MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
http://www.comsoc.org/~mmc

MMTC Communications - Frontiers

Vol. 12, No. 1, January 2017

CONTENTS

Message from MMTC Chair

Dear MMTC colleagues:

The GC 2017 Symposium on CSSMA is one of the main venu for colleagues to disseminate recent research results in area of communication software, services and multimedia related topics. With the rapid advances in 5G wireless communication technology, innovations in communication software architecture like virtualization, and new immersive and 3D media applications like VR and AR, there are many exciting breakthroughs in technology and applications. CSSMA symposium is seeking for submissions in the following areas:

+ Multimedia Applications and Services
    • Multimedia delivery and streaming over wired and wireless networks
    • Cross-layer optimization for multimedia service support
    • Multicast, broadcast and IPTV
    • Multimedia computing systems and human-machine interaction
    • Interactive media and immersive environments
    • Virtual Reality and Augmented Reality Communication Systems
    • Multimedia content analysis and search
    • Multimedia databases and digital libraries
    •Converged application/communication servers and services
    • Multimedia security and privacy
    • Multimedia analysis and social media
+ Network and Service Management and Provisioning
    • Multimedia QoS provisioning
    • Multimedia streaming over mobile social networks and service overlay networks
    • Service creation, delivery, management
    • Robust content identification and hashing, content de-duplication and delivery acceleration
    • Virtual home environment and network management
    • Charging, pricing, business models
    • Security and privacy in network and service management •
    • Cooperative networking for streaming media content Next Generation Services and Service Platforms
    • Location-based services
    • Social networking communication services
    • Mobile services and service platforms
    • Home network service platforms
    • VoP2P and P2P-SIP services
+ Software and Protocol Technologies for Advanced Service Support
    • Ubiquitous computing services and applications
    • Networked autonomous systems
    • Communications software in vehicular communications
    • Web services and distributed software technology
    • Software for distributed systems and applications, including smart grid and cloud computing services
    • Peer-to-Peer technologies for communication services
    • Context awareness and personalization

Zhu Li, Steering Committee Chair of MMTC
University of Missouri, Kansas City,USA, lizhu@umkc.edu

Adlen Ksentini, IRISA, France, adlen.ksentini@irisa.fr

EMERGING TOPICS: SPECIAL ISSUE ON IMMERSIVE MEDIA and AR Technologies

Guest Editors: Tasos Dagiuklas
London South Bank University

tdagiuklas@lsbu.ac.uk

Recent developments of head-mounted hardware and virtual immersion apparatuses enable access to a continuous Virtual Reality (VR) and Augmented Reality (AR) experience, stimulating a wide range of senses to imitate the sensorial feeling of the virtual part of the experience, while also providing the means for a more versatile information integration and access. The concept of Pervasive Augmented Reality defines a continuous and universal augmented interface to information of the physical worlds. Several works have been conducted on modelling behaviour of AR/VR objects and on the idea of semantic description of 3D worlds.

This special issue of E-Letter focuses on the recent progresses of immersive media and augmented reality technologies. It is the great honor of the editorial team to have four contributions, from both academia and industry, to report their innovations for developing novel methodologies and solutions in addressing the challenges.

In the first article entitled, "Creating and broadcasting video-based multi-platform experiences ", Llobera *et. Al* have presented an end-to-end production, delivery and rendering pipeline for offline content production which specifically addresses immersive experience. The authors present the solution that have been designed and developed in the context of H2020 European Project Immersia TV. The key challenges presented in the paper, are the following: Delivery of content that is consistent across head-mounted displays giving freedom to the end-user to engage with content in one or another device, content adaptation to the specificities of each device, both in terms of content format and interactive input and integration of video with and further augment traditional TV and second screen consumer habits.

Apostolakis *et al* have presented a tele-immesrive game experience in the second article, entitled *"*Beyond Online Multiplayer: Sharing and augmenting Tele-Immersive (TI) 3D game experiences with multiple geographically distributed users*"*. As digital games often push the technological limits to advance the maturity of technologies for other sectors to closely follow, this paper presents an online multiplayer games that will contribute to further advance research towards efficient and robust user 3D reconstruction, networking, data compression and augmented reality rendering techniques. The paper examines ways of enhancing the current TI infrastructure to support more fully immersed participants' reconstructions, addressing the inherent networking and data compression challenges. Furthermore, it is foreseen to launch thorough Quality of Experience (QoE) and Quality of Service (QoS) studies to identify elements needing improvement and scientifically ground the benefits of Tele-Immersive and AR bring to multiplayer online games.

In the third article entitled, "Of Standards and Herrings: Tales of Technology and Tumult ", Polys has presented high-level sketch of the actors, motivations, and dynamics in the 3D content industry, the WWW, and standards organizations. The paper makes a survey on Web and interactive 3D technology and offer reflections on the dynamics of communities, standards, and industry.

In the fourth article entitled, "Overview of Standardization Efforts for Augmented Reality", Kosmides *et. Al* have presented an overview of the standards in augmented reality. The authors have presented the main benefits of developing standards for AR applications, as well as the key relevant activities that are carried out by international organizations for composing standards in the field of AR.

Tasos Dagiuklas is a leading researcher and expert in the fields of Internet and multimedia technologies for smart cities, ambient assisted living, healthcare and smart agriculture. He is the leader of the newly established SuITE research group at the London South Bank University where he also acts as the Head of Division in Computer Science. Tasos Dagiuklas received the Engineering Degree from the University of Patras-Greece in 1989, the M.Sc. from the University of Manchester-UK in 1991 and the Ph.D. from the University of Essex-UK in

1995, all in Electrical Engineering. He has been a principle investigator, co-investigator, project and technical manager, coordinator and focal person of more than 20 internationally R&D and Capacity training projects with total funding of approximately £5.0m from different international organizations. His research interests include Smart Internet Technologies, Media Optimization across heterogeneous networks, QoE, Virtual Reality, Augmented Reality and cloud infrastructures and services. His research has been received more than 1250 citations by researchers (google scholar). He has served as Vice-Chair for IEEE MMTC QoE WG and Key Member for IEEE MMTC MSIG and 3DRPC WGs. He is serving as Associate Technical Editor for IEEE Communications Magazine and is Member of the Editorial Boards on EURASIP on Image and Video processing and on ELSEVIER Signal Processing: Image Communications journals. He has served as Guest Editor in many scientific journals. He is a reviewer for journals such as IEEE Transactions on Multimedia, IEEE Communication Letters and IEEE Journal on Selected Areas in Communications. He has served as technical committee member of numerous international conferences and workshops including ICC, GLOBECOM, ICME and ICIP. Dr Dagiuklas is a Senior Member of IEEE and Technical Chamber of Greece.

Creating and broadcasting video-based multi-platform experiences

Joan Llobera, Juan A. Núñez° and Sergi Fernández Langa¥*

** i2cat Foundation, joan.llobera@i2cat.net*

° i2cat Foundation, juan.antonio.nunez @i2cat.net

¥ i2cat Foundation, sergi.fernandez@i2cat.net

1. Introduction

The majority of TV consumers now watch TV programs in a multi-display environment [1]. 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 [2]. 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.

ImmersiaTV[1] is an European H2020 Research and Innovation action that is redesigning the production and delivery process to offer a new form of broadcast multi-platform video. This novel form of multi-platform video is based on the following assumptions:

1. The arrival of head-mounted displays (HMD) to the living room further requires the delivery of content that is consistent across displays, giving freedom to the end-user to engage with content in one or another device. This means that content must be 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), but the end-user should be free to adopt one or another device at different moments of the program being broadcasted.
2. This new form of broadcast multi-platform video must seamlessly integrate with and further augment traditional TV and second screen consumer habits. For TV, this means that content can be consumed simply by sitting on the couch and watching, without further input, and that the audiovisual language used follows established conventions. For tablets and smartphones, it means that user input and social media integration must work seamlessly with the specificities of each device.

To address these requirements, in the first year of the project we have designed and implemented an end-to-end production, delivery and rendering pipeline for offline content production which specifically addresses these needs. In this short article we outline the design principles, production and delivery methods as well as the tools used and developed for this purpose.

Section 2 outlines the design principles adopted, and related academic work. Section 3 introduces the elements constituting the ImmersiaTV end-to-end pipeline and, where there exist, commercial alternatives. Section 4 introduce our main conclusions and further planned work within ImmersiaTV to address some of the limitations detected. The content examples are adapted from the first pilot of ImmersiaTV.

2. Design principles and related work

2.1 Synchronous multi-platform playout

To create content for all devices, we need to create content that is adapted to each of them, and play it synchronously [3]–[5] . To play synchronized content, we have adapted emerging standards[2] and Gstreamer's version of the Precision Time Protocol (IEEE 1588)[3], as done, for example, in [6]. 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. In other terms: the audience is still able to watch TV sitting on their couch, or tweet comments about it. However, the audience can also use immersive displays to feel like being inside the audiovisual stream, or use tablets and smartphones to explore these omnidirectional videos, or even, in the future, to zoom in, or share portions of it through social media.

---

[1] www.immersiatv.eu

[2] http://www.etsi.org/deliver/etsi_ts/103200_103299/10328601/01.01.01_60/ts_10328601v010101p.pdf

[3] https://gstreamer.freedesktop.org/data/doc/gstreamer/head/gstreamer-libs/html/GstPtpClock.html

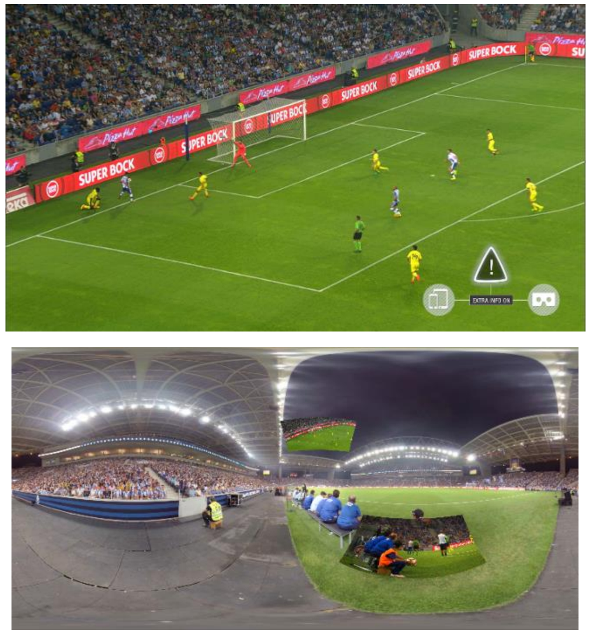


Figure 1: Top: an image typical of a traditional TV showing a football match. An insert informs the consumer that content is also available for tablets and HMDs. Bottom: a capture of an omnidirectional video with inserts of traditional cameras. This content is delivered synchronized with the main TV stream. Image courtesy of Lightbox (www.lightbox.pt).

Figure 2: 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).

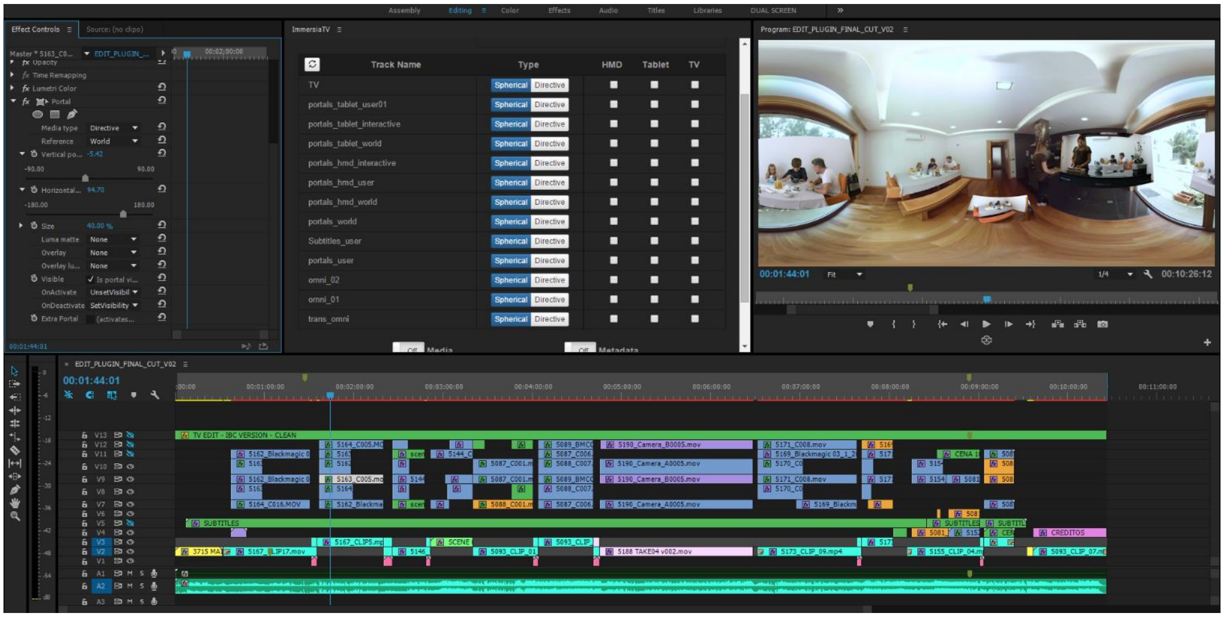


Figure 3: The Adobe Premiere ImmersiaTV panel, shown at the center, allows defining omnidirectional and directive (i.e., traditional) tracks, as well as which devices does each track target. The inserts added to the omnidirectional view, shown at right, can be edited with the ImmersiaTV Portal Effect, whose widgets are shown at the left. Image courtesy of Lightbox (www.lightbox.pt).

2.2 Portals

The idea of portals is inspired from the homonymous and famous videogame Portal.[4] In the context of streaming omnidirectional video, we introduce the idea of portals as video inserts that can be rendered in the HMD. These portals can be portions of other omnidirectional videos, which allows introducing basic interactive storytelling techniques such as scene selection or forking paths. These portals can also be inserts of traditional or directive videos. These traditional video inserts allow reintroducing classical audiovisual language that is not possible to render solely with omnidirectional videos, such as close-ups, slow motion, shot-countershot, etc. (see also Fig. 2). These strategies will not avoid the necessary precautions needed for shooting omnidirectional video [7], but they offer more options to integrate classic audiovisual conventions, particularly in a TV-centered consumption context..

3. An end-to-end pipeline

Designing and implementing a broadcast audiovisual production chain is challenging due to the diversity of processes, technologies and production practices that it requires. In this section we outline the main solutions, either adopted or implemented, for our purpose, with examples from the first pilot of ImmersiaTV.

[4] http://store.steampowered.com/app/400/

3.1 Capture

The creation of content that is both omnidirectional and traditional requires shooting simultaneously in both content formats. Preliminary tests with separate shootings for omnidirectional and traditional cameras revealed it was unfeasible to synchronize two different shootings, even when the actors in the scene were repeating the same actions. The solution found by the production team 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. This combined with an omnidirectional capture rig composed of 6 GoPro 3 Black Rig cameras allowed capturing simultaneously traditional and omnidirectional footage. However, for a joint shooting, we must address the fact that omnidirectional cameras capture the whole visual field, and therefore would show the traditional camera and the film crew behind it. This is not problematic for sports or music events, but it goes strongly against the conventions of fiction or documentary.

3.2. Edition

Dedicated stitching tools such as video-stitch studio by Video-stitch, or Autopano by Kolor, allow stitching video footage captured with camera rigs in order to create omnidirectional video. Tools for omnidirectional video edition, such as CaraVR[5] and Mettle's suite[6] allow further post-production. However, 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. 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. Further work should refine and expand the possibilities to introduce interactive capabilities, in order the user's input affects the media being rendered or other aspects of the experience.

3.3 Distribution

The media encoding uses readily available and methods, concretely 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[7].Video decoding uses the Gstreamer library[8].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.[9] Content publication is performed through a custom built website which allows triggering media conversion, as well as monitoring progress on media encoding and publishing content generating a list of content parsed at the delivery stage.

3.4 Delivery

Metadata parsing is done with a custom parser, which also generates the appropriate geometry and provides the DASH player with the appropriate DASH manifests. Delivery of omnidirectional video requires integrating the user's input with the experience, in order to select the right portion of the omnidirectional view. We are currently exploring two different software architectures for media playout. A first version is based on a custom integration of Gstreamer and the Unity3D game engine which we have released open source[10]A second version is based entirely on web technologies, and content is rendered within a browser. We are using custom code combined with three.js[11] for the geometry generation and shaka-player[12] for DASH rendering.

We have not yet completed systematic evaluation of these implementations, initial benchmarking tests shows both players, when executed on a Samsung Galaxy s6 with android 6.0 can play one 4K video. However, the additional processing power required for metadata parsing imposes reductions on the playout quality. In addition, synchronization across devices necessarily imposes seeking through a video stream being consumed, which

[5] https://www.thefoundry.co.uk/products/cara-vr/
[6] http://www.mettle.com/
[7] https://gpac.wp.mines-telecom.fr/mp4box/
[8] https://gstreamer.freedesktop.org/
[9] http://server.immersiatv.eu/public_http/metadata/ImmersiaTV.html
[10] https://www.assetstore.unity3d.com/en/#!/content/59897
[11] https://threejs.org/
[12] https://github.com/google/shaka-player

becomes slower at higher rates. Rendering for mobile-based HMD, either cardboard or Samsung GearVR, also imposes a double rendering process (one for each eye), which further decreases performance. Therefore, in practice we can currently reproduce synchronously video up to 4096x2048, bitrate of 50 Megabits per second (Mbps) and 30 frames per second on a Samsung Galaxy S6, but this resolution drops to 1024x512, bitrate of 2.3 Mbps and 25 frames per seconds when VR rendering is required. The browser-based rendering, in practice, can reproduce up to 2048x512, with 12Mbps of bitrate and 30frames per second. It seems clear that the limits in the quality that we can deliver is determined by hardware processing load, rather than bandwidth limitations. Further work will, first, develop more precise performance tests and, second, implement strategies to improve the data throughput that can be processed with off-the-shelf hardware, particularly when different media streams need to be synchronized in one device and where the orchestration of these media streams is affected by the consumer's input.

4. Conclusions and further work

The availability of open source tools as well as the existence of APIs to develop plugins for proprietary editing software makes possible, today, to design, implement and demonstrate an innovative end-to-end broadcast pipeline from scratch in less than a year. For our case, it required, approximatively, 8 software engineers with different profiles, and a mean dedication of 6 months each. Further work will be needed to optimize the media quality delivered, as well as to develop examples of content exploring more exhaustively the possibilities in terms of interaction that such a platform offers. Further work to adapt this paradigm to live production will also require implementing live production tools, as well as integrating a live MPEG-DASH encoding service. Other aspects missing, which we would like to address in the future are: access control for content conversion and publication, tiling techniques to optimize bandwidth consumption [8], [9], as well as social media integration on tablets and smartphones to further integrate with existing consumer habits.

Acknowledgements

Developing an end-to-end pipeline requires competent and motivated programmers. In ImmersiaTV, we acknowledge the contributions of Szymon Malevski and Wojtiech Kapsa, from PSNC, as well as Ibai Jurado, Isaac Fraile, David Gómez, Einar Meyerson and Juan Gordo from i2cat Foundation. David Cassany and Dr. Xavier Artigas, formerly at i2cat Foundation, also contributed in the early stages of the project. Pau Pamplona from i2cat Foundation and Maciej Glowiak from PSNC also helped with project management.

References

[1] T. Neate, M. Jones, and M. Evans, "Interdevice Media: Choreographing Content to Maximize Viewer Engagement," *Computer*, vol. 49, no. 12, pp. 42–49, 2016.

[2] R. C. Nee and D. M. Dozier, "Second screen effects: linking multiscreen media use to television engagement and incidental learning," *Convergence*, p. 1354856515592510, Jul. 2015.

[3] M. Montagud, F. Boronat, H. Stokking, and R. van Brandenburg, "Inter-destination multimedia synchronization: schemes, use cases and standardization," *Multimedia Systems*, vol. 18, no. 6, pp. 459–482, 2012.

[4] B. Rainer and C. Timmerer, "Self-organized inter-destination multimedia synchronization for adaptive media streaming," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 327–336.

[5] V. Vinayagamoorthy, R. Ramdhany, and M. Hammond, "Enabling Frame-Accurate Synchronised Companion Screen Experiences," in *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, 2016, pp. 83–92.

[6] A. Veenhuizen and R. van Brandenburg, "Frame accurate media synchronization of heterogeneous media sources in an HBB context," in *Media Synchronization Workshop*, 2012.

[7] A. Sheikh, A. Brown, Z. Watson, and M. Evans, "Directing attention in 360-degree video," presented at the International Broadcasting Convention / The Best of IET and IBC 2016-17, 2016, vol. 8, pp. 43–47.

[8] R. Skupin, Y. Sanchez, C. Hellge, and T. Schierl, "Tile Based HEVC Video for Head Mounted Displays," presented at the IEEE International Symposium on Multimedia (ISM), 2016.

[9] O. A. Niamut, E. Thomas, L. D'Acunto, C. Concolato, F. Denoual, and S. Y. Lim, "MPEG DASH SRD: spatial relationship description," in *Proceedings of the 7th International Conference on Multimedia Systems*, 2016, p. 5.

Dr. Joan Llobera is an electrical engineer and a virtual reality scientist. He obtained a double diploma in Electrical Engineering (Universitat Politècnica de Catalunya and Télécom Paris) and holds 2 Masters, one in cognitive sciences (EHESS) and another in software (UPC). In 2012 he obtained a PhD at the Universitat de Barcelona on the topic of Stories in Virtual Reality and worked as a postdoc researcher at the Ecole Polytechnique Fédérale de Lausanne. In 2015 he joined the i2CAT Foundation.

Mr. Juan Antonio Núñez is a telecommunications engineer specialized in image and sound. He obtained a master in Audiovisual Communication Systems by the Universidad de Málaga. He worked six years for Tedial, a software developer for Media Asset Management and a Media IT solutions integrator. In 2016 he joined the i2CAT Foundation.

Mr. Sergi Fernández is a computer scientist and holds a Master in Artificial Intelligence by the Universitat Politècnica de Catalunya since 2007, obtained researching the field of Natural Language Processing. Since May 2008, he has given support to the coordination of i2CAT's Media & Internet Unit and led i2CAT's involvement in different national and European projects. Since 2012 he is the director of the Media & Internet Unit of i2CAT. He previously coordinated the European project TV-Ring (www.tvring.eu) and currently coordinates the H2020 Project ImmersiaTV.

Beyond Online Multiplayer: Sharing and augmenting Tele-Immersive 3D game experiences with multiple geographically distributed users.

Konstantinos C. Apostolakis, Nikolaos Zioulis, Alexandros Doumanoglou, Dimitrios S. Alexiadis, Dimitrios Zarpalas and Petros Daras, Senior Member, IEEE

Information Technologies Institute, Centre for Research & Technology Hellas, Thessaloniki, Greece

kapostol; nzioulis; aldoum; dalexiad; zarpalas; daras@iti.gr

1. Introduction

Tele-immersion involves the combination of different technologies, such as tele-conferencing, tele-presence and virtual reality, to allow users to meet up, interact with and co-manipulate a shared digital environment [1]. Currently, the challenge lies in providing cost-efficient platforms for enabling remote interaction among multiple collaborators who could be located in variously distributed geographical locations. Developing an efficient TI platform is no trivial task, however. TI technology employs computer vision for image and/or depth acquisition and 3D reconstruction, information theory for data compression and coding, networking techniques for data exchange, between remote sites and computer graphics for rendering and 3D visualization. The majority of the workload is invested into generating realistic user representations that carry over lifelike characteristics of an individual (such as facial expressions and body deformations) into the digital world. This contributes to user immersion, and immersive tendency has been shown to contribute to both presence (i.e. the sense of being there) and flow (the sensation of being involved in an action), elements vital to the perceived enjoyment and replay value of video games [2]. Tele-immersion research aims to advance virtual reality and networking, the two fundamental pylons of the game industry's projected market growth. Games have always been an innovation driver. The video games industry has an increasingly significant impact on our everyday lives, as it develops technologies and interfaces which gradually overtake the mass market. Examples include Artificial Intelligence (AI), real-time networking, marker-less motion capturing interface devices and most recently, Virtual and Augmented Reality (VR/AR). As the consumer's market currently sees an ongoing arms race between major players in the digital games industry with respect to Virtual Reality (VR) game technology, researchers already see past contemporary use cases in order to transcend from VR into Tele-Immersive (TI) technologies and applications. Therefore, TI-related advancements will be at the forefront of research activities to come in the foreseeable future, and as a result, will be exploited and marketed by the video games industry as a continuation of the current VR trend. Since games are considered the forerunners of innovative content, services and business models of a growing digital economy, developing successful tele-immersive games to demonstrate the efficiency of the underlying technology will pave the way for other sectors (e.g. education, design, navigation and medicine) where the digital revolution has not yet started.

In this work we will thus describe a low-cost, portable 3D tele-immersive platform that focuses on ease of use and rapid deployment that enables real-time 3D reconstruction of users, who become the actual game characters of a multiplayer online 3D game titled *Space Wars*. The spectator mode of the game is further enhanced with AR combining the physical environment with the tele-immersive elements of the virtual space, thus allowing users to elicit a greater sense of co-presence and reality judgment of the digital game world and its remote participants. The remainder of this paper is organized as follows: Section 2 will present the *Space Wars* multiplayer game and Capture the Flag (CTF) mechanics. Section 3 will then describe the overall TI platform used to enhance the game, shining more light on the digital 3D reconstruction of users, procedures related to efficient networking between remote TI-stations and 3D rendering complemented by AR enhancements and how they contribute to the overall user experience. The paper concludes with a short discussion in Section 4.

2. Tele-immersive multiplayer games

Seeing how the games industry can be a significant innovation driver, we developed *Space Wars*, a tele-immersive multiplayer game to act as a proof of concept for a portable, tele-immersive platform (see Section 3). The game itself connects two remote players, shown riding on hoverboards and enables competition against one another in order to capture holographic flags placed in a symmetrical, sci-fi, distant future arena. We embedded tele-immersive capabilities into the game to improve the overall Quality of Experience (QoE), mainly the game's immersive tendency and users' feeling of co-presence. Within *Space Wars*, we employed one of the most well-known challenge

archetypes and integral component in a variety of successful game genres, the Capture the Flag (CTF) type of multiplayer match [3]. In this mode of play, a team of cooperating players needs to work together to *protect* their flag, positioned at a pre-defined location on the game map (which serves as a spawning point for players who have been defeated), while at the same time attempting to *capture* the opposing team's flag. The player capturing the opposing flag is tasked with returning it to his/her own team's spawning ground, while the rest of the team has to *escort* and protect the flag bearer until he/she reaches the designated location [3]. Successfully carrying the opposing team's flag to the destination point awards the whole team one point. The team whose flag has been captured needs to eliminate the player holding their flag to avert the opposing team from scoring.
Embedding the CTF archetype with tele-immersive properties provides a significant visual enhancement, allowing players to physically enter the digital game world and actually see their photorealistic 3D representations riding the digital hoverboards. This visualization novelty uniquely offered through TI cements a direct, real-time link between the real and digital game world. Additionally, we implemented a motion-control gesture-based interface that allows users to navigate the environment (leaning the torso to the left or right while bending knees in a "surfer's posture"), attack the other player (pull one hand back to charge, then extend forward to release a projectile attack) and capture the enemy flag (navigating over the flag and see it latch onto the player's hoverboard), thus exploring intuitive aspects of user movement to interact with the game world, while simultaneously maintaining seamless, real-time visual communication and consistency with the other player. *Space Wars* was successfully demonstrated to live audiences during several public events receiving high acclaim from participating users. Snapshots of these events, as well as the game's screenshots can be seen in Figure 1.

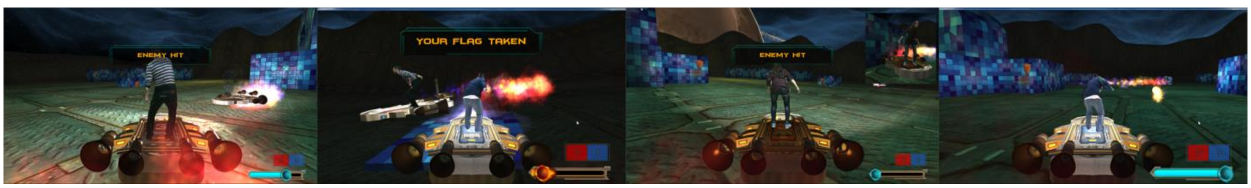


Figure 1: *Space Wars* screenshots of players engaging in CTF multiplayer matches.

3. Portable TI-Platforms for multiplayer gaming interactions
Figure 2 shows the overall architecture of the prototype portable TI platform for sharing multiplayer game experiences in *Space Wars*. We follow Smolic's distinction in describing the multidisciplinary nature of the depicted platform in the following paragraphs, promptly addressing *capturing and reconstruction*, *data compression and transmission* and *rendering* [4].

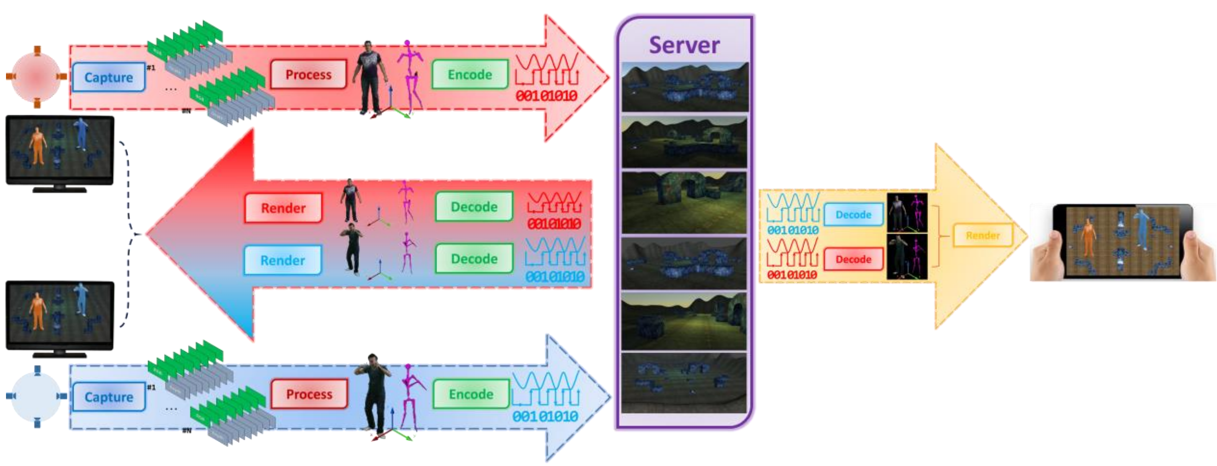


Figure 2: TI-Platform Architecture

3.1. Capturing and Reconstruction
As mentioned in the previous Section, TI stations optimized for gaming should be portable and easily deployable, taking cost-efficiency into account to effectively attract consumer interest. Towards this goal, the proposed platform uses low-cost commercially available RGB-D sensors (namely, the Kinect for Xbox One gaming hardware), and utilizes an easy-to-use, multi-sensor calibration scheme, which allows for coarse sensor positioning around a pre-

defined capturing space, without further restrictions or custom configurations. Unlike its predecessor, the current configuration of the Kinect for Xbox One gaming hardware device limits the number of sensors allowed to be connected on a single platform to just one sensor per computer. This is the primary reason for requiring multiple client machines to operate the sensors. This limitation is compensated for by the sensor's superior quality of both the depth and RGB cameras, which stream data to the central hub at a minimum 15FPS rate through local Ethernet connections. The game capture site comprises 4 Kinect sensors placed in a circular trajectory of radius 2.5-3m, at approximately 90◦ angle intervals around the capturing area, and 5 PCs, one for each sensor and one acting as a central hub collecting and processing the data. Software-based synchronization methodologies on the capturing and data acquisition sides, based on the central PC broadcasting a signal to all clients to receive the latest captured frames from all sensors, are applied to compensate for the sensors' lack of hardware triggering. This approach has been optimized to guarantee time consistency between data acquired from all viewpoints to a point where desynchronization becomes noticeable only if users within the capturing area start moving significantly very fast. The platform is externally calibrated using a fully automated calibration method, involving a simple calibration structure be built out of commercially available packing boxes with printed paper QR code markers glued on their sides[13]. The box totem is then used as a registration anchor, and is positioned roughly at the center of the capturing space once, prior to the game session. The process uses SIFT [5] correspondence establishment between the acquired views of the calibration totem and its virtual CAD 3D model counterpart. The overall calibration process takes only a couple of minutes to complete, making the platform more robust against outside intervention (e.g. spectator crowds accidentally knocking off a single sensor, etc.) and more friendly towards inexperienced users.

Real-time full-3D geometry reconstruction of the players inside the capturing area is implemented using a stream of spatially and temporally aligned color and depth frames collected by the central processing hub. The 3D user mesh geometry is generated by mapping depth frame pixels corresponding to the captured user (e.g. *"foreground pixels"*) to 3D vertices, using the intrinsic depth camera's parameters and the extrinsics matrix of each sensor, which was calculated during the calibration process described above. Subsequently, a manifold, watertight mesh is generated using the Fourier Transform-based volumetric reconstruction method described in [6], [7], to approximate the user's 3D geometry. A multi-texturing approach is then used to embed appearance to the reconstructed mesh. The RGB values extracted from the frame viewpoints doubling as texture maps are blended following a weighted approach [8]. The resulting textured mesh can then be treated as any other 3D asset within a game engine IDE, such as Unity[14]. All processes are implemented on CUDA-enabled GPUs to ensure near real-time performance.

3.2. Data Compression & Transmission

The captured data at the clients mentioned in the previous paragraph are encoded following an intra-compression scheme to minimize transfer latency. An open-source, lossless algorithm is used for coding the depth data [9] while HD color textures are compressed using the JPEG standard codec [10]. Towards an end-to-end tele-immersive pipeline to achieve real-time integration of player captured 3D reconstructed meshes with the 3D virtual environment of the *Space Wars* game, a strong interrelation between the data format, compression scheme, transmission layer and networking architecture need to be accounted for. In this respect, the *Space Wars* TI multiplayer network is server-based rather than peer-to-peer, to account for scalability and multiple participants (players, as well as spectators, a common practice in multiplayer games). Each TI client station therefore transmits data to the server, who is responsible for synchronizing the game state and transmitting it back to all connected users. Efficient compression of the 3D meshes utilizes intra-frame static mesh codecs via OpenCTM [11], [12] allowing for the platform to be scalable to network conditions [13].

3.3. Rendering

As previously mentioned, textured 3D meshes of reconstructed users can be rendered within the Unity game engine IDE. Using a standard OpenGL-based graphics pipeline, each mesh can be seamlessly integrated into the game environment and considered within scene lighting calculations to cast and receive shadows on other virtual objects. Advanced shaders can also be written to embed different effects onto the users' appearance, such as outlining and vertex displacement. In one such case, *Space Wars* uses a custom written Cg shader to cause the user mesh to "explode" into thousands of particles when hit by an enemy projectile.

Recently, the consumer's market has seen a substantial growth with respect to heterogeneity of end-users' devices, ranging from home consoles to smartphones and Virtual, or Augmented Reality (AR) headsets. The Unity game engine is already equipped to support the latest head-mounted devices (HMD), such as Microsoft's HoloLens. Game

[13] Instructions and virtual CAD model are available at http://vcl.iti.gr/spacewars/.

[14] https://unity3d.com/

developers are thus called upon to explore new features and ways in which games can be played, and interactions among players can be fostered. AR is becoming very popular in the gaming field and as such, we address its use in tele-presence applications to support captivating real-time interactions of remote users in tele-immersive games. We incorporated the Vuforia[15] platform to offer a means of spectating a *Space Wars* match in progress through the user's tablet or HoloLens goggles. Through registration between the 3D game world and an overhead orthographic view of the rendered game environment, we create real-life image markers that can be augmented with virtual elements, effectively enhancing the feeling of presence. In Space Wars' case, the 3D walls, player bases and flags emerge on top of a printed piece of cardboard resembling the game level's floor. Player avatars, hoverboards and projectiles can be seen navigating this augmented environment at any possible viewing angle. The size of the printed marker determines the size of the virtual objects being rendered, thus allowing various amounts of comfort for viewers to spectate the action on a tabletop or on the floor. This feature is displayed in Figure 3.

Figure 3: Augmented *Space Wars* terrain rendered from different view angles on textured paper marker.

4. Discussion

Tele-immersive systems and augmented reality are emerging as technological enablers that will soon be embraced by various fields incorporating innovative solutions to enhance their services and business models. As digital games often push the technological limits to advance the maturity of technologies for other sectors to closely follow, our work demonstrates the potential of portable, cost-efficient TI stations, hopefully pioneering a new era of online multiplayer games that will contribute to further advance research towards efficient and robust user 3D reconstruction, networking, data compression and augmented reality rendering techniques. We hope (and foresee) an increased take-up of such technologies to provide aid in service of human healthcare and education as a result of gaming companies striving to achieve dominance in a groundbreaking new market for digital TI games. Our own future work will examine ways of enhancing the current TI infrastructure to support more fully immersed participants' reconstructions, addressing the inherent networking and data compression challenges. Furthermore we aim to launch thorough Quality of Experience (QoE) and Quality of Service (QoS) studies to identify elements needing improvement and scientifically ground the benefits of TI and AR bring to multiplayer online games.

References

[1] R. Singh, and A. K. Singh. "*Transcending from virtual reality into tele-immersive technologies and applications: a perspective*", in ACM Ubiquity, p. 6, June 2008.

[2] D. Weibel, and W. Bartholomäus. *"Immersion in computer games: The role of spatial presence and flow"*, in International Journal of Computer Games Technology, p. 6, 2011.

[3] J. B. Rocha, S. Mascarenhas, and R. Prada. *"Game mechanics for cooperative games"*, in ZON Digital Games, pp. 72–80, 2008

[4] A. Smolic. *"3D video and free viewpoint video from capture to display",* Pattern recognition, vol. 44, 2011.

[5] D. G. Lowe. *"Object recognition from local scale-invariant features"*, in Proc. of the 7th IEEE ICCV, 1999, vol. 2.

[6] M. M. Kazhdan. *"Reconstruction of solid models from oriented point sets"*, in 3rd Eurographics SGP, 2005.

[7] D. Alexiadis, D. Zarpalas, and P. Daras. *"Real-time, full 3-D reconstruction of moving foreground objects from multiple consumer depth cameras"*, IEEE Trans. on Multimedia, vol. 15, no. 2, pp. 339–358, 2013.

[8] D. Alexiadis, D. Zarpalas, and P. Daras. *"Real-time, realistic full-body 3D reconstruction and texture mapping from multiple kinects"*, in Proc. IEEE IVMSP, 2013.

[9] Y. Collet. *"Lz4: Extremely fast compression algorithm"*, code.google.com, 2013.

[10] G. K. Wallace. *"The JPEG still picture compression standard"*, Comm. of the ACM, vol. 34, pp. 30–44, 1991.

[11] M. Geelnard. *"OpenCTM, the open compressed triangle mesh file format"*, 2010.

[12] D. Alexiadis, A. Doumanoglou, D. Zarpalas, and P. Daras. *"A case study for tele-immersion communication applications: From 3D capturing to rendering"*, in IEEE VCIP, 2014.

[13] S. Crowle, A. Doumanoglou, B. Poussard, M. Boniface, D. Zarpalas, and P. Daras. *"Dynamic adaptive mesh streaming for real-time 3d teleimmersion"*, in Proc. of the 20th Int. Conf. on 3D Web Technology, 2015.

[15] https://www.vuforia.com/

Konstantinos C. Apostolakis is a software engineer and independent game developer, working as a research assistant for the Information Technologies Institute of the Centre for Research and Technology Hellas (CERTH). He received his university degree in Computer Science from the Computer Science Department of the Aristotle University of Thessaloniki in September 2009. He has been involved in European Research Projects and has co-authored 15 scientific papers for international journals and conferences. His interests include game design & development, game engine programming, accelerated computer graphics, 3D character modeling & animation and augmented reality applications.

Nikolaos Zioulis is an Electrical and Computer Engineer working in the Information Technologies Institute (ITI) of the Centre for Research and Technology Hellas (CERTH) since October 2013. He obtained his university degree from the Electrical and Computer Engineering department of Aristotle University of Thessaloniki in 2012. His interests include 3D processing and 3D graphics, particularly performance oriented real-time computer vision. He has been involved in four (4) European Projects, tackling areas and fields such as motion capturing, motion evaluation, 3D reconstruction, multi modal capturing and acquisition, spatial and temporal registration from heterogeneous sensors, algorithmic design, quality analysis and tele-immersion platforms.

Alexandros Doumanoglou is an Electrical and Computer Engineer working in the Informatics and Telematics Institute since April 2012. He received his university degree from the Electrical & Computer Engineering department of the Aristotle University of Thessaloniki in 2009. His research interests include computer vision, pattern recognition, 3D reconstruction, image recognition, 3D Graphics and GPGPU computing.

Dimitrios Alexiadis is an Electrical and Computer Engineer who received the Diploma degree and the Ph.D. degree in Electrical and Computer Engineering, in 2002 and 2009, respectively, both from the Aristotle University of Thessaloniki - Greece. Since March 2011, Dr Alexiadis has been working in the Informatics and Telematics Institute (I.T.I), as a postdoctoral research fellow. He has (co)authored eighteen (18) high-impact journal papers and over twenty (20) international conference papers. His main research interests include still and moving image processing, face recognition, 2-D and 3-D motion estimation, structure from motion and stereo/multi-view image processing.

Dimitrios Zarpalas is an Electrical and Computer Engineer (2003, Aristotle University of Thessaloniki) with an MSc in computer vision (2006, The Pennsylvania State University, USA) and a PhD in Medical Informatics (2014, the Medical School, Aristotle University of Thessaloniki). Since 2007, he joined the Information Technologies Institute of the Centre for Research and Technology Hellas. His current research interests include real time tele-immersion applications, 3D computer vision, AR technologies, 3D object recognition, 3D motion capturing and evaluation, while in the past has also worked in indexing, search and retrieval and classification of 3D objects and 3D model watermarking.

Petros Daras is a Principal Researcher Grade A', at the Information Technologies Institute of the Centre for Research and Technology Hellas. He received the Diploma in Electrical and Computer Engineering, the MSc degree in Medical Informatics and the Ph.D. degree in Electrical and Computer Engineering all from the Aristotle University of Thessaloniki, Greece in 1999, 2002 and 2005, respectively. He is the head researcher of the Visual Computing Lab coordinating the research efforts of more than 35 scientists. His research interests include 3D media processing and compression, multimedia indexing, classification and retrieval, annotation propagation and relevance feedback, bioinformatics and medical image processing. He has co-authored more than 160 papers in refereed journals and international conferences, and has been involved in more than 30 national and international research projects.

Of Standards and Herrings: Tales of Technology and Tumult

Nicholas F. Polys
Virginia Tech, npolys@vt.edu

1. Introduction

This letter provides a high-level sketch of the actors, motivations, and dynamics in the 3D content industry, the WWW, and standards organizations. Market Churn has cost investors and consumers. Here using historical perspective, we consider and debunk common falsehoods in the thinking about virtual reality and mixed and augmented reality technology. We parse out the alphabet soup of standards and their organizations and consider how they are scoped. If the dream of portable, shared information spaces is to become a reality, we must cooperate and play well with others in the Web ecosystem.

The last two decades have witnessed several revolutions in consumer electronics and multimedia production. Central to these revolutions has been the relentless march of hardware innovation: improved CPU and network speed, RAM sizes, and rendering speed (GPUs) put the supercomputers of the 1980s in the hands of consumers in the 2000s. While the computer hardware industry has seen several consolidations and disruptive events over the years, it seems remarkably stable when compared with the software and media sectors over this time. In that time, numerous technologies and companies have come and gone.

As proprietary platforms and bankruptcies proliferated, so much virtual world content and destinations were consigned to the bit-bucket, never to be run again. Authors and enterprises suffered from the loss of assets. Venture capitalists suffered from the loss of investment. The boom-and-bust cycles of the 3D Internet also have hurt the consumer and the market in general as fragmentation and obsolescence crippled any mass adoption and growth [1]. Despite these setbacks, 3D content and applications are not only surviving, but flourishing in specific domains where durability, portability, and interoperability are mission-critical.

In this paper, I survey the last 20+ years of Web and interactive 3D technology and offer reflections on the dynamics of communities, standards, and industry. Considering the current state of affairs through the lens of history can provide lessons and direction as to how we can move forward to realize the promise of this new media. In this first Section, I motivate and scope and consideration for this special issue article. Section 2 describes key communities of stakeholders and Section 3 discusses technology ecosystems and the scoping of various specifications and standards; Section 4 provides a historical narrative and reflection on market forces. Section 5 details the standards organizations and Section 6 considers the paths forward for Mixed and Augmented Reality standards.

1.1 The Web

It is hard to understate the impact of open networked information technologies on our modern life. Certainly some readers started their professional life with typewriters and the paper economy and can appreciate this. As the 1990s unfolded, a great optimism lit up society as the World Wide Web (WWW) began to grow and connect communities around the world. The development of technology specifications and standards such as HTTP and HTML enabled a whole economy and ecosystem to be built out of networked information. Hailed as a 'great equalizer' and promising a 'democratization of information', the Web quickly became a cental means for communication and access for a wide variety of users and communities. HyperText and multimedia presented exciting new possibilities for cross-linking information across digital works.

1.2 Cyberspace

The idea of immersive virtual worlds and shared information spaces may have began with Sutherland's paper about 'the Ultimate Display' [2]. However, the notion of a parallel digital world of spatialized information really reached public consciousness through the literary works of William Gibson [3] and Neil Stephenson [4]. Although dystopian in their tone, the potential of networked digital spaces sparked many early technologists, especially youth, who considered it their generation's platform to change to world… or create any number of 'virtual' alternatives.

1.3 Interactive 3D

In these early days, the computing technology for graphics rendering was limited. In the 1950s, the math behind perspective drawing was being worked out and the first computer graphics examples were coming out of engineering applications such as aerospace design and ergonomics. Early interactive drawing algorithms were concerned with the speed of hidden surface removal for rendering 3D objects with 2D lines [5]. Today, GPUs have hundreds of parallel processors and the programmable shader pipeline enables richness and realism to imagery at interactive rates. Dedicated graphics accellerators are now showing up on mobile devices, bringing the rendering power of the workstation to new form-factors: handheld devices or head-mounted displays. Over this amazing trajectory of hardware improvement, software and data formats proliferated. The most persistent has been the internationally-ratified ISO Standards of Virtual Reality Modeling Language (VRML) and Extensible 3D (X3D)[6], which we will describe below.

2 Stakeholder Communities

Considering the growth of the Web and the 3D content industry since the mid 1990s, there are many stakeholders operating throughout the ecosystem. This section reviews some of the main concerns of each and their view of platforms.

2.1 Content Creators

The expressive and creative power of interactive 3D proved both attractive and valuable to many. The flush of excitement sparked a flood of applications and worlds from science, education, archaeology, and simulation to multi-user social worlds. For those whose property and services are actually 3D objects or 3D systems with behaviors, the 3D content itself is the currency. The durability of their assets is critical as well as the security of their intellectual property. The only thing more expensive than creating 3D content is creating it twice (or more) when the technology changes or the company dies.

2.2 Technologists

The computer industry (here hardware and software producers) continually push innovation and novel solutions from real-time rendering techniques to algorithms for big data. Whether it is a prototype or product, architects and developers play a critical role. The evaluation and adoption of new languages and tools typically happens here; while technology strategies can be based on long-term or sustainability thinking, typically they are governed by ease-of—use or short-term rewards.

2.3 Business

Clearly, there is a lot of investment in VR and MAR companies. The current cycle of business boom presents many new (or renewed) opportunities. We should remind ourselves how much money was spent and lost through the hubris of 3D startups, the duplicated effort, and the mis-steps of big corporations. While there are a number of business models in the wide industry of 3D, the players themselves rarely grasp the ecosystem they are in, and rarely play well with others.

2.4 Consumers

Lastly, it is important to recognize this group as stakeholders since the number of Web users continues to skyrocket. End users face many challenges in getting into 3D, from the usability of the user interface to the babel-ization and fragmentation of required installations. While some-early adopters are accessing and experiencing 3D content on the Web, they are often using several different programs, each with its special value (and limitations). The trend toward empowering consumers to be producers face the same challenges for mass market success.

3. Technology Ecosystems

3.1 Content Models

The first ecosystem to discuss is the content model. HTML has a content model. 3D assets (for VR or MAR) must also be described in structured way- in what is commonly known as a scene graph. A scene graph *is* a content model; for example, describing what methods or parameters are supported for different nodes in the scene graph, how Groups and Transforms work in a hierarchy, how lights are scoped, where does texture coordinates or per-vertex color information go. The content can be encoded in multiple ways and programmatically accessed with different languages (next sections).

3.2 Encodings

Data can be encoded in many ways. The encoding determines several important outcomes, especially file size and parsing requirements. For example, a scenegraph may be encoded in a plain text encoding (like utf-8), as an XML document, or in binary. Different encodings can be useful in different circumstances; consider hand editing is easy in a text file but an XML can be validated, but binary may be better for publishing models.

3.3 APIs & Language Bindings

The Application Programming Interface (API) is the set of methods or functions supported by a computer program, software object or hardware device. Typically, APIs are used to structure a system and how it communicates and integrates with other programs. APIs may be bound to different languages; for example, one might program a scene graph with Javascript, Java, Python, or C++.

4. Market Mentalities

Interested readers are referred to the main conferences in the field, where a rich historical record of technology evolution and revolution can be found. Most notable are the ACM SIGGRAPH / Eurographics conference on Web3D, IEEE VR and IEEE ISMAR.

Herring Number 1: "Content standard S is slow"

Such statements are Red Herrings because it is the implementation of the standard that has a runtime, not the standard itself. A second difficulty with such a statement is that slow is a relative term and rarely is the difference specified. There are many factors that determine runtime, for example, what data is being included or represented and how it is encoded.

Herring Number 2: "We need a killer app for 3D"

This statement first hinges on the assumption that such a thing could be defined and second, that 3D as a media type would develop similarly to prior examples of 'killer applications' such as email, spreadsheets, or GPS. In the first case, speakers are commonly referring to something that drives mass market consumers: lots of users or revenue. They often ignore the fact that within certain vertical markets, 3D has already established its central value (it is the killer itself).

Herring Number 3: "Company C is the future of the 3D internet"

This is related to the prior herring and is usually a sign that Company C is a subscriber to the 'Highlander Myth': that there can be only one. This is essentially marketing since there is no evidence from any current (or dead) company that such a unilateral approach could work with the Internet, the Web, and their ecosystems.

Herring Number 4: "Technology T is open"

This is a herring often invoked when considering proprietary technology adoption. 'Open' is not the same as 'non-proprietary', 'non-patented', or 'royalty-free'. Being standardized and published openly is no guarantee of intellectual property exposure; look for explicit licensing and open-source implementations.

5 International Standards

Standards are different from specifications in that they are typically developed, evaluated, and ratified by a community and a cooperative process. This section provides an overview of relevant organizations and their standards.

5.1 International Organization for Standardization/International Electrotechnical Commission (ISO-IEC)

The Joint Technical Committee 1 (JTC1) of ISO and IEC: Information Technology was created in 1976 (http://www.iso.org/iso/jtc1_home.html). Its relevant SubCommittees:

- **SC 24:** Computer graphics, image processing and environmental data representation. This is the path Web3D Consortium standards travel (via a Cooperative Agreement; declared royalty-free).
- **SC 29**: Coding of audio, picture, multimedia and hypermedia information. This is the path Motion Pictures Experts Group (MPEG) standards travel. MPEG standards may be patent-bearing.
- **Joint Technical Working Group on Mixed and Augmented Reality** (SC24 + SC29) are standardizing a Reference Model for MAR; this work is declared to be royalty-free.

5.2 World Wide Web Consortium (W3C)

The W3C was founded in 1994 by members of MIT, CERN, and DARPA. They have produced and standardized the content models and APIs that underpin the World Wide Web: HTML, XML, CSS, SVG, and many others. The W3C promotes a royalty-free policy. The Web3D Consortium has an official Liaison relationship with the W3C. https://www.w3.org

5.3 Web3D Consortium

The not-for-profit Web3D Consortium was founded in 1996. Virtual Reality Modeling Language (VRML), Extensible 3D (X3D), Humanoid Animation (H-Anim) are openly published and freely available. They are also declared royalty-free. X3D is several standards covering the Scene Graph, Encodings, and the API respectively. X3D is defined with a modular, component-based architecture that has been extended for CAD, Volume Rendering. Like VRML, X3D is specified with no assumptions about display platform or user interface devices; because of this abstraction, X3D can be rendered with OpenGl, WebGl, DirectX, or even Flash. The X3D scene graph can be encoded in utf8, XML, or binary. The Scene Access Interface (SAI) is the API.for X3D runtimes; Javascript and Java are standardized and several more (C++, Python) demonstrated. http://www.web3d.org/

5.4 Khronos

Founded in 2006, Khronos is a not-for-profit consortium that develops and administers the OpenGL API and related technologies such as as WebGL. While typically focused on the low-level hardware-software rendering layer, Khronos has ventured into content models and encoding standards: Collada (XML-encoded content model) was standardized there. https://www.khronos.org/

5.5 Open Geospatial Consortium (OGC)

OGC was founded in 1994. Recent standards include WebMapService (WMS), WebFeatureService (WFS). The Keyhole Markup Language (KML) was standardized through OGC. Recent Liaison relationships with the Web3D Consortium have produced Web3D Service (Web3DS), which supports X3D and Collada for 3D GIS portrayal. The OGC also standardized Augmented Reality (ARML) describes a content model and API for location-based models using KML and Collada. ARML's front matter declares it contains patented technology. http://www.opengeospatial.org/

5.6 Motion Picture Experts Group (MPEG)

Founded in 1988, this group focuses on the encoding and transmission of multimedia content (video, audio, and 3D data). The codecs and other specifications produced may carry patents and royalty requirements. MPEG-4 Part 11 adopted portions of the VRML and X3D scene graph and included a Binary Format for Scenes (BIFS).

5.7 National Electrical Manufacturer's Association (NEMA)

This organization is the current home to the DICOM standardization activities ('Digital Imaging and Communications in Medicine'). Medical imaging, media storage, application services and display functions are covered by the standard. The description of 3D surfaces in DICOM matches the X3D mesh data types. The Web3D Consortium has an official Liason relationship with DICOM. http://dicom.nema.org/

5.7 Institute of Electrical and Electronics Engineers Standards Association (IEEE-SA)

Has an association that has develops a broad range of standards generally supporting the network: from wireless protocols to medical device communications. https://standards.ieee.org/

6 Mixed Reality: Status & Future Work

6.1 Status

Reflecting on the lessons of history, we can see great savings and efficiencies through the evolution and adoption of standards. The stakeholders in VR and MAR must be educated and demand standards and invest in improving them. The durability, portability, and interoperability of the X3D standard has been shown in multiple implementations over a decade and half with applications such as environmental monitoring [6], medical and volume rendering[7], networks [8], spectrum visualization [9], cities [10], and multi-user mirror worlds [11][12]. Proving the concept early, ARToolkit [13] supported ISO-IEC VRML as the source for 3D models as augmentation. VRML assets and environments still run today and faster than ever. Research into Mixed Reality using X3D for specification began in 2006 [14] with annual innovations from around the world [15] [16]. Work continued, for example, matching lighting [17] and depth [18] between the virtual and real scene. Recent member activities include the extension of the X3D standard for MAR applications, specifically proposing new nodes for the Camera and Tracking Sensors, BackDropBackground and adding additional projection parameters to the Viewpoint node.

6.2 Future Work

While this is a tale of technology tumult, there is one clear lesson from both the lost opportunities and gritty survivors: Cooperate or Die; build partnerships, scope compatible standards, build the ecosystem of Cyberspace.

ACKNOWLEDGMENT

The author wishes to thank all the volunteers and contributors to international technology standards.

REFERENCES

[1] Polys, Nicholas F. "Recall the Lost Frontiers of Virtual Worlds." Letter to the Editor *Communications of the ACM* 54.5 . 2011.

[2] Sutherland, Ivan E. "The ultimate display." *Multimedia: From Wagner to virtual reality*. 1965.

[3] Gibson, William Neuromancer. 1984.

[4] Stephenson, Neal Snow Crash. 1992.

[5] Weiler, Kevin, and Peter Atherton. "Hidden surface removal using polygon area sorting." *ACM SIGGRAPH computer graphics*. Vol. 11. No. 2. ACM, 1977.

[6] Polys, Sforza, Hession, Munsell. "Extensible Experiences: Fusality for Stream and Field". In *Proceedings of the 21th International Conference on 3D Web Technology* (Web3D '16). ACM, New York, NY, USA. 2016.

[7] Polys, N. and A. Wood. "New Platforms for Health Hypermedia." Issues in Information Systems **13**(1): 40-50. 2012.

[8] Peter J. Radics, Nicholas F. Polys, Shawn P. Neuman, and William H. Lund. "OSNAP! Introducing the open semantic network analysis platform", Proc. SPIE 9397, Visualization and Data Analysis. 2015.

[9] Nikita Sharakhov, Vuk Marojevic, Ferdinando Romano, Nicholas Polys, and Carl Dietrich. (2014). "Visualizing real-time radio spectrum access with CORNET3D". In *Proceedings of the Nineteenth International ACM Conference on 3D Web Technologies* (Web3D '14). ACM, New York, NY, USA, 109-116. 2014.

[10] Ji-Sun Kim, Nicholas Polys, and Peter Sforza. "Preparing and evaluating geospatial data models using X3D encodings for web 3D geovisualization services". In *Proceedings of the 20th International Conference on 3D Web Technology* (Web3D '15). ACM, New York, NY, USA, 55-63. 2015.

[11] Nicholas F. Polys, Benjamin Knapp, Matthew Bock, Christina Lidwin, Dane Webster, Nathan Waggoner, and Ivica Bukvic. "Fusality: an open framework for cross-platform mirror world installations". In *Proceedings of the 20th International Conference on 3D Web Technology* (Web3D '15). ACM, New York, NY, USA, 171-179. 2015.

[12] Tilden, D., A. Singh, N. F. Polys, and P. Sforza. "Multimedia mashups for mirror worlds", *Web3D '11 Proceedings of the 16th International Conference on 3D Web Technology*, Paris, ACM. 2011.

[13] Kato, I. Poupyrev H., Mark Billinghurst, and Ivan Poupyrev. "Artoolkit user manual, version 2.33." *Human Interface Technology Lab, University of Washington*. 2000.

[14] Polys, Nicholas F., & Ray, Andrew. "Supporting Mixed-Reality Interfaces through X3D Specification". Workshop on Mixed-Reality Interface Specification, *Proceedings of IEEE Virtual Reality*, IEEE Press.2006.

[15] Jung, Yvonne, et al. "Enhancing X3D for advanced MR appliances." *Proceedings of the twelfth international conference on 3D web technology*. ACM, 2007.

[16] Polys, Nicholas and Brutzman, Don and Steed, Anthony and Behr, Johannes. *Future Standards for Immersive: Report on the IEEE VR 2007 Workshop.* IEEE Computers Graphics & Applications Vol. 28, Number 2, IEEE Computer Society, 2008.

[17] Franke, Tobias, and Yvonne Jung. "Precomputed radiance transfer for X3D based mixed reality applications." *Proceedings of the 13th international symposium on 3D web technology*. ACM, 2008.

[18] Franke, Tobias, et al. "Enhancing realism of mixed reality applications through real-time depth-imaging devices in X3D." *Proceedings of the 16th international conference on 3D web technology*. ACM, 2011.

Nicholas F. Polys is Director of Visual Computing with Virginia Tech Research Computing Group and Affiliate Research Professor in the Department of Computer Science. He has developed interactive 3D graphic content and systems since 1997. His research interests lie at the center of graphics and Human Computer Interaction: the intersection of visualization, virtual environments, and perception. After his undergraduate research in Cognitive Science at Vassar College (1996), he jumped into the networked information space of the WWW developing audio, visual, and 3D assets and software. His doctoral work at Virginia Tech (2006) examined perceptual cues and layout behaviors for Information-Rich Virtual Environments for desktop to immersive platforms. He is a member of ACM, IEEE Computer Society, and the Web3D Consortium. He is a co-author of the international standard (ISO-IEC) Extensible 3D (X3D) and elected Director and President of the Web3D Consortium.

Overview of Standardization Efforts for Augmented Reality

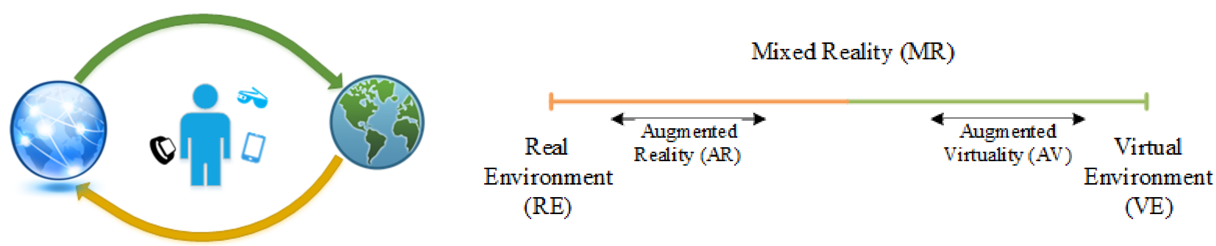
Pavlos Kosmides, Konstantinos Demestichas and Evgenia Adamopoulou
Institute of Communication and Computer Systems, National Technical University of Athens, Greece, {pkosmidis, cdemest, eadam}@cn.ntua.gr

1. Introduction

The last two decades have witnessed the rapid growth of Information and Communication Technologies (ICT), including the revolution of handheld mobile devices (smartphones), cloud services and infrastructures, high-speed wireless/mobile communications, and the Internet of Things. This progress has led to a parallel increase in both the number and expectations of mobile users, leading to a continuous struggle to meet and exceed user satisfaction. In this context, researchers and industry have invested a lot on combining the physical and the digital worlds, in order to make a leap in user experience improvement –away from the conventional desktop- or smartphone- based interaction paradigm.

In this light, Augmented Reality (AR) is one of the most promising technologies to achieve this combination. AR integrates computer-generated objects in the real environment and allows real-time interactions [1]. Unlike Virtual Reality (VR), AR allows users to have direct contact with the real world, but it also augments users' surrounding with virtual computer-generated objects. This is one of the major differences between AR and VR, since in VR the user is immersed into a completely virtual environment, interacting with non-real objects. AR has the potential to revolutionize the way users experience information and interact with machines, other people and the environment, in the same way that that Moore's law and cloud infrastructures have immensely improved processing power, and 4th generation/wireless standards have boosted telecommunications.

AR technologies need to satisfy three main characteristics: (i) combination of real and virtual objects in a real environment, (ii) user interaction in real time, and (iii) perfect alignment between virtual and real objects [2]. Specifically, in an AR context, there is a constant interaction between the digital and real world (Figure 1(a)). Data from the real world are usually collected from sensors, while the interaction with the digital world is made using application services that allow the user to have direct contact with the real environment. To this extent, handheld devices, and especially smartphones, are an excellent platform for developing AR services, while their wide adoption has increased industry's interest in investing on AR applications [3], like Pokémon GO [4], layar [5] and others (Figure 2).

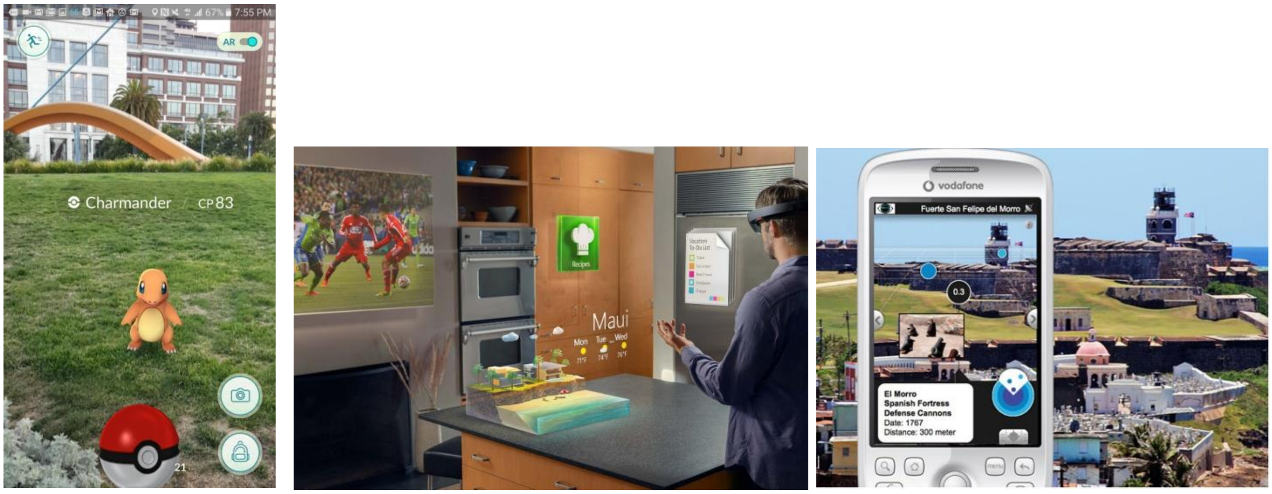


(a) (b)

Figure 1: (a) Interaction in the Augmented Reality ecosystem, (b) Mixed Reality and Reality-Virtuality Continuum

To further clarify the difference between the real and the virtual environments, the authors in [6] introduce the Reality-Virtuality (RV) continuum (Figure 1(b)). Although both the real environment (RE) and virtual environment (VE) may exist as separate entities, they are also considered as opposite poles of the RV continuum, creating a spectrum where AR, Augmented Virtuality (AV) and Mixed Reality (MR) can be defined. From Figure 1(b), it becomes apparent that AR covers a portion of the RV continuum that is close to, but not exclusively part of, the real environment.

The first appearance of the term Augmented Reality was made in the early 90s in [7], where a head-mounted display was used in order to assist workers to put together wiring harnesses on boards. According to [8], it was not until the late 90s that AR became an active research field, as various International Workshops and Symposia emerged. The main technical components that are used to create AR are: (i) display technologies in order to combine the real and virtual worlds (e.g. smartphones, Microsoft HoloLens [9], eMacula contact lenses [10]), (ii) tracking sensors and other approaches to acquire a user's position and orientation, and (iii) user interface to achieve real-time interaction.



(a) (b) (c)

Figure 2: Augmented Reality applications: (a) Pokémon GO [4] – chasing Charmander, (b) Microsoft HoloLens [9] – kitchen use case scenario, (c) layar [5] – using EyeTour application.

AR has met great response in many areas of users' daily life. One of the most important fields of AR application is education. A literature review on using AR in educational studies in order to support learning and teaching, is presented in [11] covering all levels of education, while a special focus on children with disabilities is given in [12]. Another interesting application of AR is made in medicine, where, for instance, in [13] the authors propose the use of AR in order to get familiar with the human anatomy in an easier manner. Although it is clear that we are in a period where AR applications are growing rapidly, standardization efforts are still an ongoing process that has attracted the attention of industry, academia and policy makers.
The rest of this paper is organized as follows. In Section 2, we present the benefits of developing standards for AR applications. In Section 3, we present the most relevant standards and corresponding organizations in the field of AR. Finally, the paper concludes in Section 4.

2. Benefits of AR Standards

The increasing growth of ICT in both hardware and software has encouraged the incubation of new AR applications with emphasis on mobile devices. The use of AR can radically change user experience and, particularly, information retrieval, as described in Section 1. However, there are still interoperability issues in the development process of AR applications. Although there are various international standards that have been developed, there is still work that needs to be done since AR requires the convergence of various technologies, which increases the interoperability challenges that must be addressed. For this reason, we are moving towards creating a *suite* of standards for using AR applications, instead of one global standard.
By investing in the establishment of robust international standards, we create a solid basis for developers, content publishers, platform and tool providers, as well as end-users, to work together. Specifically, the enhancement of existing international standards or the creation of new standards aids in the following ways:

- Provide a universally accepted sustainable development framework that assures cross-platform AR content exchange.
- Prepare a baseline for interoperability between manufacturers and content providers.
- Ease the development of AR applications.
- Assist the hardware industry in creating devices that can support AR content for various use cases, increasing their appeal to end-users.
- Reduce the investment risk by minimizing the implementation cost.
- Grant access to AR publishers and developers to currently deployed and proven emerging technologies, decreasing time to market.
- Make available standard interfaces and encodings in order to access content from various sources (e.g. user data, proprietary data, government sources).
- Assist the generation of Quality Assurance measures, in order to build better AR applications.

3. Standardization Efforts

As described in Section 2, we are moving towards the creation of a suite of standards for AR applications. To this end, it is worth noting that there has been an increase lately in the number of groups and organizations that work toward establishing frameworks or extending already existing standards related to AR. In this Section, we present some of the most influential groups and organizations that have been contributing in AR standardization.

3.1. Khronos Group activities

One of the pioneers is the *Khronos* group, a non-profit organization that is focused on creating royalty-free, open standards for authoring, accelerating and accessing visual computing on a wide variety of operating systems and devices. Khronos group and its member organizations (consisting of both business and academia) have also invested time on building interoperable APIs and formats to facilitate the development of AR applications. Specifically, focus is given on AR standards covering needs relevant to visual computing, 3D asset handling, sensor processing and browser interaction.

3.2. Open Geospatial Consortium activities and the Augmented Reality Markup Language

An attempt to develop standards that make complex spatial information and services accessible and useful to all kinds of applications is made by the *Open Geospatial Consortium (OGC)*. Related to AR, a corresponding standardization effort is the Augmented Reality Markup Language (ARML) 2.0, which provides an interchange format for Augmented Reality applications to describe and interact with objects in an AR scene, with a focus on mobile, vision-based AR.

3.3. Open Mobile Alliance activities and MobAR

AR has also intrigued the *Open Mobile Alliance (OMA)*, one of the most prevalent organizations for mobile services. OMA's mission is to set targets and to distinguish market mobile service enablers, in order to encourage worldwide users to adopt mobile data services, by ensuring that users will have service interoperability across devices, locations, providers and operators. Regarding the AR standardization efforts, OMA has invested in creating a framework that enables mobile AR services, the MobAR (Mobile Augmented Reality) v1.0. The main functionalities of the MobAR are described in Table 1.

3.4. International Organization for Standardization activities and the Augmented Reality Continuum

The *International Organization for Standardization (ISO)* has also initiated the processes for creating international standards for AR. In particular, ISO has introduced Working Group 9 (WG 9), a WG for the Augmented Reality Continuum (ARC) concepts and reference model. Another related initiative of ISO is the SC24, which is a Standards Committee responsible for producing standards for computer graphics (e.g. X3D, OpenGL), image processing and environmental data representation.

3.5. Moving Picture Experts Group activities and the Mixed and Augmented Reality standard

As part of the ISO/IEC, the *Moving Picture Experts Group (MPEG)* working group is working on the MAR (Mixed and Augmented Reality) standard, which aims at defining the main concepts, a system architecture and a set of minimum functions and performance levels for building a MAR system. Besides this, another standard developed from MPEG is the Augmented Reality Application Format (ARAF), which is a combination of other MPEG standards, for enabling real-time 2D/3D hybrid content.

3.6. World Wide Web Consortium activities and the Augmented Web

Similarly, an Augmented Reality Community Group was created by the *World Wide Web Consortium (W3C)*, in order to contribute to discussions about the intersection of AR and the Web, also known as Augmented Web. W3C standards that are embraced by the Augmented Web are HTML5, Geolocation, DeviceOrientation, DeviceMotion, WebGL, WebAudio, Media Capture & Streams and WebRTC. This group's vision is to assist the seamless integration of the disparate technologies that are developed from other relevant working groups and standards bodies.

3.7. Web3D consortium activities and the Extensible 3D format

Finally, an attempt to utilize and extend the capabilities of Extensible 3D (X3D) to support AR, with respect to other AR standards (e.g. W3C Augmented Reality Community Group, OGC ARML, ISO-SC24 and Khronos), is made by the *Web3D* Consortium. X3D is a 3D graphics web-based standard, which enables real-time communication using animation, user interaction and networking. Thus, by extending this standard, Web3D aims at providing a file format specification and architecture for representing and communicating 3D objects, events, behaviours and environments for an AR ecosystem.

3.8. Synopsis – Taxonomy

In Table 1, we present a summary of the aforementioned standardization efforts. However, since AR standardization is still an ongoing process, there are even more efforts with various topics including, but not limited to, content formats, media and audio APIs, graphics, location, hardware and communication. In order to monitor the progress across this wide variety of standards development organizations (SDO) and industry groups, an AR Community for Open and Interoperable Augmented Reality Experiences was established [14]. The AR Community's goal is to bring together all interested SDOs and communities by providing a neutral platform on which stakeholders can explore ways to coordinate their activities and at the same time establish collaborations.

TABLE 1. Standardization efforts per organization.

Organization	**Activity Name**	**Short Description**
Khronos group	Open CL, Open GL and Open GL|ES	Visual computing
	glTF and Collada	3D asset handling
	StreamInput, OpenKCam and OpenVX	Sensor processing
	WebGL and WebCL	Browser interaction
OGC	ARML 2.0	Interchange format for AR applications focusing on mobile, vision-based AR
OMA	MobAR 1.0	Functionalities: personalization and management of AR Content, user interactivity with AR content, network and client APIs, security and privacy
ISO	WG 9	Working Group for the Augmented Reality Continuum (ARC) concepts and reference model
	SC24	Standards Committee for computer graphics, image processing and environmental data representation
MPEG	MAR	Definitions of the main concepts, a system architecture and a set of minimum functions and performance levels for building a MAR system
	ARAF	Combination of other MPEG standards, for enabling real-time 2D/3D hybrid content
W3C - Augmented Reality Community Group	HTML5, Geolocation, DeviceOrientation, DeviceMotion, WebGL, WebAudio, Media Capture & Streams, WebRTC	Augmented Web, i.e. intersection of AR and the Web
Web3D	X3D	Extensible 3D modeling and scene composition

4. Conclusion

Augmented Reality is a technology that emerged during the past decades and is already part of our daily life, promising to transform user interaction. However, there is still great need for establishing internationally accepted standards in order to further strengthen AR's evolution. In this paper, we have presented the main benefits of developing standards for AR applications, as well as the key relevant activities that are carried out by international organizations for composing standards in the field of AR. It is worth noting that such standards benefit all entities across the AR value chain, including developers, content publishers, platform and tool providers, as well as end-users.

References

[14] Ronald Azuma, "*A survey of augmented reality*", Presence: Teleoperators and virtual environments, 1997.

[15] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre, "*Recent advances in augmented reality*", IEEE computer graphics and applications, 2001.

[16] Pasquale Daponte, Luca De Vito, Francesco Picariello, and Maria Riccio, "*State of the art and future developments of the Augmented Reality for measurement applications*", Measurement, 2014.

[17] "Pokemon GO", http://www.pokemongo.com/. Accessed Jan. 2017.

[18] "Layar", https://www.layar.com/. Accessed Jan. 2017.

[19] Paul Milgram, and Colquhoun Herman, "*A taxonomy of real and virtual world display integration*", Mixed reality: Merging real and virtual worlds, 1999.

[20] Thomas Caudell, and David Mizell, "*Augmented reality: An application of heads-up display technology to manual manufacturing processes*", in Proceedings of the Twenty-Fifth Hawaii International Conference, 1992.

[21] D.W.F. van Krevelen, and Ronald Poelman. "*A Survey of Augmented Reality Technologies, Applications, and Limitations*", The International Journal of Virtual Reality, 2010.

[22] "Microsoft HoloLens", https://www.microsoft.com/microsoft-hololens/en-us. Accessed Jan. 2017.

[23] "eMacula contact lenses", http://www.emacula.io/home/invest. Accessed Jan. 2017.

[24] Murat Akçayır, and Gökçe Akçayır, "*Advantages and challenges associated with augmented reality for education: A systematic review of the literature*", Educational Research Review, 2017.

[25] Chien-Yu Lin, Hua-Chen Chai, Jui-ying Wang, Chien-Jung Chen, Yu-Hung Liu, Ching-Wen Chen, Cheng-Wei Lin, and Yu-Mei Huang, "*Augmented reality in educational activities for children with disabilities*", Displays, 2016.

[26] Siti Salmi Jamalia, Mohd Fairuz Shiratuddina, Kok Wai Wonga, and Charlotte L. Oskamc, "*Utilising mobile-augmented reality for learning human anatomy*", Procedia-Social and Behavioral Sciences, 2015.

[27] "AR Standards", http://www.perey.com/ARStandards/. Accessed Jan. 2017.

Pavlos Kosmides received his Diploma degree in Information & Communication Systems Engineering in 2008 from the University of the Aegean (Samos). He is currently pursuing a Ph.D. degree in the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA). His research interests lie in the fields of distributed context management, context inference, intelligent transport systems and heterogeneous wireless networks. He is a member of the Technical Chamber of Greece.

Konstantinos Demestichas received his Dipl.-Ing., as well as his Ph.D. degree in Telecommunications, from the School of Electrical and Computer Engineering, NTUA. He is currently working as a project coordinator at the Institute of Communication and Computer Systems of NTUA. His primary research interests include data analytics and cognitive mobile communications. He has over 100 publications in the aforementioned fields, and has actively been involved both as a researcher and reviewer in several EU research projects.

Evgenia Adamopoulou received her Diploma, as well as her Ph.D. degree in Mobile Communications, from the School of Electrical and Computer Engineering of the NTUA. She is currently working as a senior research associate at the Institute of Communication and Computer Systems of NTUA. Her primary research interests include machine learning and human-machine interaction. She has over 90 publications in the aforementioned fields, and has actively been involved both as a researcher and reviewer in several EU research projects.

CFP: Special Issue for IEEE Transactions on Multimedia
Video over Future Networks: Emerging Technologies, Infrastructures, and Applications

The current Internet faces numerous challenges as a platform for media delivery: the high bandwidth demand is fueled by online video streaming services and the advent of ultra-high-definition (UHD) video; widespread use of social media fosters instant sharing of user-generated video; rising consumer interests in augmented and/or virtual reality (AR/VR) underscore the need to support richer media forms with lower latency. In the meantime, design and experimentation of the Internet itself is evolving at its own pace, as demonstrated by recent advances in software-defined networking (SDN), network function virtualization (NFV) technologies, and information-centric networking (ICN) --- technologies with potentially far-reaching impacts on the future Internet. This special issue aims to highlight research works that investigate future Internet technologies through the prism of its most prevalent application: video distribution. Intriguing research questions abound: how can named-data-networking (NDN) support live video streaming? What is the most efficient distribution mechanism for social sharing of user-generated video? What are proper performance metrics for novel networked multimedia applications based on augmented/virtual reality?

We invite submissions of high-quality papers on either original research or survey/overview, which have not been published previously. Topics of interest include, but are not limited to:

- Dynamic resource provisioning for video distribution over SDN-enabled networks
- Network-assisted rate adaptation in SDN-enabled networks
- In-network caching and caching for mobile video delivery
- Video distribution over information-centric networking (NDN) architectures
- Cost and economic models for video distribution over future networks
- Network support for emerging novel applications, e.g., based on augmented reality (AR)
- Networking and distributed systems for augmented reality (AR)
- Distribution of ultra-high-definition (UHD) video over next-generation networks
- Distribution of user-generated media content over future Internet
- Network and cloud support for real-time video analytics
- Integration of video distribution and multimedia computing

Prospective authors should submit an electronic copy of their complete manuscript to http://mc.manuscriptcentral.com/tmm-ieee. Authors are encouraged to contact guest editors for the appropriateness of their topics. Please indicate in the cover letter that the manuscript is intended for the special issue on "*Video over Future Networks: Emerging Technologies, Infrastructures, and Applications*".

Important Dates

Manuscript Due:	**January 15, 2017**	**Final Decisions:**	**June 1, 2017**
First Round of Reviews:	**March 15, 2017**	**Final Paper Due:**	**July 1, 2017**
Revision Due:	**May 1, 2017**	**Publication Date:**	**September 15, 2017**

Guest editors

Xiaoqing Zhu, xiaoqzhu@cisco.com, Cisco Systems Inc., USA
Shiwen Mao, smao@ieee.org, Auburn University, USA
Mahbub Hassan, mahbub.hassan@unsw.edu.au, University of New South Wales, Australia
Hermann Hellwagner, hellwagn@itec.uni-klu.ac.at, Klagenfurt University, Austria

MMTC Communication - Frontier CFP: Content-Driven Communications and Computing for Multimedia in Emerging Mobile Networks

Due to continuing advances in wireless communications and mobile devices, we are entering an era of rapid expansion in multimedia applications and services called as multimedia -based services, such as video streaming (Youtube, Netflix) and content sharing (Instagram, Snapchat). Obviously, these multimedia -based services are content-centric and dominant driving forces behind the expansion. Today's mobile network architectures, however, are designed to be connection-centric, which have become a barrier to meet the diverse application requirements and the quality expectation of the end users. Although the bandwidth and data rate increase, current mobile networks are still facing poor user experience and low service quality.

A paradigm shift is needed to meet the proliferation of content-centric services. On the one hand, the developments of multimedia transmission systems and services call for new understanding and evaluation of user's perceived quality of experience (QoE). There is tremendous demand for objective, online QoE prediction and monitoring for content-centric services. The large-scale data sets of online video streaming and content sharing have made it possible to reveal the true relationship between QoE and traditional metrics as well as network conditions. On the other hand, the increasing data rates, bandwidths, as well as processing capabilities of base stations and user end devices lays the foundation for efficient and massive content distribution. It is envisioned that content-driven communications and computing technologies will break the bottleneck of current connection-centric network architectures, and lead to a clean-slate redesign of network architecture.

Topics of interest include, but are not limited to:

Content-driven heterogeneous multimedia networks architecture
Content-driven software-defined multimedia networks architecture
Content-driven communications technologies in wireless multimedia networks
Content-driven computing technologies in mobile multimedia networks
Content-driven caching for multimedia in wireless networks
QoE-aware content delivery for multimedia in mobile networks
QoE-aware 5G multimedia network architecture/techniques and performance analysis
QoS-QoE relationship modeling
Real experiments and testbeds for QoE evaluation
Efficient video codec design with guaranteed QoE

GUEST EDITORS

Tao Jiang
Huazhong University of Science and Technology, P. R. China

Wei Wang
Huazhong University of Science and Technology, P. R. China

Cheng Long
Queen's University Belfast, UK