ImAc: Enabling Immersive, Accessible and Personalized Media Experiences

Mario Montagud
I2CAT Foundation
Barcelona, Spain
mario.montagud@i2cat.net

Juan A. Nuñez
I2CAT Foundation
Barcelona, Spain
juan.antonio.nunez@i2cat.net

Isaac Fraile
I2CAT Foundation
Barcelona, Spain
isaac.fraile@i2cat.net

Sergi Fernández
I2CAT Foundation
Barcelona, Spain
sergi.fernandez@i2cat.net

Abstract
The integration of immersive contents and consumption devices within the TV landscape brings new fascinating opportunities. However, the exploitation of these immersive TV services is still in its infancy and ground-breaking solutions need to be devised. A key challenge is to enable truly inclusive experiences, regardless of the sensorial and cognitive capacities of the users, their age and language. In this context, ImAc project explores how accessibility services (subtitling, audio description and sign language) can be efficiently integrated with immersive media, such as omnidirectional and Virtual Reality (VR) contents, while keeping compatibility with current standards and technologies. This paper provides an overview of the project, by focusing on its motivation, the followed user-centered methodology and its key research objectives. The end-to-end system (from production to consumption) being specified, the envisioned scenarios and planned evaluations are also briefly described.

Author Keywords
Accessibility Services, Immersive Media, Multi-Screen Scenarios, Spatial Audio, Subtitles, Virtual Reality.

ACM Classification Keywords
H.5. Information interfaces and presentation (e.g., HCI).

ImAc: Immersive Accessibility

Goal: Seamless integration between accessibility services and immersive media, by enabling personalization and keeping compatibility with current technologies and standards.

Immersive Contents: High-Resolution Media, 360º Video, and Spatial Audio.

Accessibility Contents: subtitling, audio description, and sign language interpreting. Assistive technologies (e.g. voice recognition, guiding mechanisms). Adapted User Interfaces (UI). Personalization.

Use Cases: Single- and Multi-Screen Scenarios, combining traditional and VR consumption devices.

Figure 1: ImAc logo.

Introduction

Research on TV-related content consumption is being intense in the last years. One of the most active research topics consists of enabling successful and coherent multi-screen experiences (e.g. [1, 8, 9]). A step beyond, the integration of omnidirectional contents and devices within the TV landscape can open the door to new fascinating scenarios (e.g. [3]).

These customizable, interactive and immersive multi-screen TV experiences can increase the users' engagement, but also bring new challenges. A key requirement of every emerging media technology and scenario is to consider accessibility, in order to ensure a proper narrative, interpretation of contents and usability, 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, while ensuring compliance with regulatory guidelines (e.g. Human Rights Obligations). In this context, subtitles play a key role in TV-related scenarios to better interpret the audiovisual contents. Previous works have targeted at proposing solutions for an adaptive and customizable presentation of subtitles, even in multi-screen scenarios (e.g. [6, 7]) and for 360º videos (e.g. [2]). However, many challenges still need to be overcome in order to efficiently integrate accessibility services (not only subtitling, but also audio subtitling, sign language interpreting and audio description) in immersive environments.

ImAc is a H2020 European project (www.imac-project.eu), October 2017-March 2020, that explores how accessibility services can be efficiently integrated with immersive media, such as omnidirectional and Virtual Reality (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 enable truly inclusive, immersive and personalized multi-screen experiences, adapted to the particular needs and/or preferences of the consumers.

Among others, ImAc aims at providing efficient solutions and/or meaningful insights to the following research questions:

- What are the requirements in order to enable truly accessible and inclusive immersive services?
- How current (immersive) technologies and systems can be augmented in order to seamlessly integrate and support accessibility services?
- What personalization features should be provided in order to meet particular users' needs and/or preferences?
- To what extent accessibility services in immersive environments contribute to a higher e-inclusion and to equal opportunities?
- What scenarios and use cases would be (more) benefited by a seamless integration between immersive and accessibility services?
- What kind of assistive technologies should be adopted?
- Which presentation modes for accessibility contents are better suited in the envisioned scenarios?

Presentation of Accessibility Contents in Immersive Environments:

A proper presentation of accessibility contents when consuming immersive media can contribute to a higher e-inclusion, but is more complex than for traditional media. On the one hand, there is more information to process and users can get overwhelmed. On the other hand, the presentation is no longer purely time-based, but involves a spatial dimension, determined either by the user's viewing point or by the direction where the main action is taking place. Therefore, proper presentation modes will be considered to assist to a higher comprehension and to guide users. The most proper presentation modes may depend on the specific users' profiles, their sensorial capacities and preferences. Therefore, adaptability and personalization become essential in this context.

- What are the comfortable viewing fields for accessibility contents, especially when using Head Mounted Displays (HMDs)?
- What benefits are provided (e.g., in terms of usability, content comprehension, level of immersion and engagement…)? How to properly evaluate and determine them?

The remaining of the paper provides a brief overview of the ImAc project. Its relationship with other related projects is indicated in Section 2. The employed methodology is described in Section 3. The end-to-end system to be specified and implemented, focusing on the consumer side, is introduced in Section 4. The envisioned scenarios, use cases and pilot evaluations are then presented in Section 5. Finally, conclusions and the planned roadmap are provided in Section 6.

Relationship with other Projects

ImAc departs from the expertise, insights and contributions from other recent related projects. On the one hand, HbbTV4all project has addressed accessibility in the emerging connected hybrid broadcast broadband media ecosystem, within the umbrella of Hybrid Broadcast Broadband TV (HbbTV) standard [4]. ImAc seeks the same success story within immersive environments. On the other hand, ImmersiaTV targets at overcoming 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, ImmersiaTV proposes backward-compatible and standard-compliant re-definitions to the end-to-end chain to make these new experiences a reality.

Methodology

ImAc is built around three 3 pillars: 1) requirements gathering, 2) development and integration; and 3) validation and dissemination. A simplified diagram of the chosen methodology is illustrated in Figure 1. It must be remarked that the project follows a user-centric approach, in which end-users, professional users 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 the accessibility requirements, desired features and scenarios, as well as the required new technologies or extensions to existing ones in order to maximize usability and deployment. The ImAc platform specification, services to be developed, in addition to the deployed pilots, will be based on the gathered insights from the user-centric activities.

The pilots and evaluations will allow validating the contributions and results, but also refining the gathered requirements, functionalities and architecture of the ImAc platform.

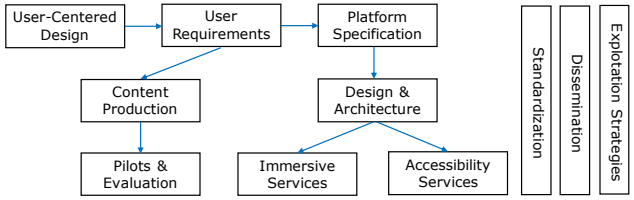

Figure 1: Methodology followed in ImAc.

End-to-End ImAc System

This section provides an overview of the end-to-end ImAc system specification and functionalities, derived

from the functional requirements gathered in the conducted user-centric activities and their transformation into technical requirements. Exhaustive details about the architecture, components and modules of the system are not provided, as they are still being currently specified. Basically, the system is divided into four parts (Figure 2): production/edition, management, preparation/distribution and consumption. The key modules / functional blocks in these parts are indicated in Figure 2, where green color indicates these components are under the umbrella of ImAc, orange indicates they have been developed in other projects (e.g. ImmersiaTV), and white that are out-of-scope of the project, but form part of the end-to-end workflow in typical systems.

Content Production: tools for the production, authoring and edition of accessibility contents, and for their integration with classical and immersive media services. It includes the definition of metadata and signalization to be provided to the Service Prover part.

Service Provider: where the management of programs is handled, including Media Asset Management (MAM), linking of additional contents to main TV programs and scheduling playout.

Content Preparation & Distribution: handles distribution via various technologies. This part covers content encoding and packaging, signalization, and distribution via DVB and IP-based CDNs (e.g. using DASH [5]).

Content Consumption: includes all required components for the presentation of the available ImAc contents, including traditional, immersive and accessibility contents, to end-users. In ImAc, traditional TV contents will be played out on connected TVs. The immersive (360º videos and spatial audio) and accessibility contents will be played out on companion screens (e.g., tablets, HMDs…) via a web-based player being developed. The player will support different presentation modes and include customization options. Different interaction functionalities and modalities will be also provided. In addition, inter-media and inter-device sync functionalities will be provided to enable a time-aligned presentation of all contents across the involved devices. A high-level overview of the components of the ImAc player is provided in Figure 3.

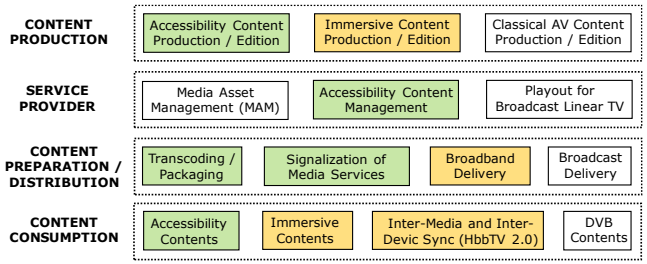

Figure 2: Main components of the ImAc system.

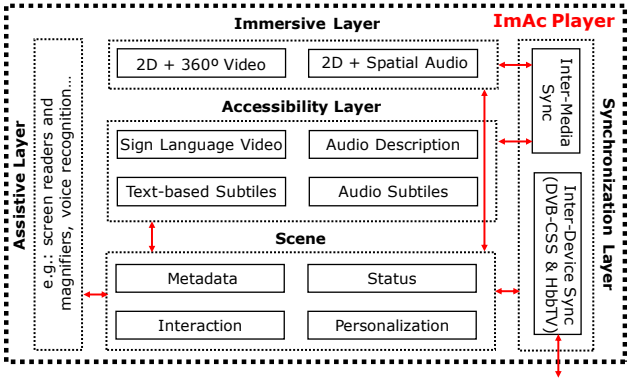

Figure 3: Main Components of the ImAc player.

Scenarios and Pilots

ImAc considers both single- and multi-screen scenarios, composed of a main TV and companion screens (e.g., tablets, smartphones, HMDs). In the latter scenarios, proper discovery, association, app launching and media sync solutions will be provided (Figure 4).

The user-centered activities will determine the contents to be created, specific scenarios to be developed and demonstration pilots to be set up (Figure 1). As part of its objectives, ImAc will perform two (large scale and open) national pilots plus a cross-national one.

In addition, in a joint initiative with ImmersiaTV, ImAc has participated in an extra pilot: an immersive recording of "Roméo et Juliette" at the Gran Teatre del Liceu Opera House, in Barcelona. The goal is to create a ground-breaking product that allows consumers enjoying an opera performance in a highly interactive, immersive, personalized and accessible manner. In particular, multiple video cameras (4 360º, 2 320º, 4 170º, and 5 directive cameras) and more than 80 audio sources (including 3D –Ambisonics– and binaural formats), distributed both on and off the stage, were used in the recording. This allows experiencing the opera performance from the preferred viewpoint, using the preferred media formats and devices, and being able to dynamically switch between viewpoints and devices. In relation to this, the presented audio matches the selected camera position and current viewpoint, providing detailed sound landscapes and thus highly immersive and realistic experiences. Finally, a customizable, adaptive and assistive presentation of accessibility contents related to the 360º videos is also supported. Regarding subtitles, different colors for each speaker can be added, and different font sizes and languages can be dynamically selected. Subtitles are, by default, presented at the bottom region, as in typical video players. However, it is also possible to present them at the top region of the player (Figure 5). This presentation option is known as super-titles or surtitles, and is typically used and preferred in musical and theatre performances. In order to properly identify the active speaker(s) and main actions while freely exploring the 360º area, different guiding mechanisms are provided, such as adding arrows, a compass, or sided text. Similar guiding mechanisms are also provided for sign language (Figure 6) and audio description. The playback of all the selected contents will be accurately synchronized, regardless on the number and types of consumption devices being used in multi-screen scenarios (Figure 7). A demo video can be watched at: https://goo.gl/xpqVPF

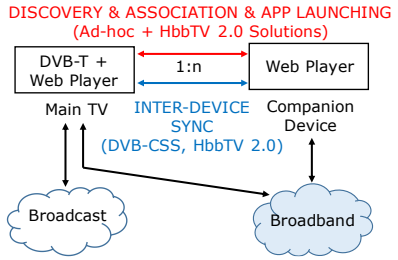

Figure 4: Functionalities in the multi-screen scenarios.

Figure 5: Presentation of Subtitles in 360º Environments.

Figure 6: Presentation of Sign Language Videos in 360º Environments.

Figure 7: Immersive, Accessible and Personalized Multi-Screen Scenarios considered in ImAc.

Conclusions and Future Work

This paper has provided an overview of the ImAc project, by focusing on its motivation, the followed user-centered methodology and its key research objectives. The end-to-end system being specified, the envisioned scenarios and planned evaluations have been also briefly described. As the project advances in time, it is expected to refine the research objectives and plan, to achieve relevant contributions and to reach a big impact.

Acknowledgements

This work has been funded by European Union's Horizon 2020 program, under agreement nº 761974 (ImAc project).

References

1. F. Boronat, D. Marfil, M. Montagud, J. Pastor, "HbbTV-Compliant Platform for Hybrid Media Delivery and Synchronization on Single- and Multi-Device Scenarios", IEEE Transactions on Broadcasting, vol. PP, no. 99, pp. 1-26, 2017.
2. A. Brown, J. Turner, J. Patterson, A. Schmitz, M. Armstrong, M. Glancy, "Subtitles in 360-degree Video", ACM TVX '17, June 2017.
3. D. Gómez, J.A. Núñez, M. Montagud, S. Fernández, "ImmersiaTV: Enabling Customizable and Immersive Multi-Screen TV Experiences", ACM MMSYS'18, June 2018.
4. Hybrid Broadcast Broadband TV (HbbTV) 2.0.1 Specification, HbbTV Association Resource Library, https://www.hbbtv.org/resource-library, July 2016.
5. ISO/IEC 23009-1: 2012. Information Technology. Dynamic Adaptive Streaming over HTTP (DASH). Part 1: Media Presentation Description and Segment Formats. April 2012.
6. M. Montagud, F. Boronat, J. González, J. Pastor, "Web-based Platform for Subtitles Customization and Synchronization in Multi-Screen Scenarios", ACM TVX '17, June 2017.
7. A. Rodriguez, G. Talavera, P. Orero, J. Carrabina "Subtitle Synchronization across Multiple Screens and Devices", Sensors, 12(7), June 2012
8. V. Vinayagamoorthy, R. Ramdhany, M. Hammond, "Enabling Frame-Accurate Synchronised Companion Screen Experiences", ACM TVX '16, June 2016.
9. C. Ziegler, et al., "On Time or Not on Time: A User Study on Delays in a Synchronised Companion-Screen Experience", ACM TVX '17, June 2017.