

IMMERSIVE MEDIA AND ACCESSIBILITY: HAND IN HAND TO THE FUTURE

Mario Montagud^{1,2}, Pilar Orero³, Sergi Fernández¹

¹i2CAT Foundation, Spain, ²Universitat de València, Spain, ³Universitat Autònoma de Barcelona, Spain

Abstract – *Advances in immersive media can provide relevant benefits to our society. The future looks promising with the development of new communication technologies, formats and hardware, together with their wide applicability areas. However, a fundamental aspect remains unexplored in this media landscape evolution: accessibility. A lack of accessibility hinders the interaction and immersion with virtual reality (VR) experiences to many consumers, becoming a barrier for global e-inclusion and equality of opportunities. Making immersive media accessible then becomes a priority and thus the research topic covered in this article. After reviewing the societal impact and regulatory framework, the article identifies current advances, limitations and challenges in this field. Then, it presents technological solutions and insights from initiatives targeted at achieving a seamless presentation of access services (subtitling, audio description and sign language) with VR360 content. Although these contributions are primarily conceived for broadcast services, their applicability spans other sectors and can pave the way for next-generation accessible VR experiences with six Degrees-of-Freedom (6DoF).*

Keywords – 360° video accessibility, immersive media, subtitling, virtual reality.

1. INTRODUCTION

The unceasing advances in information and communication technologies (ICT) are profoundly impacting our society and daily lives, providing relevant benefits in a wide set of sectors. On the one hand, advances in communication technologies not only enable new forms of adaptive, ubiquitous and hybrid content delivery, but also new combinations of interaction possibilities, between content modalities, devices and users. On the other hand, apart from advances in terms of compression and resolution, novel media formats are being conceived to represent the content and information. In particular, the recent explosion of immersive media technologies, like Virtual Reality (VR), is opening the door to new fascinating opportunities and revenue models, replacing and/or augmenting the media services based on classic audiovisual formats (e.g., TV, web, conferencing), bringing a paradigm shift from passive consumption to rich interaction and immersion.

In this context, 360° videos, a.k.a immersive or VR360 videos, have become a simple and cheap, yet effective and hyper-realistic, way to provide VR experiences. VR360 videos can be enjoyed via traditional devices (like PCs, laptops, smartphones) or VR devices (like Head Mounted Displays, HMDs). Especially when watched on VR devices, VR360 videos can provide unique experiences, placing viewers at the center of omnidirectional scenes, allowing them to freely look anywhere. This gives

viewers the illusion of actually ‘*being there*’ in the captured environment, and greatly improves both immersion and engagement compared to the use of traditional media.

The media industry is becoming increasingly aware of the opportunities that VR can enable in terms of communication, entertainment and learning, among other relevant benefits. Indeed, major platforms, content providers and cultural institutions are exploring how VR360 videos can provide more innovative, hyper-personalized and immersive experiences to their audience [1]. This trend is also a reality within the broadcast sector. According to a recent report issued by the European Broadcasting Union (EBU) [2], 49% of its members are devoting efforts into the production and integration of immersive content and services.

However, many challenges still need to be addressed in order to efficiently integrate immersive content into current media services, and to fully exploit its potential. While significant effort is being devoted to specifying adequate technologies and formats, exploring new storytelling solutions and ensuring backward compliance with traditional resources and services, there is still a fundamental aspect that remains unexplored: accessibility. The media sector is currently offering a plethora of (standard) accessibility solutions and guidelines for classic audiovisual services. However, this is not yet true when it comes to immersive media. Despite the

growing interest in, and adoption of, VR360 experiences, research studies on, and thus solutions for, accessibility in immersive media are limited so far. This hinders the interaction of a significant percentage of the population with VR experiences.

Everyone has the right to access and comprehend the information provided by media services in general, and by immersive experiences in particular. Therefore, accessibility becomes a priority, not only to adhere to current worldwide regulations, but also to reach a wider audience and contribute to equal opportunities and global e-inclusion. Immersive experiences need to be inclusive across different languages, addressing not only the needs of consumers with hearing and vision impairments, but also of those with cognitive and/or learning difficulties, low literacy, newcomers, and the aged. Accessibility in immersive media is the research topic addressed in this article, with a particular focus on the broadcast sector, and on achieving a seamless integration between the presentation of access services (subtitling, audio description and sign language) and VR360 content (360° videos and spatial audio).

As a proof of relevance of this topic, the paper initially highlights its societal impact, by reflecting on the need for providing accessible media services and by outlining the regulatory framework that aims to guarantee the accessibility of every multimedia service, and monitor the fulfillment of these demands and obligations. After that, the paper discusses advances, limitations and challenges towards making immersive media accessible, with a special focus on the broadcast sector. Finally, the paper presents novel contributions from initiatives in this direction, in terms of technological solutions (end-to-end platform, extensions to current technologies and formats, interaction modalities) and recommendations (presentation modes, guiding methods, evaluation methodologies). Although these contributions are mostly focused on the broadcast sector, their applicability spans other sectors (e.g., learning, culture and heritage [1]), and can pave the way for next-generation accessible VR experiences, with six Degrees-of-Freedom (6DoF), like gaming and virtual meetings.

2. SOCIETAL IMPACT

Technological advances typically provide significant social and financial benefits, but may also raise barriers for some citizens. No homogeneous profiles for consumers exist, and

factors like disabilities, ageing and capabilities (e.g., skills) influence the interaction and engagement with immersive technologies and experiences.

The study in [1] provides statistics about the percentages of the worldwide population with some form of audiovisual impairment, and about the increasing ageing process. Ageing is closely related to disability rates, as older people may develop age-related impairments, such as visual, hearing, physical and cognitive ones [1].

Disabilities, sensorial impairments and ageing have functional, social, emotional and economic impacts. This involves a barrier to properly access and interpret the information, which is a right of all citizens, in all sectors and services. In addition, linguistic and cultural capabilities need to be also taken into account. Society has become global, and people commonly travel or move to different countries for professional and/or leisure reasons, so language and culture may also become a barrier.

All these factors have an impact in the media accessibility field, which has traditionally mostly focused on an audience with disabilities, but now there is a need to take into account other audiences, such as the elderly, non-native speakers, illiteracy, etc., to enable a full democratic participation in the current society where technologies and media play a key role. E-inclusion thus becomes a priority of worldwide governments, and this is also the case in the EU, as proposed by the Single Digital Market (SDM) strategy that aims to guarantee a seamless exchange of media assets across countries and communities. Likewise, the United Nations Centre for Regional Development (UNCRD) [3] provides a framework for action in a number of articles highlighting the interrelations between ageing and disability, especially in article 9 (accessibility).

In this context, two important United Nations (UN) conventions have legal repercussions within the scope of media accessibility. The Convention on the Protection and Promotion of the Diversity of Cultural Expressions [4] identified the elderly as an audience sector to be protected against exclusion. This is further expanded in the Convention of Rights of Persons with Disabilities [5], since the elderly may not be classified as disabled, but are the highest number of citizens with hearing and/or sight loss. These two UN conventions have been developed in two Directives and one Act affecting all EU countries:

- 1) The EU Directive on the Accessibility of Websites and Mobile Applications [6]. This is based on Web Content Accessibility Guidelines (WCAG) 2.0 [7] guidelines, and references EN 301 549 [8] as the standard which will enable websites and apps to meet accessibility requirements.
- 2) The Audiovisual Media Services Directive (AVMSD) [9]. This governs a coordination of national legislations on audiovisual media, addressing key issues like new technological developments, preserving diversity, protecting the whole spectrum of consumers, and safeguarding media pluralism.
- 3) The European Accessibility Act [10]. This is a law that aims to make media products and services more accessible. It takes the form of a Directive, which is legally binding for all EU Member States.

These pieces of legislation demand accessibility solutions in all media services, products and events, including the offered content and information. This therefore also applies to immersive media. Monitoring and control strategies are being established by public sector organizations to ensure these accessibility demands are successfully met.

3. STATE-OF-THE-ART

3.1 Related projects in the broadcast sector

Typically, accessibility has been only considered after a technology has matured to meet the demand of the mass market, and to fulfill legal requirements. This approach, however, results in significant disadvantages and barriers for the population with disabilities, the aged, and those who do not speak or read the language of the media content. However, when it comes to novel and not-yet mainstreamed services, the development of accessibility solutions becomes more challenging. It is due to the fact that accessibility not only becomes an issue of sensorial disabilities, but of capabilities, skills and usability aspects [11]. However, considering accessibility since the early development of novel technologies contributes to more effective solutions, a wider adoption and less financial burden.

Different research projects have targeted the provision of accessibility solutions within the broadcast sector, as soon as new TV-related technologies or formats have been developed. The DTV4ALL project addressed access service provision when TV moved from analogue to digital

in Europe [12]. Results from this project are still valid in some regions across the world where the shift has not taken place yet (e.g., Sweden). Then, the HBB4ALL project [13] explored new accessibility possibilities in the emerging hybrid broadcast broadband media ecosystem, within the umbrella of the Hybrid Broadcast Broadband TV (HbbTV) standard [14]. This hybrid multiplatform ecosystem enables to augment broadcast services with additional broadband content, to be consumed on the same main screen as the broadcast content or on companion screens (i.e., tablets and smartphones), in an interactive and personalized manner. HBB4ALL targeted at leveraging these features in order to seamlessly integrate the presentation of access services (subtitling, audio description and sign language) with traditional broadcast programmes. In the same context of hybrid media delivery, the ImmersiaTV project developed an end-to-end platform to enable customizable and immersive multiscreen TV experiences [15]. By adopting novel storytelling formats, the contributions from ImmersiaTV allowed augmenting traditional broadcast TV programmes with multilocation 360° views and interactive content (e.g., notifications, picture-in-picture videos or images) delivered via broadband.

Most recently, the ImAc project [16] has aimed at following the same success story as in the DTV4ALL and HBB4ALL projects, but focusing on immersive media formats (i.e., VR360 videos and spatial audio), and also exploiting the expertise from the ImmersiaTV project.

3.2 Research on accessibility in immersive media

As for every media content, VR360 videos need to be accessible. VR360 videos include audio and visual elements which, at the same time, may contain linguistic and non-linguistic components. A user meaning-making process is possible thanks to the intersection and interaction between all these diverse components, which contribute to a comprehensive information understanding. Therefore, alternatives must be provided to cater for the needs of users who cannot access all audiovisual components. This is commonly provided by means of access services, like subtitling, audio subtitling, audio description and sign language. As reviewed in [1], while scarce research on proposing subtitling solutions for VR360 content exists, that is yet not the case for the other access services. The study in [1] also highlights existing

challenges towards making immersive media accessible. These challenges are not only related to the specification of technological aspects and presentation modes for the access services, but to adopt the most appropriate interaction modalities (e.g., voice control, gestures), assistive technologies (e.g., zoom, guiding methods) and user interfaces (UIs). Moreover, the work in [17] provides a set of 14 tools to enhance the interaction of people with low vision with VR environments, by means of augmentati.

4. IMAC: END-TO-END PLATFORM FOR IMMERSIVE ACCESSIBILITY

ImAc (Immersive Accessibility) is an European H2020 project (www.imac-project.eu) that has explored how accessibility services and assistive technologies can be efficiently integrated with immersive media, focusing on VR360 video and spatial audio.

The project has been carried out by a cross-disciplinary consortium, composed of international experts in the fields of content creation and service provisioning (e.g., broadcasters), multimedia systems and accessibility, with the goal of contributing with appropriate solutions and guidelines to the identified limitations and needs in this field. The project has adopted a user-centric methodology where end users, professionals and stakeholders are involved at every stage of the project through the organization of workshops, focus groups, tests, and the attendance at events. This elicited a bottom-up approach to determining

the accessibility, interaction and personalization requirements. The insights from the user-centric activities in turn determined the necessary technological solutions to meet these requirements, and the specific services and scenarios to be provided.

An overview of the end-to-end platform developed in ImAc to achieve the targeted goals is provided in Fig. 1. The platform includes all the necessary parts, components and steps from media ingest/authoring to media consumption. It has been designed with the premise of keeping backward compliance with current technologies, formats, infrastructures and practices in the broadcast ecosystem. Next, its key parts and components are briefly explained with the goal of providing a proof of concept of how worldwide deployed services, like TV and online video, can be extended to accommodate accessible immersive experiences. For each of the components and parts, potential standardization opportunities and contributions are indicated.

4.1 Content production / authoring

The scenarios envisioned in ImAc include classic audiovisual content to be broadcast, and both immersive and accessibility content are to be also delivered via broadband (see Fig. 1). However, only the authoring and production of access service content has been addressed in the project, assuming that the classic and immersive content has already been produced (e.g., by using existing tools, like the ones in [15]), although their preparation for media delivery is supported.

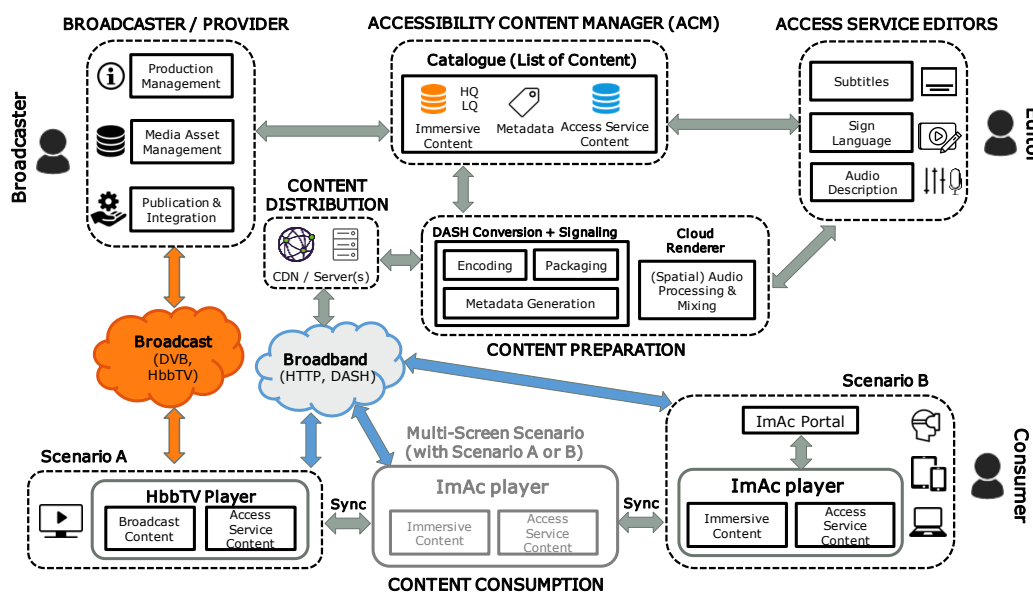


Fig. 1 – Overview of an end-to-end platform

Accordingly, the content production part of the platform includes a set of (web-based) tools for the ingest of immersive VR360 content, the production/editing of access service content (including subtitles, audio description and sign language videos), and their integration with the available VR360 content. These tools provide the required metadata to signalize the properties of the authored media content. The supported features are not detailed, but can be inferred from the features provided by the developed VR360 player [18] (described later). The access service editors are provided under a Software as a Service (SaaS) distribution model.

In this context, the ImAc project has contributed to the specification of standard-compliant extensions to current technologies and formats. Two examples can be highlighted. First, the IMSC (Internet Media Subtitles and Captions) subtitle format has been extended to indicate where the speaker is in the 360° space. IMSC is a subset of Timed Text Markup Language (TTML) for distribution of subtitles, drawing the attention of, and being adopted by, many standardization bodies, like the World Wide Web Consortium (W3C). Second, a new ad-hoc, but standard-compliant, solution has been defined to be able to indicate the speaker's name and position together with the associated sign language video, as well as to be able to dynamically show/hide the video window based on the signer's activity.

4.2 Service provider

This part of the platform includes different components for the management and cataloguing of content, and triggering its publication, either by associating it to scheduled (broadcast) TV programs or by posting it on specific websites. One key component in this part is the Accessibility Content Manager (ACM), developed to enable the upload of VR360 content (and optionally its related assets, like personalized covers/thumbnails and existing access service content), to catalogue it, to link it with the editing tools, and finally to trigger the publication of the immersive and accessibility content for final consumption.

4.3 Content preparation / distribution

This part of the platform includes components for preparing the available content for an appropriate distribution. These components are mainly in charge of (multi-quality) content encoding, segmentation and signaling.

The project has focused on media delivery through broadband Content Delivery Networks (CDNs), via the Dynamic Adaptive Streaming over HTTP (DASH) standard [19]. However, the platform is also prepared for the delivery of broadband content via DASH, but as an enrichment of conventional broadcast Digital Video Broadcasting (DVB) services, by leveraging the features provided by the HbbTV standard [14].

In this context, ImAc has contributed with the release of open-source software components and open APIs to connect them with the content production tools. In addition, ImAc has explored the specification of standard-compliant extensions to delivery technologies and signaling solutions. One example is the design of solutions to enable interactive and immersive multiscreen scenarios, which can be relevant for the HbbTV standard. Another example is the extension of the DASH technology to signalize the availability of all considered access services, beyond the classic and immersive media content elements. This can be relevant for the DASH technology specified by the Moving Picture Experts Group (MPEG) standardization body.

4.4 Content consumption

This part is mainly composed of a portal to select the desired content (see Fig. 2) and a web-based player (see Fig. 3) for the personalized presentation of the immersive (VR360 video and spatial audio) and access service (subtitles, audio subtitles, audio description and sign language) content. For visual elements, the player supports different presentation modes such as:

- User-reference: the visual elements (e.g., subtitles, sign language, UI) are always visible and follow the user's viewing direction.
- Scene-reference: the visual elements are attached to specific regions of the 360° space (e.g., fixed angles, position of the speaker), so they can be outside the user's field of view during media consumption.
- Auto-positioning: the visual elements or the user's field of view can be dynamically positioned, based on predefined metadata generated at the production phase.

For instance, Fig. 4 includes an example of the presentation of user-referenced subtitles with a radar as the guiding method, and Fig. 5 of the presentation of a user-referenced sign language video with the speaker's name and arrows to indicate where the speaker is in the 360° space.

The player also includes support for different spatial presentation modes for (3D) audio description and audio subtitling such as:

- Classic mode: no positioning.
- Static mode: from a fixed point in the scene.
- Dynamic mode: coming from the direction of the action.

The player can be run on traditional consumer devices (e.g., connected TVs, PCs, laptops, tablets and smartphones) and on VR devices (e.g., HMDs). It also supports different interaction modalities, like via the mouse, keyboard, touch screen, gyroscope, VR controls, or even via voice.

This part of the ImAc platform also includes the proper technology to enable multiscreen scenarios in a synchronized and interactive manner, using both fully web-based and HbbTV-compliant technology. More details about the player and its supported features can be found in [18].

A demo video of the portal and player features can be watched at: <https://bit.ly/2Wqd336>. Their current version and the available VR360 and access service content can be accessed via this URL: <http://imac.i2cat.net/player/>. The source code can be downloaded from: <https://github.com/ua-i2cat/ImAc>.

In this context, the ImAc project has explored the specification of standard solutions for subtitle presentation modes (e.g., within W3C), and for appropriate UIs and interaction modalities (e.g., within International Organization for Standardization, ISO).

In general, the contributions from the ImAc project can be also of interest to groups of the International Telecommunication Union (ITU). Examples are: 1) the Intersector Rapporteur Group Audiovisual Media Accessibility (IRG-AVA), which studies topics related to audiovisual media accessibility and aims at developing draft Recommendations for “Access Systems” that can be used for all media delivery systems, including broadcast and broadband services; 2) the ITU-R Study Group 6 (SG 6), which addresses broadcasting scenarios, including vision, sound, multimedia and data services; and 3)

ITU-T SG16 – Multimedia, which includes a work item on presentation of sign language. All these groups and related activities are being monitored to get updated about latest news and contributions and to assess the feasibility of presenting contributions.

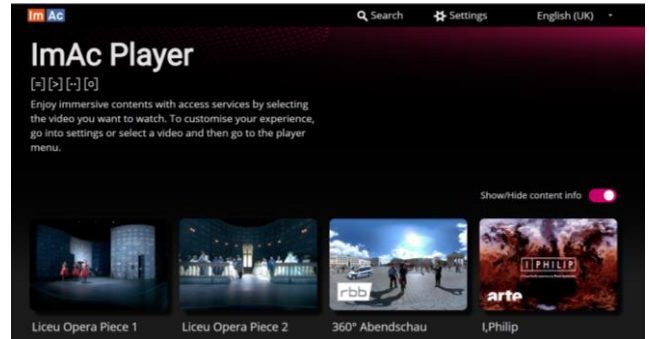


Fig. 2 – ImAc portal

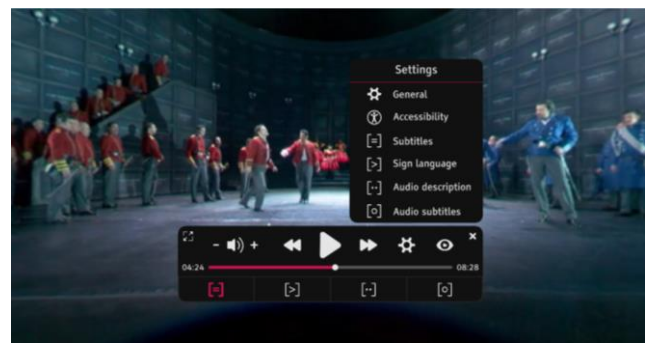


Fig. 3 – ImAc player / menu



Fig. 4 – Presentation of subtitles with radar indicator



Fig. 5 – Presentation of sign language video, with arrows and speaker's name (and color)

Table 1 provides a summary of key contributions and insights from the ImAc project, categorized by the involved platform components and parts.

Table 1 – Contributions and insights from the ImAc project

Platform components	Contributions / Recommendations
Content production / Authoring	<ul style="list-style-type: none"> – Extension to media formats (e.g., IMSC subtitles, accompanying metadata for sign language videos). – Narrative and storytelling (e.g., scripting and spatial presentation modes for audio description, positioning of subtitles) – Content creation guidelines – Presentation modes (e.g., spatial modes for audio description, non-continuous sign language video based on speaker's activity)
Service provider	<ul style="list-style-type: none"> – Service description and signaling
Content preparation / Distribution	<ul style="list-style-type: none"> – Content encoding, mixing and packaging – Content description and signaling
Content consumption	<ul style="list-style-type: none"> – Media interpretation – Interaction and hyper-personalization – Presentation modes for visual elements (user-reference, scene-reference, auto-positioning). – Presentation modes for 3D audio (gains, spatial combinations, scripting modes). – Interaction modalities (e.g., adaptive UIs, voice control) – Assistive technologies (e.g., voice control, augmentation, guiding methods) – Synchronization (intra-media, inter-media, inter-device). – Multiscreen scenarios (HbbTV, web) – Evaluation methodologies

5. CONCLUSION

Providing cutting-edge technology to all citizens is essential to achieve global e-inclusion. The use of VR technologies and, in particular, of VR360 content can provide great benefits in a wide variety of sectors. However, as in any (multimedia) service, accessibility becomes an essential requirement to achieve universal access to the information, and effective content comprehension by everyone.

The specification, development and evaluation of accessibility solutions for novel technologies, like VR, is more challenging than for traditional technologies. This is due to the fact that no other comparable accessibility solutions exist and the consumers are not yet used to these novel technologies. However, considering accessibility since the early development of novel technologies contributes to more effective solutions and to a

wider and earlier adoption, without handicaps for any citizen. This can also contribute to avoiding the proliferation of isolated and closed add-on solutions, by designing standard-compliant end-to-end solutions that encompass all stages from media authoring to media consumption.

This paper has presented an overview of the contributions of the ImAc project towards seamlessly integrating accessibility services within immersive media. Standardization possibilities and contributions in this context have been identified as well. Although the presented contributions are primarily focused on VR360 content and on the broadcast sector, their applicability spans to other areas, and can pave the way for next-generation accessible immersive experiences, considering volumetric media and 6DoF VR scenarios, even in multi-user settings.

ACKNOWLEDGEMENT

This work has been partially funded by the European Union's Horizon 2020 program, under agreement number 761974 (ImAc project). Work by Mario Montagud has been additionally funded by the Spanish Ministry of Science, Innovation and Universities with a Juan de la Cierva – Incorporación grant (with reference IJCI-2017-34611), and by Fundación BBVA with a “Beca Leonardo a Investigadores y Creadores Culturales 2018”. Work by Pilar Orero has been additionally funded by the Catalan Government with the funded TransMedia Catalonia research group, with reference 2019SGR113.

REFERENCES

- [1] M. Montagud, P. Orero, A. Matamala, “Culture 4 All: Accessibility-Enabled Cultural Experiences through Immersive VR360 content”, *Personal and Ubiquitous Computing*, 2020, DOI: 10.1007/s00779-019-01357-3.
- [2] Virtual Reality: How are public broadcasters using it? 2017. Retrieved from: <https://www.ebu.ch/publications/virtual-reality-how-are-public-broadcasters-using-it> European Broadcasting Union (EBU) Report. Last accessed in December 2019.
- [3] United Nations Centre for Regional Development (UNCRD) <https://www.un.org/development/desa/disabilities/disability-and-ageing.html>. Last accessed in December 2019.

- [4] UNESCO (2005) The Convention on the Protection and Promotion of the Diversity of Cultural Expressions <https://en.unesco.org/creativity/convention/texts>
- [5] UN (2006) Convention of Rights of People with Disabilities <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>
- [6] Directive on the Accessibility of Websites and Mobile Applications, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L2102>
- [7] WACAG 2.0 <https://www.w3.org/TR/WCAG20/>
- [8] Mandate 376 <http://mandate376.standards.eu/standard/technical-requirements/#9>
- [9] Audiovisual Media Services Directive (AVMSD), <https://ec.europa.eu/digital-single-market/en/audiovisual-media-services-directive-avmsd>
- [10] European Accessibility Act <https://data.consilium.europa.eu/doc/document/PE-81-2018-INIT/en/pdf>
- [11] B. Agulló, P. Orero and A. Matamala, "From disabilities to capabilities: testing subtitles in immersive environments with end users". *Hikma* 17, pp. 195-220, 2018.
- [12] T. Itagaki, T. Owens, P. Orero, Digital TV accessibility – Analogue switch off in Europe and Africa. *IST-Africa Week Conference*, 2016.
- [13] P. Orero, C.A. Martín, M. Zorrilla, HBB4ALL: Deployment of HbbTV services for all, *IEEE BMSB'15*, Ghent (Belgium), June 2015.
- [14] Hybrid Broadcast Broadband TV (HbbTV) 2.0.2 Specification. 2018. HbbTV Association Resource Library, <https://www.hbbtv.org/resource-library>, February 2018.
- [15] J. A. Núñez, M. Montagud, I. Fraile, D. Gómez, S. Fernández, ImmersiaTV: an end-to-end toolset to enable customizable and immersive multi-screen TV experiences, *Workshop on Virtual Reality, co-located with ACM TVX 2018*, Seoul (South Korea), June 2018.
- [16] M. Montagud, I. Fraile, J. A. Núñez, S. Fernández, "ImAc: Enabling Immersive, Accessible and Personalized Media Experiences", *ACM TVX 2018*, Seoul (South Korea), June 2018.
- [17] Y. Zhao, E. Cutrell, C. Holz, M. R. Morris, E. Ofek, A. D. Wilson. "SeeingVR: A Set of Tools to Make Virtual Reality More Accessible to People with Low Vision", *ACM CHI 2019 (CHI '19)*, Glasgow (UK), May 2019.
- [18] M. Montagud, I. Fraile, E. Meyerson, M. Genís, S. Fernández, "ImAc Player: Enabling a Personalized Consumption of Accessible Immersive Content", *ACM TVX 2019*, June, Manchester (UK).
- [19] I. Sodagar, "The MPEG-DASH standard for multimedia streaming over the internet", *IEEE multimedia*, 18(4), 62-67, 2011.