Usability and Functionality Assessment of an Oculus Rift in Immersive and Interactive Systems Using Voice Commands

Valéria Farinazzo Martins[1], Paulo N.M. Sampaio[2],
Fernanda da S. Mendes[1], André Santos Lima[1],
and Marcelo de Paiva Guimarães[3(✉)]

[1] Computing and Informatics Program,
Mackenzie Prebisterian University, Sao Paulo, SP, Brazil
valfarinazzo@hotmail.com, nandayoko@hotmail.com,
andresantoslima@gmail.com
[2] Computing and Systems Graduate Program,
Salvador University (UNIFACS), Salvador, BA, Brazil
pnms.funchal@gmail.com
[3] Brazilian Open University(Federal University of São Paulo)/Master Program
of Faculty Campo Limpo Paulista, São Paulo, SP, Brazil
marcelodepaiva@gmail.com

Abstract. Virtual Reality has provided means to improve interface and user interaction with computational systems. This improvement has affected directly the degree of immersion and user interaction with computer-generated synthetic environments, through the utilization of human senses such sight, hearing, touch, smell and taste. This work aims to present a usability assessment and the functionalities of an interactive and immersive virtual reality system based on the utilization of the panoramic Oculus Rift device and interaction through voice commands. The virtual environment adopted to be integrated to the solution is the Google Street View through which user can visualize and navigate on. Besides assessment, this paper also introduces the hardware and software integration carried out, as the different phases of the development of the virtual environment.

Keywords: Virtual reality · Oculus rift · Voice commands · Immersion · Interaction

1 Introduction

Currently, the technologies applied in the development of application interfaces aim at turning users' experience to be as natural as possible [1–5]. For this purpose, the user's intuitive knowledge about the physical world has been considered in order to determine how this user interaction with the interface should be carried out. Virtual Reality (VR) is a relevant research area in the development of interfaces that explore users' experiences since it allows user interaction and their immersive participation with the simulation, either through visualization devices, touch or movement capture [6].

Indeed, VR does not constrain users only to simulation within real world, instead it amplifies the physical and temporal perception of users, enriching information and, thus, user's experience. For instance, it enables users to explore a volcano in eruption or navigate within a cell. VR has been a breakthrough within different knowledge areas: if we consider Engineering, for instance, reducing or avoiding the construction of mis-designed physical prototypes, or through the simulation of inaccessible or dangerous environments; when it comes to health environments, it is possible to provide realistic data analysis, patient monitoring and surgeries simulation. Clearly, the success of utilization of these environments depends on different factors, among them usability.

The goal of this paper is to assess usability and functionality of an Oculus Rift [7] in immersive and interactive environments using voice commands. For this purpose, an immersive application has been developed in order to provide navigation within Google Street View [8] – application that enables users to visualize streets and avenues. This application applies the user's head tracking in order to indicate which direction the user is visualizing, and user's voice to move around – a natural user interface.

Besides the usability assessment, this paper also presents the hardware and software integration carried out, as the development phase of the virtual environment. In order to achieve the proposed goals of this project the following activities were executed:

- Configuration, installation and deployment of the Oculus Rift device;
- Development of the voice-controlled immersive and interactive application;
- Execution of tests with the targeted audience, and;
- Analysis of the assessment results.

Some contributions related to the utilization of Oculus Rift can be found in the literature. Reiners [9] applied an Oculus Rift to have users immersed in an occupational health and safety learning environment. The experiments propose relevant activities and missions to be carried out by the user, focusing on his ability to identify high risk situations and react accordingly. Hoffman [10] applied this device as a support to help burn treatment. Once conventional methods (medication and therapies) are not effective for disinfecting wounds, for stretching burnt members and for stretching affected skin, which is painful for the patient. In this context, psychological intervention can also be helpful as a complement to standard analgesic medication prescribed. Therefore, patients can apply Oculus Rift in order to navigate within an ice customized environment while he is submitted to the painful procedures required. Since the patient's brain has been occupied navigating within virtual environment, he forgets the pain he is under. The paper presents an experiment carried out with an 11-year-old boy who had several electrical and flash burn injuries in his head, shoulders, arms and feet. He went through a 20 min session without VR in the first day, another VR session on the second day and a session without VR on the third day. On the day the patient had VR applied on the treatment he claimed the pain dropped from unbearable to mildly bearable.

The remaining of the paper is organized as follows: Sect. 2 describes the development phases of the application; Sect. 3 discusses the results obtained after running users tests; and, finally, Sect. 4 presents some final consideration of this work.

2 Development of a Voice-Controlled Immersive and Interactive Environment

The application adopted as basis for this implementation was Google Street View, which allows the panoramic visualization of a location in 360 degrees (horizontal) and 290 degrees (vertical), and also supports interaction using mouse and keyboard. The integration of this application with Oculus Rift and voice-commands provide users with an immersive interaction with any location anywhere in the world without leaving home. Therefore, it is also possible to visualize Google pictures of landscapes experiencing the immersive feeling of being there which is provided by the virtual environment. The voice commands were associated with action within the environment, such as go, back, back to search page.

The application has been developed using a prototyping technique [11, 12]. The development phases were iterative and in each development phase a new functionality was identified and aggregated to the system. Figure 1 depicts the development life cycle adopted:
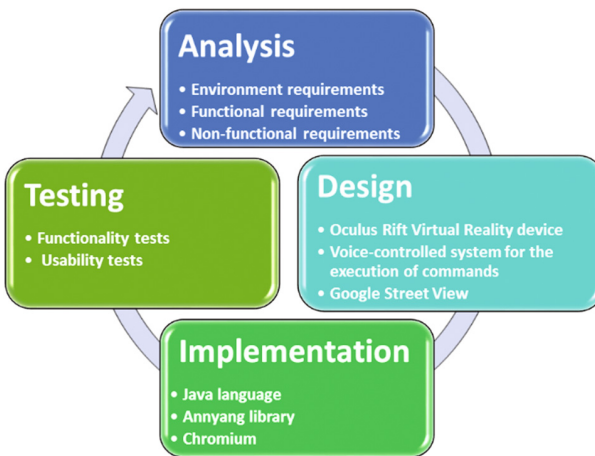
Fig. 1. The development life cycle

1. Analysis: definition of environment functional and non-functional requirements;
2. Design: definition of visualization and interaction device, and application design;
3. Implementation: the code is written and the hardware is configured;
4. Testing: testing functionality and usability.

2.1 Analysis

In order to specify the functionalities of the system (application and virtual environment) the following requirements were identified:

Environment Requirements: related to the physical environment where the application is being deployed:

– The environment should be kept in silence in order to ease voice capture by the microphone;
– User should be seat in a safe and steady place in order to avoid a possible fall since once the Oculus Rift is being deployed user loses sight and notion of the real physical space, and consequently loosing cognitive balance;
– The user cannot make sudden movements with his/her head in order to reduce sickness or dizziness.

- **Functional Requirements:** related to the functionalities of the system.

– The system should be executed on-line in order to access Google Street View;
– The addresses collected through voice-commands should be translated to a string which will be inserted in a text field;
– All the commands should be spoken in English.

- **Non-functional Requirements:** related to the response time when the application is executed. The interval in seconds should be as small as possible between the voice-command reception and the response presentation in order to provide a fast information retrieval and a straightforward presentation without interruptions, longer delays guaranteeing content integrity.

2.2   Design

Immersion is one of the important features of VR [6, 13], which can be provided by the deployment of a head mounted display. In this project the immersion is provided by the utilization of an Oculus Rift (Fig. 2). This device applies 360 degrees head motion tracking in order to allow users to look around as they would do in real life, providing a realistic experience.

Fig. 2. Oculus rift

The application was hosted on Web servers using technologies such as Java servlets and Javascript. The tools applied were:

- The Oculus Software Development Kit – SDK v0.3.2 [14];
- The version of Google Street View [8] applied was the Oculus Street View, which is a version designed for this device;
- The library Annyang [15] was applied for voice recognition. This Javascript library is oriented to web applications. Nevertheless, voice training is not applied in order to improve the degree of voice liability;
- The Google Chromium WebVR [16] was the Web browser applied, which is a version of Google Chrome oriented to the execution of Oculus Rift.

2.3   Implementation

This project was developed within 7 months, most of which was spent with studies on how to develop an application for Oculus Rift.

Figure 3 illustrates the operational flow for the developed application. Initially, the browser (in this case Chromium WebVR) presents a search page. In this page, there is a text field which will map voice commands into a text string. The commands are captured as voice through a microphone, as if user were typing it and clicking on the mouse. Thus, without the need to use hands to interact with the application, the user just pronounces the command and the local address, for instance, "Go to Wall Street New York". The voice command is captured by Annyang library and is sent to a servlet - a Java program that extends the capabilities of a server. This Annyang servlet processes information and presents it as an output text (string), which is then sent to another servlet that applies the Google Map's API called Geocode in order to map address into coordinates. At last, these coordinates are sent to Google Street View in order to present the required location on Oculus Rift. From this moment on, the user is free to navigate wherever he wants using voice commands. If user provides a non-specific address, for instance, "Go to Disney", then a pair of unknown coordinates is presented to the user.
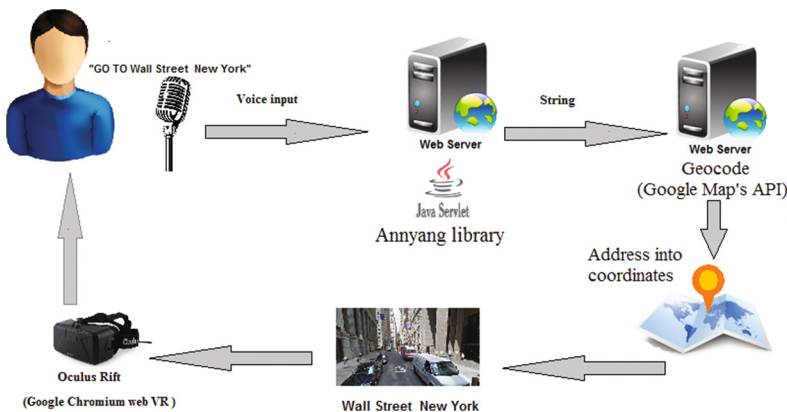

Fig. 3. Operational schema of the system

2.4   Testing

The functionality and usability tests were carried out with 15 users in three different locations: at the place where the group who developed the application lives, at the company one of the members of the group works and at the university where they study/work at. Initially, everyone filled out a pre-test questionnaire, which aimed at identifying the users' profile (age, sex, education attainment and technological literacy). In order to visualize the desired location, the user speaks some of the pre-defined commands, and the destination address. For instance, consider the command "go to" and the address "Wall Street New York". After this moment, the user navigated freely in the chosen street. In the end, all the users answered the post-testing questionnaire with questions related to functionalities and usability. The Likert scale [17] has been adopted (Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree) in the questionnaire. Currently, it takes the application around 2 s from the moment the user triggers his request (he/she speaks the desired action) until its execution, which is acceptable, according to the user tests. If the physical location where the experiment is being carried out presents considerable noise from other sources, this duration may surpass 5 s and, thus, increases the probability of erroneous recognition of words by the application.

Table 1 depicts the post-testing questionnaire, which is applied for the application usability assessment, considering user experience, navigation dynamics, and potential execution errors. The questionnaire was filled out after the utilization of the application.

Table 1. Application usability assessment

System Utilization Experiences					
Evaluate the utilization of the system according to the criteria below marking with an X					
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
This application has a graphical interface pleasant and understandable					
It is easy to understand what we have to do within the application					
The interface is very intuitive					
It was easy to learn how to use the application					

(*Continued*)

Table 1. (*Continued*)

System Utilization Experiences					
Evaluate the utilization of the system according to the criteria below marking with an X					
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The response time is enough (time between address recognition and The location rendering)					
The application works properly as for the recognition of my pronounciation					
When I hesitate (saying words like "mmmh", "aaahh"), the system ignores capturing only the words that define the address					
When I'm navigating on Street View the screen freezes					
When I turned my head the image gets blurred					
There are some color distortions when the 3D image is rendered					
It is not possible to visualize the image completely					
It is easy to navigate within the application using the Oculus Rift					
I would use the application to visit places before going there physically					
I felt dizzy when using the Oculus Rift					
Considerations about the application					
Positive aspects about the application:					
Negative aspects about the application:					
Suggestions about the application:					

3   Results

Through the results of the application of the pre-test questionnaire was possible to determine the users profile: people between 14 and 46 years old, within which most of the users were between 20 and 29 years old, representing half of the subjects. Most of them with undergraduate and graduate degree, using computers more than 15 h weekly and who already had experience with digital games. This profile also identified that the users already apply Google Maps and Street View to search addresses.

As for the obtained results of the application of the usability questionnaire (post-testing) it is possible to determine that:

- 54 % of the users strongly agree and 46 % of the users agree that the application has a pleasant and understandable graphical interface;
- 54 % of the users strongly agree, 33 % of the users agree and 13 % of the users were not sure if it was easy to understand what to do within the application;
- 47 % of the users strongly agree, 47 % of the users agree and 6 % of the users were not sure if the graphical interface is intuitive;
- 67 % of the users strongly agree, 20 % of the users agree, 6 % of the users disagree and 7 % of the users did not answer if it was easy to learn how to use the application;
- 34 % of the users strongly agree, 54 % of the users agree, 6 % were not sure and 6 % of the users strongly disagree that the performance of the system feedback is enough;
- 47 % of the users strongly agree, 6 % of the users agree and 34 % of the users were not sure if the application works properly in order to recognize the pronunciation;
- 27 % of the users strongly agree, 27 % of the users agree, 34 % were not sure and 12 % of the users did not answer if the system ignored correctly when they hesitated saying words like "mmmhh" and "Aaaah";
- 6 % of the users agree, 6 % of the users were not sure, 34 % of the users disagree, 41 % of the users strongly disagree and 12 % did not answer if when they are navigating on Street View the screen freezes;
- 2 % of the users agree, 12 % were not sure, 40 % of the users disagree and 36 % of the users strongly disagree that when they turn their head the image gets blurred;
- 6 % of the users strongly agree, 20 % of users were not sure, 37 % of the users disagree and 37 % of the users strongly disagree about having colors distortions when the 3D images are rendered;
- 12 % of the users agree, 12 % of them were not sure, 20 % of the users disagree and 40 % of them strongly disagree that they cannot visualize the images completely;
- 60 % of the users strongly agree, 34 % of the users agree and 6 % of the users were not sure if it was easy to get around within the virtual environment using the Oculus Rift;
- 60 % of the users strongly agree and 40 % of the users agree that they had the feeling of being physically present at the required location;

- 54 % of the users strongly agree, 20 % of the users agree, 20 % of the users were not sure and 6 % of the users disagree about using the application to visit places before going physically;
- 12 % of the users strongly agree, 27 % of the users agree, 6 % of the users not sure, 27 % of the users disagree and 28 % of them strongly disagree that they felt dizzy when using the device.

Through the previous information it is possible to conclude that for this group of users:

- The application has a pleasant and understandable graphical interface;
- It is easy to understand what it is possible to do with the application;
- The graphical interface if intuitive;
- It is easy to learn how to use the application;
- The response time of the application is enough;
- The voice recognition system works properly when the user hesitates saying words like "mmmhh" and "Aaaah" ignoring these words and capturing only important words such as the required address;
- The image does not freeze within Street View;
- When the user turns his head the image does not get blurred;
- There are no color distortions when the 3D image is rendered;
- It is always possible to visualize the image completely;
- It is easy to get around within the virtual environment using the device Oculus Rift;
- The system provides the feeling of immersion (being physically present) within the required location;
- Most of the users would use the system to visit the places before going there physically;
- The feeling of dizziness varies from person to person, some persons felt it other not.

4   Conclusions

The basic human senses (e.g., hearing, touch, taste, smell and sight) have been widely explored for the proposal of simulated virtual reality environments in order to provide immersion (where users feel integrated with the environment), interaction (user's actions affect the environment) and engagement (users feel engaged with the activities during simulation).

This work presented the development of the application that integrates the Oculus Rift with the Street View application. The application applies a voice recognition system to understand the required locations, contributing for a better human-computer interaction, since users' hands and body are free within the physical space to navigate within the virtual environment. Tests with 15 users were carried out and the results of this group were quite satisfactory.

According to the information collected from a group of users, it would be interesting the utilization of an audio software as for future works. Two options were proposed, the first one a pleasant background sound would be presented for relaxing

the user navigates within the virtual environment with the required location; in the second one, the background noise of a large city (cars passing by honking and some persons talking) in order to simulate the physical presence.

It is clear the challenge of deploying new technologies for developers. Although being innovative, this project also presented some challenges and drawbacks, such as the lack of support of browsers.

References

1. Jain, J., Lund,A., Wixon, D.: The future of natural user interfaces. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2011), pp. 211–214. ACM, New York (2011). http://dx.doi.org/10.1145/1979742.1979527
2. Delimarschi, D., Swartzendruber, G., Kagdi, H.: Enabling integrated development environments with natural user interface interactions. In: Proceedings of the 22nd International Conference on Program Comprehension (ICPC 2014), pp. 126–129 (2014). ACM, New York. http://dx.doi.org/10.1145/2597008.2597791
3. Del Ra, W.: Brave NUI world: designing natural user interfaces for touch and gesture by Daniel Wigdor and Dennis Wixon. SIGSOFT Softw. Eng. Notes **36**(6), 29–30 (2011). http://dx.doi.org/10.1145/2047414.2047439
4. Seow, S.C., Wixon, D., Morrison, A., Jacucci, G.: Natural user interfaces: the prospect and challenge of touch and gestural computing. In: CHI 2010 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2010). ACM, New York, pp. 4453–4456 (2010). http://dx.doi.org/10.1145/1753846.1754172
5. Carter, M., Allison, F., Downs, J., Gibbs, M.: Player identity dissonance and voice interaction in games. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY 2015), pp. 265–269. ACM, New York. http://dx.doi.org/10.1145/2793107.2793144
6. Azuma, R.T.: A survey of augmented reality. Presence Teleoperators Virtual Environ. **6**(4), 355–385 (1997). Earlier version appeared in Course Notes #9: Developing Advanced Virtual Reality Applications, ACM SIGGRAPH 1995 (Los Angeles, 6–11 August 1995), 20-1 to 20-38
7. Davis, B.A., Bryla, K., Benton, P.A.: Oculus Rift in Action, 1st edn, p. 440. Manning Publications, Greenwich (2015)
8. Google Street View. https://www.google.com/maps/streetview/. Accessed 02 July 2016
9. Reiners, T., Teras, H., Chang, V., Wood, L.C., Gregory, S., Gibson, D., Teras, M.: Authentic, immersive, and emotional experience in virtual learning environments: the fear of dying as an important learning experience in a simulation. In: 23rd Annual Teaching and Learning Forum, Perth, Australia, pp. 1–14, 30–31 January 2014
10. Hoffman, H.G., Meyer III, W.J., Ramirez, M., Roberts, L., Seibel, E.J., Atzori, B., Patterson, D.R.: Feasibility of articulated arm mounted oculus rift virtual reality goggles for adjunctive pain control during occupational therapy in pediatric burn patients. Cyberpsychol. Behav. Soc. Netw. **17**(6), 397–401 (2014)
11. Smith, M.F.: Software Prototyping: Adoption, Practice and Management, p. 216. McGraw-Hill, London (1991)
12. Sommerville, I.: Software Engineering, p. 816. Pearson, New York (2015)

13. Sherman, W.R., Craig, A.B.: Understanding Virtual Reality: Interface, Application, and Design. The Morgan Kaufmann Series in Computer Graphic, 1st edn, p. 608. Morgan Kaufmann, San Francisco (2002)
14. Developer Center (2016). https://developer.oculus.com/. Accessed 02 July 2016
15. Annyang! Speech Recognition that just works (2016). https://www.talater.com/annyang/. Accessed 01 Oct 2016
16. WebVR: Bringing Virtual Reality to the Web (2016). http://webvr.info/. Accessed 02 Mar 2016
17. Likert, R.: A technique for the measurement of attitudes. Arch. Psychol. **140**, 1–55 (1932)