A Framework for Adaptive Delivery of Omnidirectional Video

Christian Timmerer; Alpen-Adria-Universität Klagenfurt / Bitmovin; Klagenfurt am Wörthersee, Austria
Ali C. Begen; Ozyegin University / Networked Media; Istanbul, Turkey

Abstract

Omnidirectional or 360-degree videos are considered as a next step towards a truly immersive media experience. Such videos allow the user to change her/his viewing direction while consuming the video. The download-and-play paradigm (including DVD and Blu-ray) is replaced by streaming, and the content is hosted solely within the cloud. This paper addresses the need for a scientific framework enabling the adaptive delivery of omnidirectional video within heterogeneous environments. We consider the state-of-the-art techniques for adaptive streaming over HTTP and extend them towards omnidirectional/360-degree videos. In particular, we review the encoding and adaptive streaming options, and present preliminary results reported in the literature. Finally, we provide an overview about the ongoing standardization efforts and highlight the major open issues.

Introduction

Universal media access (UMA) [1] as proposed in the late 90s, early 2000s is now reality. It is very easy to generate, distribute, share, and consume any media content, anywhere, anytime, and with/on any device. These kinds of real-time entertainment services – specifically, streaming audio and video – are typically deployed over the open, unmanaged internet and account for the the majority of the internet traffic [2]. A major technical breakthrough and enabler was certainly the HTTP adaptive streaming (HAS) resulting in the standardization of MPEG Dynamic Adaptive Streaming over HTTP (DASH) [3].

In the past we have seen multiple approaches enabling a truly immersive media experience, and recently we have seen an increasing number of proposals related to virtual reality (VR) applications and, specifically, omnidirectional (360-degree) video delivery. Current deployments focus on adopting traditional HAS techniques and, thus, build on top of the existing adaptive streaming ecosystems [4].

Omnidirectional or 360-degree video allows the user to change her/his viewing direction while consuming the video, resulting in a more immersive experience than consuming traditional video content with a fixed viewing direction. Such video content can be created and consumed using a plethora of devices making it attractive for various use cases. Creation devices are available at different price ranges targeting various industries and consumption devices range from smartphones and desktop computers to head-mounted displays (HMD). When using an HMD to watch such content, the viewing direction can be changed by head movements. On smartphones and tablets, the viewing direction can be changed by touch interaction or by moving the device around thanks to built-in sensors. On a desktop computer, the mouse or keyboard can be used for interacting with the omnidirectional video.

In this paper, we present a framework for the adaptive delivery of omnidirectional video over the internet considering the recent efforts in both standardization and research. The major goal is to reduce storage and bandwidth requirements and – at the same time – to enable dynamic adaptation towards *(i)* network conditions (e.g., available bandwidth), *(ii)* device characteristics (e.g., HMD, smart phone, TV/STB), and *(iii)* user interactivity (e.g., selected viewport). Based on these findings we discuss adaptive delivery options with the aim to transmit the best possible content configuration under the given conditions while preserving the users' Quality of Experience (QoE). We discuss several options including pros and cons to be used in practical deployments and provide an overview about current adaptive streaming challenges in this domain. Finally, we highlight the ongoing standardization efforts in this context, specifically MPEG-I and the omnidirectional media format (OMAF) as well as its binding to MPEG-DASH.

Framework Architecture

The the building blocks of the proposed framework are depicted in Figure 1. The *content creation* comprises capture and acquisition, which are typically achieved through multiple sensors (*e.g.*, multiple cameras including microphones specifically arranged within a predefined rack), and includes various editing and processing steps such as stitching or region-wise packing in order to accommodate a given projection format. Various projection formats are proposed and, in practice, equirectangular projection (ERP) and cubemap projection (CMP) are currently supported. The functionality of the *server* typically includes *(i)* encoding (*i.e.*, reducing the bitrate for bandwidth efficient streaming while preserving quality), *(ii)* encryption (*i.e.*, enabling digital rights management), and *(iii)* encapsulation (*i.e.*, providing flexible access to the media data) in order for the subsequent *network* component to enable *(a)* storage, *(b)* delivery, and *(c)* distribution. Finally, the *client* component enables the consumption of omnidirectional content that eventually leads to an immersive media experiences. This component typically comprises multimedia systems aspects such as decapsulation, decryption, and decoding as well as the actual rendering and means for interactivity, which may also include various processing steps.

The focus of this paper is on encoding and delivery, specifically adaptive streaming. However, we also address certain client aspects as some of them are related to encoding and streaming. An overview of virtual reality and MPEG-DASH is also provided in [5]. The existing adaptive streaming options can be roughly clustered into three categories, namely: *(1)* traditional, viewport-agnostic streaming, which adopts the state-of-the-art techniques for omnidirectional content without further optimization; *(2)* viewport-adaptive streaming [6], which provides multiple versions (*i.e.*, additional representations) for predefined viewports allowing for more efficient encoding of the selected view-

port; and *(3)* tile-based streaming [7, 8], which uses the tiling technique of modern video codecs to efficiently address storage and bandwidth requirements. The choice of streaming approach has a direct impact on the encoding strategy and its options. Tile-based streaming is currently considered as the most advanced and promising approach and, thus, we focus on tile-based streaming in the remainder of this paper.

Encoding Options for Omnidirectional HAS

According to a recent video developer report [9], AVC is still the dominant video codec, but others (i.e., HEVC, VP9, AV1) are also used or considered to be used in the near future. Additionally, these other codecs support tiles, which is a prerequisite to implement efficient omnidirectional HAS solutions. Tiles (*e.g.*, in HEVC [10]) divide a picture into independent, rectangular regions that allow for better parallel processing, improvements with respect to delivery in terms of maximum transfer unit, and region-of-interest (RoI) based coding. In particular, the latter is an important feature for tile-based streaming as it allows to deliver tiles within the user's current viewport in higher quality compared to those tiles outside the user's current viewport.

A drawback of tiling, however, is the reduced coding efficiency as the picture is divided into smaller regions, which are encoded independently and, thus, only small regions can be considered for intra prediction or motion compensation. On the other hand, the more tiles the more flexibility and increased granularity for a given viewport. Therefore, a tradeoff between bitrate, quality, and flexibility is needed. In general, we can cluster the existing approaches into those using uniform [7] and non-uniform tiling [11]. The former uses equally-sized tiles across the entire picture whereas the latter adopts larger tiles around the equator and smaller at the poles. Interestingly, both approaches achieve similar results with only minor differences.

Figure 2 depicts the tiling overhead for different uniform tiling configurations. As expected, the coding efficiency decreases with an increasing number of tiles. The more tiles the higher the granularity that allows for a better match with the user's viewport. Further details about the encoding configuration and evaluation methodology can be found in [7].

Another differentiation can be made based on tiles having all the same versus mixed resolutions (*i.e.*, some tiles have higher whereas others have lower resolution). Using mixed resolutions reportedly shows better performance [12]. In particular, it adopts a non-uniform tiling approach, which divides the picture into tiles representing foreground, background, top, and bottom. Foreground and background cover larger areas than top and bottom. Additionally, background, top, and bottom are encoded at lower resolution but higher signal-to-noise ratio (SNR) quality than foreground tiles which provides 10% bitrate reduction compared to when encoding background/top/bottom at the same resolution but lower SNR quality. Tiling can be even further optimized as suggested by Xiao *et al.* [13] that estimates per-tile storage costs and solves an integer linear program (ILP) to determine an optimal, potentially non-uniform tiling. Results show that this approach reduces bandwidth requirements up to 73% compared to adaptive streaming without tiling and up to 44% compared to uniform tiling methods.

A number of papers also look into objective and subjective evaluations. An objective evaluation framework is proposed in [14], which is suitable for various projection formats (*i.e.*, equirectangular projection (ERP), cubemap projection (CMP), equal-area projection (EAP), octahedron projection (OHP)) and also suggests area weighted spherical PSNR (AW-SPSNR) as a new metric. Subjective evaluation and consequently QoE evaluations are rare in the literature, but first steps have been made. Xu *et al.* [15] propose a subjective visual quality assessment (VQA) methods for panoramic videos. Additionally, they propose overall and vectorized Differential Mean Opinion Score (DMOS) metric to quantify the overall and regional quality reduction of impaired panoramic videos, respectively. Interestingly, the subjective studies reveal that subjects consistently prefer looking at the center of the front region of panoramic videos, which is refuted by Rai *et al.* [16] as they found that it is between 14 and 20 visual degrees from the center of the viewport. Another subjective evaluation has been conducted by Singla *et al.* [17], which evaluates UHD and HD using HMD at different bitrates and reveals that 8 MBit/s is sufficient for UHD content as higher bitrates do not necessarily increase QoE.

Finally, publicly available datasets are needed to enable reproducible research. For example, Xu *et al.* [15] also provides with a dataset whereas others primarily focus on the actual dataset such as the ones from Corbillon *et al.* [18] and Rai *et al.* [19]. The latter also includes accurate eye-tracking data allowing for more detailed evaluations than those only focusing on head/device movements.

Omnidirectional HAS

In this section, we discuss omnidirectional HAS. As stated previously, the adaptive streaming of omnidirectional content can be clustered into three categories, namely: *(1)* traditional, viewport-agnostic streaming; *(2)* viewport-adaptive streaming; and *(3)* tile-based streaming, which are examined below in more detail.

Overview of the Approaches

Traditional, viewport-agnostic streaming currently assumes that the omnidirectional content is available using the equirectangular projection (or cubemap projection) and encoded with existing, state-of-the-art HAS workflows. That is, the entire 360-degree video is encoded and streamed as if it is a traditional video. At the client, the rendering is performed by special libraries that transform the content from the projection format to the spherical domain and enable its consumption. While this is a simple and straightforward approach, bandwidth is wasted as the content, which cannot be consumed by the end user (*e.g.*, what happens outside the viewport), is still streamed, decoded and rendered on the device mostly at the same quality as the content within the visible area (viewport). Obviously, this is a suboptimal approach, but it surprisingly works in practice.

Viewport-adaptive streaming [6] adopts the basic principles of HAS and extends them to omnidirectional content by simply adding additional representations corresponding to predefined viewports. These representations are added on top of those used for different resolutions and bitrates and, thus, significantly increase the total number of representations available for HAS (*e.g.*, with 10 representations from traditional HAS and a viewport-specific representation for every 30° would result in 120 representations). However, the viewport-specific representations are
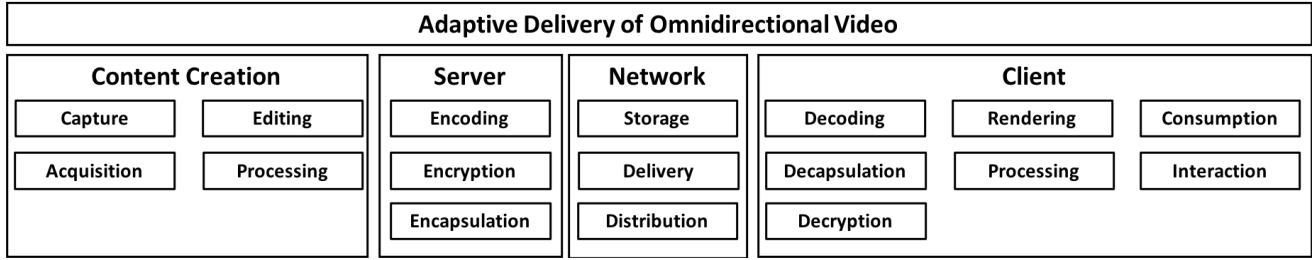
Figure 1. Framework functional building blocks from content creation to consumption.
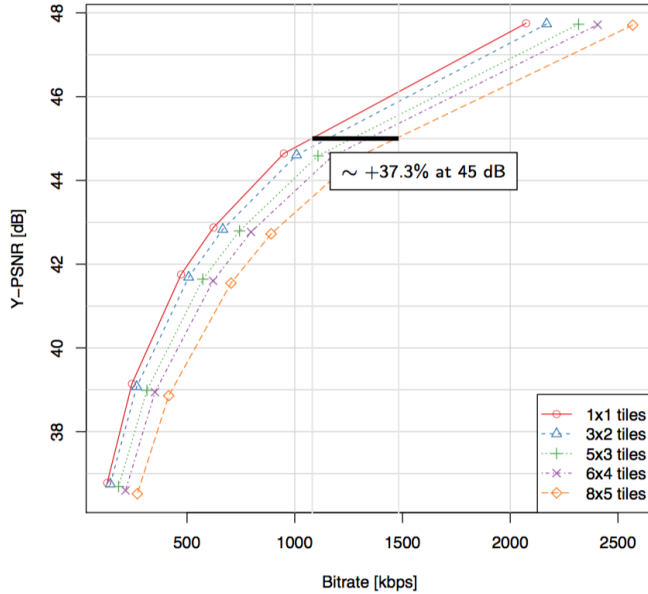

Figure 2. Tiling overhead for sequence 'Assassins Creed' at resolution 1920x960 (adopted from [7]).

encoded with a given viewport in mind and, thus, have typically higher quality around the center of the viewport and lower quality outside the viewport, utilizing an RoI-based coding technique or projection format. This representation is requested based on the user's viewport (*i.e.*, feedback from the device's orientation sensors) and in case of user interactivity (*e.g.*, looking right, left, back) the user is still able to consume the content, but at a lower quality. Consequently, a different viewport representation is selected and delivered towards the client. In general, this approach is also straightforward to implement as shown in [20] using a pyramid projection format, but also suffers from the increased storage and content delivery network (CDN) requirements (due to the additional representations added during content provisioning). Additionally, it offers limited flexibility, which is defined by the number of viewports provided. The more viewports provided the higher the flexibility, but also the higher the overhead with respect to storage and CDN requirements. The bandwidth requirements for the actual streaming session are reduced compared to traditional, viewport-agnostic streaming, but viewport-adaptive streaming provides higher quality for a given viewport.

Finally, **tile-based streaming** adopts the tiling technique of modern video codecs as described in the previous section [7, 8]. It allows the adaptive streaming client to request individual tiles

from different quality representations or no tiles at all. For example, tiles within the current viewport can be requested in the highest possible quality, neighboring, adjacent tiles with gradually lower quality, and other tiles being totally outside of the user's viewport not at all. It is noted that different strategies will lead to different results. In case the client decides not to request certain tiles, bandwidth requirements can be reduced dramatically, but this strategy also imposes the risk of a blank screen when a user suddenly decides changing his viewport. Thus, a tradeoff is required to determine which tiles should be requested in which quality based on the given conditions (*e.g.*, available bandwidth).

Figure 3 shows viewport-PSNR (V-PSNR) results from [7] showing that tile-based streaming could increase the displayed quality significantly compared to other approaches. In particular, *monolithic* refers to traditional, non-tiled streaming with different segment sizes (1s, 2s, 4s) and basically shows the coding efficiency for these different segments sizes. Interestingly, *tiles monolithic* adopts a similar streaming strategy as monolithic streaming but uses tiles (*i.e.*, all tiles in the same quality), which expectedly shows lower coding efficiency due to the tiling overhead. *Tiles with full delivery basic* refers to the case where the tiles within the user's viewport are streamed in the highest possible quality and all other tiles are streamed in the lowest possible quality. It shows that quality could be more than 40% better than for non-tiled streaming. Further details about the setup and methodology can be found in [7].

Adaptive Streaming Challenges

Increased number of HTTP requests. Dividing pictures into tiles and allowing the client to request them individually from different representations based on the client's context (*e.g.*, orientation, bandwidth) may increase the number of HTTP requests significantly. One approach to reduce the number of requests is adopting HTTP/2 server push in combination with specific request parameters describing which tiles from which representation should be delivered to the client for a single request [21]. However, this would break a fundamental assumption of HAS as it requires additional server-side functionality as opposed to simply serving HTTP requests.

Low latency streaming. Latency in HAS has been always an issue not only for live streaming but even more for omnidirectional streaming due to increased user interactivity. An obvious approach is reducing the segment length, but this leads to reduced coding efficiency which is already an issue due to tiling. An evaluation regarding the overhead and performance is provided by Bouzakaria *et al.* [22], which utilizes HTTP chunked transfer encoding and Gradual Decoding Refresh (GDR) encoding in com-
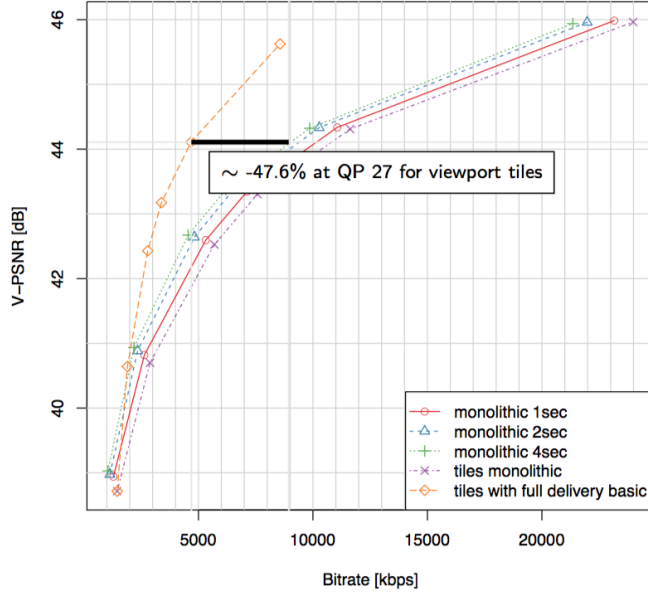
Figure 3. *Monolithic streaming compared to tile-based streaming: 'Explore The World', 6x4 tiles, resolution: 3840x1920, including head movement (adopted from [7]).*

Figure 4. *Overview of the standards development organizations currently active in the VR space (adopted from [29]).*

bination with the ISO base media file format. Results indicate that latency can be reduced to 240ms within local area networks but, in practice, one has to consider CDNs and larger-scale deployments. Recently, the MPEG common media application format (CMAF) has been ratified [23], which enables requesting chunks of segments for reduced latency. The effect of CMAF chunks on omnidirectional content has yet to be studied.

Viewport prediction. A natural improvement for the adaptive delivery of omnidirectional video would be viewport prediction based on the previously recorded user interactions (*i.e.*, head movements, eye tracking, *etc.*). This could be useful in the context of prefetching, caching, and content distribution in general. Unfortunately, current viewport prediction methods are limited in terms of how much they can predict into the future, which again is related to segments size, buffering at the client, and latency in general. However, it is expected that artificial intelligence (AI)-based approaches such as machine learning can be used for such purposes as soon as sufficient data is available. An example approach is described in [24].

Quality of experience. The QoE of adaptive delivery of omnidirectional video is mostly in its infancy, but the first approaches and results have been reported in the literature. The majority of papers target the encoding options for omnidirectional HAS as described in the previous section. One of the first papers addressing the QoE of omnidirectional HAS is provided by Schatz *et al.* [25], which specifically targets the impact of stalling events when consuming 360-degree content on HMDs and traditional TV sets. The main findings are that such kind of assessment is not trivial leading to more open issues than actual recommendations. In general, the expectation would be similar as for the traditional HAS, *i.e.*, no stalls at all (if possible).

Recently, some interesting approaches for improving omnidirectional HAS have been proposed which we would like to re-
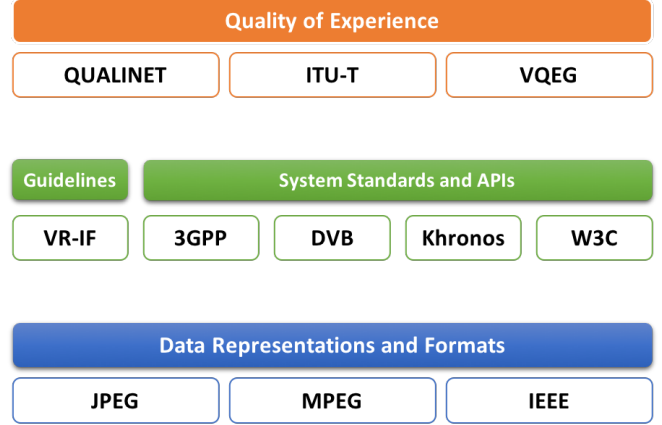
view briefly. Corbillon *et al.* [26] describe an approach to determine the optimal set of 360-degree videos for viewport-adaptive streaming based on so-called quality emphasized regions (QERs), which allow for up to 45% bandwidth savings. Nasrabadi *et al.* [27] suggest using scalable video coding (SVC) for tile-based omnidirectional HAS, which addresses rebuffering issues and shows minor bandwidth savings compared to non-scalable tile-based streaming. Additionally, SVC allows for a more efficient in-network caching which could help reducing the distribution costs. Finally, Xie *et al.* [28] propose another tile-based approach, which uses user's head movements as probability events to prefetch tiles. They propose a probabilistic model of viewport to calculate the viewing probability of tiles and then formulate an optimization problem, which minimizes the quality distortion and spatial variability of quality (based on PSNR) given the total transmission bitrate budget. Results show approximately 39% gain on viewport PSNR and about 46% reduction on spatial quality variance compared to the existing approaches.

Overview of the Standardization Efforts

An overview of the standardization efforts is described in [29] and further detailed for MPEG OMAF [30] and video coding in [31]. Figure 4 depicts an overview of standards development organizations (SDOs) (including support organizations) working in the VR space. We can group them into three categories as follows (bottom-up): *(1)* Data representation and formats providing basic tools to be adopted directly by the industry or referenced by other SDOs (or parts thereof); *(2)* guidelines, system standards, and application programming interfaces (APIs) typically providing so-called system specifications including end-to-end aspects *(3)* QoE addressing the perceived quality as experienced by the end users of such applications and services. At the time of writing of this paper, each SDO has been at different stages and up to date information can be found in [32].

The MPEG OMAF [30] is one of the first standardized formats available in this domain. The scope of OMAF comprises the *(i)* coding, *(ii)* storage, *(iii)* delivery, and *(iv)* rendering of omnidirectional media, including video, images, audio, and timed text. It defines *(a)* a coordinate system; *(b)* projection (currently ERP and CMP) and rectangular region-wise packing methods; *(c)* stor-

age of omnidirectional media and the associated metadata using the ISO base media file format; *(d)* encapsulation, signaling, and streaming of omnidirectional media (*e.g.*, with DASH or MMT); and *(e)* media profiles and presentation profiles that provide interoperable and conformance points. Further details can be found in [31, 33].

Conclusions

The tremendous growth of research articles in the area of omnidirectional or 360-degree videos calls for a scientific framework identifying the buildings blocks. In this paper, we proposed such a framework and discussed encoding and streaming in more detail. We showed that novel approaches such as tile-based encoding and streaming can significantly reduce bandwidth compared to traditional streaming. Additionally, we highlighted the state-of-the-art in this domain (traditional, viewport-adaptive, tile-based streaming) including their challenges.

Future work will certainly focus on low-latency streaming (possibly in connection with the latency promise of future 5G networks), viewport prediction (based on machine/deep learning), QoE (datasets, methodology, metrics), and combinations thereof.

Acknowledgments

References

[1] R. Mohan, J. R. Smith, and Chung-Sheng Li, Adapting Multimedia Internet Content for Universal Access, IEEE Transactions on Multimedia, 1, 1. (1999)

[2] Sandvine. 2016 Global Internet Phenomena Report: Latin America & North America (2016). Online: `http://sandvine.com/`.

[3] I. Sodagar, The MPEG-DASH Standard for Multimedia Streaming Over the Internet, IEEE MultiMedia, 18, 4. (2011)

[4] C. Timmerer, M. Graf, C. Mueller, Adaptive Streaming of VR/360-degree Immersive Media Services with high QoE, 2017 NAB Broadcast Engineering and IT Conference (BEITC), Las Vegas, NV, USA. (2017)

[5] D. Podborski, E. Thomas, M.M. Hannuksela, S. Oh, T. Stockhammer, and S. Pham, Virtual Reality and DASH, IBC Content Delivery - technical paper, Amsterdam, The Netherlands. (2017)

[6] X. Corbillon, G. Simon, A. Devlic and J. Chakareski, Viewport-Adaptive Navigable 360-degree Video Delivery, 2017 IEEE International Conference on Communications (ICC), Paris, France. (2017)

[7] M. Graf, C. Timmerer, and C. Mueller, Towards Bandwidth Efficient Adaptive Streaming of Omnidirectional Video over HTTP: Design, Implementation, and Evaluation, ACM Multimedia Systems (MMSys), Taipei, Taiwan. (2017)

[8] C. Concolato, J. Le Feuvre, F. Denoual, E. Nassor, N. Ouedraogo and J. Taquet, Adaptive Streaming of HEVC Tiled Videos using MPEG-DASH, IEEE Transactions on Circuits and Systems for Video Technology, PP, 1. (2017)

[9] Bitmovin, Inc., 2017 Video Developer Report. (2017) Available: `https://bitmovin.com/2017-video-developer-report/`

[10] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth and M. Zhou, An Overview of Tiles in HEVC, IEEE Journal of Selected Topics in Signal Processing, 7, 6, pg. 969. (2013)

[11] A. Zare, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, HEVC-compliant Tile-based Streaming of Panoramic Video for Virtual Reality Applications, ACM on Multimedia Conference (MM), Amsterdam, The Netherlands. (2016)

[12] I. D.D. Curcio, H. Toukomaa, and D. Naik, Bandwidth Reduction of Omnidirectional Viewport-Dependent Video Streaming via Subjective Quality Assessment, Workshop on Multimedia Alternate Realities (AltMM), Mountain View, CA, USA. (2017)

[13] M. Xiao, C. Zhou, Y. Liu, and S. Chen, OpTile: Toward Optimal Tiling in 360-degree Video Streaming, ACM on Multimedia Conference (MM), Mountain View, CA, USA. (2017)

[14] X. Xiu, Y. He, Y. Ye, B. Vishwanath, An Evaluation Framework for 360-degree Video Compression, IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA. (2017)

[15] M. Xu, C. Li, Y. Liu, X. Deng and J. Lu, A subjective visual quality assessment method of panoramic videos, IEEE International Conference on Multimedia and Expo (ICME), Hong Kong. (2017)

[16] Y. Rai, P. Le Callet and P. Guillotel, Which saliency weighting for omni directional image quality assessment?, International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany. (2017)

[17] A. Singla, S. Fremerey, W. Robitza, P. Lebreton, and A. Raake, Comparison of Subjective Quality Evaluation for HEVC Encoded Omnidirectional Videos at Different Bit-rates for UHD and FHD Resolution, Thematic Workshops of ACM Multimedia 2017, Mountain View, CA, USA. (2017)

[18] X. Corbillon, F. De Simone, and G. Simon, 360-Degree Video Head Movement Dataset, ACM on Multimedia Systems Conference (MMSys), Taipei, Taiwan. (2017)

[19] Y. Rai, J. Gutiérrez, and P. Le Callet, A Dataset of Head and Eye Movements for 360 Degree Images, ACM on Multimedia Systems Conference (MMSys), Taipei, Taiwan. (2017)

[20] E. Kuzyakov and D. Pio, Next-generation video encoding techniques for 360 video and VR, Blog post. (2016). Online: `https://code.facebook.com/posts/637561796428084`

[21] S. Petrangeli, V. Swaminathan, M. Hosseini, and F. De Turck, An HTTP/2-Based Adaptive Streaming Framework for 360° Virtual Reality Videos, ACM on Multimedia Conference (MM), Mountain View, CA, USA. (2017)

[22] N. Bouzakaria, C. Concolato and J. Le Feuvre, Overhead and Performance of Low Latency Live Streaming using MPEG-DASH, International Conference on Information, Intelligence, Systems and Applications (IISA), Chania, Greece. (2014)

[23] K. Hughes, D. Singer, (eds.), Information technology – Multimedia application format (MPEG-A) – Part 19: Common media application format (CMAF) for segmented media, ISO/IEC 23000-19. (2017)

[24] C.-L. Fan, J. Lee, W.-C. Lo, C.-Y. Huang, K.-T. Chen, and C.-H. Hsu, Fixation Prediction for 360° Video Streaming in Head-Mounted Virtual Reality, Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV), Taipei, Taiwan. (2017)

[25] R. Schatz, A. Sackl, C. Timmerer and B. Gardlo, Towards Subjective Quality of Experience Assessment for Omnidirectional Video Streaming, Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany. (2017)

[26] X.Corbillon, A. Devlic, G. Simon, and J. Chakareski, Optimal Set of 360-Degree Videos for Viewport-Adaptive Streaming, ACM on Multimedia Conference (MM), Mountain View, CA, USA. (2017)

[27] A. T. Nasrabadi, A. Mahzari, J. D. Beshay, and R. Prakash, Adaptive

360-Degree Video Streaming using Scalable Video Coding, ACM on Multimedia Conference (MM17), Mountain View, CA, USA. (2017)

[28] L. Xie, Z. Xu, Y. Ban, X. Zhang, and Z. Guo, 360ProbDASH: Improving QoE of 360 Video Streaming Using Tile-based HTTP Adaptive Streaming, ACM on Multimedia Conference (MM), Mountain View, CA, USA. (2017)

[29] C. Timmerer, Immersive Media Delivery: Overview of Ongoing Standardization Activities, IEEE Communications Standards Magazine, 1, 4, pg. 71-74. (2017)

[30] B. Choi, Y.-K. Wang, M. M. Hannuksela, Y. Lim, A. Murtaza, (eds.), Information Technology – Coded Representation of Immersive Media (MPEG-I) – Part 2: Omnidirectional Media Format, ISO/IEC 23090-2 Final Draft International Standard (FDIS). (2017)

[31] R. Skupin, Y. Sanchez, Y.-K. Wang, M. M. Hannuksela, J. Boyce, M. Wien, Status of 360 degree Video Coding and Delivery, IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA. (2017)

[32] C. Timmerer, VR/360 Streaming Standardization-Related Activities, 2nd version, Blog post. (2017) Online: `https://multimediacommunication.blogspot.com/2017/04/vr360-streaming-standardization-related.html`

[33] Omnidirectional Media Format, web site & slide deck. (2018) Online: `https://mpeg.chiariglione.org/standards/mpeg-i/omnidirectional-media-format`

Author Biography

Christian Timmerer is an Associate Professor with Alpen-Adria-Universität Klagenfurt, Austria, and his research focus is on immersive multimedia communication, streaming, adaptation, and quality of experience. He has authored over 190 publications in his research area and was the General Chair of WIAMIS 2008, QoMEX 2013, ACM MMSys 2016, and PV 2018. He participated in several EC-funded projects, notably, DANAE, ENTHRONE, P2P-Next, ALICANTE, SocialSensor, and the COST Action IC1003 QUALINET. He also participated in ISO/MPEG work for several years, notably, in the areas of MPEG-21, MPEG-M, MPEG-V, and MPEG-DASH. He is a Co-Founder of Bitmovin and CIO and the Head of Research and Standardization. Find out more about him at http://blog.timmerer.com.

Ali C. Begen recently joined the computer science department at Ozyegin University, Turkey. Previously, he was a research and development engineer at Cisco, where he designed and developed algorithms, protocols, products and solutions in the service provider and enterprise video domains. Currently, in addition to teaching and research, he provides consulting services to industrial, legal, and academic institutions through Networked Media, a company he co-founded. Begen has a PhD in electrical and computer engineering from Georgia Tech. He received a number of scholarly and industry awards, and he has editorial positions in prestigious magazines and journals in the field. He is a senior member of both the IEEE and ACM. In January 2016, he was elected distinguished lecturer by the IEEE Communications Society. Further information on his projects, publications, talks, and teaching, standards and professional activities can be found at http://ali.begen.net.