Convergence between TV and VR: Enabling Truly Immersive and Social Experiences

Mario Montagud, Juan A. Núñez, Themistoklis Karavellas, Ibai Jurado, Sergi Fernández
i2CAT Foundation
C/ Gran Capità 2-4, Edifici Nexus I, Barcelona, Spain
{mario.montagud, juan.antonio.nunez, t.karavellas, ibai.jurado, sergi.fernandez}@i2cat.net

ABSTRACT
A seamless convergence between the connected TV and Virtual Reality (VR) ecosystems can open to the door to new fascinating scenarios. This paper initially introduces three EU H2020 projects (ImmersiaTV, ImAc and VRTogether) addressing relevant and complementary aspects towards achieving this integration. Next, an overview of the VRTogether project is provided. This project aims at enabling truly social VR experiences between remote users in domestic scenarios, based on photorealistic immersive content and on TV-related scenarios. Although the project is still in an early stage, its key objectives and requirements derived via user-centric activities are briefly introduced. Finally, the envisioned use cases and planned evaluations, together with the infrastructure and component architecture being deployed to support these scenarios, are also presented.

Author Keywords
Copresence, Immersive TV, Immersive Virtual Environments, Social VR, Togetherness, Virtual Reality.

ACM Classification Keywords
H.5.1 Information interfaces and presentation (e.g., HCI). Multimedia Information Systems; H.1.2 User/Machine Systems: Human Factors

INTRODUCTION
Virtual Reality (VR) can bring significant benefits in many scenarios, such as e-learning, e-tourism and Immersive TV. In particular, a convergence between connected TV ecosystems and VR contents and devices can open to the door to new fascinating scenarios. In this context, VR can contribute not only to increase the users' engagement, but also to increase their feeling of immersion, co-presence and networked togetherness (i.e., the feeling of being together in common virtual scenarios, while apart).

In the last years, our research group has been intensively researching towards achieving a seamless integration between the TV and VR areas. These research efforts have been in part supported by the participation in, and coordination of, three H2020 European projects (see logos in Figure 1).

Figure 1. VR-related projects in which authors are involved.

First, ImmersiaTV project (http://www.immersiatv.eu/) targets at overcoming the existing challenges to enable *customizable and immersive multi-screen TV experiences*. By considering the current heterogeneity in terms of contents and consumption devices, with a special focus on omnidirectional media (e.g. 360º video and spatial audio), ImmersiaTV proposes backward-compatible and standard-compliant re-definitions to the end-to-end chain to make these new experiences a reality. In particular, novel forms of content capturing, production, storytelling, encoding, distribution and consumption are being created. These contributions allow providing different versions and formats of audiovisual and VR contents, which can be interactively selected and presented on multiple devices in multi-screen scenarios, in a coherent, intuitive and personalized manner (see Figures 2 and 3). Demo videos of ImmersiaTV can be watched at: https://goo.gl/WCF4J3

Second, ImAc project (www.imac-project.eu) explores how accessibility services can be efficiently integrated with immersive media, such as omnidirectional and VR contents, while keeping compatibility with current standards and technologies. The idea of ImAc is not to consider accessibility as an afterthought, but as a key aspect in the specification and deployment of end-to-end immersive systems and services. It involves overcoming existing limitations in current technologies and systems in order to not only enable customizable and immersive experiences, but also *truly inclusive experiences*. Accessibility is a key requirement to ensure an effective narrative and interpretation of immersive contents, as well as a proper usability of the immersive services, regardless of the sensorial and cognitive capacities of the users, their age, language, and/or other specific impairments and difficulties.

This will contribute to a global e-inclusion, offering equal opportunities of access to the whole consumers' spectrum, and complying with worldwide regulations (e.g. Human Rights obligations). Demo videos of ImAc can be watched at: https://goo.gl/Kf8zoB and https://goo.gl/ckAn3o

Third, VRTogether (http://vrtogether.eu/) aims at offering ground-breaking *truly social VR experiences* between users located in remote domestic scenarios, based on photorealistic immersive content, and in a cost-effective manner. On the one hand, the project considers the assembling of an end-to-end workflow integrating state-of-the-art technologies and off-the-shelf components. On the other hand, innovative solutions and optimizations for key technological and creative aspects along this end-to-end workflow will be devised. Different contents and scenarios will be created with the goal of conducting both controlled experiments and open pilots for: i) assessing the proper performance of the designed and adopted technological components; ii) exploring the suitability and optimizing different users' reconstruction and content formats; iii) assessing what evaluation methodologies and metrics are best suited in such environments; and iv) determining the benefits provided by such technologies and scenarios.

Our participation in the workshop will be based on sharing the outcomes from these three projects, and the lessons learned during their evolution. We believe it is very appropriate forum to experience rich interactions and discussions with researchers and practitioners interested in how VR can enrich the TV ecosystem, and explore potential synergies and collaborations. The three presented projects fit within the scope of the workshop, and contribute to complementary aspects of how TV can meet VR, and vice versa. Nonetheless, VRTogether seems to be the project with a stronger relationship with the workshop themes. Therefore, the rest of the paper concentrates on VR-Together, by providing an overview of its objectives, requirements, envisioned evaluations and scenarios, and the infrastructure and platforms to be deployed. Although the project is still in an early stage (started in October 2017), we believe such an overview can raise interesting discussions and interactions during the workshop.

VRTOGETHER: OBJECTIVES AND MILESTONES

VRTogether pursues the next five main objectives:

Objective 1. Develop and integrate new media formats that deliver high quality photo-realistic content and create a strong feeling of co-presence in coherently integrated experiences.

Objective 2. Adapt the existing production pipeline to capture and encode multiple media formats and integrate them with state-of-the-art post-production tools.

Objective 3. Re-Design the distribution chain in order to allow an effective and scalable orchestration and delivery of the considered innovative content formats.

Objective 4. Explore and specify appropriate Quality of Experience (QoE) metrics and evaluation methodologies to determine the quality and benefits of these new social VR experiences.

Objective 5. Maximize the impact across content creators, producers, distributors, tooling companies, service providers and, particularly, the general audience.

In line with these objectives, the next key milestones are expected during the evolution of the project:

- **Outside-in point cloud capture setup**. Photorealistic capture of actors based on point-clouds that allow parallax, usable in a production environment, for human replicants.
- **Point-cloud encoding**. A system to distribute efficiently point clouds, integrated with video and 3D mesh streaming.
- **High quality scenario capturing**. Using either photogrammetry or manual modelling and photo texturing to provide high quality and realistic capturing of scenarios.
- **Home video capture setup**. Live video capture that can be seamlessly blended with the content being delivered, and being usable in domestic environment.
- **(Live) Orchestration and rendering of distributed streams**. Signalization, delivery and synchronization solutions to enable coherent virtual shared spaces between distributed users, involving a variety of audiovisual streams.

Figure 2. Envisioned multi-screen scenarios in ImmersiaTV.

Figure 3. Multi-screen consumption experience for an opera use case, including multi-location omnidirectional videos and spatial audio.

VRTOGETHER: USER-CENTRIC ACTIVITIES

VRTogether is built around three 3 pillars: 1) requirements gathering; 2) development and integration; and 3) validation and dissemination. The project adopts a user-centric design, in which end-users, professionals and stakeholders are involved since the beginning of the project through the organization of workshops, focus groups, evaluations, and the attendance to events. This allows determining with higher precision what are the functional and non-functional requirements to be provided, mainly in terms of the scenarios to be considered, the component architecture to be developed and integrated, and the features to be provided and evaluated.

From the user-centric activities, a list of requirements has been derived, although it is still being extended and refined. The requirements can be classified into general requirements, which are common for the whole VRTogether ecosystem, and specific requirements for each of the three iterative pilots considered in the project (described in next section). A summary of the key general requirements for VRTogether is provided in Table 1.

Copresence. End users (equipped with the corresponding capture and visualization systems) should be able to be immersed in a virtual space and engage in real-time face-to-face social activities. Copresence should lead to both self-awareness (i.e. embodiment) and other-awareness, enabling natural communications, social interactions, coherence and responsiveness.
Distributed shared experiences. End users should be able to be immersed in shared virtual spaces from distributed locations.
Place and plausibility illusion. End users inside a virtual space should have the feeling of being in the physical space depicted in the VR content.
Content adaptability and blending. Different formats of high-quality (photorealistic) VR contents should be supported, for both the virtual space and end-users' representations, and these contents should be seamlessly blended (i.e. perceived as a whole).
Adaptive media delivery. Media delivery should be dynamically adapted, based on the capabilities and conditions of the network and involved devices.
Synchronization. Accurate (temporal and spatial) synchronization between media contents and locations must be guaranteed to enable coherent shared media experiences.

Table 1. Key general requirements for VRTogether.

VRTOGETHER: EXPERIMENTS AND PILOT SCENARIOS

The partners of VRTogether will carry out different types of evaluations, focused on assessing different aspects of the project. In particular, four categories of experiments are foreseen, each addressing relevant dimensions for VRTogether:

- **Technological dimension**. This category of experiments mainly focuses on viability and performance. Among other objectives, this category comprises: content distribution, latency, synchronization, optimization of computational resources, replacing the Head Mounted Displays (HMDs) by the users' faces, etc.
- **Content dimension**. It explores what types and/or combinations of content formats, for both the virtual environments and end-users' representation, are feasible and provide the best results.
- **QoE dimension**. Experiments with users to determine QoE-related aspects, such as: their perception when using different combinations of contents and scenarios, in different conditions; usability aspects; their interest in such scenarios; and their overall satisfaction.
- **Psychological dimension**. The experiments under this category go beyond the previous dimension, by analyzing psychological aspects, such as the feeling of being there, the feeling of being together, and naturalness. For such purposes, users' gaze, head movements, physiological signals, interaction patterns, activities, gestures and behaviors, etc., may be taken into account.

These experiments will be conducted in local and distributed scenarios, and will be targeted at assessing both Quality of Service (QoS) and QoE aspects, determining the most proper metrics to be employed and methodologies to follow, determining the need for specifying additional ones. The insights from these experiments will serve as a basis for the main project activities in terms of evaluation: the pilots. Indeed, the pilots will the project checkpoints to evaluate the creative and technical challenges addressed towards the creation of truly realistic social VR experiences, and validate their usefulness. These partial experiments can make use either of ad-hoc contents and scenarios created specifically for them or of the ones created for the pilots.

Pilot Scenarios

The project has initially scheduled three pilots, aiming at demonstrating and validating complementary aspects, with an increased level of complexity, but keeping a coherent storyline across them. The first pilot mainly focuses on communication and interaction features between remote participants while conducting an activity or sharing an experience together, the second one addresses scalability issues, and the third one focuses on increasing the interaction possibilities. Next, these three pilots are briefly presented:

Pilot 1 (July–September 2018). **Police investigation theme**. The first pilot focuses on an offline production of an intrigue scene where a police officer interrogates a series of murder suspects. The idea consists of extending that virtual scenario with the addition of two seated characters, associated to two remote captured users, placed next to each other and being separated by a one-way mirror to the interrogation room (see Figure 4). This way, the users can be involved in a unique virtual experience, having the possibility of sharing their impressions (e.g., about the authority of the crime, or about the experience itself!), by naturally communicating (hearing and seeing each other and themselves), as if they were located in a common physical location. The specific key requirements for Pilot 1 are indicated in Table 2.

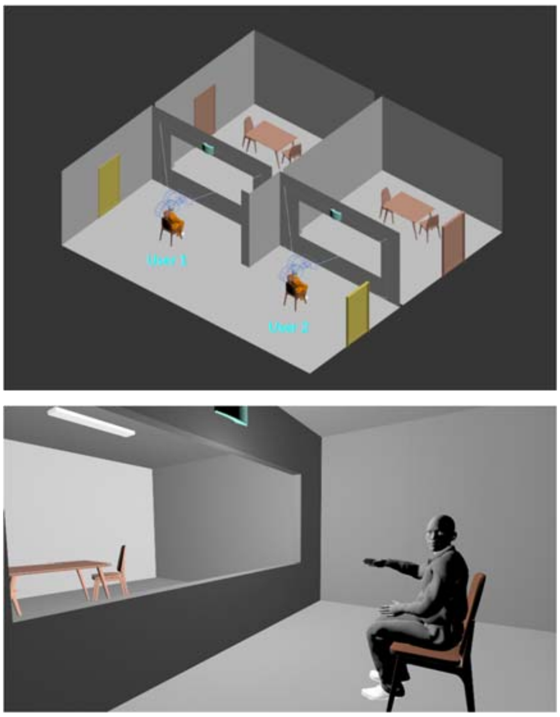

Figure 4. Interrogation scene (up) and position of the characters or users' representation (down).

User's Reconstruction and Character Rendering. The scene must integrate the rendering of characters obtained through user's reconstruction techniques.
Gaze. Rendered characters should be able to retarget their gaze according user's viewpoint.
Basic end user movements. Users will be seated, but can rotate their head and have certain level of translation capacity (3DoF+).

Table 2. Key specific requirements for Pilot 1.

Pilot 2 (June–August 2019). **Live news**. This pilot will consist of a live production of multi-source immersive contents, with the ability of bringing users to the location where TV broadcasted news occur, thus maximizing the connectedness between (multiple) remote users (i.e., the audience) and the news. The goal is to make the audience to feel like being there together where the news actually happen. Key challenges in these scenarios rely on providing a feeling of closeness and an empathic connection between the audience and the content, while maximizing the feeling of immediacy and realism. The additional key requirements for Pilot 2 are indicated in Table 3.

Live Services. The system must provide support for the live delivery of photorealistic immersive VR environments.
Facial Expressions. Sufficient detail to see facial expressions should be available in the end-user and character representations.
Higher Scalability. The system must support up to 10 simultaneous end-users (in different rooms/locations).

Table 3. Key specific requirements for Pilot 2.

Pilot 3 (May–July 2020). **Interactive Fiction**. This pilot seeks to demonstrate how the VR-Together platform, in a custom-designed content production process, can allow for a novel form of content where users meet and blend within an interactive and immersive experience. Herein, users will become active characters within the story plot being rendered, be able to move, interact with the environment and even manipulating it. Therefore, the goal will be to explore to what extent and how a fiction scenario can be rendered in VR, while still allowing remote users to actively intervene in the scene being broadcasted, preserving the feeling of being there together, consistence and credibility, as well as maximizing place illusion and plausibility. The additional key requirements for Pilot 3 are indicated in Table 4.

Active Intervention. End users can become characters within the story plot being rendered, and be able to interact with the environment and manipulate it.
Movement. End users can move (translation), enabling 6DoF.
Interactive storytelling. The system will integrate interactive storytelling engines.

Table 4. Key specific requirements for Pilot 3.

VRTOGETHER: INFRASTRUCTURE & ARCHITECTURE

The iterative user-centric activities being carried out in VRTogether serve to gather requirements and take decisions about the platform (technical aspects) and the use cases (creative aspects). The consortium has devoted efforts on deploying a permanent distributed User Lab to vertebrate such activities. The User Lab will be composed of three *lab nodes* (see Figure 5), with the required (capturing, processing and rendering) equipment and communication infrastructure to not only being able to perform the required experimental tests, but also serve as a demonstrator of the project contributions to interested agents. The three lab nodes will be located in Spain (Barcelona, at i2CAT premises), The Netherlands (Amsterdam, at CWI premises), and Greece (Thessaloniki, at CERTH premises). In addition to these lab nodes, all project partners also have at their disposal lab facilities with partial functionalities of the VR-Together platform, required for their planned specific experiments.

Finally, high-level overviews of the related component and hardware architectures for the VRTogether platform are provided in Figures 6 and 7, respectively. Their relationships can also be inferred from the notations used in both figures.

CONCLUSIONS

This paper has initially provided an overview of three ongoing projects that address relevant and complementary aspects to achieve a seamless and effective integration between the TV and VR ecosystems. Then, it has focused on presenting and discussing the key objectives of VRTogether project, and the requirements that have been, and are still being, derived via a plethora of user-centric activities. The evaluations (including both specific experiments and open pilots) to be conducted, the available infrastructure for that, as well as the component and hardware architectures to be setup, have been also briefly introduced. Although the project is still in an early stage, and no major outcomes have been achieved yet, we believe such an overview and research plan can lead to rich interactions and discussion during the workshop. Overall, VRTogether is a very challenging project, from which we expect high-impact contributions and a huge reach.

ACKNOWLEDGMENTS

This work has been funded by European Union's Horizon 2020 program, under agreement nº 762111 (VRTogether).

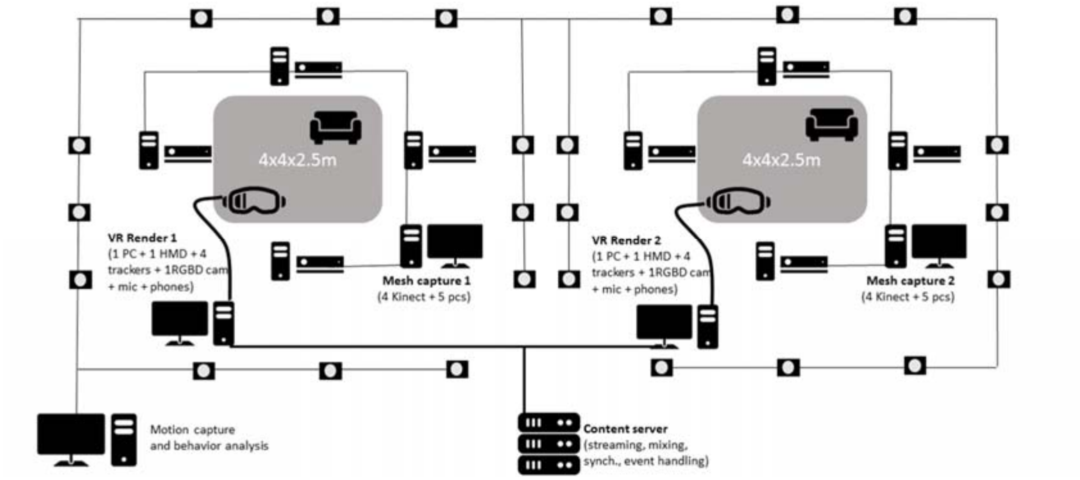

Figure 5. Composition of each of three lab nodes that form part of the distributed VR-Together User Lab.

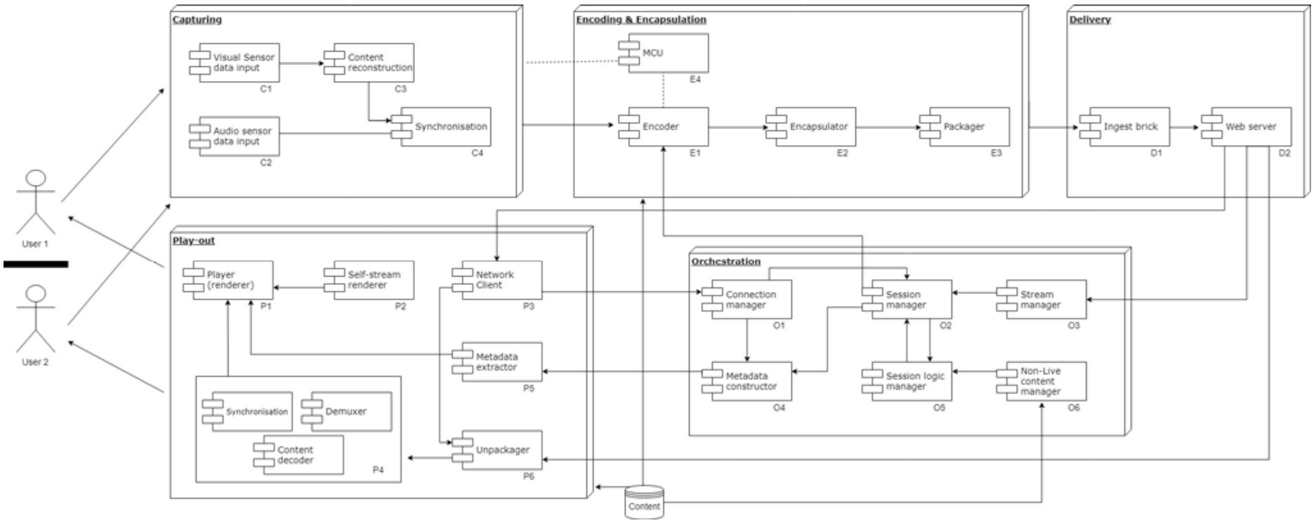

Figure 6. Component Architecture for the VRTogether platform.

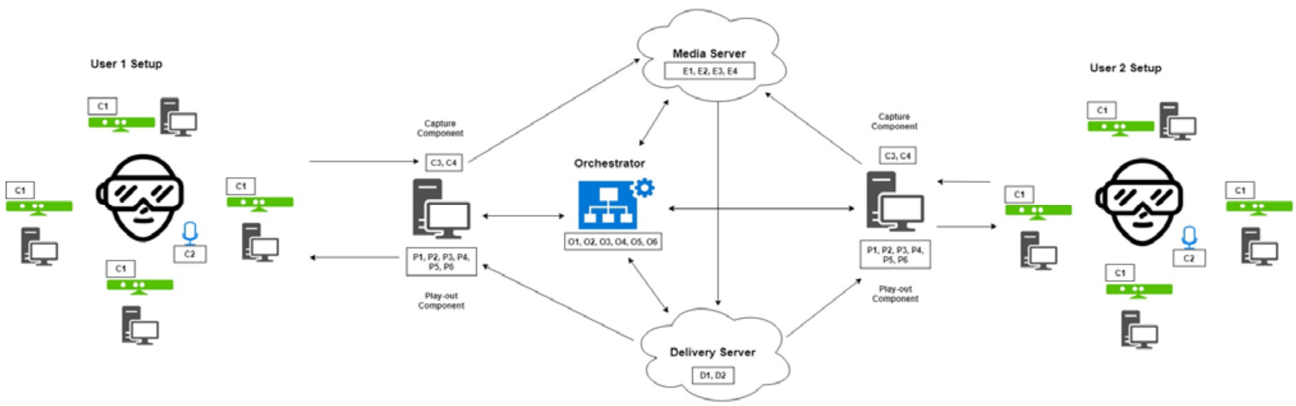

Figure 7. Hardware Architecture for the VRTogether platform.