ImmersiaTV

Grant Number	688619
Period	1/2016-6/2018
Budget	3.8M€
Funding organism:	H2020 (EC)

Abstract

ImmersiaTV will create a novel form of broadcast omnidirectional video content production and delivery that offers end-users a coherent audiovisual experience across head mounted displays, second screens and the traditional TV set, instead of having their attention divided across them. This novel kind of content will seamlessly integrate with and further augment traditional TV and second display consumer habits. ImmersiaTV will assemble an end-to-end toolset covering the entire audiovisual value chain: immersive production tools, support for omnidirectional cameras, including ultra-high definition and high dynamic range images, and adaptive content coding and delivery, and demonstrate it through 3 pilot demonstrations addressing both on-demand and live content delivery

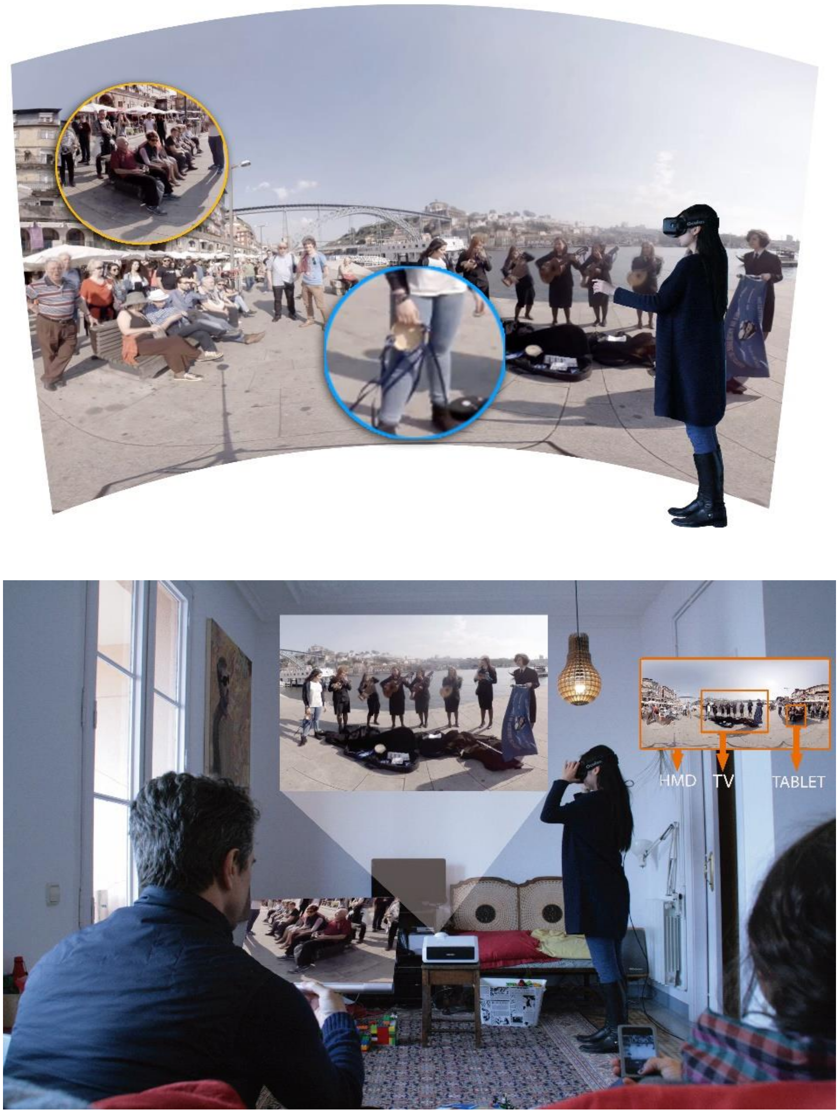

Figure 1: on top, Detail of the HMD view with an illustration of the portal approach. Bottom, . Illustration of the home set environment, with synchronized access across devices

Why ImmersiaTV?

The majority of European TV consumers now watch TV programs in a multi-display environment. Second screens -mostly smartphones, tablets or laptops- are generally used to check information not directly related to the events in the TV content being watched. As a result, the attention of the audience is generally divided between these different streams of information. Broadcasters have tried to orchestrate all these different rendering platforms to complement each other consistently. However, their success is limited, and this limited success is due, at least in part, to the very different formats in which information is delivered (web-based texts, mobile apps, traditional broadcast television…)

In this context, the arrival of immersive head-mounted displays to the consumer market introduces new possibilities, but also poses new challenges. Immersive displays impose radically different audience requirements compared to traditional broadcast TV and social media. They require a constant, frequently refreshed, omnidirectional audiovisual stream that integrates sensorimotor information. This means that, at minimum, the visual perspective rendered changes consistently with changes in head position and rotation. In addition, immersive displays challenge the conventions of traditional audiovisual language. For example, cuts between shots, which constitute the very basic fabric of traditional cinematic language, do not work well in immersive displays. From a user perspective, omnidirectional TV offers a new user experience and a different way of engaging with the audiovisual content.

This project will create new forms of digital storytelling and broadcast production that, by putting omnidirectional video at the center of the creation, production and distribution of broadcast content, delivers an all-encompassing experience that integrates the specificities of immersive displays, and the feeling of "being there", within the contemporary living room. We propose a new form of broadcast omnidirectional video that offers end-users a coherent audiovisual experience across head mounted displays, second screens and the traditional TV set, instead of having their attention divided across them. This new experience will seamlessly integrate with and further augment traditional TV and second screen consumer habits. In other terms: the audience will still be able to watch TV sitting on their couch, or tweet comments about it. However, by putting omnidirectional content at the center of the creation, production and distribution processes, the audience will also be able to use immersive displays to feel like being inside the audiovisual stream.

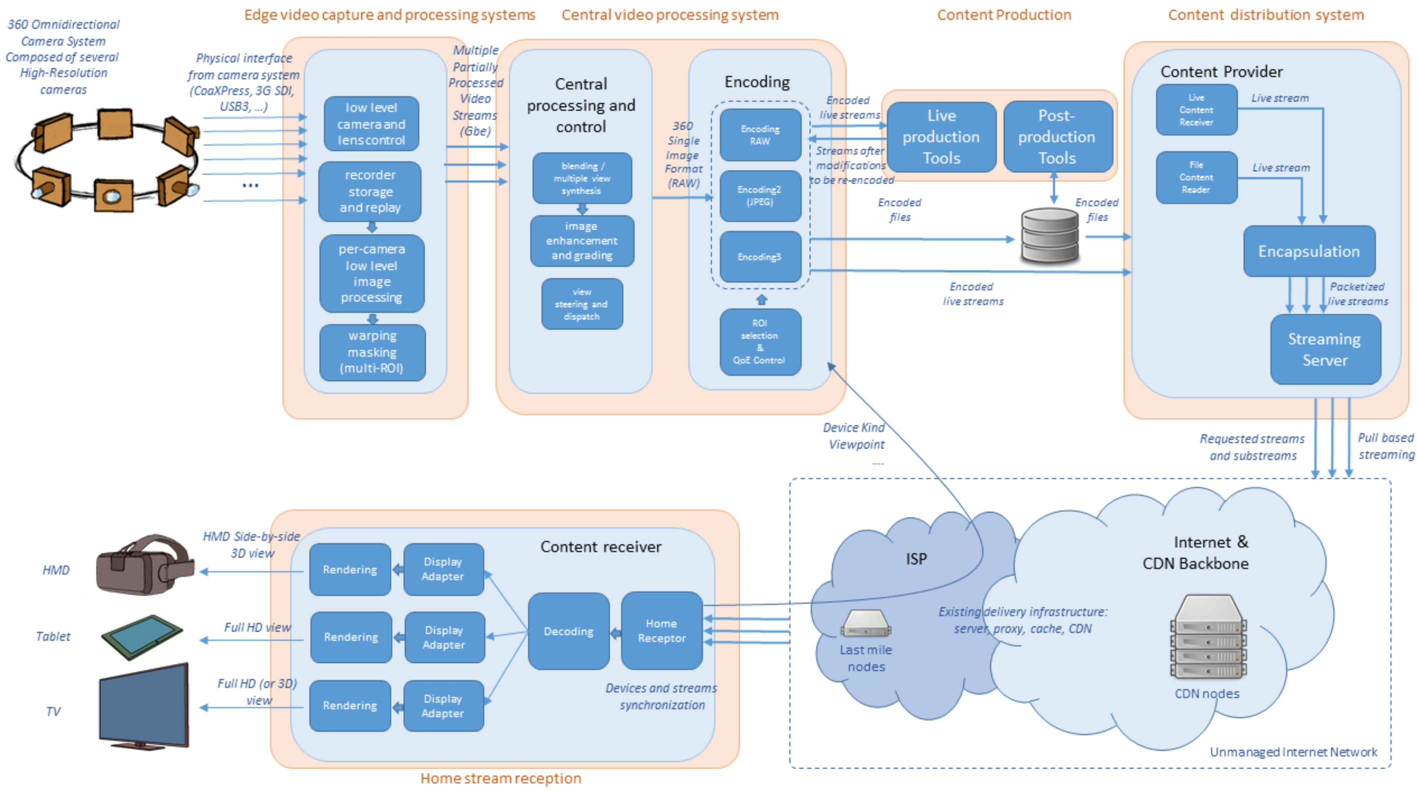

Figure 2. The ImmersiaTV Immersive Broadcast Platform.

Team

Project Coordinator

i2cat Foundation
Mr Sergi Fernandez
sergi.fernandez@i2cat.net

Tooling Companies	Research Institutes	Production Companies
IP-Based production pipeline	Stitching	Omnidirectional video production
Omnidirectional production tools	Innovative Codecs	Broadcasting
	User Evaluation	
VideoStitch	ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE	vrt
Cinegy	Institute of Bioorganic Chemistry Polish Academy of Sciences	LIGHTBOX
	iMinds	

Demonstration pilots to test an end-to-end system

ImmersiaTV will pilot an innovative **end-to-end system** covering the entire audiovisual value chain to enable a novel form of creative audiovisual storytelling based on omnidirectional video. The project will encompass immersive production tools, support for omnidirectional cameras, adaptive content coding and distribution mechanisms, and immersive (HMD) & second screen visualisation. ImmersiaTV will demonstrate the use of its end-to-end system in real production and distribution scenarios via 3 pilots addressing the production of documentaries and live sports event.

Specific Objectives

To demonstrate the feasibility of this novel approach for the creation, production, broadcast and display of omnidirectional video, ImmersiaTV has the following objectives:

- **OBJ1. Create a new cinematographic language** where the specificities of immersive displays are taken into account, and which conciliates immersive paradigms with traditional storytelling techniques.
- **OBJ2. Extend the production pipeline** to create omnidirectional content for a multi-platform environment.
- **OBJ3. Re-design the distribution chain** to address the specific technical challenges that omnidirectional content imposes in terms of capture, compression, distribution, reception, and rendering.
- **OBJ4. Maximize the quality of the end-user experience**, across devices, and within the technical limitations of existing production structures, distribution facilities and reception devices to create an optimal immersive experience.
- **OBJ5. Maximize the impact of the ImmersiaTV solutions** within the ecosystem of content creators, broadcasters, and consumers.

Specific Contributions

To achieve the previous objectives, ImmersiaTV will work to progress beyond the state of the art in the following domain areas.

Omnidirectional video capture	Immersia TV provides a new approach to video capture. At this time, there is no solution for automatic seamless stereoscopic full omnidirectional capture, but the market has been introducing different options: single-lens 360° cameras, custom rigs for multiple cameras and professional omnidirectional cameras. ImmersiaTV is developing a system for video capture, replay and processing based on a distributed architecture that is ground up designed for omni-directional video. The architecture consists of a novel kind of "edge" processors, that combine raw camera data storage and replay and low level image processing and camera control found commonly in cameras or recorders, with per-camera sub-tasks in the stitching process, now responsible for the need of a heavy compute cluster linked with the cameras via expensive long-range very high bandwidth links. The resulting partially processed video data with the ImmersiaTV capture system is ad-hoc coded to allow transport over affordable long-range links to be blended on a central processing unit that is much lighter than in current practice.
Immersive content production tools	ImmersiaTV is developing new and adapting existing content production tools, both for live and offline editing. Most of the current commercial solutions have been conceived from a classical video edition perspective, and therefore do not support efficiently omnidirectional video editing. This process becomes even tougher when multiple omnidirectional streams are managed jointly with regular streams. This leaves a short range of options for editing omnidirectional contents in offline scenarios. In this context, ImmersiaTV proposes a custom production workflow from the only commercial tool that currently exists for stitching and editing omnidirectional video live streams, Vahana VR, developed by the company Video-Stitch. This software solution is currently being combined with Cinegy's software solutions, capable of managing a large amount of video sources in offline and live production setups. ImmersiaTV complements these tools with low-level advanced algorithms in stitching, coding and distribution to provide a first real software based solution to edit complex production scenarios integrating omnidirectional video editing and distribution in both offline and live scenarios.
Encoding /Decoding of Omnidirectional video	ImmersiaTV is developing new codecs adapted to omnidirectional video. The encoding of omnidirectional video sequences currently requires sphere-to-plane mapping, since currently available off-the-shelf video compression algorithms where designed for rectangular video formats. ImmersiaTV proposes a solution based on state-of-the-art knowledge on this subject, providing codecs that go beyond the assumption of rectangularity. These novel coding schemes will also be adapted to a multi-platform multi-resolution reality through foveated coding, a technique that allows having

	different resolutions on a single image. The goal will be to balance dynamically the available bandwidth with the resolution, and select the resolution differently in particular regions of interest defined from the user's focus on a particular field of view of the scene being watched.
Content Delivery and Reception	Video distribution is one of the main challenges in Internet traffic, as it is one of the most resource demanding content format, not only in terms of bandwidth, but also concerning Quality of Service and Quality of Experience for users. In ImmersiaTV, we are performing a deep analysis of novel coding and decoding techniques to come up with a technology capable to satisfy the project requirements at 3 levels: in a broadcast production environment, in the main distribution infrastructure, and within the end-user local network, where several devices need to synchronize around the television set. Insofar, the main approach followed has been the use of CDNs on its two main variables (edge servers and edge networks). HTTP is well adapted to these solutions, despite it was not conceived to deliver video services. This fact brought market to take over other variants, such as HTTP file based streaming and Enhanced Streaming over UDP, but there is not still a leading option guiding these process. There are also challenges in the end-user local network: the content format developed in ImmersiaTV assumes that several devices synchronize around the television set. The distribution of omnidirectional video among synchronized devices in a local network requires taking into account that different devices will have a different performance regarding video processing. This fact implies that within the transmission pipeline additional buffering is required to ensure that all contents are visualized at the same time in all devices, smoothly switching from one to another.
Visualization and interaction	How should omnidirectional contents be consumed? What visualisation and interaction principles allow a good end-user experience? Head mounted displays seem to be naturally adapted to deliver and visualise immersive videos. However, end-users will still want to experience television with the language and conventions they are used to. They will also want to consume content through mobile phones and tablets. Broadcast content formats are not ready for this multiplatform reality. The central purpose of ImmersiaTV is to demonstrate a novel kind of audiovisual content format, compatible with immersive displays, coherent across devices, and ready for broadcast distribution. Moreover, natural technology evolution (e.g.: new devices launching) demands a holistic approach that could work with existing devices, but also with those that will be commercialized in a near future. Last but not least, interaction is also key in ImmersiaTV and it opens a huge range of possibilities for narrative and new experiences for users. The project will put great effort to find out what are these possibilities and how to adapt (and demonstrate) them to the TV industry.

Progress beyond the state-of-the-art

To illustrate the innovation potential of this project, we further detail the different fields in which Immersia TV will provide progress beyond the state-of-the-art:

1. Omnidirectional video capture

Omnidirectional video capture has a long history[1]. Many rigs, including one or more cameras, either without or with one or multiple planar or curved mirrors, in combination with customized or off-the-shelf, lenses, have been proposed in literature or built for specific projects[2]. Rigs have been proposed both for monoscopic and stereoscopic omnidirectional video[3] [4]. At this time, there is no solution for automatic seamless stereoscopic full omnidirectional capture yet. We can distinguish three basic groups of omnidirectional cameras on the market:

- **Single-lens 360° cameras**. There are several relatively inexpensive 360° degrees camera devices on the market (e.g. Kodak SP360[5], Ricoh Theta[6], Bublcam[7] or Giroptic 360cam[8]) which are able to produce omnidirectional video for consumer and stream it over wireless network interfaces. Some of such small and lightweight omnidirectional cameras are being announced regularly in crowd-funding campaigns, but did not yet hit the stores. The quality of video coming from such compact devices is usually sufficient for mobile applications, but cameras are not able of producing high resolution or high frame rate video for more demanding users and immersive entertainment
- **Custom rigs for multiple cameras**. Custom rigs for inexpensive cameras such as Gopro (360 heros[9] or freedom 360 usual*y) are very popular. They provide high resolution a*nd a good frame rate, in a light-weight, compact, unthettered and affordable package, but they lack inter-camera synchronisation and central exposure and white balancing control. More advanced rigs are usually still custom-build rigs, with high end cinema cameras such as ARRIflex or RED, or based on machine vision cameras, high end to very low cost board level depending on budget, skills and desired image quality.
- **Professional omnidirectional cameras.** Some companies are working on more **advanced omnidirectional camera systems** constructed using multiple high-resolution cameras places on specialized rig. Such ready-made omnidirectional camera heads include the point grey ladybug. Cameras are perfectly synchronised and sensors are consistently controlled, but often frame rate and/or resolutionare not adequate. Examples of such cameras are JauntVR[10], Panono Ball Camera[11], Panocam 3D[12], Samsung 360 Camera[13], Camargus (Qamira)[14].

In this context, ImmersiaTV will deploy its own approach and capture set, taking into account the existing approaches in the market. In other terms: to exploit the current rapid evolution of sensor and camera hardware technology, data interfaces, optics, stitching algorithms in the context of omnidirectional video broadcasting, ImmersiaTV will not focus on building its own rigs, or stitching algorithms, but rather on developing a video capture, replay and processing distributed architecture that is ground up designed for omni-directional video. The architecture consists of a novel kind of "edge" processors, that combine raw camera data storage and replay and low level image processing and camera control found commonly in cameras or recorders, with per-camera sub-tasks in the stitching process, now responsible for the need of a heavy compute cluster linked with the cameras via expensive long-range very high

---

[1] Shree K.Nayar, Terry Boult, Omnidirectional Vision Systems: 1998PIReport
[2] Shree K.Nayar, Catadioptric Omnidirectional Camera
[3] Terry Boult, Remote Reality via Omnidirectional Imaging
[4] T.E.Boult,R.J.Micheals,M.Eckmann,X.Gao,C.Power,S.Sablak, Omnidirectional Video Applications
[5] Kodak SP360, http://kodakpixpro.com/Americas/cameras/activeCam/sp360.php
[6] Ricoh Theta, https://theta360.com/en/
[7] Bublecam, http://www.bublcam.com/
[8] Giroptic 360cam, http://www.giroptic.com/
[9] Gopro 360 heros, http://www.360heros.com/
10 JauntVR, http://www.jauntvr.com/
11 Panono Ball Camera, http://jonaspfeil.de/ballcamera
12 Panocam 3D, http://www.panocam3d.com/camera.html
13 Samsung 360 camera, http://gizmodo.com/samsung-built-its-own-360-degree-camera-for-streaming-v-1657936437
14 Camarcus now owned by Qamira, http://qamira.com/

bandwidth links. The resulting partially processed video data with the ImmersiaTV capture system is ad-hoc coded to allow transport over affordable long-range links to be blended on a central processing unit that is much lighter than in current practice.

Omnidirectional video has been used in the past for TV broadcast purposes on first or second screen (e.g. in the iMinds explorative television ICON project in 2011-2012), but resolution and frame rate was perceived too low, stitching too cumbersome, processing too slow and capture and processing equipment too costly and delicate. ImmersiaTV will result in capture infrastructure design that addresses all these issues enabling omnidirectional video to become mainstream in the broadcast and videoconferencing world[15].

2. Immersive content production tools

The creative editing of omnidirectional video is still a quite challenging undertaking. Off the shelf software to edit offline video, such as Adobe Premiere, or Final Cut X, offer limited support for these. It is possible to edit omnidirectional video through post-production software suites like Adobe AfterEffects, or Nuke, but this process is far from efficient. It is very time-consuming, and unpractical to use. Moreover, traditional editing environments, typically composed of a preview screen and an output screen, are ill-suited to edit omnidirectional video. Editing is therefore often done with "flattened" images, or with portions of the field of view. Moreover, mainstream production tooling has not been designed to easily combine several video streams, both omnidirectional and not omnidirectional, in a coherent experience. As a result, despite it is feasible to produce such content offline, it is unpractical, and it is not possible with the software solutions currently existing to produce such content live.

The technical needs raised by the creative concept underpinning the ImmersiaTV content, where portals need to appear, grow, or move as a result of both the content creator and the end-users behavior, will steer the technical development of production tools. In this context, it is worth highlighting that the ImmersiaTV consortium, will develop its approach from the only commercial tool that currently exists for stitching and editing omnidirectional video live streams, Vahana VR, will made available through the company Video-Stitch. This will also allow customizing concrete aspects of this software solution if the need appears. In this context it is worth highlighting that Video-stitch was started as the natural continuation of creative projects involving omnidirectional video. By that time, the offline editing of omnidirectional video required massive amounts of manual labor for video stitching, and ad-hoc video processing scripts. The limitations found in this creative process, and the technical solutions developed, were the backbone on which Video-stitch was started. Similarly, Cinegy was started to address the need for a software solution capable of managing a large amount of video sources. ImmersiaTV will leverage this background expertise and radically increase the efficiency to create its defining novel content formats.

An additional advantage of the ImmersiaTV consortium is that it also includes access to low-level advanced algorithms in stitching, coding and distribution, and efforts to integrate these within the existing production tools. This will further allow a tighter integration of the end-to-end broadcast pipeline, and increase even further the quality and efficiency of content production. The ImmersiaTV consortium is therefore uniquely placed to deliver an innovative tool allowing to manage several video streams, both omnidirectional and non-omnidirectional, in coordination with standardization bodies and leading academic researchers, as well as testing of these by professional creatives with large experience in the audiovisual industry.

[15] Fernando De la Torre Carlos Vallespi Paul E. Rybski Manuela Veloso Takeo Kanade , Omnidirectional Video Capturing, Multiple People Tracking and Identification for Meeting Monitoring

3. Encoding /Decoding of Omnidirectional video

Encoding of omnidirectional video sequences, which represent time-varying 360° environments, requires sphere-to-plane mapping, at least when using off-the-shelf video compression algorithms originally designed for rectangular video data. Typical mapping schemes that have been explored so far include cylindrical projection[16] [17] [18] [19], Cube-map projection[20], and Rhombic-dodecahedron mapping [21]. However, the efficiency of these mapping schemes has not been studied in depth especially in terms of their influence on coding efficiency. Inappropriate mapping approaches may lead to insufficient sampling of omnidirectional video frames causing aliasing. In-depth analysis and experiments needs to be conducted to evaluate spherical parametrization and their influence when off-the-shelf coding schemes are used.

In addition, relatively large amount of data in omnidirectional video sequences challenges the current video compression approaches in terms of real-time processing, especially real-time encoding. In order to deliver real-time streaming of 360° omnidirectional video or its parts, low complexity and low latency constraints needs to be taken into account and implemented into encoder while maintaining the high video quality level. One way to face the real-time processing constraints is a foveated encoding, where region of interest of user is estimated and a priority saliency map determining the progressive hierarchical coding of the content is computed. Several solutions to estimate user attention exist in the literature mostly for high definition resolution [22] and rarely for ultra-high definition resolution.[23] However, there is no salient model or solution to estimate user attention in the omnidirectional video or its parts.

4. Content Delivery and Reception

Nowadays several studies[24] state that multimedia data is the most significant part of the data flowing over the Internet, with a radical increased in recent years. However, this recent growth might not be the most relevant fact: what is remarkable is that the future estimations on Internet traffic suggest that video content over the internet is expected to keep growing dramatically. In fact, video delivery is now and it still will be a challenge in terms of efficiency, scalability, QoS and QoE. In order to accelerate Internet and provide scalability for mainstream Internet services (media delivery, web apps and HTTP traffic in general) CDNs platforms were born. Main CDNs solution to this problem relies over two different approaches:

- Edge Hosts: Based on adding infrastructure (mainly hosts or servers) at the edge of the internet network, inside ISP premises.
- Edge Networks: Based on adding huge central deployment (big CPDs at the internet backbone) were edge networks (ISP providers) have a direct peering network.

This is the way CDNs tend to deploy content closer to the end user and in consequence reduce network throughput variability and network congestion. These approaches are feasible and well adapted to HTTP based services, however

---

[16] D.-S. Lee, B. Erol, J. Graham, J. J. Hull, and N. Murata, “Portable meeting recorder,” in Proc. 10th ACM Int. Conf. Multimedia , 2002, pp. 493–502

[17] C. Grünheit, A. Smolic, and T. Wiegand, “Efficient representation and interactive streaming of high-resolution panoramic views,” in Proc. Int. Conf. Image Processing, 2002, vol. 3, pp. 209–212

[18] I. Bauermann, M. Mielke, and E. Steinbach, “H.264 based coding of omnidirectional video,” in Proc. Int. Conf. Computer Vision and Graphics, 2004, pp. 22–24

[19] A. Smolic and D. McCutchen, “3DAV exploration of video-based rendering technology in MPEG,” IEEE Trans. Circuits Syst. Video Technol. , vol. 14, no. 3, pp. 348–356, Mar. 2004

[20] K.-T. Ng, S.-C. Chan, and H.-Y. Shum, “Data compression and transmission aspects of panoramic videos,” IEEE Trans. Circuits Syst. Video Technol. , vol. 15, no. 1, pp. 82–95, Jan. 2005

[21] J.-S. Lee and T. Ebrahimi, “Perceptual video compression: A survey," Selected Topics in Signal Processing, IEEE Journal of 6(6), 684{697 (2012).

[22] Z. Wang, L. Lu, and A. C. Bovik, “Foveation scalable video coding with automatic fixation selection," Image Processing, IEEE Transactions on 12(2), 243{254 (2003)

[23] M. Rerabek, H. Nemoto, J.-S. Lee, T. Ebrahimi, “Audiovisual focus of attention and its application to Ultra High Definition video compression”, Proc. SPIE 9014, Human Vision and Electronic Imaging XIX, 901407 (February 25, 2014)

[24] http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html

HTTP was not designed for video services. In live video streaming technologies these two approaches shown emerged recently and are taking over the market:

- HTTP file based streaming: based video segmentation into files and serve them via HTTP.
- Enhanced Streaming over UDP: based on adding reliability to UDP at application level (e.g. scalable coding,Forward Error Correction techniques, retransmissions of lost packets, etc.) and at network level (multipath streaming and/or other advanced SDN features).

However mature and older technologies like RTMP must be considered still; as there is quite a good, functional and stable infrastructure already supporting this technologies (it is a supported technology in many commercial CNDs and delivery services), and they still represent an option to get quickly into the market. But is also relevant to note that majors in video delivery such as Netflix, YouTube or even Adobe (inventors of the above mentioned RTMP) are clearly moving to HTTP open standards like MPEG-DASH at the same time they are starting to provide UHD services.

	HTTP	**ENHANCED UPD**	**RTMP**
SERVER COMPONENTS	Open	Specialized	Specialized
SECURITY AND IP PROTECTION	Available, but lack of unified approach	Specialized	Available
CHACHING	Native	Specialized	Specialized
MULTICAST SUPPORT	None	Supported	Supported
PLUGIN AND PLAYERS PENETRATION	Emerging	Specialized	Extensive
FIREWALL OR PROTOCOL/PORT ISSUES	None	NAT transversal and firewall restrictions	Firewall restrictions
VARIABLE BITRATE	No impact at server side	No impact if scalable coding techniques are used	Susceptible to data spikes

Table 1 Feature comparison between HTTP, UDP and RTMP

UDP streaming, as noted in the table has to cope with two relevant issues and barriers, reliability and scalability. Reliability in UDP usually is faced by the use of FEC technologies and some sort buffering to enable retransmission of lost data which forces both complex and costly (in computation terms) streaming servers and clients. Technically speaking basic scalability issues using UPD might be easily solved by using multicast transmissions so each representation of the stream would only be transmitted once over the Internet, but multicast support in OTT networks is not widely extended and many ISP doesn't support it.

However there several ongoing initiatives like Automatic Multicast Tunnelling draft standards[25] already supported by major network equipment vendors such Cisco[26] and Juniper[27], 5G and C-RAN emerging technologies[28] [29] or SDN like approaches[30] and the omnipresent switch from IPv4 to IPv6; all growing and expanding technologies that enhance multicast support. Multicast streaming over UDP might have a relevant role in future networks as they get gradually updated infrastructure and management techniques, this is the approach considered strategic like Akamai, who recently acquired Octoshape[31], a reference in the novel media delivery systems. In the comparison table, maybe the most relevant fact related to HTTP streaming is that is reliable, runs over wide spread technologies without requiring specialized infrastructure and can easily implement some adaptation features by using smart clients. This are the main reasons behind the industrial interest in the new open standard for HTTP streaming MPEG-DASH, key actors of the

[25] https://tools.ietf.org/html/draft-ietf-mboned-auto-multicast-18
[26] http://www.cisco.com/c/en/us/td/docs/routers/crs/software/crs_r4-3/multicast/configuration/guide/b_mcast_cg43xcrs/b_mcast_cg43xcrs_chapter_011.pdf
[27] http://www.juniper.net/documentation/en_US/junos12.1/topics/example/mcast-amt.html
[28] http://eeweb.poly.edu/~elza/Publications/icassp08.pdf
[29] https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf
[30] http://link.springer.com/article/10.1007%2Fs10922-014-9322-8
[31] http://www.akamai.com/html/about/press/releases/2015/press-040615.html

media delivery (Akamai, Netflix, Adobe and Cisco among many others) are currently member of the Dash Industrial Forum[32] promoting the usage MPEG-DASH. With the expansion of HTTP 2.0 these technologies are expected to get relevant improvements related to QoS because of properly addressing the Head-Of-Line (HOL) issue[33] and improving HTTP caching possibilities.

As it has been already noted ImmersiaTV targets multiple devices, expected to provide a simultaneous coherent experience (synchronization and adaptations of flows) to the users. For in-house live streaming (from the media centre to the display devices) different techniques will be analysed, this is a pretty different scenario, as in will be a controlled LAN that includes Ethernet and WiFi as the main network technologies to consider. Streaming only small portions of the global sphere to tiny displays must be considered while keeping interactivity and the capability to move around over whole sphere. For such a purpose very low latency streaming is needed, as the media centre must react accordingly to the user movements in real-time. Extensions to the current real time streaming strategies might be considered here the RTSP extension proposed by the Osaka University and Osaka Institute of Technology[34]. Other options might be focused on ROI estimation so only the ROI is transmitted to the display, or the whole stream is transmitted but just the ROI gets the full resolution and quality. The approach will be in strict relation with the novel coding and decoding techniques adopted by the ImmersiaTV.

5. Visualization and interaction

Historically, Immersive Displays have been used mostly to render virtual reality involving computational graphics, generated with methods similar to 3D games or 3D animation movies (for example [35]). The creation of Omnidirectional Video has been a technical possibility for long, but not its live delivery. It is only now that the confluence of affordable hardware for immersive display, live delivery of omnidirectional streams and improvements in omnidirectional video hardware enables the delivery of video-based omnidirectional content. However, content formats are not ready for this reality.

The central purpose of ImmersiaTV is to demonstrate a novel kind of audiovisual content format, compatible with immersive displays, coherent across devices, and ready for broadcast distribution. This requires progress on several aspects. Omnidirectional video distribution has already been demonstrated[36] [37] .However, these attempts do not address the problem of synchronization among devices in a local network. Despite there exist approaches to guarantee synchronization of multimedia streams between devices, [38] [39], existing approaches do not take into account the specifics of immersive displays, where the highest priority should always be to keep a very fast sensorimotor loop between head movements and rendering of the immersive display. These constraints require an update of the approach where "the slower device wins", where the less powerful device is taken as the temporal reference to ensure synchronization. In other terms: the extra need for temporal buffering and the larger field of view needed by immersive displays might impose needs to consider such devices "slower" than less powerful devices, just to be able to guarantee the extra buffering needed to address the demanding requirements of immersive displays. In addition,

[32] http://dashif.org/members/

[33] http://www.streamingmedia.com/Articles/Editorial/Featured-Articles/HTTP-2.0-and-DASH-Planning-Tomorrows-Improved-Video-Delivery-98653.aspx

[34] http://onlinelibrary.wiley.com/doi/10.1002/ecj.10375/pdf

[35] Spanlang, B., Normand, JM., Borland, D., Kilteni, K., Giannopoulos, E, Pomés, A, González-Franco, M., Perez-Marcos, D., Arroyo-Palacios, J., Navarro Muncunill, X., Slater, M. How to build an embodiment lab: achieving body representation illusions in virtual reality Front. Robot. AI, 27 November 2014

[36] https://littlstar.com/

[37] Also, for medical purposes http://lnco.epfl.ch/realism

[38] Bang, Y., Han, J. and Lee, K. Yoon, J. Joung, J, Yang, S. Rhee, J-K K., (2009) Wireless Network Synchronization for Multichannel Multimedia Services}, Proceedings of the 11th International Conference on Advanced Communication Technology - Volume 2, ICACT'09 1073-1077

[39] Rautiainen, M., Aska, H., Ojala, T., Hosio, M., Makivirta, A., & Haatainen, N. (2009, June). Swarm synchronization for multi-recipient multimedia streaming. In Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on 786-789

current synchronization methods [40] [41] do not address the problem of device discovery: a novel device, not existing at the time of the implementation of the software, might impose different parameters on the overall synchronization mechanism. We will therefore take inspiration of solutions proposed for musical devices in live environments[42] to address synchronization in a way that enables device discovery.

In addition to synchronization, the display also needs to blend seamlessly with interaction with the available video streams. Despite there have also been some exploratory studies on interaction with omnidirectional video[43] [44], to the best of our knowledge there has not been an explicit attempt to:

- Develop an audiovisual language based on combining several video streams, both immersive and non-immersive, which enables creating narratives adapted to immersive displays.
- Adapt the rich audiovisual language of traditional TV to this novel kind of display, where the use of travellings, cuts, and complex camera movements would provoke sickness if rendered in an omnidirectional field-of-view.
- Explore the possibilities of interaction within such immersive display narratives
- Deliver an experience coherent across devices.

Overall, by developing a software solution that addresses, on one hand, the adaptation of audiovisual language and interaction and, on the other side, synchronization among devices in a distributed way, ImmersiaTV will deliver a unique update of the broadcast model, particularly adapted to the reality of the contemporary living room, where immersive displays arrive to join traditional TV and second displays.

[40] Mahmood, A., & Gaderer, G. (2009, October). Timestamping for IEEE 1588 based clock synchronization in wireless LAN. In Precision Clock Synchronization for Measurement, International Symposium on Control and Communication, 2009 1-6

[41] Mughal, M.A., Juhlin, O (2014) Context-dependent software solutions to handle video synchronization and delay in collaborative live mobile video production. Personal and Ubiquitous Computing 18 (3) 709-721

[42] Bergstrom, I., & Llobera, J. OSC-Namespace and OSC-State: Schemata for Describing the Namespace and State of OSC-Enabled Systems. Nime Conference 2014

[43] Benko, H., and Wilson, A. D. Multi-point interactions with immersive omnidirectional visualizations in a dome. In Proc. of the ACM Int. Conference on Interactive Tabletops and Surfaces, ITS (2010), 19–28

[44] Rovelo Ruiz, G.A. Vanacken, D., Luyten,K. Abad,F. Camahort, E. Multi-viewer gesture-based interaction for omni-directional video Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 4077-4086