Data Fragment: Virtual Reality for Viewing and Querying Large Image Sets

Theophilus Teo*
University of South Australia

Mitchell Norman†
University of South Australia

Matt Adcock‡
CSIRO
University of South Australia

Bruce H. Thomas§
University of South Australia

ABSTRACT

This paper presents our new Virtual Reality (VR) interactive visualization techniques to assist users querying large image sets. The VR system allows users to query a set of images on four different filters such as locations and keywords. The goal is to investigate if a VR platform is preferred over a non-VR platform for viewing and querying large image sets. We employed an HTC Vive and a traditional desktop screen to represent VR and non-VR platforms. We found users preferred the VR platform over the traditional desktop screen.

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities—

1 INTRODUCTION

We present our Virtual Reality (VR) visualization system for viewing and querying a large set of images with meta data. The paper examines different techniques of interacting inside a virtual environment [4], as well as manipulating the 3D data objects in order to assist with image selection analytics. VR holds the potential as a means of exploring large images sets or photographic collections in an intuitive manner. We believe the use of a person's proprioception will enhance their ability to explore large image sets.We introduce *Data Fragments* an image aggregation visualization technique that groups up to 24 images into a single 3D object. We performed a user study to evaluate the viability of the application running in VR (HTC Vive) and non-VR (traditional desktop display). The main contributes of this paper are: (1) a new method for viewing and filtering large image collections and (2) a user study to validate the new method we have developed.

2 BACKGROUND

Immersive Analytics has suggested VR technologies could become an inexpensive and effective tools for business and the public [1]. Large image sets with meta data such as ImageNet [2] are now common and accessible. CAVE-based VR [3] has been exploited for the exploration of large image corpuses, with two modes of interactive exploration: one was a steering wheel visualization and the second mode was a logical hierarchy of visual concepts. Venson et al. [5] determined VR was an effective diagnostic tool in fracture identification with medical images.

3 VR SYSTEM FOR LARGE SETS OF IMAGES

We developed a new VR application and new interaction techniques to visualize large sets of images in a 3D virtual world. The different VR-based and non VR-based platforms will generate different user experience with the different display devices and input controls. We

---

*e-mail: teoty004@mymail.unisa.edu.au

†e-mail:normf001@mymail.unisa.edu.au

‡e-mail:Matt.Adcock@data61.csiro.au

§e-mail:bruce.thomas@unisa.edu.au

(a) Data fragment 17 images
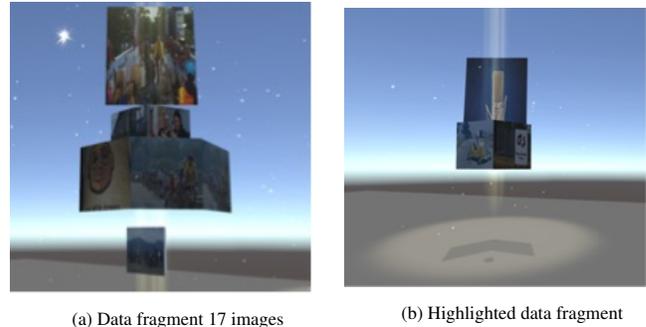
(b) Highlighted data fragment

Figure 1: Data Fragment

leverage Unity 3D's cross-platform infrastructure to support this capability. The image's meta data is leveraged as a means of querying the different images.

3.1 Software Design

The system is constructed with Unity 3D running on versions 5.3.4 and 5.4, with a JSON file as the data source, SteamVR for running HTC Vive, SteamVR Unity Toolkit, and the C# programming language. The image set's meta data for each image is retrieved from a JSON file. The system is designed to work with image sets defined by a JSON file. Future investigations are required to automatically accomindate different JSON files. The example application presented in this paper employs a data set that was retrieved from the Australian Government's data.gov.au website. The data is a set of 6,326 metadata entries from stories featured by the Australian Broadcasting Corporation (ABC). We developed a data set of images from the ABC with 14 meta data items, such as state the article originated from, city or township featured in the article, ABC station that published the article, and story publication date.

3.2 Data Fragment

The images and their meta data will be grouped together and displayed as a set of Data Fragments that will be represented as the individual data inside the 3D world. In order to visualize images that share common attributes, we designed a 3D object with multiple sides of pictures to represents the data entity with common attributes; we refer the 3D objects as *Data Fragments*. Figure 1a shows a Data Fragment with 17 related data images. The main image of a Data Fragment will always appear on the top layer whereas the related images will appear on the other three layers. There can be one to four layers for a Data Fragment. The number of layers determines the quantity of the related data that can be displayed, and when one layer fills, a new layer is added. This provides a quick understanding of the amount of related data on a Data Fragment without highlighting other images, as well as to provide easy distinguish when there is two set of Data Fragments with a similar quantity of related data (such as a set of 8 Related Data and a set of 9 Related Data). A limitation of our Data Fragment is that it can only show up to 24 related images at a time with four layers (Layer 1 - one image, Layers 2, 3, and 4 - up to eight images). This is

because the user may not able to view the main image properly, as with a Data Fragment with 80 related images with ten layers.

In order to assist the analytics for observing and searching large image sets, the images are positioned in a meaningful orientation. The user applies a *Filter* to specify which parameters are applied to the Data Fragment positioning calculation. Our current system supports four Filters: State, Place, Station and Year. Each of the Data Fragment will be placed in a range of x, z coordinate where y coordinate (height) will be fixed. The value of x, z coordinates will be constrained by the size of *Floor*, the space allowed for the user to move inside the virtual world. The center of the Floor is the origin. The Data Fragment is allocated with same Filter type as a group and will only spawn on a range of a degree from the origin. The calculation for the spawn point is based on the Filter *parameter size*$/360°$. For example, the State parameter consists of six states ($60°/state$).

We support the highlighting of selected Data Fragments that share a common parameter (See Figure 1b). The application allows the user to show the related Data Fragments of a specific Data Fragment by highlighting them when the desired one is being selected. The Data Fragment that has been highlighted can be easily distinguished from the ray of light projected directly above it.

3.3 VR Interaction

When all of the Data Fragments are loaded into the virtual world, the user can either use the mouse for the Non-VR platform or the HTC Vive controls for the VR platform to manipulate them. The remainder of this section will focus on the VR interactions, as the system leverages the standard Unity 3D interaction techniques. The user interface displays the meta data for the most recently selected image and the filter that was currently being used. We leveraged the HTC Vive Controllers as our main input devices. The Table below shows the list of controls for each of the function of controlling the system. The Vive system allows users easily look and move around inside the virtual world by movement of their head and walking.

Controls	HTC Vive Input device/button
Walk around (XZ axis)	Head display
Look around (XYZ axis)	Head display
Teleport to a coordinate	Left controller trigger button
Point at a Data Fragment	Right controller trigger button
Change filter	Left controller touchpad

The user interface for the VR system supports information panels for the last select image, and the current filter is shown floating above each controller (Figure 2a and 2b). A user can also change the filter or finish a task by using the buttons that are indicated by the graphical overlay on the left controller (Figure 2b). For selection of an image, the user presses the trigger button on the right controller which shows a pointer line to indicate the image that is currently being selected. The trigger on the left controller is used for teleporting and will display an indicator to show where the user will land by releasing the trigger.

4 USER STUDY

We performed a user study in order to evaluate the preference of using the VR platform and the non-VR platform for image querying and viewing. We engaged 20 participants aged between 18 to 40 to participate in the study. Participants were required to search for a Data Fragment with a specific parameter inside the virtual world that consists of up to 100 Data fragments. There were four tasks for each of the VR and non-VR experiments.

The quantitative results showed user performed better on the desktop user interface. The participants through the questionnaire stated they preferred the VR setting to the desktop. The participants found the experience novel and interesting. The feedback indicated most of the participants having difficulties with the controls in VR,

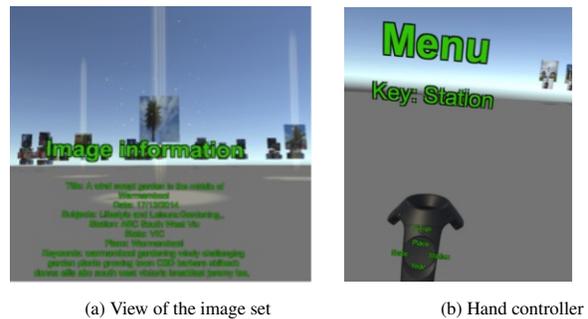(a) View of the image set   (b) Hand controller

Figure 2: VR user interface

as they are unfamiliar to them. The participants did enjoy the movement features of the HTC Vive, as it accurately tracked the user's movements. This helped participants to be able to look at a direction easily since the traditional desktop had a limited field of view and they are required to hover the mouse to an edge in order to look on the direction. In addition, the visual feedback that shows a beam from the Vive controller inside the virtual world when pointing or teleporting is also one of the favorite for the participants as it gives a depth perception, while the mouse is limited to a 2D grid that has a limited 3D visual feedback when pointing at a Data Fragment.

Regarding the user interface design, most of the participants had difficulties reading the words, as the layout are inconsistent. This is mainly due to some of the Data Fragments appeared to have more details than the others. The larger amount of information will shrink all of the text in order to fit into the interface. Two participants mentioned: *For short term tasks, non VR will be better as it allows someone to sit down and start using it but for long term tasks, VR has a significant advantage because of the features it provided.*

5 CONCLUSION

In summary, we compared the efficiency of user analyzing large image sets in an HMD VR platform against a non-VR platform. We evaluated the speed and accuracy of task solving from people without analytical background by performing a user study. This result confirmed our hypothesis that people would prefer using the VR platform over the non VR platform. In the future, we would like to develop a method to represent the Data Fragment that has more than 24 related images and to allow for collaboration between multiple users.

REFERENCES

[1] T. Chandler, M. Cordeil, T. Czauderna, T. Dwyer, J. Glowacki, C. Goncu, M. Klapperstueck, K. Klein, K. Marriott, and F. Schreiber. Immersive analytics. In *IEEE Big Data Visual Analytics (BDVA)*, pages 1–8, 2015.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *IEEE CVPR*, pages 248–255, 2009.

[3] S. Khanwalkar, S. Balakrishna, and R. Jain. Exploration of large image corpuses in virtual reality. In *ACM Multimedia*, MM '16, pages 596–600, 2016.

[4] H. Slay, M. Phillips, R. Vernik, and B. Thomas. Interaction modes for augmented reality visualization. In *Asia-Pacific Symposium on Information Visualisation*, pages 71–75, 2001.

[5] J. E. Venson, J. Berni, C. S. Maia, A. M. da Silva, M. d'Ornelas, and A. Maciel. Medical imaging vr: Can immersive 3D aid in diagnosis? In *ACM Virtual Reality Software and Technology*, pages 349–350, 2016.