Production and Delivery of Video for Multi-device Synchronized Playout

Juan A. Núñez
i2cat foundation
Barcelona, Spain
juan.antonio.nunez@i2cat.net

Szymon Malewski
PSNC
Poznan, Poland
szymonm@man.poznan.pl

Sergi Fernández
i2cat foundation
Barcelona, Spain
sergi.fernandez@i2cat.net

Joan Llobera
i2cat foundation
Barcelona, Spain
joan.llobera@i2cat.net

ABSTRACT
In the contemporary living room, the audience’s attention is often divided between TVs, second screens and, increasingly, head mounted displays. To address this reality, ImmersiaTV is a H2020 European project which is redefining the end-to-end broadcast chain: production, distribution and delivery. It is built on two ideas: multi-platform synchronous content playout, and orchestrated videos rendered in the head-mounted display as interactive inserts, which allow introducing basic interactive storytelling techniques (scene selection, forking paths, etc.) as well as classical audio-visual language that is not possible to render with 360 videos (close-ups, slow motion, shot-countershot, etc). We demonstrate our pipeline for offline production, distribution and synchronized playout.

Author Keywords
Inter-device synchronization, content production, innovative content formats

ACM Classification Keywords
H.5.1. Multimedia Information Systems: Video; Artificial, augmented and virtual reality.

INTRODUCTION
The majority of TV consumers now watch TV programs in a multi-display environment [2]. Second screens – most often smartphones - are generally used to check information not directly related to the events in the TV content being watched. Broadcasters have tried to orchestrate these different platforms, and there is reason to believe this contributes to user engagement [3]. However, their success has been limited. This might be caused, at least in part, by the different formats of content being used: mobile apps show graphics and text similar to web content, while TV renders a continuous stream of audiovisual content. The arrival of virtual reality displays to the living room further increases the need for consistent experiences across displays. To address this fact, we have designed and implemented an end-to-end production, delivery and rendering pipeline for offline content production specifically designed to create experiences which give freedom to the end-user to engage with content in one or another device, and are still adapted to the specificities of each device, both in terms of content format and interactive input (or lack of interaction, for the case of the TV). We demonstrate such an approach with an example documentary showing the day life of a kid attending the Porto Football School.

CAPTURE
The creation of content that is both omnidirectional and traditional requires shooting simultaneously in both content formats. The solution found was to use two BlackMagic Micro Studio Camera 4k micro-cameras for the traditional shooting, which could be hidden or, if visible, removed in post-production with a reasonably small amount of effort.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

TVX’17 Adjunct, June 14-16, 2017, Hilversum, Netherlands
© 2017 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-5023-5/17/06.
http://dx.doi.org/10.1145/3084289.3089908

Figure 1: Top: a camera setup to record traditional and omnidirectional video simultaneously. Bottom: a schematic diagram of possible directive inserts located within the omnidirectional video. Image courtesy of Lightbox (www.lightbox.pt).
EDITION
We are not aware of an editing tool targeting synchronous rendering across devices. To address this fact, we have designed and implemented a plugin for Adobe Premiere Pro. The ImmersiaTV plugin shown in figure 3 allows defining the inserts that are placed within an omnidirectional video. It also allows selecting which tracks should be rendered in each of 3 possible devices (TV, tablet or HMD). It works both with Mac and Windows, and we have validated that, after going through a tutorial, can use it to create multi-platform content.

DELIVERY
The media encoding uses H.264 and AAC encoding, and adaptive bitrate streaming based on MPEG-DASH (ISO/IEC 23009-1:2014). Encoding is implemented as a cloud service, running on a Linux server using the Dockers virtualization tool as well as MP4Box from Gpac’s MP4Box for MPEG-DASH multiresolution encoding. Video decoding uses the Gstreamer multimedia framework. The additional metadata required for playout, which relates audiovisual streams with devices (i.e., allows selecting different streams for TVs and tablets), as well as to define interaction and media orchestration requirements, follows closely the format of MPEG-DASH manifests, and its XML specification is publicly available. Content publication is performed through a transcoding and content publishing server controlled with a custom built web interface. It allows triggering media conversion, as well as monitoring progress on media encoding and publishing a list of content, which is then parsed by the player to access the media assets.

INTER DEVICE PLAYOUT
To play synchronized content [1,4,5], we have adapted the DVB-CSS protocol as defined in:
http://www.etsi.org/deliver/etsi_ts/103200_103299/103286/01/01.01_60/ts_10328601v010101p.pdf
and Gstreamer’s version of the Precision Time Protocol (IEEE 1588), as described in:
https://gstreamer.freedesktop.org/data/doc/gstreamer/head/gstreamer-libs/html/GstPtpClock.html
We have also embraced the use of omnidirectional video for HMDs and smartphones, in order to allow the user to visualize the scene in different directions, thanks to a tool integrating Unity3D and GStreamer, publicly available at:
https://www.assetstore.unity3d.com/en/#!/content/59897
https://github.com/ua-i2cat/gst-unity-bridge

REFERENCES
1. Mario Montagud, Fernando Boronat, Hans Stokking, and Ray van Brandenburg. 2012. Inter-destination multimedia synchronization: schemes, use cases and standardization. Multimedia Systems 18, 6: 459–482.
2. Timothy Neate, Matt Jones, and Michael Evans. 2016. Interdevice Media: Choreographing Content to Maximize Viewer Engagement. Computer 49, 12: 42–49.
3. Rebecca Coates Nee and David M Dozier. 2015. Second screen effects: linking multiscreen media use to television engagement and incidental learning. Convergence: 1354856515592510
4. Benjamin Rainer and Christian Timmerer. 2014. Self-organized inter-destination multimedia synchronization for adaptive media streaming. In Proceedings of the 22nd ACM international conference on Multimedia, 327–336.
5. Vinoba Vinayagamoorthy, Rajiv Ramdhany, and Matt Hammond. 2016. Enabling Frame-Accurate Synchronised Companion Screen Experiences. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, 83–92.

Figure 2: The Adobe Premiere Pro ImmersiaTV panel, shown at the centre, allows defining omnidirectional and directive (i.e., traditional) tracks, as well as targeted tracks. The inserts added to the omnidirectional view, shown at right, can be edited with the ImmersiaTV Portal Effect, whose widgets are shown at the left. Image courtesy of Lightbox (www.lightbox.pt).