How Do New Visual Immersive Systems Influence Gaming QoE?
A use case of serious gaming with Oculus Rift

Isabelle Hupont, Joaquin Gracia, Luis Sanagustín, Miguel Angel Gracia
Multimedia Technologies Division
Aragon Institute of Technology
Zaragoza, Spain
{ihupont, jgracia, lsanagustin, mgracia}@itainnova.es

Abstract—The recent introduction in the market of low cost yet high fidelity head-mounted displays with stereoscopic 3D perspective such as Oculus Rift opens the door to novel virtual reality experiences for gaming. However, QoE metrics and methodologies for this kind of platforms are still unexplored. In this paper, a comparative study on how these novel displays impact gaming QoE with regard to conventional 2D PC screens is presented. Through an experiment where we invited 22 users to play with a virtual forklift driving serious game in both environments, we demonstrate that Oculus broadly increases the sense of immersion in the 3D world as well as perceived game usability. Affective factors are also deeply impacted by this platform that shoots up amazement, astonishment and excitement levels. Nevertheless, a worrying aspect is the high percentage of persons that report feelings of nausea after wearing the goggles.

Keywords—gaming QoE; immersive virtual reality; serious games; human factors;

I. INTRODUCTION

Humans have been trying to create a display that mimics the experiences their eyes see and their brains interpret since cavemen scratched figures on rock walls. Nowadays, our culture is fascinated with virtual settings, indistinguishable from reality, which we can augment at our whim. In the pursuit of this ideal, researchers have developed various virtual reality (VR) systems. However, creating high fidelity virtual environments is a difficult endeavor.

Initial VR systems often failed to live up to the hype, and were considered to be low-quality and cartoonish. In particular, the visual elements were often jerky and did not respond quickly to the users movements [1]. Some early systems went some way to address these concerns, however surround screen approaches that produced highly immersive environments were exceptionally costly and certainly beyond the reach of all but the most dedicated of research teams, letting alone consumers [2].

Much of the hardware that is necessary to make VR a believable and cost-effective proposition has become much more accessible over the past few years. With the rise of high quality small displays, accelerometers and other hardware that is necessary to make VR effective, prices have come down, while at the same time, the devices have become many times more capable. Nowadays, the revolutionary introduction of low cost yet high fidelity head-mounted displays (HMDs) with stereoscopic 3D perspective such as Oculus Rift [3] or the upcoming Project Morpheus [4] opens the door to novel VR experiences and interaction paradigms for gaming.

However, Quality of Experience (QoE) metrics for this kind of displays are still undefined and unexplored. The few existing works that compare virtual games in PC screens vs. more immersive displays still focus on measuring task performance rather than more subjective QoE factors [5, 6]. The study by Lugrin et al. [7] offers more QoE-oriented results (namely about sense of presence and engagement), suggesting that users perceive overwhelming subjective preference for immersive versions of the same video game, but they performed their work with CAVE-like environments instead of HMDs.

On the other hand, low cost VR HMDs have been proven to be useful for various application domains. In particular, they have been shown to be a successful environment for learning [8, 9], for the visualization of engineering assemblies [2] or for interacting in virtual shopping [10]. More oriented to human factors, the work by Hoffman et al. demonstrated that Oculus Rift environment can elicit a strong illusion of arm presence and reduce pain in burn patients [11]. However all these studies still provide results in the form of performance-oriented metrics instead of QoE-oriented measures.

This paper explores how novel visual immersive systems affect gaming QoE. In particular, three different QoE factors have been studied: perceived presence (i.e. the sense of “being-in” the computer generated environment [12]), perceived usability and emotions. To that end, a user study (N=22) was performed. Users interacted with the same serious game (a forklift simulator) in two different environments: a PC with conventional 2D display and Oculus Rift. Self-reported data based on questionnaires were collected and findings about the impact of the environment in QoE are presented.
II. CASE STUDY: THE FORKLIFT SERIOUS GAME

A. Game Purpose and Description

The purpose of the forklift serious game is to enable students to acquire the basic skills for handling a forklift before start driving a real one, and thus reducing possible dangers or damages caused by inexperience and insecurity. The simulator is part of an official course for obtaining the forklift driver license. It consists of several training practices, from the most basic (such as the forklift startup protocol) to the more advanced (e.g. carry loads from compact shelves into a truck).

To increase the sense of realism, and thus favor the student to become familiar with every aspect related to the forklift, we prioritized to model the virtual world by mimicking as many details as possible of the real forklift and its usual environment: instrument panels, cabin, rear mirror, wheels, forks, types of shelves, loads and their positions, etc. (Fig. 1). Physics, materials and textures were also carefully included. An avatar representing the user is also present in the serious game.

As previous studies support the notion that motion fidelity is even more important than visual fidelity [13], especial attention has also been paid to mimic the interaction between the user and the virtual forklift. For that purpose, a gaming steering wheel with control buttons and pedals were used as interaction devices (Fig.1, bottom right), and the movement behavior and reactions of the real forklift was closely simulated in terms of speed, rotation axis, response delays, etc.

B. The Two Different Environments: PC vs. Oculus Rift

The simulator is currently compatible and running in two different environments: a PC with a conventional 2D computer screen and it also has a version for the Oculus Rift HMD (Fig. 2).

1) Conventional 2D display (PC)

In the conventional 2D display scenario, the simulated virtual world is projected onto a flat computer screen with a field-of-view (FoV) of 75 degrees. The refresh rate of the computer generated images is 60Hz. Users can change the 3D scene point of view, and therefore observe their virtual environment, by using the computer mouse. The screen size is 55 inches with a resolution of 1920x1080 pixels.

Since the user is sitting in front of the monitor at an average distance of 50-75 centimeters, he/she is fully aware of the real surrounding environment and can therefore “physically see” the interaction devices (wheel, pedals, buttons).

2) HMD Oculus Rift

In the Oculus Rift environment, the simulated virtual world is projected into a HMD by generating separate 2D images for each eye (left and right), which allows the perception of a final 3D view. This 3D view has a FoV of 100 degrees. Considering that the horizontal viewing angle of a human being is around 120 degrees [14], that means that Oculus Rift covers 83% of our FoV. This stereoscopic vision of the virtual world helps to measure more accurately volumes and to position more easily objects in the 3D space.

In that case, the forklift game has been developed with Oculus Development Kit version 2 and using a refresh rate of 75Hz. The HMD includes orientation and absolute position sensors so that the head movements are transmitted to the computer that changes the user point of view in the virtual scene correspondingly. The resolution of the device is 960x1080 pixels per eye.
As the HMD completely covers the user’s eyes, he/she is not able to watch the real surrounding environment. On the one hand, this fact allows to increase the sense of immersion but, on the other hand, the user loses any kind of link with the real world (he/she cannot see his/her real hands nor the “physical” interaction devices, etc.). To decrease the latter effect, self-awareness techniques are used. An avatar (i.e. a virtual representation of the user) accurately maps each of his/her movements to the virtual world: the position of the hands over the wheel, as well as the feet and the head pose when the user looks around (Fig. 3).

III. EVALUATION SETUP AND METHODOLOGY

The forklift serious game was already evaluated successfully in previous user tests from a learning (task-oriented) point of view. In this work the objective is different: the aim here is to study how the Oculus platform impacts three different QoE factors: perceived presence, perceived usability and emotions. This section presents the test protocol followed, participants description and the different evaluation metrics that were collected.

A. Study protocol

The test protocol was as follows. Participants were firstly welcomed and given general instructions by the test leader about the procedure of the experiment. Each user had then to fill in a short pre-test questionnaire, with questions about user profile (age, gender, profession and vision problems). Some additional information was also collected, related to the gaming behavior of the respondents: the frequency in which they use to play games, the type of video games they interact with (platform games, first person games, puzzle games, social games, etc.) and their previous experience with virtual immersive environments. Then the user was asked to perform a simple forklift driving task in a small virtual circuit, both in a PC with conventional 2D display and Oculus environments (in random order, c.f. Fig. 4). After each interaction in each environment, he/she was asked to answer another questionnaire compiling questions about the three QoE factors to be evaluated: presence, usability and emotions. A more detailed description of the questions and metrics used to evaluate these factors is provided in Section III.C. The test duration was around 20-25 minutes total per user.

B. Participants description

22 persons participated in the study. Out of the 22, 10 are male and 12 are female. The average age of the test participants is 35 (Standard Deviation SD=7.37). 8 users wear glasses or contact lenses, and used them during the test, but no other strong vision problems were detected. 7 of the participants are working or studying in the field of video games, multimedia or human-machine interaction and can thus be considered as more “expert” users.

Regarding previous experiences with video games and immersive environments, 23% of the participants play games frequently (once per week or more). Their favorite are puzzle (36%), platform (32%) and social (23%) games. 50% of the users had never tried immersive VR systems (such as VR goggles or CAVEs) before, 36% had used them occasionally and only 14% have frequent access to these platforms (all of them match the “expert” users category).

C. Evaluation measures and data analysis

After the interaction of the user with each platform (Oculus and PC screen), a mixture of standardized questionnaires was used to evaluate each of the three QoE factors targeted:

- In order to gauge perceived usability, users completed the System Usability Scale (SUS) [15]. The SUS is a well-established, standard usability scale for measuring perceptions of usability and has been shown to be reliable even with small sample sizes [16]. It consists of 10 questions in a Likert-scale that together, and in correspondence with the methods described by Sauro and Lewis [16], provides a single composite score reflecting the whole usability of the system. Moreover, we have also computed descriptive statistics (e.g. mean, SD, interquartile range IQR) from Likert-scales of individual SUS questions.

- To gather affective information, three constructs from the Differential Emotions Scale (DES) [17] were used. They are measured on a 5-point scale ranging from 0 (“Not at all”) to 4 (“Extremely”). These constructs are Fear (3 items - scared, fearful, afraid), Surprise (3 items - surprised, amazed, astonished) and Interest (3 items -
attentive, concentrating, alert). We also added an extra question using the same scale to ask the user to what extent he/she felt nauseated during the interaction. Once again, descriptive statistics from 5-points scales were obtained. To conclude the emotional part, the Pick-A-Mood (PAM) pictorial affective scale was included [18].

- Finally, to collect insights about the perceived presence inside the virtual environment users filled in the Presence Questionnaire (PQ) by Witmer and Singer [19]. This test measures in what degree users feel as they are actually part of the experienced environment. The questionnaire contains 19 questions in a 7-point scale, that allow to obtain sub-scores in terms of realism, possibility to act, quality of interface, possibility to examine and self-evaluation of performance by applying a scoring formula [19]. Also, a final score compiling the whole perceived presence of the system can be computed.

Next section presents the evaluation results obtained for each platform (Oculus vs. PC screen) and discusses the main findings resulting from the study.

IV. FINDINGS

As expected, perceived presence scores show up that Oculus increases the sense of immersion in the virtual world: the overall presence score in a 0-100 scale is 81 for Oculus while it only reaches 63 for conventional 2D display (Fig. 5). Moreover, Oculus PQ sub-scores are broadly higher for each category. Users especially perceive more realism (increment of 25%) and better possibilities to act (increment of 17%) and examine (increment of 21%) the virtual world through Oculus.

![Fig. 5. Detailed and overall perceived presence scores and confidence intervals with a 95% confidence level for each environment (values are normalized in a scale of 100 for comparison purposes).](image_url)

Fig. 5. Detailed and overall perceived presence scores and confidence intervals with a 95% confidence level for each environment (values are normalized in a scale of 100 for comparison purposes).

Figure 6 presents the results obtained for perceived usability. Once again, every individual usability aspect is perceived more positively using Oculus than traditional 2D screen, and so reflects the overall perceived usability score (81.5 for Oculus vs. 73 for 2D screen in a scale of 100). The only exception is “need of technical support”; probably due to the novelty of the Oculus device and the fact that users are not technically used to it yet: many of them asked for help to correctly put the device on their head and start playing. In any case, users reported that Oculus made the video game much easier to learn and use, and expressed a higher willingness to play it again in the near future. From these results, it can therefore be concluded that the level of immersion directly impacts perceived usability.

![Fig. 6. Detailed and overall perceived usability scores and confidence intervals with a 95% confidence level for each environment (values are normalized in a scale of 100 for comparison purposes).](image_url)

Felt moods and emotions are also strongly different after interacting with each environment (Fig. 7). Oculus help to stir users’ interest in the video game (from an interest score of 68 for 2D screen to 78 for Oculus in a 0-100 scale). The reported levels of surprise (amazement, astonishment) and excitement are also largely higher for Oculus: there is a score increment of 39% for surprise and of 23% for the excitement PAM pictogram. This finding may be partly explained by the recentness, freshness and “originality” of the system and more longitudinal studies would be needed to build strong conclusions about the affective impact of VR HMDs in gaming.

The most alarming point to highlight about the impact of Oculus in gaming QoE is related to the feeling of nausea users experience. The mean score for “feeling nauseated” factor in a scale of 100 is 36 for Oculus vs. 14 for 2D screen. Interestingly, we also found out a clear correlation between age and reported sense of dizziness (Fig. 8, top). Our older users were also the less familiar with technology, video games and immersive systems. As Oculus have not completely penetrated yet the big consumer market, it can therefore be envisioned that if HMDs become more and more popular in the next years, users will start to get used to them and therefore nausea will be reduced by training. This hypothesis seems to be in line with the relationship found between the feeling on nausea and the previous experience of the users with virtual immersive environments (Fig. 8, bottom). But, once again, longitudinal tests are needed to deeply study this effect. In fact, it may be that physiological factors related to age might also influence nausea, and that this cannot be reduced by training.
V. CONCLUSIONS, STUDY LIMITATIONS AND FUTURE WORK

In this paper a comparative study on how the novel low cost head-mounted Oculus Rift immersive display impacts gaming QoE with regard to conventional 2D PC screens has been presented. 22 users were asked to interact with the same forklift driving serious game in both environments. After analyzing the information reported through questionnaires in terms of three different QoE factors: perceived presence, perceived usability and emotions, interesting conclusions have been drawn. Oculus broadly increases the sense of immersion in the 3D world. Users perceive more realism and naturalness in the virtual scene itself and also in the way they move through the 3D environment and explore it. This perceived higher level of immersion has also been demonstrated to be directly proportional to perceived usability: users find out the serious game much easier to play, learn and use. Affective factors are also deeply impacted by this platform: amazement, astonishment and excitement levels shoot up after the interaction through Oculus. Nevertheless, the high percentage of persons that report feelings of nausea after wearing the VR goggles is worrying.

A set of open questions result from this work. Firstly, it should be quantified to what extent the positive emotions elicited by Oculus are a consequence of the novelty and originality of a device that is not yet part of our daily life. Secondly, the evolution of the sense of dizziness over time should be also monitored to check whether users are able to “get used” to the platform or not. In a near future we actually plan to drive longitudinal QoE tests to answer these questions. We also plan to introduce physiological sensors in our experiments (EEG and ECG signals) to obtain automatic measures of brain cognitive load and emotions.

Acknowledgment

This work has been partly funded by the Aragon Government’s Department of Education, University, Culture and Sports. The authors would also like to thank all the participants in the user tests.

References
