Ethics for assistance: phone: (902)494-1462, email: Catherine.connors@dal.ca.

"*I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered to my satisfaction. I hereby consent to take part in the study. However, I understand that my participation is voluntary and that I am free to withdraw from the study at any time.*"

Participant	**Researcher**
Name: ____________________________	Name: ____________________________
Signature: _________________________	Signature: _________________________
Date: ______________________________	Date: ______________________________

"I understand and consent that my participation in the experiments will be audio and video recorded for the purpose of analysis. I understand that this is a condition of participation in the study, and I understand that this audio and video record will not be used in the publication or presentation of results."

Participant	**Researcher**
Name: ___________________________	Name: ___________________________
Signature: ________________________	Signature: ________________________
Date: _____________________________	Date: _____________________________

If you are interested in seeing the results of this study, please check below and provide your email address. We will contact you with publication details that describe the results.

"I would like to be notified by email when results are available via a publication."
[If this option is chosen, please include a contact email address: __________________]

Participant	**Researcher**
Name: ___________________________	Name: ___________________________
Signature: ________________________	Signature: ________________________
Date: _____________________________	Date: _____________________________

Appendix D2: Informed Consent for the Second-group Study

3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices

Principle Investigator:

Xiaoting Hong, MCS Thesis Student, Faculty of Computer Science (x.hong@dal.ca)

Dr. Stephen Brooks, Faculty of Computer Science (sbrooks@cs.dal.ca)

Contact Person:

Xiaoting Hong, MCS Thesis Student, Faculty of Computer Science (x.hong@dal.ca)

We invite you to participate in a research study being conducted by Xiaoting Hong and Dr. Stephen Brooks at Dalhousie University. There are 24 positions available in total. This study is entirely voluntary and you can withdraw from it anytime you wish. Your academic (or employment) performance evaluation will not be affected in any way by whether or not you take part in the study. If you have any further questions about the study, please contact Xiaoting Hong.

The study is about evaluating 3D object selection methods on mobile devices and it could take approximately 40 minutes. The researcher will guide you through the beginning with an introduction to the study. You will be given several tasks related to the usability of the application. During the study, your voice, times and finger interaction with the tablet will be recorded and your whole body (including head and hands movements) will be videotaped after you log in. Although there is a chance that may not achieve some tasks, the researcher will always be close to providing guidance and encouragement. After finishing each task, you will need to fill out a multiple choice post-task questionnaire (6 in total) and participate in a short interview with the researcher. The researcher will be always available to answer any questions you might have.

All personal data from the study will be kept strictly confidential. We will use pseudonyms (e.g., P1, P2) to denote you and all your direct quotes. Your pseudonyms will be random and not related to your name. All the data will be kept in a secure location under confidentiality.

Every participant will receive compensation of $10 and it will be given directly from the researcher after the study. The compensation will be given even if the participant does not finish the study.

In the event that you have any difficulties with, or wish to voice concerns about, any aspect of your participation in this study, you may contact Catherine Connors, Director, Office of Research Ethics Administration at Dalhousie University's Office of Human Research Ethics for assistance: phone: (902)494-1462, email: Catherine.connors@dal.ca.

"*I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered to my satisfaction. I hereby consent to take part in the study. However, I understand that my participation is voluntary and that I am free to withdraw from the study at any time.*"

Participant	**Researcher**
Name: ___________________________	Name: ___________________________
Signature: ________________________	Signature: ________________________
Date: ____________________________	Date: ____________________________

“I understand and consent that my participation in the experiments will be audio and video recorded for the purpose of analysis. I understand that this is a condition of participation in the study, and I understand that this audio and video record will not be used in the publication or presentation of results.”

Participant	**Researcher**
Name: ___________________________	Name: ___________________________
Signature: ________________________	Signature: ________________________
Date: _____________________________	Date: _____________________________

If you are interested in seeing the results of this study, please check below and provide your email address. We will contact you with publication details that describe the results.

“I would like to be notified by email when results are available via a publication.”
[If this option is chosen, please include a contact email address: __________________]

Participant	**Researcher**
Name: ___________________________	Name: ___________________________
Signature: ________________________	Signature: ________________________
Date: _____________________________	Date: _____________________________

Appendix E: Background Questionnaire

1. Gender: ☐ Male ☐ Female

2. Age:

3. Participant ID:

4. How familiar are you with mobile devices (tablet or phone)?

☐ Very familiar ☐ Somewhat familiar ☐ Not familiar

5. How familiar are you with touch-screen devices?

☐ Very familiar ☐ Somewhat familiar ☐ Not familiar

6. How familiar are you with 3D models?

☐ Very familiar ☐ Somewhat familiar ☐ Not familiar

7. Have you ever interacted with 3D models on a computer?

☐ A lot ☐ Sometimes ☐ Not at all

8. Do you have experience on interacting with 3D models on a tablet or other mobile devices?

☐ A lot ☐ Some ☐ A little ☐ None

9. Do you have experience with tag clouds?

☐ A lot ☐ Some ☐ A little ☐ None

Appendix F1: Post-task Questionnaire for the First-group Study

Please respond to the following statements using the given scale:

1	2	3	4	5
Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

1. The controlling interface of the application is easy to understand

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

2. It was easy to finish the task without any difficulty

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

3. It was easy for me to find the targeted objects

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

4. It was easy to browse the models

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

5. It was easy for me to identify the differences and similarities between the target and other objects

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

6. The layout is "visually pleasing"

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

7. The orientations (viewpoints) of the objects were informative

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

8. The arrangement of the 3D objects was understandable

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

Appendix F2: Post-task Questionnaire for the Second-group Study

Please respond to the following statements using the given scale:

1	2	3	4	5
Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

1. It was easy to finish the task without any difficulty

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

2. It was easy for me to track the targeted objects

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

3. The movement of objects is easy to follow

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

4. The movement of objects is understandable

Cloud View	1	2	3	4	5

Grid View	1	2	3	4	5

5. The movement of objects is predictable

Cloud View	1	2	3	4	5
Grid View	1	2	3	4	5

Appendix G1: Interview Questionnaire for the First-group Study

1. Did you find the task easy in the Cloud View / Grid View approach? Why?

2. What difficulties did you have with the Cloud View / Grid View approach?

3. What advantages and disadvantages does each technique have?

4. Which system you prefer to browse the 3D object dataset? Why?

5. What improvements can our Cloud View / Grid View technique benefit from in the future?

____

6. Have you ever had experience browsing 3D models in the Cloud View way? How did you find these techniques in comparison with other techniques you may have used?

Appendix G2: Interview Questionnaire for the Second-group Study

1. Did you find the task easy with the Staggered / Non-staggered animation? Why?

2. What kind of strategy you have developed to track/find the targets with the Staggered / Non-staggered animation?

3. When this strategy was challenged with the Staggered / Non-staggered animation?

4. Which type of transitions you prefer? Why?

Appendix H: Participant Payment Receipt

My signature below confirms that I received an amount of 10 CAD from Xiaoting Hong as an honorarium payment for participating in the "3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices" research project.

I understand this honorarium is taxable income and it is my responsibility to claim it on my income tax as Dalhousie University will not be issuing a T4A for this payment.

Name (please print): ______________________________________

Signature: ______________________________________

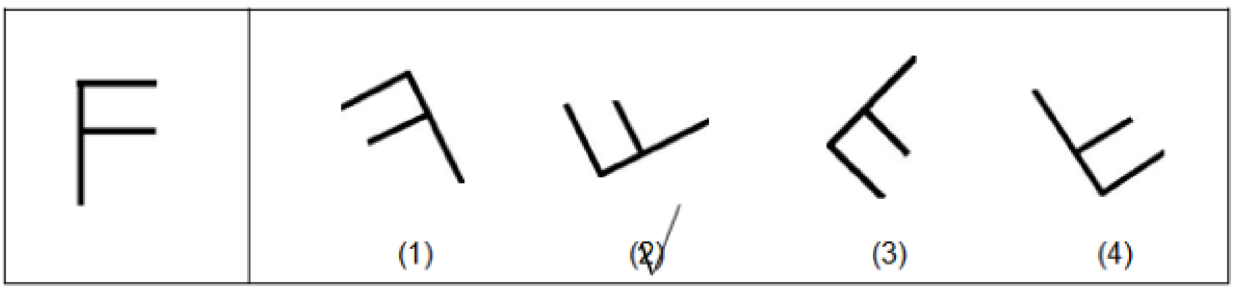
Date: ______________________________________

Appendix I: Spatial Ability Test

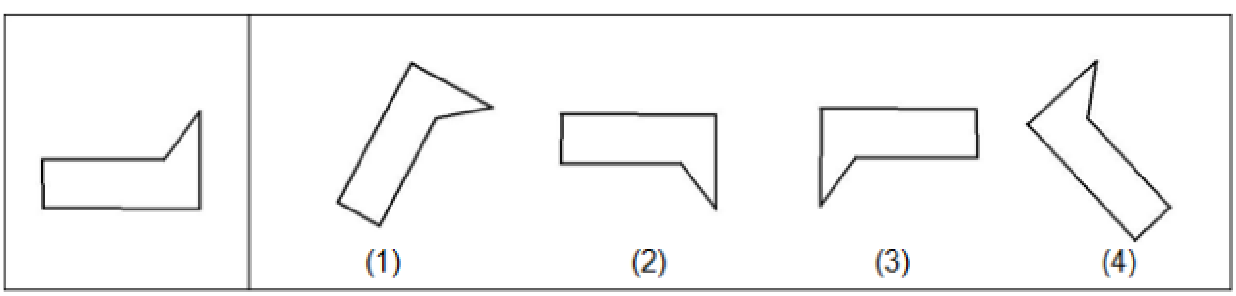
Spatial Ability Test: Mental Rotation

Instruction: In the right side, which one is the same with the left target after rotating? Please tick the right answer.

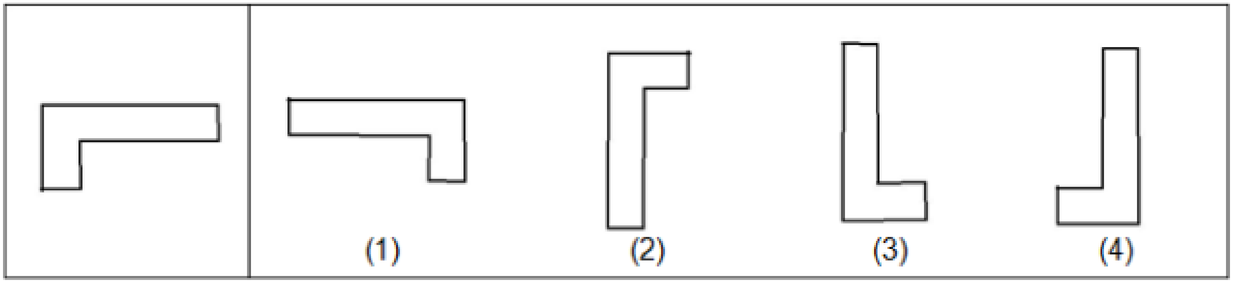
Sample:



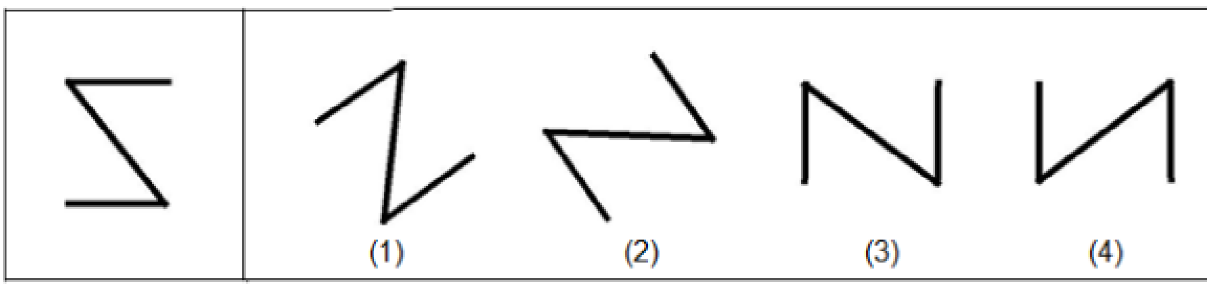
Practice:



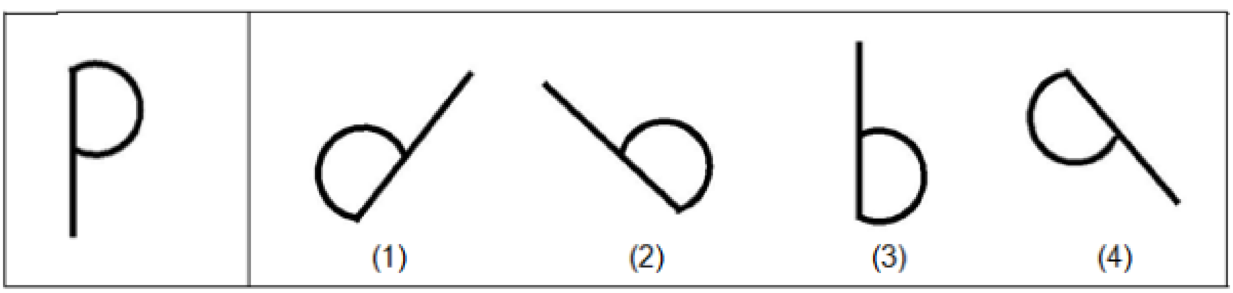
Question1:



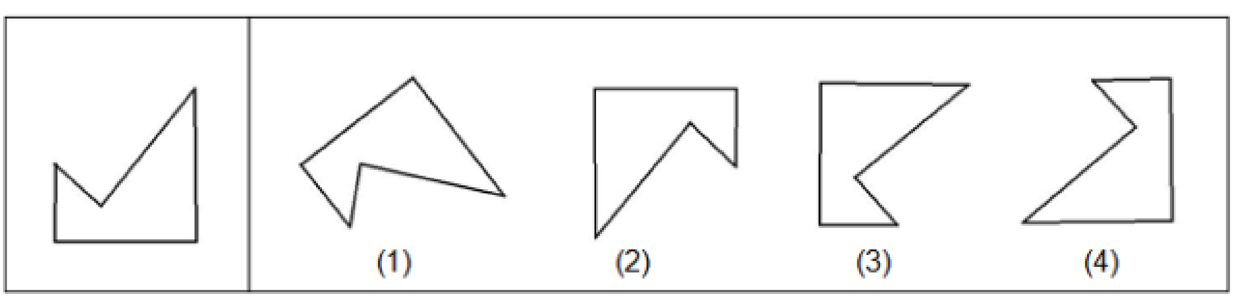
Question 2:



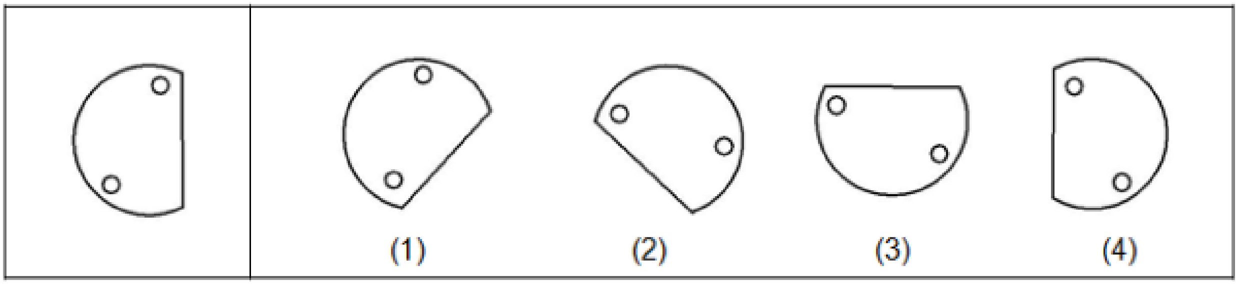
Question 3:



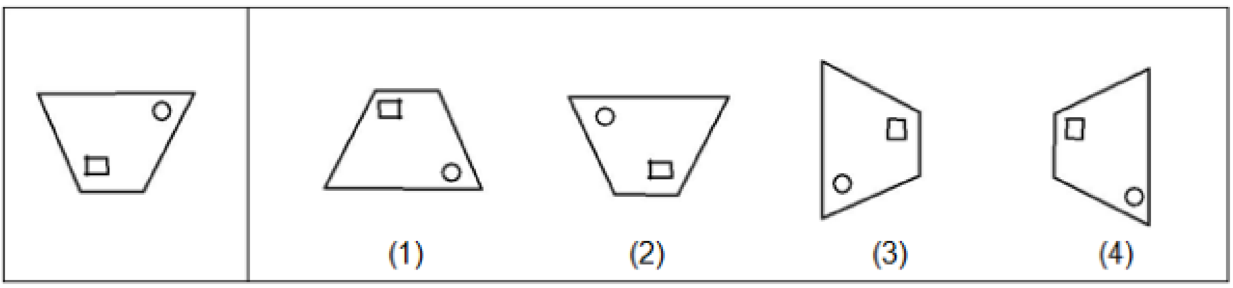
Question 4:



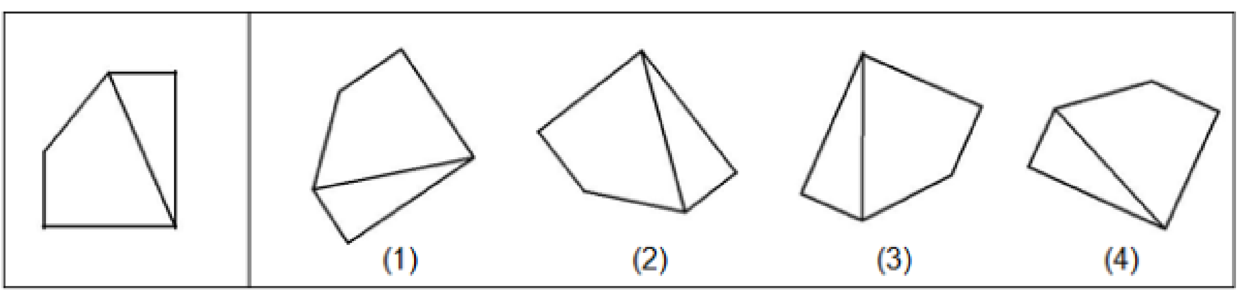
Question5:



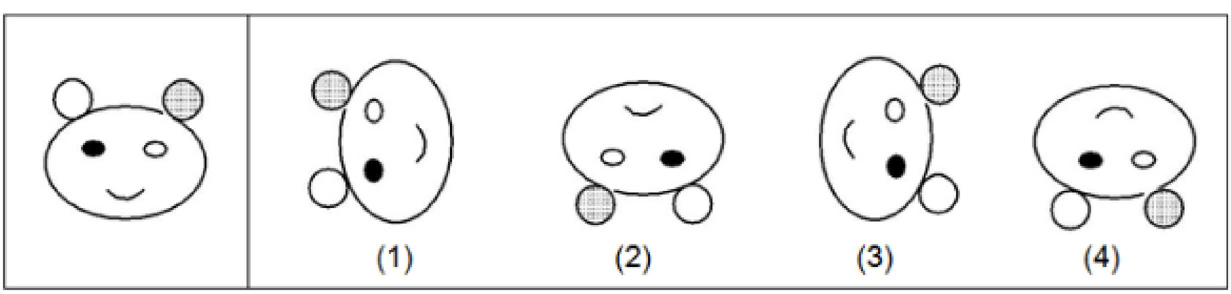
Question 6:



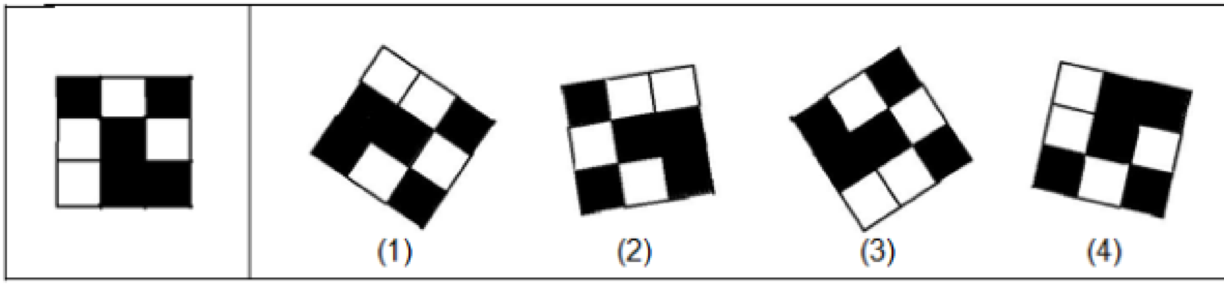
Question 7:



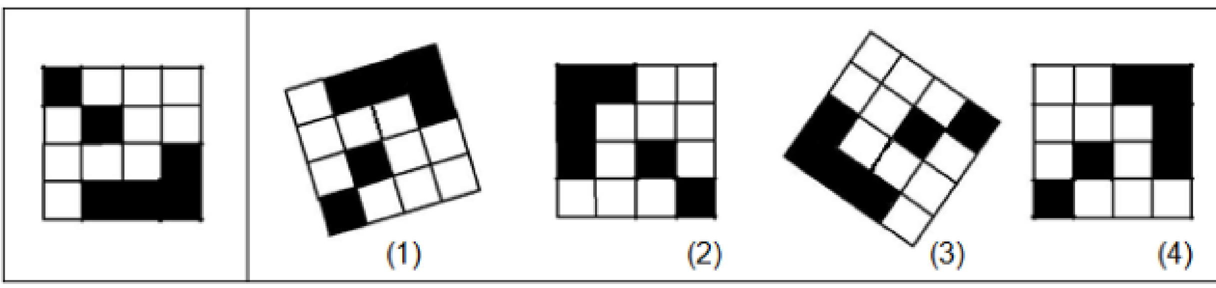
Question 8:



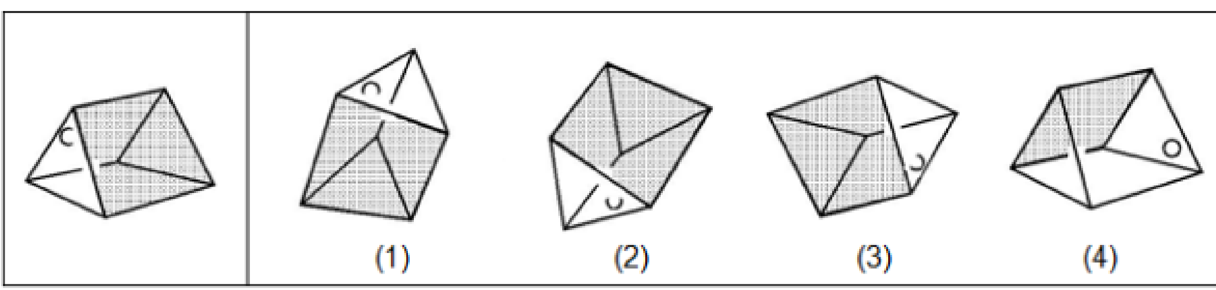
Question 9:



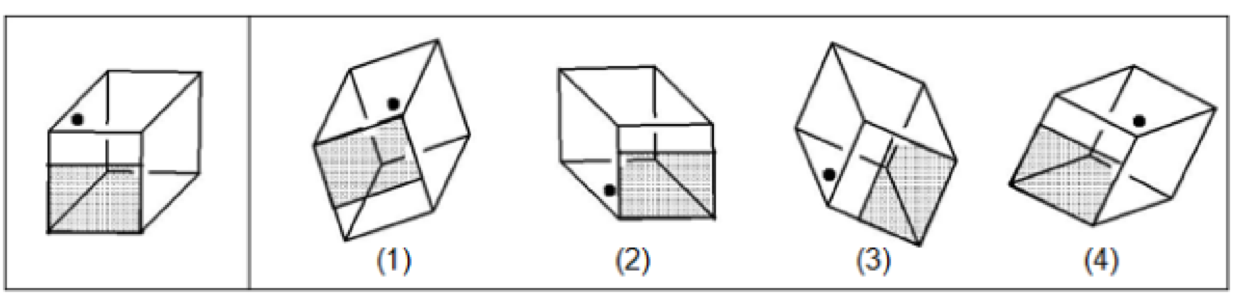
Question 10:



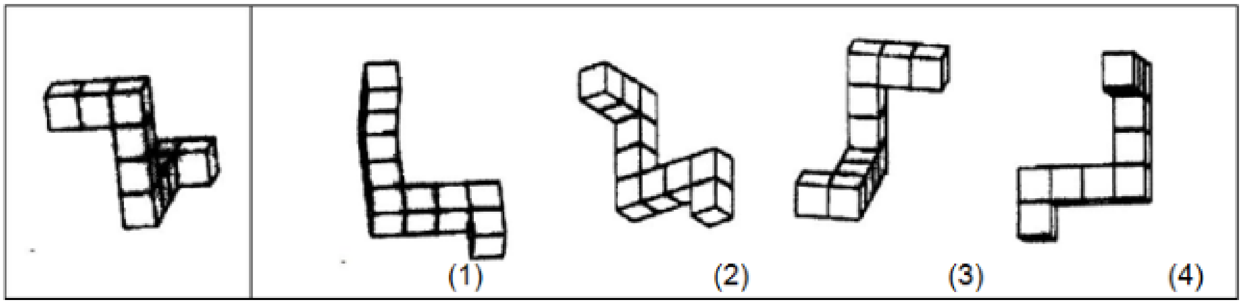
Question 11:



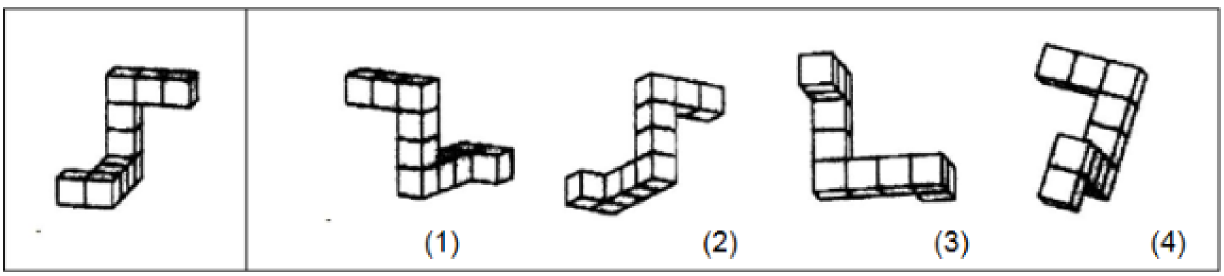
Question 12:



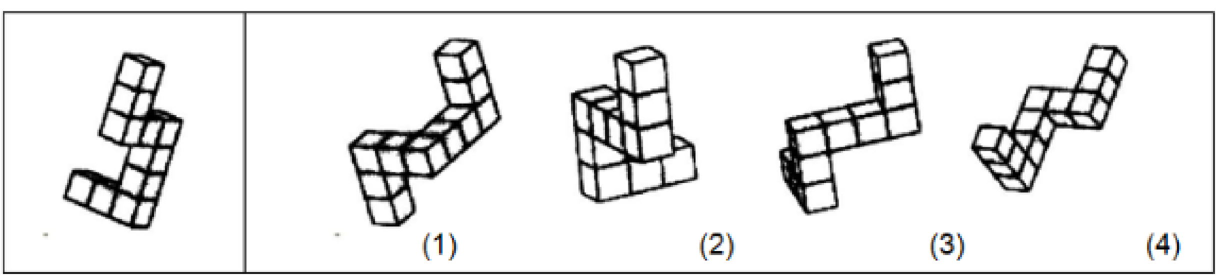
Question 13:



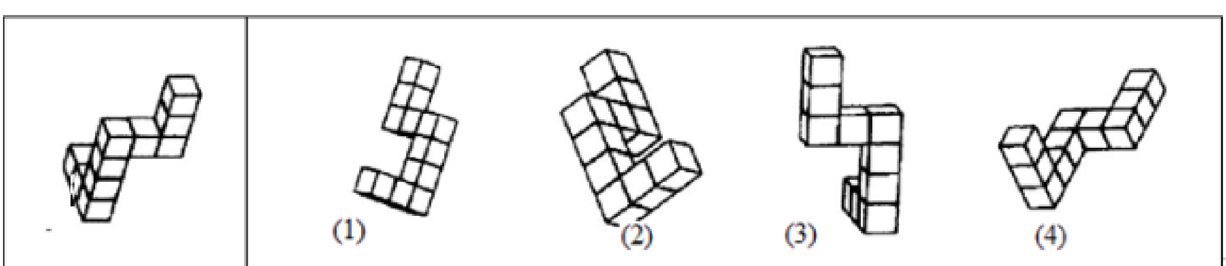
Question 14:



Question 15:



Question 16:



Appendix J: Research Ethics Board Approval Letter

Social Sciences & Humanities Research Ethics Board
Letter of Approval

June 08, 2015

Ms Xiaoting Hong
Computer Science\Computer Science

Dear Xiaoting,

REB #: 2015-3567
Project Title: 3D Object Clouds: Viewing Virtual Objects in Interactive Clouds on Mobile Devices

Effective Date: June 08, 2015
Expiry Date: June 08, 2016

The Social Sciences & Humanities Research Ethics Board has reviewed your application for research involving humans and found the proposed research to be in accordance with the Tri-Council Policy Statement on *Ethical Conduct for Research Involving Humans.* This approval will be in effect for 12 months as indicated above. This approval is subject to the conditions listed below which constitute your on-going responsibilities with respect to the ethical conduct of this research.

Sincerely,
Dr. Valerie Trifts, Chair