

Overview of Motion-to-Photon Latency Reduction for Mitigating VR Sickness

Yeongil Ryu and Eun-Seok Ryu*

Department of Computer Education, Sungkyunkwan University
Seoul, South Korea

[e-mail: esryu@skku.edu]

*Corresponding author: Eun-Seok Ryu

*Received January 6, 2021; revised February 24, 2021; revised April 14, 2021; accepted June 10, 2021;
published July 31, 2021*

Abstract

For several years, virtual reality (VR) and augmented reality (AR) technologies have been improving. However, some hurdles remain that slow down the distribution of VR and AR devices, such as head-mounted display (HMD), and related consumer content. One issue is VR motion sickness, which has been experienced by users using 360 degree VR content via HMD. This paper discusses the related international standardization work that classifies the factors causing VR sickness, and proposes the process for VR sickness level evaluation.

Among the factors causing VR sickness, many research institutes regard minimizing MTP (Motion-to-Photon) latency as the key enabler to mitigate VR sickness. Thus, this paper introduces research trends of MTP latency measurement and MTP latency mitigation. This paper categorizes the research on MTP latency measurement into 2 categories of hardware-based approach and software code-level approach. The 2 approaches have different pros and cons depending on use-case, purpose, and architecture of each multimedia system. The pros and cons are addressed in this paper. Additionally, the research on mitigating MTP latency with diverse strategies such as proactive computing, caching, and edge server technology is explained, and compared to conventional technologies, shows improved performance.

Keywords: Motion-to-Photon latency, tile, video coding, viewport, VR sickness, 360 degree videos

1. Introduction

Recent developments in the 360 degree video technology for virtual reality (VR) and augmented reality (AR) applications are key next-generation factors in the field of multimedia technology. Therefore, the Moving Picture Experts Group (MPEG), an international standard organization, launched the MPEG-I (Coded Representation of Immersive Media, ISO/IEC 23090) project in 2016 to develop standards for immersive media technologies, i.e., 3 degrees of freedom (3DoF), 3DoF+, 6DoF, and light fields. The project is expected to be completed by 2021 [1, 2].

Conversely, several factors have been slowing down the distribution of VR and AR devices, such as head-mounted display (HMD), and related consumer content. One issue is VR motion sickness, which has been experienced by users using 360 degree VR content via HMD [3]. Users experience motion sickness while using 360 degree VR content, owing to the disharmony between human motion and the senses from sensory organs in skin and muscles, eyes, and vestibular organs (Fig. 1).

IEEE, another international standard organization, launched the IEEE 3079 standard (IEEE standard for HMD-based VR sickness reduction technology) [4]. Unlike MPEG-I, IEEE 3079 standard focuses on the reduction of VR motion sickness. The IEEE 3079 standard defines the factors that affect the user's VR sickness, and suggests the process for analyzing and evaluating the user's VR sickness level in a quantitative domain. The standardization project for the IEEE 3079 was started in 2017, and it is expected to be finalized by 2021. Fig. 2 shows the timeline of the standardization of IEEE 3079 and related standard.

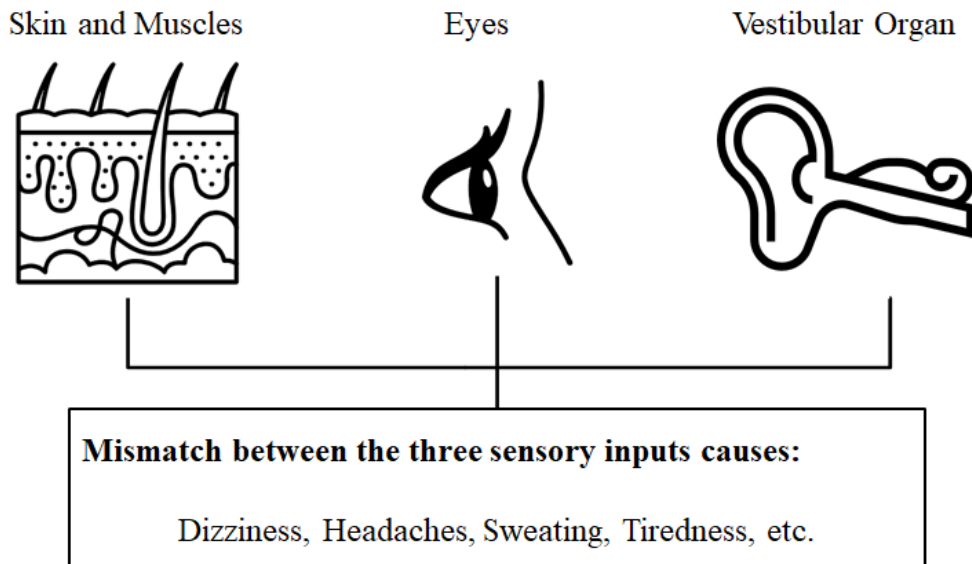


Fig. 1. Three sensory organs, and the cause of motion sickness.

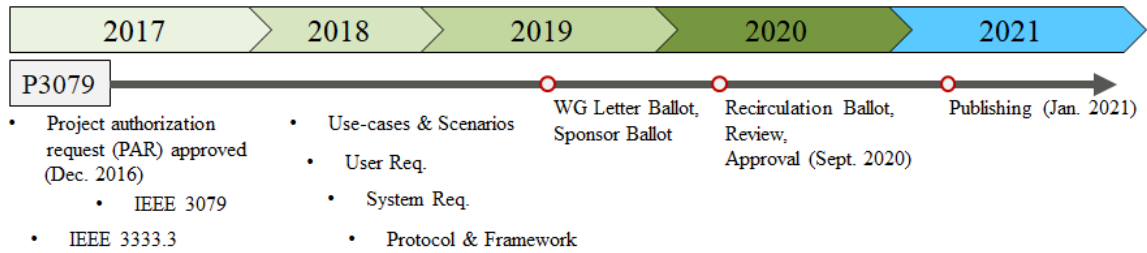


Fig. 2. The IEEE 3079 standardization project.

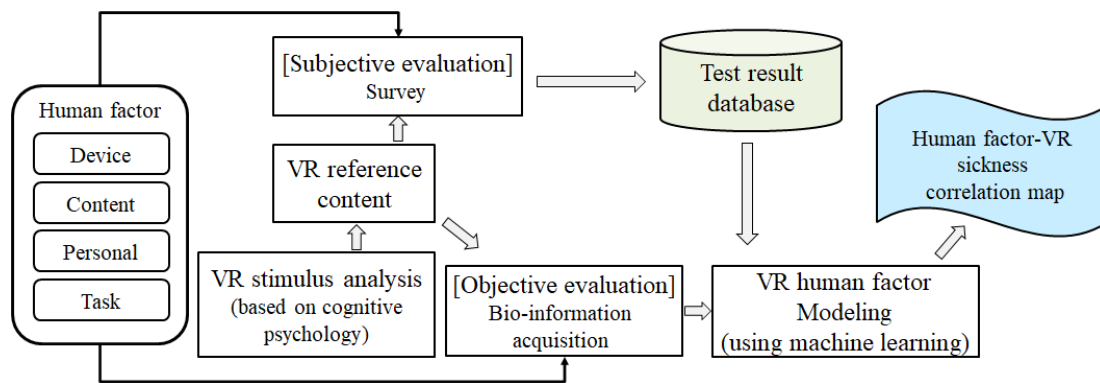


Fig. 3. VR sickness level evaluation process in IEEE 3079.

The factors that affect user VR sickness include the continuity of viewpoint, 3D sound, stitching optimization, latency minimization, resolution and flicker optimization, as shown in [Table 1](#). IEEE 3079 surveyed the correlation between the factors and user motion sickness. IEEE 3079 used the result of the survey for subjective motion sickness evaluation and user bio-information for objective motion sickness evaluation. Based on the subjective and objective evaluation tools, IEEE 3079 proposed the VR sickness evaluation process and measurement metric in a quantitative domain. [Fig. 3](#) depicts the VR sickness evaluation process of IEEE 3079, which can be useful for the VR sickness level of diverse VR multimedia systems with HMD and can be an indicator inducing VR sickness mitigation.

Among the factors distinguished by IEEE 3079, Motion-to-photon (MTP) latency significantly induces the disharmony of the aforementioned senses when users experience 360 degree VR content via HMD. MTP refers to the delay between the motion of the user, and the corresponding video shown on the display [5]. A lengthy MTP latency time intensifies VR motion sickness. [Fig. 4](#) shows that MTP latency on HMD-based VR systems is caused by delays in motion sensor input and computation, motion sensor data communication, application computation, rendering computation, and display refresh [6]. Depending on the use of various applications, delays in application computation may refer to delays in video data decoding computation and communication. In addition, delays in the display refresh depend on the refresh rate of the HMD display. [Fig. 5](#) shows that regular displays with refresh rates of (60 and 30) Hz have average delays of (8.3 and 16.6) ms, respectively; in the worst-case scenario, the refresh rates for (60 and 30) Hz can be as bad as (16.6 and 33.3) ms, respectively.

Table 1. Factors causing VR sickness.

VR content design	Management for VR scene	VR scene capture
<ul style="list-style-type: none"> • Continuity of viewport • Placement of Scene structure • 3D sound 	<ul style="list-style-type: none"> • Optimization of viewport rotation • Optimization of scene complexity • FoV adjustment • Placement of user interface • Visual flow • VR fidelity • Frame reference 	<ul style="list-style-type: none"> • Optimization of stitching • Rig structure
Features of HMD device	User's human factor	Management for Runtime environment
<ul style="list-style-type: none"> • Minimization of latency • Optimization of frame rate • Optimization of stereo 3D • Optimization of resolution • Display types • Flicker optimization 	<ul style="list-style-type: none"> • Gender/age • Previous experience • Sensitivity to motion sickness • Tolerance for VR experience • VR sickness control 	<ul style="list-style-type: none"> • Motion platform sync • Vertical synchronization • Clinical trial

As explained above, MTP latency has different characteristics depending on the multimedia system. Researchers have developed techniques to measure and reduce MTP latency, and such technologies are very important for the expansion of the VR and AR markets.

The remainder of the paper is structured as follows. Sections 2 and 3 discuss research trends on techniques and technologies to measure and reduce MTP latency, while Section 4 concludes the paper.

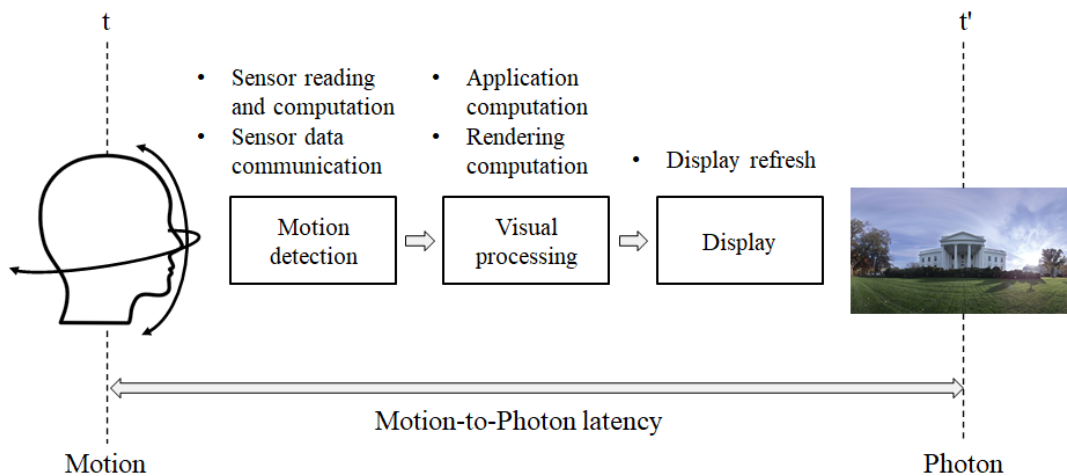


Fig. 4. MTP latency in HMD systems and its causes.

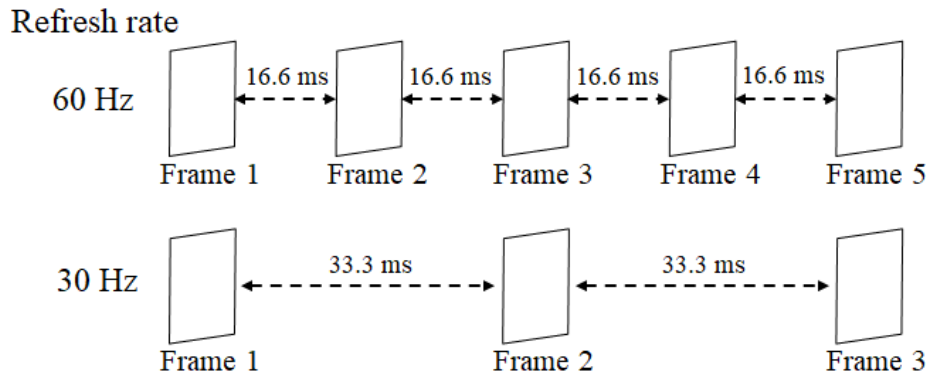


Fig. 5. Output delays depending on the display refresh rate.

2. Trends of Research on MTP Latency Measurement

Several studies have considered different techniques to measure MTP latency, depending on the use and features of the target multimedia system. The studies can be classified into two categories: (1) studies that design and implement a hardware-based MTP latency measurement system, and (2) studies that suggest an MTP latency measurement technique based on a software code level. Sections 2.1 and 2.2 review research for the aforementioned categories, and discuss the limitations of the studies.

2.1 Time Sequential MTP Latency Measurement System for HMD Developed by Sogang University

In 2018, Choi et al. at Sogang University developed a hardware-based MTP latency measurement system using a device that mechanically mimics the motions of the human head model [7]. The proposed system comprises a motion generator that creates motions similar to human head movements and sightline changes based on the neck motion data of the user [8], and a photodetector system that measures the final delay of the HMD screen using photodiode modules.

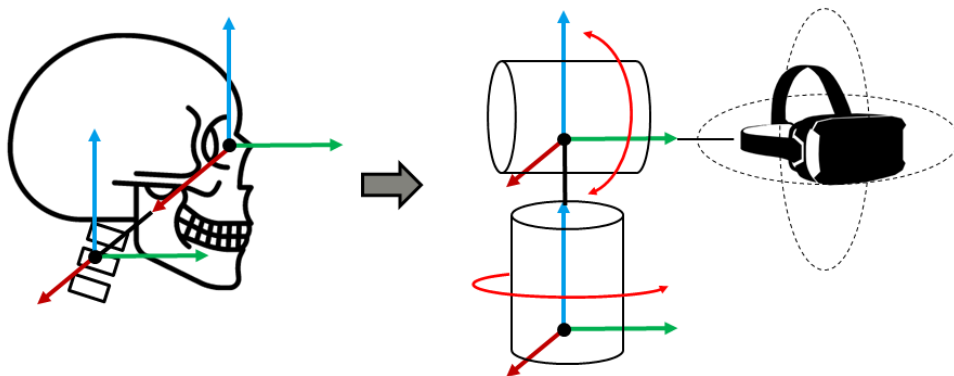


Fig. 6. Kinematic modeling of the motion generator.

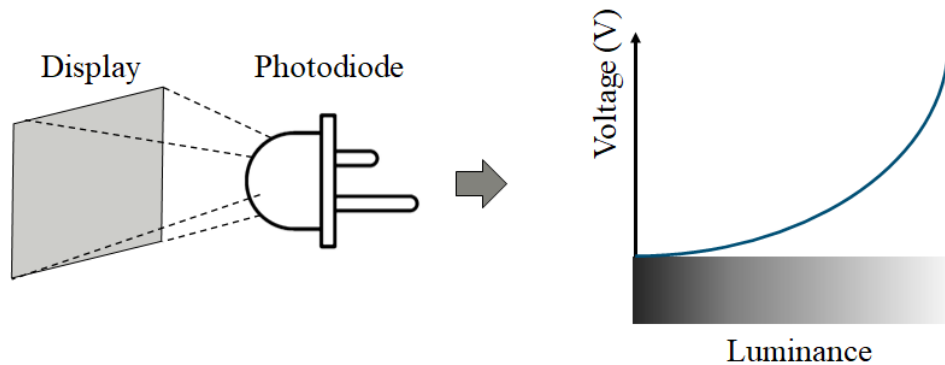


Fig. 7. Mechanism of a photodiode.

The motion generator uses servo motors to implement yaw and pitch rotations of the neck of the user. However, to simplify the system, the module does not consider roll rotation, which is less frequent. **Fig. 6** shows that the motion generator mimics the joints and links in the human head using servo motors; the system uses forward kinematics [9, 10] to generate motions based on predefined user movement scenarios.

The photodetector system comprises photodiode modules and a display. A photodiode is a semiconductor diode that generates output voltages proportional to the brightness of the input light (**Fig. 7**) [11]. The system uses such attributes of photodiodes to calculate the output voltage corresponding to the brightness of the input light. The output voltage can be reverse-converted to calculate brightness as well.

The proposed system measures the MTP latency through the following processes: First, it configures a motion sequence that corresponds to a pre-defined scenario. Next, the sequence is entered into the motion generator and photodetector system. The motion generator creates physical movements based on the input motion sequence, and measures the resulting physical rotation angles. The output of the photodetector system is four square objects placed on a playing image.

The brightness level of the square objects is determined by finding the corresponding motion sequence according to the mapping table listing the pre-defined brightness–rotation angle. **Fig. 8** shows that the brightness of the two right-hand-side objects represents the yaw angle, while the brightness of the two left-hand-side objects represents the pitch angle. When these four square objects are overlaid on the screen, the photodiode modules receive the brightness information of the output objects to generate voltages that correspond to the brightness levels. Through this process, the photodetector system obtains the current rotation angle information output on the display.

The proposed system compares the timing of the physical rotation angles measured by the motion generator, and that of the rotation angles measured by the photodetector system, thereby calculating the MTP latency in real time. **Fig. 9** shows the system mechanism.

The MTP latency measurement system proposed by Choi et al. at Sogang University is different from those proposed in other studies. This system generates physical motions based on pre-defined scenarios and measures MTP latency in real time. However, the study does not consider natural scene content, and focuses on computer-generated content. Another shortcoming is that the proposed system only supports the measurement of 2 DoF (yaw and pitch).

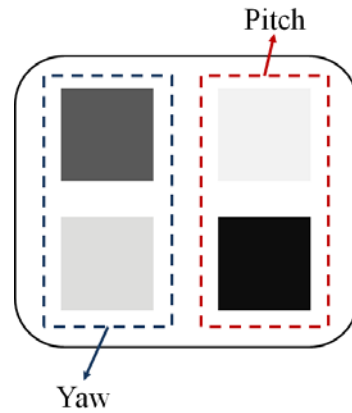


Fig. 8. Square objects that are placed on a scene to indicate yaw and pitch rotation angles.

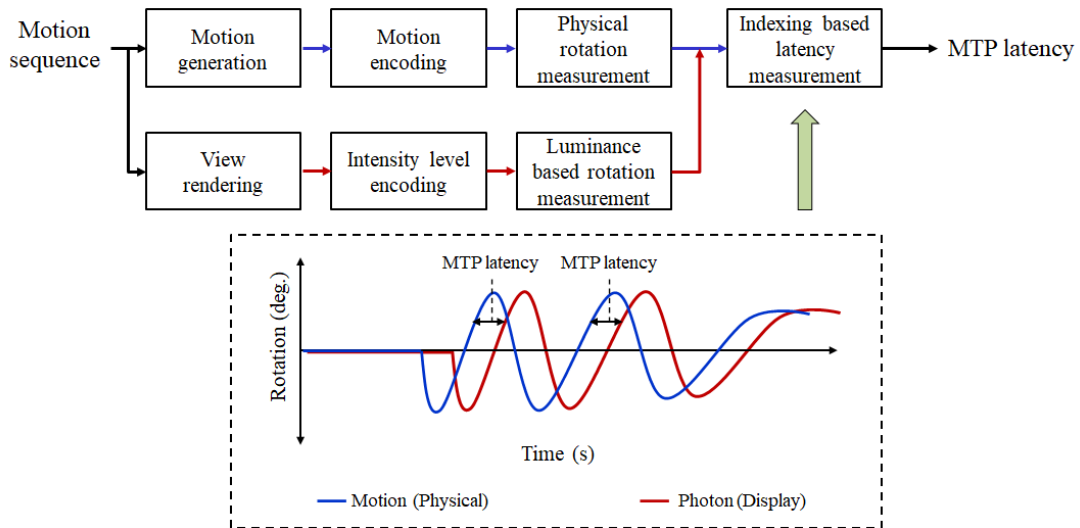


Fig. 9. Block diagram of the MTP measurement system.

2.2 Software Code-Level MTP Latency Measurement by Fraunhofer HHI

Fraunhofer HHI, the German research institute, suggested a software code-level MTP latency measurement technique while developing a low-latency cloud-based streaming system for volumetric videos in 2020 [12]. The proposed system does not directly transmit volumetric videos from the server to the client [13]. Rather, the server renders the 3D view of the volumetric video that corresponds to the current viewport of the user, after which the rendered 3D view is converted to 2D images that are compressed and transmitted to the client. This method allows the client to offload 3D video processing into the server, thereby reducing the load on the client, and increasing the MTP latency. To reduce the MTP latency, the Fraunhofer HHI considered a software code-level approach, which can yield efficient processing, to measure the MTP latency.

For accurate MTP latency measurement, the study defined MTP latency in the volumetric video streaming system as follows:

$$T_{mtp} = T_{server} + T_{network} + T_{client} \quad (1)$$

T_{server} , $T_{network}$, and T_{client} are defined in Eqs. (2)-(4), respectively:

$$T_{server} = T_{rend} + T_{enc} \quad (2)$$

$$T_{network} = T_{up} + T_{down} + T_{trans} \quad (3)$$

$$T_{client} = T_{dec} + T_{disp} \quad (4)$$

Equation (2) refers to the delay time occurring on the server side, where T_{rend} refers to the time required to render the volumetric images corresponding to the current viewport of the user; and T_{enc} refers to the time required to convert the 3D view of the rendered volumetric video to 2D images, and then encode using a video codec, such as H.264/AVC [14] or H.265/HEVC [15].

Equation 3 refers to the total delay time that occurred during the server–client data communication, where T_{up} and T_{down} represent the propagation delay between the server and the client; and T_{trans} refers to the time required to transmit the data.

Equation (4) refers to the delay time on the client side, where T_{dec} is the time required to decode the received video data; and T_{disp} refers to the delay time for outputting the decoded picture on the screen.

Fig. 10 shows a schematic of how Eq. (1) is applied to a cloud-based volumetric video streaming system. In the study, the system makes software code-level recordings of time R_i , when the client application requests frame F_i corresponding to the viewport of the user, and time D_i , when the requested frame is received and output to the display to measure the MTP latency based on Eq. (1) and **Fig. 10**. Then, to measure the MTP latency in a simplified manner, T_{mtp} is obtained from $D_i - R_i$.

The MTP latency measurement technique of Fraunhofer HHI is different from the technique developed by Choi et al. at Sogang University; the Fraunhofer HHI technique is easier to implement than the hardware-based MTP latency measurement method, and considers delays occurring on the network.

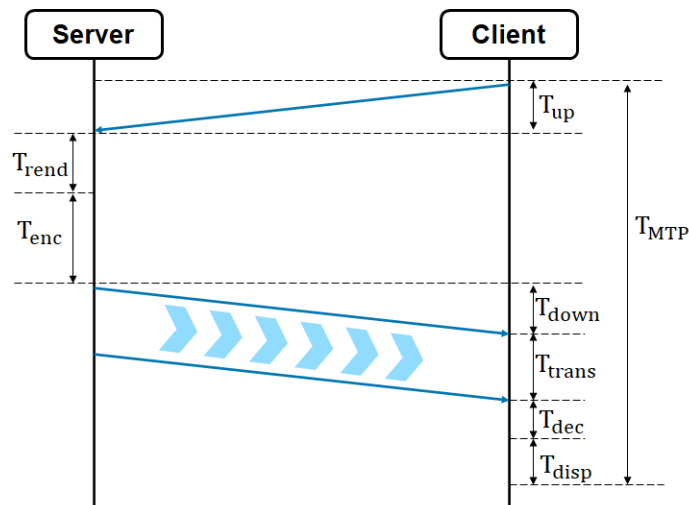


Fig. 10. Elements of MTP latency in a volumetric video streaming system.

3. Trends of Research on MTP Latency Reduction

Along with the MTP latency measurement, MTP latency reduction is an essential element for the quality of experience (QoE) in immersive 360 degree VR content. Key MTP latency reduction technologies include mmWave communication, mobile edge computing, and proactive caching. **Table 2** lists information about the requirements for providing highly reliable VR services, and key technologies necessary to satisfy each requirement [6]. The following sections discuss studies based on edge computing and proactive caching technology that support low-latency computing.

Table 2. Requirements and key technologies for highly reliable VR services.

Requirements	Technology enablers
High capacity	<ul style="list-style-type: none"> • mmWave communications
Low-latency computing	<ul style="list-style-type: none"> • Edge computing • Proactive computing and caching
Low-latency communication	<ul style="list-style-type: none"> • Multi-connectivity • Multicasting
Reliable communication	<ul style="list-style-type: none"> • Multi-connectivity
Scalability	<ul style="list-style-type: none"> • Edge computing • Multicasting

3.1 Adaptive Cropping Technique by AT&T Labs Research

In 2019, AT&T Labs Research in the US proposed a technique that reduces the MTP latency by proactively caching technique and edge servers [16]. The proposed technique reduces the MTP latency during the server–client streaming of 360 degree game videos. Rather than streaming the entire 360 degree view, the system crops only the field of view (FoV) of the user for lower latency and computation. However, when the MTP latency is high, the margins surrounding the FoV are cropped, and additionally transmitted with the video data of FoV. In doing so, the system can output the video data of the additionally transmitted margin areas when the viewport of the user is outside the current FoV, thereby increasing the QoE of the streaming service, even when MTP latency is high, as shown in **Fig. 11**.

The size of the additionally transmitted margin areas is determined based on the current MTP latency. It is determined in real time based on the pre-defined table analyzing the MTP latency and required margin sizes. This table deals with MTP latencies of (50, 100, 300, and 500) ms, providing test results that indicate the optimal size of viewpoint that can cover the user's movement at a specific MTP latency. Then, it maps the size of the margin areas that can fully (100 %) cover the corresponding scope of movements.

Compared to the conventional method of streaming the entire 360 degree view, the proposed method reduced the bandwidth by 80 %. Moreover, this enables 8K live VR streaming.

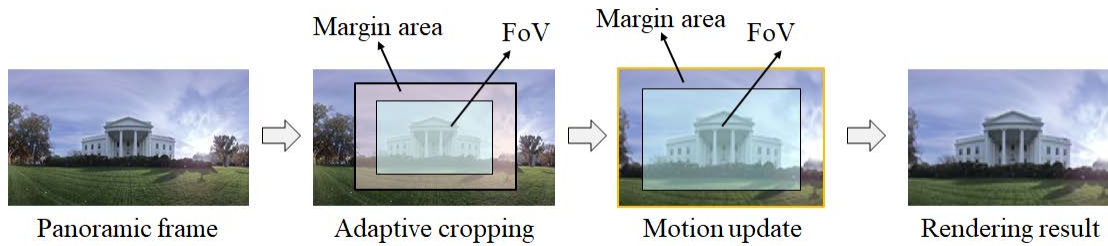


Fig. 11. Adaptive cropping technique considering MTP latency.

3.2 MTP Latency Reduction Technique for Tile-Based 360 degree Video Streaming Systems by UT Dallas

In 2018, UT Dallas in the US proposed a caching technique for tile-based 360 degree video streaming systems in edge computing architecture [17]. Fig. 12 shows that tile-based video streaming divides the video into multiple tiles, and streams only necessary tiles [18, 19, 20, 21]. Compared with the conventional streaming of 8K, 16K, and other ultra-high-definition videos, this method can highly reduce the required network bandwidth and computational complexity [22, 23, 24].

The proposed technique pre-analyzes the movement patterns of a user for each video content, thereby prioritizing each tile based on the frequency with which each tile is included in the FoV. To reduce the MTP latency, the video tiles are cached on the edge server based on priority.

The study analyzed the movement patterns of the user for each content section using the user viewpoint information dataset [25] in 360 degree videos, which was presented by Lo et al. in 2017. Using the analysis results, the proposed system configures a heatmap of each tile in given intervals (Fig. 13). The heatmap represents the possibility that each tile can be included in the FoV. The proposed technique uses this information to prioritize caching tiles with a higher possibility of being included in the FoV, thereby reducing the MTP latency.

The study conducted tests that compare the cache hit rate of the proposed method with that of generally used caching techniques, such as least recently used (LRU), and least frequently used (LFU) [26], caching methods. The test results indicated that compared with LRU and LFU, the proposed method had (20–40) % higher hit rates.

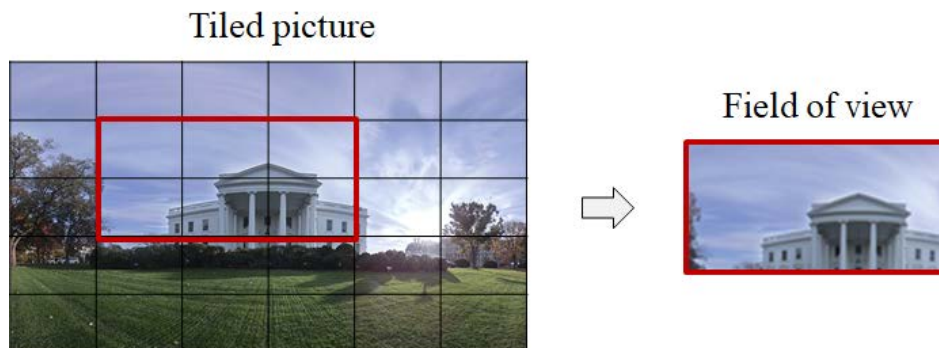


Fig. 12. Example of tile-based video streaming.

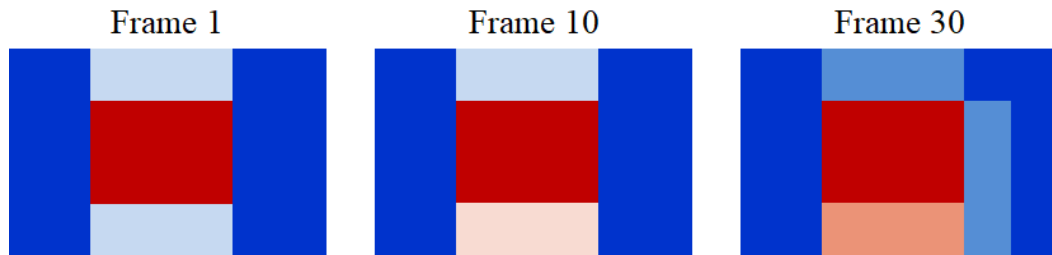


Fig. 13. Example of user motion pattern heatmap.

3.3 Saliency Map-based Bit-rate Adaptation for VR streaming by Sungkyunkwan University

Lee et al. at Sungkyunkwan University proposed the motion-constrained tile set (MCTS)-based 360 degree VR streaming system with bit-rate adaptation technique using saliency map in 2019 [27]. The tile-based video streaming method is a highly efficient technique for VR multimedia systems with HMD, because it transmits only the video data in the user's viewport. Therefore, compared to the conventional approach that transmits the entire pictures, a tile-based approach can achieve higher video quality at limited network bandwidth [28].

Although the tile-based approach has advantages in 360 degree video streaming, the prediction in the process to decode tiled video has a few hurdles. This is because when the reference video data was outside previous viewports at a client side, the reference video data of the tiles in the current viewport can be lost. Therefore, the server side in 360 degree video streaming systems needs to consider the coding structure that restricts the available search range for inter-picture prediction at the encoder level.

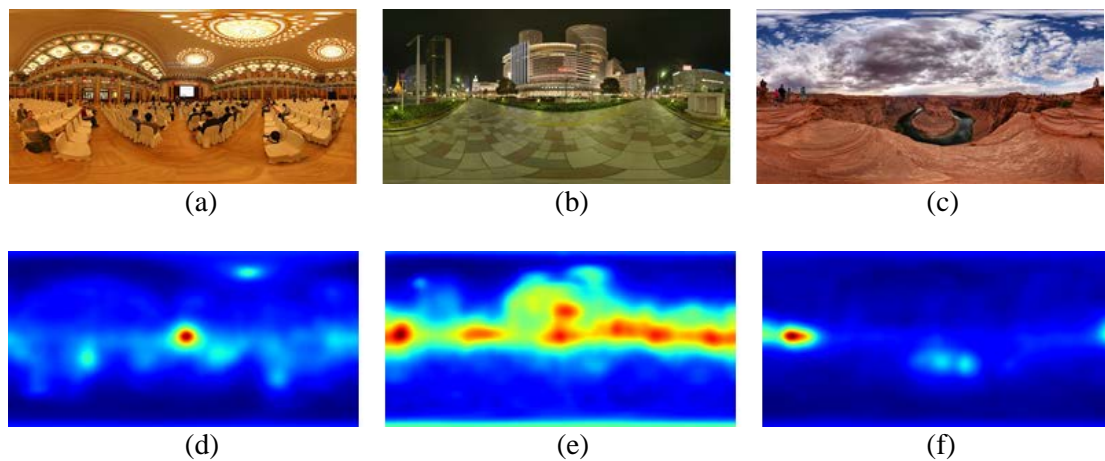


Fig. 14. Example of saliency map. (a)-(c) L Test image in the *Salient360!* Dataset. (d)-(f): Corresponding saliency map prediction from the CNN-based saliency model.

To solve this problem in tile-based 360 degree video streaming systems, MCTS was introduced. MCTS technique controls the search range for reference during temporal prediction to be in the viewport of the previous picture. This coding tool enables independent

decoding of each tile at the client side, regardless of the movement of the user's viewport. Using MCTS achieves dramatic Bit-rate reduction in the entire streaming process, although the coding efficiency of each tile is slightly lower than the conventional coding scheme, owing to the restriction of the search range in temporal prediction [29, 30]. Son et al. implemented and tested the MCTS with HM (HEVC test model software) and SHM (SHVC test model software) in 2018 [31]. According to Son et al., MCTS in the tile-based video streaming system with 9 tiles achieves a bit-rate reduction of about 48 %, compared to the conventional HM [15] and SHM [32]. However, many factors, such as the number of tiles and the size of viewport, can change this result.

To achieve improved performance, Lee et al. employs saliency detection technique [33, 34, 35], as well as MCTS for viewport-dependent video streaming systems, such as 360 degree video streaming with HMD. The saliency detection in the research finds a probability distribution of user eye fixation information based on a convolutional neural network (CNN). Fig. 14 shows an example of saliency map with the *Slient360!* dataset [36], which is a platform for evaluating the performance of saliency mapping models. The authors assigned priorities to each tile. For example, tiles that have a higher probability are assigned higher priority and more bits than others. In doing so, the importance of each tile can be known, and more bits can be assigned to more important tiles for encoding. Table 3 outlines a brief procedure for the assigning of priority and bits based on a given network bandwidth and saliency map.

Table 3. Proposed procedure of the quality adaptation based on saliency map.

```

BW: the available bandwidth [Mbps]
RL: quality level QP {22, 27, 32, 37} ∈ L representation
 $\tau_i^{R_L}$ : each tile with bitrate  $\tau_{bit_i}^{R_L}$ 
 $\tau$ : list of tiles sorted by saliency map in each tile,  $\forall \tau_i \in \tau$ 
Budget ← BW -  $\sum \tau_{bit_{min}}$  % initialization for minimum bitrate
while  $\tau \neq 0$  do
  for  $R_L \ni L' \leq L$  do
    if  $\tau_{bit_i}^{R_L'} - \tau_{bit_i}^{R_L} \leq \textit{Budget}$  then
       $\tau_i^{R_L} - \tau_i^{R_L'}$  % quality assignment
      Budget ← Budget -  $(\tau_{bit_i}^{R_L'} - \tau_{bit_i}^{R_L})$ 
    else
       $\tau \leftarrow \tau - \{\tau_i\}$ 
    end if
  end for
  i ← i + 1
end while

```

Table 4 shows that at 4×4 , 6×3 , 8×4 , and 12×6 tiling, the proposed method achieves a BD-rate reduction of an average (17.2, 17.9, 18.9, and 20) %, respectively. The proposed method focuses on maximizing video quality at a limited network bandwidth, and it can be considered to achieve MTP latency optimization by the proactive computing technique.

Table 4. Bit-rate reduction of the proposed MCTS technique with saliency map

Category	BD-rate (%)			
	4 × 4	6 × 3	8 × 4	12 × 6
High motion	-18.3	-19.1	-20.5	-22.1
Low motion	-16.2	-16.8	-17.4	-18.0
Average	-17.2	-17.9	-18.9	-20.0

4. Conclusion

This paper discussed hurdles for immersive 360 degree videos and VR/AR applications, and the related international standardization work was explained, which classifies the factors causing VR sickness, and proposes the process for VR sickness level evaluation.

Among the factors causing VR sickness, many research institutes regard MTP latency as the key enabler to mitigate VR sickness. Thus, this paper introduced the research trends on MTP latency measurement, and categorized the research on MTP latency measurement into 2 categories of hardware-based approach, and software code-level approach. Furthermore, the pros and cons of each research effort were discussed.

Additionally, this paper introduced the key enablers for the highly reliable VR multimedia system, and research on mitigating MTP latency. The research employs proactive computing, caching, and edge server technology, and compared to conventional technologies, shows improved performance.

Acknowledgement

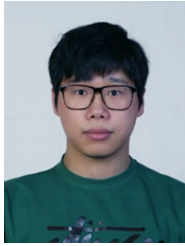
This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2019R1A2C1010476).

References

- [1] M. Domański, O. Stankiewicz, K. Wegner, and T. Grajek, "Immersive visual media – MPEG-I: 360 video, virtual navigation and beyond," in *Proc. of 2017 International Conference on Systems, Signals and Image Processing*, Poznan, pp. 1-9, 2017. [Article \(CrossRef Link\)](#)
- [2] M. Wien, J. M. Boyce, T. Stockhammer, and W. Peng, "Standardization Status of Immersive Video Coding," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 5-17, 2019. [Article \(CrossRef Link\)](#)
- [3] R. Azuma, Y. Baillet, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Comput. Graph. Appl.*, vol. 21, no. 6, pp. 34-47, 2001. [Article \(CrossRef Link\)](#)
- [4] IEEE Std 3079-2020, IEEE, 2020. [Article \(CrossRef Link\)](#)
- [5] B. Iribe, "Virtual reality-A new frontier in computing," in *Proc. of the AMD Announces 2013 Developer Summit*, San Jose, CA, USA, 2013. [Article \(CrossRef Link\)](#)
- [6] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Toward Low-Latency and Ultra-Reliable Virtual Reality," *IEEE Network*, vol. 32, no. 2, pp. 78-84, 2018. [Article \(CrossRef Link\)](#)
- [7] S. W. Choi, S. Lee, Seo, M. W. Seo, and S. J. Kang, "Time sequential motion-to-photon latency measurement system for virtual reality head-mounted displays," *Electronics*, vol. 7, no. 9, pp. 171, 2018. [Article \(CrossRef Link\)](#)

- [8] M. W. Seo, S. W. Choi, S. L. Lee, E. Y. Oh, J. S. Baek, and S. J. Kang, "Photosensor-based latency measurement system for head-mounted displays," *Sensors*, vol. 17, no. 5, pp. 1112, 2017. [Article \(CrossRef Link\)](#)
- [9] P. M. Sharkey, D. W. Murray, and J. J. Heuring, "On the kinematics of robot heads," *IEEE Transactions on Robotics and Automation*, vol. 13, no. 3, pp. 437-442, 1997. [Article \(CrossRef Link\)](#)
- [10] J. Funda, R. H. Taylor, and R. P. Paul, "On homogeneous transforms, quaternions, and computational efficiency," *IEEE Transactions on Robotics and Automation*, vol. 6, no. 3, pp. 382-388, 1990. [Article \(CrossRef Link\)](#)
- [11] SM05PD Mounted Photodiodes – SM05 and SM1 Compatible Specifications. [Article \(CrossRef Link\)](#)
- [12] S. Gül, D. Podborski, T. Buchholz, T. Schierl, and C. Hellge, "Low-latency cloud-based volumetric video streaming using head motion prediction," in *Proc. of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pp. 27-33, 2020. [Article \(CrossRef Link\)](#)
- [13] S. Shi, and C.-H. Hsu, "A Survey of Interactive Remote Rendering Systems," *ACM Comput. Surv.*, vol. 47, no. 4, pp. 1-29, 2015. [Article \(CrossRef Link\)](#)
- [14] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, 2003. [Article \(CrossRef Link\)](#)
- [15] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012. [Article \(CrossRef Link\)](#)
- [16] S. Shi, V. Gupta, M. Hwang, and R. Jana, "Mobile VR on edge cloud: a latency-driven design," in *Proc. of the 10th ACM Multimedia Systems Conference*, pp. 222-231, 2019. [Article \(CrossRef Link\)](#)
- [17] A. Mahzari, A. Taghavi Nasrabadi, A. Samiei, and R. Prakash, "FoV-aware edge caching for adaptive 360 video streaming," in *Proc. of the 26th ACM international conference on Multimedia*, pp. 173-181, 2018. [Article \(CrossRef Link\)](#)
- [18] Y. Bao, H. Wu, T. Zhang, A. A. Ramli, and X. Liu, "Shooting a moving target: Motion-prediction-based transmission for 360-degree videos," in *Proc. of IEEE International Conference on Big Data*, pp. 1161-1170, 2016. [Article \(CrossRef Link\)](#)
- [19] J. He, M. A. Qureshi, L. Qiu, J. Li, F. Li, and L. Han, "Rubiks: Practical 360-Degree Streaming for Smartphones," in *Proc. of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 482-494, 2018. [Article \(CrossRef Link\)](#)
- [20] D. Ochi, Y. Kunita, A. Kameda, A. Kojima, and S. Iwaki, "Live streaming system for omnidirectional video," in *Proc. of IEEE Virtual Reality*, pp. 349-350, 2015. [Article \(CrossRef Link\)](#)
- [21] F. Qian, B. Han, L. Ji, and V. Gopalakrishnan, "Optimizing 360 video delivery over cellular networks," in *Proc. of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*, pp. 1-6, 2016. [Article \(CrossRef Link\)](#)
- [22] M. Graf, C. Timmerer, and C. Mueller, "Towards Bandwidth Efficient Adaptive Streaming of Omnidirectional Video over HTTP: Design, Implementation, and Evaluation," in *Proc. of the 8th ACM on Multimedia Systems Conference*, pp. 261-271, 2017. [Article \(CrossRef Link\)](#)
- [23] M. Hosseini and V. Swaminathan, "Adaptive 360 VR Video Streaming: Divide and Conquer," in *Proc. of 2016 IEEE International Symposium on Multimedia*, San Jose, pp. 107-110, 2016. [Article \(CrossRef Link\)](#)
- [24] P. R. Alface, J. Macq and, N. Verzijp, "Interactive omnidirectional video delivery: A bandwidth-effective approach," *Bell Labs Technical Journal*, vol. 16, no. 4, pp. 135-147, 2012. [Article \(CrossRef Link\)](#)
- [25] C. Wu, Z. Tan, Z. Wang, and S. Yang, "A dataset for exploring user behaviors in VR spherical video streaming," in *Proc. of the 8th ACM on Multimedia Systems Conference*, pp. 193-198, 2017. [Article \(CrossRef Link\)](#)

- [26] H. Chen and Y. Xiao, "Cache access and replacement for future wireless Internet," *IEEE Communications Magazine*, vol. 44, no. 5, pp. 113-123, 2006. [Article \(CrossRef Link\)](#)
- [27] S. Lee, D. Jang, J. Jeong, and E. S. Ryu, "Motion-constrained tile set based 360 degree video streaming using saliency map prediction," in *Proc. of the 29th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pp. 20-24, 2019. [Article \(CrossRef Link\)](#)
- [28] de la Fuente, Y. Sánchez, R. Skupin, and T. Schierl, "Video processing for panoramic streaming using HEVC and its scalable extensions," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5631-5659, 2017. [Article \(CrossRef Link\)](#)
- [29] A. Zare, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, "HEVC-compliant Tile-based Streaming of Panoramic Video for Virtual Reality Applications," in *Proc. of the 24th ACM international conference on Multimedia*, pp. 601–605, 2016. [Article \(CrossRef Link\)](#)
- [30] C. Feldmann, C. Bulla, and B. Cellarius, "Efficient stream-reassembling for video conferencing applications using tiles in HEVC," in *Proc. of the International Conferences on Advances in Multimedia*, pp. 130-135, 2013. [Article \(CrossRef Link\)](#)
- [31] J. Son, D. Jang, and E. S. Ryu, "Implementing motion-constrained tile and viewport extraction for VR streaming," in *Proc. of the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pp. 61-66, 2018. [Article \(CrossRef Link\)](#)
- [32] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20-34, 2016. [Article \(CrossRef Link\)](#)
- [33] D. D. Salvucci, and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proc. of the symposium on Eye tracking research & applications*, pp. 71-78, 2000. [Article \(CrossRef Link\)](#)
- [34] X. Dong, Y. Yan, M. Tan, Y. Yang, and I. W. Tsang, "Late Fusion via Subspace Search with Consistency Preservation," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 518-528, 2019. [Article \(CrossRef Link\)](#)
- [35] E. Vig, M. Dorr, and D. Cox, "Large-scale optimization of hierarchical features for saliency prediction in natural images," in *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 2798-2805, 2014. [Article \(CrossRef Link\)](#)
- [36] J. Gutiérrez, E. J. David, A. Coutrot, M. P. Da Silva, and P. L. Callet, "Introducing UN Salient360! Benchmark: A platform for evaluating visual attention models for 360° contents," in *Proc. of 2018 Tenth International Conference on Quality of Multimedia Experience*, pp. 1-3, 2018. [Article \(CrossRef Link\)](#)



Yeongil Ryu is a Ph.D. course student at the Department of Computer Education in Sungkyunkwan University (SKKU). He received his B.S. and M.S. degrees in computer engineering from Gachon University in 2016 and 2018, respectively. His current research interests include the video compression standards and 6 DoF immersive video processing.



Eun-Seok Ryu is an Associate Professor at the Department of Computer Education in Sungkyunkwan University (SKKU), Seoul, Korea. Prior to joining the University in 2019, he was an Assistant Professor at the Department of Computer Engineering in Gachon University, Seongnam, Korea, from 2015-2019. From 2014-2015, he was a Principal Engineer (Director) at Samsung Electronics, Suwon, Korea, where he led a multimedia team. He was a Staff Engineer at InterDigital Labs, San Diego, California, USA, from Jan. 2011 to Feb. 2014, where he researched and contributed to next generation video coding standards such as HEVC and SHVC. From Sep. 2008 to Dec. 2010, he was a Full-time Visiting Research Scientist II at GCATT (Georgia Centers for Advanced Telecommunications Technology) in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia USA. In 2008, he was a Research Professor at the research institute for information and communication technology in Korea University, Seoul, Korea. His research interests are in the area of multimedia computing systems that includes video source coding and wireless mobile systems. He received his B.S., M.S., and Ph.D. in computer science from Korea University in 1999, 2001, and 2008, respectively. He is a senior member of the IEEE and a member of the ACM.