第 24 回日本バーチャルリアリティ学会大会論文集（2019 年 9 月）

Experiential Introduction About Objects Using AR-VR System

Xiangyi Gao[1)], 小木哲朗[2)]

1) 慶應義塾大学大学院　システムデザイン・マネジメント研究科　（〒223-8526 神奈川県横浜市港北区日吉 4-1-1, xiangyi.gao@keio.jp）

2) 慶應義塾大学大学院　システムデザイン・マネジメント研究科　（〒223-8526 神奈川県横浜市港北区日吉 4-1-1, ogi@sdm.keio.ac.jp）

Abstract: This research proposes an object detection-based AR-VR system which recognizes objects and delivers VR experience as information. As an implementation, tourism promotion purpose system was developed and exhibited. Users can select local specialties through object detection AR and experience the VR worlds which describe those objects. The evaluation indicates that this system is effective as an information system.

Key Word: Object detection, AR, VR website, User interface

1. Introduction

Online search has already become a major tool for people to receive information. However, when objects of interest are in front of a person, he or she still needs to search and view the information on a separate device. In this case, information and objects are independent, and a person may not understand the information intuitively. AR (Augmented reality) is a solution to enhance the relationship between objects and information. With AR technology, digital information can be overlaid onto the physical world, and information can be placed near objects to emphasized which object the information belongs to. Although AR strengthen the relationship between objects and information, the physical space to display information is limited. An effective way to deliver information through AR system should be proposed. According to researches [1][2], virtual reality (VR) is an effective method to deliver information. In this research, we propose an object detection-based AR-VR system, which delivers VR information after user selecting objects in AR. With this system, user can not only experience a VR environment for information, but also understanding which object the information is describing. As an implementation, a promotion event utilizing this system was held at Haneda Airport to introduce Shintomi-cho, a town in Miyazaki prefecture.

2. Concept

The concept of object detection-based AR-VR system is that when a user sees an object, he or she can immerse into and experience the VR world which describes the object. To promote Shintomi-cho, when a person sees a lychee, which is the local specialty, he or she will be taken to the actual lychee farm in Shintomi-cho where the lychee was picked, as shown in Figure 1. The person will learn information about the lychee while exploring the lychee farm and become interested in Shintomi-cho.

Figure 1: System Concept. When one sees the lychee, he or she will immerse into the lychee farm

3. Method

In this research, head mounted display (HMD) is utilized to bring users to VR worlds from physical objects. The HMD is built with an iPhone Xs and a head mount device HOMiDO PRIME that conforms to the Google Cardboard v2 which has a button to tap the screen. The system is implemented as an iOS application and consists of two phases of operation. The first phase is object detection AR, where a user with HMD can see-through with camera captured images, and detection results of objects are shown in the user's view as AR tagged bounding boxes. The second phase is selection and VR information delivering, where the user can select one of the detected objects through aiming the cursor and pressing the button on head mount device, and a WebVR website linked to the selected object is accessed. When the button is pressed again, the system changes back to the first phase again. The system layout is shown in Figure 2.

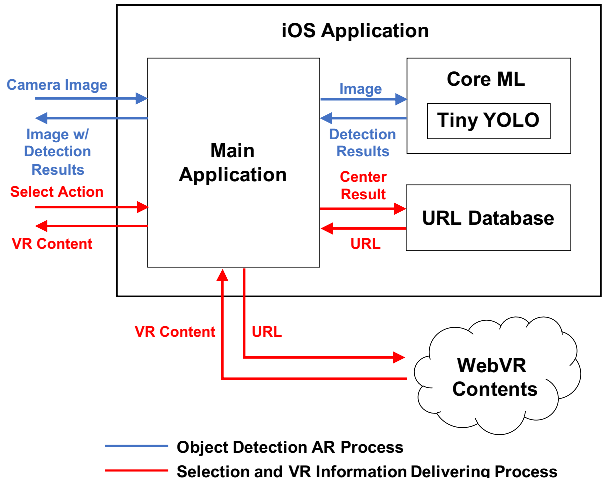

Figure 2: System Layout

3.1 Object Detection AR

The video see-through is achieved by displaying camera captured images side-by-side on the iPhone screen. Along with the video see-through process, on-device object detection is processed. The object detection function is developed utilizing Tiny YOLO algorithm. YOLO [3][4] is a deep learning-based object detection algorithm, and Tiny YOLO is the simplified version with less layers in convolutional neural network. Since the object detection function is implemented on a mobile device, Tiny YOLO which requires less computing power is preferred. We trained the model to detect three objects which are Shintomi-cho's lychee, beach scene printed panel, and morning market scene printed panel. The trained neural network model is integrated into the iOS application with Apple's Core ML [5]. The object detection results, which are tags and bounding boxes, are also displayed side-by-side on top of video see-through, as shown in Figure 3 and 4.

3.2 Selection and VR Information Delivering

The selection is performed through aiming and pressing button. If the center of the image which is marked by a cursor is located inside the desired object's bounding box, and the smartphone screen is tapped, the object is selected. Since pressing the button on the head mount device can trigger a tap action on screen, the user wearing HMD can perform object selection without taking it off.

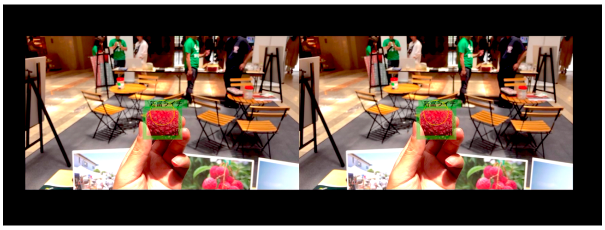

Figure 3: Object Detection AR

Figure 4: User Using the System

After the selection action, the object identifier is matched in the URL database, and a specific WebVR website is accessed. Customized WebVR websites are built on A-Frame, which is a web framework for building VR content. 360-degree videos taken with THETA V are used to construct VR environments. Since our system is trained to detect three classes of objects, which are Shintomi-cho's lychee, beach scene printed panel, and morning market scene printed panel, we build three VR contents which are actual Shintomi-cho lychee farm, beach, and morning market, corresponding to the detectable objects. The pairs of object and VR content are shown below in Figure 5, 6 and 7.

Figure 5: Lychee to Lychee Farm

Figure 6: Beach Scene Panel to Beach

Figure 7: Market Scene Panel to Market

4. Evaluation

This system was exhibited at Haneda Airport during the promotion event of Shintomi-cho, a town in Miyazaki prefecture. The evaluation was conducted through a survey and 60 participants answered the survey questionnaire after experiencing the system. The gender and age variance distribution of participants is shown in Figure 8. The participants in good variance answered four questions about object recognition AR experience, four questions about VR contents experience, and three questions about over all experience.

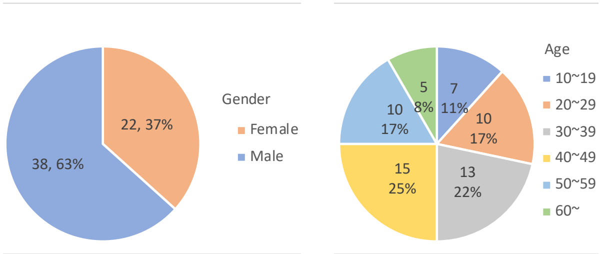

Figure 8: Gender and Age Distribution

4.1 Object Detection AR and VR Contents Evaluation

Same four questions about object detection AR and VR contents experiences were answered: whether the user did not feel sickness, spatial unawareness, delay, and narrow visual field. The answers were in five levels Likert scale (-2=Disagree, -1=Slightly Disagree, 0=Neither, 1=Slightly Agree, 2=Agree). The evaluation results are similar for both functions. The results are shown in Figure 9 and Figure 10.

The positive mean results indicate that users are comfortable with both AR and VR functions of this system regarding to sickness, spatial unawareness, delay, and narrow visual field. However, since the standard deviations of all four criteria for both functions are not small, the user experience of object detection AR and VR on the smartphone depends on each individual.

4.2 System Experience Evaluation

Three questions about the entire system experience were answered: whether the system is effective as information system, whether the system is interesting, and willingness of using the system in the future. The answers were in five levels Likert scale (-2=Disagree, -1=Slightly Disagree, 0=Neither, 1=Slightly Agree, 2=Agree). The results are shown in Figure 11. The positive mean and low standard deviation result for each question indicate that most users think the object detection-based AR-VR System is effective and interesting as an information system, and they are willing to see more information presented in VR through object detection AR.

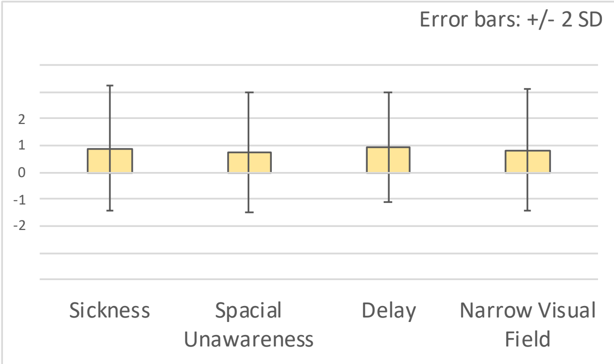

Figure 9: Object Detection AR Experience Evaluation

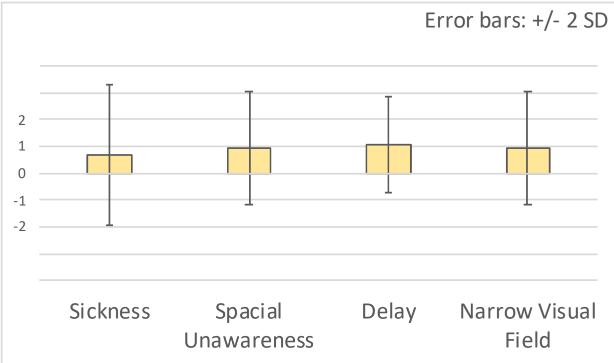

Figure 10: VR Contents Experience Evaluation

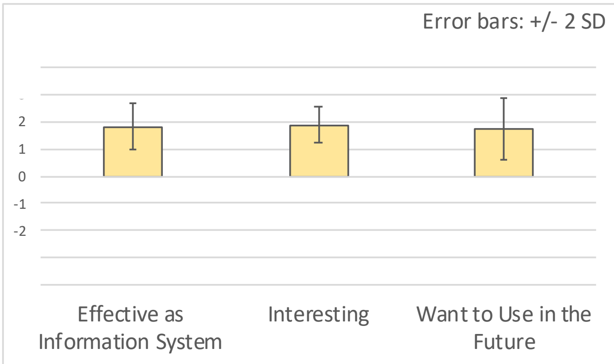

Figure 11: Over All Evaluation

5. Conclusion

This paper set out to propose and develop a VR information delivering AR system based on object detection. Furthermore, we implemented this system practically in order to promote a town in Miyazaki prefecture. Users can experience the VR worlds about real objects they are interested in for information. The effectiveness and potential of this system are verified through evaluation.

References

[1] Slavova, Yoana, and Mu Mu. "A Comparative Study of the Learning Outcomes and Experience of VR in Education." *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018.

[2] Usui, Shoko, et al. "Development and Evaluation

Experiment of Display Media Using VR for Art Appreciation Learning." *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, 2017.

[3] Redmon, Joseph, et al. "You Only Look Once: Unified, Real-Time Object Detection." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[4] Redmon, Joseph, and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[5] M. Hollemans, "YOLO with Core ML and MPSNNGraph," https://github.com/hollance/YOLO-CoreML-MPSNNGraph