Spatial Immersion versus Emotional Immersion, Which is More Immersive?

Chenyan Zhang, Andrew Perkis, Sebastian Arndt
Department of Electronic Systems,
Norwegian University of Science and Technology (NTNU) (Trondheim, Norway)
chenyan.zhang@ntnu.no, andrew.perkis@iet.ntnu.no, sebastian.arndt@ntnu.no

***Abstract*—In this paper, we defined and measured two types of immersion, namely, spatial immersion and emotional immersion, on flat 2D screen displays and using a 33-item questionnaire. Our results show that emotional immersion is significantly more immersive than spatial immersion in terms of sense of "being there", time perception, realism, sense of engagement, emotional aspects, sensory cues, etc. Spatial immersion is almost as immersive as emotional emotion in terms of attention and image motion, and spatial immersion is more immersive than emotional immersion in terms of spatial dis-orientation. Our results also show that there are individual differences in the perception of immersion. Finally, we linked parameters of measurement with QoE influencing factors, in an aim to bridge the gap in the QoE assessment and modelling of immersive experience in storytelling.**

Keywords—Spatial immersion; Emotional immersion; QoE influencing factors; Immersive Storytelling

I. INTRODUCTION

With the notion of Immersive Media Technology Experiences (IMTE) [1] gaining increasing momentum, the quality assessment and modelling of IMTE becomes an urgent task, due partly to the fact of the increasing pervasiveness of immersive media and technologies, partly to the enhanced design and innovation requirements from the producers' side and the heightened user experience, usability and enjoyment requirements from the users' side. Among the many influencing factors of QoE, the ability of a multimedia system to immerse the users or audience is an important defining factor in designing and evaluating the multimedia systems, and this is particularly true in digital storytelling. However, the extent to which a storytelling content can be marked as "immersive" and the criteria of how and why an immersive scenario is defined have not reached a definitive consensus, to the extent that even attempts to clarify these gaps seem sparse and few [2]. This is particularly true given the fact that in immersion literature only a limited amount of measurements are discussed without distinguishing what the nuanced types and degrees of immersion are there in the reality.

In this paper we are distinguishing between two types of immersion: spatial immersion and emotional immersion. We are not trying to do a mutually exclusive distinction or a full range classification here. What we are trying to do is to pick up some influencing factors of immersion (such as spatiality and emotion) to experiment with and discuss them in terms of their impact on immersive experiences.

Spatial immersion refers to the type of immersion triggered and maintained by the spatial qualities of the virtual environment. In spatial immersion, the immersive effect of the virtual environment can be achieved by the deliberate manipulation of a few spatial compositions of the scene, such as swift zoom-in and zoom-out, abrupt change of camera angles, or the whirling sensation of on-the-fly sky-diving shots. All these filmic techniques achieve a phenomenological experience as if the user could walk into the virtual environment and touch and feel the spatial mise-en-scène. This kind of illusory experience created by the spatiality of the virtual environment can be dizzying and dis-orientating, depending on the user's own physiological coping abilities, yet the stimulation and adrenaline of a bewildered sensory-motor system offers increased excitement and sense of presence in storytelling [3, 4].

Emotional Immersion is the type of immersion when the user feels emotionally aroused and absorbed by the narrative content of the story. Different from spatial immersion, emotional immersion does not necessarily allow users to feel the "bodily presence" into the scene, but allow them to be cognitively identified and emotionally empathized with one of the characters of the story or avatars in the game world.

We want to investigate the following research questions:

- What characteristics of storytelling lead to spatial immersion, and what characteristics of storytelling lead to emotional immersion? Can spatial immersion and emotional immersion be measured with QoE assessment methodologies, and how?
- Do spatiality and emotion increase the degree of immersion in storytelling? If so, which specific quality or qualities of them play a significant role in immersion? And which one is leading to greater degree of immersion, the spatial immersion or emotional immersion?
- Are subjective and qualitative QoE assessment methodologies appropriate and sufficient for the measurement of immersion? What are the advantages and disadvantages of these methodologies, particularly when measuring spatial immersion and emotional immersion?

II. RELATED WORK

Immersion is a complex phenomenon that demands multiple levels of neuro-psychological involvement such as perception, attention and emotion. Its mechanisms and development trajectories during the spectatorship processes of games and storytelling have far from been fully investigated and thoroughly understood, though substantial fruitful attempts have been made in the past. These previous works, though rudimentary and sometimes implicit, provide the foundations for our theoretical formulation and empirical study.

A. Immersion and its related concepts

There are several similar concepts associated with immersion, namely, *flow*, *presence*, *engagement*, *engrossment, cognitive absorption, narrative involvement, puppetry,* and *transportation* [5, 6, 7, 8, 9]. In the study presented by this paper, we consider them as similar concepts with slightly different scopes. Thus, we are trying to obliviate the nuanced differences among them, and consider them as homogeneous or synonymous towards one central theme of "immersion".

The definitions of immersion are diverse and multi-faceted, yet the most celebrated one is that of Janet Murray's, which is "the pleasurable experience of being transported to an elaborately simulated place" and "the sensation of being surrounded by a completely other reality that takes over all of our attention and our whole perceptual apparatus" [10]. This definition points out several key aspects of immersion, such as transportation, simulation, sensation, attention and perception. Full immersion has been described by Brown & Cairns [11] as being a feeling of "entirely cut off from reality." Similarly, Turner et al. [12] consider immersion as "being positively associated with the degree of technologically-mediated sensory richness that facilitates isolation or decoupling from the real world." Coomans & Timmermans [13] defines immersion as a feeling of being

QoMEX2017 – Erfurt, Germany; 978-1-5386-4024-1/17/$31.00 ©2017 IEEE

deeply engaged in a "make-believe world as if it is real." This belief, according to Salen & Zimmerman [14], is termed as "immersive fallacy" when "the virtual reality is so complete that the player truly believes that he or she is part of an imaginary world". Furthermore, Brown and Cairns [11] depict immersion as "a Zen-like state where your hands just seem to know what to do, and your mind just seems to carry on with the story." Another generally agreed-upon definition of immersion is borrowed from Sense of Presence - "the feeling of 'being there'" [15]. In particular, presence is defined by Witmer & Singer [16] as a "subjective experience of being in one place or environment, even when one is physically situated in another."

B. *Types of immersion:*

There are a number of attempts to classify immersion, among which were *Diegetic* and *Non-Diegetic* immersion [17], *Diegetic* and *Situated* immersion [18], *Mental* and *Physical* immersion [19], *Sensory*, *Challenge-based* and *Imaginative* immersion [20], *Perceptual* and *Psychological* immersion [21], *Sensory-Motoric*, *Cognitive* and *Emotional* immersion [22], *Sensory*, *Fictional* and *Systemic* immersion [23], *Spatial, Narrative*, *Strategic* and *Tactical* immersion [24], *Spatial*, *Ludic*, *Narrative* and *Social* immersion [25], *Egocentric* and *Exocentric* immersion [26], *Physical presence* and *Self-presence* [27], and *Visceral* and *Vicarious* immersion [28]. Generally, these can be summarized into two types of immersion: 1) ***Embodied*** immersion: non-diegetic, situated, physical, sensory(-motoric), perceptual, spatial, egocentric, and visceral; and 2) ***Empathetic*** immersion: diegetic, mental, imaginative, fictional, cognitive & emotional, psychological, narrative, exocentric, vicarious, and self-presence. Our criterion is the "self-other" dichotomy of mediated identity, i.e., whether a user is mindful of the "bodily experience" of the Self, or the cognitive and/or emotional experience of the Other (the other self) in the virtual environment.

1) ***Embodied*** *immersion (non-diegetic, situated, physical, sensory(-motoric), perceptual, spatial, egocentric, visceral):*

This category of immersion allows the user to feel the "bodily presence" into the virtual environment, as if they could physically enter into the story and even interact with the virtual objects. Embodied immersion deploys the recruitment of sensory-motoric adjustment and control over the virtual environment instead of cognitive-emotional processing of the story. For instance, physical immersion is described as "bodily entering into a medium" [19] or "the possibility of 'entering' a computer-simulated VR environment and of interacting 'physically' with the objects inside it, of receiving their responses in real time," [29], or "synthetic stimulus of the body's senses via the use of technology" [19]. Sensory immersion is concerned with the audiovisual presentation of a game, utilized in such a way as to overpower any external, real-world sensory information [20, 28]. Perceptual immersion is the extent to which the game experience "monopolizes" the player's senses [21]. This can be accomplished by "blocking as many of the senses as possible to the outside world and making it possible for the user to perceive only the artificial world, by the use of goggles, headphones, gloves, and so on [17]." Spatial immersion has been characterized by Witmer & Singer [16] as the "subjective experience of being in one place or environment, even when one is physically situated in another."

2) ***Empathetic*** *immersion (diegetic, mental, imaginative, fictional, cognitive & emotional, psychological, narrative, exocentric, vicarious, self-presence):*

This category of immersion allows the user to form a mental representation of the real world experience, and project it into the virtual environment, using either imagination, or other psychological, cognitive or emotional apparatus as a vehicle. For instance, mental immersion is considered as being "engaged to the point of suspending disbelief in what they are experiencing" [19]. Imaginative immersion is "the sensation of being mentally absorbed by a story, its world, or its characters" [20]. Psychological immersion suggests the player's engagement in the game using their imagination or their sense of "mental absorption" [17]. In the game world, emotional immersion represents a "positive empathic link toward the avatar" [22]. Narrative immersion is "a state of intense focus on a narrative, elicited by a strong sense of place and the joy of exploration and brought about by emotional attachment to characters" [30]. In a role-player survey, 82.9% of participants confirmed that they have experienced the emotional immersion to "identify so strongly with one's character that it becomes one's primary identity" [31]. In summary, as Witmer & Singer suggest, "When identifying with a character in a book or movie, individuals tend to put themselves in the character's place, and in a sense, experience what that character experiences" [16].

We position our research within the theoretical framework of the above two general categories of immersion. Our study aims to measure *embodied* and *empathetic* immersion, and compare which one is more immersive. To ease understanding among wider populations, we use their lay language definitions: *spatial* immersion and *emotional* immersion. Our hypothesis is:

H: Emotional immersion is more immersive than spatial immersion.

This hypothesis is based on the consideration that spatial immersion evokes mainly sensory-motoric re-adjustment, whereas emotional immersion involves cognitive re-adjustment, thus the latter leads to greater level of neuro-psychological activation.

III. EXPERIMENT

A. *Experiment procedures*

1) *Participants:*

A total of 45 participants were recruited in the study. They were randomized into two groups. The inclusion criteria were normal vision, mental alertness and fair level of English ability. Before the experiment, they were asked to fill in a demographic sheet about their age, gender, educational background, English level, experience with virtual environment and online videos. The demographic conditions of the participants are as follows: age ranges from 21 to 65 years old (mean: 34, median: 29). Education background ranges from high school to Doctorate degree. There were 23 male participants and 22 female participants. All participants are current employees or students of Norwegian University of Science and Technology (NTNU), and they are from 21 nationalities, ranging from Europe (58%), Asia (36%), Africa (4%), and North America (2%). The majority of the participants are active users of online video services: 27% of the participants use online video services on a regular basis (1-3 times a week), and 60% of the participants watch online videos almost every day.

2) *Procedures*

Before the experiment, they were asked to fill in an 18-item questionnaire about their tendencies to immersion [32]. Then the first group (TV group, 23 participants) were required to sit comfortably in a dim dark laboratory environment to watch two pieces of storytelling content of 7 minutes each on a 40-inch LCD display one by one. According to ITU recommendation, the viewing distance is four times the height of the screen [33]. Then the participants were asked to fill in a 33-item questionnaire after each content, which measures their immersive experiences in watching the videos. The second group (tablet group, 22 participants) were following exactly the same procedures, the difference being that this group were watching the same storytelling contents on a 10-inch Android tablet screen. The participants were required to naturally hold the tablet at a position comfortable for them. In the TV group, audio was presented via external PC loudspeakers. The tablet group was exposed to the audio via the integrated loudspeakers of the device. In both cases,

the audio was set to a general comfortable level of loudness by the experimenter and was maintained the same for all participants. The intention of dividing the participants into a TV group and a tablet group was to measure the different impact of devices. All the questionnaires were filled out with paper and pen.

3) Stimulus materials:

The two pieces of storytelling content are characterized as either "spatial" or "emotional", where the spatial content is an architectural visualization of a residential project, and the emotional content is a fairy-airy love story with elements of humor and surprise. Both contents were selected from YouTube under the Creative Commons License, by matching the definitions and criteria of spatial immersion and emotional immersion as closely as possible. Both videos were of 720 HD in resolution and had been downloaded to local hard drive while playing to ensure that the bitrates are roughly equal. A typical frame from each content sequence is presented in Figure 1 and Figure 2. We randomized the order of which content appearing first to offset the order effect. To minimize the potential language barriers, there is no speech of any language in both contents, and there is only background music in each content. It is worth noting that there is not any narrative element in the spatial content.

Figure 1: A frame of the spatial content

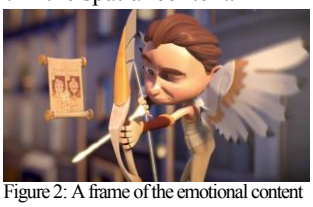

Figure 2: A frame of the emotional content

4) Instruments:

Our 33-item main questionnaire was developed by synthesizing previously validated questionnaires of presence and immersion [34, 16], such as those of Kim & Biocca's, Gerhard et al.'s, I Group, Reality Judgment and Presence, Presence, Dinh et al.'s, Witmer & Singer's, etc. We named our questionnaire as "The Spatial and Emotional Immersion Questionnaire" (SEIQ), which is available in our online repository [35]. We use a 5-point Likert scale as the answer to the questions. The answer can be either from 1 (strongly disagree) to 5 (strongly agree), or from 1 (not at all) to 5 (extremely), depending on how we formulate the question. For device-wise differences, we use a between-subjects design, and for content-wise differences, we use repeated measures of a within-subjects design, thus this experiment is a mixed design combining both between-subjects and within-subjects design methods.

B. Results

1) Comparing overall immersiveness

We select Question 27: "*There were moments during the virtual environment experience when I felt completely focused on the task or environment*" as the markup question to show if the participant has, at any moment during the experiment, reached full immersion. We calculated the arithmetic averages for the spatial content and the emotional content in Excel, and results show that the average for the spatial content in the TV group is 3.17, the same content in the tablet group is 3.27; and the average for the emotional content in the TV group is 3.65, and the same content in the tablet group is 3.68. This result shows that device-wise difference between TV and tablet is not significant, but content-wise, the emotional content is significantly more immersive than the spatial content across devices.

To reinforce this claim, we further performed a two-way mixed-design repeated measures analysis of variance (ANOVA) for the same question in SPSS, using "content" as the within-subjects independent variable and "device" as the between-subjects factor independent variable. Results show that content-wise difference is significant (p=0.012, p<0.05), whereas device-wise difference is not significant (p=0.787, p>0.05). And for device x content interaction, the p value is 0.839. Later, as we proceed further with other items of the questionnaire, we find that two-way ANOVA yields similar results, i.e., device-wise difference is not significant between TV and tablet. This can be explained that the immersive effects of large and small screen sizes are offset by the viewing distance, i.e. tablet allows for an intimate viewing distance, whereas for visual comfort, TV requires the viewers to sit at a certain distance from the screen. And these offset the psychological impact of screen sizes to immersion. Previous arguments also support the views in both camps, for instance, that a larger screen would provide a higher level of immersion due to the 'cinematic' atmosphere that it provides, or that the smaller screen would force the player to focus harder on a smaller space and so be much less aware of their surroundings [36]. Considering the above facts, and for the sake of simplifying the matter, in later parts of the paper we no longer distinguish between devices – we only perform analysis on the content-wise differences here. The nuanced device-wise differences will be discussed later in another paper.

Other questions that define the general aspects of immersiveness include: **Q1**. "*In the virtual environment I had the sense of 'being there'*"; **Q3**. "*When the video ended, I felt like I came back to the 'real world' after a journey*"; **Q4**. "*The story came to me and created a new world for me, and the world suddenly disappeared when the video ended*"; **Q5**. "*During the story, I NEVER forgot that I was in the middle of an experiment*"; **Q6**. "*During the story, my body was in the room, but my mind was inside the world created by story*"; ***Q8***. "*I was involved in communication and the experimental task to the extent that I lost track of time*"; **Q9**. "*My senses were completely engaged during the experience*"; and **Q10**. "*I was completely captivated by the virtual world*." For each question, we calculated the arithmetic averages for the degree of immersiveness of the spatial content and the emotional content, and we performed ANOVA to show if the difference is significant. The results are shown in Table 1:

	Q1	***Q3***	***Q4***	***Q5***
Spatial	2.80	2.56	2.42	3.64
Emotional	3.42	3.38	3.42	2.87
p value	0.002	0.001	<0.0005	<0.0005
	Q6	***Q8***	***Q9***	***Q10***
Spatial	2.60	2.17	2.89	2.31
Emotional	3.56	3.02	3.33	3.33
p value	<0.0005	<0.0005	0.012	<0.0005

Table 1: Comparison of both contents in terms of overall immersiveness

All data shown in Table 1 consistently reflect the fact that overall, the emotional content is significantly more immersive than the spatial content in terms of sense of "being there", time perception, and sense of engagement, which is also consistent with result of the markup question Q27 (Note that Q5 is a negative statement).

2) Specific components of immersion

a) Realism

Q2 is about realism and graphic fidelity: "*In general, the virtual world appears realistic to me*." **Q7** is also about realism, but more of a subjective feeling: "*During the story, the story-generated world was more real or present for me compared to the 'real world'*". **Q24** is about an even higher level of vividness: "*To what extent what you experienced in the virtual world was congruent to other experiences in the real world?*" Q2 is about external realism (perceived match to the actual world), Q24 is about narrative realism (perceived coherence in the story) [37], and Q7 is somewhere in-between. The averages and p values are computed for each question and shown in Table 2:

	Q2	***Q7***	***Q24***
Spatial	3.04	2.00	2.89
Emotional	2.67	2.78	2.60
p value	0.078	<0.0005	0.099

Table 2: Comparison of both contents in terms of realism and vividness

The spatial content is an ultra-realistic architectural visualization of familiar living conditions, whereas the emotional content is an animated cartoon which is more of a metaphoric nature. The results suggest that although graphically the spatial content appears more photo-realistic, the emotional content that contains a story is much closer to the real-life experience. In other words, the spatial content is slightly more immersive in terms of external realism and narrative realism considered separately, yet comprehensively taken together, the emotional content allows much greater feeling of realism.

b) "Bodily presence"

Q15: "*To what extent did you feel like you 'went into' the virtual world, and you almost forgot about the world outside?*" and **Q16**: "*I feel that I could have reached into the virtual world and grasped an object*" are about "bodily presence". For Q15, the average for the spatial content is 1.98, and for the emotional content is 2.82. ANOVA shows the difference is highly significant (p<0.0005). For Q16, the average for the spatial content is 2.07, and for the emotional content is 2.16. ANOVA shows the difference is not significant (p=0.651, p>0.05). The above results suggest that emotional content allows the users to feel more "bodily presence" into the story than the spatial content, but not necessarily to the extent of feeling more able to interact with the virtual objects. This is counter-intuitive to our presumption, and this will be discussed in the Discussions section.

c) Spatial dis-orientation:

Q17 is about spatial disorientation: "*To what extent did you feel disoriented or confused in the Virtual Environment?*" The average for the spatial content is 1.93, and for the emotional content is 1.53. ANOVA shows the difference is significant (p=0.012, p<0.05), suggesting that the spatial content does allow the users to feel more disoriented and require more sensory-motoric re-adjustment of the users.

d) Emotional aspects:

Q18: "*I believe that the virtual world was able to induce emotions*", **Q19**: "*To what extent did the virtual world make you feel emotions (anxiety, sadness, happiness, etc.)?*", **Q20**: "*To what extent did you feel emotionally involved in the virtual experience?*", **Q21**: "*How exhilarated did you feel after the experience?*" and **Q25**: "*To what extent did you get bored while experiencing the virtual world?*" are about the emotional aspects of immersion. We've computed the average and p value for each question, which is shown in Table 3:

	Q18	***Q19***	***Q20***	***Q21***	***Q25***
Spatial	2.53	1.89	1.86	1.84	2.96
Emotional	4.02	3.35	3.15	2.80	1.42
p value			<0.0005		

Table 3: Comparison of both contents in terms of emotional aspects

All these indicate that the emotional content is able to induce or elicit greater amount of high arousal emotions than the spatial content. (Note that Q25 is a negative statement). The spatial content only has a limited or moderate effect on emotion, but the emotional content does exert a great deal of effect on the participant's subjective feelings.

e) Attention:

Q22: "*To what extent did the experience imply a mental effort to you?*", **Q23**: "*To what extent did you have to pay a lot of attention about what was going on in the virtual world?*", **Q26**: "*To what extent did events occurring outside the virtual environment distract from your experience in the virtual environment?*" And **Q31**: "*How much did the visual display quality interfere or distract you from performing assigned tasks or required activities*?" are about attention and immersion. For Q22 and Q23, ANOVA shows the difference is not significant (For Q22: p=0.368, p>0.05; For Q23: p=0.903, p>0.05). For Q26 and Q31, ANOVA shows it's on the verge of achieving statistical significance (For Q26: p=0.175, p>0.05; For Q31: p=0.066, p>0.05). The above results suggest spatial content and emotional content, while immersed in, require similar amount of attentions and mental effort, yet when people are emotionally immersed, attentively and mentally they have slightly better ability to ignore distractions, given the same background conditions.

f) Sensory cues

Q11: "*To what extent did you feel bodily sensations in the virtual world (wind, heat, cold, etc.)*?", **Q12**: "*How much did the visual aspects of the environment involve you*?", **Q13**: "*How much did the auditory aspects of the environment involve you?*" and **Q28**: "*To what extent was the information provided through different senses in the virtual environment (e.g., vision, hearing) consistent?*" are about sensory cues. The averages and p values are shown in Table 4:

	Q11	***Q12***	***Q13***	***Q28***
Spatial	1.82	2.71	2.56	3.13
Emotional	2.02	3.4	3.46	3.84
p value	0.184	<0.0005	<0.0005	<0.0005

Table 4: Comparison of both contents in terms of sensory cues

The above results suggest that emotional content enables significantly greater degree of immersion that allows the users to feel more consistency of the sensory cues.

g) Image motion:

Q29: "*How compelling was your sense of objects moving through space*?", **Q30**: "*How compelling was your sense of moving around inside the virtual environment*?", **Q32**: "*How completely were you able to actively survey or search the environment using vision*?" and **Q33**: "*How closely were you able to examine objects*?" are about image motion. The averages and p values are shown in Table 5:

	Q29	***Q30***	***Q32***	***Q33***
Spatial	2.95	2.84	3.49	3.31
Emotional	3.28	2.93	3.71	3.49
p value	0.035	0.912	0.138	0.229

Table 5: Comparison of both contents in terms of image motion

The above results show that both the spatial content and the emotional content could enable very great sense of image motion in the virtual space, which leads to the sense of presence. And although emotional content scores slightly higher, the difference is not significant.

3) Other findings:

Generally speaking, both the spatial content and the emotional content are able to elicit a certain degree of immersion, and the emotional content is significantly more immersive than the spatial content. However, there are 9 participants (20% of all participants, 7 male and 2 female) who find the spatial content is, at varying degrees (from slightly to extremely) more immersive than the emotional content. By contrast, almost all female participants find the emotional content very immersive. However, male participants' attitudes towards the spatial content are also greatly polarized – some male participants report that they prefer more violent scenes than merely peaceful and serene spatial demonstration.

Another finding is that no matter how we randomize the order of the content, a fair portion of the participants (25 participants, 56% of all participants) always find the second content more immersive, be it spatial or emotional, suggesting that immersion may need time to induce or elicit.

Finally, we find that immersive experience is quite an idiosyncratic experience, and is strongly associated with the participant's demographic conditions and immersive tendencies, but not entirely dependent on the nature of the task. For instance, a male participant finds the spatial content significantly more immersive because he is studying civil engineering as a major; and a female participant finds the spatial content slightly more immersive because she is actively looking for a new apartment lately. And the level of immersion to both contents is strongly

correlated with the participant's immersive tendencies – People with low immersive tendencies can be indifferent to both contents, and those with high immersive tendencies find both contents highly immersive.

IV. Discussions

Admittedly, the psychological mechanism of immersion is still very poorly understood, let alone any attempt to evaluate immersion in the QoE domain. However, designing media technologies that could deeply engage and retain users is one of the primary objectives of researching into Quality of Experience. Thus, immersion is a key experiential factor in evaluating multimedia systems and particularly in storytelling. We reiterate this point at the beginning of the paper because we consider this cannot be over-emphasized. By studying and distinguishing between spatial immersion and emotional immersion, we've made the first step of tapping into the psychological processes of immersion.

The QoE assessment and modelling of immersive experience is a fledgling yet promising topic, fledgling because immersive experience touches upon many of the intangible features of human experience which are rather difficult to sense, capture, interpret and/or interact with, let alone its quality assessment and modelling; promising because immersion is a major psychological mechanism in media enjoyment which, if properly unveiled, can lead to significant improvement and innovation in the value creation of media production and consumption.

From the system perspective, the capability of a technical system "to deliver an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant" has been defined as immersion [38]. The system QoE factors influencing immersion include [39]: pictorial realism and graphic fidelity, image motion, screen size, visual cues (spatial and object cues), sensory effects (wind, vibration, light effects), etc. These have all been investigated in our study. We have also investigated human QoE factors such as attention, time perception, emotions, etc. and contextual factors such as awareness of environmental distractions, interface fidelity, etc. All these offer new insights to the QoE assessment and modelling of immersive experiences.

The components and dynamics underlying media enjoyment (QoE) have been found as a dependent variable of personality traits, individual differences, mood, content characteristics, etc. As a result, it has been characterized as a multidimensional construct conditioned by affective components in its first place [39]. In particular, emotional enjoyment has been found closely linked to entertainment as a media effect, where media provides a means to "escape to a fantasy world where emotions can be experienced" [40]. Thus, emotion is a primary component that determines the QoE in immersive storytelling. And emotion is triggered by cognitive primers, thus emotional content containing a narrative or story has better capability to immerse audience. Also, "drama" is an important defining element in eliciting immersion [16].

From the individual difference perspective, male participants seem to be more inclined to applaud the spatial content. This can be explained by the fact that men are primarily spatial visualizers [41], and thus have better spatial intelligence or spatial skills [42]. And spatial intelligence is a human factor in QoE to evaluate user's adaptation to the virtual environment. Our study also shows that female participants empathize better with the virtual character than male participants, which is consistent with previous findings that in general women are more empathetic than men [43] and in the virtual environment females empathize with virtual characters more readily than males [44], and this leads to higher level of presence. Length of exposure to the media environment is also a contextual factor in QoE that determines degree of immersion, and this can explain why a fair portion of the participants always find the second content more immersive.

Now we can address or answer the research questions:

- Drama or narrative is the primary factor that enables immersion, and this overshadows other immersive effects such as photo-realistic graphic fidelity or disorientating or dazzling spatial effects. Without certain dramaturgical structures in the storytelling content, it is difficult to elicit high degree of sense of presence among users, even if it demonstrates significant levels of other experiential qualities in storytelling, immersion, or QoE.
- Subjective and qualitative method such as questionnaire is a cost-effective and easily quantifiable method in QoE assessment and modelling. Adapted to the immersive storytelling scenario, it's a viable means of measuring immersiveness, but in terms of more accurately measuring the users' cognitions, emotions and behaviors through neuro-psycho-physiological data we would say it is not adequate. For instance, some participants complain that a 5-point Likert scale is not enough to measure their nuanced degree of immersiveness. In such cases we would assume that a 9-13 point scale would be more appropriate and accurate. And some participants report that after watching the second video, they had a desire to scale up or scale down the answers to the first video. And this is perhaps a major dilemma of a within-subjects design using questionnaires.
- For spatial immersion, ideally it would be better to use a Head-Mounted Display (HMD) to allow for much greater degree of "bodily presence" in the virtual environment. But we settled with the current experimental design because we wanted to measure and compare immersiveness under equal experimental conditions, i.e., using flat 2D screen displays for both contents. Previous studies have measured and compared egocentric immersion (HMD-based) and exocentric immersion (screen-based), and concluded that egocentric immersion is positively associated with better task performance in a chess game [26]. Thus, with HMD, the results might be different, and this is pending further investigation.
- Our results show that emotional immersion allows the users to feel more "bodily presence" into the virtual environment. Previous studies suggest that film enjoyment is a function of narrative experience, perceived realism and transportability [37]. Transportation theory also suggest that presence is a coherent process of transportation (spatial immersion) and identification (emotional immersion) in a narrative media enjoyment [9]. Thus we may conclude (tentatively) that spatial immersion and emotional immersion are an integral process in viewing a narrative content, and spatial immersion is an antecedent process immediately followed by emotional immersion, i.e. people firstly feel "bodily presence" into the scene, and then are emotionally empathized and cognitively identified with character. And this identification process further reinforces the feeling of transportation. This also suggests that emotional immersion is a higher level immersion above spatial immersion, i.e. being spatially immersed does not necessarily mean one is emotionally immersed, yet emotional immersion is always on the premise of spatial immersion. In other words, identification always incorporates elaborated transportation.

V. Conclusion

In conclusion, our research provides empirical evidence of measuring spatial and emotional immersion that are previously only theoretically or conceptually discussed. Interestingly, the phenomena of spatial immersion and emotional immersion have been extensively discussed in previous literature, but they haven't been systematically classified and empirically validated as we did in this paper, which is important for future research in designing better immersive technologies and experiences. We have proved the hypothesis that emotional immersion is more immersive than spatial immersion, *tempus esse*. We also linked various

components of the measurement to QoE influencing factors, in an aim to bridge the gap in the QoE assessment and modelling of immersive experience in storytelling.

The practical implications of our research are multi-fold:

i) In the majority of current VR/AR/MR applications, tremendous emphases are still placed on enhancing spatial immersion and spatial experience, whereas the drama or narrative elements are significantly missing. This to a great degree hampers the enjoyment of full immersion. Future directions of these immersive technologies should make efforts to combine more engaging storytelling techniques with the existing spatial elements to allow for better immersive experience.

ii) Since emotion is major component in eliciting immersion. Future design of immersive technologies should benefit from other cognate areas such as affective computing, intelligent interaction, affective neuroscience, etc. This cross-fertilization will bring fruitful progress by complementing and broadening the existing QoE influencing factors research and in turn enhancing the experiential qualities of immersive technologies.

iii) Emerging visual reality technologies such as JPEG Pleno may have the capacity to revolutionize our perception towards spatially immersive virtual environment by playing with spatial qualities of the image such as depth-of-field, focus, and observer's perspective [45]. These would greatly enhance the effect of spatial immersion, then we might need to re-evaluate our topic when these technologies become mature and popular.

iv) In addition, there is still a lot of space to enhance spatial immersion by future display technologies, enhanced bodily interaction technologies, larger Field-of-View (FOV), enhanced graphic simulation and motion realism technologies, etc. So, technologies are rapidly advancing. With these promising future technologies comprehensively enhancing our spatial perception, it might be still too early to make the judgment that spatial immersion is less immersive.

v) Since there are meta-cognition and meta-emotion, and considering that emotional immersion is a higher level immersion, is there meta-immersion, i.e. immersion embedded in immersion(s)? This can be an interesting area to explore in Mixed Reality (MR) when designing more complex, intricate or elaborate immersive experiences.

In summary, both spatiality and emotion have vast potentials for immersiveness. The future immersive technologies should try to harness the powers and exploit the potentials of spatiality and emotion to produce an alternate reality that stimulates all our senses. and engages all our attentional resources and perceptual apparatus, i.e. an "ultra immersion".

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 643072, Network QoE-Net.

REFERENCE

[1] Perkis, A. (2013). "Quality of Experience (QoE) in Multimedia Applications, SPIE Newsroom, 27 February 2013. [Online]. Available: http://spie.org/x92222.xml

[2] Perkis, A. (2013). A QoE cross layer approach to model media experiences. IEEE COMSOC MMTC E-Lett, 8(2), 6-8.

[3] Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. Frontiers in psychology, 2, 395.

[4] Bouchard, S., St-Jacques, J., Robillard, G., & Renaud, P. (2008). Anxiety increases the feeling of presence in virtual reality. Presence: Teleoperators and Virtual Environments, 17(4), 376-391.

[5] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. International journal of human-computer studies, 66(9), 641-661.

[6] Brown, E., & Cairns, P. (2004, April). A grounded investigation of game immersion. In CHI'04 extended abstracts on Human factors in computing systems (pp. 1297-1300). ACM.

[7] Quintero Johnson, J. M., & Sangalang, A. (2017). Testing the explanatory power of two measures of narrative involvement: An investigation of the influence of transportation and narrative engagement on the process of narrative persuasion. Media Psychology, 20(1), 144-173.

[8] Calvillo-Gámez, E. H., & Cairns, P. (2008). Pulling the strings: A theory of puppetry for the gaming experience. In Conference proceedings of the philosophy of computer games (pp. 308-323).

[9] Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. Communication Theory, 14(4), 311-327.

[10] Murray, J. (1997). Hamlet on the Holodeck. Cambridge, MA: MIT Press p.68-73.

[11] Brown, E., & Cairns, P. (2004). A grounded investigation of immersion in games. In Proceedings of the ACM Conference on Human Factors in Computing Systems, Vienna, Austria: CHI 2004 (pp. 1297-1300). New York, NY: ACM Press.

[12] Turner, S., Huang, C. W., Burrows, L., & Turner, P. (2016). Make-Believing Virtual Realities. In Digital Make-Believe (pp. 27-47). Springer International Publishing.

[13] Coomans, M. K. D., & Timmermans, H. J. P. (1997). Towards a taxonomy of virtual reality user interfaces. In Proceedings of the International conference on Information Visualisation (IV97) (pp. 279-284). Los Alamitos, CA: IEEE Computer Society.

[14] Salen, K., & Zimmerman, E. (2004). Rules of Play: Game Design Fundamentals. Cambridge MA: MIT Press.

[15] Heeter, C. (1992). Being there: The subjective experience of presence. Presence: Teleoperators & Virtual Environments, 1(2), 262-271..

[16] Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence, 7(3), 225-240.

[17] McMahan, A. (2003). Immersion, Engagement and Presence: A Method for Analysing 3D Video Games. In M. J. P. Wolf & B. Perron (Eds.). The Video Game Theory Reader (pp. 67-86). New York, NY: Routledge..

[18] Taylor, L. N. (2002). Video Games: Perspective, point-of-view, and immersion. (Master's Thesis). University of Florida. Retrieved from http://www.laurientaylor.org/research/taylor_l.pdf

[19] Sherman, W. R., & Craig, A. B. (2003). Understanding Virtual Reality: Interface, Application and Design. CA: Morgan Kaufmann Publishers

[20] Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. In S. D. de Castell & J. Jenson (Eds.), Worlds in play: International perspectives on digital games research (pp. 15–27). New York, NY, USA: Peter Lang Publishing.

[21] Carr, D., Buckingham, D., Burn, A., & Schott, G. (2006). Computer Games: Text, Narrative and Play. Cambridge, UK: Polity Press.

[22] Björk, S., & Holopainen, J. (2005). Patterns in Game Design. Hingham, MA: Charles River Media, Inc.

[23] Arsenault, D. (2005). Dark waters: Spotlight on immersion. In proceedings of the Game-On North America 2005 Conference (pp. 50–52). Ghent, Belgium: Eurosis.

[24] Adams E., & Rollings A. (2006). Fundamentals of game design. NJ, USA: Prentice Hall.

[25] Thon, J. N. (2008). Immersion revisited: on the value of a contested concept. Extending Experiences-Structure, analysis and design of computer game player experience, 29-43.

[26] Slater, M., Linakis, V., Usoh, M., Kooper, R., & Street, G. (1996, July). Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In ACM virtual reality software and technology (VRST) (Vol. 163, p. 72). New York, NY: ACM Press.

[27] Lee, K. M. (2004). Presence, explicated. Communication theory, 14(1), 27-50.

[28] Curran, N. (2013). The psychology of immersion and development of a quantitative measure of immersive response in games. Doctoral dissertation, University College Cork.

[29] Civitarese, G. (2008). Immersion versus interactivity and analytic field. International Journal of Psychoanalysis, 89(2), 279-298.

[30] Ryan, M. L. (2003). Narrative as virtual reality: Immersion and interactivity in literature and electronic media. Baltimore, MD, USA: The Johns Hopkins University Press.

[31] Harviainen, J.T., 2007. Information, Immersion, Identity: The Interplay of Multiple Selves during Live-Action Role-play. Journal of Interactive Drama, 2, 1

[32] Witmer & Singer (1996). Immersive tendencies questionnaire

[33] BT2022, ITU. R. (2012). General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays. Int. Telecommun. Union.

[34] Van Baren, J., & IJsselsteijn, W. (2004). Measuring presence: A guide to current measurement approaches. Deliverable of the OmniPres project IST-2001-39237.

[35] http://www.iet-multimedialabs.org/wordpress/wp-content/uploads/2017/02/SEIQ.pdf

[36] Thompson, M., Nordin, A. I., & Cairns, P. (2012, September). Effect of touch-screen size on game immersion. In Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers (pp. 280-285). British Computer Society.

[37] Bilandzic, H., & Busselle, R. W. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. Communications, 36(1), 29-50.

[38] Slater M (2003) A note on presence terminology. Presence-Connect, pp 1–5

[39] Galloso, I., Feijóo, C., & Santamaría, A. (2015). Novel approaches to immersive media: from enlarged field-of-view to multi-sensorial experiences. In Novel 3D media technologies (pp. 9-24). Springer New York.

[40] Sherry JL (2004) Flow and media enjoyment. Commun Theory 14:328–347.

[41] Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. Memory & cognition, 33(4), 710-726

[42] Sacau, A., Laarni, J., & Hartmann, T. (2008). Influence of individual factors on presence. Computers in Human Behavior, 24(5), 2255-2273.

[43] Norscia, I., Demuru, E., & Palagi, E. (2016). She more than he: gender bias supports the empathic nature of yawn contagion in Homo sapiens. Royal Society open science, 3(2), 150459.

[44] Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. Human communication research, 31(4), 511-537

[45] Ebrahimi, T., Foessel, S., Pereira, F., & Schelkens, P. (2016). JPEG Pleno: Toward an Efficient Representation of Visual Reality. IEEE MultiMedia, 23(4), 14-20.