Natural User Interaction Requires Good Affordance When Using a Head-Mounted Display

Takayuki Miura, Shoma Urakawa, Mitsuharu Isojima, Jianing Yu, Akihito Yoshii, Tatsuo Nakajima
Department of Computer Science and Engineering
Waseda University
Tokyo, Japan
emails: {t.miura, shoma_urakawa, mitsuharu_isojima, yu_jianing, a_yoshii, tatsuo}@dcl.cs.waseda.ac.jp

***Abstract*—Natural user interface (NUI) devices have become commercially common in various types of Virtual reality (VR)-based services such as video games and public attractions. Recently, a type of head-mounted display (HMD) such as Oculus Rift also attracts the industry with the possibility of developing new types of emerging VR-based services. In this paper, we report that appropriate affordances are necessary to use respective NUI devices, particularly when a user wears an HMD, which implies that different affordances are necessary for different NUI devices. We have developed a preliminary case study to use different NUI devices, where a user wears an HMD and navigates the interaction with the case study service. We conducted an experiment to investigate the relationship between the affordances and the NUI devices to extract useful insights to develop future VR-based services that use NUI devices and HMDs. The results suggest that it is important to consider the differences of NUI devices for the affordance design to navigate VR-based services.**

Keywords-NUI; HMD; VR; Affordance; Mental model.

I. INTRODUCTION

Recently, virtual reality (VR) technologies have revived because inexpensive and practical commercial head-mounted displays (HMD) such as Oculus Rift are used [12]. Thus, various types of VR-based services [13] are easily developed and can be commercially used. For example, in Japan, some VR-based attractions using an HMD have attracted many people [14].

To offer desirable interactive users' experiences, natural user interaction (NUI) devices such as Microsoft Kinect [15] or Leap Motion [16] are widely used. Many VR-based games have been developed, and they assume to use these NUI devices because these devices offer an immersive user experience through the natural interaction, which we use in our daily life, such as hand gesture and arm movement [2], with the virtual world without prior knowledge. However, it is assumed that using the current NUI devices causes a gap between the ideal situation and reality. In particular, using NUI devices with an HMD may cause a new problem. Yang and Pan have reported that MS Kinect fails to track a user's body when the user does not have enough experience with an HMD [11]. Additionally, Sabir, Stolte, Tabor, and O'Donoghue show that poor performance when using NUI devices has been found if the users have little practice with the NUI devices [5].

As claimed in [9], considering the affordance is effective for VR-based services when used with an HMD. Thus, we believe that those types of problems occur because different NUI devices require proper affordances to use them, particularly an HMD. 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 soon become a more serious problem. We must also investigate what types of affordance are appropriate. We discuss two types of affordance: inherent and augmented affordances, which we defined based on the inherent and augmented feedforward, as proposed in [10]. Furthermore, we would like to study whether different features of NUI devices affect the appropriateness of the types of affordance. Our research question is that a different NUI device requires a different affordance. This research question is the foundation in the research area of VR, HMD, and NUI devices, but only few studies were mentioned because of the rapid development of this research area.

In this paper, we have developed a simple VR-based service to investigate the above issues as a case study and demonstrate how we can design proper affordances for respective NUI devices. The extracted insights from our experiment are useful for designing future NUI devices and VR-based services.

This paper consists of the following sections. Section II shows the background of our study and the issues that we must investigate in our study. In Section III, we explain some issues of designing affordances and how we tackle these issues. Section IV presents some related work of this study. Section V illustrates a prototype service as a case study that we developed to investigate our research question. Section VI presents our conducted experimental design, and Section VII shows the results and discussions of the experiment. Section VIII presents the conclusion and future direction of our study.

 ISBN: 978-1-61208-452-7

II. BACKGROUND

A well-designed service requires a good mental model for navigation [3]. Traditional VR-based services typically use special and dedicated NUI devices that are developed only for the services [1]. Thus, the NUI devices usually fit well for the mental models of the services. However, in ubiquitous computing environments, we would like to adopt cheap and available NUI devices, such as MS Kinect and Leap Motion, to easily deploy the new VR-based services.

One of the potential pitfalls is that the NUI devices may cause gaps between the mental model and the assumption of the NUI devices, although it is desirable that these services can be used with any NUI device to increase the portability. For example, a user can navigate services with the movement of his/her entire body using MS Kinect, whereas he/she must assume to use only one hand to navigate the services when using Leap Motion.

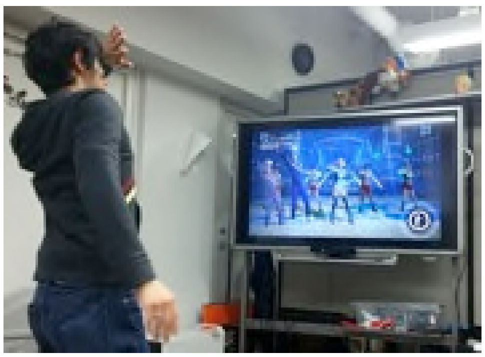


Figure 1. A Play Scene of Dance Evolution

Some concrete examples of using NUI devices are "Dance Evolution" [17] and "Nike+ Kinect Training" [18], where the video games are played on Xbox 360 using MS Kinect. "Dance Evolution" is a dance game, as shown in Figure 1, and "Nike+ Kinect Training" is an exercise game, as shown in Figure 2. In both games, the movement of each player's body is tracked by MS Kinect, and some players can compete for scores in the games. The players easily play the games by following the visual instructions on the screen, but when we assume that MS Kinect is replaced by Leap Motion, no player may play the games well because the presented visual instructions do not afford the operations of the games with Leap Motion.

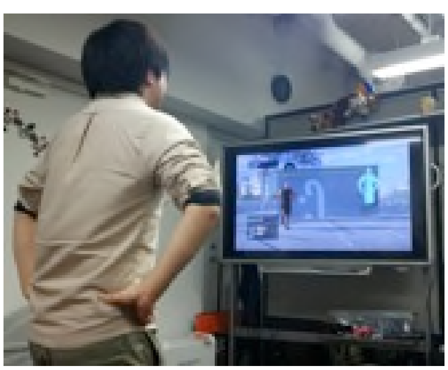


Figure 2. A Play Scene of Nike+ Kinect Training

This problem does not apply in these games because the games were developed to be currently operated only with MS Kinect. However, for general-purpose VR-based services that can be operated using several types of NUI devices, to overcome the issues of using various types of NUI devices for VR-based services, we require a new solution to use a VR-based service.

III. AFFORDANCES AND OUR RESEARCH GOAL

Our solution is to offer affordances to help construct the mental model to navigate the VR-based services. In this paper, we use the term "affordance" with the meaning of "perceived affordance". A typical affordance is the knob of a door; we usually know how we can open the door without any instruction when we look at the form of the doorknob.

In other words, affordances are the functions that provide the critical clues required for their proper operation. Additionally, affordances can be used to navigate human behavior [6]. However, the following research questions remain: whether respective affordances are necessary for different NUI devices, and what types of affordances must be offered.

In this study, based on [10], we define two types of affordances: inherent affordance and augmented affordance. The inherent affordance makes us understand how we use a VR-based service based on the UI elements' shapes, positions, etc. The definition of the affordance is widely used when designing daily objects [3]. The augmented affordance uses images or words to make us understand how we use a VR-based service. Investigating these two types of affordances enables us to extract useful insights when designing affordances for future VR-based services.

IV. RELATED WORK

Terrenghi, Kirk, Sellen, and Izadi show that each interface creates a different affordance in [8]. In this paper, the authors asked the participants to perform a puzzle task and a task to sort photos, where each task was performed with the two following methods: using physical puzzle pieces or photos and using their digital forms, which could be operated through a touch panel, as shown in Figure 3 and 4. The result of their study is that even with identical tasks, the affordances of the respective interfaces appear differently.

In [7], Shin, Kim, and Chang studied the usability of two devices in VR-based services with HMD. They asked the participants to play a race game and an action game as shown in Figure 5 and 6, respectively, with two different types of controllers: Hydra, which must be grasped to play the game, and MS Kinect. The results of the experiment show that even in identical games, the difference in controller devices affects the impression that the users feel.

Figure 3. Puzzle Task in Two Styles [8]

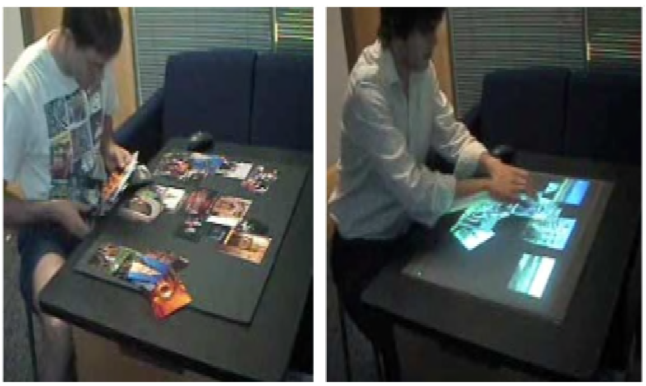


Figure 4. Sorting Photos in Two Styles [8]

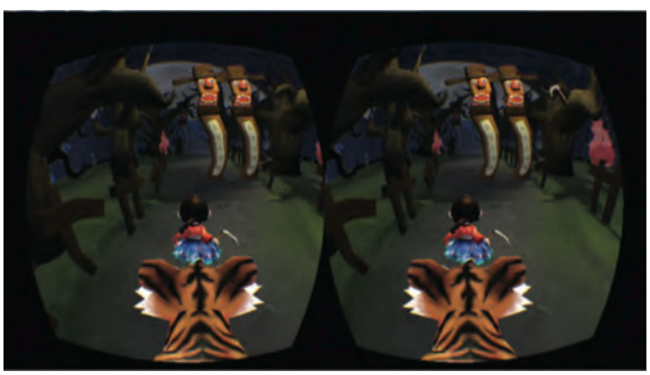


Figure 5. Race Game [7]

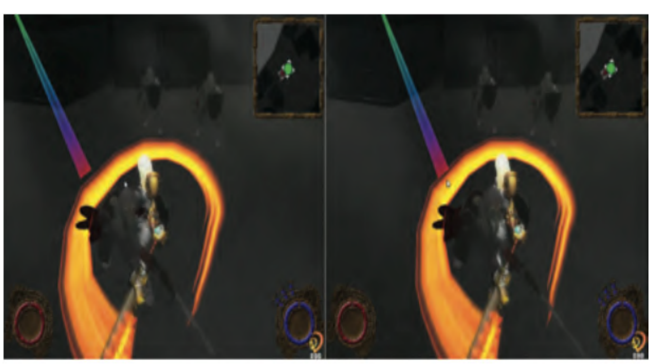


Figure 6. Action Game [7]

Thus, it is desirable to offer different affordances for each controller device to increase the usability of VR-based services when an HMD is used.

V. A CASE STUDY

We have developed a VR-based photo viewer service as a case study, which is illustrated in Figure 7, to demonstrate the proposed ideas.

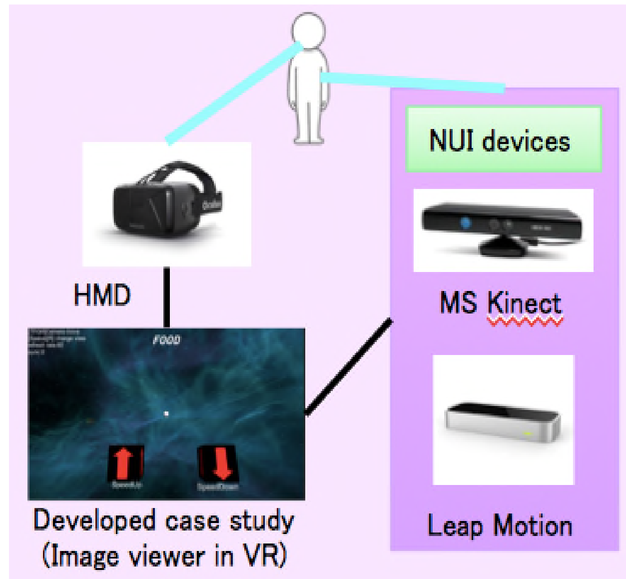


Figure 7. Construction of the image viewer we developed

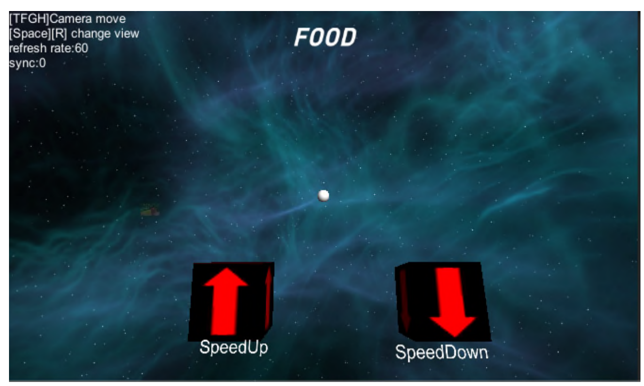


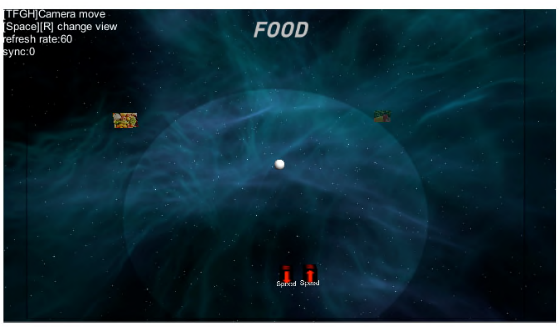
Figure 8. No Affordances for both MS Kinect and Leap Motion

Figures 8, 9 and 10 show the screen captures of the services that are displayed in an HMD. The small white sphere shown on the screen represents a cursor that a user can navigate using his/her motion. Several photos rotate around in user's sight. At the bottom part of the screen, there are two arrow objects, which indicate "Speed up" and "Speed Down". A user adjusts the photo's moving speed by putting the moving cursor on these objects. By putting the cursor on a photo for a time period, the size of the photo is enlarged.

Inherent

Augmented

Figure 9. Inherent and Augmented Affordances for MS Kinect



Inherent

Augmented

Figure 10. Inherent and Augmented Affordances for Leap Motion

The current case study assumes to use either MS Kinect or Leap Motion as NUI devices. We designed the inherent and augmented affordances for the respective NUI devices. The service can be used without presenting these affordances.

MS Kinect tracks the positions of a user's body. In this case study, the position of a user's right hand is captured to move the cursor. Conversely, Leap Motion tracks the positions of the joints of a user's hand. In this case study, the position of the back of a user' hand is detected to move the cursor. Figure 8 is a screenshot when no affordance is shown. Figure 9 shows the screens that represent the inherent and augmented affordances for MS Kinect, and Figure 10 presents those for Leap Motion.

In the experiment of this case study, we conducted user studies for the following five combinations.

(1) *No affordance + MS Kinect or Leap Motion:* the "Speed Up" and "Speed Down" objects are shown at the bottom of the screen.

(2) *Leap Motion + Inherent affordance:* The "Speed Up" and "Speed Down" objects are represented with smaller sizes than those in the service with no affordance, and the region that a user can move his/her hand is visualized.

(3) *Leap Motion + Augmented affordance:* A picture of a hand and the sentence "Right Hand here" are displayed at the bottom of the screen.

(4) *MS Kinect + Inherent affordance:* The positions of "Speed Up" and "Speed Down" objects are shown on the top and bottom right side of the screen.

(5) *MS Kinect + Augmented affordance:* A picture of a hand and the sentence "Use Right Arm Widely" are displayed at the bottom of the screen.

VI. EXPERIMENT DESIGN

In this study, we performed an experiment to investigate the above combinations. A participant selects two photos in each combination. The word "select" indicates enlarging the photos by putting the cursor on the photos for a period of time. In this experiment, we investigated the differences when there is an affordance or not and when two types of affordances are presented. We also investigated the situations when the participant knows what NUI device he/she uses and when he/she does not know which NUI device is used.

In this experiment, twelve participants with ages of 21-54 participated. Figure 11 shows one actual scene during the experiment.

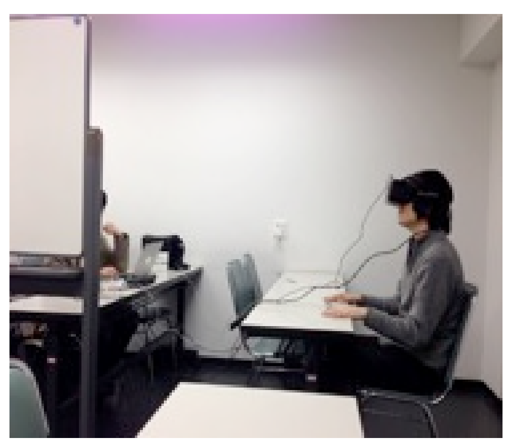


Figure 11. A Scene in Our Experiment

 ISBN: 978-1-61208-452-7

The experiment for one person took approximately 20-40 minutes. We conducted the semi-structured interview for them after the experiment.

VII. Design Implication

In the experiment, most participants basically showed positive attitudes for our proposed approach, and said, "*The presentation of the affordance helps me to navigate the service.*" In the section, we presented some more detailed answers in the interview conducted after the experiment described in the previous section.

We asked the participants "*Why did you consider that the affordance helps you?*" Some of them answered *"How, what, and where to move my hand or arm could be understood easily by the affordance"* Also, when asking a question "*Why did you prefer the inherent affordance of MS Kinect?*" Some participants who prefer the inherent affordance for MS Kinect answered "*The inherent affordance for MS Kinect is effective because the position of the arrow objects indicates the necessary action for my right arm*"

One opinion for the merit of the inherent affordance for Leap Motion is that the operative range of the device is easy to understand. However, some participants said, "*The inherent affordance is not good for me*", and they presented its reason as follows; *"I cannot understand the meaning of the affordance, and it confused me."* When comparing the inherent affordance for MS Kinect and Leap Motion, most participants said the affordance for Leap Motion is better. In terms of the augmented affordance, when asking the participants *"Comparing MS Kinect with Leap Motion, do you think which augmented affordance was easy to understand?",* Most participants answered *"The difference between the affordances for MS Kinect and Leap Motion is small because the augmented affordances for both devices are similar*"

In the opinion about the inherent and augmented affordance, many participants said the augmented affordance is better than inherent one, because words and images are understandable in an easier way. Additionally, the inherent affordance offers the better effect when the participants know which NUI device is currently used, whereas the augmented affordance has a better effect when they do not know which NUI device is used.

Ideally, a VR-based service should offer a proper mental model regardless of the NUI devices, but in reality, it is difficult to navigate the service without knowing which NUI device is used. Some participants said that the inherent affordance was good, but the others said that it was not good. We think that the variations are caused by whether they consider that they can intuitively understand the offered affordances. We hypothesize that the inherent affordance for Leap Motion was preferred to the affordance for MS Kinect because the participants easily understood the visualization of the affordance, whereas the movable region of their hands was limited in Leap Motion case. This result may indicate that some features of NUI devices affect the difficulty of understanding particular types of affordances. Additionally, most participants feel that the differences between the augmented affordances for MS Kinect and Leap Motion are small because the images and sentences in the augmented affordance are easily understood for most participants. Thus, the variations among individuals are small.

When comparing the inherent affordance with the augmented one, many participants claimed the augmented affordance is better. We analyzed the reason of it is that the question in the interview asked only whether the affordance was easy to understand or not. The inherent affordance uses only objects' shapes or positions, so in terms of human abilities for understanding the world, words and images have significant advantages because these are useful tools that can be used for explaining the world. We consider that asking *"Did you think this affordance is suitable for the guidance?"*, does not mean not only the difficulty of understanding but also the accident in understanding, and the results may differ.

Finally, the participants that knew which NUI device was used during the experiment preferred the inherent affordances because they could construct proper mental models before they found the affordances. Thus, they preferred the inherent affordance that required a lower recognition load. Additionally, the participants that did not know which NUI device was used preferred the augmented affordance because the affordance that helped them to easily construct the mental model is desirable for them.

VIII. Conclusion and Future Direction

Recently, HMD and VR have become attractive options to develop emerging new entertainment services and attractions. In particular, video games will use them to offer more immersive game experiences. NUI devices allow the games to be naturally interacted. However, there is no sufficient discussion on how to offer affordances for different NUI devices. This paper shows that each NUI device requires a different affordance when the VR-based services are used with HMDs.

In the next step, we will investigate a more systematic design guideline for affordances based on the insights of the current experiment. For example, as shown in [4], using multiple types of affordances together may offer a better result because of different effects. We will also attempt to discuss how to use other NUI devices to expand our current insights. In addition, we should consider other VR-based services to deepen our study.

In the future, many types of NUI devices will be available to develop advanced VR-based services. However, if the developers must consider different affordances for respective NUI devices, it may become troublesome.

References

[1] T. Mazuryk and M. Gervautz, "Virtual Reality History, Applications, Technology and Future", TR-186-2-96-06, Institute of Computer Graphics and Algorithms, Vienna, University of Technology, 1996.

 ISBN: 978-1-61208-452-7

[2] A. Macaranas, A. Antle, and E. B. Riecke, “Three Strategies for Designing Intuitive Natural User Interfaces”, In Proceedings of Extended Abstracts of the Designing Interactive Systems (ACM DIS) Conference, 2012.

[3] D. Norman, “The Design of Everyday Things”, The MIT Press; revised and expanded, 2014.

[4] D. Norman, “Affordance and Design”, http://www.jnd.org/dn.mss/affordances_and.html, accessed 09/29/2015.

[5] K. Sabir, C. Stolte, B. Tabor, and S. I. O’Donoghue, “The Molecular Control Toolkit: Controlling 3D Molecular Graphics via Gesture and Voice”, In Proceedings of the International Symposium on Biological Data Visualization, 2013, pp. 49-56.

[6] M. Sakamoto and T. Nakajima, “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, 2015., pp. 251-260.

[7] S. Shin, S. Kim, and J. Chang, “An Implementation of the HMD-Enabled Interface and System Usability Test”, ISSSG 2014, 2014, pp. 183-193.

[8] L. Terrenghi, D. Kirk, A. Sellen, and S. Izadi, “Affordances for Manipulation of Physical versus Digital Media on Interactive Surface”, HCI 2007 Proceedings, Novel Navigation, 2007, pp. 1157-1166.

[9] T. R. Corte, M. Marchal, G. Cirio, and A. Lecuyer “Perceiving affordances in virtual reality: influence of person and environmental properties in perception of standing on virtual grounds”, Virtual Reality vol.17, no.1, Springer-Verlag London, 2013, pp. 17-28.

[10] S. A. G. Wesveen, J. P. Djajadiningrat, and C. J. Overbeeke, “Interaction Frogger: A Design Framework to Couple Action and Function through Feedback and Feedforward”, In Proceedings of the 5th conference on Designing interactive systems, 2004, pp. 177-184.

[11] X. Yang and L. Pan, “Navigating the Virtual Environment Using Microsoft Kinect”, accessed 12/08/2015.

[12] http://www.oculus.com/, accessed 12/07/2015.

[13] http://share.oculus.com/, accessed 12/07/2015.

[14] http://weekly.ascii.jp/elem/000/000/278/278687, accessed 12/07/2015.

[15] http://www.microsoft.com/en-us/Kinectforwindows/, accessed 12/07/2015.

[16] https://www.leapmotion.com/, accessed 12/07/2015.

[17] http://www.xbox.com/ja-JP/Marketplace/SplashPages/danceevolution, accessed 12/07/2015.

[18] http://www.xbox.com/ja-JP/Marketplace/SplashPages/nike-kinect-training, accessed 12/12/2015.