Designing Affordances for Virtual Reality-Based Services with Natural User Interaction

Takayuki Miura, Akihito Yoshii, and Tatsuo Nakajima[✉]

Department of Computer Science and Engineering, Waseda University,
Tokyo, Japan
{t.miura,a_yoshii,tatsuo}@dcl.cs.waseda.ac.jp

Abstract. The progress of new technologies makes virtual reality (VR) easier, and inexpensive head-mounted displays (HMDs) accelerate the development of advanced VR services. The advantage of the technologies is to offer the better immersion of a VR world, to use natural user interface (NUI) devices and to operate them with a user's gesture. The problem to use NUI devices for VR-based services is to take into account various types of NUI devices. We need to consider how a user to use NUI devices in a proper way, in particular, when an HMD limits his/her view of the real world. Our approach to solve the problem is to use an affordance that offers implicit information how to navigate VR-based services. However, there are a little researches to investigate the relationship between different types of NUI devices and the design of proper affordances to navigate VR-based services. The paper provides some insights how respective types of NUI devices influence affordance design to navigate VR-based services. For extracting effective insights, we have developed two VR-based services, where we discuss two types of operating methods and three types of affordances for respective operating methods are examined. Also, we chose two types of NUI devices for navigating the VR-based services. Then, we conducted some experiments to extract some insights as a guideline to develop future VR-based services. The result shows that the differences among respective NUI devices may not significantly influence affordance design, however have strong effects on understanding how to navigate VR-based services. The understanding also affects how each user prefers which NUI devices, because NUI devices require us to use gesture to navigate the services, but the intuition that differs in each individual is important.

Keywords: Natural user interface · Virtual reality · Affordance

1 Introduction

Recently, as shown in [2, 6], virtual reality (VR) technologies have revived due to the progress of new technologies that make it easy to develop advanced VR-based services like programming platforms: *Unity*[1] and inexpensive and practical commercial head-mounted displays (HMD): *Oculus Rift*[2]. These technologies make it easy to

---

[1] Unity - Game Engine: http://japan.unity3d.com/, accessed 2015/12/22.
[2] Oculus Rift | Oculus: https://www.oculus.com/en-us/rift/.

develop various types of the VR-based service, and make it possible to use VR-based technologies in commercial purposes.

When offering desirable interactive user experiences, natural user interaction (NUI) devices like *Microsoft Kinect (MS Kinect)*[3] or *Leap Motion*[4] are widely used. Many VR-based games are already developed, and they assume to use these devices, because these devices offer an immersive user experience through the natural interaction [3] with the virtual world without prior knowledge. However, the assumption to use the current NUI devices causes a gap between the ideal expectation and the reality. In particular, using NUI devices with an HMD may cause a new problem. For example, Yang and Pan reported that *MS Kinect* fails to track a user's body when the user does not have sufficient experiences with an HMD [11].

When using NUI devices, it is usually assumed that a user can easily find where the devices are and how to navigate them, but the devices cannot be seen when the user wears an HMD. In computing environments, various commodity NUI devices will be used to develop new VR-based services; thus the described issues will become a more serious problem soon. One approach for overcoming the problem is using affordance [1], which is the information for ensuring the proper operation and can be used to navigate human behavior [7]. We also need to investigate what types of affordance is appropriate for respective NUI devices. We actually discuss three types of affordance — *inherent, image* and *sentence affordances* by referring to [10]. Furthermore, we like to research whether different features of NUI devices influence the appropriateness of the types of affordance. Thus, the research question in this paper is that a different NUI device needs a different affordance.

In this paper, we demonstrate how we can design proper affordances for respective NUI devices. The insights extracted from our experiment are useful for designing future NUI devices and VR-based services. We have developed a VR-based photo viewer service and a shooting game application for demonstrating the proposed ideas as case studies. The current case studies assume to use either *MS Kinect* or *Leap Motion* as NUI devices. We have designed three affordances for the respective NUI devices. *MS Kinect* tracks the positions of a user's body. In this case study, the position of a user's arms is captured for moving the cursors. Conversely, *Leap Motion* tracks the position of the joints of a user's hands. In this case study, the position of a user' hands is detected to move the cursors.[5]

2   Related Work

Terrenghi et al. claimed that the difference of user interfaces requires different affordances [9]. In their research, the participants were imposed to perform a puzzle task and a photo sorting task with physical and digital pieces and photos. Figures 1 and 2 are the scenes to perform the tasks. The result showed that even in the same task, the difference

---

[3] Xbox 360 – Kinect: http://www.xbox.com/ja-JP/kinect, accessed 2015/12/26.

[4] Leap Motion: https://www.leapmotion.com/?lang=jp, accessed 2015/12/22.

[5] A preliminary result was reported in [4], and the paper enhances insights discussed in the paper.
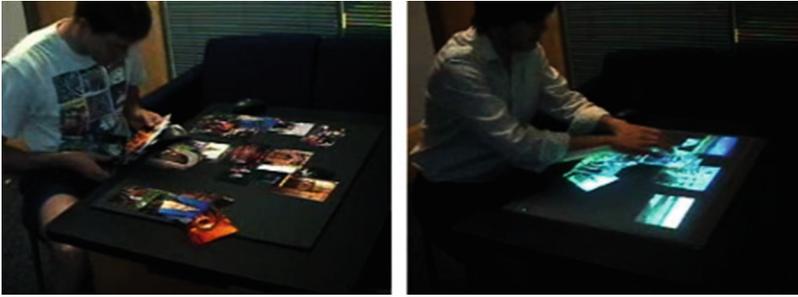
Fig. 1. Photo sorting tasks [9]

Fig. 2. Puzzle tasks [9]

in physical and digital makes participants' action different. This means that the differences in interfaces require different affordances.

Shin et al. showed the difference in devices needs different affordances [8]. In this research, they asked participants to play a VR-based application while a user wears an HMD, using *Hydra Controller*[6] and *MS Kinect* for operating the application. Figure 3
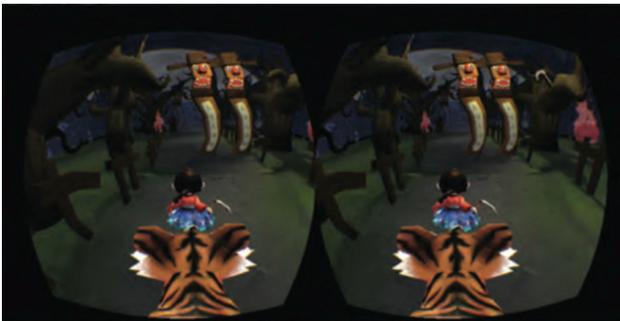

Fig. 3. The VR-based application with HMD [8]

shows the scene of the application. The results showed that different devices cause different problems even in the same VR-based application when a user wears an HMD. This means that in a VR-based application that uses an HMD, different devices may need different affordances.

3   A VR-Based Service and its Affordance Design

3.1   VR Frameworks, HMD, NUI Devices and Operating Objects in VR-Based Services

For developing a VR-based service, we adopted *Unity4.6.5*[7], that is a platform for easily creating a VR-based service. We also adopted *Oculus Rift*[8] as an HMD.

In our research, we selected two NUI devices: *MS Kinect* and *Leap Motion*. *MS Kinect* can track the entire whole body of a user, but the detection error is bigger than *Leap Motion.* Although *Leap Motion* makes accurate tracking possible rather than *MS Kinect*, it is only able to track a user's hands or small objects. By tracking a user's arm by *MS Kinect* or hands by *Leap Motion,* we arrange some objects in a VR world that performs the same movement as the user's arms or hands: we call them arm objects or hand objects. These arm/hand objects are used to operate the objects in the VR-based services.

3.2   An Overview of a VR-Based Service

When we considered VR-based services, we decided to create two types of services for extracting more insights from the experiment for them. For investigating extreme different cases, we have developed a VR-based service that is easy to understand named *Image Planet*, and the one is difficult to understand named *Shooting Game*. Also, for extracting more useful insights from the experiment, we created two types of operating method that are specialized and the general-purpose to the service: the former is called the *specialized operation* and the latter the *general-purpose operation*.

Image Planet as shown in Fig. 4 is a VR-based service developed as an example to offer simple operations. In this service, many images are floating over and rotating around a user. The background of the service is like the universe, so respective images look like stars. A user can select and expand the images by his/her operation on them, and also control the movement of the images through the panels for controlling them. As the *specialized operation*, a user uses a red line like a laser pointer attached to an arm object or a hand object to point the images or panels, and as the *general-purpose operation*, a user needs to move his/her arm/hand object over the images or panels. In this service, a user can use only his/her right arm or hand, and choose the proper pace to operate the objects that a user likes. This makes it possible to understand how to operate objects easily.

---

[7] Unity - Game Engine: http://japan.unity3d.com/, accessed 2015/12/22.

[8] Oculus Rift | Oculus: https://www.oculus.com/en-us/rift/, accessed 2015/12/26.

Fig. 4. Image planet (Color figure online)

Shooting Game as shown in Fig. 5 is the service developed as an example to offer complex operations. In this service, a user attacks and destroys enemies in the VR world, while guarding from the enemies' attack. In the *specialized operation*, a user operates a gun by his/her right arm object or right hand object, and attacks enemies. Also he/she operates a translucent wall for guarding from enemies' attack through their left arm/hand object. In the *general-purpose operation*, a user needs to move his/her arm/hand objects over enemies to attack them or over a panel that controls the wall to guard. This service needs to use both arms or both hands to attack enemies and guarding from them. It is also an important design intention to make a user hasty through the enemies' attack. The aim of the service is to investigate insights how complex operations influence affordance design.

Fig. 5. Shooting game

3.3 Affordance Design

For affordance design, we define three types of affordances, *inherent affordance*, *image affordance*, and *sentence affordance* based on the discussions described in [10].

Inherent affordance is an affordance that uses an object's shape, color, positions and so on, without directly offering images or sentences as the information for supporting the proper operation in a VR-based service, and this means that the intuition of a user is important for understanding how to navigate the service. This affordance is similar to the *perceived affordance* that Norman explained in [5], and is widely used in daily objects' design. In this case study, we implemented the affordance as follows; we designed an arm/hand object as the 3D models of an arm or hand, for making a user to understand that they can operate the arm/hand object by his/her own arm or hand. In particular, in *Leap Motion*, we arrange the object that represents the detectable spatially limited region for the interaction by the NUI device because the detectable region for the interaction of *Leap Motion* is very small. We also try to separate *inherent affordance* with the operating methods for the objects representing the affordances. *Inherent affordance* used for the *specialized operation* is called *high-inherent affordance,* and one used for the *general-purpose operation* is called *low-inherent affordance*. Figure 6 shows one example of *inherent affordance*.

Fig. 6. Inherent affordance

Image affordance is an affordance that uses images to offer the information for operating objects representing the affordance properly. An example used in the real world is a pictogram, like a sign used for the emergence exit. In our approach, we use this kind of image that explains how to use a user's arms or hands to operate the arm/hand objects in a VR-based service, and how to navigate the service. Different from *inherent affordance*, the arm/hand objects in a VR-based service are not the 3D models of arms or hands, but a white sphere. On the other hand, similar to *inherent affordance*, we classify *image affordance* into two types by operating method. The affordance with the *specialized operation* is *high-image affordance*, and one with the *general-purpose operation* is *low-image affordance*. Figure 7 presents an example of *image affordance*.
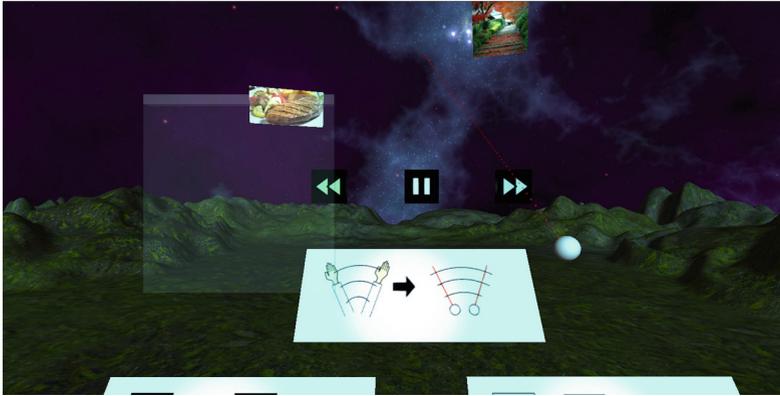
Fig. 7. Image affordance

Sentence affordance is an affordance that offers the sentences to represent the information for operating an affordance properly. The sentences are often used for explaining how to operate the affordance. In our approach, we offer the sentences to explain how to operate the affordance properly as *sentence affordance*. Similar to *image affordance*, arm/hand objects are represented as white spheres, and also we classify the affordance into two types of operating method. The affordance with the *specialized operation* is named *high-sentence affordance*, and one with the *general-purpose operation* is *low-sentence affordance*. Figure 8 shows an example of *sentence affordance*.

We have developed two VR-based services to offer two types of the operating method, and three types of affordances that support two operating methods. Also, in respective affordances, two NUI devices: *MS Kinect* and *Leap Motion* are adopted. Therefore, we totally conducted 24 patterns in the experiment of our approach as described in the next section.

Fig. 8. Sentence affordance

4   Experiments

In our experiment, we asked participants to perform tasks for respective VR-based services. Thus, there are 12 patterns for each, then we asked them to answer the questionnaires about whether they feel that these affordances are easy to understand to operate objects representing them. The tasks conducted in the experiment are as follows; in *Image Planet*, a participant selects and expands three photos that they like. In *Shooting Game*, a participant attacks and destroys three enemies while guarding from the enemies' attack. In our experiment, 17 participants whose ages are between 22 and 29 participated. The experiment for each person took about 2 h. We conducted the semi-structured interview for each participant after the experiment. Figure 9 shows one scene in the experiment.
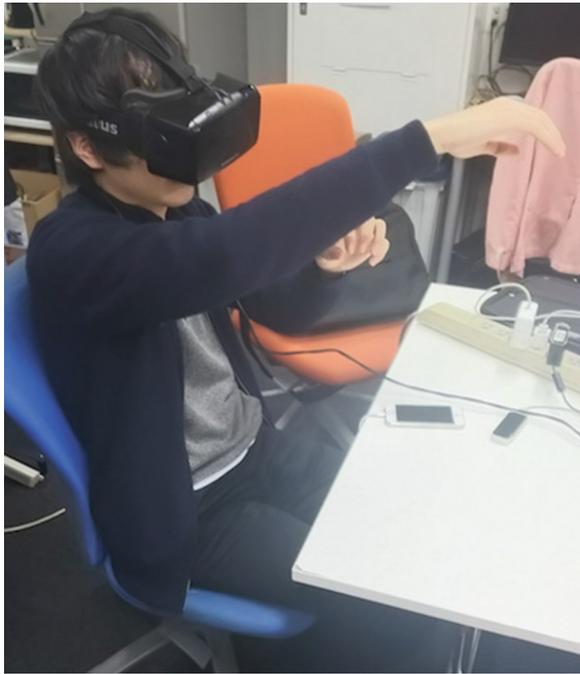

Fig. 9. A scene in our experiment

In the questionnaire, we asked the participants "*Did you think this affordance is easy to understand the meaning of an affordance in the service?*", and also "*Did you think this affordance is appropriate to express the meaning of an affordance in the service?*". The first question is the main question to extract participants' opinions explicitly for our research questions in this research, but we consider the question is not sufficient, because the question does not consider the whole cognitive load of a user. The cognitive load, that prevents a user from understanding the meaning of an

affordance, indicates not only the easiness to the understanding, but also the trouble-someness for the understanding, for example the troublesomeness of reading the sentences in an affordance. Therefore, we need to ask the second question that is the question for investigating a user's whole cognitive load.

Before conducting the experiment, we had a hypothesis that *image* and *sentence affordances* have no significant differences for respective NUI devices, and for *inherent affordance*, *Leap Motion* is preferred than *MS Kinect*. This is because that we thought images and sentences are very easy to understand, and there are no rooms to specialize the influence of the differences in NUI device, but *inherent affordance* needs a user's intuition for understanding the meaning of an affordance, thus *Leap Motion*, which offers a very accurate and good feedback is preferred. Also, we thought that the differences in NUI devices more influence on affordance design as operating the object that represents an affordance becomes more difficult. However, actually the result of the experiment shows that our hypothesis is not always true. The result shows that *image* and *sentence affordances* do not have much differences according to the differences appeared in NUI devices, and also *inherent affordance* is not significantly affected by the differences appeared in NUI devices. Moreover, the different types of VR-based services does not significantly influence the affordance design that needs to reflect the differences in NUI devices.

In the interview in the experiment, many participants said that *MS Kinect* is not good in terms of the precision, and *Leap Motion* is very good rather than *MS Kinect*, but also, some participants said "*I prefer MS Kinect because moving my arm is very intuitive for me, and it helped me to understand how to operate objects representing an affordance*", and "*I dislike Leap Motion because the detectable region for the inter-action is too narrow for me.*" This opinion expresses that some participants think that the precision is not important, but the detectable region for the interaction is more significant. On the other hand, some participants said "*I like Leap Motion because I need not to move my hand widely, and it is very intuitive.*" Similar opinions were appeared in most patterns in the experiment. For these reasons, we consider that the differences appeared in NUI devices are not a big factor to influence affordance design, but to influence the overall understanding how to operate objects representing an affordance, and also the understanding the meaning of an affordance differs in each individual. NUI devices use a user's gesture for the interaction, and this means that the intuition of a user is more important, where the intuition is differed for each individual largely. This insight can be used for the decision of what NUI devices should be used for respective VR-based services. We must consider not only what are the efficient and easy to use functions of NUI devices for operating an object representing an affordance, but also what functions each user prefers. Therefore, the usability and the comforta-bility need to be taken into account independently when designing good affordances. In the future, many types of inexperience NUI devices will be appeared, so we may be able to select NUI devices in terms of both usability and comfortability.

We also had a hypothesis about the easiness and the appropriateness for under-standing the meaning of an affordance. We thought that *sentence affordance* is prob-ably the most preferable in terms of the easiness for understanding the meaning how to operate objects representing affordances, and the next is *image affordance*, then *high-inherent affordance*, and the worst is *low-inherent affordance*. In terms of the easiness

of understanding the meaning of an affordance, a sentence has a significant advantage because it can explain how to operate an object representing an affordance concretely, and the next is an image. Comparing *high-inherent* and *low-inherent affordance*, the *specialized operation* is more understandable than the *general-purpose operation*, so *low-inherent affordance* was thought as the worst. As an early hypothesis, we thought that VR-based services with more complex operations have more significant influences on affordance design. However, the result showed that only some parts of the hypothesis are corrected. Similarly, as we expected, *sentence affordance* is the best, and *low-inherent affordance* was the worst. But, when actually comparing *image affordance* with *high-inherent affordance*, there is not so big difference between them. We consider that this is because the preference for *high-inherent affordance* can be changed in each individual. The understandability of *inherent affordance* depends on the intuition of a user, and it means that the individual difference influences the understandability of an affordance largely. We consider the reason of the result that most participants prefer *high-inherent affordance* used in the experiment in many cases.

In terms of the appropriateness of an affordance, our hypothesis is that *high-inherent affordance* may be the most preferable, and the next is *image affordance*, and *sentence affordance* is the worst. This is because we thought that a sentence that may cause serious troublesomeness in reading requires the heaviest cognitive load, and the cognitive load to understand an image is lighter than a sentence. Also, we assumed that *high-inherent affordance* is the lightest in terms of a user's cognitive load. However, the result is actually as follows; *sentence affordance* is the best when considering the gap between the easiness and appropriateness, but comparing the appropriateness in respective affordances, in all patterns, the effect of *sentence affordance* is almost the same as the effect of other affordances. We consider that there are two reasons for the result. The first is that the easiness is just one element to increase a user's cognitive load. Some participants said "*Sentence affordance allows me accurate understanding the meaning of an affordance rather than inherent and image affordances.*" The easiness for understanding the meaning of the *sentence affordance* avoids the decrease of its appropriateness. The second causes due to an individual's difference. For some participants, a difference between the easiness and appropriateness is not so important, but other participants said that they differ very much. Before the experiment, we considered that for every participant, the difference between the easiness and appropriateness is significant, so the aspect decreases the appropriateness of *sentence affordance*.

From the results in terms of the easiness and appropriateness, we are able to investigate the combination of affordances. It means that if easy to understand is more important, *inherent* or *image affordance* should be used to reduce the cognitive load, but when a developer expects that a user needs to understands precisely how to operate objects representing affordances in VR-based services, *sentence affordance* should be used, even when a user's cognitive load will be increased. Another way is to use these affordances together. As the result showed, how a user feels the difference between the easiness and appropriateness differs in each user. So offering multiple affordances at the same time and making a user to select his/her preferable affordance can be a promising way to design better affordances.

5 Conclusion

Recent development of a VR-based service that is based on common VR developing platforms and HMDs, which significantly increases the immersion of a VR world, requires NUI devices for enhancing the immersion. When using NUI devices, however, the proper operation of the devices should be offered, in particular, with HMDs. One approach to support the proper navigation is to offer an affordance, but there is also a new problem because there are a little researches to investigate the relationship between the differences in NUI devices and affordance design. Considering the future progress of NUI devices and the increase of their usages in many commercial VR-based services, more researches are necessary to extract sufficient insights for developing better VR-based services.

This paper is willing to offer a guidance how the differences in NUI devices influence affordance design. In this research, we have developed two VR-based services, two types of the operating method and three types of affordance for each ways of operations, and also we used two NUI devices: *MS Kinect* and *Leap Motion*. The result of the experiment to explore our approach showed that the differences in NUI devices may not significantly influence affordance design, but have significant effects on the overall understanding how to operate objects representing affordances in VR-based services. In particular, the understanding the meaning of an affordance differs in each individual. In an aspect of what affordance should be used for operating the objects representing affordances, when easy to understand is important, *inherent* or *image affordance* should be chosen to reduce a user's cognitive load, but when a developer wants a user to precisely understand the meaning how to operate objects representing affordances, *sentence affordance* should be adopted.

The conclusion in this research is that affordance design may not be influenced by the differences in NUI devices, but the difficulty of the operation for objects representing affordances in VR-based services, and the differences in NUI devices significantly influence the overall understanding of the meaning of an affordance.

In the next step, we need to investigate other types of NUI devices. More experiments may make us extracting more useful insights. Also, we need to investigate the reasons why people do not like some images and sentences or how the images and sentences are difficult to understand.

References

1. Corte, T.R., Marchal, M., Cirio, G., Lecuyer, A.: Perceiving affordances in virtual reality: influence of person and environmental properties in perception of standing on virtual grounds. Virtual Reality **17**(1), 17–28 (2013)
2. Linowes, J.: Unity Virtual Reality Projects. Packt Publishing, Birmingham (2015)
3. Macaranas, A., Antle, A., Riecke, E.B.: Three strategies for designing intuitive natural user interfaces. In: Proceedings of Extended Abstracts of the Designing Interactive Systems (ACM DIS) Conference (2012)
4. Miura, T., Urakawa, S., Isojima, M., Yu, J., Yoshii, A., Nakajima, T.: Natural user interaction requires good affordance when using with a head-mounted display. In: Proceedings of the Eighth International Conferences on Advances in Multimedia (2016)

5. Norman, D.: The Design of Everyday Things. MIT Press, Cambridge (2014). revised and expanded
6. Parisi, T.: Learning Virtual Reality: Developing Immersive Experiences and Applications for Desktop, Web, and Mobile. O'Reilly Media, Sebastopol (2015)
7. Sakamoto, M., Nakajima, T.: In search of the right abstraction for designing persuasive affordance towards a flourished society. In: Proceedings of the 9th International Conference on Design and Semantics of Form and Movement, pp. 251–260 (2015)
8. Shin, S., Kim, S., Chang, J.: An implementation of the HMD-Enabled interface and system userbility test. In: ISSSG 2014, pp. 183–193 (2014)
9. Terrenghi, L., Kirk, D., Sellen, A., Izadi, S.: Affordances for manipulation of physical versus digital media on interactive surface. In: HCI 2007 Proceedings, Novel Navigation, pp. 1157–1166 (2007)
10. Wesveen, S.A.G., Djajadiningrat, J.P., Overbeeke, C.J.: Interaction frogger: a design framework to couple action and function through feedback and feedforward. In: Proceedings of the 5th Conference on Designing Interactive Systems, pp. 177–184 (2004)
11. Yang, X., Pan, L.: Navigating the virtual environment using microsoft kinect. In: Proceedings of the 12th International Conference on Construction Application of Virtual Reality (2013)