THE WALL OF MOMENTS: AN IMMERSIVE EVENT EXPERIENCE AT HOME

R. Bauwens, T. Deboosere, M. Matton, L. Overmeire

Vlaamse Radio- en Televisieomroeporganisatie, Belgium

ABSTRACT

In order to create an immersive experience for people who cannot attend an event, we have created The Wall of Moments prototype. User-generated content is a key element: captured right in the action makes it often very unique. To increase the authenticity of the experience, this user-generated content is synchronised with professional content (in time as well as place) and presented in a vibrant way to the end user. The application has been filled with content and evaluated during a test shoot in the Marconi studio at the VRT campus. This paper describes the developed prototype in detail, discusses the lessons learned and provides some gained insights regarding immersive event experiences.

INTRODUCTION

Within the ICoSOLE project, different (cost-effective) ways of capturing spatially outspread events are being considered. A core research challenge is to synchronise all the content in time and space to create an immersive experience for people who cannot attend an event.

ICOSOLE

Main goals

ICoSOLE[1] is a collaborative project to research and develop tools and services for covering spatially spread out live events using hybrid broadcast-Internet technologies. The approach uses a wide variety of sensors, ranging from mobile consumer devices over professional broadcast capture equipment to panoramic and/or free-viewpoint video and spatial audio, to optimise an immersive experience for the consumer. The project has developed methods for providing incentives to a live audience to contribute high-quality content as well.

1. In order to combine a number of different capture sources, ICoSOLE has investigated professional and consumer capture devices, including mobile (and moving) sensors, based on metadata and content analysis. Methods for fusing visual and audio information into a coherent data representation have been developed, which enable video and audio for virtual viewer/listener positions to be rendered.
2. ICoSOLE has built a prototype networked platform for streaming live high-quality audiovisual content from mobile capture devices to content acquisition, processing and editing services, and to distribute it to viewers via broadcast (including second screen), web and to mobile devices. This work includes the development of efficient

tools for media production professionals to select, configure and review the content sources being used.

3. To ensure this end-to-end chain, ICoSOLE has created methods to capture, extract and annotate metadata during the production process and to integrate this metadata throughout the entire production chain to the end user.

Consortium

Developing a novel tool chain for the immersive coverage of live events, supporting a wide range of devices for both content capture and consumption, requires a consortium that includes the scientific- technical to deploy beyond state of the art technologies, has knowledge of the requirements of the different types of users involved, and access to both professional users of the technologies and consumers. A complete list of the consortium members can be found in the appendix.

DEVELOPMENT OF THE WALL OF MOMENTS

Role within the ICoSOLE project

The Wall of Moments prototype fulfils the playout side of the ICoSOLE project. Together with related prototypes for capturing and filtering, these prototypes cover the whole project flow, which is shown in the Figure 1 below.

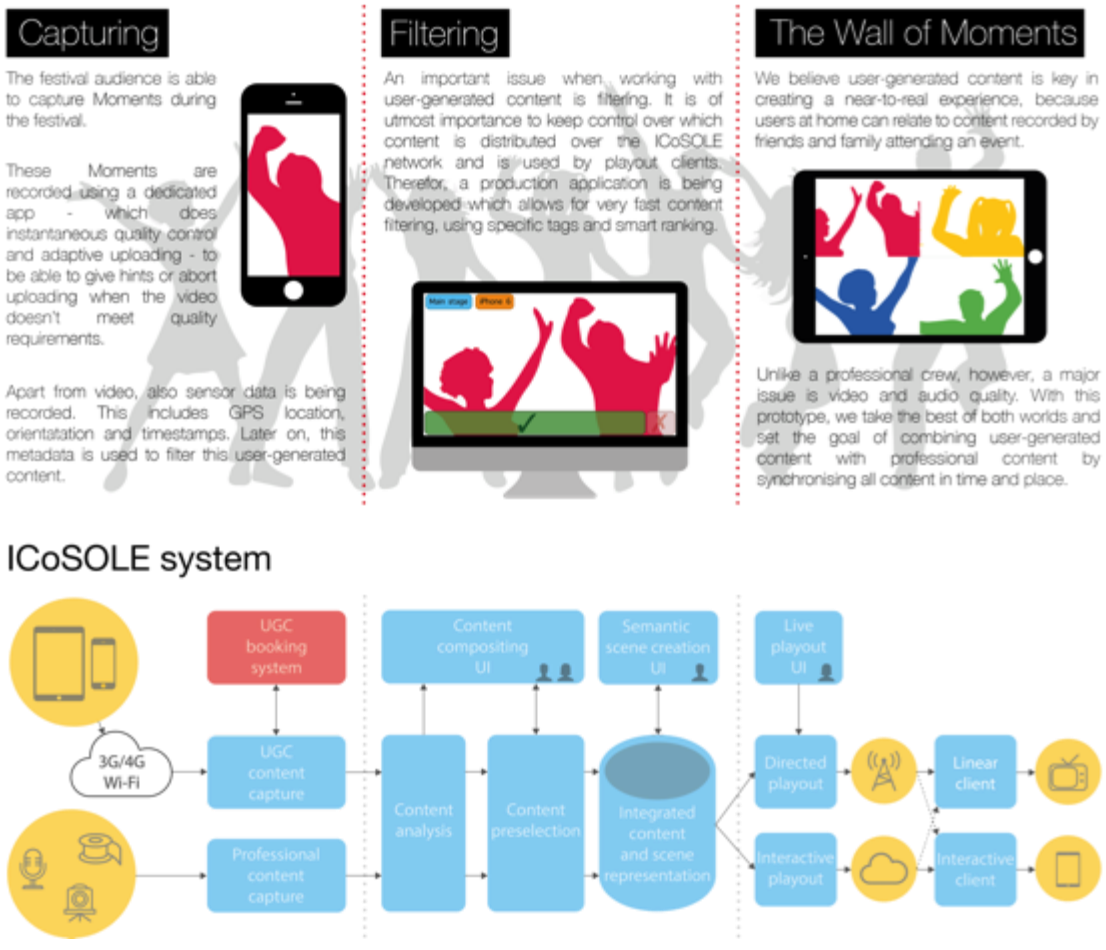

Figure 1 – ICoSOLE overview

How it came to be

A lot of content is captured in the ICoSOLE project. Firstly, novel camera setups such as omnidirectional cameras are used. This can benefit the users at home by allowing them more control over the event scene, as well as the production team, giving them an overall overview over the scene. Secondly, the audio scene is captured using a wide range of microphones, e.g. audience microphones, individual microphones on every band member in a group and an Eigenmike [2]. This way, users take more control, mixing their own audio scene or experiencing binaural audio.

A third content source is footage captured by the audience. The Wall of Moments uses this user-generated content to broaden and elevate the event experience for people at home.

A potential issue with user-generated content, however, is its inferior quality (especially the audio) compared with professional content. This is not a big problem when e.g. shared on social media, but it is a different case when used in a professional production. By talking to directors from VRT who attempted to mix user-generated content with professional content in a recording, it was identified that this was a very hard thing to do; this is because they are very different kinds of content. The visual transitions between the two genres can be extremely uncomfortable for an audience to watch.

On the other hand, user-generated content often offers the most exciting and timely content, giving a first-row experience from within the crowd. People at home have more connection with content from their peers because it is personal. This makes user-generated content absolutely essential for re-creating a real-life experience.

As mentioned before, a lot of high-quality equipment such as 4K and omnidirectional cameras has been used in the project. Our goal was to find the right intersection between the professional and user-generated content to enhance the user experience at home.

The project started with the premise that we wanted user-generated content to capture ambience, not (video/audio) quality. When we want to deliver high quality material, we switch to a stream that can guarantee this (e.g. a professional production or omnidirectional stream). We also aimed to build an application that would connect users with each other and with peers at the event, without inventing yet another social network; instead, the project made use of existing social media. Lastly, we required the audience at the festival to contribute some added value. They, after all, have to capture the videos for our prototype *Wall*.

With our first ideas as a starting point, we created a mockup (Figure 2) –using YouTube content– and invited colleagues from a wide range of services at VRT to test it out. We gathered a lot of input and used it to create the first prototype *Wall of Moments*.

Figure 2 – Mockup for ‘The Wall of Moments’

THE WALL OF MOMENTS

The different aspects of the prototype are outlined below.

Moments

As the term ‘user-generated content’ is too general for our interpretation, we decided to name the clips made by users differently. We chose the term ‘Moment’, because it explains the emotionally charged volatility we look for in this content in a simple, elegant way. These *Moments* tell a story, show something unique, and catch something timely that a professional camera would probably miss.

The Wall

The *Wall* (Figure 3) arranges the latest and most interesting *Moments* in a mosaic and is highly personalized: the social network of the users is taken into account to decide what is shown. A user can zoom into a *Moment* on *the Wall* by clicking or tapping. The *Moment* is then played full screen and some controls are shown (Figure 4).

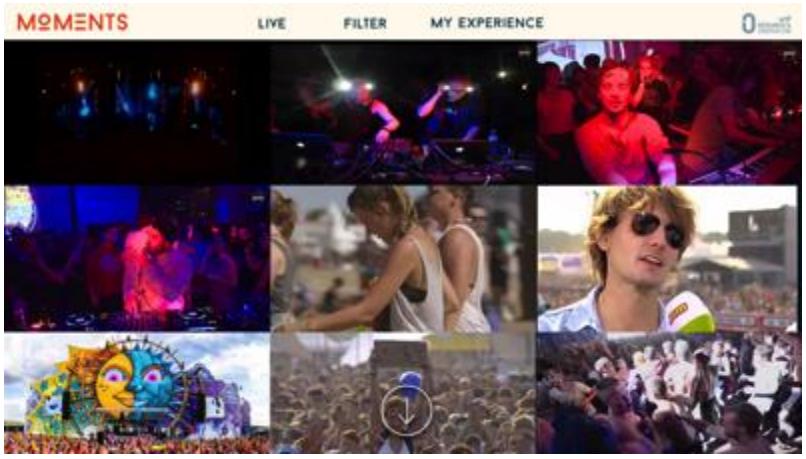

Figure 3 – The Wall

Synchronisation

We believe user-generated content is key in creating a near-to-real experience, because users at home can relate to content recorded by friends and family attending the event. Unlike a professional crew, however, a major issue with user-generated content is video and audio quality. Notwithstanding, a professional crew is more prone to missing the most emotional and exciting moments in the crowd due to accessibility issues (among other things). To circumvent this problem, we take the best of both worlds and set the goal of combining user generated content with professional content.

Figure 4 – Synchronisation

All the content that has been captured by a device that is part of the ICoSOLE system, is annotated in several ways (e.g. timecode, location, orientation). This way, content can be mapped precisely in time and place of the event in the ICoSOLE core.

This makes it possible to play the professional content along with the opened Moment (picture-in-picture, when available), increasing the authenticity of the experience. When a user wants to start watching the event from there on, he can switch to the professional content to get the audience's view.

At the top, a timeline is shown. Every other interesting *Moment* (for this user) is annotated, and s/he can jump to a particular one with just a click or a tap. At some points, clusters of dots are formed, possibly indicating something interesting happened at that point in time. This encourages a more social navigation through the concert, instead of a linear song-by-song navigation.

My Experience

When a user opens a *Moment*, s/he can add it to *My Experience* (Figure 5). Essentially, this is a personal favourites list.

To make this a really interesting feature, a personal aftermovie –based on this list– is created by adding an intro, outro and background music. This feature will be enhanced along the way: users will be able to mix in professional content, add titles and more.

Figure 5 – My Experience

Live

Within the application, a section is reserved for watching a live performance. For now, it only shows the current live performance. In later stages, a timeline will be added to this stream, which will show *Moments* popping in near-to-live as they are picked up by the ICoSOLE backend.

Technologies

We are developing *the Wall of Moments* as a web application to support as many (mobile) devices as possible, using the Angular.JS application framework [3] (based on and extended with HTML5 [4], JavaScript and CSS3). Communication with the backend is handled through a JSON [5] REST interface, which is developed by consortium partner JRS. For live and professional video playback, a DASH-enabled video player by consortium partner Bitmovin [6] has been integrated.

Synchronisation is key for this prototype. JRS is developing tools for using captured metadata to precisely determine the time and place of user-generated content. To achieve this, GPS location and timestamps are used, as well as audio analysis of the professional

and user-generated content. All content and metadata is subsequently added to the ICoSOLE core system, currently being developed by Technicolor.

More information on the technologies is available on the ICoSOLE website [1].

TEST SHOOT - MARCONI MOMENTS

Test setup

A first small-scale test shoot –Marconi Moments– was organized in the Marconi studio at the VRT campus last October. Professional content was captured using a wide variety of devices:

- Four TV cams, of which a live director's cut was made
- Four omnidirectional cameras: one on stage, and three in the venue on the left, right and back
- Two GoPro cameras, placed above the two screens to measure people's interest
- Six microphones places in the audience

Moreover, everyone in the audience was able to record short clips with their own smartphone and send them in to a backend we provided. After uploading was completed, they were placed in a queue and shown on a screen next to the stage a short time thereafter, arranged in a two-by-two matrix (Figure 6). This has proven to be a popular feature: if we only account for user-contributed videos, we are able to reproduce 89% of the second concert. The collected content now serves as a data set to test and refine our algorithms before going (near-to) live.

Figure 6 – Test setup at the Marconi Moments

Survey results

Following the test shoot, we organized a small survey for the audience. Some interesting insights:

- 73% used the capture app to upload Moments
- 51% favoured the screens next to the stage
- 66% would consider using such an application on a real festival

FIELD TRIAL - DRANOUTER FESTIVAL

The next test of the Wall of Moments will be the near-to-live implementation at the Dranouter Festival (August 7th – 9th); and we will show results from this in any presentation of this paper. To make this system work here, some improvements are currently ongoing:

- We are integrating the capture tool developed by other consortium partners (JRS and Bitmovin) with *the Wall of Moments*. As this is a native application, we will have more control over which content is captured and when (e.g. we can restrict it to only live streaming, GPS data capture etc., as required).
- From the results and experiences gained so far, the Wall of Moments is being rewritten from the ground up. In the prototype we used last year, all concurrent video streams were played simultaneously. This forces mobile devices to switch from hardware to software decoding for video, which is far less energy inefficient, consumes more processing power and thus impacts battery life greatly. Moreover, software decoding is not allowed on iOS devices, which makes the prototype unable to run on the platform. In the new version, we are using keyframes to render the Wall, which solves the incompatibility issue and is more energy efficient. A drawback of this approach is the increased network throughput, but lowering the frame rate and resolution can reduce this issue.
- We are also deploying a new prototype to filter user-generated content. Because our next demo is taking place at a real festival, we are leaving the closed environment and have to guarantee that no inappropriate content is making it to a large screen at the festival or to the people watching at home. The generated content will be fed (near-to-live) into the ICoSOLE core system and will subsequently be automatically tagged, using the gathered metadata. We believe a final human verification is crucial and rely on automatic analysis only to rank user-generated content (by pushing valid content upwards in the content chain) and to make suggestions. We are using it as a tool, rather than a surrogate. In future versions, we plan to use distributed filtering, which allows a group of people to collaboratively accept or decline user-generated content.

CONCLUSION AND FURTHER RESEARCH

With the first end-user focused prototype in the ICoSOLE project, the project consortium made great progress in elevating the festival experience for users at home and underway. As the project has developed, we have intensified this experience even more by improving the operation of the prototype system, and implementing new techniques. Furthermore, by organising frequent field tests, we have gained a better understanding of problems we will run into when attending larger events: copyright issues and filtering of inappropriate content are among the identified challenges.

We showcased *the Wall of Moments* prototype on our own media conference, Media Fast Forward 2014, where reactions were wildly positive. We also demonstrated the *Wall of Moments* at a consortium partner's conference, BBC Sound Now & Next 2015, which is a two-day event on innovation in sound production and broadcasting. During these events, we made interesting contacts regarding content sychronisation and network transmission technologies, which in time could lead to further collaborative efforts. We have been

approached by several event organisers as well, who are keen to integrate prototypes for their own festivals. Therefore, we are considering business models to make this prototype available as a service for these event organisers in a later stage of the project.

This creates opportunities for the field trials planned for 2016, because we are thinking about applying the ICoSOLE technology to other kinds of events as well, e.g. sports events. Field trials in this sector will provide further insights into improvements of applications for end users in terms of usability, design and features. By putting our prototypes to the test from very early on, we are able to adapt them very quickly and efficiently.

Following the field trial at Dranouter in August of 2015, we will make a thorough evaluation of the entire ICoSOLE system, to resolve issues and to give direction for subsequent field trials. During presentation of this paper, we plan to share the results and insights, and our plans for the future, which will hopefully lead to even more vibrant and lively event experiences in 2016 and beyond.

ACKNOWLEDGEMENTS

This work is part of the ICoSOLE project, funded by the European Commission through the 7th Framework Programme (FP7/2007-2013) under grant agreement no. 610370.

REFERENCES

1. EC Project FP7/2007-13, Immersive Coverage of Spatially Outspread Live Events project website.

   http://icosole.eu

2. Eigenmike microphone array.

   http://www.mhacoustics.com/products

3. AngularJS open-source web application framework.

   https://angularjs.org

4. HTML5 markup language.

   http://www.w3.org/TR/html5

5. JSON (JavaScript Object Notation) language.

   http://json.org

6. Bitdash suite of MPEG-DASH clients.

   http://www.bitmovin.net/bitdash- mpeg- dash- player.

APPENDIX

The consortium consists of following partners:

1. **Joanneum Research (JRS)** is a non-profit organisation concentrating on applied research with a highly qualified staff of more than 400 people. Services include specifically geared research tasks for small- and medium-sized companies, complex interdisciplinary national and international assignments as well as tailored techno-economic consulting.
2. **Technicolor (DTO)**, with its approximately 20,000 employees worldwide, is one of the world's leading technology enterprises in the media industry. Technicolor is focused on its content creator customer base, leveraging its outstanding market position, the strength of the Technicolor brand with film and television studios, its in-home delivery assets (set-top boxes and gateways) and its research and licensing capabilities. The main business areas comprise solutions for digital content/media as well as for audio/video networks in the professional and consumer domains.
3. **De Vlaamse Radio- en Televisieomroeporganisatie (VRT)** is the public broadcaster of the Flemish Community in Belgium. Its mission is to reach the largest possible number of viewers and listeners with a diversity of programs that arouse and fulfill their interests. To accomplish this, VRT owns 3 TV channels (Eén, Ketnet, Canvas), 5 radio stations (Radio1, Radio2, Klara, Studio Brussel and MNM) and 3 web portals (deredactie.be, sporza.be and cobra.be).
4. **iMinds** was founded in 2004 by the Flemish government as an independent research institute to stimulate ICT innovation. It is an internationally recognized multidisciplinary research centre (looking at technological, legal, business and sociological aspects) and enables accelerated development and exploitation of new ICT products and services in strategic sectors in Flanders (Belgium).
5. **Bitmovin (BIT)** is an Austrian technology providing company (SME) founded in 2012 as a spin-off of the Alpen-Adria-Universität Klagenfurt and the FP7-ICT project ALICANTE. Its principal expertise is in the domain of information and communication technologies for audiovisual networks, in particular for HTTP streaming services and embedded multimedia applications. Additionally, Bitmovin provides consulting and standardization services, specifically towards ISO/IEC MPEG.
6. **The British Broadcasting Corporation (BBC)** is a large broadcasting organisation producing radio and television programmes in a number of studio centres of varying size throughout the United Kingdom. In addition, it operates one of the most popular web sites in Europe. BBC R&D is a world-leading centre for media production and broadcasting technology, providing the BBC with a competitive advantage through technology development and expertise.
7. **Tools at Work (TaW)** offers IT-solutions for small companies up to corporate clients as well as educational customers from K-12 to higher education. The main focus of business is the creative market. Tools at Work provides solutions and services for agencies, graphic designers, media companies, publishers, video and audio studios and professional broadcasters.


2015