

Special Paper

방송공학회논문지 제25권 제7호, 2020년 12월 (JBE Vol. 25, No. 7, December 2020)

<https://doi.org/10.5909/JBE.2020.25.7.1052>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

Configuration of Supplemental Tile Sets based on Prediction of Viewport Direction for Tile-based VR Video Streaming

Eun-bin An^{a)}, A-young Kim^{a)}, and Kwang-deok Seo^{a)‡}

Abstract

As the market demand for immersive media increases, an efficient streaming method is required in consideration of network conditions while maintaining the user's immersive experience. Accordingly, transmitting a viewport with relatively high-quality, such as tile-based streaming, is mainly used. But there still remains a lot of technical challenges, such as quickly providing a new viewport in high-quality according to the gaze. To solve the aforementioned problem, in this paper, we propose a method of configuring and transmitting a supplemental tile set through the predicted direction, and a range of stable utilization of the transmitted supplemental tile set.

Keywords : VR video streaming, tile-based streaming, prediction of viewpoint direction

1. Introduction

Augmented Reality (AR) and Virtual Reality (VR), kinds of immersive media, have been greatly growing in the market worldwide. In particular, it has been facilitated by the launch and development of various devices such as smart-phone-based VR head-mounted displays (HMD) and stand-alone HMDs. Needless to say, VR video is now streamed

to the user's device over wired/wireless networks just like the ordinary video streaming service. However, since VR video is usually much larger than 4K/8K video in resolution, the limited bandwidth causes a big problem for immersive experience when the VR contents are streamed. Therefore, various techniques have been studied in order to deliver VR video efficiently in terms of bandwidth utilization^[1].

Dynamic Adaptive Streaming over HTTP (DASH)^[2], like the MPEG Media Transport (MMT)^[3], is designed to transport large volumes and multiple media. Therefore, it is a suitable technology for the streaming of VR video. MPEG-DASH defines Media Presentation Description (MPD) and segment format. Also, the adaptive bitrate streaming can be used with MPEG-DASH to divide one multimedia file into one or more segments and transport

a) Division of Computer and Telecommunications Engineering, Yonsei University

‡ Corresponding Author : Kwang-deok Seo

E-mail: kdseo@yonsei.ac.kr

Tel: +82-33-760-2788

ORCID: <http://orcid.org/0000-0001-5823-2857>

※ This research was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1D1A1B0747065).

· Manuscript received October 12, 2020; Revised November 30, 2020; Accepted November 30, 2020.

them. Spatial Representation Description (SRD)^[4], which is supported by MPEG-DASH, extends MPD to describe spatial relationship information. It makes enable to stream only spatial sub-parts of a video to display on various devices.

Viewport adaptive streaming (VAS)^[5], which is usually combined with MPEG-DASH SRD^[6,7], is an efficient streaming method using characteristic of VR video contents. VAS has been studied in several streaming approaches, such as tile-based streaming^[8,9,10] and Quality Emphasized Region (QER) based streaming^[11]. The tile-based VAS approach adopts the tile structure and motion constrained tile sets (MCTS), which are described in high efficiency video coding (HEVC)^[12], to divide the video into multiple tiles and encode them in various qualities. And then, the tiles corresponding to the viewport are streamed with high-quality and other tiles with low-quality.

With the advent of the tile-based streaming method, it is possible to stream VR video more efficiently than the conventional approaches. However, many challenges are still remaining to stream the VR video while maintaining the immersive experience. Predicting where the user's gaze will be stay, particularly, is an important problem. Many researchers are studying to implement the efficient and accurate gaze prediction algorithms^[13,14]. A method of viewport prediction based on sound localization information introduced by Jeong et al.^[15], and Zou et al.^[16] proposed a method to improve transmission efficiency of 360-degree video by calculating the gaze probability of each tile based on CNN model for viewpoint prediction. Furthermore, Feng et al.^[17] proposed the viewport prediction using content-based motion tracking and dynamic user interest, and the proposed method was considered in the live mobile video streaming for 360-degree video, especially. Using a long short-term memory (LSTM) encoder-decoder network, recently, Jamali et al.^[18] presented to predicts future viewpoint positions of up to 4 seconds.

If the tile-based streaming system is implemented based

on MPEG-DASH, the quality switching between tiles in the current segment according to the gaze movement are not flexible until the segment is completely consumed. In other words, even if the quality switching between tiles occurs as a result of the viewport position change, the client should wait until the current segment is fully consumed for the requested high-quality tile set. The delay caused by the above mentioned issue could increase motion-to-photon (MTP) latency, and lead to a fatal problem to disturb immersive experience^[19,20].

In this paper, we propose a method to predict the direction where the viewport will be located and to transmit the supplemental high-quality tile set corresponding to the predicted direction. Furthermore, we analyze the movable range of the viewport and the utility of the proposed additionally transmitted tiles.

II. Supplemental Tile Set

1. Prediction of the gaze direction

Since Quality of Experience (QoE) is a significant factor for high-quality immersive media experience, the guarantee of Quality of Experience (QoE) is a significant factor in the consumption of the immersive media. Furthermore, the MTP latency, which means a certain time gap between the user's movement and the triggered change of display during the process of using HMD, also affects to the QoE when streaming the immersive media. For the tile-based streaming, the client should be immediately receiving the high-quality tile set at the same time as the viewport position changes while keeping the QoE. However, it is physically impossible to achieve such coincidence due to network problem and limited device performance, etc. Therefore, we have been studying predicting the direction of the viewpoint as a way to compensate for this problem.

As shown in Figure 1, the next position of the viewpoint

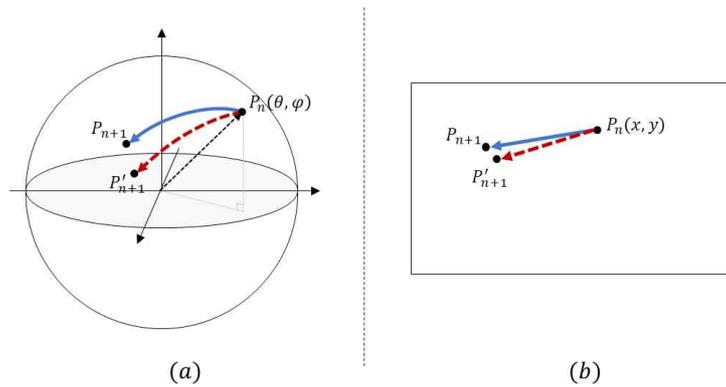


Fig. 1. Gaze prediction in (a) virtual sphere and (b) equirectangular

is predicted by one or more previous positions. In order to determine the exact prediction, it is necessary to collect the previous positions over a period of time. Figure 1(a) illustrates the position $P_n(\theta, \varphi)$ and $P_{n+1}(\theta, \varphi)$ at the moment n and $n+1$. In addition, a dotted line represents the predicted movement of viewport from $P_n(\theta, \varphi)$ to $P'_{n+1}(\theta, \varphi)$ in virtual sphere space, whereas the solid line means the actual trace of viewport movement from $P_n(\theta, \varphi)$ to $P_{n+1}(\theta, \varphi)$. Figure 1(b) shows the projected equirectangular image from the sphere in Figure 1(a).

When the error E formulated from Equation (1) converges to 0, the predicted viewpoint $P'_{n+1}(\theta, \varphi)$ becomes an accurate prediction. $p'_{n+1}(\theta, \varphi)$ can be predicted by the prediction function $pred()$ at time n using previous view-

point data obtained from P_1 to P_{n-1} .

$$E = |P' - P| \tag{1}$$

$$P'_{n+1} = Pred(P_1 : P_{n-1}) \tag{2}$$

It takes a lot of cost to predict the future point P'_{n+1} as an accurate value. On the other hand, it requires relatively low calculation cost to simply predict the direction to the next viewpoint using the previous viewpoint.

Figure 2(a) shows the moment when P'_{n+1} is predicted by the previous viewpoint P_{n-1} and the current viewpoint P_n . At this time, the event is generated periodically when the viewpoint changes, and occurs in a very short time. If the P is assumed to be generated for each event, the pre-

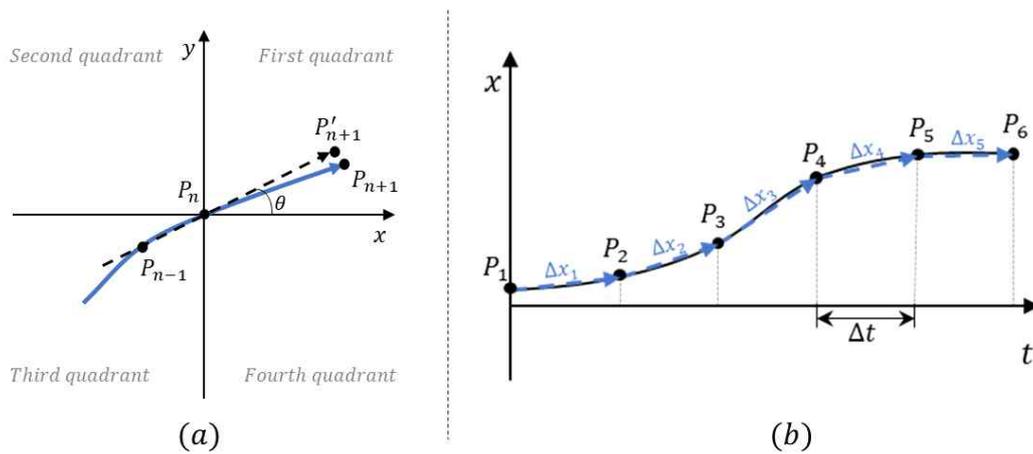


Fig. 2. Movement of viewpoint on (a) equirectangular plane and (b) x-domain

dicted viewpoint P'_{n+1} is similar to the next actual viewpoint P_{n+1} . In other words, a movement from P_{n-1} to P_{n+1} can be seen as a linear movement at a constant speed. Figure 2(b) illustrates how the viewpoint position changes in \mathcal{X} domain. If the event occurrence period Δt is short enough, the distance between adjacent viewpoint positions does not differ significantly.

It can be denoted as Equation (3), where $\Delta P_n = P_n - P_{n-1}$ and Δt is the period of the event occurrence. Equation (4) is to obtain the direction to the next viewpoint from the current viewpoint. As a result, the quadrant Q which is out of the four quadrants can be determined based on the current viewpoint position through Equation (5).

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta p_n}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta p_{n+1}}{\Delta t} \quad (3)$$

$$\theta = \arctan\left(\frac{\Delta y}{\Delta x}\right) \quad (4)$$

$$Q = \begin{cases} 1, & \text{if } x > 0 \text{ and } y > 0 \\ 2, & \text{else if } x < 0 \text{ and } y > 0 \\ 3, & \text{else if } x < 0 \text{ and } y < 0 \\ 4, & \text{otherwise} \end{cases} \quad (5)$$

2. Configuration of Supplemental Tile Set

By the predicted quadrant N , a supplemental tile set consists of the tiles in the direction where the viewport is fac-

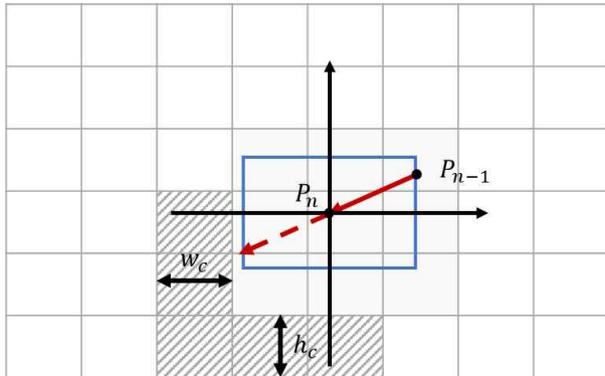


Fig. 3. Configuration of the supplemental tile set

ing, and delivered together in high-quality with the tile set of the current viewport. Figure 3 shows that the diagonally hatched area is considered a supplemental tile set regarding the previous viewpoint P_{n-1} . The tiles corresponding to the diagonal hatched area are consumed as high-quality at time n . Also, the range of supplemental tile set can be adjusted depending on the network conditions and type of content.

Figure 4 shows how the segment duration and the arrival time of the tile quality change request affects to the client's immersive experience. When the viewpoint moves from P_{n-1} to P_{n+1} , as shown in Figure 4(a), there are three delivery scenarios as illustrated in Figure 4(b).

In Figure 4, the case A is not employing a supplemental tile set, but the case B and C adopt the supplemental tile set. In the case A and B, the request of tile quality switching and response are accomplished during k -th segment consumption. Therefore, regardless of the adoption of the supplemental tile set, the tile quality is switched flexibly and does not affect to the user's immersion. Figure 4 also shows the delivery scenario according to the case where the requested switching is received from the video composed of "48-grid 8x6 uniform tiles".

In the case C, the request is delivered during k -th segment consumption, but the segment delivered after k -th segment is fully consumed. Hence, the tile quality will be changed at $(k+2)$ -th segment, but not at the targeted $(k+1)$ -th segment. Even if the tile quality does not change at the targeted $(k+1)$ -th segment, the viewport area is in all high-quality at $(k+1)$ -th segment because supplemental tiles are provided as high-quality. Otherwise, the user experiences poor quality, since low-quality tile is shown in the edge of the viewport. However, QoE cannot be guaranteed if the gaze moves too fast so that the viewpoint gets out of the supplementary tiles. Thus, it is necessary to contemplate consideration for the movable range of the user's head movement.

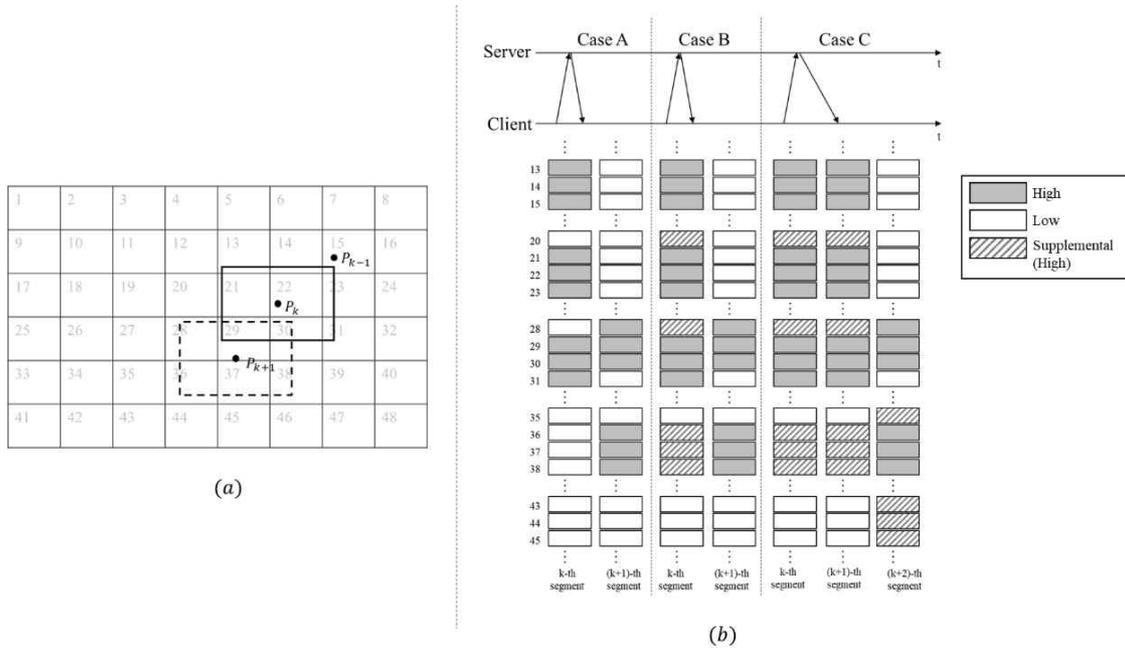


Fig. 4. Delivery scenario according to the case where the requested switching is received from the video composed of "48-grid 8x6 uniform tiles"

III. Consideration of supplemental tile set transmission

1. Influence of segment duration

Even though the direction of gaze is successfully predicted and the supplemental tiles corresponding to the predicted direction is transmitted together with the tiles corresponding to the viewport, it is difficult to guarantee suitable quality during the event due to the duration of DASH segment or the velocity of the HMD.

In Figure 5, at time t_{req} , the viewport hits the contour

of another tile and the client requests to switch the quality of the changed viewport. At this time, the client is consuming the k -th segment. Upon receiving the request, the server delivers high-quality tiles corresponding to the viewport and to the predicted direction back to the client. But there are three different points when the client receives the response and it has completely different results depending on the cases. In other words, depending on when the points of response to the request is reached from the server, it could be a big deal for QoE.

In the case of ①, as shown in Figure 5, there is no

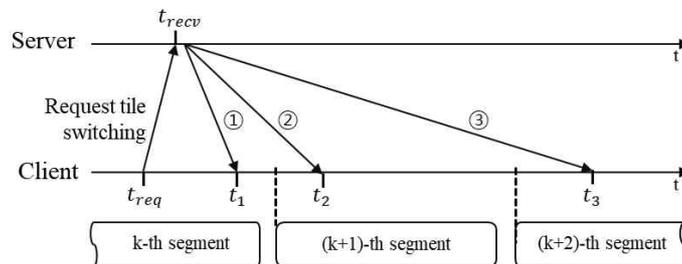


Fig. 5. Scenario of tile switching request

problem for consuming high-quality viewport using tiles prepared in advance through direction prediction in the next segment, (k+1)-th segment, because the switching request is received during consuming the current segment, k-th segment. On the other hand, in the case of ② and ③, The (k+1)-th segment must be consumed to accept the current switching request. As a result, there could be additional delay for segment duration in the worst case. When the viewport moves out of the range of the transmitted supplemental tile set, the user consumes low-quality content and experiences a huge loss of immersion.

2. Movable range of viewport position

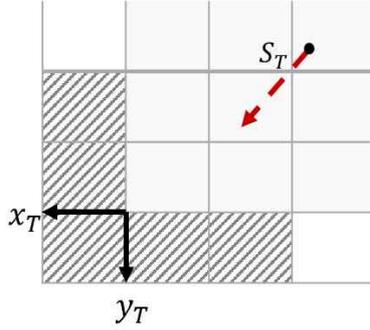


Fig. 6. Gaze movement (x_T, y_T) for T duration. The shaded and diagonally hatched areas are the tiles for viewport and supplemental tiles, respectively.

In Figure 6. S_T is the distance for the viewpoint to move during the segment duration T . At this time, the movement of viewpoint assumes a continuous and uniform motion of straight line. S_T can be expressed as broken up into two separate horizontal, x_T , and vertical, y_T , directions to simplify calculations as shown in Equation (6).

$$S_T = x_T + y_T \quad (6)$$

where x_T and y_T are the distance of viewpoint moved on the x -axis and y -axis, respectively. C is the number of moments in which a tile change request occurs when

the viewpoint crosses the boundary of the tile. Meanwhile, other conventional notations are listed in Table 1.

Table 1. Explanation of conventional notations

Notation	Meaning
T	DASH Segment Duration
$N \times M$	The number of Tiles (row x column)
$W \times H$	Resolution (Width x Height)
$C = C_x + C_y$	The number of tile conversions
$w \times h$	Size of tile (Width x Height)

When C_x or C_y exceeds 2, it can be seen that the additionally transmitted tile set exceeds the range that can be covered in a segment. Thus x_T and y_T are represented by Equation (7) and (8), if the supplemental tile set is successfully worked.

$$x_T = w + w_c \quad (7)$$

$$y_T = w + w_c \quad (8)$$

where w_c and h_c are additional range for supplemental tile set by the user definition, each determined by a multiple of the tile's width and height. Also, the and could be set to expand the supplemental tile set in consideration of network conditions.

If a video of 3840x2160 resolution is divided into 6 by 4 tiles and there is no user definition w_c and h_c , the HMD could move within 60° left and right, and 45° up and down to fully utilize the function of the supplemental tile set. That is, the viewport can move 640 pixels to the left and right, and 480 pixels up and down in an equirectangular image, in which only the middle of the equirectangular plane is adopted considering the distortion. Likewise, the video divided by 10x8 tiles can be horizontally moved 36°, and 22° vertically. It can be seen that the technique of transmitting these supplemental tiles is a great advantage to QoE if there is not much movement to turn the head back within 1 second.

IV. Experimental Results

The supplemental tile set proposed in this paper is meaningful in that additional time can be obtained to provide the tiles of high-quality in viewport despite the user's movement. Figure 7 shows the movable range of the viewport in the horizontal direction depending on how many tiles the video is divided into when the FoV is 90°.

As the number of tiles increases, the tile size becomes relatively small, which generally decreases the movable range. The size of the movable range can be different depending on where the current viewport is located. The minimum and maximum movable range are the value d of Figure 8(a) and (b), respectively.

In Figure 7, the dashed-line corresponds to the conventional tile-based streaming, whereas the solid-line represents case of adopting supplemental tile set. Through the results, it can be seen that the maximum movable range is wider with the supplemental tile set applied than without. If the supplemental tile set is expanded, the movable range also increases. Furthermore, the minimum movable range with supplemental tile set and the maximum movable range without supplemental tile set are just the same, and the in-

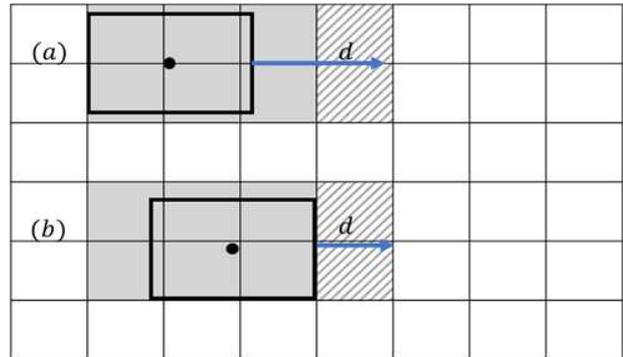


Fig. 8. (a) The maximum movable range and (b) the minimum movable range

creased movable range is obtained as the difference between them. Consequently, this difference becomes a solution which avoids to disturb immersion even though the user's movement is fast.

However, the time it takes to reach the low-quality varies depending on the HMD's angular velocity. Figure 9 shows the point where the high-quality tile set ends from the current viewport position in the horizontal direction. That is, this graph represents the time it takes from the moment of entering into the low-quality tiles to the time when the user's immersion decreases.

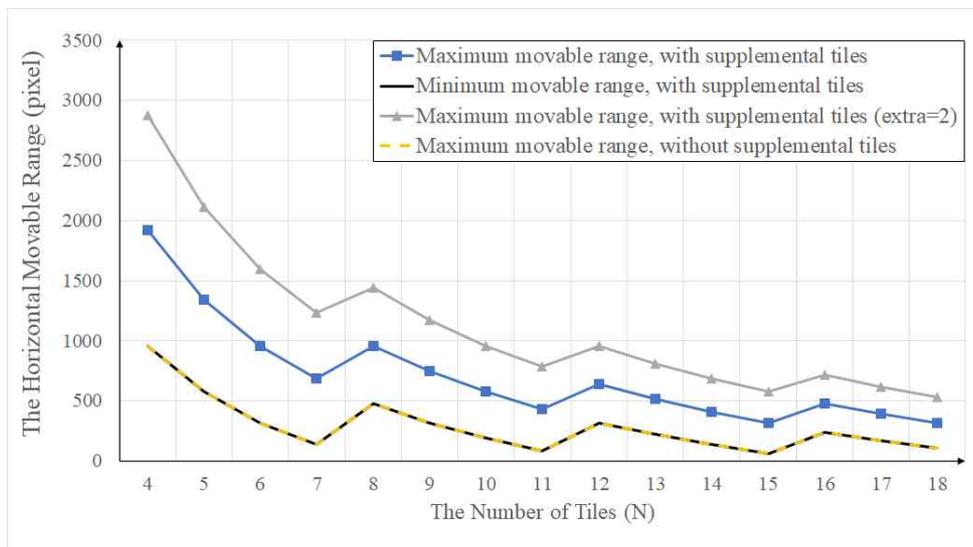


Fig. 7. The horizontal movable range according to the number of tiles

The HMD's angular velocity $2\pi/1\text{sec}$, $2\pi/3\text{sec}$, and $2\pi/6\text{sec}$ can be expressed as fast, normal, and slow, respectively. If the movement of HMD is slow, the time it takes to reach the low-quality tiles would be good enough. The slower the speed, the more stable the supplemental tile set can be used. Of course, the larger the supplemental tile set, the more time can be gained. For instance, there is a video divided horizontally into 8 tiles. If a primary supplemental tile set is configured and the angular velocity is $2\pi/3\text{sec}$, the time to reach low-quality is saved from a maximum of 0.375 to a minimum of 0.75 seconds. But if the supplemental tile set is increased by a tile, the time is saved to 1.125 seconds. In addition, if more tiles and larger size of the supplemental tile set are provided, the bandwidth is used more efficiently because the area of supplemental tile set becomes relatively smaller than that of dividing the tiles into smaller numbers.

As the quality conversion of tiles is limited according to the size of the segment duration, it is necessary to check the utilization range of the supplemental tile set according to the duration. In Figure 10, the time to reach the low-quality tile can be obtained depending on the angular

velocity of the user's HMD, where the N is the number of row tiles and the extra means increasing the number of tiles from primary supplemental tile set in horizontal direction. At this time, there are several reasons for considering only the minimum movable range in Figure 7 as an analysis in Figure 10. Firstly, it does not account for times of differences to reach low-quality tiles depending on the location of the viewport. Secondly, as mentioned in Section III. 2, the quality conversion request occurs as soon as the viewport enters into the supplemental tile set. In other words, the next segment has to be received until the viewport enters and exits to and from the supplemental tile set. Therefore, the user must move slower than the speed at any arbitrary point on the graph to not see the low-quality tiles.

If the segment duration is 1 second, the time to reach low-quality tiles should be larger than the segment duration 1 second. It means that the user must move slower than the angular velocity at the value of 1 second to ensure the user's immersion experience without consuming the tile located outside of the supplemental tile set. Therefore, if the average angular velocity of the HMD movement during video playback is $2\pi/3\text{sec}$ in the horizontal direction, the

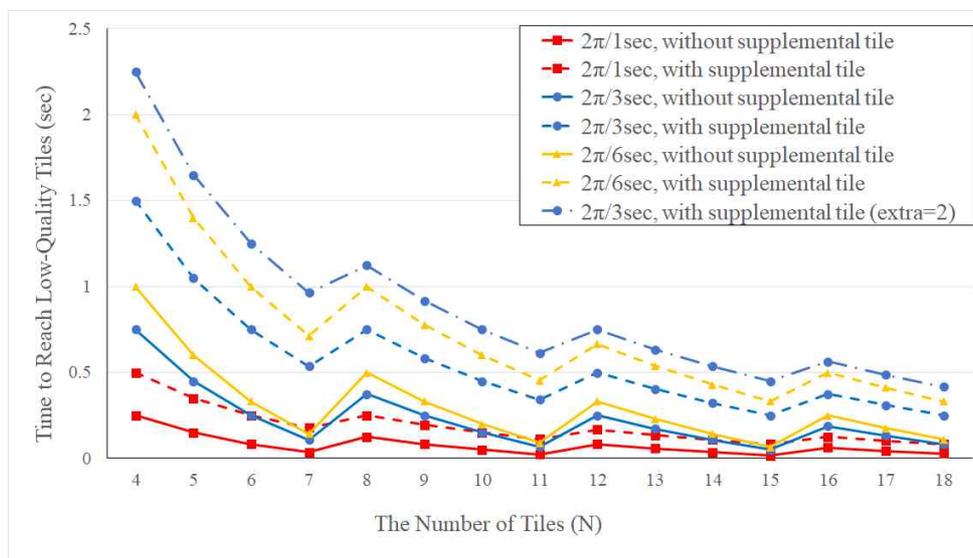


Fig. 9. The time to reach low-quality tiles depending on the number of tiles in horizontal direction

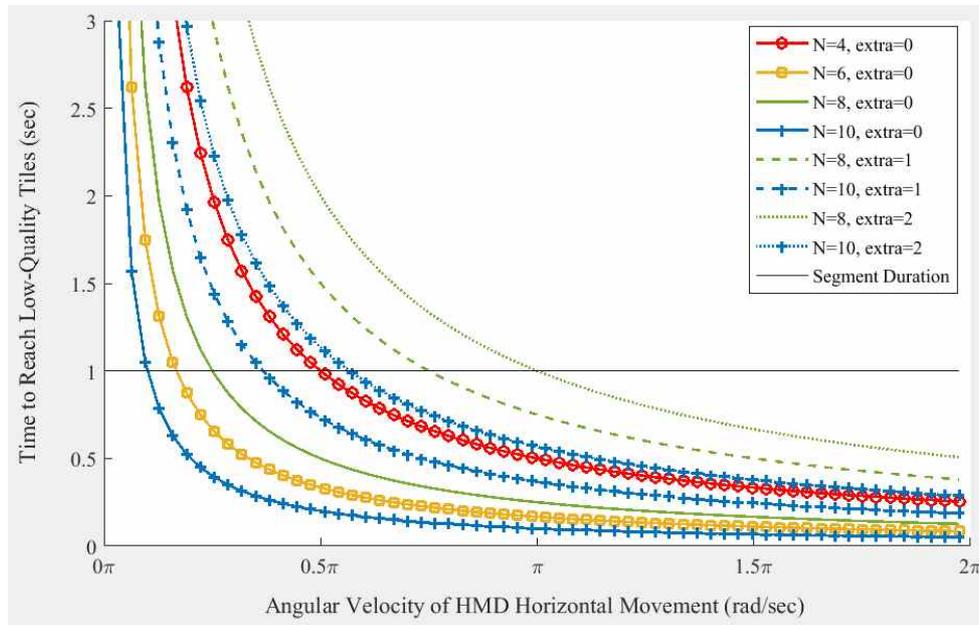


Fig. 10. The time to reach low-quality tiles depending on the angular velocity of the HMD of tiles in horizontal direction

video could be split into 4 or less fewer tiles for encoding. And it could even be considered setting the supplemental tile set one step larger and splitting it into 10 tiles. It is possible to perform more efficient transmission by determining how much fast movement of user is required in the video and applying an appropriate tile size for the proposed supplemental tile set.

V. Conclusion

In this paper, we propose a method of predicting the direction of viewpoint movement and configuring a supplemental tile set corresponding to the predicted direction. In addition, we investigated how to efficiently perform tile-based streaming in consideration of the DASH segment duration and the number of divided tiles. The proposed method guarantees a high-quality viewport corresponding to the user's movement despite the limited circumstances. However, if it is split into more tiles or if the HMD is moving instantaneously and rapidly, and the shorter the dash

segment duration is, it could be disadvantageous to switch the quality of tiles. It is necessary, consequently, to study a more efficient and accurate viewpoint prediction technique. A study on the application of MCTS suitable for the characteristics of the content and the network condition should also be conducted. The effectiveness of the proposed algorithm was proved through extensive simulations.

References

- [1] M. Zink, R. Sitaraman, K. Nahrstedt, "Scalable 360° Video Stream Delivery: Challenges, Solutions, and Opportunities", *Proceedings of the IEEE*, vol. 107, no. 4, pp. 639-650, 2019.
- [2] ISO/IEC 23009-1:2019, "Information Technology – Dynamic Adaptive Streaming over HTTP (DASH) – Part 1: Media Presentation Description and Segment Formats", 2019.
- [3] ISO/IEC 23008-1:2017 "High efficiency coding and media delivery in heterogeneous environments - MPEG-H Part 1: MPEG Media Transport (MMT)", 2017.
- [4] O. A. Niamut, E. Thomas, L. D'Acunto, C. Concolato, F. Denoual, S. Y. Lim, "MPEG DASH SRD: spatial relationship description", *Proceedings of the 7th International Conference on Multimedia Systems*, Klagenfurt, Austria, pp. 1-8, 2016.

- [5] K. K. Sreedhar, A. Aminlou, M. M. Hannuksela, M. Gabbouj, "Viewport-adaptive encoding and streaming of 360-degree video for virtual reality applications", 2016 IEEE International Symposium on Multimedia (ISM), pp. 583-586, 2016.
- [6] M. Hosseini, V. Swaminathan, "Adaptive 360 VR video streaming based on MPEG-DASH SRD", 2016 IEEE International Symposium on Multimedia (ISM), pp. 407-408, 2016.
- [7] L. D'Acunto, J. Van den Berg, E. Thomas, O. Niamut, "Using MPEG DASH SRD for zoomable and navigable video", Proceedings of the 7th International Conference on Multimedia Systems, pp. 1-4, 2016.
- [8] J. Le Feuvre, C. Concolato, "Tiled-based adaptive streaming using MPEG-DASH", Proceedings of the 7th International Conference on Multimedia Systems, pp. 1-3, 2016.
- [9] A. Zare, A. Aminlou, M. M. Hannuksela, M. Gabbouj, "HEVC-compliant Tile-based Streaming of Panoramic Video for Virtual Reality Applications", Proceedings of the 24th ACM international conference on Multimedia, Amsterdam, The Netherlands, pp. 601 - 605, 2016.
- [10] C. Concolato, J. Le Feuvre, F. Denoual, F. Maze, E. Nassor, N. Ouedraogo, et al., "Adaptive Streaming of HEVC Tiled Videos Using MPEG-DASH", IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 8, pp. 1981-1992, 2018.
- [11] X. Corbillon, G. Simon, A. Devlic, J. Chakareski, "Viewport-adaptive navigable 360-degree video delivery", 2017 IEEE international conference on communications (ICC), pp. 1-7, 2017.
- [12] ISO/IEC 23008-2:2020, "Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding", 2020.
- [13] A. Van Rhijn, R. Van Liere, J. D. Mulder, "An analysis of orientation prediction and filtering methods for VR/AR", IEEE Proceedings. VR 2005. Virtual Reality, 2005., pp. 67-74, 2005.
- [14] X. Hou, S. Dey, J. Zhang, M. Budagavi, "Predictive Adaptive Streaming to Enable Mobile 360-degree and VR Experiences", IEEE Transactions on Multimedia, 2020.
- [15] E. Jeong, D. You, C. Hyun, B.-S. Seo, N. Kim, D. H. Kim, et al., "Viewport prediction method of 360 VR video using sound localization information", 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 679-681, 2018.
- [16] J. Zou, C. Li, C. Liu, Q. Yang, H. Xiong, E. Steinbach, "Probabilistic Tile Visibility-Based Server-Side Rate Adaptation for Adaptive 360-Degree Video Streaming", IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 1, pp. 161-176, 2019.
- [17] X. Feng, V. Swaminathan, S. Wei, "Viewport Prediction for Live 360-Degree Mobile Video Streaming Using User-Content Hybrid Motion Tracking", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 3, no. 2, pp. 1-22, 2019.
- [18] M. Jamali, S. Coulombe, A. Vakili, C. Vazquez, "LSTM-Based Viewpoint Prediction for Multi-Quality Tiled Video Coding in Virtual Reality Streaming", 2020 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-5, 2020.
- [19] M. Fiedler, H.-J. Zepernick, V. Kelkkanen, "Network-induced temporal disturbances in virtual reality applications", 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), pp. 1-3, 2019.
- [20] M. S. Elbamby, C. Perfecto, M. Bennis, K. Doppler, "Toward low-latency and ultra-reliable virtual reality", IEEE Network, vol. 32, no. 2, pp. 78-84, 2018.

Introduction Authors



Eun-bin An

- Aug. 2016 : B.S. degree, Division of Computer and Telecommunications, Yonsei University
- Mar. 2017 ~ currently : Ph.D. Candidate, Division of Computer and Telecommunications, Yonsei University
- ORCID : <https://orcid.org/0000-0001-7681-4682>
- Research interests : Visual communication, Real-time streaming, 360/VR video



A-young Kim

- Feb. 2016 : B.S. degree, Division of Computer and Telecommunications, Yonsei University
- Mar. 2016 ~ currently : Ph.D. Candidate, Division of Computer and Telecommunications, Yonsei University
- ORCID : <https://orcid.org/0000-0002-3793-1365>
- Research interests : Visual communication, Real-time streaming, 360/VR video

Introduction Authors



Kwang-deok Seo

- Feb. 1996 : B.S., Department of Electrical Engineering, KAIST
- Feb. 1998 : M.S., Department of Electrical Engineering, KAIST
- Aug. 2002 : Ph.D., Department of Electrical Engineering, KAIST
- Aug. 2002 ~ Feb. 2005 : Senior research engineer, LG Electronics
- Sep. 2012 ~ Aug. 2013 : Courtesy Professor, Univ. of Florida, USA
- Mar. 2005 ~ currently : Professor, Yonsei University
- ORCID : <http://orcid.org/0000-0001-5823-2857>
- Research interests : Video coding, Visual communication, digital broadcasting, multimedia communication system