J Multimodal User Interfaces
DOI 10.1007/s12193-016-0236-5

ORIGINAL PAPER

Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift

Paola Salomoni[1] · Catia Prandi[1] · Marco Roccetti[1] · Lorenzo Casanova[2] · Luca Marchetti[2] · Gustavo Marfia[1]

Received: 14 April 2016 / Accepted: 20 December 2016

Abstract Research efforts, in the area of virtual reality, have mostly concentrated on the design and implementation of devices supporting user presence in immersive environments. Much can be done, however, to further improve the experience of wearing a head mounted display, in order to forget being at home while playing a computer game or navigating a virtual environment. In fact, with the widespread availability, in the near future, of head mounted displays (such as oculus rift) not only the power of gaming platforms will increase, but also the support to the construction of graphical user interfaces, which more closely resemble reality. Hence, although it may be exciting to navigate a virtual space wearing a head mounted display, user presence still depends on the functional fidelity of virtual objects. Simply said, presence perception is highly influenced by the opportunity of interacting realistically with the environment at hand. Hence, many graphical user interface (GUI) components should be re-engineered adopting a diegetic approach, as simple adaptations may result unnatural, awkward or out of place. Such fact, urges, for example, re-thinking how users receive and interact with information in virtual worlds. Along this direction of work, this contribution moves a step forward analyzing the role of diegetic interfaces. In particular, it describes the design, the implementation and the performance, to the eyes of its users, of three different diegetic interfaces, tested utilizing the oculus rift display. The results that are here presented, although referring to three specific cases, are of wider scope as they have been obtained for two GUI components that are adopted throughout many different entertainment virtual environments, namely a shell and a global control interface.

Keywords Diegetic · Stereoscopic three-dimensional interfaces · Oculus rift · Presence

✉ Catia Prandi
catia.prandi2@unibo.it

1 Department of Computer Science and Engineering, Alma Mater Studiorum University of Bologna, Mura Anteo Zamboni 7, Bologna, Italy

2 Studio Evil S.r.l., Via Barontini 20/A, Bologna, Italy

1 Introduction

We here present a study where entertainment virtual environment interfaces are re-thought in the light of the emergence of head mounted displays (HMDs). The traditional tension between diegetic and non-diegetic is approached through an analysis of their suitability to provide a natural and immersive environment to a user.

It is possible to recall many examples where a given technology has been available for long, raising low to mediocre levels of interest, exploding in popularity only at a specific point in time (e.g., tablets, gestural interfaces, etc.). This may certainly be the case of head mounted displays: it is possible to envision a near future where HMDs, now mainly employed for gaming purposes, may exponentially widespread and be used in many different applications. Such expectation is also built upon the knowledge that different HMDs are now available on the market.

For example, many HMDs have been available for a while now (the first VR headset, Forte VFX1 was announced at CES in 1994), however it is widely expected that Oculus Rift, released on March the 28th, 2016, will boost the use of such devices in the consumer market. Other expected strong players on this segment are: HTC Vive, developed by HTC and Valve Corporation, released on April the 5th, 2016, and the OSVR HDK2, an open source HMD for OSVR (Open Source Virtual Reality), designed by Razer and Sensics. Finally, Sony has scheduled the launch of the PlayStation VR on October the 13th, 2016.

A different kind of HMD is the Samsung Gear VR 4, a mobile virtual reality headset developed by Samsung in collaboration with Oculus. Such HMD is composed by a Samsung Galaxy device, that acts as the headset display and processor, and the Samsung Gear VR, which functions as the controller. This device is more accurate than its low cost counterpart: the Google Cardboard, developed by Google, but available for both Android OS and iOS.

The described industrial progress is very interesting, as recent studies corroborate the fact that the interactions supported by HMDs, i.e., user-tracking, stereoscopic visuals and wide fields of view, have a significantly higher impact on the presence perception of a user than other immersive system features [1]. Presence, however, is not only influenced by the immersion provided by hardware components, but also by software ones [2]. An important role is played by the graphical user interface (GUI), where graphical realism and functional fidelity to fictitious three-dimensional spaces is required to make a player feel at ease.

Now, the opportunity of soon being able to enjoy the use of HMDs has obviously excited game developers. Old titles are being adapted and new ones are being invented to be ready to hit the market [3,4]. However, while a wealth of research has been carried out on analyzing the impact of different hardware systems and stimuli on perceived presence, much less has, instead, been undertaken to understand the impact of the adoption of different GUI solutions. Such a type of analysis is, however, urgently needed, as the first problem many developers have already been, or will be soon, approaching is understanding how traditional desktop PC-based interfaces can be optimized to make the most of devices like Oculus Rift. In other words: how should interfaces be designed to allow users to take full advantage of HMDs?

As part of this process, for example, game developers do not only need to re-adapt the game itself, but also all the graphical components which serve the purpose of controlling the game (e.g., menus, etc.). Two trends have, to this date, emerged when approaching such a problem. A first approach has simply transferred controls, super-imposing them at some depth of the three-dimensional world. Although the easiest solution, such a type of approach risks breaking the effect of presence that is so hardly sought. A second approach amounts to mapping such components into a physical counterpart, taking advantage of the opportunities set by HMDs. These two trends have given birth to two different approaches: the design of non-diegetic and diegetic interfaces.

The concept of diegesis was defined by Aristotle as "the telling of the story by a narrator", in contrast to mimesis, defined as "imitation properly speaking" [5]. Over the years the definition of diegesis, and in particular of diegetic, has evolved, and has been reconsidered in different areas (such as cinema, storytelling and music) with the meaning of "within the narrative" [6]. In this context diegetic refers to artistic elements that are part of the action and can be perceived by the characters, rather than be just experienced by people that watch.[1] This definition of diegetic can be naturally adapted to the computer interface context [7]. To be more specific, diegetic interfaces represent scenes where all of the objects seen by a user belong to the virtual world and are experienced firsthand. As the player moves around the world, all of its fictional characters, objects, sounds and environments included into the diegetic interface are naturally perceived as being a part of the narrative. On the other hand, non-diegetic interfaces including audial and visual elements provide information concerning an entertainment virtual environment as a contextual but external piece of information [8,9].

Obviously, one might think that diegetic interfaces should be adopted in all possible GUIs. This would be highly desirable. In fact, while, for example, a plethora of PC video games are well-known to implement diegetic interfaces (some examples are: Far Cry 2, Dead Space, Metro 2033), games built for HMDs are expected to be as successful as they will be capable of providing realistic experiences and of better meeting the requirement of diegetic interfaces. Their use with stereoscopic three-dimensional interfaces must, however, be scientifically evaluated taking into account different user experience components [2]: (a) perceived 3D image, (b) game GUI quality, (c) game integration, (d) presence, and, (e) simulator sickness. Even when GUI components are designed adopting a diegetic approach, their shape, color, position or trajectories may be poorly perceived, degrading the general user experience. Hence, studies providing insights regarding how different GUI components may be appreciated through HMDs are eagerly sought, as they may pave the way to many different implementations.

On this stage, this paper contributes to the scientific literature of user experience with three-dimensional stereoscopic interfaces, describing how three original diegetic interfaces, immersed into a 3D world, have been welcomed by a set of users. A detailed analysis of the application of different diegetic/non-diegetic approaches applied to two specific types of interfaces is presented: two different menus and a weapon ammunition control. Although specific, the three interfaces that have been considered are representative of two well known classes of interfaces which are found in many different entertainment virtual environments: shell interfaces (i.e., interfaces which are typically encountered before entering the virtual world or during a pause) and global control interfaces (i.e., interfaces which provide status information). Due to this fact, the findings obtained during the analysis of these three interfaces can be put to good use in a large set of scenarios.

[1] http://dictionary.cambridge.org/dictionary/english/diegetic.

Finally, to the best of the authors' knowledge, first along such line of research, this work is relevant for the multimodal interaction community as it investigates the use of diegetic interfaces by relating graphical components and gestures, exploiting HMDs and external controllers and devices.

The remainder of the paper is organized as follows. Section 2 briefly surveys the concepts of sense of presence and immersion in a virtual reality setting, while Sect. 3 and its subsections describe the case study of interest, along with its design and implementation. Section 4 describes and discusses the obtained results, both from a quantitative and qualitative perspective, drawing general guidelines for the application of a diegetic approach to stereoscopic three-dimensional scenarios, whereas, finally, Sect. 5 concludes the paper.

2 Immersion and presence in 3D worlds: a review

Often confused as representing the same phenomenon, immersion and presence are in fact two different sides of the same coin [1]. In brief, immersion is defined as a technological quality of media, while presence as a psychological user experience. Due to their importance for this work, such concepts are further detailed in the following.

Technically speaking, immersion could be achieved removing sensations coming from the real (outer) world, replacing them with stimuli originating from a virtual environment [10]. With this view in mind, the more realistic the stimuli provided by technology (e.g., HMDs, near-eye displays, wide screens, projection wall systems, etc.), the better the feeling of being immersed. In essence, immersive devices should be: extensive, surrounding, inclusive, vivid and matching.

A common trend in literature presents presence as a complex and multidimensional concept, which can be generally explained as a sense of being there [11–13]. A more in-depth definition is that the sense of presence is the illusion of a *non-mediated experience*, meaning that a user fails to recognize that s/he is actually interacting with a piece of technology, while enjoying that virtual experience. Presence has been analyzed to the point of identifying the distinct dimensions which characterize it [14], specifically:

- **Spatial presence**, the sense of being physically located in a virtual environment. This variable is typically affected by the perceived quality of three-dimensional images and GUIs.
- **Self presence**, the capability of identifying, not distinguishing, the user from the avatar. This, instead, closely relates to the ability of performing actions (and perceiving the effect of such actions) in a natural way.
- **Social presence**, feeling the presence of other (virtual) users, who are perceived as different human beings. This depends on how realistically users can communicate among each other in the virtual world, on how they appear in different and distinct ways, on how they move and behave differently.

An interesting definition provided by Wirth et al. envisions spatial and self presence as part of a two-step process, where a person at first builds a spatial model of the surrounding environment, then accepting the mediated environment as his/her primary self-reference, through the construction of a perceived self-location and a perceived possibility to act [15]. As it will be clear in Sect. 3, this work concentrates on user experience in terms of spatial and self-presence, not considering, instead, social presence.

Now, the relationships between immersion and presence may be succinctly explained as follows: immersion is a technology-related goal sought by virtual environments (and the devices which support them), whereas presence is a psychological, perceptual and cognitive consequence of immersion. In simple words: presence can be thought as a psychological perception of being in or existing in an immersive environment.

A recent study has examined the effects of immersive system technologies on user experiences of spatial and self presence [1]. Grounding their analysis on the review of a substantial corpus of scientific contributions in this area, the authors concluded that increased levels of user-tracking, use of stereoscopic visuals, and wider fields of view of visual displays are significantly more impactful than improvements to other immersive system features, including quality of visual and auditory content. As the same authors point out, their study corroborates preceding findings by Reeves and Nass, which observed little correlation between fidelity of visuals and user attention, visual recognition memory (i.e., the ability to recognize items in their visual environment) or subjective experience. In other words, to this date, it appears that the most important technological factors, which most influence user experience, are those which support the most accurate action and interaction fidelity, rather than physical resemblance.

To conclude this section, it is worth noticing that while a wide scientific literature exists where virtual reality interfaces are evaluated in terms of immersion and presence [16–18], no experimental study (to the best of the authors' knowledge) has been conducted, so far, to assess the impact of diegetic interfaces with the use of a HMD such as Oculus Rift. In [19] the authors discuss the use of a HMD for 3D navigation in a virtual reality with a desktop PC, without investigating the benefits of immersion, and therefore not discussing the role of diegetic/non-diegetic interfaces.

Within this very vast field, the contribution of this work is that of assessing the improvement (if any), in terms of immersion, obtained adopting a diegetic approach for the design of three exemplars of two very widely used GUIs, two shell interfaces (SIs) and a global control interface (GCI).

3 Diegetic interfaces design case study

The approach followed for the design of the case study, as anticipated, has been that of working with a real system. While such an approach has provided the opportunity of assessing the user experience with a true virtual environment, it has also set a few difficulties.

The first one amounted to be able to work with the HMD of interest, Oculus Rift, as it was not available on the market at the time of writing. This first problem has been solved resorting to an alpha version of this device, Oculus Rift DK2, specifically released for experimentation and development.

The second problem, this time of technical nature, raised as a consequence of the unavailability, even for developers, of Oculus Touch, the built in tracking system that will be shipped together with Oculus Rift. It has been solved implementing an alternative tracking system, fully integrated with Oculus Rift, using the Leap Motion controller and the Razer Hydra devices (as described in Sect. 3.1).

Finally, a last difficulty amounted to carefully design the testing scenario, which included selecting the graphical elements of interest and the virtual environment for testing, as well as formulating the questions asked to the virtual world users.

The following subsections explain how all of the mentioned problems have been solved, starting from the problems of hardware nature and finishing with a presentation of the survey results obtained for the test-cases of interest.

3.1 Hardware and software system architecture

We designed the case study with the main goal of measuring immersion, testing with a set of users the results in terms of: (1) presence (spatial and self-presence), and, (2) comfort (i.e. effectiveness of control and simulator sickness) obtained adopting a diegetic and a non-diegetic approach in interface design.

The the field trials presented in this work have been conducted using:

- **Oculus Rift DK2**. This is a HMD, which acts as an input-output device. In particular, it displays a three-dimensional scene resulting from the processing of the inputs taken from its internal sensors and its external infrared camera.[2] Three-dimensional stereoscopic images are displayed on a 14 cm wide monitor, where images are non-completely overlapped in correspondence of the two human eyes (just as in reality, the left eyes sees an extra area on the left and the right one an extra area on the right). Image resolution is 1920 × 1080, 960 × 1080 per eye. Embedded sensors (gyroscopes, accelerometers and magnetometers) constantly follow head motions, adjusting the displayed image accordingly.
- The **Leap Motion** controller. This controller tracks hand gestures utilizing an optical tracking system.[3] It can track all 10 fingers of the two hands at the same time, with a fingertip position detection accuracy of approximately 0.01 mm at a frame rate of 300 fps, within an area of 24 cm^2. This is performed utilizing three infrared light emitters and two infrared cameras.
- The **Razer Hydra** device. This device tracks the positions of its two 3D magnetic trackers with a resolution of 1 mm and 1 deg.[4] In order to do so, its base station exploits a magnetic field which is only 20 times weaker than the earths natural magnetic field, allowing it to operate even in the presence of objects such as credit cards, hard disk drives and speakers.

As anticipated, due to the unavailability of Oculus Touch, Leap Motion and Razer Hydra, have been exploited to provide the users involved in the experiments the ability of touching and feeling the surrounding environment with their hands. In particular, the virtual environment and the support for interaction with the described devices has been implemented for the Microsoft Windows operating system, utilizing the C# programming language and resorting to the following software modules:

- Unity3D 4.6.1f1 PRO;[5]
- Oculus Rift DK2 libraries which include the:
  - Oculus Rift Runtime module 0.4.4-beta, to connect the hardware devices;
  - Oculus SDK 0.4.4-beta, to develop with the Oculus Rift;
  - Unity 4 Integration 0.4.4-beta, to integrate the virtual environment engine created with Unity3D.
- Leap Motion SDK version 2.2.3.25970 together with Unity Pro Asset version 2.2.0.23475, for the Leap Motion controller integration;[6]
- Sixense SDK, version 062612, and the Sixense Unity Plugin, version 1.0.2, for the Razer Hydra integration.[7]

2 https://www.oculus.com/en-us/dk2/.

3 https://www.leapmotion.com.

4 http://www.razerone.com/minisite/hydra.

5 https://www.unity3d.com.

6 https://developer.leapmotion.com.

7 http://sixense.com/windowssdkdownload.

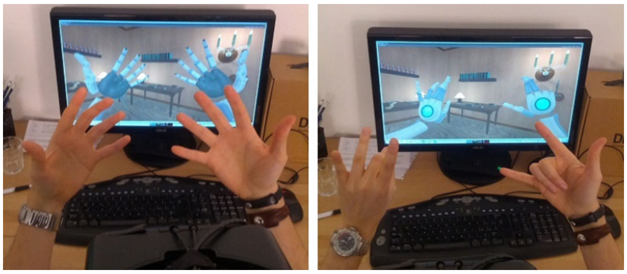

Fig. 1 Oculus rift and leap motion integration

An example of the integration of these components is shown in Fig. 1. In this particular figure, a user who utilizes Oculus Rift can see how his/her hands move in the virtual world set in front of him/her. In essence, the integration of these components let users experiment with a virtual world, moving, touching, grabbing and performing actions, ensuring a realistic haptic feedback.

3.2 Methodology

With the aim of assessing the diegetic and the non-diegetic interfaces in terms of immersion, we designed and carried out a user experience evaluation measuring presence and comfort. In particular, the global hypotheses we intended to assess is that a diegetic interface, related to the entertainment virtual environment and exploited with a HMD, is more immersive considering three dimensions: (1) presence (spatial and self-presence), (2) effectiveness of controls and (3) simulator sickness.

The experimentation campaign, although related to computer gaming, has been designed to assess, above all, the presence and comfort of user experience in a virtual environment. For this reason no gaming narrative has been developed, but a virtual space, i.e., a closed room, where, in isolation, users could navigate, without being influenced by: (1) the flow of a game, intended as “the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it” [20]; (2) the engagement due to gaming dynamics or by the presence of other users, defined as “a persistent and pervasive affective cognitive state not focused on any particular object, event, individual or behaviour” [21]. It is clear from these definitions that flow and engagement are directly related with, and consequence of, the game narration. Instead, our study focused only on those interface elements which are independent from a game. In this way, the assessment can be performed avoiding the influence of gaming flow.

Nonetheless, the graphical elements which have been selected for experimentation are representative of widely utilized components in many different games, in order to be able to produce an analysis, which may result useful for future development experiences. With this in mind, the choice fell on two common interface elements that can be found in many different games, a shell interface (SI) and a global control interface (GCI). In fact, shell interfaces can be easily found in any game, as they are displayed at the beginning of the game, before game start (or during its pause). Global control interfaces are also very common, as they are typically employed to return the status of an internal variable of the game (e.g., ammunition rounds, battery life, etc.). In addition, the shell interface and the global control interface were found particularly interesting as they differ in many ways. Many experiences exist, in video gaming, where diegetic approaches are taken for the design of GCIs, as it is very often possible to find ways of making them fit in virtual worlds. Trivial examples are given by the speed of a racing car, a mentioned before, but also the number of rounds available on a gun, etc. In essence, it is a common practice to provide such a type of information as realistically as possible within a virtual narrative. The same cannot be said, instead, for SIs, as they usually provide outer world information (e.g., rankings), which is more hardly displayed within the game flow.

In order to provide a significant comparison, non-diegetic versions of the selected user interfaces have been implemented along with the diegetic ones. In particular, two versions of a GCI interface, the ammunition status of a gun, have been implemented. Also two interfaces have been implemented, for each of the two SIs. For the first one, a non-diegetic version (a simple menu displayed before entering the virtual room) and a quasi-diegetic one (a player could interact with it with his/her hands, but the interface was still implemented as an object not belonging to the physical semantics of the virtual world). Finally, for the second interface fully diegetic/non-diegetic versions have been implemented, as it will be described in the following subsections.

Finally, the preparation of a set of survey questions which aimed at assessing two particular aspects of the two interfaces, their: (a) presence, and, (b) comfort. While many words have been spent discussing user presence, very little has been instead said about a user’s comfort. Such variable is of particular interest, as users interact with the virtual world utilizing Oculus Rift, Leap Motion controller and Razer Hydra device. Hence, an assessment of the comfort level reached with the controls has been carried out. In addition, Oculus Rift, as all other HMDs, may provoke a feeling of sickness. The personal sensitivity to simulator sickness differs from person to person (just as some people may feel a sense of nausea on a boat, while others do not). However, it has been shown that different controls can lead to different degree of simulation sickness, even forcing the users to abort the use of the headset [22,23].

Fig. 2 Virtual room

Fig. 3 SI1: non-diegetic version

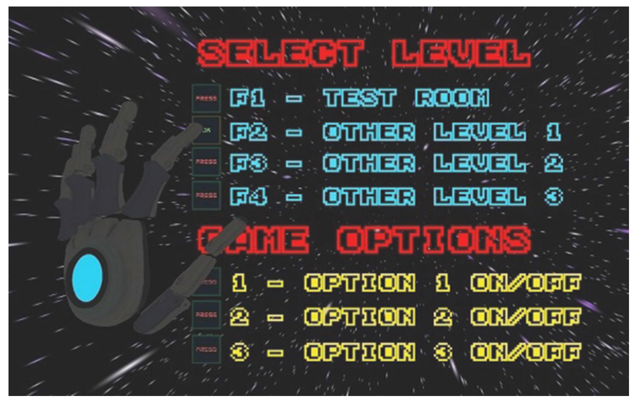

Fig. 4 SI1: quasi-diegetic version

3.2.1 Virtual environment setting

While on one side it could be convenient to test the proposed diegetic solutions on a real gaming system, as anticipated, the dynamics of a game may distract a user from providing an assessment, which well describes his/her experience. Presence and comfort, rather than other variables, remain those of interest. It is hence important to take design choices, which limit the interference of other components (e.g., engagement and flow), which may still be of interest for many designers, but which fall out of the scope of this work.

For this reason, a simple virtual space, i.e., a virtual room filled with some interactive objects, has been created as shown in Fig. 2. The floor, the ceiling and all walls, are all rigid bodies, which stop other objects and the player from passing through: no matter the physical force applied on them, they do not move. A user entering the room finds the shelves, a sword-holder and the lights on the walls, pedestals on the floor and a light with a rotating fan on the ceiling. While inside the room, a user can walk around, tactilely interact with all unanimated objects just as in a real world, without being disturbed by any active presence. Hence, a user entering the room can pick and hold any object (e.g., a sword, a gun, a tablet), and operate it following its normal utilization dynamics (e.g., swing the sword, fire the gun, switch on and use the tablet).

The three-dimensional models, textures and materials have been partially created from scratch and have been partially downloaded, charge and royalty free, from the TurboSquid 3D objects repository.[8]

3.2.2 Test cases

Starting with the SIs, the following graphical components have been implemented. The first SI, here denoted SI1, amounts to the interface shown in Fig. 3, in its non-diegetic form. As expected, SI1 reports the typical gaming information that may be found before entering a gaming arena. The quasi-diegetic version of SI1 is, instead, provided in Fig. 4. In this case the term quasi-diegetic is coined as the interface, in this particular case, does not perfectly fit the flow of the virtual world, while, nonetheless, a user can tactually interact with it touching and moving its words, as shown in the figure.

A different approach has been, instead, taken with the second shell interface, namely SI2. Its non-diegetic version is shown in Fig. 5, the menu simply serves the purpose of providing different options while in pause. With this interface, a user can reload a scene, load the main menu or quit. Figure 6 provides its diegetic version. Please note that, unlike the quasi-diegetic version of SI1, the diegetic version of SI2 fully integrates with the virtual environment. In fact, it is implemented on a virtual tablet, sitting in the virtual room. That tablet, in the virtual world, can be physically grabbed utilizing the Razer Hydra device held in one hand, in the real world, and operated with the fingers of the other hand as they are tracked by the Leap Motion controller.

Finally, the two versions of the global controller interface are provided in Figs. 7 and 8. It is important to notice that the GCI in Fig. 7 is not statically displayed, but shown when the user hits the space bar on the keyboard.

8 http://www.turbosquid.com.

Fig. 5 SI2: non-diegetic version

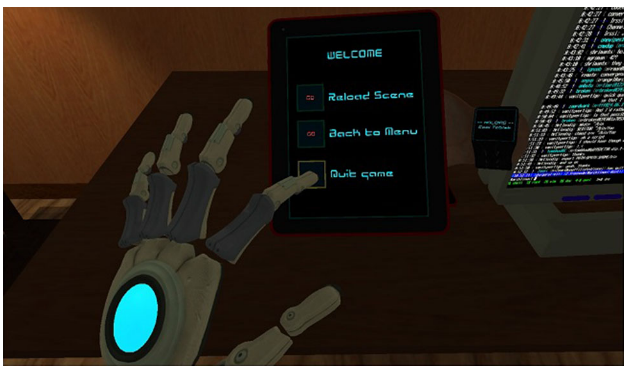

Fig. 6 SI2: diegetic version

Fig. 7 Global control interface: non-diegetic version

Just as for SI2, the diegetic version of this interface exploits a very popular device, almost all people are accustomed with: a wristwatch. While moving in the virtual world and playing with the gun, a user can check how many rounds are still available looking at his/her watch's display. In the real world this is implemented resorting to the Razer Hydra device, as shown in Fig. 8.

3.2.3 Participants

Ten participants, six males and four females, ages ranging between 20 and 56 (average 29.2), have tested the three interfaces.

The participants were young-adults, recruited voluntarily, with the status of students and employees, with in common a great passion for video games.

Fig. 8 Global control interface: diegetic version

The number of participants to the experiments (10 users) has been chosen as a trade-off between the necessity of acquiring sufficient feedback data to analyze the interface and the length and depth of the evaluation phase. The correctness of involving 10 users in the evaluation phase is also corroborated by the state of the art, where such number has repeatedly proven to be sufficient to discover over 80% of existing design problems [24,25].

With the exception of the oldest participant, 56 years old, who was used to play on a desktop computer, all other users normally played video games using one of the popular gaming consoles (e.g., Xbox, Nintendo, PlayStation, etc.), for an average time of 24 h a week (ranging from a minimum of 18 to a maximum of 32 h a week).

During the evaluation, we asked the participants to perform specific actions while interacting with all of the three interfaces, and in particular, the users have been requested to: (a) select an option on SI1 and SI2, and, (b) aim and shoot at a few targets and read the amount of ammunition left in the gun from the GCI.

3.2.4 Metrics of interest and survey construction

Field trials were evaluated through a questionnaire, developed building upon findings of previous studies in this area.

In particular, the survey questions have been formulated considering the work by Jennett et al. on measuring the experience of immersion in entertainment virtual environments [18]. The authors of this work presented two different questionnaires to evaluate and discuss a number of issues in defining and measuring the sense of presence and immersion. Such questionnaires have been cited by many recent papers in this field, however, most of these works present trials conducted on games in general, while this contribution concentrates on a specific virtual environment which could be explored with the use of a HMD [26–29]. Focusing on HMDs, a specific investigation of the benefits of three-dimensional stereoscopic vision has been carried out in [30]. In this study, a first group of users was asked to play with 3D stereo displays, while a second one used a traditional 2D monitor. Qualitative results were gathered using two different sets of questions: (1) a first set, derived from [18], related to the entertainment environment itself, and, (2) a second set specifically devoted to evaluate the overall experience of the 3D experience. The latter set introduced evaluation criteria for 3D games, but without focusing on the nature of the user interface (diegetic, non-diegetic). Takatalo et al. in [31] presented an evaluation of the user experience with 3D stereoscopic games resorting to a psychological research framework. This methodology evaluates both the technical game components (e.g., mechanics, story) and the psychological determinants of the user interface (e.g., cognitions, emotions, motivations). It is used by the authors to compare the user experience in playing the same game both with a 2D and a 3D output. Also in this case, the user experience analysis concentrates on user interactions with the game rather than focusing on the effectiveness of the user interface.

Constructing on the above-mentioned literature, a custom questionnaire has been built to evaluate the experience of immersion while comparing diegetic and non-diegetic approaches. The focus of the evaluation is clearly on the virtual interface. As anticipated, the virtual room, and the tools it contained, used for the test trials, was not placed inside an entertainment narrative, amounting instead to a space where users could move, performing specific actions (as details in Sect. 3.2.3).

Such limited context of action is reflected in the questionnaire, which limits its assessment to these specific components, rather than to the whole virtual environment. The 10 questions that compose the poll, derived from the ones proposed in [18] and adapted to interface evaluation, consider three main aspects:

1. **Presence**, measured in terms of how, to what extent, users really felt into a different place. In order to measure such aspect, 6 questions have been adapted from [18]. In particular, the questionnaire contains 4 positive (*I could interact with the room as if I were in the real world*, *I felt I was interacting with the room environment*, *I felt detached from the outside world*, *I moved around the room as I wanted* and 2 negative questions *I wanted to know what was happening around me*) and (*I was consciously aware of being in the real world whilst interacting*).
2. **Effectiveness of controls**, in terms of easiness of use. Such aspect has been resorting to the following three questions: *I was unaware I was using any control* (positive), *I found myself so involved I was unaware I was using controls* (positive) and *Controls were not easy to use* (negative).
3. **Simulator sickness**, which has been simply assessed with the following: *I felt sick while interacting with the interface*.

All questions have been evaluated using a five level Likert item: strongly disagree, disagree, neither agree nor disagree, agree, strongly agree.

In addition, as experimented in [18], a final question asked an overall impression concerning the presence feeling (*How immersed did you feel?*). This question, in analogy with [18], could be answered returning a value between 1 and 10, 1 corresponding to *not at all* and 10 to *very much*.

Finally, in order to assess all three interfaces for each of the two approaches (diegetic and non-diegetic), the questionnaire has been replicated 6 times for each user. In addition, to prevent the users from losing context while answering, they have been asked to fill the corresponding questionnaire section right after completing their interactions with each interface.

The questionnaire was designed to assess the hypotheses that a diegetic interface in a virtual environment, navigated with a HMD, is more immersive in terms of presence and effectiveness, despite the possibility of simulator sickness symptoms.

4 Results and discussion

4.1 Quantitative results

Survey questions are listed in Fig. 9. To provide a first general view of the results, Table 1 shows the values (average and the standard deviation) obtained with the responses to each of the 11 questions for both the diegetic and the non-diegetic interfaces, where those values were averaged over all ten players and three different interfaces. Numerical values have been obtained mapping *Strongly agree* into +2 and *Strongly disagree* into −2 for questions 1, 2, 4, 5, 7 and 8. This pattern has been, instead, reversed mapping *Strongly agree* to -2 and *Strongly disagree* to +2 for questions 3, 6, 9 and 10, as these specific questions have been implemented to reduce the acquiescence bias.

A graphical summary of Table 1 (only for questions from 1 to 10) is provided in Fig. 10, emphasizing that the diegetic interface (red line) received, in general, higher (better) scores compared to the non-diegetic one (blue line), in terms of presence, usability and motion sickness. Nonetheless, some of these values were negative even with a diegetic interface in use. Not surprisingly, the most negative numbers are observed with SI1 (Fig. 11), for which a quasi-diegetic interface, rather than a fully diegetic one, was implemented. However, negative values have been returned also for the other interfaces. A possible interpretation of this fact is that the user input equipment may have been unnatural to some players.

Nonetheless, it is worthwhile noticing that: (a) usability in general, and the use of controls in particular, generally

Considering the **interface you just tested**, please rate how far you would agree with the statements below:

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
I could interact with the room as if I were in the real world					
I felt I was interacting with the room environment					
I was consciously aware of being in the real world whilst interacting					
I felt detached from the outside world					
I moved around the room as I wanted					
I wanted to know what was happening around me					
I was unaware I was using any control					
I found myself so involved I was unaware I was using controls					
Controls were not easy to use					
I felt sick while I was interacting with the interface					

Please, now rate how immersed you felt:

	0	1	2	3	4	5	6	7	8	9	10
How immersed did you feel? (0 not at all immersed, 10 very immersed)											

Fig. 9 Interface assessment questions

Table 1 Diegetic/non-diegetic scores, average and standard deviation

Question	1	2	3	4	5	6	7	8	9	10	11
Non-diegetic interface											
Avg.	−0.97	−0.63	−0.97	−0.57	0.13	−0.20	−1.07	−1.23	0.20	0.43	4.33
Std. dev.	0.72	1.10	0.72	1.01	0.68	1.10	0.78	1.01	1.19	1.14	1.63
Diegetic interface											
Avg.	0.87	1.23	−0.60	0.53	1.00	−0.33	−0.40	−0.20	1.00	1.27	7.70
Std. dev.	0.78	0.50	0.81	0.90	0.64	1.09	1.00	1.06	0.79	0.94	1.39

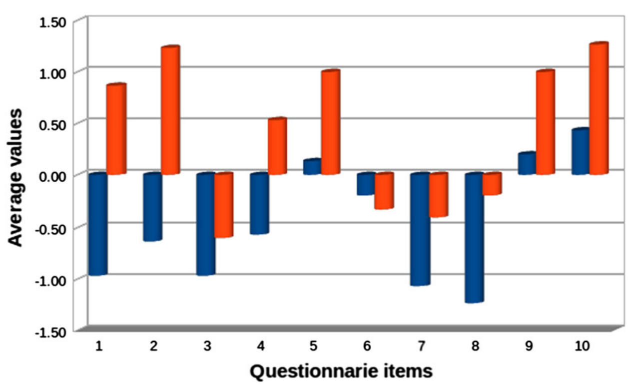

Fig. 10 Diegetic (*red*) vs. non-diegetic: a graphical comparison (reversing values of negative items)

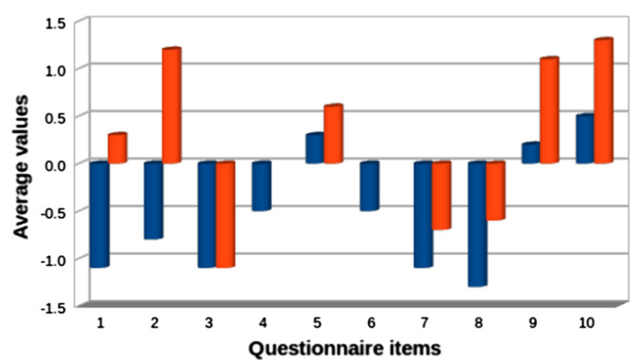

Fig. 11 SI1: quasi-diegetic (*red*) vs. non-diegetic (reversing values of negative items)

improved with diegetic interfaces, and (b) motion sickness has not been reported as a problem. What really strikes as odd instead are the responses to question 6 (*I wanted to know what was happening around me*), where the non-diegetic interface slightly surpassed the diegetic one. A possible explanation is that the question could have been misunderstood, with many players misinterpreting the use of *around me* (virtual or real environment?). Figures 11, 12 and 13 provide the average scores per each different user interface.

So far, we have provided some descriptive statistics that highlight the supremacy of the diegetic interface in term of user immersion. However, a statistically well-grounded evaluation is needed to compare the two different approaches. In particular, we are interested to understand whether the different feedback obtained from users experimenting the diegetic interfaces and the non-diegetic ones really resulted from different levels of immersion. Specifically, we carried out Friedman's test to confirm that the returned datasets are related to two different interfaces. The H0 null hypothesis and the alternative hypothesis follow.

H0: *No difference in terms of immersion emerges when using a diegetic or a non-degetic interface.*

H0: *Using a diegetic interface results different from using a non-diegetic one in terms of immersion.*

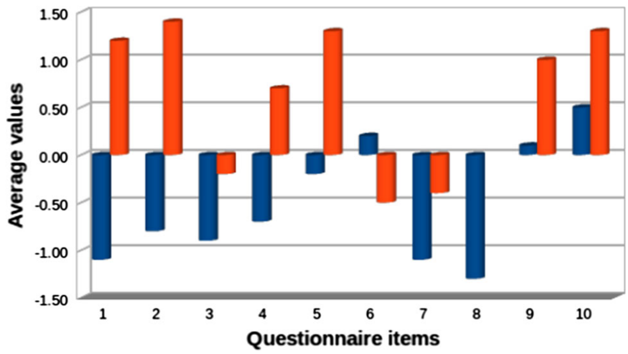

Fig. 12 SI2: diegetic (*red*) vs. non-diegetic (reversing values of negative items)

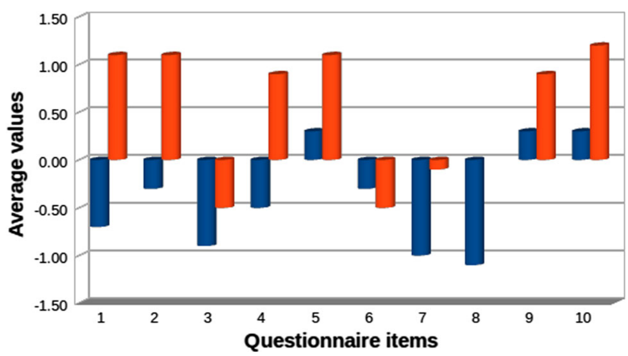

Fig. 13 GCI: diegetic (*red*) vs. non-diegetic (reversing values of negative items)

The run of this tests on our dataset let us reject the H0 hypothesis: the two interfaces really result different in terms of immersion. In fact, Friedman's test returned a critical threshold value of 3.8 for a p value = 0.05. For SI1, we obtained a Chi-square value of 6.4 and for SI2 and GCI a Chi-square value of 10, leading to the rejection of H0 confirming that the use of a diegetic interface has a significant (and not accidental) effect on the perception of immersion.

4.2 Qualitative results

In order to expand the extent of our results, we recorded the comments expressed by the 10 involved users, before and after their use of the interfaces.

In particular, we adopted the "thinking-aloud" method to gather data during the evaluation that was conducted, simply asking the test participants to think aloud [32].

From the qualitative point of view, four of the most interesting outcomes of the questionnaire are: (a) the impact of diegetic approach on presence, (b) the use of virtual smart objects to convey options and info, (c) the role of hardware controls in using the two interfaces, and, finally, (d) the motion sickness effect.

Quantitative data shows that users perceived the diegetic interface as more immersive than the non-diegetic one. In particular, considering all questions related to immersion, an average value for the non-diegetic interface of -0.46 (between disagree and neutral) vs. an average value for the diegetic one of +0.39 (between neutral and agree) was obtained. Such values have been computed from the 6 questions related to presence (values are inverted for the negative guards). It is worth noting that while the average value for the non-diegetic interface is negative, for the diegetic one this value is positive, but not very high.

A feeling of presence with the diegetic interface better emerged, instead, from the users' comments. Some of the users explicitly criticize the non-diegetic interface, labeling it as *unnatural, poorly immersive* and *fictitious*. For example, participant #1 commented:

> *The non-diegetic interface penalized natural interactions with the VR room.*

In a similar way, participant #3 expressed his perception of the non-diegetic interface:

> *…the contrast between the realistic view of the room and the menu floating in the air completely destroys the sense of immersion!*

Some users compared their experience with the two interfaces, expressing their preferences for the diegetic one. For example participant #7 affirmed:

> *In the first room*, i.e., the non-diegetic one, *you are aware, at all times, of not being physically there, while in the second room*, the diegetic one, *you feel in a real place*.

Finally some users emphasized how much they felt immersed using the diegetic interface defining it *immersive*, *natural* and *captivating*. This is the case of participant #8, who said:

> *With the diegetic interface I felt as thrown into another world*,

Controlling the menu and displaying the information through a virtual smart object was considered more natural and simple than using a traditional non-diegetic interface. For example, participant #2 said:

> *Interacting with everyday objects, like the tablet and the watch, makes me feel more immersed in the virtual context compared to the non-diegetic interface, the interaction was really more captivating!*

The opinion expressed by participant #7 on this aspect was instead intriguing and unexpected.

> *The use of smart watches and pads makes the interaction more natural and easy, but while pushing the player to interact with smart objects makes the expe-*

rience compelling I am wondering whether the use of too many external objects may render the game more difficult to play in general

The issue raised by participant #7 surely deserves further investigations as it poses a complex challenge to the designers of diegetic UIs, nonetheless we felt comfortable with the opinion of participant #4 who instead said:

Menus in the non-diegetic interface reset immersivity, while the tablet on the contrary enforces it!

As anticipated in Sect. 4, another controversial aspect that emerged is related to the use of controls. Three questions directly investigated the effectiveness of the control system and the average values (adjusting the negative guards) were -0.7 for the non-diegetic interface (almost disagree) and +0.13 for the diegetic interface (almost neutral). Such mild unsatisfaction is confirmed by qualitative results. Some users, as for example participant #9, expressed a general negative feeling while interacting with controls. He declared:

I felt uncomfortable with these control (sometimes even in the diegetic case).

Some users were particularly uneasy with controls using the non-diegetic interface. For example participant #5 affirmed:

In the non-diegetic room …controls are very difficult to use, awkward, definitively unnatural.

It is worth noting that the lack of immersion produced by controls had an impact on the whole evaluation. For example participant #4 said:

Sometimes controls penalized the naturalness of interactions and the sense of immersion,

and participant #5 affirmed:

…in some points, controls break immersivity!

All solutions were based on two different input devices, which have been integrated with Oculus Rift in order to improve the sense of immersion. Clearly, efforts in this direction were not sufficient, indicating that the headset requires an appropriate control system in order to be used at its best. Possibly, Oculus Touch will be the key element to solve both of these issues.

Finally, motion sickness was not perceived as a critical problem during trials, despite the fact it has been frequently cited as a potential problem in using a HMDs. Results from the questionnaire (*I felt sick while interacting with the virtual world*) state that motion sickness is stronger while using the non-diegetic interface (average value is -0.43, between neutral and disagree) than with the diegetic interface (average value is -1.27, between disagree and strongly disagree). Only one comment addressed this topic, it was made by participant #10, who concluded his comments concerning the non-diegetic interface saying:

Doing a wide rotation, I felt a bit sick.

5 Conclusion

This paper has presented an assessment of diegetic and non-diegetic user interfaces to be used with head mounted displays. In particular, among the many possible interface types, it has focused on two very widely used ones, shell interfaces and global control interfaces. The results that are here presented show a general appreciation of users for fully-diegetic interfaces (opposed to non-diegetic and even quasi-diegetic ones), while confirming the criticality of the role of the controls, even when diegetic approaches are adopted. A possible interpretation of this fact is that the user input equipment may have been unnatural to some players, thus indicating a further evaluation of the proposed interfaces should be carried out as soon as the Oculus Touch device will be released.

As a positive note, unexpectedly, motion sickness has not significantly emerged, showing that improvements that have recently been made with HMDs are limiting this unwanted effect. Indeed, we cannot definitely exclude that this result was influenced by the use of the Leap motion controller and the Razer Hydra device. Also in this case, a further evaluation, maybe using Oculus Touch, is desirable.

Considering the preliminary results obtained in this case study, we can assume that it will be interesting to investigate the same questions using a real game.

In conclusion, this work opens the path for new directions of work which include, but are not limited to: (a) an analysis of further interface type and solutions, and, (b) an extension of the proposed experiments to other control systems, such as Oculus Touch.

References

1. Cummings JJ, Bailenson JN (2016) How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. Media Psychol 19(2):272–309
2. Schild J, Bölicke L, LaViola JJ Jr, Masuch M (2013) Creating and analyzing stereoscopic 3d graphical user interfaces in digital games. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 169–178, ACM
3. Kuchera B (2015) Oculus offering a $10 million budget for indie games on the oculus rift. http://www.polygon.com/2015/6/11/8767033/oculus-offering-a-10-million-budget-for-indie-games-on-the-oculus-rift. Accessed 15 Oct 2016
4. Te Z (2015) Oculus rift vs. morpheus vs. vive VR. http://www.gamespot.com/articles/oculus-rift-vs-morpheus-vs-vive-vr/1100-6427162/. Accessed 15 Oct 2016
5. Kirby JT (1991) Mimesis and diegesis, foundations of aesthetic theory in plato and aristotle. Helios 18(2):113–128

6. Cecchi A (2010) Diegetic versus nondiegetic: a reconsideration of the conceptual opposition as a contribution to the theory of audiovision. Worlds of audio vision. http://www-5.unipv.it/wav/pdf/WAV_Cecchi_2010_eng.pdf. Accessed 15 Oct 2016
7. Galloway AR (2006) Gaming: essays on algorithmic culture, vol 18. University of Minnesota Press, Minneapolis
8. Soyluçiçek S (2016) The association of typography with form and content in digital games. Glob J Humanit Soc Sci, 3:216–223
9. Saunders K, Novak J (2012) Game development essentials: game interface design. Cengage Learning, Boston, Massachusetts, United States
10. Mestre D, Fuchs P, Berthoz A, Vercher J (2006) Immersion et résence. Le traité de la réalité virtuelle. Ecole des Mines de Paris, Paris, pp 309–338
11. Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. J Comput Mediat Commun 3(2):0. doi:10.1111/j.1083-6101.1997.tb00072.x
12. Hartmann T, Klimmt C, Vorderer P (2010) Telepresence and media entertainment. In: Immersed in media: telepresence in everyday life, pp 137–157
13. Tamborini R, Bowman ND (2010) Presence in video games. In: Immersed in media: telepresence in everyday life, pp 87–109
14. Brown E, Cairns P (2004) A grounded investigation of game immersion. In: CHI'04 extended abstracts on human factors in computing systems. ACM, pp 1297–1300
15. Wirth W, Hartmann T, Böcking S, Vorderer P, Klimmt C, Schramm H, Saari T, Laarni J, Ravaja N, Gouveia FR et al (2007) A process model of the formation of spatial presence experiences. Media Psychol 9(3):493–525
16. Tamborini R, Skalski P (2006) The role of presence in the experience of electronic games. In: Playing video games: motives, responses, and consequences. Routledge Taylor & Francis Group, UK, pp. 263–281
17. Skalski P, Tamborini R, Shelton A, Buncher M, Lindmark P (2010) Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. New Media Soc 13(2):224–242
18. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, Walton A (2008) Measuring and defining the experience of immersion in games. Int J Hum Comput Stud 66(9):641–661
19. Santos BS, Dias P, Pimentel A, Baggerman J-W, Ferreira C, Silva S, Madeira J (2009) Head-mounted display versus desktop for 3d navigation in virtual reality: a user study. Multimed Tools Appl 41(1):161–181
20. Csikszentmihalyi M (2013) Flow: the psychology of happiness. Random House, Inc, New York
21. Schaufeli WB, Salanova M, González-Romá V, Bakker AB (2002) The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. J Happiness Stud 3(1):71–92
22. Llorach G, Evans A, Blat J (2014) Simulator sickness and presence using hmds: comparing use of a game controller and a position estimation system. In: Proceedings of the 20th ACM symposium on virtual reality software and technology. ACM, pp 137–140
23. Moss JD, Muth ER (2011) Characteristics of head-mounted displays and their effects on simulator sickness. Human Factors J Hum Factors Ergon Soc 53(3):308–319
24. Faulkner L (2003) Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 35(3):379–383
25. Hwang W, Salvendy G (2010) Number of people required for usability evaluation: the 10±2 rule. Commun ACM 53(5):130–133
26. Brockmyer JH, Fox CM, Curtiss KA, McBroom E, Burkhart KM, Pidruzny JN (2009) The development of the game engagement questionnaire: a measure of engagement in video game-playing. J Exp Soc Psychol 45(4):624–634
27. Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence Teleoperators Virtual Environ 7(3):225–240
28. Boyle EA, Connolly TM, Hainey T, Boyle JM (2012) Engagement in digital entertainment games: a systematic review. Comput Hum Behav 28(3):771–780
29. Nijhar J, Bianchi-Berthouze N, Boguslawski G (2011) Does movement recognition precision affect the player experience in exertion games?. In: International conference on intelligent technologies for interactive entertainment. Springer, pp 73–82
30. LaViola JJ Jr, Litwiller T (2011) Evaluating the benefits of 3d stereo in modern video games. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 2345–2354
31. Takatalo J, Kawai T, Kaistinen J, Nyman G, Häkkinen J (2011) User experience in 3d stereoscopic games. Media Psychol 14(4):387–414
32. Lewis C (1982) Using the thinking-aloud method in cognitive interface design, report rc 9265. IBM Research, Yorktown Heights