Enabling User-centric Assessment and Modelling of Immersiveness in Multimodal Multimedia Applications

SAM VAN DAMME, MARIA TORRES VEGA, and FILIP DE TURCK, IDLab, Department of Information Technology (INTEC), Ghent University - imec, Belgium

Multimodal, immersive systems are the latest development within the field of multimedia. These systems emulate the senses by means of omnidirectional visuals, 360° sound, motion tracking and touch simulation to create a feeling of presence in the virtual environment. They have the potential to substitute physical interactions in application domains such as training (Industry 4.0) or e-health (tele-surgery). However, the COVID-19 pandemic has shown that they are not ready, as they still have room for improvement in terms of network streaming quality, usability and the users' feeling of presence. In addition, these systems can induce feelings of dizziness, nausea etc. (i.e. cybersickness). These factors therefore have an important impact on the user's total immersion. In this work, we therefore propose that immersiveness can be devised from measuring four aspects, namely: presence (i.e. the feeling of being "in" the environment), cybersickness, network related Quality-of-Experience (QoE) and the usability of the application. Therefore, a two-dimensional user-centric approach on the assessment and modelling of immersiveness is proposed. These dimensions include (i) subjective and objective assessment of presence, cybersickness, usability and QoE and (ii) real-time modelling of immersiveness. Furthermore, a proof-of-concept is envisioned including two use cases. As such, we believe that this position paper will significantly advance the state of the art on immersive systems and multimedia in general.

Additional Key Words and Phrases: Immersive multimedia, Presence, Cybersickness, Usability, Quality-of-Experience, Assessment, Modelling

1 INTRODUCTION

Augmented Reality (AR) and Virtual Reality (VR) multimodal experiences are the latest evolution within multimedia applications [27]. These systems aim at emulating (certain) senses as accurately as possible to create a realistic virtual or augmented environment. By means of omnidirectional video or volumetric media, i.e. point clouds, combined with 360° sound and motion tracking, the users are given the feeling that they truly are "in" the environment. Nowadays, these systems are even further improved by means of haptic feedback. The recent *SenseGlove Nova*[1], for example, provides a unique combination of force and vibrotactile feedback and motion tracking. Other companies such as *TeslaSuit*[2] even provide full body suits providing touch and force feedback, motion capture and biometrical monitoring, therefore even further improving the immersive potential. These immersive systems are already being adopted in the entertainment sector to some extent, e.g. for gaming and video content (e.g. Playstation VR[3], Netflix VR[4]…). Moreover, they open good opportunities for sectors with more societal and economical impact. Future applications in industry (Industry 4.0), healthcare [5] (e.g. tele-surgery), and education (remote learning) will highly benefit from this AR/VR multi-sensor systems. However, the current CoViD-19 health situation has taught us that these immersive multimodal systems are still not fully ready to substitute real human contact, despite the rise of 5G and its accompanying possibilities. Pinpointing the exact factors of influence of increasing/decreasing immersiveness, however, is extremely difficult as it depends on a large range of conditions and parameters [26].

[1]https://www.senseglove.com/product/nova/
[2]https://teslasuit.io/the-suit/
[3]https://www.playstation.com/nl-be/explore/playstation-vr/
[4]https://play.google.com/store/apps/details?id=com.netflix.android_vr&hl=nl

IMX '22 - Doctoral Consortium, 2022, Portugal

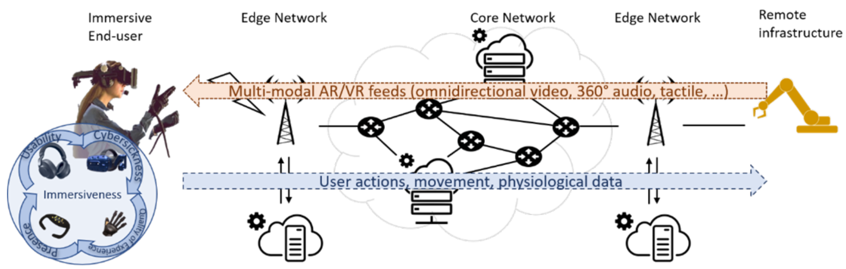

Fig. 1. Immersive session in the form of a remote presence scenario.

In this PhD work, I hypothesize that immersiveness can be understood from four different factors[12, 15] (Figure 1). These are Quality-of-Experience (QoE), cybersickness, presence and usability. *QoE* is defined as *the degree of delight or annoyance of the user of an application or service, where its focus lies on the delivered quality* [30]. *Cybersickness* denotes the physical discomfort resulting from the use of the immersive system [22]. *Presence* is the state of consciousness that indicates the psychological sensation of being inside the virtual environment [15]. It is typically influenced by how natural and intuitive the user can interact with and/or navigate in the environment, as well how "real" the environment appears in terms of visual, auditive and haptic quality and the synchronization between them [15]. Finally, *usability* is expressed as the ability to perform a certain task within a reasonable amount of time, e.g. picking up an object, or locating a certain place, rather than the numerical quality scoring [28]. Each of them have been evaluated by means of a posteriori questionnaires. Such are the case of the well-know Motion Sickness Susceptibility Questionnaire (MSSQ) [9] and Virtual Reality Neuroscience Questionnaire (VRNQ) [13] for cybersickness or the Igroup Presence Questionnaire (IPQ) [25] for presence.

By means of a joint analysis of QoE, cybersickness, presence and usability, this work proposes a multidimensional approach to enable real-time assessment of immersiveness for multimodal AR/VR applications. The remainder of this paper provides first a thorough description of the related work in section 2. Then, the two pinpointed research questions are outlined in section 3. The current status of the research is provided in section 4. Finally, section 5 provides some conclusion and room for future research.

2 RELATED WORK

As stated above, the research proposed in this project hypothesizes that immersiveness can be assessed as the combined action of four different aspects, namely *QoE, cybersickness, presence and usability*. Most of the existing research has evaluated these aspects independently. Therefore, this analysis provides an overview of the most recent and relevant approaches and evaluates the few examples of multi-objective assessment approaches.

The most straightforward manner to understand the effects of multimedia feeds on users' *QoE* has traditionally been subjective evaluations [32] on a few dozen of people with varying backgrounds. For more traditional applications such as 2D video, testing conditions have been standardized by the International Telecommunication Union (ITU)[32]. These tests are performed by offering multiple sequences with different degrees of impairment to the subjects, after which the users grade the quality on a certain scale. The average score over all subjects is called the Mean Opinion Score (MOS). As the degree of immersiveness of the application rises, however, assessing the QoE of the end-user becomes more difficult. In 360° videos, for example, one must take into account that a user only sees a limited part of the 360° hemisphere (the viewport) at each instant. Therefore, users might watch different portions of the video during playback, which makes it rather difficult to compare quality scores among users and to combine them to a single average score per video. At this point in time, no generalized assessing methodology exists to solve this issue[36]. It becomes even more complex when additional sensory inputs, and tactile feedback in particular, are added to the experience. Due to the complex combination of sensorial data types that influence the user, MOS becomes rather infeasible to define the quality. Research towards the actual design of such performance tests is currently scarce, however.

Given the diversity of symptoms (disorientation, headache, dizziness, nausea), a broad range of methods for detecting and assessing the severity of *cybersickness* have appeared over the years[11]. Questionnaires are the oldest and the most common form of detection, where the Simulator Sickness Questionnaire (SSQ) is the most popular option still used nowadays [4]. The SSQ consists of a set of questions regarding the severity of symptoms on a scale of 0–3. Scores are computed for three categories (nausea, oculomotor, and disorientation). Even if broadly used, the SSQ has several drawbacks. Several symptoms contribute to more than one category. It can lead to over-sensitivity and bias depending on the user taking the test. Moreover, it takes too long to complete, therefore making some participants to lose their attention. Finally, as it was originally created for pilots in the air force[22], it lacks the generality for AR/VR multimodal applications. Lately, the VRNQ [13] has been developed, which evaluates the overall perception based on 20 questions divided in four subdomains of five questions each. Here, the subdomain on Virtual Reality Induced Symptoms and Effects (VRISE) is representative for cybersickness by evaluating symptoms of nausea, disorientation, dizziness, fatigue and instability on a 7-point scale. Other subdomains are User Experience (UX) ($\approx$ QoE + presence) and Game Mechanics + In-Game Assistance ($\approx$ usability). The VRNQ shows a lot of potential for subjective evaluation of multimodal symptoms, but is not yet enhanced with the evaluation of haptic feedback, nor is the scenario for AR experiences instead of VR included. As such, it is necessary to further improve and update the VRNQ to provide the required benchmark for novel AR/VR multi-modal applications.

Presence has also been assessed by means of questionnaires. The most known and used approach is the IPQ [25] that covers Spatial Presence, Involvement and Realism by means of 14 questions. Its biggest advantage is that it is the only questionnaire on presence that was subject to a confirmatory factor analysis of its psychometric properties. In addition, it is one of the most widespread evaluation methods due to its translation in a plethora of languages [37]. Furthermore, some questions of the UX component of the VRNQ can be added for subjective presence evaluation as well.

The concept of *usability* can be interpreted in the context of playability for gaming environments. Here, it is defined as *the degree to which specified users can achieve specified goals with effectiveness, efficiency and, especially, satisfaction and fun in a playable context* [21]. In other words, usability's goal is to measure the ability of the user to perform an action. Multiple studies have been performed in the gaming community in order to define heuristics suited for assessment of playability in video games [6]. As such, we envision usability in this project by means of a fully data-driven approach, to avoid intrusion of the application. Different metrics will be used to assess performance, such as the time needed to complete the task, the number of attempts before success, the average accuracy (e.g. hitting a target etc.) [17].

While subjective evaluations provide the most accurate manner to assess the perception of the user, they cannot be run in real-time. Regarding *QoE*, most research has been conducted on each of the individual feeds of the system. On the visual, limited attempts exist to expand traditional 2D video metrics towards 360° and holographic content, such as the Point cloud Quality Rating (PQR) for point clouds [1] or Weighted-to-Spherically-uniform Peak Signal-to-Noise Ratio (WS-PSNR) [31] and 360° Video Multimethod Assessment Fusion (VMAF) [19] for standard omnidirectional video content. During the last years, some additional studies have been performed concerning the application of traditional video metrics for quality assessment of point clouds by means of so-called projection-based methods, such as cube projections, binary patterns, color-based approaches and the adaptation of traditional Structural Similarity Index Measure (SSIM) for use with point clouds by Alexiou et al. [3]. In addition to this, we presented a thorough correlation analysis of both No-Reference (NR) and Full-Reference

(FR) objective metrics to MOS for multiple volumetric media streaming scenarios in one of our previous works [35]. By means of Region-of-Interest (ROI) selection, the applicability of these classic visual metrics is improved.

For auditory feeds, there also exist a handful of metrics. In terms of perceived quality of sound, most metrics focus on the quality of speech sequences [16]. However, ViSQOLAudio and AMBIQUAL [16] for more general, traditional audio and ambisonics (i.e. a full sphere audio surrounding technique), respectively, are worth mentioning. Recently, an improved, open-source version for production usage, called ViSQOL v3, has been released [7]. The haptic feed is the least explored path of the three senses. The limited amount of haptic-related, objective metrics is based on generic metrics for evaluation of signal quality, e.g. Mean Squared Error (MSE) or Signal-to-Noise Ratio (SNR). Haptic Perceptually Weighted Peak Signal-to-Noise Ratio (HPWPSNR) [24] is probably the most known example of a haptic objective metric. Another, more recent one is the Haptic Structural Similarity Index Measure (HSSIM) [27], which is an adaptation of the classic SSIM for video quality estimation to the specific case of haptic feedback signals. These approaches are rather limited in amount and, more importantly, stay within the limits of one particular sensor channel. The influence of multiple feeds on each other remains unexplored [27].

In terms of objective *cybersickness* assessment, the path of physiological signals is mainly followed at the moment. Physiological signals or biometrics have the potential to provide a solution both objective and non-intrusive. Early research has shown certain levels of correlation between cybersickness (by means of the SSQ) and electrocardiogram (ECG), blood pressure, electrogastrogram (EGG), and respiration [22]. Other options are skin temperature, Galvanic Skin Response (GSR), eye movement and electroencephalogram (EEG). However, there is still plenty of research to be done in this topic to explore the link between cybersickness, physiological signals and subjective evaluation. Combining multiple detection methods could improve the accuracy of detection. Machine Learning (ML) has recently been used to create multi-method real-time models of cybersickness. Most of the approaches use a broad range of parameters to classify an experience as having cybersickness or not [4]. These approaches, while being promising, still lack the granularity required to make real-time changes in the virtual environment. Therefore, there is a need for integrated metrics able to detect the granular level of cybersickness in real-time.

Objective modelling approaches for *presence* are scarce in literature. The approaches that do exist are often based on an analysis of the verbal and non-verbal behavioural patterns of the user, such as their linguistic complexity over time, inter-pausal duration, navigation paths, body movements etc. [14]. Once again, ML techniques such as Support Vector Machines (SVMs) or Random Forests (RFs) are used to further model the input data based on a subjective ground truth [18].

As of today, no specific studies exist specifically on the objective modelling of usability in multimodal systems. However, objective measurements of user performance and usability are often embedded within other user studies, including the time to complete the task[14], task accuracy [23] etc.

3 RESEARCH QUESTIONS AND OVERALL ARCHITECTURE

In our proposed research path, we focus on two main research questions: (i) How can we accurately assess the user’s level of immersiveness, both subjectively and data-driven? (ii) How can we create an objective benchmark for immersiveness and use this for real-time modelling and monitoring? The envisioned approach to answer both questions is visually illustrated in Figure 2. As explained before, immersiveness will first be assessed modular in terms of network QoE, cybersickness, presence and usability by means of subjective questionnaires, physiological measurements and data-driven performance metrics. Afterwards, ML approaches will be implemented to combine these factors in a so-called *Overall Immersive Index*, which can be used as a benchmark for real-time modelling. More details on the two research questions are provided in the following sections.

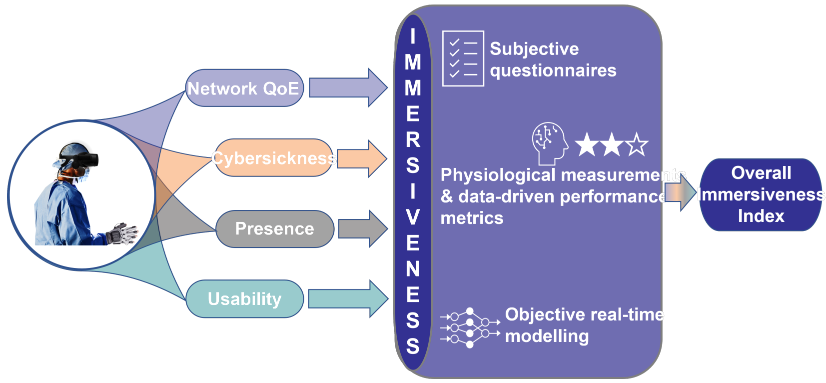

Fig. 2. The envisioned assessment and modelling approach of Overall Immersiveness

3.1 How can we accurately assess the user's level of immersiveness, both subjectively and data-driven?

Immersive, multimodal systems are relatively new to the multimedia scene. As a result, user perception of environments with multiple sensory inputs is a research topic still largely unexplored in literature. In order to establish a human-centric, client-side immersiveness management mechanism, it is of great interest to be able to understand the user's perception of the application [30]. To this end, we propose to interpret immersiveness as the combination of QoE, cybersickness, presence and usability as illustrated in Figure 2. As such generic, task-independent user tests will be needed that cover each of these four aspects to provide an overall degree of immersiveness. We therefore propose a solution to this question consisting of five stages. The first four tackle each of the individual aspects and the fifth deals with the creation of an *Overall Immersiveness Index*.

To this end, first a data-driven usability evaluation will be designed. Therefore, a set of task-independent, objective metrics will be defined that can measure to what extent the user is able to perform the envisioned task. These will include the time needed to complete the task, the number of attempts before success, the dropout rate (users that do not reach the goal) etc.

Second, a QoE subjective evaluation will be performed. In current evaluations in the state-of-the-art, users provide one single score for the whole quality of the experience (e.g. the well-known MOS [33]). This will not be enough for the case of multimodal experiences. Therefore, this analysis envisions to modify standard subjective questionnaires to include not only an overall quality, but questions regarding individual feeds. Furthermore, it has been shown that performing one overall questionnaire at the end of the experience biases the user to perceive the last part of the experience as the overall experience. To avoid this, the users will be asked to provide answers to single-question questionnaires after certain specific events (throughout the immersive experience). As such, each of these questionnaires will cover a certain aspect of the current environment (i.e., visual quality, audio quality, haptic) and task while the experience is not disturbed.

Third, both presence and cybersickness will be subjectively assessed based on questionnaires subsequent to the experience. These will be based on the IPQ [25] and VRNQ [13] respectively. Part of the research will include the further adaptation of these questionnaires to the multimodal nature of the presented multimedia experiences. In addition, objective measurements of cybersickness will be concluded by means of the physiological signals presented earlier.

Furthermore, research will be performed to combine the assessment methodologies on usability, QoE, presence and cybersickness. As such, a (set of) important subject derived assessment(s) can be provided in order to combine them into an overarching measurement and assessment methodology of overall immersiveness. User clustering can optionally be performed to combine users that show the same level of susceptibility to one of the aspects to assess to what extent these add to the total perception.

As this research merely focuses on applications within industry, healthcare, education and remote collaboration, we mainly aim at user profiles within academia and the private market that are familiar with and/or open to adopting new types of technology. As such, the primarily focus of the user tests will lay on users aged 18-45 with distributions in ethnic background and gender approximately uniform. The latter is especially important as subjects identifying as female tend to be more susceptible to cybersickness [8]. In addition, as age tends to play a role in the occurrence of cybersickness as well, a minority of subjects will be selected above the preferred age range. Participants will be gathered through announcements on publicly available mailing lists and social media, as well as ad valvas and via word of mouth within the academic setting.

3.2 How can we create an objective benchmark for immersiveness and use this for real-time modelling and monitoring?

The second research question focuses on the creation of objective, mathematical models that grasp the multiple factors of immersiveness. To tackle this research, we base our approach on prior experience on QoE modelling for gaming content in which we created an objective benchmark for game video streaming purposes, as well as a lightweight curve-fitting model for real-time quality estimation [34]. For it, three subsequent phases will be defined: (i) correlation analysis, (ii) creation of objective benchmarks and (iii) realization of real-time, client-based, lightweight quality estimation models. Throughout these three phases, we envision a plethora of sensors and metrics that will act as an input for each of them: (i) network-related (i.e. bitrates, delays, packet loss, jitter...), (ii) application-layer (NR/Reduced-reference (RR)) Quality-of-Service (QoS) metrics and (iii) physiological data (i.e. EEG, ECG, GSR, blood pressure, eye-tracking...).

The first phase will include a thorough correlation analysis on how different system, application and local-network parameters relate to the four components of immersiveness described before. These will reveal the most important network conditions, system parameters and application characteristics to monitor and will serve as input to a quality model. In our previous research [34], Pearson Linear Correlation Coefficient (PLCC) and MSE were applied to this end. In addition, clustering mechanisms will be researched to investigate whether multiple types of users can be derived that are susceptible to different influence factors.

In a second phase, objective immersiveness metrics will be selected/designed with the purpose of accurately assessing the user's overall perception. As such, no explicit participation of human subjects is needed to evaluate the performance of real-time, lightweight quality models. We hypothesise that a combination of per-feed objective metrics, combined with the physiological signals, will provide the most reliable metric for this benchmarking purpose [34]. To this end, per-feed quality metrics will be calculated: (i) PQR[2] or WS-PSNR[29] and 360° VMAF [20] (depending on the content, i.e. point-clouds or synthetic) for the visual feed, (ii) AMBIQUAL [16] for audio and (iii) HSSIM [10], SNR and MSE for the tactile feeds. How to combine these metrics is a topic of thorough research, in which ML-based solutions, such as Artificial Neural Networks (ANNs) and/or SVMs can play an important part.

As the immersiveness estimation needs to take place in (near-) real-time on the client-side, lightweight prediction algorithms are required. Once again, ML algorithms will act as a supporting tool to relate low-complexity QoS and physiological parameters to the objective benchmarks [30].

As such the three presented phases of this research direction will result in a fully objective mechanism for modelling of immersiveness. A set of highly influential, lightweight parameters on each of the four components will be derived, as well as an objective benchmark that accurately mimics subjective perception in absence of human evaluators. Both will be related one to another by means of ML techniques in order to allow for real-time lightweight quality measurements.

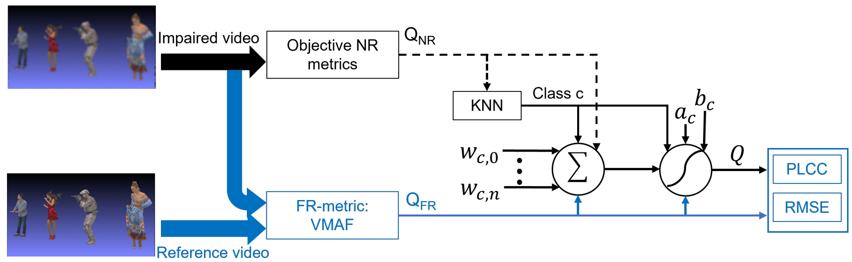

Fig. 3. Block diagram of the point-cloud QoE assessment methodology. The parts in black are used at client-side for real-time quality assessment. The blue parts are added at server-side for training and evaluation purposes.

Table 1. Average PLCCs and Root Mean Squared Errors (RMSEs) per video of the proposed point cloud QoE assessment solution, compared with the cases with and without clustering and sigmoidal mapping.

(a) PLCC

	Without clustering	**With clustering**
LR	0.826	0.942
LR + sig. map.	0.869	**0.977**

(b) RMSE

	Without clustering	**With clustering**
LR	0.114	0.095
LR + sig. map.	0.199	**0.077**

4 CURRENT STATUS

Our preliminary research on the implementation of the aforementioned approach for the particular case of QoE for streamed point-cloud videos has already shown promising results. As shown in Figure 3, we used a K-Nearest Neighbours (KNN) clustering mechanism to pre-classify a collection of point cloud videos based on a set of calculated NR metrics (ie. blur, noise, blockiness and Spatial Information (SI)). Next, a per-class Linear Regression (LR) and sigmoidal mapping are calculated by minimizing MSE towards the VMAF benchmark. As illustrated in Table 1, both the clustering and the sigmoidal mapping show to be beneficial in comparison with the cases without, with an average PLCC up to 0.977 and an RMSE down to 0.077 on a zero-to-one scale after evaluation on an adaptive streaming point-cloud dataset consisting of sixteen source videos and 453 sequences in total. As modelling is performed at video-level, however, it is worth further exploring whether these findings still hold on a per-Group-Of-Pictures (GOP) or even a per-frame level in further research.

In order to accurately evaluate the proposed modelling and assessment methodologies, it is important to create a generic evaluation infrastructure in which a multitude of use cases can be plugged in. This experimental testbed will be created between the IoT-lab at the iGent building of Zwijnaarde and the Art & Science lab of De Krook in Ghent, as both are already connected by a 10 Gbps communication link (Figure 4). During testing, users will be equipped with an *Oculus Quest 2* Head-Mounted Display (HMD), headphones supporting ambisonics and the *SenseGlove Nova* haptic gloves including vibrational, tactile and force-feedback. In addition, motion cameras will be installed for motion tracking. The use case scenarios are chosen such that the user needs sufficient time for completion. As such, each of the four immersiveness components can be thoroughly monitored. Additional characteristics of the evaluation, such as the duration of the session, the presented stimuli etc., will be chosen

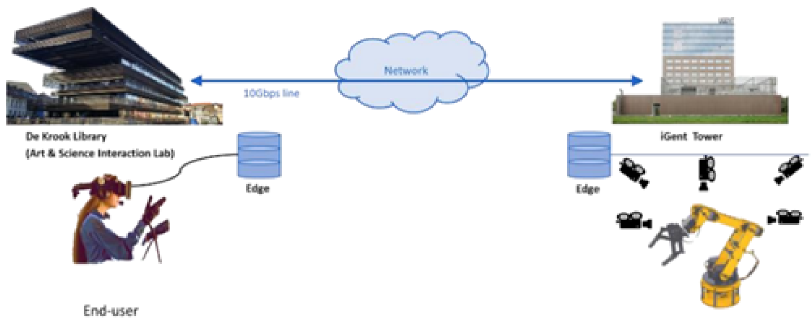

Fig. 4. The envisioned experimental testbed

depending on the use case at hand. After each session, the user's feeling of presence and cybersickness will be assessed by means of the previously discussed questionnaires. As such, a comparison between these and the output of the objective models can be performed to estimate the performance of the application. In addition, non-intrusive data-collection related to usability (e.g. time to complete a task, number of attempts...) and cybersickness (e.g. EEG) will be performed during testing by means of an *Emotive Epoc X*[5] headset and the *Empathica E4* [6] wristband.

Based on this testbed, two use cases are put forward to provide proof-of-concept and to demonstrate the resulting immersiveness models in a realistic, real-time environment. First, a VR scenario will be prepared, in which the user can virtually exercise a simple set of tasks, chosen such that they represent some key characteristics of the system in terms of usability. These include (i) the guidance of a deformable ball through a maze, (ii) a "Simon says" alike game that measures reaction speed and (iii) a wire loop game mainly focusing on accuracy. As such, the availability of the system in terms of basic, fundamental tasks (i.e. accurately moving objects, reacting to specific events...) for virtual training scenarios in industry and education can be assessed.

The second scenario will implement a remote presence system, such that the real-time performance of the system can be illustrated. The testbed will be designed such that the user is able to control a humanoid robot based on visual and haptic feedback over a network connection. As such, the simple task of locating an object, picking it up and placing it in a basket should be performed where the same evaluation mechanisms as for the VR scenario are in place. As a result, the influence of delay on the system can be evaluated, especially for more critical tasks in environments with stringent timing conditions such as real-time robotic surgery.

5 CONCLUSION

Traditional subjective and objective perception measurements do no longer fulfill the requirements of immersive, multimodal multimedia. In this position paper, we have revisited this traditional stance by describing immersiveness as a four-factor concept consisting of presence, cybersickness, QoE and usability. Based on the presented two-dimensional user-centric approach, the main contribution of this work is threefold. First, a concrete approach is presented for both subjective and data-driven assessment of the four key components of immersiveness in future multi-modal multimedia. Second, we illustrate our envisioned methodology towards the real-time modelling of immersiveness using network-related and application layer metrics as well as physiological data as an input. To support this vision, we illustrated by means of some preliminary results on point cloud QoE modelling. Third, we present two proof-of-concept scenarios, i.e. VR virtual training and remote presence, to make the aforementioned

[5]https://www.emotiv.com/epoc-x/
[6]https://www.empatica.com/research/e4/

ideas more tangible. As such, we believe this work provides a valuable addition to the state-of-the-art of subjective assessment and modelling and immersive multimedia in general.

ACKNOWLEDGMENTS

Sam Van Damme and Maria Torres Vega are funded by the Research Foundation Flanders (FWO), grant numbers 1SB1822N and 12W4819N, respectively.

REFERENCES

[1] Evangelos Alexiou and Touradj Ebrahimi. 2017. On subjective and objective quality evaluation of point cloud geometry. In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Erfurt, Germany, 1–3. https://doi.org/10.1109/QoMEX.2017.7965681

[2] Evangelos Alexiou and Touradj Ebrahimi. 2017. On subjective and objective quality evaluation of point cloud geometry. In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Erfurt, Germany, 1–3. https://doi.org/10.1109/QoMEX.2017.7965681

[3] Evangelos Alexiou and Touradj Ebrahimi. 2020. Towards a Point Cloud Structural Similarity Metric. In *2020 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*. IEEE, London, UK, 1–6. https://doi.org/10.1109/ICMEW46912.2020.9106005

[4] Muhammad Shahid Anwar, Jing Wang, Wahab Khan, Asad Ullah, Sadique Ahmad, and Zesong Fei. 2020. Subjective QoE of 360-Degree Virtual Reality Videos and Machine Learning Predictions. *IEEE Access* 8 (2020), 148084–148099. https://doi.org/10.1109/ACCESS.2020.3015556

[5] Thiago Braga Rodrigues, Ciarán Ó Catháin, Noel E O'Connor, and Niall Murray. 2020. A Quality of Experience assessment of haptic and augmented reality feedback modalities in a gait analysis system. *PLoS One* 15, 3 (2020), e0230570.

[6] Alberto Chang, Lourdes Montalvo, and Freddy Paz. 2019. Usability and Playability Heuristics for Augmented Reality Video Games in Smartphones. In *Design, User Experience, and Usability. Practice and Case Studies*, Aaron Marcus and Wentao Wang (Eds.). Springer International Publishing, Cham, 19–29.

[7] Michael Chinen, Felicia S. C. Lim, Jan Skoglund, Nikita Gureev, Feargus O'Gorman, and Andrew Hines. 2020. ViSQOL v3: An Open Source Production Ready Objective Speech and Audio Metric. In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Athlone, Ireland, 1–6. https://doi.org/10.1109/QoMEX48832.2020.9123150

[8] Simon Davis, Keith Nesbitt, and Eugene Nalivaiko. 2014. A Systematic Review of Cybersickness. In *Proceedings of the 2014 Conference on Interactive Entertainment* (Newcastle, NSW, Australia) *(IE2014)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/2677758.2677780

[9] John F. Golding. 2006. Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences* 41, 2 (2006), 237–248. https://doi.org/10.1016/j.paid.2006.01.012

[10] Rania Hassen and Eckehard Steinbach. 2018. HSSIM: An Objective Haptic Quality Assessment Measure for Force-Feedback Signals. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Sardinia, Italy, 1–6. https://doi.org/10.1109/QoMEX.2018.8463365

[11] Andras Kemeny, Jean-Rémy Chardonnet, and Florent Colombet. 2020. *Reducing Cybersickness*. Springer International Publishing, Cham, 93–132. https://doi.org/10.1007/978-3-030-59342-1_4

[12] Juno Kim, Wilson Luu, and Stephen Palmisano. 2020. Multisensory integration and the experience of scene instability, presence and cybersickness in virtual environments. *Computers in Human Behavior* 113 (2020), 106484. https://doi.org/10.1016/j.chb.2020.106484

[13] Panagiotis Kourtesis, Simona Collina, Leonidas A. A. Doumas, and Sarah E. MacPherson. 2019. Validation of the Virtual Reality Neuroscience Questionnaire: Maximum Duration of Immersive Virtual Reality Sessions Without the Presence of Pertinent Adverse Symptomatology. *Frontiers in Human Neuroscience* 13 (2019), 13 pages. https://doi.org/10.3389/fnhum.2019.00417

[14] Ochs Magalie, Jain Sameer, and Blache Philippe. 2018. Toward an Automatic Prediction of the Sense of Presence in Virtual Reality Environment. In *Proceedings of the 6th International Conference on Human-Agent Interaction* (Southampton, United Kingdom) *(HAI '18)*. Association for Computing Machinery, New York, NY, USA, 161–166. https://doi.org/10.1145/3284432.3284452

[15] Jesus Mayor, Laura Raya, and Alberto Sanchez. 2021. A Comparative Study of Virtual Reality Methods of Interaction and Locomotion Based on Presence, Cybersickness, and Usability. *IEEE Transactions on Emerging Topics in Computing* 9, 3 (2021), 1542–1553. https://doi.org/10.1109/TETC.2019.2915287

[16] Miroslaw Narbutt, Andrew Allen, Jan Skoglund, Michael Chinen, and Andrew Hines. 2018. AMBIQUAL - a full reference objective quality metric for ambisonic spatial audio. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Sardinia, Italy, 1–6. https://doi.org/10.1109/QoMEX.2018.8463408

[17] Steve Nebel and Manuel Ninaus. 2019. New Perspectives on Game-Based Assessment with Process Data and Physiological Signals. In *Game-Based Assessment Revisited*, Dirk Ifenthaler and Yoon Jeon Kim (Eds.). Springer International Publishing, Cham, 141–161.

https://doi.org/10.1007/978-3-030-15569-8_8

[18] Magalie Ochs, Jeremie Bousquet, and Philippe Blache. 2019. Multimodal Cues of the Sense of Presence and Co-Presence in Human-Virtual Agent Interaction. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents* (Paris, France) *(IVA '19)*. Association for Computing Machinery, New York, NY, USA, 218–220. https://doi.org/10.1145/3308532.3329438

[19] Marta Orduna, César Díaz, Lara Muñoz, Pablo Pérez, Ignacio Benito, and Narciso García. 2020. Video Multimethod Assessment Fusion (VMAF) on 360VR Contents. *IEEE Transactions on Consumer Electronics* 66, 1 (2020), 22–31. https://doi.org/10.1109/TCE.2019.2957987

[20] Marta Orduna, César Díaz, Lara Muñoz, Pablo Pérez, Ignacio Benito, and Narciso García. 2020. Video Multimethod Assessment Fusion (VMAF) on 360VR Contents. *IEEE Transactions on Consumer Electronics* 66, 1 (2020), 22–31. https://doi.org/10.1109/TCE.2019.2957987

[21] Janne Paavilainen. 2017. Playability: A Game-Centric Definition. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 487–494. https://doi.org/10.1145/3130859.3131306

[22] Lisa Rebenitsch and Charles Owen. 2016. Review on cybersickness in applications and visual displays. *Virtual Reality* 20 (2016), 101–125. Issue 2. https://doi.org/10.1007/s10055-016-0285-9

[23] Tyler Rose and Karen B. Chen. 2018. Effect of levels of immersion on performance and presence in virtual occupational tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62, 1 (2018), 2079–2083. https://doi.org/10.1177/1541931218621469 arXiv:https://doi.org/10.1177/1541931218621469

[24] N. Sakr, N.D. Georganas, and J. Zhao. 2007. A Perceptual Quality Metric for Haptic Signals. In *2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games*. IEEE, Ottawa, ON, Canada, 27–32. https://doi.org/10.1109/HAVE.2007.4371581

[25] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603 arXiv:https://direct.mit.edu/pvar/article-pdf/10/3/266/1623697/105474601300343603.pdf

[26] Jean-Christophe Servotte, Manon Goosse, Suzanne Hetzell Campbell, Nadia Dardenne, Bruno Pilote, Ivan L. Simoneau, Michèle Guillaume, Isabelle Bragard, and Alexandre Ghuysen. 2020. Virtual Reality Experience: Immersion, Sense of Presence, and Cybersickness. *Clinical Simulation in Nursing* 38 (2020), 35–43. https://doi.org/10.1016/j.ecns.2019.09.006

[27] Meryem Simsek, Adnan Aijaz, Mischa Dohler, Joachim Sachs, and Gerhard Fettweis. 2016. 5G-Enabled Tactile Internet. *IEEE Journal on Selected Areas in Communications* 34, 3 (2016), 460–473. https://doi.org/10.1109/JSAC.2016.2525398

[28] Lea Skorin-Kapov, Martín Varela, Tobias Hoßfeld, and Kuan-Ta Chen. 2018. A Survey of Emerging Concepts and Challenges for QoE Management of Multimedia Services. *ACM Trans. Multimedia Comput. Commun. Appl.* 14, 2s, Article 29 (may 2018), 29 pages. https://doi.org/10.1145/3176648

[29] Yule Sun, Ang Lu, and Lu Yu. 2017. Weighted-to-Spherically-Uniform Quality Evaluation for Omnidirectional Video. *IEEE Signal Processing Letters* 24, 9 (2017), 1408–1412. https://doi.org/10.1109/LSP.2017.2720693

[30] Maria Torres Vega, Cristian Perra, Filip De Turck, and Antonio Liotta. 2018. A Review of Predictive Quality of Experience Management in Video Streaming Services. *IEEE Transactions on Broadcasting* 64, 2 (2018), 432–445. https://doi.org/10.1109/TBC.2018.2822869

[31] Huyen T. T. Tran, Nam Pham Ngoc, Cuong Manh Bui, Minh Hong Pham, and Truong Cong Thang. 2017. An evaluation of quality metrics for 360 videos. In *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, Milan, Italy, 7–11. https://doi.org/10.1109/ICUFN.2017.7993736

[32] International Telecommunication Union. 1999. ITU-T Rec. P.910: Subjective video quality assessment methods for multimedia applications.

[33] International Telecommunication Union. 2006. ITU-T Rec. P.10: Vocabulary for performance and quality of service.

[34] Sam Van Damme, Maria Torres Vega, Joris Heyse, Femke De Backere, and Filip De Turck. 2020. A low-complexity psychometric curve-fitting approach for the objective quality assessment of streamed game videos. *Signal Processing: Image Communication* 88 (2020), 115954. https://doi.org/10.1016/j.image.2020.115954

[35] Sam Van Damme, Maria Torres Vega, and Filip De Turck. 2021. A Full- and No-Reference Metrics Accuracy Analysis for Volumetric Media Streaming. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, Montréal, Canada, 225–230. https://doi.org/10.1109/QoMEX51781.2021.9465420

[36] Jeroen van der Hooft, Maria Torres Vega, Tim Wauters, Christian Timmerer, Ali C. Begen, Filip De Turck, and Raimund Schatz. 2020. From Capturing to Rendering: Volumetric Media Delivery with Six Degrees of Freedom. *IEEE Communications Magazine* 58, 10 (2020), 49–55. https://doi.org/10.1109/MCOM.001.2000242

[37] Jacinto Vasconcelos-Raposo, Maximino Bessa, Miguel Melo, Luis Barbosa, Rui Rodrigues, Carla Maria Teixeira, Luciana Cabral, and António Augusto Sousa. 2016. Adaptation and Validation of the Igroup Presence Questionnaire (IPQ) in a Portuguese Sample. *Presence* 25, 3 (2016), 191–203. https://doi.org/10.1162/PRES_a_00261